BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

BRNO UNIVERSITY OF TECHNOLOGY

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION



PROCEEDINGS OF THE 22ND CONFERENCE

2016

Název:	Proceedings of the 22nd Conference STUDENT EEICT 2016
Garant	prog. Ing. Jarmila Dědková, CSc.
Editor:	doc. Ing. Vítězslav Novák, Ph.D.
Vydavatel:	Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií
Vydáno v roce:	2016
Vydání:	první

Za obsahovou a jazykovou úpravu odpovídají autoři.

ISBN 978-80-214-5350-0

Obsah

Předmluva

Středoškolské projekty

Radvanský, M.
Arduino Based AC Wattmeter 2
Kolouškova, I.
Determination of the Brain Tumor Type with Magnetic Resonance Techniques 5
Širůčková, K.
Evaluation of the Most Effective MRI Methods for a Diagnosis of the Brain Tumors 9
Benda, D.
Remote Control for Steam Engines Models 12
Andrla, O.
Retinal Vessel Segmentation in Fundus Images 15

Bakalářské projekty

Elektronika a komunikace

Zedka, R.
Analog Phase Shifters of Fractional Order 19
Kaděra, P.
Energy Harvesting Demonstrator for GSM Band 22
Hudec, D.
Flood attacks generation 25
Fišer, O.
Integrated Multichanel Audio Amplifier for Home Cinema 28
Pařízek, T.
Mini-Whip Antenna 31
Kovaľová, S.
The Design and Realisation of Metamaterial 34
Matoušek, M.
VHF SSB Receiver with Low Consumption 37

Zpracování signálů, obrazu a dat

Giertl, J.
Arduino Based Serial Bus Decoder 41
Valíčková, V.
Bioimpedance Signal Processing 44
Kulíková, B.
Calculation of the Disparity Map of Stereo Color Images 47
Kozlová, D.
Medical Image Segmentation Based on Graph Cut with Shape Prior 50
Nejedlý, P., Virgala, J.
Physiocrate: A SignalPlant Toolbox for Respiratory, Blood Pressure and EMG Signal Analysis 53

Automatizace a silnoproudá elektrotechnika

Majchrák, F.
Belt Filling Line 57
Benda, D.
Electronic Parts of Small Photovoltaic Power Plant 60
Doležal, L.
Propulsion System for Electric Scooter 63
Hnát, M.
Robocar The Autonomous Driving 66
Zajac, T.
Battery Electric Vehicle Charging and Operating Costs Analysis 70
Pinďák, M.
Multipurpose electroplating power supply 73
Veselý, P.
Three-Phase Inverter for High Speed Induction Motor 76

Biomedicínské inženýrství

Mívalt, F.
A Technique for Parameter Eestimation of Equivalent-Circuit Models 80
Vařečka, M.
Active Forearm Prosthesis 83
Svoboda, M.
Active Prosthetic Hand 86

Sobotková, M.
Microcontroller Audiostimulation for Brain FMRI Examinations 89
Blažej, S.
Motion Detection of Small Objects 92
Serečunová, S.
Music Volume Adjustment Based on Distance Detection 95
Demjan, M.
LED Pipetting Indicator 98

Mikroelektronika a komunikační technologie

Sládeček, M.
An External User Interface With USB 102
Poloch, M.
Laser Target for Simulating Shooting Sport 105
Krejčí, P.
Printed Planar Inductor 108
Kloda, T.
Connection Establishment between SDON Controller and OF Agents within OXC 112
Sikora, V.
Design of Optical Wavelength Selective Switch 115
Dorazin, D.
Fiber Optic Michelson Interferometer 118
Vaněk, S.
Normalized Inter-class Variance (NICV) 121

Magisterské projekty

Elektronika a komunikace

Gajdoš, A.
Electronically Reconfigurable Frequency Filters 125
Dulesov, E.
Identification of Spectral Components Significance 128
Gallo, M.
Implementation error coding for NG-PON2 in matlab 132
Novák, M.
Improvement of Bit Error Rate in Free Space Optical Link 135
Hradňanský, V.
Multimode VHF Receiver 138
Janík, L.
Network Interface for Microblaze 141
Polešáková, Z.,
Pronciple of the New AdjustedArchitecture of R-2R D/A Converter 144
Sládok, O.,
Pseudo-differential Filter Using DDCCs 147
Štůsek, M.,
Wireless Control of Embedded M2M Devices Using Unlicensed ISM Band 150
Hladký D.
Designing a Directional Acoustic Transmitter 153
Marek M.
Comparison of Multi-objective Optimization Methods 157

Zpracování signálu, obrazu a dat

Musil, D.
Algorithm for Detection of Positive and Negative Text 161
Pacura, D.
Invisible Image Watermarking using Frequency Domain Feature Points 164
Honc, L.
Order Analysis 168
Rajnoch, D.
Person Following by Spot Light Using Camera 171
Kašpar, J.
Plagiarism Detection in Program Codes Using Mapping Technique 174
Majtán, M.
Trainable Image Segmentation Using Deep Neural Networks 177

Automatizace a silnoproudá elektrotechnika

Skalský, M.
Closed-Loop Interferometric Fiber-Optic Gyroscope with Unpolarized Light 181
Kunz, J.
Field Instrument with HART Interface 184
Hasík, S.
System for MEMS Gyroscope Testing 187
Gál, M.
Universal Measurement Interface for Digital Audio Signal 190
Vícha, A.
Impact of Surface-Mounted Permanent Magnets Imperfections on the Machine 193
Szekeres, V.
Rotor Slot Optimization of an Induction Motor 196
Veteška, M.
Development of the Versatile Industrial Manipulator 199

Mikroelektronika a technologie

Toufar, M.
Analysis of Application Flux and Solder Paste on PCB for BGA Components 203
Vala, M.
Analysis of Defects on PCB Using X-RAY, 3D SW CERA and Micro-Sections 206
Novák, L.
Bootloader for Sci-Trace 209
Jelínek, M.
Calibration of the FBG Measuring System by Laser Interferometer 212
Bartek, T.
Driver Utilizing HDMI Interface for Modular LED Displays 215
Jaššo, K.
Effect of Compaction Pressure to the Electrochemical Properties of the Electrodes for Li-S Accumulators 218
Koporec, L.
Effect of Plasma Oxidation of Carbon Working Electrode on Its Electrochemical Response 221
Eliáš, J.
FPGA Module for Fast Waveform Generation and Synchronous Data Acquisition 224
Appel, D.
Optimization of Switching Regulated DC-DC Converter Design 227
Dušek, J.
Reactive Magnetron Sputtering of the Aluminium Oxide Coating 231
Gottwald, P.
Solar Power Inverter 234
Starčok, T.
Streamlining the Process of Cleaning PCB after Removing BGA 237

Biomedicínské inženýrství

Vičar, T.
Active Contours Initialization for Cell Tracking in the Images from Holographic Microscope 241
Vaněčková, T.
Alignment-free Methods for Classification of Metagenomic Data 244
Kupková, K.
Alignment-free Visualization of Metagenomic Data by Genomic Signal Processing 247
Fajmon, F.
Electrooculography in Marketing 250
Bernátek, P.
Pulse Visualization in Videosequence of Face 253
Kosková, M.
The Simulation of Effect of MR Pulse Sequences on the Measurement Object Temperature 256

Komunikační technologie

Důbrava, M.
Data flow Control in Autonomous Systems 260
Grabovský, Š.
Emulator Transmission Parameters of Data Networks 263
Sejkora, P.
MPT and MPTCP Mechanisms in Data Network 266
Cymorek, P.
Simulation of Static Grant Allocation in NG-PON2 Networks 270
Rajnoha, M.
Warehouse Modeling Using Graphical User Interface 273

Doktorské projekty

Elektronika a komunikace

Vychodil, J.
60 GHz Time Domain Channel Sounder 277
Nawfal Al-Zubaidi R-Smith
Accuracy Comparison of Some 2D Numerical Inverse Laplace Transform Methods 282
Dvořák, J., Langhammer, L.
Adjustable Fractional-Order High-Pass Filter with Transconductance Amplifiers 288
Kartci, A.
Current-Mode Multifunctional Filter Utilizing Losses of Inductor Simulator 293
Pospíšil, M.
Digital Predistortion Experiments with MATLAB-controlled Software Defined Radio 298
Mrnka, M.
Directive Dielectric Resonator Antenna 303
Bartoš, M.
Educational Model of Synthetic Aperture Sonar Using Range-Doppler Algorithm 308
Cupal, M.
H-plane SIW Howrn Antenna for On-body Communication 313
Hubka, P.
Linearly Polarized HMSIW U-slot Antenna for Off-body Communication 318
Mencáková, K.
Matrix Solution of Difference Linear Systems Arising in the Theory of Digital Filters 323
Kotol, M.
Neural Model of Transmission Channel in 60 Ghz ISM Band 328
Krutílek, D.
Numerical Analysis of the Lightning on Small Aircraft 333
Polák, J., Langhammer, L.
Single-Ended Frequency Filter in Practical Realization with Digitally Adjustable Current Amplifier 338
Vašina, P.
Textile SIW Rectangular Ring-slot Antennas for Off-body Communication 343
Levek, V.
The Description of the Program for Factorization of Large Numbers 348

Zpracování signálů, obrazu a dat

Novosadová, M.
Sparse Signal Recovery via Convex Optimization 354
Báňa, J.
Analysis of Speech Signals in VoIP Transferred Over a Network Channel 359
Smíšek, R.
ECG signal classification based on SVM 365
Němcová, A.
Evaluation of Distortion of ECG Signals after Compression 370
Klejmova, E.
Evaluation of Methods for AR Coefficients Estimation Using Monte Carlo Analysis 375
Hrabina, M.
Implementing Gunshot Detection Algorithm on Digital Signal Processor 380
Mego, R.
Instruction Mapping Process on the VLIW Architectures 385
Zach, O.
Let the Crowd be Heard - A Crowdsourcing-based Study on Video Coding 390
Kufa, J.
Predefined Quality Profiles in Different Versions of HEVC Encoder - Influence of PC Architecture of Processor 395
Maršánová, L.
Rules for Determination of Expected P Wave Type in ECG Signals 400

Kybernetika a automatizace

Obšilová, L.
Comparison of the Methods Used to Low Level DC Voltage Measurement 406
Klečka, J.
Continuous Realization of Monocular Structure from Motion 411
Kozovský, M.
Control Platform for Multiphase Motors 416
Klusáček, J.
Distributed Framework for Testing Machine Learning Methods 421
Glos, J.
FMU Toolbox for Matlab/Simulink 426
Davídek, D.
Image Processing Algorithms Visualiser 431
Nováček, P.
Modern Methods for Image Quality Assessment 436
Mynář, Z.
Power Optimized Control of Synchronous Reluctance Machine 441
Čala, M.
Suppression of Resonances from Vibration Signal using Vold-Kalman Filter 446
Gábrlík, P.
Transformation of UAV Attitude and Position for Use in Direct Georeferencing 451
Arm, J.
Web Control of the Robotic Arm 456

Silnoproudá elektrotechnika a elektroenergetika

Ishkova, I.
Analysis of Eccentricity and Broken Rotor Bar Faults Effect on the Induction Motor Spectrums of Current and Stray Flux 462
Zelenka, M.
Calculation and Measurement of Thermostatic Bimetal Deflection in Molded Case Circuit Breaker 467
Vyčítal, V.
Electricity Load Forecasting Using Auto-Regressive and Artificial Neural Network Model 472
Šimek, D.
Experimental Analysis of Low Voltage Arc Between Carbon Electrodes with Unknown Cores 477
Klíma, J.
Finite Element Analysis of High Speed Induction Motor With Axially Slitted Rotor 482
Kroupa, M.
Frequency Characteristic of the Induction Machine 487
Ondrejček, V.
Laboratory Workbench for Educational Purposes 492
Toman, M.
Thermal Model of Induction Machine and its Practical Verification 497
Bulín, T.
Problems with Magnetic Measurement of 41CrMo4 Steel 502
Bárta, J.
The Eddy Current Losses of High-Speed Solid Rotor Induction Machine as Function of the Radial Air-Gap Lenght 507
Novotný, F.
The Process of Cross Section Generation for Reactor Core Simulations 512

Biomedicínské inženýrství a bioinformatika

Kiska, T.
Acoustic Analysis of Sentences Complicated for Articulation in Patients with Parkinson's Disease 518
Koščová, P.
Comparative Analysis of Common and Unique Targets in Drug Resistant Strain of Borrelia Burgdorferi ..523

Potočňák, T.
EEG Signal Analysis based on EMD and Discrete Energy Separation Algorithm 528
Galáž, Z.
Potential of Prosodic Features to Estimate Degree of Parkinson's Disease Severity 533
Galáž, Z.
Quantification of Prosodic Impairment in Patients with Idiopathic Parkinson's Disease 538
Němcová, A., Maršánová, L., Smíšek, R.
Recommendations for ECG Acquisition Using BITalino 543
Dvořáková, L.
Retrospective-gating Myocardial T1 Mapping in Rats with Doxorubicin Cardiomyopathy 548
Hesko, B.
Segmentation of the Common Carotid Artery Intima Media Using Active Contour Models 553
Sedlář, K.
Signal Based Feature Selection for Fast Classification of Sequences in Metagenomics 558

Komunikační technologie

Kenyeres, M., Novotný, B.
Dependency of the Convergence Rate Mean Extent of Variation on the Repetitions Number in Weakly Connected Topologies 564
Kenyeres, M., Novotný, B.
Dependency of the Convergence Rate Mean Extent of Variation on the Repetitions Number in Strongly Connected Topologies 569
Fujdiak, R., Mašek, P.
In-Depth Analysis of Smart City in Modern Age 574
Čučka, M.
Pulse Generator for Φ-OTDR in VPI photonics 579
Vélim, J.
Relative Permittivity Measurement of 3D-textile 584
Člupek, V.
Strong Unilateral Authentication of Low-cost Devices Involved in Internet of Things in Smart Homes 589

Mikroelektronika a technologie

Bayer, R.
Analysis of Whirling of Gas Flow within the Specimen Chamber of ESEM 595
Bílek, M., Hlavatá, P.
Comparison of Results Optained Using Monte Carlo and Ansys Fluent in Analysis Differentially Pumped Chamber 600
Kerndl, M.
Energy Harvesting Using Nanogenerators in Iot 605
Otáhal, A., Crha, A., Šimek, V.
Enhanced Method OF Through-hole Copper Plating 610
Gottwald, T.
The Influence of the Binder on the Final Capacity of the Negative Electrode Material Carbon CR5995 for Lithium-ion Batteries 615
Jahn, M.
Gel Polymer Electrolytes for Use in Lithium-ion Batteries 619
Dvořák, V.
High Precision Reference Position Generator for Motor Control in FPGA 625
Strachala, D., Hylský, J., Kadlec, M.
Organic-Inorganic Hybrid Solar Cells 630
Přikrylová, K.
Influence of Anodization Parameters on Morphology of Titania Nanotubes 635
Pytlíček, Z.
Modelling of Gas Diffusion In The 3-D Structured Metal Oxide Nanofilm 640
Znbill, L.
Low Voltage Converter for Organic Tranducers in Energy Harvesting 645

Šetka, M.
Galvanic Deposition of Gold Nanorods via AAO Template for VOCs Sensing 650
Skácel, J., Otáhal, A.
Sensitive Analysis of IR Heat Emitter for Soldering 655
Pejović, J.
Separation and Detection of Bioconjugated Quantum Dots Using on a Chip Electrophoresis 660
Polievková, E.
Shape Controlled Synthesis of Silver Nanoparticles with Antimicrobial Effect 665
Lednický, T.
Solid-state Synthesis of Gold Nanoparticles on Dimpled Aluminum Surfaces and Their Transfer to Transparent Substrates 670
Kynclová, H.
Study of Electrochemical Behaviour of Gold Nanostructured Surfaces 675
Kuparowitz, T.
Supercapacitor Parameters Dependence on Charging Duration 680
Novotný, V., Vala, R.
The Analysis of Solder Preforms in Surface Mount Assembly 685
Mojrová, B.
Impact of the SiNx Thickness on Passivation Quality and Contact Resistivity of Silicon Solar Cell 690
Salma Bay Abo Dabbous
Ultra Low Power Tunable Transconductor 695
Kuchtová, Z.
Wireless Communication Platform IQRF – a Case Study 700
Veverka, D.
Wireless Sensor Triggering Using Bluetooth Low Energy Technology 707

Teoretická elektrotechnika, fyzika a matematika

Sliž, J.
A Comparison of Edge-Based and Region-Based Segmentation Performed with PDEs 713
Jurčík, M.
Automated Electrochemical Etching Prototype for Cold Field Emission Cathodes 718
Smetana, B.
Computer-assisted Search of the Roots of the Polynomial 723
Lontrasová, A.
Epoxy Resin with Aaluminium Oxide Nanofiller and its Dielectric Properties 728
Vážanová, G.
Global Solutions to Linear Delayed Differential Equations 733
Kaspar, P.
Local Electron-Sample Interaction During Scanning Electron Microscopy on Organic and Metallic Objects 738
Škvarenina, Ľ.
Microstructural Defects of Solar Cells Investigated by a Variety Diagnostics Methods 743
Demchenko, H.
Optimization of Linear Differential Systems with a Delay by Lyapunov's Direct Method 748
Staněk, D.
Rough Sets on State Spaces of Automata 753
Podshivalov, A.
Signal Filtration for Detecting Electromagetic Emission 758
Šafařík, J.
Solution of a Weakly Delayed Difference System 763
Klimešová, M.
Stability of the Zero Solution of Stochastic Differential System 768
Prokopyeva, E.
The Influence of the Type of Electrodes on the Biological Impedance Measurement 773

Ladies and gentlemen,

You have just opened the Proceedings of the 22nd edition of the **STUDENT EEICT 2016** competition. The abbreviation in the name of this event stands for **E**lectrical **E**ngineering, **I**nformation and **C**ommunication **T**echnologies, which describe the priority specializations of the organizing faculty – the Faculty of Electrical Engineering and Communication of the Brno University of Technology.

The conference was organized in the framework of specific research at the Faculty of Electrical Engineering and Communication under support from Ministry of Education, Youth and Sports. The competition is aimed to motivate students in working on industrial and grant projects, which is more difficult than a simple defending a diploma thesis. Industrial and grant projects require a significant invention of the researcher and a huge portion of her (his) creativity. Naturally, the perfect engineering processing of the project is the matter of course.

In these proceedings, you can find papers describing the students' competitive projects. These projects were reviewed by academicians of the organizing faculty, were corrected and rearranged by authors in order to meet suggestions of reviewers, and now, they are ready to be defended in front of commissions consisting of industrial specialists, university people and representatives of students.

This year is successful from the organizers point of view. There have been about 197 papers reviewed in the first stage. The JobFair event is now very important part of the conference focused on our industrial partners with 39 participants in this year. And finally, the conference has obtained many significant media partners.

As already indicated, industrial partners play a very significant role in the frame of the competition. The organizing faculty are highly interested to meet companies' demands both in the education of their potential new employees (today's students of faculties) and in their research. Therefore, the organizers highly appreciate the goodwill of companies to nominate members of commissions evaluating students' projects, to participate in publishing proceedings of the competition, and to financially support the competition.

I hope the 22nd edition of the **STUDENT EEICT 2016** competition is successful for all the participants. Students cannot only win prices but also can hear the opinion on their development from industrial specialists. They can build contacts with companies. Companies can meet excellent students and offer them further cooperation or even job. And teachers can compare results reached by their students with results of students supervised by their colleagues.

I would like to thank all students and participants who contributed to the success of the competition. I believe that all of them have gathered much experience, inspiration and motivation for their further research work and that they will retain pleasant memories of a frank and open atmosphere of the competition.

Prof. Ing. Jarmila Dědková, CSc.

Dean of the Faculty of Electrical Engineering and Communication

Středoškolské projekty

ARDUINO BASED AC WATTMETER

Martin Radvanský
Secondary Technical School (3), Brno, Purkyňova 97
E-mail: radvansky.martin@sspbrno.cz

Supervised by: Jaroslav Nesvadba
E-mail: jaroslav.nesvadba@sspbrno.cz

Abstract: This work deals with the design and manufacture of the AC wattmeter. It is controlled by the Arduino Nano and could be connected to 3 phase AC system. It is designed to be included in the DIN rail of the house central switchboard for monitoring of power consumption. Device contains build–in RS 485 interface for connection to the smart house control system. The device is designed to use separate current and voltage transformers for safety reason.

Keywords: Arduino, AC wattmeter, smart device, Microcontroller

1 ÚVOD

Domácnosti jsou dnes vybaveny velkým množstvím spotřebičů elektrické energie a jednou z cest, jak snížit výdaje v domácnosti je snažit se snížit její spotřebu. Existuje celá řada spotřebičů, které se předhánějí v tom, kolik "A" je na jejich energetickém štítku. Abychom mohli ke snížení spotřeby energie přistoupit, je třeba napřed spotřebu energie v domácnosti přesně poznat. Každá domácnost je připojena k rozvodné síti pomocí elektroměru, kde se dají potřebné informace získat. Opisování údajů a jejich sledovaná je ovšem časově náročná činnost, která většinou nejde nijak automatizovat.

Účelem tohoto projektu je navrhnout střídavý wattmetr, schopný měřit spotřebu pro 3 samostatné okruhy nebo fáze a tyto dále odesílat do nadřízeného řídícího systému.

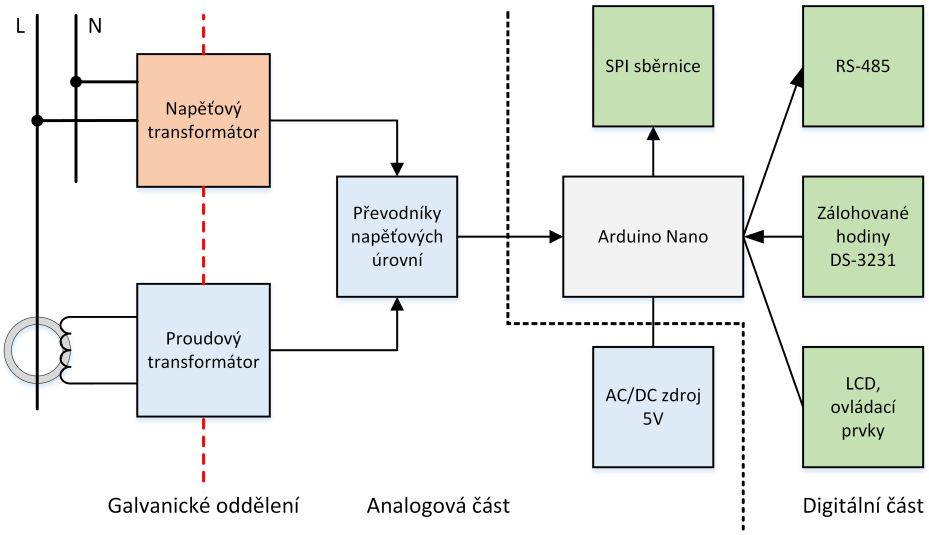
2 NÁVRH AC WATTMETRU

Při návrhu wattmetru se vycházelo z následujících požadavků:

- Zařízení musí být bezpečné pro obsluhu.
- Jednoduchá montáž v domovním rozvaděči
- Připojení do řídícího systému na bázi PLC.
- Podpora měření ve více fázích nebo okruzích.
- Využití Arduina Nano, jako řídící jednotky wattmetru.
- Co nejmenší náklady.

Při vývoji zařízení se jsme se inspirovali dostupnýni projekty, jako je například OpenEnergyMonitor [1]. Navržené zařízení vychází z jejich principů, liší se v analogové části převodu střídavého napětí na napětí měřitelné AD převodníkem Arduina. Blokové schéma wattmetru je znázorněno na obrázku 1.

Wattmetr je rozdělen do dvou hlavních částí:



Obrázek 1: Blokové schéma zařízení

- **Analogová část:**

  V této části se řeší připojení na rozvod nízkého napětí, při zachování maximální možné bezpečnosti a převod malého střídavého napětí takovým způsobem, aby bylo možné jeho měření AD převodníkem Arduina.

  - Navržené řešení obsahuje použití napěťových transformátorů, které převádějí nízké síťové napětí na malé napětí 6V.
  - Dále jsou použity proudové transformátory s měřícím rozsahem 30A/1V.
  - Pro převod napěťových úrovní je na rozdíl od většiny používaných řešení použito operačních zesilovačů, které vydělené vstupní napětí posunou o referenční stejnosměrnou složku 2,5V. Vstupní dělič napětí tvořený víceotáčkovým odporovým trimrem slouží pro kalibraci zařízení.
  - Pro napájení zařízení je použit spínaný zdroj malého napětí 5V, sloužící jak pro analogovou, tak pro digitální část přístroje.

- **Digitální část:**

  Digitální část wattmetru je tvořena mikrokontrolérem, jehož hlavním úkolem je převádět napětí z napěťových a proudových transformátorů na digitální hodnoty a následně vypočítávat okamžitou spotřebu pro každý okruh samostatně. Mezi jeho další funkce patří odesílání hodnot do nadřízeného systému a interakce s uživatelem.

  - Ústředním prvkem digitální části je deska Arduino Nano [2], která byla zvolena pro své malé rozměry. Z této desky je využito 6 AD linek sloužící pro převod vstupního napětí v rozsahu 0–5V do digitální podoby pomocí interního převodníku.
  - Měřená data jsou zobrazována na připojeném LCD displeji a komunikace s uživatelem je možná pomocí rotačního enkodéru, ve funkci navigačního prvku pohybu po menu.
  - Zařízení obsahuje modul s obvodem DS3231 připojený na sběrnici I2C, který je zdrojem hodin reálného času, zálohovaný pomocí baterie a umožňuje odesílaná data doplnit o časovou hodnotu.
  - Data jsou odesílána pomocí modulu obsahujícího převodník na sběrnici RS-485.

– Pro další možné rozšíření funkčnosti lze využít sběrnice SPI (např. moduly IQRF pro bezdrátové odesílání měřených dat).

3 OŽIVENÍ ZAŘÍZENÍ A OVĚŘENÍ FUNČNOSTI

Oživení hardware zařízení nepředstavuje žádný problém a v případě použití kvalitních součástek by mělo pracovat ihned po sestavení. Nejdůležitější částí oživení zařízení je jeho správná kalibrace, což již není triviální úkol. Samotná kalibrace se musí provádět pro každý okruh samostatně a skládá se z kalibrace napětí přiváděného na AD převodník mikrokontroléru z proudových a napěťových transformátorů. Přesnost provedené kalibrace má velký vliv na dosaženou celkovou přesnost měření spotřeby.

- Kalibrace napětí v okruhu se provádí pomocí autoransformátoru, kdy na základě známého napětí (230V) na vstupu měřicího transformátoru, nastavíme pomocí víceotáčkového odporového trimru špičkové napětí na 1V a to za pomocí osciloskopu. Na vstupu mikrokontroléru tak dostáváme napětí v rozsahu 1,5V až 3,5V. Kolísání síťového napětí se projevuje odchylkami v řádu desítek mV a jsou těsně nad hranicí měřitelnosti pomocí AD převodníků v Arduinu (4,8mV).
- Pro kalibraci napětí pocházejících z proudových měřících transformátorů je použito stejné metody, jen výstupní napětí kalibrujeme při konstantním měřeném proudu na odporové zátěži. Při kalibraci byl použit proud 1,5A a nastavované špičkové napětí bylo 100mV. Výstupní napětí pro proud 15A tak odpovídá 1V špičkového napětí a na AD převodník tak přivádíme napětí v rozsahu 1,5V až 3,5V. Hranice 15A byla zvolena z důvodu jištění zásuvkových okruhů pomocí jističů na 16A.

Program v mikrokontroléru provádí výpočet okamžitého výkonu a jeho průměrování v 5s intervalech, kdy na konci intervalu je odesílána průměrná spotřeba v okruhu po sběrnici RS-485, protokolem MODBUS RTU.

Při otestování funkčnosti bylo provedeno měření spotřeby při odporové zátěži (100W žárovky) a to bylo porovnáno s hodnotou získanou pomocí měřiče spotřeby Voltcraft Energy Logger 3500 [3]. Výsledek měření pomocí navrženého zařízení byl o 23,5% nižší než uvedeným přístrojem Voltcraft. Vzhledem k nákladům na zařízení ve výši přibližně 720 Kč je uvedená hodnota přijatelná.

4 ZÁVĚR

Zařízení bylo otestováno a porovnáno s dostupným přístrojem určeným pro měření a záznam spotřeby. I když výsledná spotřeba byla naměřena přibližně o 24% nižší než v případě komerčního zařízení, je výsledek vzhledem k nákladům dostatečný. Přesnost navrženého zařízení by bylo možné zvýšit lepší kalibrací a ověřením, zda proudové a napěťové transformátory mají lineární průběh výstupního napětí v závislosti na proudu. Přesnost výpočtu spotřeby je úzce svázána s dobře provedenou kalibrací a také přesností použitých přístrojů.

REFERENCE

[1] Guide. OpenEnergyMonitor. [online]. 26.3.2016 [cit. 2016-03-26]. Dostupné z: https://openenergymonitor.org/emon/guide

[2] Board Nano. Arduino. [online]. 26.3.2016 [cit. 2016-03-26]. Dostupné z: https://www.arduino.cc/en/Main/ArduinoBoardNano

[3] Energy Logger 3500. TD-er's Wiki. [online]. 11.9.2011 [cit. 2016-03-26]. Dostupné z: http://wiki.td-er.nl/index.php?title=Energy_Logger_3500

DETERMINATION OF THE BRAIN TUMOR TYPE WITH MAGNETIC RESONANCE TECHNIQUES

Ivana Koloušková

Gymnázium Matyáše Lercha, Žižkova 55, Brno

E-mail: ivcakolouskova@seznam.cz

Supervised by: Jan Mikulka

E-mail: mikulka@feec.vutbr.cz

Abstract: This thesis discusses the determination of primary and secondary brain tumours. The aim is to find significant differences between these two pathological tissues, using image characterization in manually segmented tumour and edema areas in T_1 and FLAIR magnetic resonance images. The selected areas of gliomas and brain metastases were described via measuring the shape and intensity descriptors. The obtained data were processed and analysed; after this stage, statistical tests of mutual distinctions were performed. The proposed tests prove the existence of 11 parameters, which exhibit different values for the primary tumour and the brain metastase groups.

Keywords: classification of tumours, intensity and shape descriptors, statistical processing of MRI

1. ÚVOD

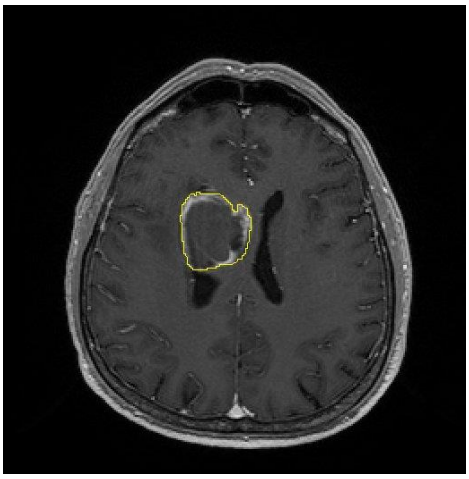
Rostoucí incidence nádorových onemocnění dnes patří k celosvětovým problémům. K diagnostice těchto nemocí výrazně přispívají moderní zobrazovací metody, jako je magnetická rezonance (MR) a její modality, přesto však někdy bývá výsledná diagnóza nepřesná. Rozlišení jednotlivých typů mozkových tumorů vyžaduje často několikanásobné snímání, následně se subjektivně posuzují pouze základní parametry jako velikost nebo intenzita postižené oblasti, a to z důvodu rychlosti a jednoduchosti diagnostiky. V rámci tohoto výzkumu byly měřeny ostatní tvarové a jasové parametry oblastí tkáně tumoru nebo edému na T1W (snímcích vážených relaxačním časem T_1) a FLAIR (snímcích vážených relaxačním časem T_2 s potlačením signálu likvoru) snímcích MR, následně byla statisticky vyhodnocena odlišnost těchto parametrů mezi skupinou pacientů s primárním mozkovým nádorem a skupinou pacientů s mozkovou metastázou.

2. METODOLOGIE VÝZKUMU

Oddělení radiologie FN Brno v Bohunicích pro tento výzkum poskytlo T1W a FLAIR snímky pacientů s nádorem mozku (viz **obrázek 1**), kteří podstoupili vyšetření MR v letech 2011 - 2014. Tyto pacienty bylo možné rozdělit do dvou kategorií: a) 8 pacientů s primárním tumorem a b) 5 pacientů s metastázou. Obě skupiny pacientů obsahovaly zástupce obou pohlaví i více věkových kategorií.

Prohlížení MR snímků pacientů umožnil program *3D Slicer*, ve kterém byly vytvořeny a vyexportovány série snímků konkrétních pacientů k dalšímu zpracování. V programu *Fiji* bylo z každé série vybráno 10 reprezentativních řezů celkovým rozložením nádoru a otoku každého pacienta. Pomocí nástroje ručního výběru byly označeny oblasti tumoru a edému na každém snímku. Poté bylo provedeno měření zvolených parametrů.

Mezi zkoumané parametry patří tvarové deskriptory (plocha, obvod, délka hlavní a vedlejší poloosy vepsané elipsy, jejich poměr a sklon elipsy v obraze, kruhovost, kulatost i celistvost zvolené oblasti) a jasové deskriptory (střední hodnota, směrodatná odchylka, modální hodnota, maximální a minimální hodnota intenzity, součet intenzit, medián intenzit, šikmost a strmost hustoty pravděpodobnosti dané oblasti).



Obrázek 1: Příklad MR snímku mozku s označením oblasti nádorové tkáně.

3. STATISTICKÉ ZPRACOVÁNÍ DAT

Z 10 hodnot pro každý parametr nádoru nebo otoku jednotlivých pacientů byly vytvořeny mediánové hodnoty ke snížení odchylky způsobené analýzou 2D namísto 3D dat. Každý pacient byl nyní reprezentován čtyřmi hodnotami pro každý zkoumaný parametr.

3.1. KRABICOVÉ GRAFY

Pro první porovnání hodnot výzkumných skupin gliom/metastáza byla využita metoda krabicových grafů. U 12 z celkových 57 krabicových grafů byly viditelné rozdíly mezi skupinami gliomů a metastáz a tyto parametry byly využity k další analýze.

3.2. INTERVALOVÉ ODHADY PARAMETRŮ

Intervalové odhady se využívají pro odhad vlastností a charakteristik základního souboru, jako např. střední hodnoty parametrů náhodného procesu. Výsledkem je interval, ve kterém se střední hodnota základního souboru bude se zvolenou pravděpodobností nacházet. Pravděpodobnost udává riziko odhadu α. Toto riziko se volí zpravidla 0,05 nebo 0,01. Zvolení rizika α získáme i spolehlivost výsledného intervalu, která je rovna hodnotě (1 - α). Zvýšením spolehlivosti odhadu (1 - α) dojde k rozšíření intervalu spolehlivosti a tím i ke snížení přesnosti tohoto odhadu a naopak. Samotné stanovení intervalového odhadu parametrů vychází z nerovnice 1:

$$x - t_{1-\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}} < u < x + t_{1-\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{2}}, \tag{1}$$

kde $t_{1-\frac{\alpha}{2}}$ značí (1 - α/2) · 100 % kvantil Studentova rozdělení se stupni volnosti $n - 1$, hodnota $\frac{s}{\sqrt{n}}$ udává odhad směrodatné chyby a výraz $t_{1-\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}}$ značí přípustnou chybu výsledného odhadu. Nejprve byly u všech 12 parametrů vypočteny intervalové odhady s 95% pravděpodobností. V případě, že se výsledné intervaly obou skupin vzájemně překrývaly, byla pravděpodobnost snížena na 90 %, u jiných parametrů bylo naopak možné zvýšit přesnost intervalu až na 99 %. Vzájemně odlišné intervaly středních hodnot se podařilo najít pro 11 parametrů.

3.3. TESTY STATISTICKÝCH HYPOTÉZ

Statistická hypotéza je určité tvrzení o parametrech základního souboru a tvoří ji dvě teze. Nulová (H_0) a alternativní (A) hypotéza, které jsou vzájemnými opaky. Test statistických hypotéz představuje přijmutí jedné hypotézy či její zamítnutí ve prospěch hypotézy druhé na hladině významnosti α. Po určení rizika α (a tím i síly tesu) se provádí výpočet testového kritéria podle rovnice 2:

$$t = \frac{X-\mu_0}{s}\sqrt{n}, \tag{2}$$

kde *x* reprezentuje výběrový průměr hodnot základního souboru, μ předpokládanou střední hodnotu základního souboru, *s* výběrovou směrodatnou odchylku dat základního souboru a *n* počet hodnot základního souboru. Toto kritérium má při platnosti hypotézy H_0 Studentovo rozdělení $t(V)$ s $(n - 1)$ stupni volnosti. Testové kritérium určuje obor přijetí i kritický obor hypotézy, daný rovnicí 3:

$$W_\alpha = \{t; t \geq t_{1-\alpha}(V)\} \tag{3}$$

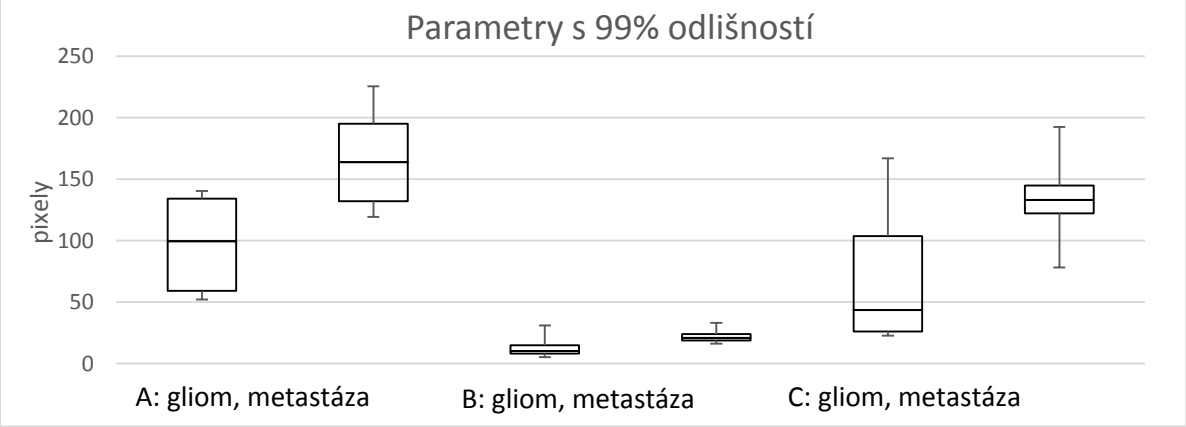
kde *t* značí testové kritérium a $t_{1-\alpha}(V)$ kvantil Studentova rozdělení. Platí-li tato nerovnost, dojde k zamítnutí H_0 ve prospěch alternativní hypotézy.

Tyto testy sloužily k ověření krajních mezí intervalových odhadů a tím i odlišnosti hodnot skupin primárních a sekundárních nádorů. To se povedlo u všech 11 parametrů na stejné hladině spolehlivosti jako jejich intervalový odhad střední hodnoty.

4. ZÁVĚR

Cílem této práce bylo nalezení signifikantních rozdílů mezi primárními tumory a mozkovými metastázami. Byly zjištěny a potvrzeny odlišnosti středních hodnot tvarových parametrů pro skupinu gliomů a skupinu metastáz.

S 90% pravděpodobností budou střední hodnoty obou skupin rozdílné u těchto parametrů: hlavní poloosa elipsy vepsané tumoru, obvod tumoru a obvod edému. Na 95 % se střední hodnoty výzkumných skupin budou lišit i u poměru plochy edému k ploše tumoru, hlavní poloosy elipsy vepsané edému, poměru obvodu edému k obvodu tumoru, poměru hlavní a vedlejší poloosy elipsy vepsané edému a kulatosti edému. S 99% pravděpodobností budou pro skupinu primárních nádorů a skupinu metastáz odlišné hodnoty u poměru hlavní poloosy elipsy vepsané edému k hlavní poloose elipsy vepsané tumoru (A), vedlejší poloosy elipsy vepsané edému (B) a poměru vedlejší poloosy vepsané edému k vedlejší poloose vepsané tumoru (C, viz **obrázek 2**).



Obrázek 2: Krabicový graf parametrů A (vlevo), B (uprostřed) a C (vpravo)

Tato práce slouží jako důkaz existence parametrů odlišujících gliomy a metastázy a tvoří podklad pro rozsáhlejší výzkum těchto rozdílů (již řešeno ve spolupráci s Radiologickou klinikou FN Brno). Přesnější rozlišení typů mozkových tumorů vede ke zlepšení lékařské diagnostiky i zkvalitnění péče a léčby pacientů s nádorovým onemocněním mozku.

PODĚKOVÁNÍ

Největší dík patří mému školiteli doc. Ing. Janu Mikulkovi, Ph.D. za jeho trpělivost, ochotu a cenné rady. Děkuji také své rodině za jejich pochopení, toleranci a stálou podporu.

REFERENCE

[1] Seidl, Z. et al.: Radiologie pro studium i praxi. 1. vydání. Praha: Grada Publishing 2012. ISBN 978-80-247-4108-6

[2] Neubauer, J., Sedlačík, M., Kříž, O. Základy statistiky, aplikace v technických a ekonomických oborech. 1. vydání. Praha: Grada Publishing 2012. ISBN 978-80-247-4273-1

[3] Multimediální pomůcky VFU Brno. Statistické tabulky: Kvantily $t_{1-\alpha/2(n)}$ Studentova *t*-rozdělení [online]. 2016 [cit. 6. 3. 2016]. Dostupné z: cit.vfu.cz/statpotr/POTR/Teorie/tabulky.htm#ttest

EVALUATION OF THE MOST EFFECTIVE MRI METHODS FOR A DIAGNOSIS OF THE BRAIN TUMORS

Kateřina Širůčková

Klasické a španělské gymnázium, GYBY

E-mail: katerina.siruckova@seznam.cz

Supervised by: Petr Marcoň

E-mail: marcon@feec.vutbr.cz

Abstract: The purpose of this project is to find out the most effective method for a diagnosis of the brain tumors utilizing different magnetic resonance imaging methods. Acquired MRI images were analyzed in two diagnostic programs, and obtained data were displayed into ratio-graphs. According to the results, volume-ration images are considered the most sufficient. The advantage of this method is that it provides doctors with significant information regarding the differences of the measured values between healthy and pathological tissue.

Keywords: medical imaging methods, tumor, magnetic resonance, volume-ratio

1. ÚVOD

Magnetická rezonance patří mezi nejvyužívanější zobrazovací techniky při podezření na patologie mozku. Včasná diagnostika jakéhokoliv nádoru je kritická, a proto je nutné maximalizovat využití obrazových modifikací a snižovat zátěž záření na tělo pacienta. Cílem tohoto článku je popis vyhodnocení obrazů mozku zachycených magnetickou rezonancí ve vybraných sekvencích. Dále ze získaných údajů diagnostikovat tumory a také určit, které obrazy jsou nejvhodnější k diagnostice. Důležitost této práce se skrývá v odhalení nejefektivnější metody mezi zobrazením MRI obrazů vážených frakční anizotropií (FA), objemovým poměrem (VR), relaxačním časem T_1, relaxačním časem T_2 a zdánlivým difuzním koeficientem (ADC) [1]. Vstupními materiály tohoto projektu jsou obrazy primárních a sekundárních nádorů, které byly získané z magnetické rezonance a vážené vybranými veličinami. Podrobnější popis mozkových nádorů a jejich dělení můžeme najít v literatuře [2].

2. PROGRAMY PRO VYHODNOCOVÁNÍ OBRAZŮ

Tomograf magnetické rezonance je v současné době velmi důležitým zobrazovacím přístrojem v medicíně. Profituje převážně z toho, že nezatěžuje tkáně škodlivým zářením, jak je tomu třeba u ionizačního záření využívaného u rentgenu. Z tomografu získáváme data, která se nadále zpracovávají ve speciálních programech. Za účelem diagnostiky tumoru jsem pro svou práci použila programy Weasis a DtiStudio.

2.1. WEASIS

Weasis je program poskytující webový přístup k radiologickým snímkům. Specificky u magnetické rezonance zobrazí všechny obrazy měřeného pacienta v příslušném formátu. K přesnějšímu nalezení tumoru v programu Weasis byly v tomto projektu použity obrazy s kontrastní látkou. Tento program byl s výhodou použit pro vykreslení obrazů vážených relaxačním časem T_1 a T_2.

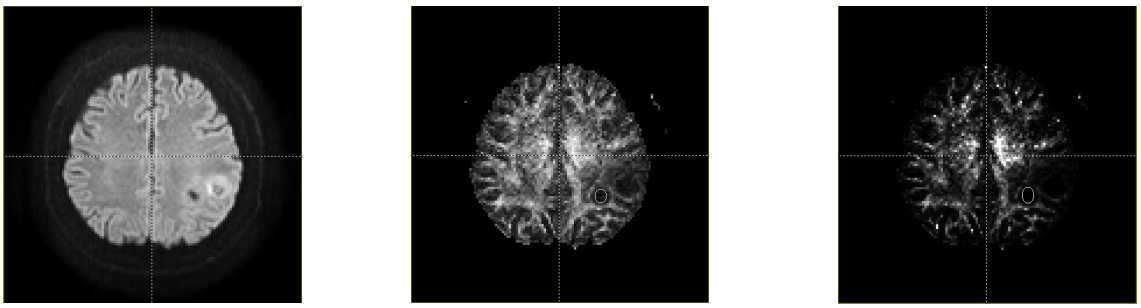
2.2. DTI STUDIO

DTI Studio je program, pomocí kterého lze zobrazit datový soubor získaný z tomografu magnetické rezonance. Je schopen vytvářet 3D obrazy mozku a také umožňuje pomocí traktografie zobrazovat nervová vlákna [1]. Po nahrání obrazů do programu lze vytvořit jak 3D rekonstrukci mozku, tak vygenerovat obrazy řezů vedené skrz axiální, koronální a sagitální rovinu. DtiStudio je určeno například pro výpočet obrazů vážených difuzním tensorem (DTI), traktografii vláken a 3D vizualizaci [3]. Program DtiStudio byl použit pro výpočet obrazů vážených FA, VR a ADC.

3. POPIS EXPERIMENTU

V praktické části jsou zkoumány rozdíly hodnot mezi zdravou a patologickou tkání v lidském mozku. Patologická tkáň může být nejjednodušeji rozdělena na nádor primární a sekundární. Vstupní data byla získána od tří pacientů s primárním nádorem a dvou se sekundárním nádorem. Každá série vstupních dat obsahovala obrazy vážené zdánlivým difuzním koeficientem (ADC), frakční anizotropií (FA), objemovým poměrem (VR) a relaxačními časy T_1 a T_2.

Následně byly vykresleny jednotlivé vážené obrazy v programech Weasis (T_1 a T_2) a DtiStudio (ADC, FA, VR – viz obrázek 1). V obrazech lze vidět oblast, kde se nachází tumor. Za účelem popisu vybraných dat byla dále provedena statistická analýza. Byly zpracovávány také obrazy vážené relaxačními časy T_1 a T_2, avšak v tomto experimentu pozbyly svojí důležitosti. Oba programy, Weasis a DtiStudio, mají obsažen modul pro statistické zpracování vybraných dat. Pro potřebu diagnostiky patologické tkáně byly ve výše uvedených programech vypočítány z vybraných dat střední hodnoty a směrodatné odchylky. Pro porovnání byly stejné statistické veličiny vypočteny z oblasti zdravé tkáně. Nutno zdůraznit, že vybraná zdravá tkáň je osově souměrná s tkání patologickou.



ADC - průměrovaný FA VR

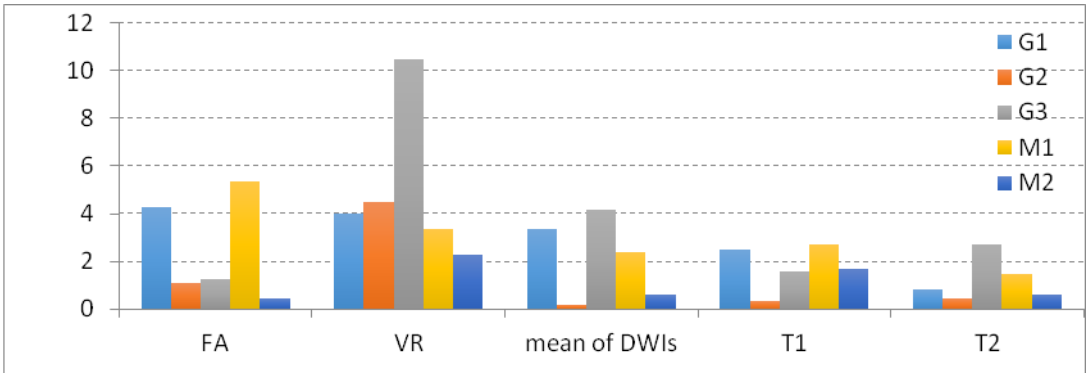
Obrázek 1: Jednotlivé vážené obrazy v DtiStudiu.

4. VYHODNOCENÍ OBRAZŮ

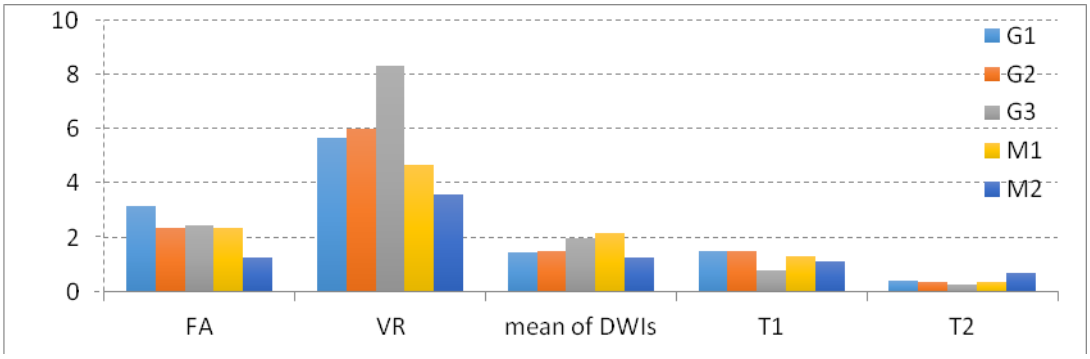
Před začátkem popisu samotného vyhodnocení obrazů, jsou vysvětleny používané zkratky v grafech: G1, G2, M1 atd. Jelikož pacienti nemohou být publikováni jménem, zvolené zkratky slouží k označení a rozpoznání pacienta. Pacienty s primárním nádorem mají označení G (jako gliom) a podle pořadí vyhodnocování mají přiřazená čísla. Označení M mají pacienti se sekundárním nádorem čili metastázou.

Ze získaných statistických údajů byly v programu Excel vytvořeny dva poměrové grafy pro směrodatné odchylky a střední hodnoty. Jde o poměr patologické tkáně vůči tkáni zdravé. To znamená, že pokud budeme porovnávat neznámou tkáň s tkání zdravou a hodnota poměru neznámé tkáně ku zdravé tkáni se bude rovnat jedné, tak můžeme konstatovat, že neznámá tkáň není patologická. Obrázek 2 je vázán k hodnotám směrodatných odchylek, oproti tomu obrázek 3 patří ke středním hodnotám. Graf směrodatných odchylek vykazuje větší odchylky od výchozí hodnoty 1, a proto můžeme říci, že je pro diagnostiku tumorů vhodnější.

Diagnostika tumoru z obrazů vážených objemovým poměrem je zvolena jako nejvhodnější, a to díky velké odchylce od hodnoty 1, kdy lze bez obav určit, že se jedná o tumor. Naproti tomu obrazy vážené relaxačními časy T_1 a T_2 nejsou vhodné, i když vizuálně je v nich tumor vidět nejlépe. Graf směrodatných odchylek na obrázku 2 obecně vykazuje větší rozdíly než graf středních hodnot na obrázku 1.



Obrázek 2: Poměrový graf hodnot směrodatných odchylek.



Obrázek 3: Poměrový graf středních hodnot.

5. ZÁVĚR

Cílem této práce bylo zjistit, která z metod zobrazování je nejvhodnější pro diagnostiku tumoru v mozkové tkáni. Za nejvhodnější je považována metoda zobrazení obrazů vážených objemovým poměrem. Tento závěr podporuje statistické vyhodnocení celého měření celkově pěti pacientů. V praxi jsou ovšem nejvíce využívány obrazy vážené frakční anizotropie. Obrazy vážené objemovým poměrem nejsou standardně používány k zobrazování v biomedicínských odvětvích, proto by bylo v budoucnu zajímavé, na tuto práci navazovat dalšími, avšak s rozsáhlejším souborem dat.

REFERENCE

[1] Mori, S.: Introduction to Diffusion Tensor Imaging, Boston, Elsevier 2006, ISBN 0-444-52828-8.

[2] *Primární a metastatické nádory mozku* [Online]. Dostupné na: http://zdravi.e15.cz/clanek/priloha-lekarske-listy/primarni-a-metastaticke-nadory-mozku-451456. [Přístup získán 20 Prosince 2015].

[3] Pavlík, J.: Difuze v kovech [online] Brno, 2013 [cit. 2016-01-10]. Dostupné z: http://www.vutbr.cz/www_base/zav_prace_soubor_verejne.php?file_id=67641. Bakalářská práce. Vysoké učení technické v Brně. Vedoucí práce Ng. EVA NOVOTNÁ, Ph.D.,Paed IGIP.

REMOTE CONTROL FOR STEAM ENGINES MODELS

David Benda

Secondary School of Technical and Technical College (3), SPŠ a VOŠT Sokolská 1, Brno

E-mail: bendadavid@seznam.cz

Supervised by: Dušan Benda

E-mail: xbenda10@stud.feec.vutbr.cz

Abstract: The aim of this project is to create a remote control for radio controlled model of steamboat. The entire device consists of two parts - a transmitter and receiver. Data transmission modules were selected from the company HopeFR communicating in a frequency band of 433 MHz. Currently the device is outfitted for steering a pair of model servo motors and five digital outputs. Simple adjustment facility can be extended as required.

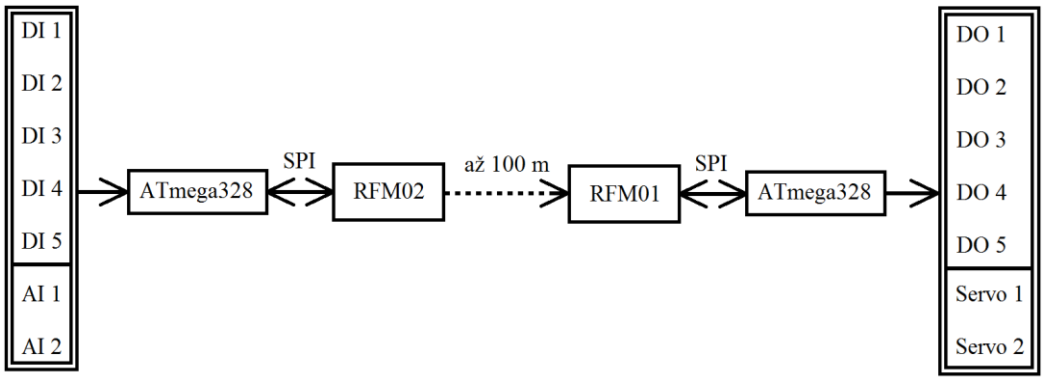
Keywords: remote control, SPI, RFM01, RFM02, ATmega328, PWM

1. ÚVOD

Většina plavby schopných modelů lodí disponuje dálkovým ovládáním. Samozřejmě lze provozovat model i bez dálkového řízení, nicméně v tomto případě model opisuje po vodní hladině obvod kruhu tj. kormidlo modelu má stále konstantní natočení. V případě neočekávaného chování (samovolné změny natočení kormidla) ztrácíme nad modelem i takto malou kontrolu. Pokud je model, respektive minimálně jeho kormidlo řízeno bezdrátově, můžeme model alespoň směrovat. Pokud koncepce parního motoru dovoluje reverzaci otáček, je možné tuto operaci opět provést pomocí modelářského servo motoru řízeného bezdrátově. Další rozšiřující periférií může být například noční osvětlení celého modelu, a to opět prostřednictvím bezdrátové komunikace.

2. KONCEPCE DÁLKOVÉHO OVLÁDÁNÍ

Blokové zapojení dálkového ovládání je uvedeno pod textem (Obrázek 2.1). O řízení komunikace na straně vysílače i přijímače se stará dvojice mikrokontrolérů firmy Atmel z řady AVR konkrétně typ ATmega328 [1]. Pro bezdrátovou komunikaci byly zvoleny moduly RFM01 [2] (přijímač) a RFM02 [3] (vysílač) od firmy HopeFR komunikující v pásmu 433 MHz. Tyto moduly komunikují s nadřazeným kontrolérem po sériové sběrnici SPI. Během návrhu vysílače a přijímače byly uvažovány dva kanály určené pro řízení modelářských servo motorů (vstupy označené jako AI 1 a AI 2, výstupy označené jako Servo 1 a Servo 2) a pětice digitálních kanálů (vstupy DI 1 až 5, výstupy DO 1 až 5) určená například pro již zmíněné osvětlení modelu.



Obrázek 2.1: Blokové zapojení vysílače a přijímače.

3. VYSÍLAČ

Úkolem vysílače, respektive řídicího mikrokontroléru (ATmega328) vysílače je neustálá kontrola a zpracování hodnot získaných z digitálních a analogových vstupů. Nashromážděná data jsou následně zpracována, normalizována, validována a posílána prostřednictvím SPI sběrnice do vysílacího modulu RFM02. Vstupní hodnota pro analogové kanály se získá pomocí potenciometru zapojeného mezi +5 V a GND. Jezdec potenciometru je připojen k příslušnému analogovému vstupu řídicímu kontroléru. Vzhledem k rozlišení AD převodníku mikrokontroléru ATmega328 (10 bitů) získáváme dostačující informaci pro následné natočení modelářského servo motoru na straně přijímače, tj. otočení potenciometru má za následek natočení serva. Pro získání stavů digitálních vstupů se využívá připojování napěťové úrovně GND (0 V) prostřednictvím tlačítek k nakonfigurovaným pinům kontroléru ATmega328. Tlačítka jsou navíc softwarově ošetřena proti zakmitávání. Aby bylo možné kontrolovat, zda jsou proměnné získané z digitálních vstupů nastaveny do logické 0, respektive logické 1, tak byly na desku vysílače umístěny ještě signalizační LED diody korespondující s aktuálně nastaveným (přijatým) stavem digitálních výstupů na straně přijímače. Celý vysílač byl nejprve navržen na nepájivém poli a následně osazen na dvouvrstvou desku plošných spojů.



Obrázek 3.1: Model lodě s otevřeným dálkovým ovladačem.

4. PŘIJÍMAČ

Obdobně jako v případě vysílače, je i v přijímači pro řízení použit mikrokontrolér ATmega328. Bezdrátovou správu zde zprostředkovává komunikační modul RFM01. Komunikace mezi modulem RFM01 a kontrolérem ATmega328 opět probíhá po sběrnici SPI. Digitální výstupy řídí unipolární tranzistory, takže je možné připojit na každý digitální výstupní kanál větší zátěž, než v situaci, kdy by byla zátěž připojena přímo na výstupu řídicího kontroléru. Nejdůležitějšími prvky, které jsou prostřednictvím přijímače řízeny, jsou modelářské servo motory. Řízení těchto pohonů probíhá prostřednictvím pulsně šířkové modulace (PWM). Frekvence opakování těchto pulsů je nastavena interním časovačem ATmega328 na frekvenci 50 Hz. Modelářský servo motor se natáčí v úhlu 0° až 90°, respektive 0° až 180° v závislosti na typu a výrobci. Pro rozsah úhlu natočení 0° až 90° odpovídá změna šířky pulsu v rozsahu 1 až 2 ms a rozsah 0° až 180° odpovídá šířce pulsu 0,5 až 2,5 ms. Obě varianty jsou v navrhované aplikaci reflektovány a změnu nastavení lze provést

prostřednictvím jednoduché úpravy programu nahraného v kontroléru ATmega328. Na straně přijímače neprobíhají již žádné výpočty pro natočení motoru, ale získané hodnoty z vysílače jsou pouze předávány internímu časovači v PWM režimu, který vytváří korektní pulsy pro řízení natočení serva. Přijímač byl obdobně jako vysílač nejprve navržen a testován na nepájivém poli a následně osazen na dvouvrstvou desku plošných spojů.

5. KOMUNIKAČNÍ PAKET A ZPŮSOB KOMUNIKACE

Komunikace mezi řídicím kontrolérem a modulem RFM01, respektive RFM02 probíhá prostřednictvím sériové sběrnice SPI. Komunikace mezi moduly je však pouze jednosměrná, tj. očekává se, že to co je nastaveno na vysílači se promítne na straně přijímače. Pro inicializaci a případné další nastavování však moduly komunikují s řídicím kontrolérem obousměrně. Datové pakety jsou posílány z vysílače s periodou 5 ms. Pro popisovanou aplikaci byl navržen vlastní komunikační paket, který se skládá celkem z 16 bajtů. Jednotlivé bajty paketu včetně vysvětlivek jsou uvedeny v obrázku pod textem (Obrázek 5.1). První dva bajty (0. a 1. bajt) tvoří hlavičku paketu. Dále následují dva bajty (2. a 3. bajt) vyhrazené pro první modelářské servo, čtvrtý bajt je vytvořen logickou operací XOR předcházejících dvou bajtů (2. a 3.) a obdobné pravidlo platí i pro druhé modelářské servo. Posledních šest datových bajtů je tvořeno stejně jako v případě přenosu dat pro modelářské servo motory s tím rozdílem, že jsou tato data vyhrazena pro přenos digitálních kanálů a zbylé nepoužité bity, respektive bajty jsou rezervovány pro možné rozšíření při posílání dalších potřebných dat. Celý paket je zakončen dvěma zakončovacími bajty. Parita v podobě logické funkce XOR se spočítá ve vysílači a následně v přijímači dojde k porovnání přijaté hodnoty logické funkce XOR s vypočtenou hodnotou z přijatých datových bajtů. V případě, že jakákoliv část paketu nevyhovuje daným požadavkům, tj. inicializační bajty (0. a 1.) nezačínají „P“ a „A“ (obdobně pro zakončovací bajty), případně kterýkoliv výsledek logické funkce XOR není validní, tak se daný přijatý paket zahazuje a čeká se na příjem dalšího korektního paketu.

P	A	Servo 1			Servo 2			DO + rezervováno			rezervováno			I	K
0x50	0x41	0xXX	0xXX	XOR	0xXX	0xXX	XOR	0xXX	0xXX	XOR	0xXX	0xXX	XOR	0x49	0x4B
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Obrázek 5.1: Struktura komunikačního paketu.

6. ZÁVĚR

Zařízení mám zabudované ve vlastnoručně navrženém a vyrobeném modelu parníku, se kterým jsem se v tomto roce (2016) přihlásil do Středoškolské odborné činnosti. Dálkové ovládání je možné použít na volném prostranství, tj. například vodní ploše na vzdálenost více než 100 m. Vzhledem ke stávající konstrukci je možné drobnými úpravami programu přidat více digitálních, respektive analogových komunikačních kanálů.

REFERENCE

[1] ATMEL. Datasheet ATmega328 [online]. 2015. Dokument dostupný na: http://www.atmel.com/images/atmel-8271-8-bit-avr-microcontroller-atmega48a-48pa-88a-88pa-168a-168pa-328-328p_datasheet_complete.pdf. [cit. 2016-03-01].

[2] HOPEFR. Datasheet RFM01 [online]. 2010. Dokument dostupný na: http://www.hoperf.com/upload/rf/RFM01.pdf. [cit. 2016-03-01].

[3] HOPEFR. Datasheet RFM02 [online]. 2010. Dokument dostupný na: http://www.hoperf.com/upload/rf/RFM02.pdf. [cit. 2016-03-01].

RETINAL VESSEL SEGMENTATION IN FUNDUS IMAGES

Ondřej Andrla

High School Student (3), Bishop Grammar School and Nursery School Brno

E-mail: ondrej.an@seznam.cz

Supervised by: Jan Odstrčilík

E-mail: odstrcilik@feec.vutbr.cz

Abstract: Segmentation of retinal vessel system plays an important role in the analysis of retinal images. In this contribution, I have used the segmentation method based on morphological image analysis. The product of this method is a final binary representation of blood vessel tree. The outcome of my work is the computer program for automatic vessel segmentation. The method was tested by standardized images from VUT database.

Keywords: fundus, retinal vessel, morphological image processing

1. ÚVOD

V rámci řešení projektu byl vytvořen uživatelský program umožňující plně automatickou segmentaci cévního řečiště ze snímků sítnice. Tato segmentace může sloužit pro účely dalšího zpracování s cílem rozpoznávání chorob postihujících cévy sítnice, jako jsou např. glaukom a diabetická retinopatie. Vyšetření retinálních cév umožňuje rozpoznat nemoc v raných stádiích, ještě než se projeví další příznaky a dovoluje tak nasazení účinnější léčby [2]. Automatizace tohoto zpracování může lékařům ušetřit velké množství práce. Prezentovaná analýza, prováděná za pomoci počítače, je založena na segmentaci cév z fotografií očního pozadí a jejich porovnání se vzorovými obrázky zdravých a poškozených očí z veřejné databáze.

2. MORFOLOGICKÉ ZPRACOVÁNÍ OBRAZU

Existuje velké množství metod, založených na různých postupech, a to od filtrace obrazu (angl. Filter-based methods) [1], přes metody založené na použití lineárních a nelineárních modelů (angl. Model-based methods) [1], až po algoritmy využívající klasifikaci obrazu (angl. Machine-learning methods) [1] .

V této práci jsem se pro segmentaci očního pozadí rozhodl využít metody morfologického zpracování obrazu [3], které by se daly zařadit do výše zmíněné skupiny filtračních metod. Tato metoda vychází z teorie množin a její základy tvoří operace dilatace a eroze. Metoda morfologického zpracování přistupuje k obrazu jako k matici. Při dilataci je k obrazu logickým součtem přičten strukturní element (matice z nul a jedniček) a při erozi je strukturní element od obrazu logicky odečten. Na těchto základních operacích jsou založeny další, jako např. morfologické otevření, což je eroze následovaná dilatací a morfologické uzavření, což je dilatace následovaná erozí.

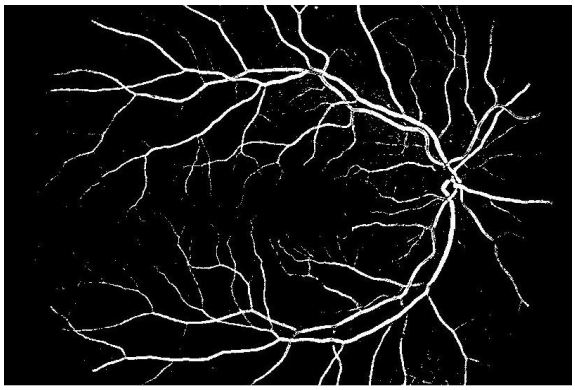
3. PRAHOVÁNÍ

Prahování (angl.. tresholding) představuje nezbytnou metodou při segmentaci obrazu [4]. Jak je vidět z obrázků 1 a 2 prahuje se šedotónový obraz a výsledkem je obraz binární. Při prahování je nejprve vytvořen takzvaný práh, který představuje číselnou hodnotu. Poté je každý pixel v obrázku porovnán s prahem. Pokud je hodnota pixelu větší nebo stejná jako hodnota prahu, stává se pixel popředím a obarví se bíle, jak je patrné na obrázku 2. Pokud je hodnota pixelu nižší než hodnota

prahu, stává se pixel pozadím a je obarven černě. Pro nalezení správné hodnoty prahu se používá Otsu metody [4], jež využívá histogramu popředí a pozadí.



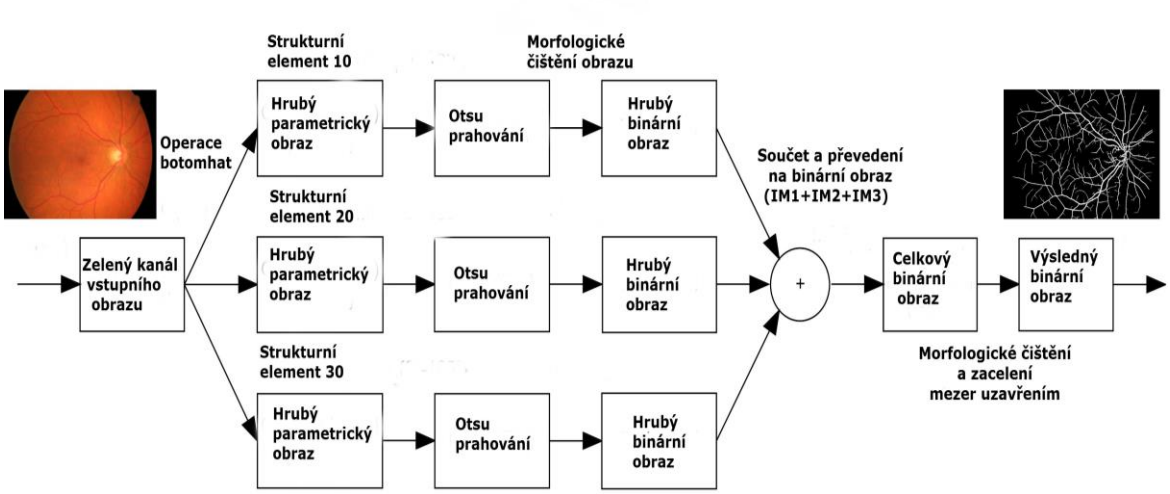
Obrázek 1: Snímek před prahováním



Obrázek 2: Snímek po prahování

4. POSTUP PRÁCE

Během práce byla nejprve z barevného RGB snímku sítnice (viz obr. 4) extrahována jeho zelená složka, která je pro další zpracování nejvhodnější z duvodu nejlepšího kontrastu mezi cévami a pozadím. Jak je patrné z blokového schématu (obr. 3), byl tento obraz následně třikrát podroben morfologické operaci *botomhat* [3], kterou představuje morfologické uzavření odečtené od původního obrazu, pokaždé s použitím většího strukturního elementu. Nejmenší strukturní element segmentoval nejužší cévy, střední prostředně široké a největší segmentoval ty nejširší. Tímto byly vytvořeny tří hrubé šedotónové obrazy (viz obr. 1). Ty byly dále upraveny prahováním, které z nich vytvořilo tzv. hrubý binární obraz, na kterém jsou cévy zvýrazněny bílou barvou, zatímco pozadí je černé (viz obr. 2). Pro prahování byla využita metoda Otsu, *viz kapitola prahování*. Pak bylo použito čištění šumu a nadbytečných fragmentů, které by působily nepřesnosti, použitím morfologické operace otevření oblasti [3]. Očištěné obrazy byly následně sečteny, jak to lze vydět ze schématu, čímž vznikl celkový binární obraz. V něm byly zaceleny defekty narušující cévy pomocí morfologického uzavření a opět byl odstraněn šum za pomoci otevření oblasti. Takto byl získán výsledný binární obraz (viz obr. 5), jež lze využít pro případnou diagnostiku chorob postihujících retinální cévy.

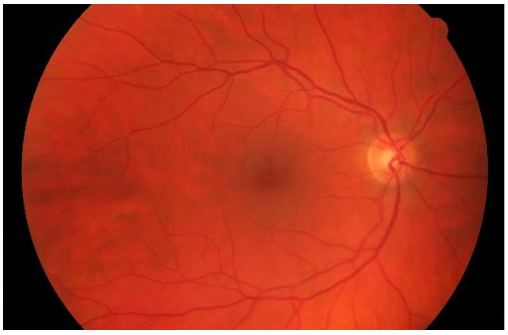


Obrázek 3: Blokové schéma graficky znázorňující jednotlivé etapy analýzy cévního řečiště.

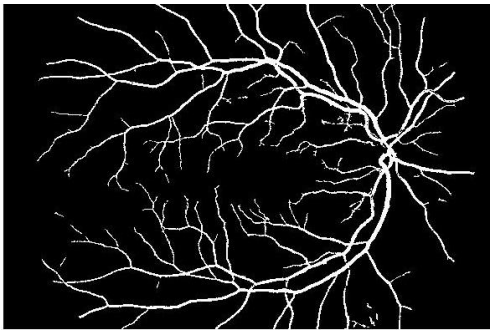
5. DISKUZE

Pomocí, v této práci, vytvořeného programu bylo segmentováno 45 snímků očního pozadí z databáze HRF (High-resolution fundus) [5], z toho 15 zdravých, 15 trpících glaukomem a 15 postižených diabetickou retinopatií.

Ohodnocení úspěšnosti metody proběhlo za pomocí standardizované databáze VUT. Výsledné obrazy byly porovnány se vzorovými snímky a na základě toho bylo určeno, které cévy jsou určeny správně, které úseky cév chybí a které artefakty byly chybně určeny jako cévy. Použitá metoda dosahovala celkové přesnosti 84% což představuje dostatečně dobrý výsledek. Celková přesnost je vyjádřena poměrem správně a chybně určených pixelů. Hodnota tedy udává, že jich 84% bylo určeno správně.



Obrázek 4: Původní snímek



Obrázek 5: Výsledný obraz segmentace

6. ZÁVĚR

V rámci práce byl vytvořen uživatelský program umožňující plně automatickou segmentaci cévního řečiště ze snímků sítnice. Pro implementaci metody bylo využito vývojové prostředí MATLAB verze R2008b a v něm implementovaná knihovna Image Procesing Toolbox. Úspěšnost metody byla ohodnocena porovnáním se vzorovými příklady z databáze VUT.

PODĚKOVÁNÍ

Tato práce byla provedena za finanční podpory Jihomoravského kraje v rámci SOČ pod vedením Jana Odstrčilíka.

REFERENCE

[1] GONZALES, Rafael C., Richard E. WOODS a Steven L. EDDINS. *Digital Image Procesing using MATLAB*. Second edition. United States of America: Gatesmark Publishing, 2009. ISBN 978-0-9820854-0-0.

[2] JOMIER, Julien, David K. WALLACE a Stephen R. AYLWARD. *Quantification of Retinopathy of Prematurity via Vessel Segmentation*. Computer-Aided Diagnosis and Display Lab The University of North Carolina at Chapel Hill, Department of Radiology 27510 Chapel Hill, USA.

[3] SOILLE, Pierre. *Morphological Image Analisis: Principless and Applications*. Second edition. Germany: Springer, 2003. ISBN 3-540-42988-3.

[4] FEICHTINGER, František. Studium vlivu parametrů vybraných segmentačních metod na analýzu obrazů buněk. Brno, 2008. Masarykova univerzita Fakulta informatiky.

[5] ODSTRCILIK, Jan, Radim KOLAR a Attila BUDAI. Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database. Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication, Brno University of Technology, Technicka 12, 61600, Brno, Czech Republic.

Bakalářské projekty

Elektronika a komunikace

ANALOG PHASE SHIFTERS OF FRACTIONAL ORDER

Radim Zedka

Bachelor Degree Programme (3), FEEC BUT

E-mail: xzedka01@stud.feec.vutbr.cz

Supervised by: Jiří Petržela

E-mail: petrzelj@feec.vutbr.cz

Abstract: This paper brings some general knowledge about linear analog fractional-order phase shifters. These useful building blocks can perform continuous phase shift with adjustable user-defined asymptotic values. Design process starts with predefined transfer function in which at least single fractional-order passive two-terminal device is needed. These elements are made as approximations in limited frequency range by using fully passive ladder structures where only resistors and capacitors are employed.

Keywords: fractional order calculus, CPE, phase shifter, two-port, two-pole, all-pass filter.

1. ÚVOD

Je známo, že existují tři druhy pasivních obvodových prvků, a to rezistor, kondenzátor a induktor. Jejich chování lze popsat matematicky pomocí lineárních a diferenciálních rovnic celistvých řádů. Jsou-li však řazeny do specifických druhů struktur, je možné, aby s omezenou přesností a kmitočtovým rozsahem aproximovaly ideální prvky, které lze popsat diferenciálními resp. integrálními rovnicemi necelistvých řádů. Tyto dvojpóly se často označují zkratkou CPE (constant phase element).

Každý CPE je přesně definován svými parametry, jimiž jsou řád α (v rozsahu od nuly do jedné), dále pak kmitočtový rozsah platnosti aproximace (určen mezními kmitočty), maximální zvlnění fázové charakteristiky a velikost impedance uprostřed pásma.

Při využití CPE v návrhu analogových posouvačů fáze je možné docílit změny fáze v rozsahu α·180 °, což rozšiřuje dosavadní konvenci celistvých násobků 180 ° pro všepropustné fázovací články.

2. POSOUVAČE FÁZE NECELISTVÝCH ŘÁDŮ S VYUŽITÍM CPE

2.1. NÁVRH CPE

Pro návrh a následnou konstrukci CPE se nejlépe osvědčila struktura paralelního spojení sériových R-C sekcí s korekčními paralelními prvky (dále jen struktura B2 viz Obrázek. 1). Ta koresponduje s metodou návrhu CPE na bázi geometrické řady [1], která byla upravena autorem textu pro více uživatelsky přívětivé parametry návrhu.

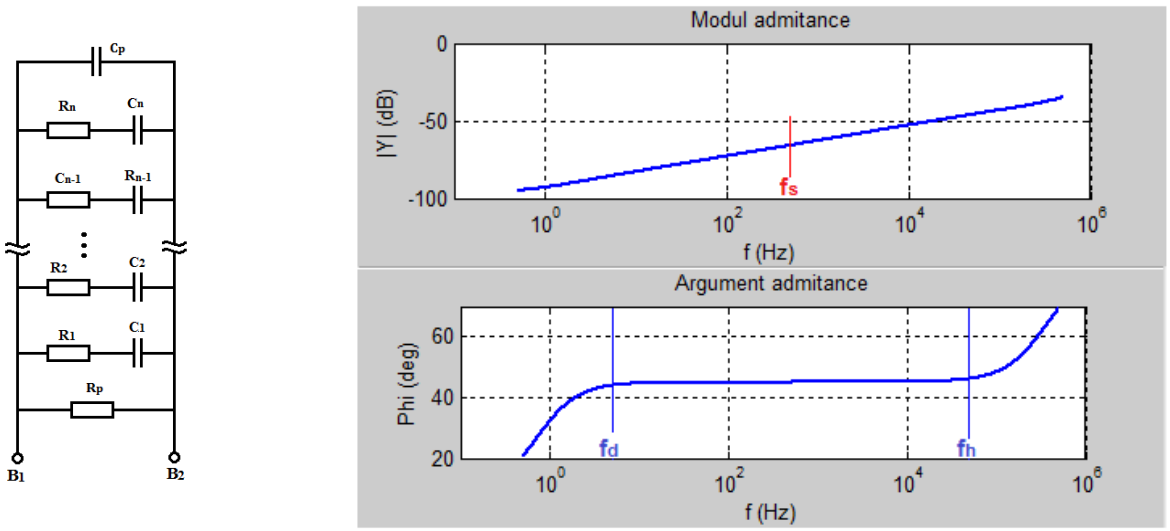
Pro jednodušší návrh zařízení s využitím CPE je možné jej popsat v kmitočtové doméně vztahem:

$$Y(s) = \frac{1}{D} s^{\alpha} \tag{1}$$

Ten je platný v omezeném kmitočtovém rozsahu, jenž je vstupním parametrem návrhu. Pro popis struktury B2 je použit vztah:

$$Y(s) = \frac{1}{R_p} + \sum_{i=1}^{n} \left(\frac{s\frac{1}{R_i}}{s+\frac{1}{R_i C_i}} \right) + sC_p \tag{2}$$

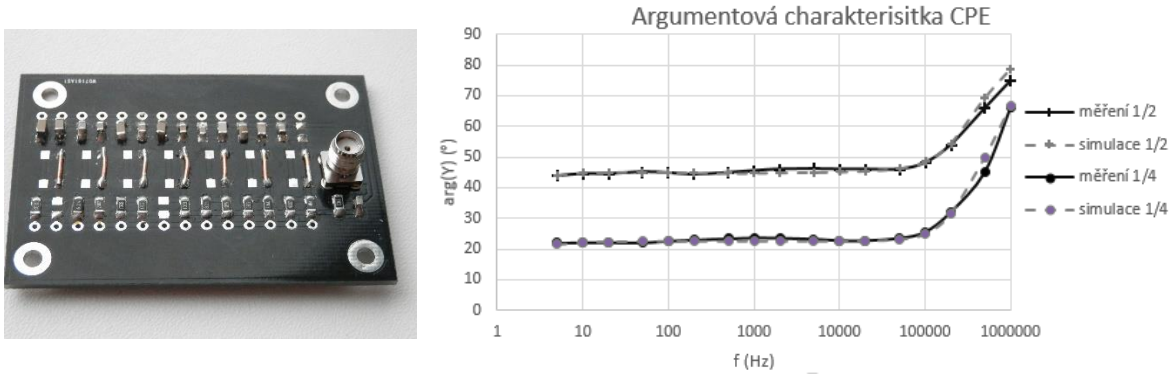
Na rozdíl od (1) je vztah (2) platný kmitočtově nezávisle, avšak neuvažuje parazitní vlastnosti součástek. Obrázek 1 graficky ilustruje vztah (2).



Obrázek 1: Struktura B2 a kmitočtové charakteristiky ideálního CPE.

2.2. REÁLNÁ KONSTRUKCE CPE

Pro realizaci byly použity rezistory SMD 1206 a kondenzátory SMD 1206 a 0805 o hmotách C0G (hodnoty do 10 nF) a X7R. Pro dosažení minimálního zvlnění fáze bylo použito 14 sekcí RC + paralelní dvojice R_p, C_p. Obrázek 2 obsahuje kmitočtové charakteristiky fáze pro řád $\alpha = \frac{1}{2}$ a $\frac{1}{4}$.



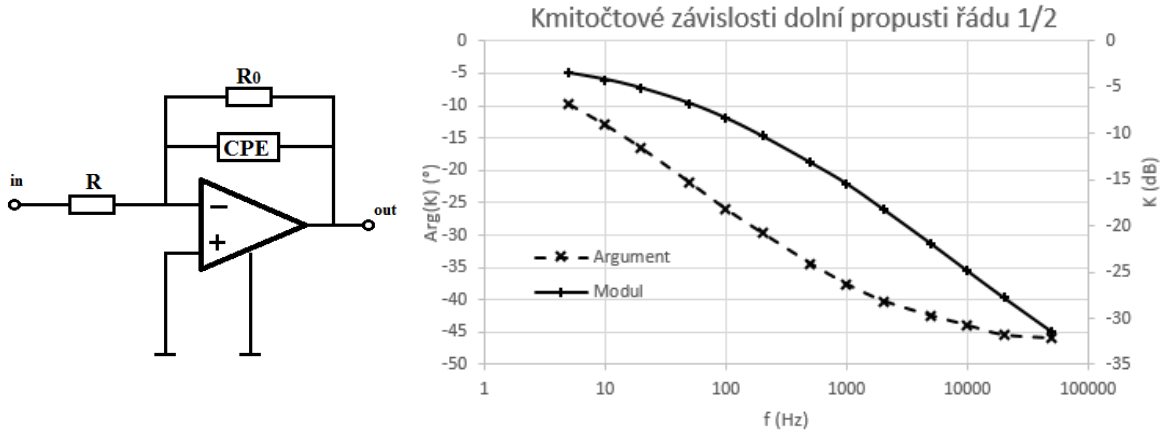
Obrázek 2: Realizace CPE a výsledky měření argumentu admitance.

Měření bylo provedeno pomocí osciloskopu měřením CPE v sérii s rezistorem, jelikož se jedná o nejvíce přímočarou metodu a nedochází k velkým systematickým chybám měření.

2.3. APLIKACE CPE V POSOUVAČÍCH FÁZE

Prvek CPE je možné využít v kombinaci s klasickými operačními zesilovači pro nízkofrekvenční aplikace, nebo s použitím moderních funkčních bloků (transkonduktory, proudové konvejory…) na vyšších kmitočtech (s ohledem na pracovní oblast CPE).

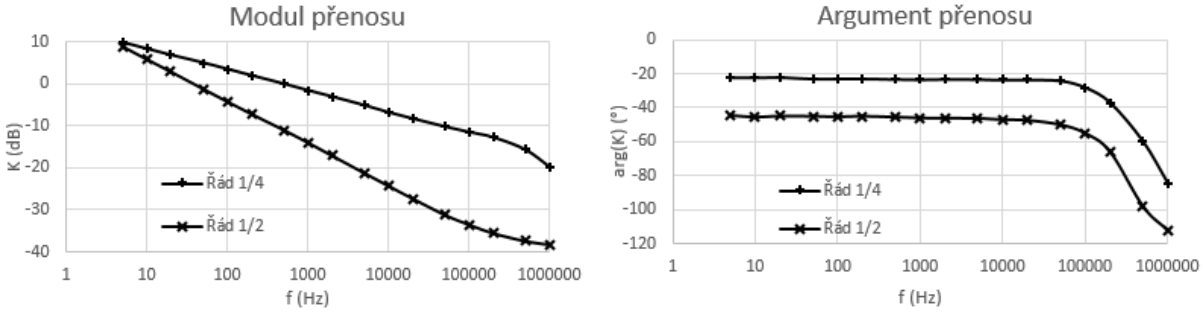
Nejjednodušší zapojení pro posuv fáze v rozsahu od 0 do -45 stupňů je aktivní dolní propust řádu ½ (viz Obrázek 3).



Obrázek 3: Dolní propust řádu ½ (reálné měření).

Na charakteristikách obrázku 3 je vidět tendence argumentu se ustálit na hodnotě $-45\,°$. Modul napěťového přenosu klesá se strmostí $-\alpha \cdot 20\ dB/dek$, čili $-10\ dB/dek$.

Další užitečnou aplikací je integrátor necelistvého řádu (viz obrázek 4).



Obrázek 4: Kmitočtové charakteristiky integrátoru řádu ½ a ¼ (reálné měření).

Fraktální integrátor resp. derivátor umožňuje konstantní fázový posuv o hodnotě $\alpha \cdot 90\,°$.

3. ZÁVĚR

Realizace obvodů s fraktální dynamikou je jistým narušením dosavadních konvencí v teorii obvodů. Jsou velmi užitečné v modelování dynamických systémů a svou úlohu plní i v oboru teorie chaosu. Jejich unikátní vlastností je však umožnění prakticky libovolného fázového posuvu signálu díky kaskádnímu spojování jednotlivých posouvačů fáze.

REFERENCE

[1] VALSA, J., DVORAK, P., FRIEDL, M. Network model of CPE. *Radioengineering*, 2011, vol. 20, issue 3, s. 59-67.

[2] PETRZELA, J. Arbitrary phase shifters with decreasing phase. In Proceedings of *37th International Conference TSP 2014*, s. 329-333.

[3] Rudolf Hilfer. *Applications of fractional calculus in physics*. World Scientific, c2000. ISBN 98-102-3457-0.

ANTENNAS DESIGN FOR ENERGY HARVESTING DEMONSTRATOR FOR GSM BAND

Petr Kaděra

Bechelor Degree Programme (3), FEEC BUT

E-mail: xkader05@stud.feec.vutbr.cz

Supervised by: Michal Mrnka

E-mail: xmrnka01@stud.feec.vutbr.cz

Abstract: The presented paper introduces the basic ideas and methods of wireless power transfer along with the design of antennas for energy harvesting demonstrator for use in GSM band with the aid of commercial available integrated circuit Powercast P2110B. The energy harvesting demonstrator should operate from - 12 dBm input power and utilize E-slot patch antenna or biplanar Yagi-Uda antenna for its work. The antennas were designed in CST Microwave studio, consequently fabricated and their input reflection coefficient measured and compared with simulation.

Keywords: energy harvesting, biplanar Yagi-Uda, E-shape patch antenna, GSM band

1. INTRODUCTION

The beginning of energy harvesting came from the idea of Nikola Tesla to transfer energy between two points without the need for physical connection to a power source. While several methods of wireless power have been introduced since Tesla´s work, including near-field magnetic resonance and capacitive coupling, only RF/microwave and laser-based systems appears to be applicable for long-range power transfer [1]. The scientist W. C. Brown made significant contribution in 1960´s with "rectennas", which are antennas able to directly rectify RF signal into DC[1]. Two major communities of researchers are today working on solar powered satellite (SPS) systems, which are focused on collecting solar energy in orbital stations around Earth and beaming it to the ground via microwave power transfer and RF identification (RFID) systems mainly used for tracking.

This paper is organized as follows. Section 2. presents the main blocks of energy harvesting system. Section 3. partially deals with the design of proper antennas and integrated circuit Powercast P2110B with supplemental parts. Section 4. presents the measured results of antennas. The last section 5. provides entire evaluation of the achieved results with suggestion to the further work.

2. ENERGY HARVESTER

The RF energy harvesting system (Fig.1) consists of an antenna or antenna array with linear or circular polarization, low losses and large bandwidth which impedance is subsequently matched by using band-pass filter that filters out high order harmonics resulting from the non-linearity of the rectifier. The rectifier is important part of the harvester which defines efficiency and sensitivity of the entire system. It is usually composed from specific topology of Schottky diodes and capacitors, e.g. the topology of charge pump allows increasing rectified DC output voltage to the higher level. In addition the bias point of the diode depends on the frequency and the level of received input power [1]. Instead of using charge pump, the step-up converter is often used in order to optimize and maximize power delivery to the storage medium such as supercapacitor or a battery. The load can be then connected to this power supply directly or in terms of the voltage regulator which provides constant voltage output for a longer time e.g. for sensitive applications [2].

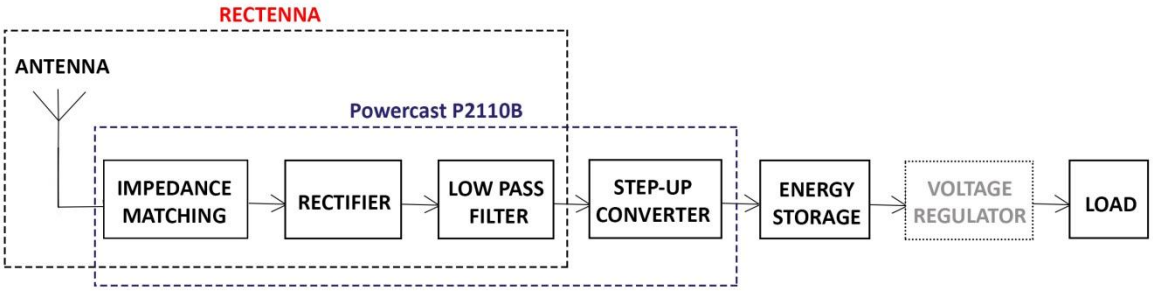


Figure 1: Block structure of RF energy harvesting system

3. ANTENNAS AND RF ENERGY HARVESTING MODULE DESIGN

The E-shape patch antenna was designed according to the approach described in [3] on the dielectric substrate Foamclad (h = 1.88 mm, ε_r = 1.25, tg δ = 0.0035) for covering the GSM band in range from 850 MHz to 985 MHz, the biplanar Yagi-Uda antenna was designed according to approach described in [4] on the dielectric substrate Arlon 25N (h = 1.524 mm, ε_r = 3.38, tg δ = 0.0025 for coverage of band from 890 MHz to 965 MHz. The simulated gain of these antennas is about 7-8 dBi vide Fig.2.

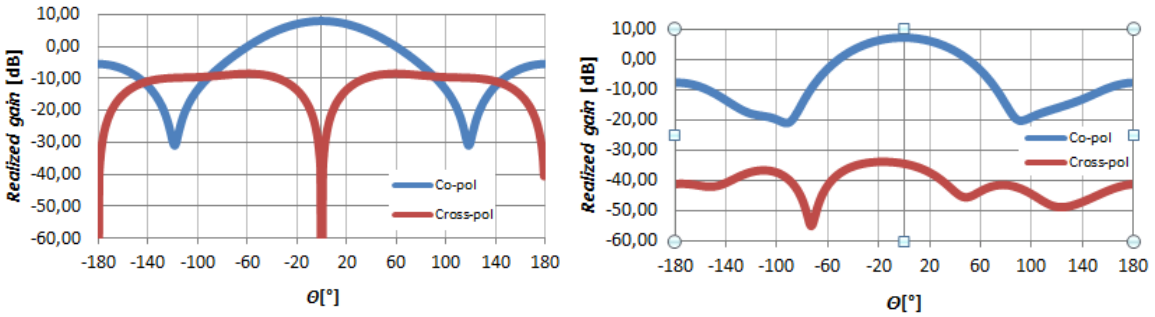


Figure 2: Simulated radiation pattern of E-shape patch antenna (left) and biplanar Yagi-Uda antenna (right) in E plane at frequency 915 MHz

The RF energy harvesting module (Fig.3) is designed to be assembled of a resistor R_SET which determines the output voltage, couple of capacitors C1, C2 with different capacities for testing the charging times and the output LED for checking activation of the circuit.

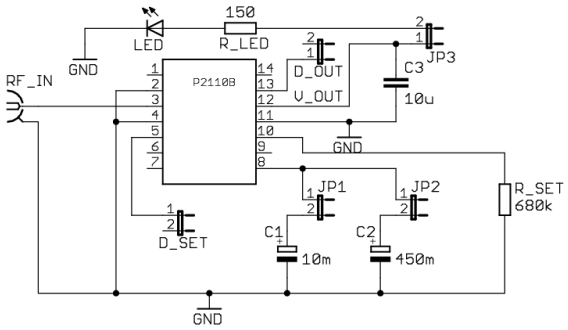


Figure 3: Circuitry of designed RF energy harvesting module

4. MEASURED RESULTS

Frequency response of the reflection coefficient (Fig.4) of previously described antennas was measured by a vector network analyzer Rohde & Schwarz ZVL13 and the measured traces of the

presented antennas have a similar shape to the simulation results, the differences could be caused by fabrication imperfections.

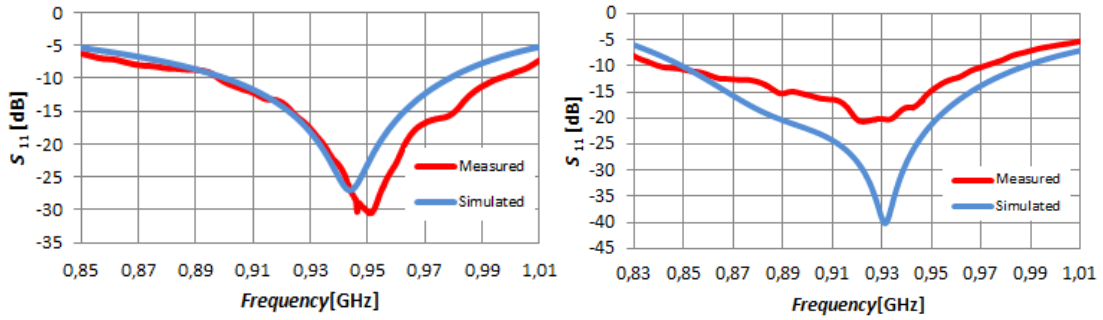


Figure 4: Measured and simulated reflection coefficient of biplanar Yagi-Uda antenna (left) and E-shape patch antenna (right)

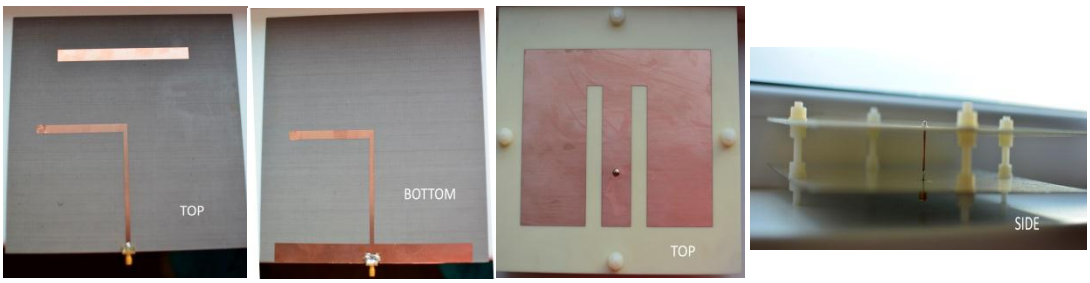


Figure 5: Fabricated biplanar Yagi-Uda (left) and E-shape patch (right) antennas

5. CONCLUSION

Two types of applicable antennas for energy harvesting demonstrator was proposed. The designed receiving antennas were manufactured, measured for the reflection coefficients and the data compared with simulation. Achieved impedance bandwidth of the E-shape antenna is 14 % and for the biplanar Yagi-Uda antenna 10.5 % for the reflection coefficient less than -10 dB (Fig.4). Manufactured prototypes of the antennas are shown in Fig.5. The energy harvesting module was designed as it is shown in Fig.3. The aim of the further work is to fabricate and check the function of the energy harvester module as the voltage power supply in different locations.

REFERENCES

[1] VALENTA, C. R., DURGIN, G. D. Harvesting Wireless Power: Survey of Energy-Harvester Conversion Efficiency in Far-Field, WirelessPower Transfer Systems. In: *Microwave Magazine*. IEEE, June 2014, vol. 15, no. 4, pp. 108-120.

[2] LEHTIMÄKI, S., LI, M., SALOMAA, J., et al. Performance of printable supercapacitors in an RF energy harvesting circuit. In: *International Journal of Electrical Power & Energy Systems*. June 2014, vol. 58, pp. 42-46.

[3] DESHMUKH, A. A., PHATAK, N. V., NAGARBOWDI, S., AHUJA, R. Analysis of Broadband E-shaped Microstrip Antennas. In*: International Journal of Computer Applications.* Oct 2013, vol. 80, no. 7, pp. 24-29.

[4] NAVARRO, E. A., BLANES, J. M., CARRASCO, J. A., REIG, C. Yagi-Like Printed Antennas for Wireless Sensor Networks. In: 2007 *Sensor Technologies and Applications*. SensorComm 14-20 Oct 2007. International Conference, pp. 254-259.

FLOOD ATTACKS GENERATION

David Hudec

Bachelor Programme (3), FEEC BUT

E-mail: xhudec27@stud.feec.vutbr.cz

Supervised by: David Smékal

E-mail: xsmeka11@stud.feec.vutbr.cz

Abstract: Proposal of a high speed packet flooding device is presented in this paper. The product is based on the FPGA (Field-programmable Gate Array) platform, developed in VHDL (Very high speed integrated circuit Hardware Description Language) and set into the NetCOPE environment. It uses an existing solution of packet generator. Once done, it should be implemented on a COMBO-80G board to serve as a network stressing tool.

Keywords: FPGA, VHDL, NetCOPE, Denial of Service, DoS, network attack, network tester, packet generator

1. INTRODUCTION

For the evaluation of current high speed packet switching networks' performance, high speed testing equipment is needed. Existing solutions are not suitable enough, as software traffic generators lack on throughput and appropriate hardware products are very expensive. FPGA platform software-hardware combination, as described in this paper, is capable of delivering good performance for a fraction of the costs spent on commercial hardware-based solutions, nevertheless providing great amount of configurability in the pack.

2. BACKGROUND

Concerns described in this section are essential to comprehend for understanding the project.

2.1. DENIAL OF SERVICE

Amongst many kinds of network intrusion, Denial of Service (DoS) attacks are one of the well-known time-proven ways of attacking a network-enabled device. Its simplicity lies in flooding with traffic so heavy the target cannot withstand, being forced to focus on excessive requests from the attacker and ceases being of any use for other, legitimate users. If the target is a server, its services will not be available for customers (for the time of the attack) and a financial loss is on the way. Great amount of data is therefore necessary. Enhancing this approach is possible by populating the attacker's base and distributing the disruptive force, thus creating a Distributed DoS.

2.2. FIELD-PROGRAMMABLE GATE ARRAY

When a computable task is not as big as to cause a standalone integrated circuit being designed for it, neither it is simple enough for conventional CPU systems to handle, FPGA boards are an option. It represents a technology of integrated circuitry designed to be configured by the customer using a hardware description language, hence programmable. The architecture contains a matrix of logic blocks, and a net of reconfigurable interconnects, allowing the blocks to be knitted together in many different ways. Thus, a unique junction of hardware-oriented solution and great configurability might be achieved, keeping it low-cost, yet effective.

Phases of *simulation*, *synthesis*, *bit stream generation* and *implementation* are what needs to be

completed before a VHDL application can run on an FPGA board.

2.3. NETCOPE

The NetCOPE framework [3] has been created to ease the development of network applications based on the FPGA platform. The firmware provided takes place between the user program core and input/output interfaces of the board used itself, making network communication programming much simpler and more transparent. As it is a living project, though, new versions of NetCOPE are being made, not necessarily keeping the backward compatibility.

2.4. HARDWARE

One COMBO-80G FPGA card is available to use, representing the powerful hardware needed. It contains eight identical RJ45 network interfaces supporting 10 Gb Ethernet technology each. The card can be mounted on a hosting computer via PCI-e bus, enabling sustainable high speed transfers between these two. Virtex-7 processing unit and 8 GB of RAM are installed.

2.5. FLEXIBLE, EXTENSIBLE, OPEN-SOURCE AND AFFORDABLE FPGA-BASED TRAFFIC GENERATOR

An existing open source packet generator called Flexible, extensible, open-source and affordable FPGA-based traffic generator (hereinafter referred to as the FEOSA generator) [1, 2] is used in this project. The solution implemented aims at great configurability, granted ability to fulfil a 10 Gb link even with the smallest packets and a high degree of modularity, hypothetically allowing for external usage, adaptation and extension. Although this is a highly functioning traffic generator, its development was conducted under conditions (2011 – 2013) different to ours (2016). This means that an older version of NetCOPE firmware was used on an earlier than the last generation of COMBO cards, resulting in various compatibility problems. Thus, implementing the FEOSA generator on the card currently available will most likely require a considerable amount of time and effort.

3. IMPLEMENTATION

This project's main goals are:

- To alter the FEOSA generator, so it can be compiled, implemented and run on modern FPGA boards. This includes remapping all ports used in the top entity file (project part responsible for mapping software-described pins on their hardware counterparts, physically present on the board), adjusting signals used in it and make it runnable on COMBO-80G card. This task is plausible.
- To transfer the environment used in FEOSA generator onto current version of NetCOPE framework. Knowledge of both versions of the software might be necessary to compare and find the discrepancies, which is in no matter guaranteed possible. Because of that, this step may prove difficult.
- To configure the generator to create DoS-like flood packets, once the previous steps are done. Due to its open-source nature and the creators having modularity in mind while developing it, this step is supposedly simple.
- To synthetize, compute and implement the final solution on the board provided. This does not stand any problem, if previous steps were completed successfully.
- To use it as a test tool, capable of stressing the device under test as much as necessary.

3.1. PROGRESS

Currently, first two phases of the plan as described above are being executed. Obstacles regarding the board differences (step 1) are time consuming to solve, yet appear that can be coped with. Concerns with the NetCOPE framework evolution are hard to grasp, providing no results or certainties whatsoever. Hard work and concentration on the matter are ahead. So far, a functioning standalone copy of the current NetCOPE environment has been successfully loaded into Vivado Design Suite [4] and examined using the test benches provided, as well as those modified by me exactly for the purpose. Same procedure was applied to the original FEOSA generator (not yet connected to NetCOPE), resulting in further understanding of its structural parts and hierarchy, along with confirming the functionalities.

The current task comprises of combining the genuine generator top file (*traffic_generator.vhd*) and the top file of the NetCOPE package (*application_core.vhd*) into a new, merged architecture. The purpose of the first one is to receive the configuration from hosting computer, set its parts accordingly, control the process of packet generation and send the outcome to the board's output interfaces. The latter one's task is to connect all the available interfaces and provide an example application. This is the code, which this project intends to replace with the contents of the traffic_generator file mentioned. The result of this combination would be able of both controlling the older entities in the generator, and communicating with the necessary newer parts of the environment. Many compatibility errors arise during this task, however all of those found so far have been managed and solved one by one. In theory, this should lead to a fully working project made from both of the parts running almost separately, despite both being controlled by one file.

4. CONCLUSION

Design and ideas presented in this paper are yet subject to development, hence it is not possible to present their results and achievements. As it seems, all the steps should be doable, meaning that a modern network high speed stress testing device / flood attack tool should be created, using an open-source software running on a widely available hardware, keeping the solution reasonably expensive.

ACKNOWLEDGEMENT

The project is being conducted at The Faculty of Electrical Engineering and Communication, Brno University of Technology, under the supervision of Ing. David Smékal. I must express my gratitude to him.

REFERENCES

[1] Groléat, Tristan. GitHub. *Hardware-traffic-generator*. [Online] 11. 3 2016. https://github.com/twisterss/hardware-traffic-generator/tree/master/hw/traffic_generator.

[2] HAL. *Flexible, extensible, open-source and affordable FPGA-based traffic generator*. [Online] 11. 3 2016. https://hal.archives-ouvertes.fr/hal-00859291/document.

[3] NetCOPE Technologies [Online] 26. 3. 2016. http://www.netcope.com.

[4] Xilinx, All programmable. *Vivado Design Suite*. 23. 6. 2016 http://www.xilinx.com/products/design-tools/vivado.html.

INTEGRATED MULTICHANNEL AUDIO AMPLIFIER FOR HOME CINEMA

Ondřej Fišer

Bachelor (3), FEEC BUT

E-mail: fiseon@gmail.cz

Supervised by: Tomáš Kratochvíl

E-mail: kratot@feec.vutbr.cz

Abstract: This paper describes complete design of amplifier for home cinema system. Described amplifier has 2 digital inputs (TOSLINK and S/PDIF) and single analog input. With this inputs it is possible to connect amplifier to a source of surround audio 5+1 signal. Paper further discuss design of correction circuits, preamplifier, power amplifier and power source. The whole is controlled by 8-bit microprocessor and has a graphics output implemented through LCD display. The goal of this paper is to design and build high quality amplifier.

Keywords: Home theatre, DSP, AES3, Surround sound, ATmega

1. ÚVOD

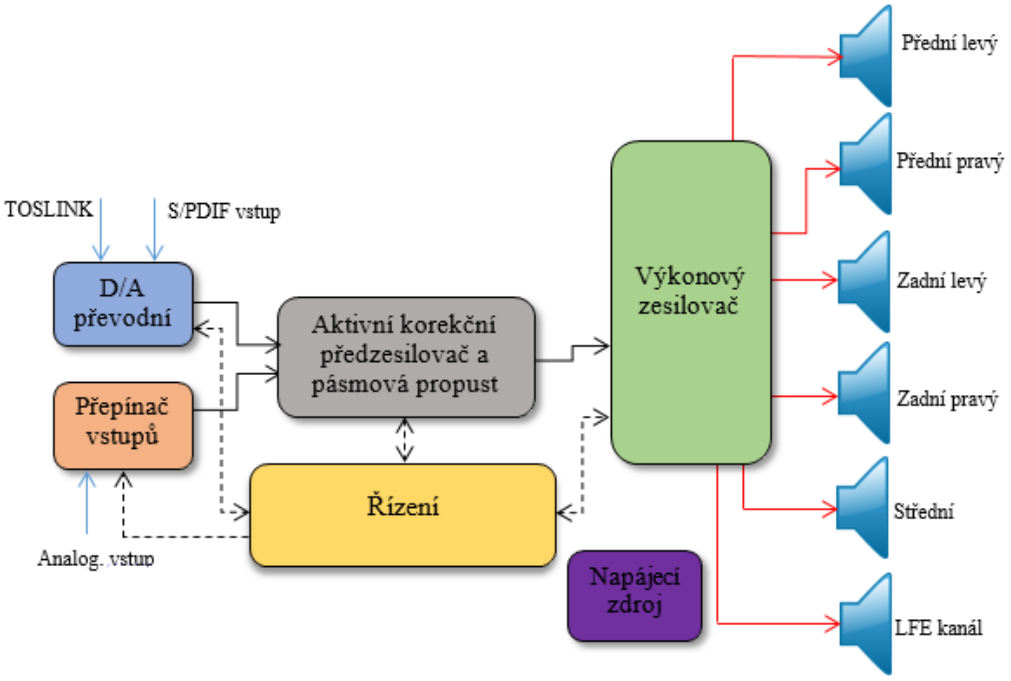
Pojem domácí kino označuje komplexní systém k reprodukci videa a zvuku ve filmové či koncertní kvalitě v domácích podmínkách. Je kladen velký důraz na to, aby divák zažil stejný nebo minimálně srovnatelný zážitek jako v kině. Systém je tvořen zpravidla kvalitním LCD televizorem s velkou uhlopříčkou nebo projektorem a velmi kvalitním zesilovačem se soustavou reproduktorů zapojených v prostorovém rozložení např. Dolby Digital 5.1. V dnešní době je na trhu nepřeberné množství „zesilovačů pro domácí kina“ v různých cenových i kvalitativních kategoriích. Mnohdy bohužel cena neznamená kvalitu.

Tento zesilovač je navržen s důrazem na co nejvěrnější reprodukci. Práce se zabývá návrhem a realizací digitálních obvodů pro zpracování prostorového šesti kanálového digitálního zvuku a převod na analogový signál následnou ekvalizací a koncovým stupněm ve třídě D v konfiguraci reproduktorů 5.1.

2. ROZBOR A REALIZACE

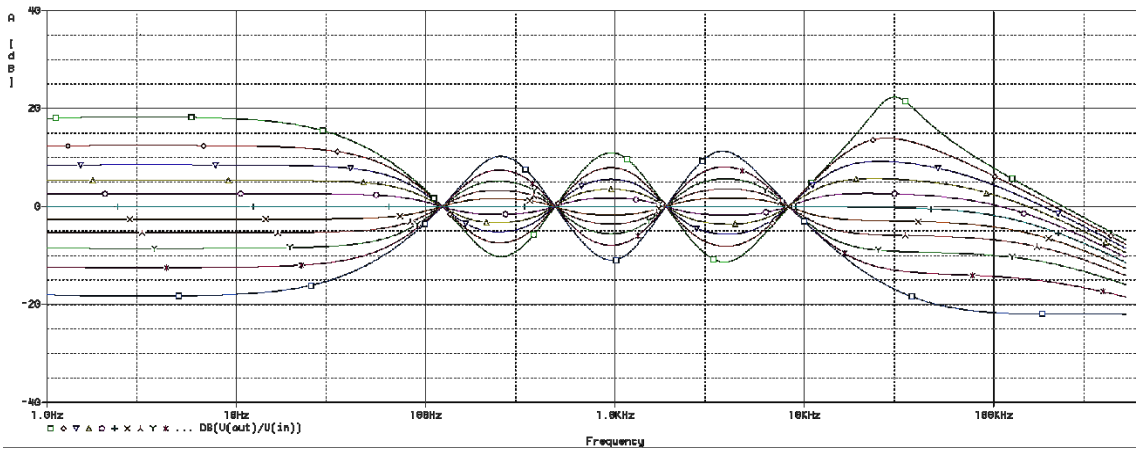
Po důkladné analýze celé problematiky [1], [2], [3], [4] a [5] bylo vytvořeno blokové schéma (viz. obrázek 1). Obvodová část D/A převodníku je tvořena digitálním signálovým procesorem STA310 [2], který zpracovává vstupní digitální signál a následně pak pomocí D/A převodníku PCM1680 [4] je převeden analogový signál. Oba vstupy (TOSLINK i S/PDIF) splňují standard IEC-61937, to znamená, že je k nim možné připojit až 6-ti kanálová audio data (vzorkovací kmitočet 44,1 kHz – 192 kHz s rozlišením 24 bitů) se ztrátovou kompresí ve formátech DTS a Dolby Digital.

Mimo digitálních vstupů je možné audio signál připojit i pomocí klasického analogového vstupu, který díky použití precizních OZ OPA4134 má impedanci cca 100 TΩ. Šířka pásma všech analogových obvodů je 70 kHz. Další klíčovou částí zesilovače je korekční 5-ti pásmový předzesilovač (viz. obrázek 2), který slouží ke zvýraznění či potlačení některých kmitočtových pásem podle potřeby posluchače a regulaci hlasitosti. Základním prvkem tohoto ekvalizéru jsou syntetické induktory [5]. Součástí korekčního stupně je také aktivní propust 3. řádu naladěná na 150 Hz pro LFE kanál.



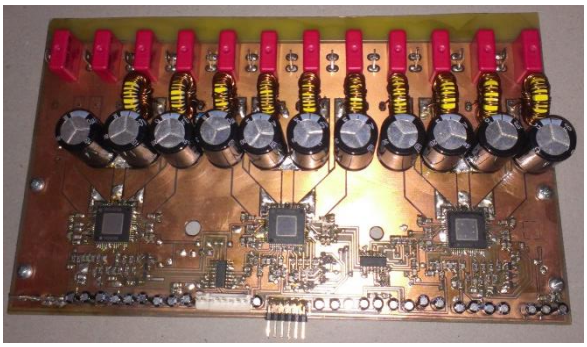
Obrázek 1: Blokové schéma zesilovače

Srdcem zesilovače je koncový stupeň tvořený třemi zesilovači TAS5630B [3] ve třídě D od společnosti Texas Instruments. Díky použití těchto integrovaných obvodů je možné dosáhnout velmi velkého výstupního výkonu 300 W/4 Ω na kanál a přitom celý koncový stupeň je relativně malý. Díky velké účinnosti nejsou nutné masivní chladiče.

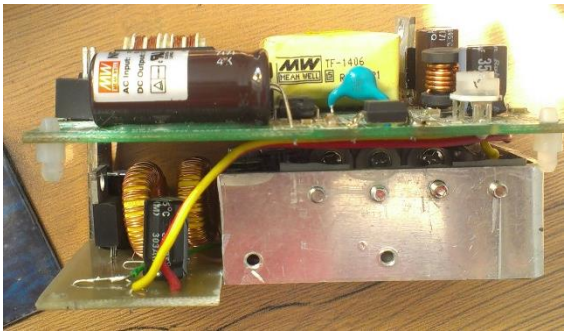


Obrázek 2: Charakteristika pěti-pásmového ekvalizéru

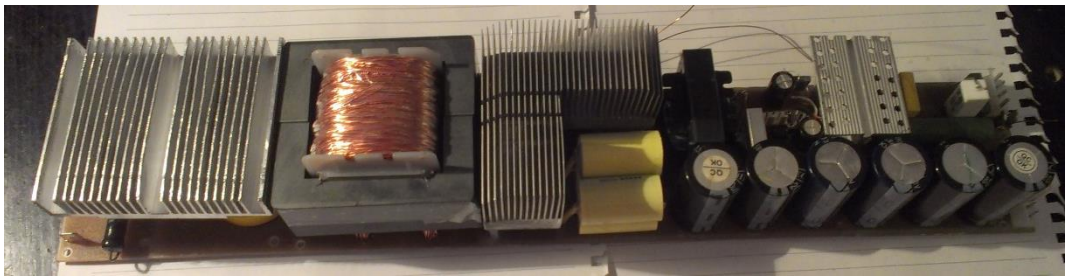
Napájení je zajištěno pomocí dvojice spínaných zdrojů. O napájení všech částí kromě koncového stupně se stará pomocný zdroj, který vytváří napětí 17 V; 12 V; 5 V; 3,9 V; 3,3 V; 2,5 V; -5 V; -3,9 V a -2,5 V. Výkonový zdroj, který napájí koncový stupeň, transformuje síťové napětí na 50 V s maximálním výkonem okolo 2400 W. Zdroj využívá topologie jednočinného propustného zdroje se dvěma spínači (IGBT tranzistory) neboli polo řízené můstkové zapojení. O řízení se stará integrovaný obvod UC3845.



a)



b)



c)

Obrázek 3:

Vlevo nahoře: deska koncového stupně a),
Vpravo nahoře: deska pomocného zdroje b),
Dole: deska výkonového zdroje c).

3. ZÁVĚR

V současné době je zcela oživena deska pomocného zdroje a výkonového zdroje (viz. Obrázek 3) včetně odladění nežádoucích rušení a dále probíhá práce na desce koncového stupně a návrh desky korekcí, která obsahuje i procesorovou část. Souběžně probíhá vývoj uživatelsky příjemného firmware pro řídící procesor a návrh dřevěného boxu na celé zařízení.

REFERENCE

[1] S/PDIF [online] – [cit. 13. prosince 2015]. Dostupné na www: https://en.wikipedia.org/wiki/S/PDIF

[2] STMicroelectronics STA310 datasheet [online]. STMicroelectronics 2003 [cit. 13. prosince 2015]. Dostupné na www: <http://datasheet.octopart.com/STA310-STMicroelectronics-datasheet-33008842.pdf>.

[3] TAS5630B [online] Texas Instrumets 2015: Analog input 300W Stereo class D amplifier with integrated feedback [cit. 13. prosince 2015]. Dostupné na www: <http://www.ti.com/product/tas5630b>.

[4] PCM1680 [online] Texas Instrumets 2008: 105dB SNR 8-Channel Audio DAC, 24-Bit, 192-kHz Sampling, Enhanced Multilevel, Delta-Sigma, Audio D/A Converter (Rev. B) [cit. 13. prosince 2015]. Dostupné na www: <http://www.ti.com/product/pcm1680>.

[5] 5-ti pásmový nf ekvalizér [online]. Jaroslav Belza – [cit. 13. prosince 2015]. Dostupné na www: <http://www.belza.cz/audio/eq5b.htm>.

MINI-WHIP ANTENNA

Tomáš Pařízek

Bachelor (3), FEEC BUT

E-mail: xpariz07@stud.feec.vutbr.cz

Supervised by: Zbyněk Raida

E-mail: raida@feec.vutbr.cz

Abstract: The thesis deals with the impedance matching of electrically very short antennas to the real resistence 50 Ω . The antenna is considered as an input of an emitter follower , which is plays the role of the impedance transformer. This follower transforms large complex input impedance, which is characteristic for electrically short antennas, a real output resistance. The output impedance of assistance was matched by matching circuitry to 50 Ω in the frequency band of operation.

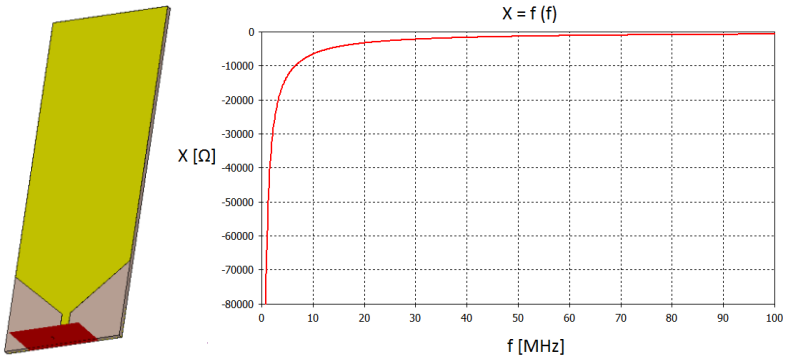
Keywords: Whip antenna, impedance matching, input impedance, emitter follower

1. ÚVOD

Vstupní impedance elektricky velmi krátkých antén má velice vysokou reaktanční složku kapacitního charakteru [1]. Vybraná elektricky krátká anténa byla simulována v programu CST Microwave studio, kde byla vyjádřena její impedance, která byla následně uvažována jako vstup emitorového sledovače. Emitorový sledovač slouží jako impedanční převodník [2], který tuto velkou kapacitní reaktanci převede na mnohem nižší hodnotu impedance, která má navíc převládající reálnou složku. Takto transformovanou impedanci je již poměrně snadné přizpůsobit k reálné hodnotě 50 Ω.

2. VSTUPNÍ IMPEDANCE PRUTOVÉ ANTÉNY VÝJÁDŘENA POMOCÍ CST

V simulátoru CST Microwave Studio byla simulována elektricky krátká prutová anténa realizována jako vodivá měděná plocha na substrátu FR-4 (Obr. 1). Délka antény byla 100 mm a její šířka se rovnala 40 mm. Měděná plocha byla na vstupu zkosena kvůli postupné změně impedance. Druhá strana substrátu byla pokryta měděnou vrstvou, jejichž délka je definovaná délkou mikropásku. Tato zemní plocha omezuje symetrické proudy a zároveň zastává funkci reflektoru.



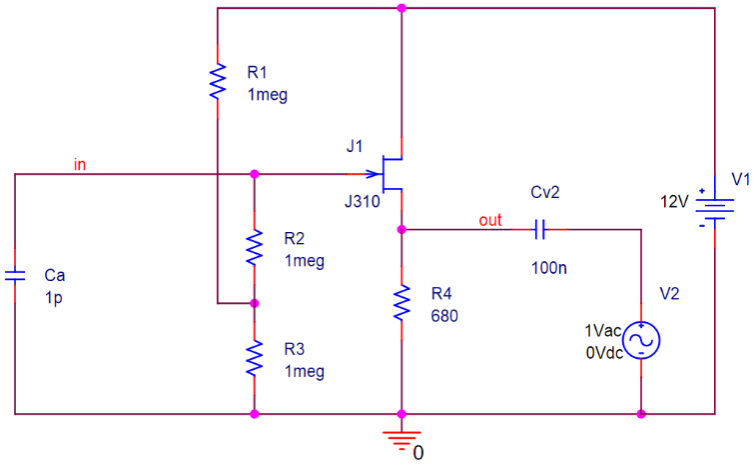
Obrázek 1: Motiv antény a výsledná vstupní reaktance antény simulovaná v CST

Vstupní impedance antény byla zjišťována ve frekvenčním rozsahu 100 kHz až 100 MHz. Reálná složka vstupní impedance je téměř nulová, proto ji lze zanedbat. Imaginární část impedance rapidně roste se snižující se elektrickou délkou, což je patrné z grafického výstupu CST (Obr. 1).

3. TRANSFORMACE VSTUPNÍ IMPEDANCE EMITOROVÝM SLEDOVAČEM

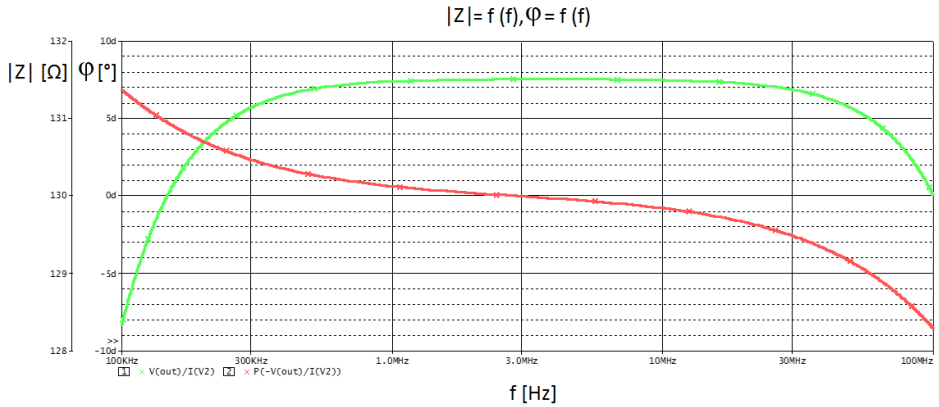
Zapojení se společným kolektorem se vyznačuje tím, že je schopno pracovat jako impedanční převodník. To znamená, že dokáže velkou impedanci na vstupu transformovat na malou hodnotu impedance na výstupu. To samé dokáže i s komplexní hodnotou impedance na vstupu. Navíc je-li na vstup připojena komplexní impedance, na výstupu je tato impedance transformována na reálnou hodnotu. Imaginární složka, která se objeví na výstupu je velice malá.

Pro simulaci bylo zvoleno zapojení emitorového sledovače s unipolárním tranzistorem J310 (Obr. 2), tento tranzistor pracuje do frekvencí cca 150 MHz, což je pro danou anténu dostačující. Jeho vstupní odpor je na nižších frekvencích v řádech gigaohmů, poté začíná klesat. Pracovní bod je nastaven odporovým děličem do hradla. Napětí U_{GS} je nastaveno přibližně na polovinu napájecího napětí. Na vstup emitorového sledovače byla připojena čistě imaginární impedance, kterou představuje kapacitor s hodnotou 1 pF. Hodnota kapacitoru byla extrahována pro střední kmitočet 50 MHz. Reálná složka byla zanedbána.



Obrázek 2: Schéma emitorového sledovače v PSPICE

V obvodovém simulátoru PSPICE byla provedena střídavá analýza ve frekvenčním pásmu 100 kHz až 100 MHz. Modul a fáze impedance na výstupu emitorového sledovače je vynesena do společného grafu na obrázku č. 3.

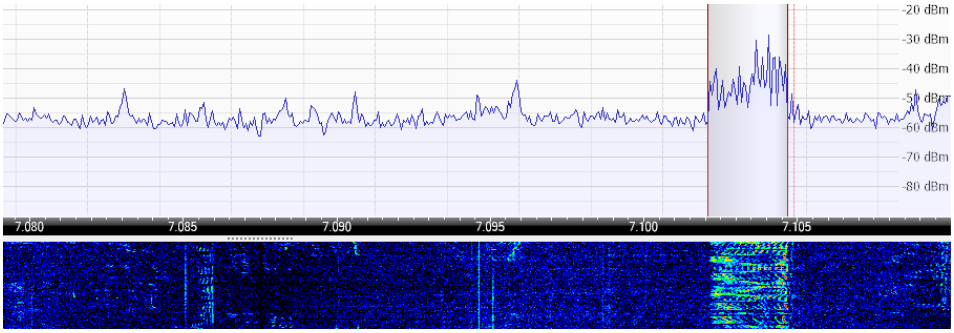


Obrázek 3: Impedance na výstupu emitorového sledovače

Zelená křivka zobrazuje kmitočtový průběh modulu výstupní impedance, červená křivka zobrazuje fázi. Modul výstupní impedance je přibližně 130 Ω. Fáze výstupní impedance se s frekvencí mění od 7° do -8°. Na frekvenci 3 MHz je fáze nulová. Následně lze pro danou frekvenci impedančně přizpůsobit k 50 Ω například vhodným „L“ článkem.

4. TESTOVÁNÍ ANTÉNY

Při testování bylo uvažováno, že takto malá anténa je převážně určena do „domácích“ podmínek, kdy není možnost využít klasických SV a KV antén. Anténa tedy byla vystrčena z okna přibližně 70 cm od stěny domu, výška nad zemí byla 5m. K testování příjmu byl využit SDR přijímač FUNcube Dongle Pro +. Pomocí SDR přijímače je možno zobrazit požadované spektrum signálu v reálném čase, také je možno určit jeho úroveň. Ihned po připojení antény bylo možné přijímat radioamatérské stanice, časové signály, AM a FM rozhlasy a další služby. Obrázek číslo 4 zobrazuje spektrum na pásmu 7 MHz přijímané Mini-Whip anténou. Ve spektru je vidět SSB signál italské radiostanice, který má špičkovou úroveň -30 dBm.



Obrázek 4: Část spektra na 7 MHz

5. DESING MINI-WHIP ANTÉNY

Mini-Whip anténa je realizovaná na substrátu FR-4, kdy byla stejně jako u simulace zvolena délka 10 cm a šířka 4 cm. Impedanční převodník je umístěn přímo na desce s motivem antény. Anténa má díky své velice krátké elektrické délce malou účinnost, proto je vhodná převážně pro příjem. Pro vyšší zisk byl za tranzistorové zapojení zařazen vf. zesilovač MAR-6, který má zisk okolo 20 dB.



Obrázek 5: Design Mini-Whip antény

6. ZÁVĚR

Tento článek popisoval využití emitorového sledovače v anténní technice jako další možný přizpůsobovací člen. Experimentálně bylo ověřeno, že anténa je vhodná pro příjem radioamatérských provozů na středních vlnách, kde bylo dosaženo srovnatelného příjmu jako s mnohem delšími drátovými anténami. Výhodou těchto antén je tedy možnost srovnatelného příjmu v podmínkách, kdy není možné využít dlouhé antény pro SV a KV. Větší účinnosti lze dosáhnout zvýšením elektrické délky antény například meandrem nebo Hilberovými křivkami.

REFERENCE

[1] BALANIS, C.A. *Antenna theory: analysis and design*. 3rd ed. Hoboken: Wiley-Interscience, 2005, xvii, 1117 s. ISBN 978-0-471-66782-7.

[2] BAKKER, R. The pa0rdt-Mini-Whip. [online]. Netherlands, 2007 [cit. 2015-11-07]. Dostupné z: http://pi4utr.nl/wp-content/uploads/download/pa0rdt_whip.pdf.

THE DESIGN AND REALISATION OF METAMATERIAL

Soňa Kovaľová

Bachelor Degree Programme (3), FEEC BUT

E-mail: xkoval11@stud.feec.vutbr.cz

Supervised by: Dušan Nešpor

E-mail: nespord@feec.vutbr.cz

Abstract: The work discusses periodic resonant structures with negative permeability (metamaterials) as an improvement of NMR tomograph resolution. The aim of work is to design and adjust magnetoinductive lens for application in nuclear magnetic resonance imaging. The software Ansys HFSS was used for finite element method analysis of designed structure.

Keywords: Metamaterial, periodic resonant structure, electromagnetic wave, permitivity, permeability, magnetoinductive lens, nuclear magnetic resonance imaging, finite element method

1. ÚVOD

Metamateriály sú kompozitné materiály, ktorých stavebnou jednotkou je mikropásikový rezonátor, s rozmermi oveľa menšími v porovnaní s vlnovou dĺžkou žiarenia, ktorým ich ožarujeme (tzv. obvod so sústredenými parametrami). Táto podmienka zabezpečí, že jav lomu bude dominovať nad rozptylom, objekt bude považovaný za homogénny. Špecifickou vlastnosťou metamateriálov je záporná permitivita a permeabilita (z ε-μ diagramu pomenovanie „left-handed“), dôsledkom čoho dochádza k odkloneniu Poyntingovho vektora (**S**) na opačnú stranu od osi rozhrania dvoch prostredí (left-handed a right-handed). Tým je spôsobený fázový posun 180° medzi fázovou a skupinovou rýchlosťou šíriacej sa vlny a z toho vyplývajúci záporný index lomu. Snellov zákon matematicky pripúšťa dve riešenia. Zápornú permeabilitu je možné realizovať pomocou prerušovaných prstencových rezonátorov, zápornú permitivitu pomocou vodivých pásikov. Superpozíciou týchto objektov vzniká left-handed materiál. Hypotetickými prejavmi týchto materiálov je napríklad reverzia Doplerovho efektu, Snellovho zákona, zámena divergentných a konvergentných efektov u konvexných a konkávnych LH šošoviek. [1],[2],[4]

2. NUKLEÁRNA MAGNETICKÁ REZONANCIA

Táto zobrazovacia technika spočíva v uložení skúmanej vzorky do magnetického poľa o indukcii $\mathbf{B_0}$ a následnom vysielaní rádiového pulzu cez vzorku pomocou kovovej cievky. Spiny atómových jadier sa odklonia od zarovnania v smere magnetického poľa, čím v jadrách vzniká magnetický moment indukujúci prúd v cievke, poskytujúci informáciu o chemickej povahe vzorky. Pomocou detekčnej cievky sa mapujú odchýlky vo fáze a frekvencii rádiových vĺn absorbovaných tkanivom spôsobené precesným pohybom častíc – vytvárajú merný magnetický tok. Rozlíšenie NMR tomografie je limitované veľkosťou budiacej a detekčnej cievky. [3]

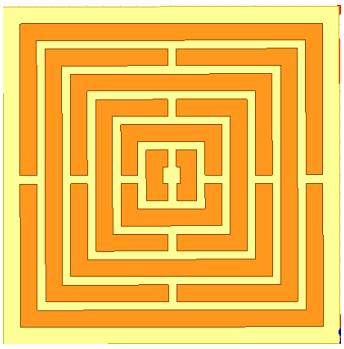
Magnetické pole je tlmené so zväčšujúcou sa hĺbkou vniku do tkaniva. Navyše, výskyt objektov s vysokou permitivitou a vodivosťou rozladuje cievku a zhoršuje činiteľa akosti Q. Vhodným prispôsobením EMG poľa môže byť dosiahnuté zlepšenie zobrazovacích schopností NMR. Riešením tohto problému sú vďaka laditeľnej permitivite a permeabilite homogénne metamateriály. Konkrétne zariadenia na úpravu EMG poľa sa nazývajú magnetoinduktívne šošovky. [4]

3. NÁVRH

Najskôr bola navrhnutá geometria základnej stavebnej jednotky – jedného rezonátora. Následne bol prispôsobený tvar a rozmery tak, aby bola dosiahnutá rezonancia na pracovnej frekvencii NMR tomografu 198,75 MHz. Rezonátor bol umiestnený do vlnovodu naplneného vzduchom, kde bol ožarovaný rovinnou EMG vlnou. Simuláciu rovinnej vlny sme dosiahli budením vlnovodu v porte 1 a následným šírením vlny do portu 2. Kolmosť vektorov **E** a **H** bola zabezpečená okrajovými podmienkami PEC (perfect electric conductor) a PMC (perfect magnetic conductor).

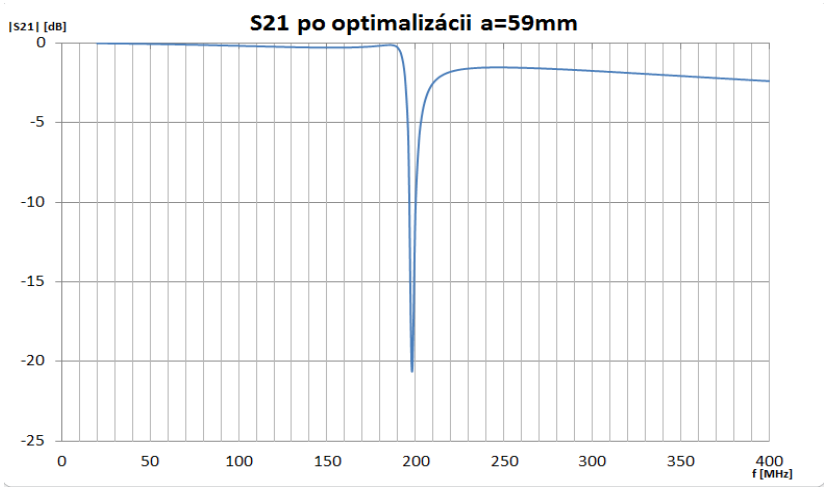
3.1. OPTIMALIZÁCIA

Boli preskúmané rôzne rozmiestnenia medených plôch rezonátora na substráte. Vzor na obrázku 1 bol zvolený pre ďalšie ladenie. Po preskúmaní vplyvu rôznych substrátov najvhodnejšie vlastnosti preukázal kremík s $\varepsilon_r = 11{,}6$ a hrúbkou $h = 2\,mm$, rezonujúci pri $f_r = 829\,MHz$. Z Thomsonovho vzťahu vyplýva, že pre zníženie rezonančnej frekvencie je potrebné zvýšiť kapacitu alebo indukčnosť rezonátora. Zavedením kondenzátora do štrbiny medzi vodivými plochami a zvyšovaním jeho kapacity bola rezonančná frekvencia znížená. Po prekročení 50 pF ($f_r = 588\,MHz$) bol pozorovaný neočakávaný jav – f_r sa viac nezmenšila, pravdepodobne kvôli transformácii prídavnej kapacity do celkovej kapacity rezonátora. Tento jav je pozorovaný v štruktúrach s komplikovanou geometriou. Zvýšenie indukčnosti nie je vhodné, spôsobilo by nárast strát a zhoršenie kvality rezonátora. Požadovaná f_r bola dosiahnutá rozmerovým ladením – dĺžka strany štvorcovej štruktúry bola zadaná parametricky. Hodnota 60 mm sa ukázala ako najvhodnejšia. Šírka medených plôch je 3 mm a šírka medzier je 0,5 mm. Rezonancia nastala na $f_r = 198\,MHz$.

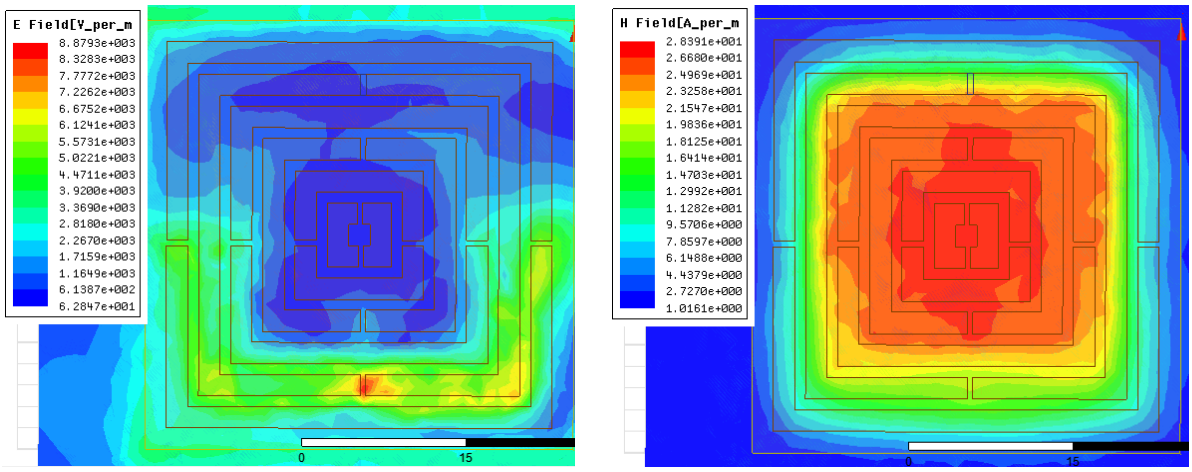


Obrázok 1: Stavebná jednotka metamateriálovej štruktúry

3.2. VÝSLEDKY ANALÝZY



Obrázok 2: Rezonancia znázornená pomocou parametra S_{21}



Obrázok 3: Rozloženie intenzity elektrického a magnetického poľa v rezonancii

4. ZÁVER

Bol zostavený numerický model základnej jednotky metamateriálovej šošovky. Prispôsobením rôznych parametrov bol naladený tak, aby rezonoval pri 198 MHz. Veľkosť vzduchového okolia v uvedenom prípade vplýva na ladenie rezonátora, zvyšuje f_r, avšak po predbežných simuláciach magnetoinduktívnej šošovky v programe Ansys HFSS sa predpokladá, že tento negatívny vplyv bude možné vykompenzovať usporiadaním rezonátorov v šošovke.

V súčasnej dobe sa pracuje na realizácii, po ktorej bude nasledovať experimentálne overenie funkčnosti návrhu. V prípade zhody s numerickým modelom bude zhotovená magnetoinduktívna šošovka tvorená skupinou rezonátorov.

REFERENCIE

[1] Cui, T.J., Smith, D.R., Liu, R. Metamaterials: Theory, Design ans Applications. In: Springer, Science+Business Media, LLC 2010. ISBN978-0-470-37704-8.

[2] Marqués, R., Martín, F., Sorolla, M. Metamaterials with Negative Parameters. John Wiley & Sons, Inc., Hoboken, New Jersey. Rok 2007. ISBN 2007017343.

[3] Nešpor, D. Numerické modelování periodických struktur. Dizertační práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Vedoucí práce Doc. Ing. Pavel Fiala, Ph.D. Brno, 2014, 90 s.

[4] Freire, M, J, Marqués, R, Jelinek, L. Experimental demostration of a μ = 1 metamaterial lens for magnetic resonance imaging. Applied Physics Letters, Departamento de Electronica y Electromagnetismo. Universidad de Sevilla.2008, vol. 93, no. 23, p. 231108 - 231108-3. ISSN 0003-6951.

VHF SSB RECEIVER WITH LOW CONSUMPTION

Martin Matoušek

Bachelor Degree Programme (3), FEEC BUT

E-mail: xmatou24@stud.feec.vutbr.cz

Supervised by: Tomáš Urbanec

E-mail: urbanec@feec.vutbr.cz

Abstract: This work is focused on proposal of SSB receiver operating at 30 MHz. The design was primarily focused on simplicity and low power consumption. This work was realized for the transmission of audio signals. SSB modulation is far more efficient in terms of the radio spectrum used. First part of this thesis describes a block diagram of a receiver. Next parts are focused on proposal of individual blocks of the receiver, especially its most important parts. Finally, the overall evaluations of the design characteristics of SSB receivers are discussed.

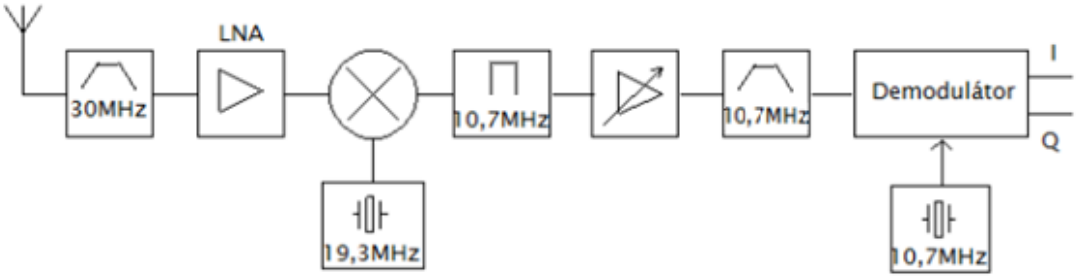
Keywords: SSB, low noise amplifier, automatic gain control, AD607.

1. ÚVOD

SSB přijímače slouží k příjmu jednoho z postraních pásem, buďto dolního postranního pásma LSB nebo horního postranního pásma USB. SSB modulace je velmi oblíbená u radioamatérů, jelikož využívá malou šířku pásma. Tato práce se zabývá návrhem SSB přijímače pracujícím na kmitočtu 30 MHz, u kterého je kladen důraz na jednoduchost a nízkou spotřebu. Velkou nevýhodou SSB modulace je však její nepoužitelnost pro vysoké přenosové rychlosti, proto je práce realizována pro přenos obecného audio signálu. Práce nejprve popisuje blokového schéma přijímače. Následně se věnuje nejdůležitějším částem návrhu jednotlivých bloků přijímače a celkovým zhodnocením vlastností navrženého SSB přijímače.

2. SSB PŘIJÍMAČ

SSB přijímač je realizován jako superheterodyn. Při návrhu byl kladen důraz na nízkou spotřebu celého přijímače a malé napájecí napětí, tak aby přijímač mohl být napájen baterií. Celé blokové schéma přímače je uvedeno na obrázku 1.



Obrázek 1: Blokové schéma přijímače.

Přijímané signály z antény vstupují přímo do pásmové propusti, která vybere pásmo signálů se kterými chceme pracovat. Nejdůležitější funkce vstupní pásmové propusti je potlačení zrcadlového kmitočtu. Zrcadlový kmitočet v našem případě vychází na 8,6 MHz. Dalším duležitým parametrem je útlum v propustném pásmu. Útlum pásmové propusti přímo souvisí s šumovým číslem. Pásmo vybraných signálů je následně zesíleno v nízkošumovém zesilovači (low noise amplifier, dále jen LNA). LNA je realizován pomocí tranzistoru BF840, který nejlépe splňoval

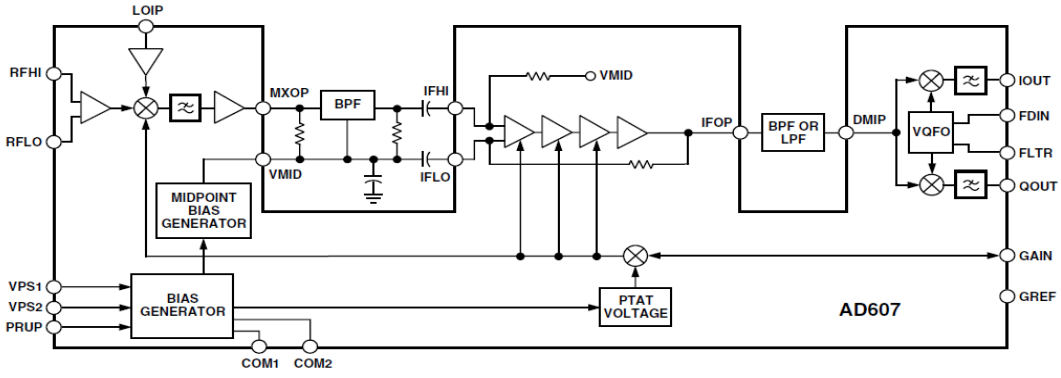
požadavky ze snadno dostupných tranzistorů pro navržení LNA. Pro zvolený tranzistor BF840 výrobce neuvádí S-parametry, proto byl návrh s výkonovým přizpůsobením vytvořen pomocí simulace v programu PSpice. Při návrhu přijímače je kladen důraz na nízkou spotřebu. Proto i při návrhu LNA bylo cílem dosažení nízké spotřeby. Spotřeba LNA byla zvolena na 2 mA, tak aby LNA bylo odolné proti rušení. Přičemž zvolení spotřeby vycházelo z katalogového listu tranzistoru. Duležitým úkolem LNA je dosáhnout dostatečného zesílení a malého šumového čísla, aby potlačil vliv šumových vlastností následujicích bloků [1].

Dále signál směšujeme na mezifrekvenční kmitočet 10,7 MHz. Na mezifrekvenčním kmitočtu hned za směšovačem je selektivní pásmová propust. Selektivní pásmová propust je realizovaná čtyřpólovým krystalovým filtrem 10,7 MHz, s jehož pomocí vybereme přesnou šířku pásma, se kterou chceme pracovat. Signál v mezifrekvenci se následně zesiluje zesilovačem s řízeným ziskem. Za zesilovačem musí být pásmová propust, pro potlačení intermodulačních produktů vznikajících v zesilovači. Zesílený signál dále vstupuje do demodulátoru, kde se převádí do základního pásma. Výstup z demodulátoru je v I a Q rovině. Přijímač pracuje s audio signálem. Proto pracujeme jen s výstupem I [1].

Příjímač obsahuje dva oscilátory, 19,3 MHz pro získání mezifrekvence a 10,7 MHz pro získání základního pásma signálu. Oba oscilátory jsou realizovány v Colpittsově zapojení. Protože je třeba přesných kmitočtů, oba oscilátory jsou řízeny krystaly. Pro realizaci byly zvoleny tranzistory BFR92. Zvolená spotřeba každého oscilátoru je 1 mA, při zvolení spotřeby se vycházelo z katalogových listů tranzistorů, tak aby byla zaručena funkčnost a stabilita při dosažení nejmenší spotřeby [2].

2.1. INTEGROVANÝ OBVOD AD607

V přijímači je použit integrovaný obvod AD607. Jeho výhodou je malá spotřeba, která se pohybuje okolo 8,5 mA. Obvod obsahuje integrovaný aktivní řízený směšovač, řízený zesilovač (atomatic gain control, dále jen AGC) a aktivní demodulátor. Regulace zesílení řízeného směšovače a AGC se provádí pomocí detektoru obálky. Zesílení můžeme regulovat pro směšovač od 1,5 do 21,5 dB a pro AGC od 7,5 – 67,5 dB, zesílení demodulátoru je pevné 18 dB. Mezi další výhody patří malé vstupní napětí signálu pro vybuzení směšovače, jehož maximální hodnota je ± 54mV (-15 dBm). Při této hodnotě směšovač pracuje v bodu jednodecibelové komprese. Blokové schéma AD607 s typickým zapojením filtrů je na obrázku 2 [3].



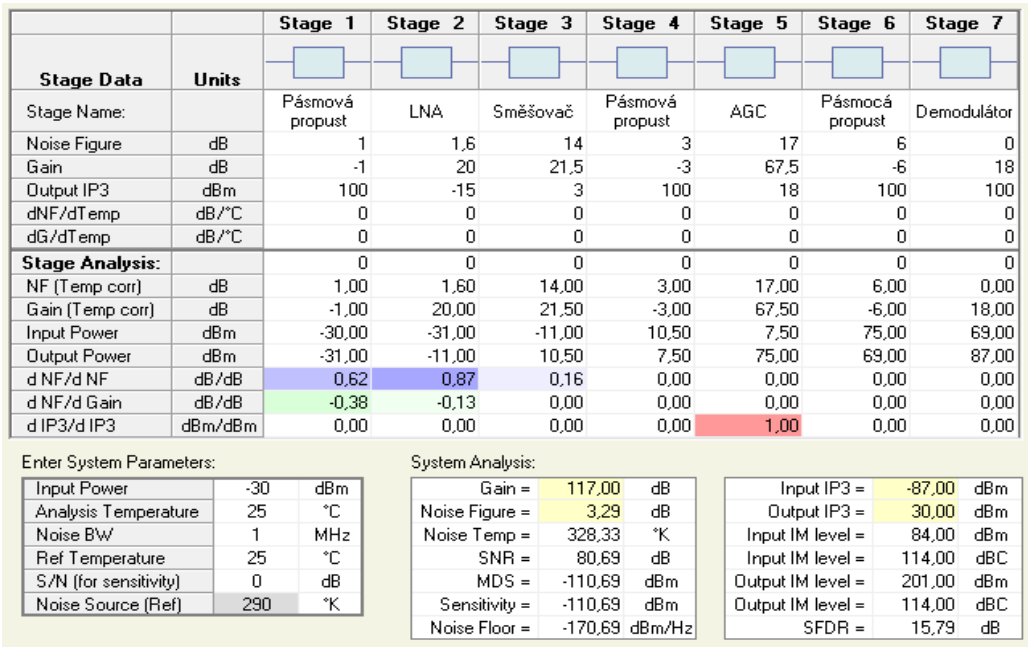
Obrázek 2: Blokové schéma AD607 [3].

Obvod AD607 využívá stejnosměrný posun pro VF signál. Hodnota posunu VF složky se zpravidla volí na polovinu hodnoty napájecího napětí. Pomocí stejnosměrného posunu VF složky lze posunout rozsah řízení napětí, pro řízení zesílení. Závislost zesílení na řídícím napětí je lineární [3].

2.2. SIMULACE CELÉHO PŘIJÍMAČE

Celkové chování přijímače je simulováno v programu AppCAD od firmy Agilent. Na obrázku 3 jsou uvedeny významné parametry jednotlivých bloků přijímače a také celkové parametry. Ze si-

mulovaných hodnot jsou vidět celkové parametry přijímače. Celkové šumové číslo určuje pásmová propust a LNA.



Stage Data	Units	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Stage 7
Stage Name:		Pásmová propust	LNA	Směšovač	Pásmová propust	AGC	Pásmocá propust	Demodulátor
Noise Figure	dB	1	1,6	14	3	17	6	0
Gain	dB	-1	20	21,5	-3	67,5	-6	18
Output IP3	dBm	100	-15	3	100	18	100	100
dNF/dTemp	dB/°C	0	0	0	0	0	0	0
dG/dTemp	dB/°C	0	0	0	0	0	0	0
Stage Analysis:		0	0	0	0	0	0	0
NF (Temp corr)	dB	1,00	1,60	14,00	3,00	17,00	6,00	0,00
Gain (Temp corr)	dB	-1,00	20,00	21,50	-3,00	67,50	-6,00	18,00
Input Power	dBm	-30,00	-31,00	-11,00	10,50	7,50	75,00	69,00
Output Power	dBm	-31,00	-11,00	10,50	7,50	75,00	69,00	87,00
d NF/d NF	dB/dB	0,62	0,87	0,16	0,00	0,00	0,00	0,00
d NF/d Gain	dB/dB	-0,38	-0,13	0,00	0,00	0,00	0,00	0,00
d IP3/d IP3	dBm/dBm	0,00	0,00	0,00	0,00	1,00	0,00	0,00

Obrázek 3: Simulace přijímače.

3. ZÁVĚR

Celý přijímač je navržen pro bateriové napájení 3,3 V. Jeho celková spotřeba by neměla přesáhnout 17,5 mA. Přijímač byl navrhován postupně podle bloků v blokovém schématu. Vstupní pásmová propust byla navržena pro anténu s impedancí 50 Ω. Návrh LNA (nízkošumového zesilovače) se nepodařilo provést s optimálními parametry. Největší problém při návrhu LNA byl výběr vhodného tranzistoru z komerčně dostupných produktů. Právě tento blok nejvíce ovlivňuje šumové číslo celého přijímače. Důvodem použití obvodu AD607 je malé napájecí napětí. Tento obvod obsahuje směšovač, řízený zesilovač a demodulátor. Výhodou integrovaného směšovače je tedy malé vstupní napětí a řízené zesílení. Dle výsledků simulovaného přijímače je celkové šumové číslo 3,3 dB. Pro zmenšení šumového čísla by bylo možné použití jiného tranzistoru. Vhodný tranzistor by měl mít shodné, případně větší zesílení na 30 MHz a nižší šumové číslo.

PODĚKOVÁNÍ

Děkuji vedoucímu projektu Ing. Tomáši Urbancovi, Ph.D. za metodickou, pedagogickou a odbornou pomoc a cenné rady při zpracování mého projektu.

REFERENCE

[1] VÁGNER, Petr. *Vysokofrekvenční technika* [online]. Skripta. Brno: FEKT: VUT v Brně, 2013 [cit. 2016-03-06].

[2] PAVLOVEC, Josef a ŠRAMAR, Jaroslav. Krystalové jednotky a oscilátory. *Amatérské Rádio* [online]. 1987 [cit. 2016-03-06]. Dostupné z: http://www.krystaly.cz/files/download/Amaterske_radio_1987-02-Krystaly.pdf

[3] *AD607: Datasheet* [online]. In: Analog Devices, 2002 [cit. 2016-03-06]. Dostupné z: http://www.analog.com/media/en/technical-documentation/evaluation-documentation/AD607.pdf

Bakalářské projekty

Zpracování signálů, obrazu a dat

ARDUINO BASED SERIAL BUS DECODER

Juraj Giertl

Bachelor Degree Programme (3), FEEC BUT

E-mail: xgiert01@vutbr.cz

Supervised by: Ondřej Krajsa

E-mail: krajsao@feec.vutbr.cz

Abstract: This article describes a development of device used for capturing data from different serial buses such as SPI, UART and I2C. Device is design as shield to support Arduino boards. The Shield includes four fully isolated inputs and five non isolated imputs for capturing data, Ethernet interface, SD card and two display interfaces witch supports up to ten displays.

Keywords: Arduino, dekoder, I2C, logic analyzer, serial busses, SPI, UART

1. ÚVOD

Dekodér sériových zberníc je zariadenie, ktoré umožní používateľovi zachytiť a zobraziť dáta, ktoré si medzi sebou posielajú pomocou sériovej zbernice napríklad procesor a analógovo-digitálny prevodník. Takýto dekodér má uplatnenie napríklad pri návrhu digitálnych systémov, kedy je možné komunikáciu odchytiť a následne verifikovať dáta.

Cieľom tejto práce je navrhnúť dekodér, ktorý by bol schopný zachytiť dáta a užívateľovi umožniť prehliadanie dát priamo na zaradení, prípadne niekoľko ďalších možností uloženia a prenosu do PC cez počítačovú sieť alebo USB.

V dnešnej dobe už existuje veľa zariadení čo dokážu dekódovať sériové dáta, napríklad drahšie digitálne osciloskopy [2]. Z lacnejších zariadení možno spomenúť obľúbene USB analyzátory, ich nevýhoda je, že vždy musia byť pripojené k počítaču, kde je spustený program, ktorý dáta spracováva.

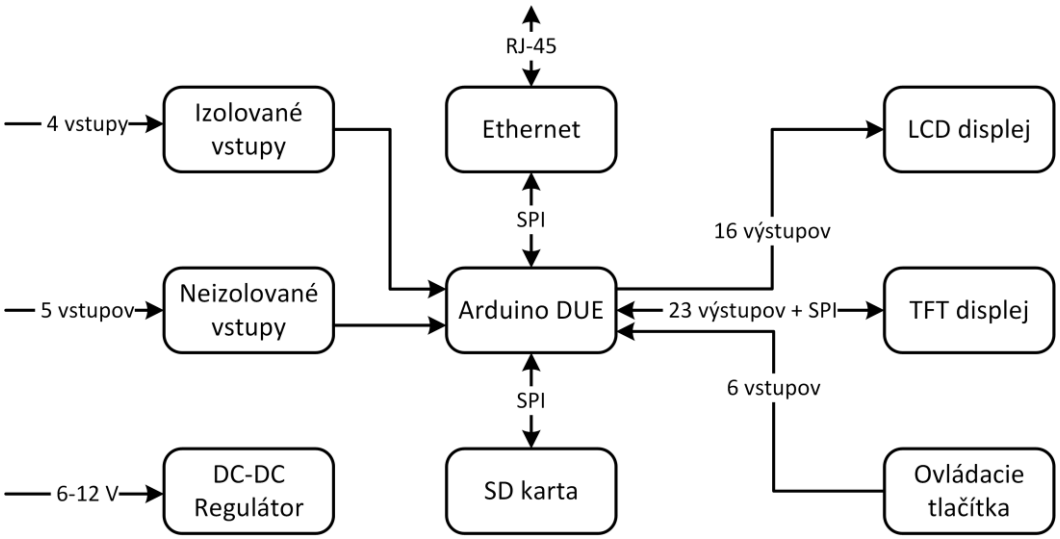
Dekodér, ktorý je popisovaný v tejto práci, je schopný dekódovať a zobrazovať dáta samostatne, bez žiadnych ďalších prostriedkov. Podporované zbernice sú SPI, I2C, UART. Ich podrobnejší popis sa nachádza v semestrálnej práci [1].

2. HARDVÉR A NÁVRH OBVODU

Ako je vidieť z obrázku 1 blokovej schémy zariadenia, celé je postavené okolo platformy Arduino DUE. Kľúčovou časťou pre dekodér je blok štyroch izolovaných vstupov, vďaka ktorým je možno bezpečne pripájať meranú zbernicu. Ďalej sú vyvedené aj neizolované vstupy pre prípadné rozšírenie. Zariadenie disponuje taktiež ethernetovým rozhraním, ktoré na fyzickej vrstve podporuje 10Ba- seT/100BaseTX ethernet.

Možnosť pripojenia dotykových TFT LCD displejov s radičom SSD1289, v rozmeroch od troch po desať palcov. Taktiež grafické LCD displeje s radičom KS0108B. Pre uloženie dát zariadenie obsahuje slot na pamäťovú kartu.

Vďaka tomu, že zariadenie je navrhnuté modulárne, je ho možné použiť nielen ako dekodér sériových zberníc, ale napríklad aj ako analyzátor sieťovej komunikácia a podobne. Záleží len na programovom vybavení použitého mikroprocesora.



Obrázok 1: Bloková schéma zariadenia

Kompletný návrh zariadenia je popísaný v semestrálnej práci [1].

3. SOFTVÉR

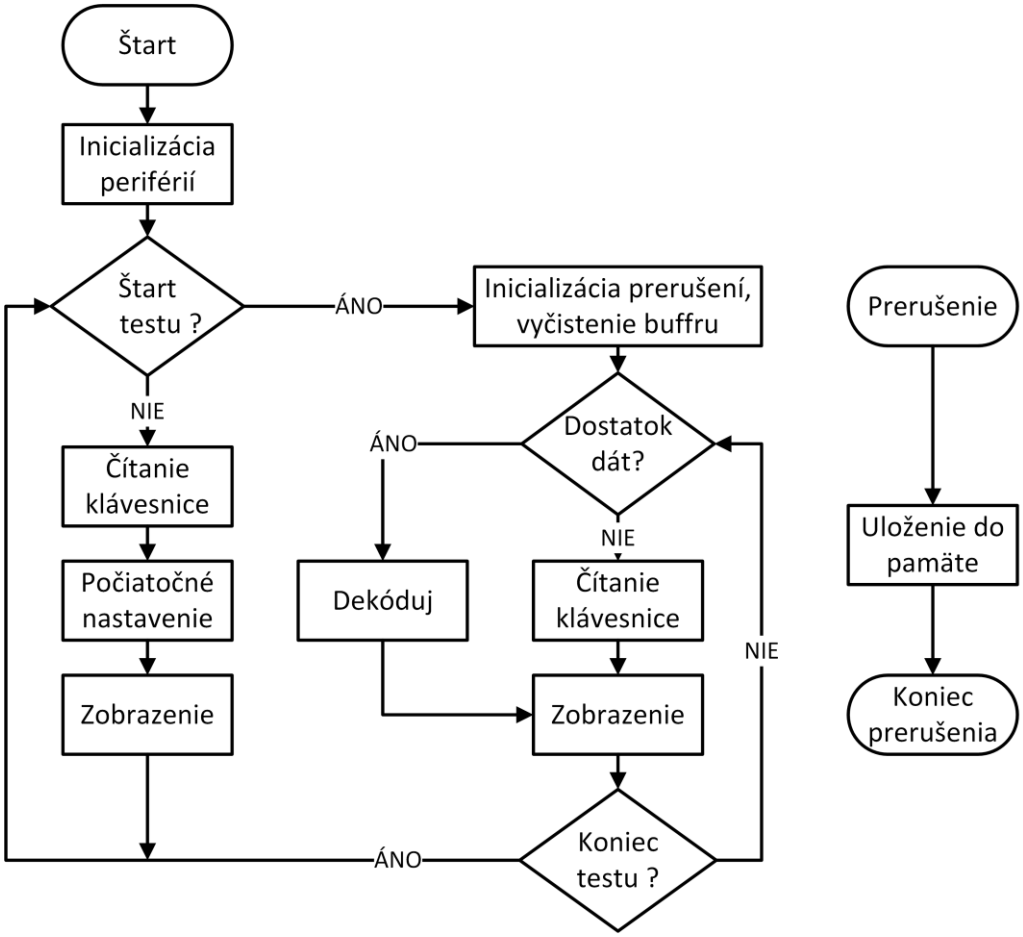
Vývojový diagram programu pre mikroprocesor je na obrázku 2. Program pracuje v dvoch základných slučkách. Po spustení zariadenia program skočí do prvej slučky, ktorá obsluhuje všetky základné nastavenia a následné spustenie dekódovania na základe nastavených parametrov. Po spustení dekódovania sa nastavia potrebné periférie a spracúvajú sa dáta. Prerušiť spracovávanie dát môže užívateľ, alebo sa ukonči automaticky po dosiahnutý nastaveného limitu.

3.1 NAČ ÍTANIE DÁT ZO VSTUPOV

Aby som dosiahol najživšiu rýchlosť a spoľahlivosť zachytenia zmeny na zbernici, rozhodol som sa použiť externé prerušenie. Vyvolané prerušenie bude volať krátku funkciu, ktorá uloží aktuálnu zmenu pinu a čas, kedy bola zmena uskutočnená. Č as je generovaný pomocou čítača implementovaného v mikroprocesori. Pri odčítaní dát pomocou prerušenia, dáta obsahujú kompletnú informáciu o sledovanom digitálnom signáli. Z pohľadu zabratej pamäte je tento spôsob tiež výhodnejší, lebo je uchová len udalosť, ktorú vyvolá zmena vstupného signálu a čas kedy nastala. Pri klasickom vzorkovaní signálu by sa zbytočne plnila pamäť rovnakými vzorkami a tento proces by bol časovo náročný pre mikroprocesor, pretože by musel periodicky odčítať hodnotu zo vstupov. Ďalší z dôvodov prečo použiť prerušenie namiesto vzorkovania je, že mikroprocesor musí dokázať v reálnom čase spraco- vať a zobraziť dáta. Tento proces je náročný na čas, a preto by bol problém zaistiť konštantný čas medzi vzorkami.

4. TECHNICKÉ PARAMETRE

Dekodér používa 32-bitový ARM Cortex-M3 mikroprocesor taktovaný na 104 MHz. Pre načítanie dát sa používa dynamicky alokovaný kruhový buffer do veľkosti 64 kB. Na neizolovaných vstupoch dekodér pracuje s 3,3 V logikou, izolované vstupy pracujú s 3,3 V a 5 V logikou. Pre zbernicu SPI je rýchlosť spracovania dát do 7 Mb/s, pre I2C do 1 Mb/s a UART do 195 kb/s. Maximálna podporovaná veľkosť Micro SD karty s FAT32 je 32 GB. Pre zobrazenie je použitý 3,2” farebný displej s dotykovou plochou. Dekodér je napájaný 12 V/1 A adaptérom.



Obrázok 2: Vývojový diagram softvéru pre Arduino

5. ZÁVER

Cieľom bolo vytvoriť univerzálne zariadenie, ktoré bude schopné spracovávať zachytené dáta zo sé- riových liniek. Doterajšia práca sa zaoberala hlavne návrhom hardvéru a z časti softvéru, kde bolo vytvorené jednoduché užívateľské rozhranie. Výhoda navrhnutého zariadenia vďaka implementova- ným modulom (Ethernet, SD karta, TFT a LCD displejom) je, že zariadenie možno použiť aj na iné účely.

LITERATÚRA

[1] GIERTL, J. Dekodér sériových sběrnic na platformě Arduino, Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2016. 38 s. Vedoucí semestrální práce Ing. OnĎrej Krajsa, Ph.D. [online]. [cit. 14. 3. 2016]. Dostupné z URL:<https://www.vutbr.cz/www_base/zav_prace_soubor.php?file_id=112213&zp_id=90425>.

[2] Keysight Technologies MSO6014A Mixed Signal Oscilloscope: [online]. [cit. 14. 3. 2016]. Dostupné z URL:<http://www.keysight.com/en/pd-723607-pn-MSO6014A/ mixed-signal-oscilloscope-100-mhz-4-scope-and-16-digital-channels? cc=CZ&lc=eng>.

BIOIMPEDANCE SIGNAL PROCESSING

Vladimíra Valíčková

Bachelor Degree Programme (3), FEEC BUT

E-mail: xvalic09@stud.feec.vutbr.cz

Supervised by: Ladislav Soukup

E-mail: xsouku08@stud.feec.vutbr.cz

Abstract: This study deals with detection of important points in derivate bioimpedance signal that may be further used in various medical applications, for example in calculation of the Pulse wave velocity. This work compares three detection methods. The best results were achieved by a detection with fixed interval. Algorithms that are dependent on constant bioimpedance and length of RR interval perform worse.

Keywords: bioimpedance, bioimpedance signal, time varying impedance, constant impedance

1. ÚVOD

Bioimpedance je snadno neinvazivně měřitelná veličina závislá na zastoupení a vlastnostech jednotlivých tkání v lidském organismu. Nejčastěji je využívána v přístrojích pro lékařskou diagnostiku, např. v kompozičních analyzátorech nebo impedanční tomografii. Cílem této práce je detekce maxim v druhé derivaci bioimpedančního signálu za účelem dalšího využití, např. v impedanční kardiografii, impedanční pletysmografii, nebo pro výpočet rychlosti šíření pulzní vlny.

2. BIIMPEDANČNÍ SIGNÁL

Bioimpedanční signál je posloupnost komplexních čísel, vyjadřující průběh impedance měřené části lidského těla v závislosti na čase. Rozeznáváme dvě složky a to statickou Z_0 a pulzativní $\Delta Z(t)$, které je možné oddělit digitální filtrací. Tyto složky popisuje tzv. paralelní válcový model bioimpedance, definovaný J. Nyboerem pro potřeby impedanční kardiografie [1].

Statická složka Z_0 je tvořena převážně impedancí kostí, svalů, tuků a reziduální (zbytkové) krve v cévách. V čase se velmi pomalu mění a to převážně díky vlivu dýchání, proto je reprezentována nízkými frekvencemi (0-0,6 Hz). Zatímco pulzativní složka $\Delta Z(t)$ vyjadřuje pulzaci krve a je reprezentována vysokými frekvencemi (0,6–15 Hz). Protože krev je jednou z nejvodivějších částí organismu, kdykoli nový ejekční objem dorazí do námi měřené oblasti, zvýší se v daném místě vodivost, to má za následek snížení hodnoty pulzativní bioimpedance [1].

3. AKVIZICE DAT

Měření bylo provedeno na vícekanálovém impedančním monitoru [2] (certifikace 12/2009, vynálezci Vondra a kol., Brno) aktuálně používaném pouze ve Fakultní nemocnici u svaté Anny v Brně pro diagnostiku stavu cévního systému. Přístroj využívá tří vysokofrekvenčních proudových zdrojů o frekvencích 49, 50 a 51 kHz pro snímání 18 kanálů bioimpedance. Každé měřené osobě byl také Peňázovou metodou kontinuálně měřen krevní tlak a 12 svodové EKG. Vzorkovací frekvence všech signálů byla 500 Hz [2].

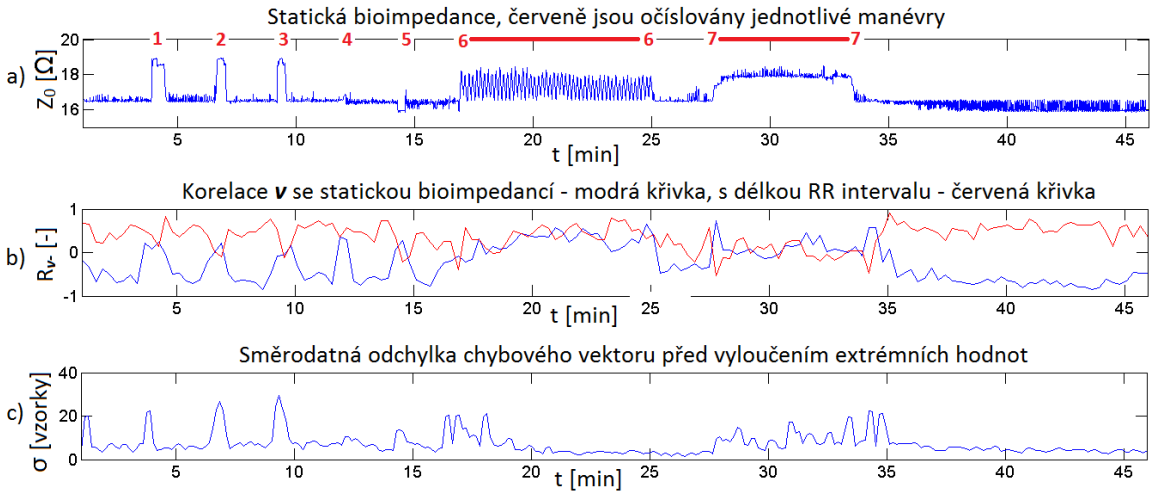
Subjekty při měření ležely na lůžku na zádech. Měřící protokol byl složen následovně: zadržení dechu, 2x Valsalvův manévr (+50 cmH_2O), 2x Müllerův manévr (+50 cmH_2O), řízené dýchání (f=0,1 Hz) a nakloněná rovina, kdy se lůžko naklonilo o cca 70°. Každý manévr byl proložen dvěma mi-

nutami klidu. Vlastní měření trvalo kolem 40 minut. Průběhy jednotlivých manévrů jsou v uvedeném pořadí označené na obr. 1a, kde je také možno vidět jejich vliv na statickou bioimpedanci [3].

4. PROGRAMOVÉ ZPRACOVÁNÍ

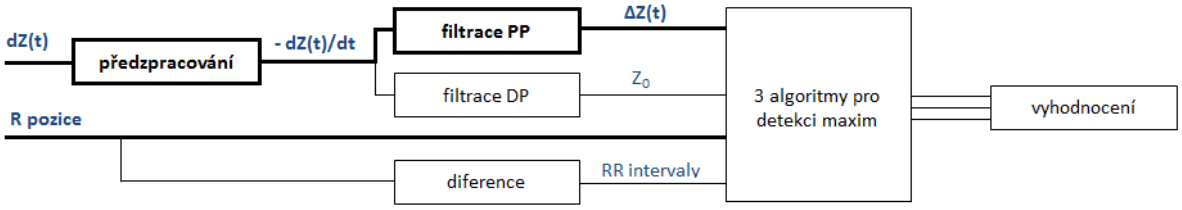
Program byl kompletně navržen v prostředí MATLAB a skládal se ze dvou částí.

První část hledala veličinu, která nejlépe koreluje s vektorem ***v*** [vzorky], tj. vektor vzdáleností poloh maxim bioimpedance od R pozic v EKG. Pozice maxim byly stanoveny metodou pevného okna a ručně zkontrolovány. Pozice R byly převzaty z ICRC (The International Clinical Research Center). Mezi hledané veličiny patřila délka RR intervalu, statická bioimpedance (střední hodnota z úseku Z_0 v intervalu mezi pozicemi R vln). Dále systolický, diastolický, střední a pulzní krevní tlak, získané z tlakové křivky. Nejlépe korelující veličinou byla Z_0, viz obr. 1b, proto byla tato veličina použita k nastavení třetího detekčního algoritmu.



Obrázek 1: Ukázka zpracování signálu

Pro hledání maxim v pulzativní bioimpedanci byly použity tři detekční algoritmy. Bioimpedanční signál byl předzpracován výběrem reálné složky, převrácením (vynásobení -1) a derivací druhého řádu. Následovala filtrace stabilní Butterworthovou IIR pásmovou propustí v oblasti 0,6 – 15 Hz pro získání pulzativní složky ΔZ(t). Samotná filtrace byla provedená funkcí filtfilt, která eliminuje fázové zkreslení. Tučně zobrazená část schématu byla stejná pro všechny tři detektory, viz obr. 2.



Obrázek 2: Blokové schéma detekční části algoritmu

1) Detektor závislý na pevném okně hledal maxima v pevném intervalu 38 až 88 vzorků za pozicí R vlny EKG.

2) Detektor závislý na délce RR intervalu hledal maxima v intervalu od 0,20 do 0,38 násobku RR intervalu EKG.

3) Ve třetím případě byla nejdříve stabilní Butterworthovou IIR dolní propustí v oblasti 0-0,6 Hz vyfiltrována statická složka bioimpedance, opět funkcí filtfilt. Následně byl stanoven koeficient tak, aby po jeho vynásobení se statickou bioimpedancí byly získána vzdálenosti pozic maxim

v pulzativní bioimpedanci od R vrcholu. Detekční okno bylo nastaveno jako interval ±60 vzorků od této pozice. Tento algoritmus byl vytvořen na základě výsledku korelací vektoru ***v*** a Z_0, viz obr. 1b.

5. VALIDACE DETEKTORŮ

Pro každého pacienta a každý detektor byl vytvořen chybový vektor ***c*** [vzorky], jako rozdíl délek RRintervalů EKG a intervalů pozic detekovaných maxim v ΔZ(t). U prvků tohoto vektoru je předpokládáno normální rozložení se střední hodnotou v nule. Nejprve byl spočítán počet prvků, průměr a směrodatná odchylka z ***c***, viz tab.1. V dalším kroku byly pomocí Grubbsova testu vyloučeny extrémní hodnoty a opět byl vypočítán průměr a směrodatná odchylka nového chybového vektoru, viz tab.1.

Tabulka 1: Průměrné hodnoty pro 10 subjektů

Typ detekčního okna	chybový vektor				chybový vektor po vyloučení extrémů		průměrná vzdálenost detekcí od R pozic [vzorky]	směrodatná odchylka vzdáleností detekcí od R pozic [vzorky]
	počet prvků [-]	Průměr [vzorky]	směrodatná odchylka [vzorky]	množství vyloučených hodnot[%]	průměr [vzorky]	směrodatná odchylka [vzorky]		
pevné	3006	-0,0024	4,9792	0,93	0,0052	4,1427	61,1247	6,3559
závislé na RR intervalu	3005	-0,0018	7,4591	1,45	0,0043	4,6443	61,1395	7,7827
závislé na Z_0	3005	-0,0015	8,7037	1,58	0,0006	4,3582	58,4080	7,7383

6. VYHODNOCENÍ

Pomocí výše uvedených detekčních algoritmů byla zpracována data od 10 osob (7 žen a 3 mužů s průměrným věkem 45 let a směrodatnou odchylkou 16,95 let) u kterých nebyly diagnostikovány závažné zdravotní problémy.

Teoreticky by detektor založený na pevném okně měl detekovat s nejnižší úspěšností, protože se nedokáže přizpůsobit dechovým manévrům, ani jednotlivým pacientům. Nejlepších výsledků by naopak měla dosahovat detekce závislá na Z_0 protože Z_0 částečně koreluje s ***v*** a protože statická bioimpedance je snímána z každého svodu, na rozdíl od EKG, které je snímáno vlastními elektrodami pouze v okolí srdce.

Avšak z tab. 1 vyplývá, že nejlepších výsledků dosahuje detekce pevným oknem. Průměrná hodnota vzdáleností pozic bioimpedance od R byla 61 vzorků, se směrodatnou odchylkou 6,36 vzorků. To znamená, že většina detekcí byla v prostřední části tohoto okna. Zbylé dva detekční postupy dosahují velmi podobných výsledků, přičemž algoritmus závislý na Z_0 je horší. Na obr. 1c můžeme vidět, že chybné detekce nejčastěji vznikaly právě v místech manévrů pacienta.

REFERENCE

[1] BAURA, Gail D. *System theory and practical applications of biomedical signals*. IEEE Press ;--Wiley-Interscience, 2002. ISBN 0471236535.

[2] US Patent Issued to INSTITUTE OF SCIENTIFIC INSTRUMENTS AS CR, V. V. I. on Oct. 27 for "Device for blood flow property measurement and method of its connection" (Czech Inventors). *US Fed News Service, Including US State News* [online]. Washington, D.C: HT Media Ltd, 2015 [cit. 2016-03-30].

[3] MATEJKOVA, Magdalena, Vlastimil VONDRA, Josef HALAMEK, Ladislav SOUKUP, Pavel JURAK, Filip PLESINGER a Ivo VISCOR. Measurement of pulse wave velocity during Valsalva and Mueller maneuvers by whole body impedance monitor. In: *Computing in Cardiology* [online]. IEEE Computer Society, 2014, **41**(January), s. 1117-1120 [cit. 2016-03-30]. ISSN 23258861.

CALCULATION OF THE DISPARITY MAP OF STEREO COLOR IMAGES

Barbora Kulikova

Bachelor Degree Programme (3), FEEC BUT

E-mail: xkulik03@stud.feec.vutbr.cz

Supervised by: Jan Klečka

E-mail: klecka@feec.vutbr.cz

Abstract: This article deals with the creation of the depth maps. The first chapter concerns with the physiology of the human space perception and the methods of displaying the 3D content which are the topics closely related to the depth map, its creation and practical usage. Subsequently, there is a chapter focused on the description of the used methods of the image processing. The fundamental chapter deals with the methods of computing the depth maps and used principles. In the practical part of the project, an experiment was made with a local method of the creation of the depth map.

Keywords: disparity map, stereoscopy, CSAD, SSIM, CGRAD

1. ÚVOD

Článek se zabývá vytvořením aplikace pro výpočet disparitní mapy. Aplikace je vytvářena v prostředí MATLAB. Disparita je rozdíl horizontálních obrazových souřadnic daného objektu mezi levým a pravým obrazem téže scény. Pomocí vnitřních parametrů kamery lze z disparity dopočítat hloubku bodu scény, tedy vzdálenost bodu od snímací kamery a jeho umístění v trojrozměrném prostoru reálného světa. Stereo vidění, a tedy i výpočet disparitní mapy, je aktuální problém. Využití nachází zejména v 3D TV, ale i tvorbě modelů objektů, které se následně dají využít v mnoha průmyslových odvětvích, jako jsou stavebnictví či strojírenství.

Vstupem pro zpracování jsou dva, případně více rektifikovaných snímků, tzv. stereo snímků, což jsou snímky téže scény pořízené dvěma fotoaparáty lišícími se polohou v horizontální ose. Druhou možností je pořízení snímků fotoaparáty v libovolné pozici, ty však následně musí být rektifikovány. Výstupem je disparitní mapa, což je snímek o stejné velikosti jako jednotlivé vstupní snímky a reprezentuje vzdálenost korespondenčních bodů v rektifikovaných souřadnicích od kamery. Vytvořená aplikace implementuje zástupce obou skupin metod pro zpracování stereo snímků a tvorbu disparitních map – lokální a globální. Základem obou metod je hledání korespondujících bodů ve dvou snímcích. Tedy bodů, které reprezentují tentýž bod scény.

2. POPIS VYTVOŘENÉHO PROGRAMU

Vytvořená aplikace umožňuje výpočet disparitní mapy pomocí tří různých přístupů

- lokální metoda založená na podobnostních metrikách,
- semiglobální metoda scanline,
- metoda založená na segmentaci a hledání výrazných bodů obrazu.

V rámci lokální metody je zjištěna cena disparit pro jednotlivé pixely obrazu. Z toho důvodu je lokální metoda základním stavebním kamenem zisku disparitní mapy. Ceny disparit mohou být následně použity v navazující semiglobální metodě scanline a v budoucnu také v globální metodě využívající algoritmus belief propagation. Pro ukázku jsou na Obrázku 1a zobrazeny získané hloubkové mapy. Pro intuitivní práci s aplikací bylo vytvořeno uživatelské prostředí aplikace (Obrázek

1b). V následujících kapitolách budou z důvodu rozsahu článku stručně popsány dvě ze tří implementovaných metod.

2.1. LOKÁLNÍ METODA

Implementovaná metoda používá základní algoritmus pro výpočet husté hloubkové mapy hledáním korespondencí bod po bodu. Algoritmus používá jako základ zjednodušenou implementací metody publikované v článku [1]. Algoritmus může být popsán vývojovým diagramem na obrázku 2a. Základními krokem je hledání správné hodnoty disparity pomocí míry podobnosti a agregace této míry podobnosti v různě velkém okně. V prvním kroku jsou definovány parametry algoritmu P_{weigh}, win_{size} a d_{max} definující maximální velikost disparity. V dalších dvou krocích je opakovaně dvakrát volána funkce *SlideImages*, která je jádrem algoritmu, v rámci něhož se provádí evaluace míry podobnosti i agregace podobnosti v určitém okolí definovaném parametrem win_{size}. Jako míra podobnosti jsou v aplikaci použity: rozdíl absolutních diferencí a podobnost dvou obrazů (Structural Similarity Index). Funkce je volána dvakrát, protože jednou jsou vyhledávány bodové korespondence ve snímku i_2 pro pixelu ve snímku i_1 a podruhé je tomu naopak. V důsledku jsou tedy získány dvě disparitní mapy d_1 a d_2. V předposledním kroku postupu je volána funkce *WinnerTakeAll*. Úloha této funkce je přiřadit každému pixelu tu nevhodnější hloubku, tedy buď hodnotu mapy d_1 nebo d_2. Rozhodnutí je provedeno na základě matic *mindiff1* a *mindiff2*, které obsahují cenu dané disparity v jedné a v druhé disparitní mapě. Jako finální hodnota disparity pro daný pixel je zvolena hodnota z té disparitní mapy, ve které je cena disparity pro daný pixel menší. V posledním kroku je provedena snaha eliminovat alespoň nějaké chyby spočívající v nespojitosti hloubkové mapy. Tohoto cíle je dosaženo pomocí průměrování hodnot v okolí daného bodu.



Obrázek 1: a) Výstupní hloubkové mapy b) uživatelské prostředí aplikace.

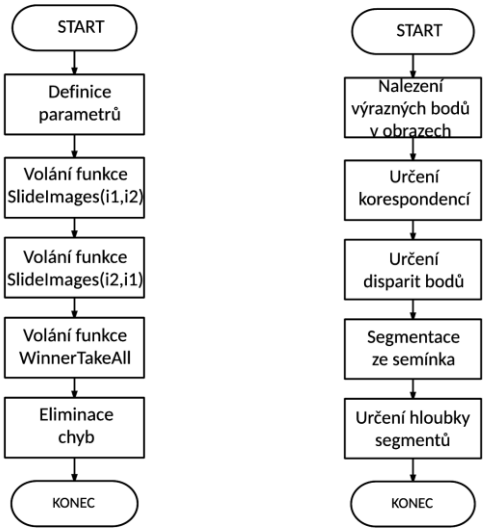
2.2. VYUŽITÍ SEGMENTACE A VÝRAZNÝCH BODŮ OBRAZU

Velké množství metod vychází z hledání bodových korespondencí pro každý pixel obrazu bod po bodu. Nalezení korespondujícího bodu pro pixel, jenž se neodlišuje od okolních pixelů, může být velmi obtížné a v důsledku nespolehlivé. Odlišným přístupem může být hledání bodových korespondencí pouze pro výrazné body obrazu. Metody pracující na tomto principu se v anglické literatuře nazývají „feature based"[2]. Výrazné body obrazu jsou takové body, které se výrazně liší od svého okolí a jsou tak dobře identifikovatelné [2]. Výrazný obrazový bod musí být nalezitelný opakovaně a spolehlivě. Z toho vyplývá, že daný bod můžeme lehce najít v obou stereo obrazech, a tím nalézt korespondující body a určit disparitu daného bodu. K hledání výrazných bodů obrazu (scény) jsou používány tzv. detektory a deskriptory. Postup lze rozdělit do následujících kroků

- Nalezení sad výrazných bodů v obou obrazech
- Výpočet deskriptoru výrazných bodů
- Porovnání deskriptorů jednotlivých bodů- určení korespondencí

Mezi nejznámější a nejčastěji používané detektory patří SIFT a SURF. Algoritmy by měly být odolné vůči natočení scény, změně měřítka, změně jasové hodnoty.

Takto nalezené korespondence však nebudou souvisle pokrývat celý obraz (scénu) a nevznikne souvislá disparitní mapa. Z toho důvodu musí být provedeny další kroky pro zisk souvislé hloubkové mapy. Logicky navazujícím krokem je segmentace scény a následné využití segmentovaného obrazu takovým způsobem, že disparita konkrétního bodu scény (výrazného bodu) je přiřazena celému segmentu, jehož je bod součástí. Získané disparity výrazných bodů lze popřípadě využít k zpřesnění nebo doplnění hloubkové mapy získané s využitím hledání bod po bodu [1]. Pro určení hloubky segmentu je možné použít různé strategie, které zde nejsou z důvodu rozsahu článku popsány. Celý algoritmus může být popsán vývojovým diagramem na obrázku 2b.



Obrázek 2: Vývojový diagram a) lokální metody b) metody využívající segmentaci.

3. ZÁVĚR

V rámci práce na bakalářské práci je vytvářena v prostředí MATLAB aplikace pro výpočet disparitní mapy. Aplikace umožnuje při vyplnění údajů o ohniskové vzdálenosti fotoaparátů a vzdálenosti fotoaparátů přepočítat disparitní mapy na hloubkové a případně vytvořit jednoduchý prostorový model scény pomoci stereogrammetrie. Aplikace v současné době nabízí výpočet hloubkové mapy třemi různými metodami a plánuji přidání globální metody. Při použití každé z metod má uživatel možnost nastavit některé z jejich parametrů a tím ovlivnit výsledek. Následně je možné porovnat získané disparitní mapy s pravdivou disparitní mapou, pokud je k dispozici. Vzhledem k těmto skutečnostem může být aplikace použita pro testování vlivu daných parametrů na kvalitu hloubkové mapy.

REFERENCE

[1] KLAUS, A., M. SORMANN a K. KARNER. Segment-Based Stereo Matching Using Belief Propagation and a Self-Adapting Dissimilarity Measure. *18th International Conference on Pattern Recognition (ICPR'06)* [online]. IEEE, 2006, , 15-18 [cit. 2. prosince 2015]. DOI: 10.1109/ICPR.2006.1033. ISBN 0-7695-2521-0.

[2] JIANG G., Yi J. "Comparison and Study of Classic Feature Point Detection Algorithm," in *Computer Science & Service System (CSSS), 2012 International Conference* on , vol., no., pp.2307-2309, 11-13 Aug. 2012 doi: 10.1109/CSSS.2012.572

MEDICAL IMAGE SEGMENTATION BASED ON GRAPH CUT WITH SHAPE PRIOR

Dominika Kozlová

Bachelor Degree Programme (3), FEEC BUT

E-mail: xkozlo03@stud.feec.vutbr.cz

Supervised by: Roman Jakubíček

E-mail: xjakub08@stud.feec.vutbr.cz

Abstract: This paper deals with a graph-based image segmentation and its improvement by using the information about the shape of the object for creating specific graph architecture (template). Improved method allows the cut to prefer more complicated structures, especially when the image contains a lot of noise and the object is hardly indistinguishable from the background. Algorithm was tested on simulated data and real CT and MRI images of vertebra and brain in 2D. Method was also extended to 3D further purposes.

Keywords: Medical image segmentation, Graph theory, Graph cut, Template

1. ÚVOD

U medicínských obrazů získaných pomocí zobrazovacích systémů, jako je počítačová tomografie či magnetická rezonance, přispívá segmentace dat ke stanovení diagnózy. Využívá se zejména pro rozlišení různých typů tkání, 3D modelování a vizualizaci struktur, jako jsou orgány, kosti, cévy i různé patologie.

Důležitou roli hrají postupy vycházející z teorie grafů, kdy je obraz převeden na jeho grafovou reprezentaci. Algoritmus klasické grafové metody spojuje uzly grafu v obrazu rovnoměrně a ekvidistantně. Tento postup ale nedovoluje preferovat určitou strukturu, což je potřebné především v případech, kdy jsou části objektu hůře rozeznatelné od pozadí. Tento problém je řešen pomocí zadaného tvaru objektu při tvorbě grafu. Pak mohou být vzorkované uzly grafu dle vytvořené šablony v obrazu rozčleněny nerovnoměrně a neekvidistantně pro preferování dané struktury.

2. REALIZACE SEGMENTAČNÍHO ALGORITMU

2.1. GRAFOVÁ REPREZENTACE OBRAZU

Hlavním bodem je utvoření váhovaného grafu $G = (V,E)$ s množinou uzlů V a množinou hran E. Uzly odpovídají pixelům obrazu a hrany spojují uzly dle zvoleného systému. Každá hrana grafu má svou váhu, která může představovat určitou míru preference, zda odpovídající pixely patří do objektu zájmu, nebo rozdíl jasů uzlů, které spojuje. V grafu je každá hrana orientovaná, tedy obsahuje stanovené pořadí uzlů. Graf je navíc směrovaný, tudíž váhování v jednom směru hrany nemusí odpovídat váhování ve směru opačném.

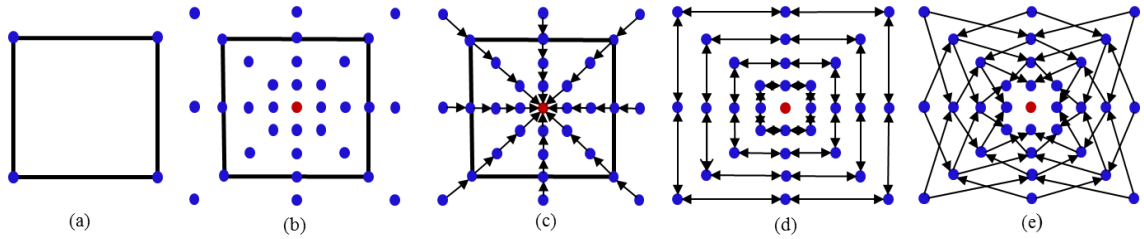
V konstruovaném grafu jsou uzly spojeny tzv. n-hranami dle systému 4-okolí ve 2D. Do grafu jsou přidány dva imaginární terminální vrcholy, zdroj s a stok t, které jsou spojeny tzv. t-hranami se všemi pixely obrazu. Výpočty vah n-hran i t-hran vychází z minimalizace segmentační energie.

Získané hodnoty t-hran a n-hran následně zpracovává „min-cut/max-flow“ algoritmus Y. Boykova a V. Kolmogorova [1], který vyhledá maximální tok sítě a zjistí minimální s-t řez grafu, tedy výsledek segmentace.

2.2. TVORBA ŠABLONY

V prvním kroku je nastaven orientovaný graf šablony z počátečního bodu definovaného uživatelem, umístěného zhruba uprostřed segmentovaného objektu. Pro tvorbu šablony je dále nutné znát souřadnice jejích krajních bodů po směru hodinových ručiček, které musí vhodně popisovat daný objekt. Zadané souřadnice jsou následně spojeny do uzavřeného obrysu, jak lze vidět na obrázku 1(a) pro čtvercovou šablonu. Poté jsou z počátečního bodu vyslány radiálně paprsky přes vytvořený obrys, podél kterých jsou vzorkovány uzly grafu, jak je znázorněno na obrázku 1(b). Všechny paprsky mají stejný počet uzlů, mezi kterými mohou být různé vzdálenosti.

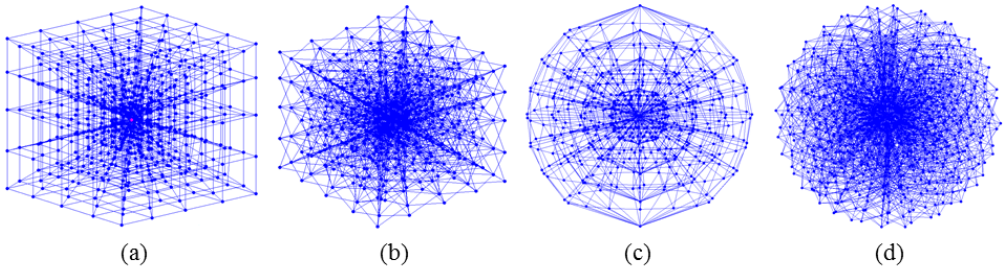
Rozmístěné uzly v grafu spojuje odpovídající množina hran. N-hrany spojují jen uzly šablony a t-hrany propojují uzly se zdrojem a stokem pro výpočet s-t řezu. Existují dva typy váhovaných n-hran, z-křivky a r-křivky. Z-křivky, znázorněny na obrázku 1 (c), jsou mezi uzly daného paprsku. R-křivky spojují uzly mezi sousedními paprsky, omezují množinu možných segmentací a určují variabilitu výsledného obrysu objektu pomocí parametru Δ_r. R-křivky šablony s parametrem $\Delta_r = 0$ a $\Delta_r = 1$ jsou zobrazeny na obrázku 1 (d) a (e). S rostoucí hodnotou Δ_r je možná větší variabilita tvaru objektu vzhledem k tvaru předdefinované šablony [2].



Obrázek 1: Konstrukce grafu čtvercové šablony. (a) Spojené rohové body šablony. (b) Uzly šablony. (c) Z-křivky. (d) R-křivky, parametr $\Delta_r = 0$. (e) R-křivky, parametr $\Delta_r = 1$.

2.3. REALIZACE VE 3D

Grafová metoda byla rozšířena do třetí dimenze pro možnost segmentace objemových dat. Algoritmus vytvoří směrovaný graf, jehož uzly odpovídají obrazovým voxelům. Množina váhovaných n-hran je utvořena dle systému 6-okolí, každý uzel je tedy spojen navíc se 2 uzly, které leží v jiných řezech 3D obrazu. Tvorba šablony je podobná, ale parně složitější oproti 2D provedení ve spojování uzlů grafu a vzorkování stran šablon. Zkonstruované šablony krychle a koule lze vidět na obrázku 2. Algoritmus je ve 3D provedení též více náročný z hlediska zpracování velkého objemu dat.



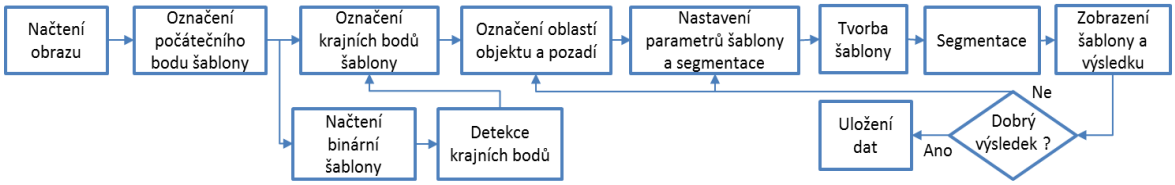
Obrázek 2: Zkonstruované 3D šablony. Krychle (a) $\Delta_r = 0$. (b) $\Delta_r = 1$. Koule (c) $\Delta_r = 0$. (d) $\Delta_r = 1$.

3. GRAFICKÉ UŽIVATELSKÉ ROZHRANÍ PROGRAMU

Neboť metoda není zcela automatická, bylo pro snadnou interakci uživatele s programem vytvořeno grafické uživatelské prostředí. Schéma programu je na obrázku 3. Nejprve je načten výchozí obraz pro segmentaci. Poté je zadán počáteční bod a následně krajní body pro tvorbu šablony, které je možné vyznačit ručně v načteném obrazu, nebo lze také použít binární vzor pro šablonu, ze kterého jsou krajní body získány detekcí hran nebo rohů a posunuty podle zadaného počátečního bodu

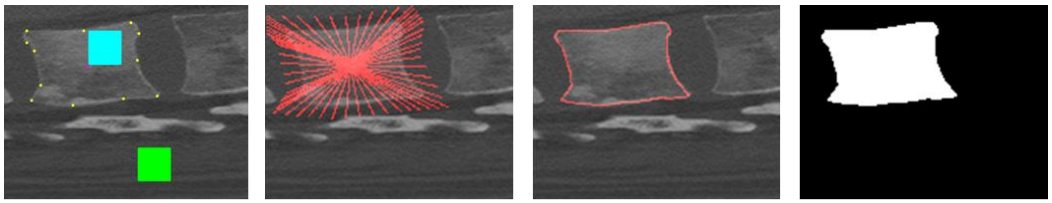
v obrazu. U šablony se dále definuje její přesah přes zadané krajní body, krok mezi jejími paprsky a body a parametr Δ_r, popsaný výše. Dále je třeba vyznačit jednu či více oblastí objektu a pozadí v obrazu. Posledním zadávaným parametrem je koeficient λ, který souvisí se segmentační energií. V případě, že je zvolena hodnota vyšší, segmentace je provedena více dle intenzity pixelů než jejich vzájemné souvislosti a naopak.

Po segmentaci jsou zobrazeny body šablony a výsledky segmentace. Při neuspokojivém výsledku lze přidat další oblasti objektu či pozadí, měnit parametry šablony i segmentace a proces opakovat.



Obrázek 3: Schéma realizovaného programu pro segmentaci.

Metoda byla testována na simulovaných datech i reálných CT a MRI snímcích obratlů v sagitálním pohledu (viz obrázek 4) a snímcích mozku s různými abnormalitami.



(a) (b) (c) (d)

Obrázek 4: Segmentace obratle. (a) Značení. (b) Body šablony. (c) Výsledný obrys. (d) Výsledek segmentace.

4. ZÁVĚR

Segmentační algoritmus byl realizován ve 2D s grafickým uživatelským prostředím a ve 3D provedení v MATLABu. Porovnány byly výsledky segmentace grafovou metodou bez a s vytvořenou šablonou na obrazech s různou úrovní Gaussovského šumu. Ze získaných výsledků lze usuzovat vyšší efektivitu segmentace při využití definované šablony, zejména při vyšší úrovni šumu v obrazu, kdy je objekt špatně rozlišitelný od pozadí a také při složitějším tvaru segmentovaného objektu.

Pro snadnou interakci uživatele s programem bylo vytvořeno grafické uživatelské rozhraní. I když metoda není zcela automatická a uživatel musí zadat různá značení a parametry, pomocí postupných úprav může dospět k výsledku požadované úrovně. Program lze tedy s výhodou využít i pro segmentace objektů obtížně rozlišitelných od pozadí. Na obrázku 4 lze vidět příklad úspěšné detekce obratle na CT snímku. Zdárně byly segmentovány i různé abnormality na CT a MRI snímcích mozku.

REFERENCE

[1] BOYKOV, Y. a V. KOLMOGOROV. An experimental comparison of min-cut/max- flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. **26**: 1124-1137. DOI: 10.1109/TPAMI.2004.60. ISSN 0162-8828.

[2] EGGER, Jan, Tina KAPUR, Thomas DUKATZ, Malgorzata KOLODZIEJ, Dženan ZUKIĆ, Bernd FREISLEBEN, Christopher NIMSKY a Yong FAN. Square-Cut: A Segmentation Algorithm on the Basis of a Rectangle Shape. *PLoS ONE* [online]. 2012-2-21, **7**(2): e31064- [cit. 2015-10-07]. DOI: 10.1371/journal.pone.0031064. ISSN 1932-6203.

PHYSIOCRATE: A SIGNALPLANT TOOLBOX FOR RESPIRATORY, BLOOD PRESSURE AND EMG SIGNAL ANALYSIS

Petr Nejedly, Jan Virgala

Bachelor Degree Program (2, 3), FEEC BUT

E-mail: xnejed07@stud.feec.vutbr.cz, xvirga00@stud.feec.vutbr.cz

Supervised by: Filip Plesinger

Institute of Scientific Instruments of the Czech Academy of Sciences, Czech Republic

E-mail: fplesinger@isibrno.cz

Abstract: The paper presented here describes the SignalPlant library for analysis of physiological signals. This library contains plugins for analysis of continuous blood pressure (BP), respiratory signals (RESP) and electromyographic signals (EMG). Its principal advantages are real-time previews of results, ease of use and a multi-thread approach. The library is available free and open- source under an MIT license as part of the SignalPlant project at https://signalplant.codeplex.com.

Keywords: SignalPlant, plugin, signal analysis, blood pressure, respiration, EMG

1. INTRODUCTION

SignalPlant [1] is a free software platform for signal analysis. Although SignalPlant is a powerful tool for signal inspection, processing and analysis, it lacks specific plugins for analysis of blood pressure (BP), respiration (RESP) and electromyographic signals (EMG). For this reason, we have developed three plugins with a user-friendly and interactive graphic user interface (GUI) allowing real-time results previewing. This is a major advantage over existing solutions such as RASCHlab [2] in which iterative parameter setting may be a time-consuming process. Here, we present the library with three new plugins extending SignalPlant to include new functionality.

2. RESPIRATION SIGNAL ANALYSIS PLUGIN

Respiratory signals may be obtained from several measuring devices such as inductance plethysmography or nasal thermistor or by deriving a respirational wave from an accelerometer placed on the chest. The respiration plugin presented here (Fig. 1) allows processing of all the mentioned signal types. It evaluates the following parameters from a RESP signal – amplitude, respiratory frequency, ventilation, and expiration and inspiration duration. It can be used in two modes. The first mode retrieves results from the whole signal, excluding artificial areas. The second mode allows analysis of selected parts of a signal, using labels as starting points and a fixed window duration (defined by the user in Fig. 1-B). RESP signal maxima and minima are localized automatically by an algorithm that finds local extremes in a floating window. This process is split into multiple threads for improved time efficiency. Signal distortion may be eliminated with a low- pass finite-impulse-response filter (FIR, Fig. 1-B). Moreover, it is also possible to manually change the positions of respiration maxima and minima in SignalPlant if the user is not satisfied with their automatic detection. The user can inspect the analyzed signal, the list of individual inspirations and expirations and the overall statistical results (switchable in Fig. 1-C) in real time as the plugin immediately recomputes the results whenever the user changes any setting. This is a major improvement over existing solutions such as [2] or [3].

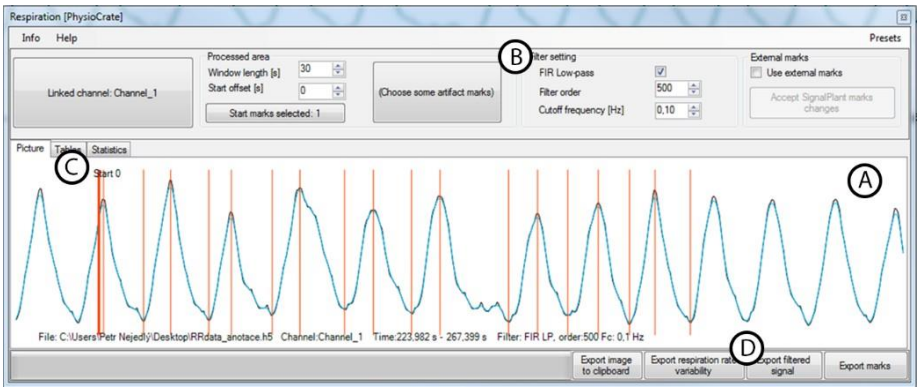


Figure 1: Respiration analysis plugin window showing preview of detected inhalations and exhalations from a respiratory record in a 30-second window started by the Start-0 mark. A – review area. The original signal is the black line, the filtered signal (FIR low-pass at 0.1 Hz, 500th order) is the blue line; minima and maxima are represented as orange vertical lines. B – parameter settings panel. C – page switch. D – export buttons.

3. CONTINUOUS BLOOD PRESSURE ANALYSIS PLUGIN

This plugin is designed for the analysis of signals obtained from invasive and non-invasive blood pressure measurements. The input signal is filtered with a filter (FIR low-pass at 10 Hz, 50th order), after which local extremes are found. These extremes must satisfy the condition that they are greater than the 90th percentile or lower than the 10th percentile in comparison with surrounding values to be accepted as systoles and diastoles. This plugin may produce heart-rate, systolic and diastolic curves as well as systolic/diastolic marks for each beat. The user interface is extremely similar to the GUI in the RESP analysis plugin.

4. ELECTROMYOGRAM ANALYSIS PLUGIN

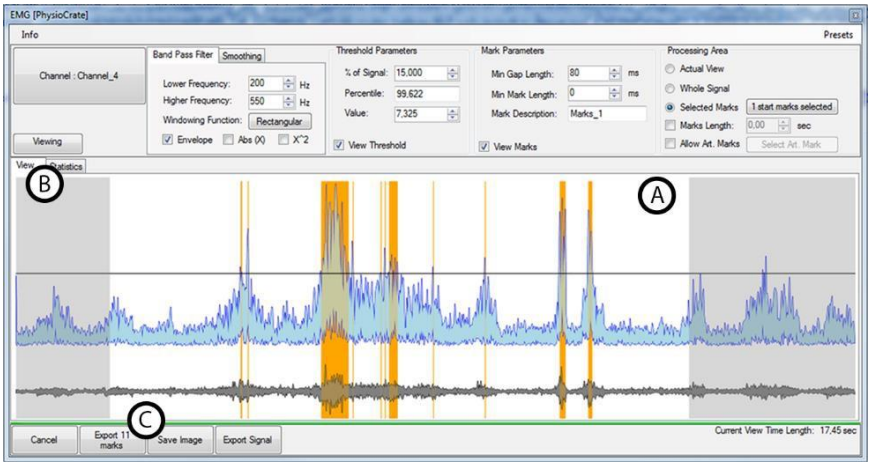


Figure 2: EMG analysis plugin window showing a preview of an EMG record of the m. corrugator supercilii activity. A – preview area. Detected areas with higher activity in the specified frequency range are marked by the orange color. The original signal (dark gray) is at the bottom. The blue signal is the resultant envelope of EMG analysis. B – parameter settings panel. C – export buttons.

The presented plugin is designed for the detection of areas in which an amplitude envelope in a given frequency range exceeds the selected threshold, as is the case in the detection of muscle activity in EMG records. The measured signal is displayed in a plugin window (Fig. 2) and the user defines the properties (Fig. 2-B) of the pre-processing chain. The signal is transformed into amplitude envelopes

using the Hilbert transform (selectable frequency range, windowing function and squaring). Consequently, it is possible to apply smoothing (selectable window size and type). These processes are split into multiple threads to speed up the computational process. The threshold value can be set in three modes. Firstly, directly as a value, secondly, as a percentage of the signal maximum, and thirdly as a user-defined percentile of the signal. Areas with an amplitude higher than the threshold value are highlighted (Fig. 2-A – orange areas) so as to be easily identified by the user. To ease the plugin manipulation, the user is given an opportunity to make his own preset for most used controls.

5. RESULTS AND CONCLUSION

Systole detection from Blood pressure plugin was validated using 18 records from MIT-BIH Polysomnographic Database [4]. We compared annotated position of R wave with detected starts of systoles. Average sensitivity was 99.33% and positive predictive value was 99.39%. Respiration plugins was validated using 42 records from CapnoBase[5][6] database. Average sensitivity was

93.27% and positive predictive value 92.80%.

We have developed interactive plugins that are going to be used in neuropsychological research for the analysis of respiration, continuous blood pressure and EMG signals. Their principal advantages over existing solutions are real-time interaction with the user, multi-thread approach and 64-bit architecture. Furthermore, presented plugins allows user to save analysis parameters that can be re-used later. Plugins allow working with file formats supported by SignalPlant (as EDF, BIN, SCOPEWIN, MAT *etc*.) and they are accompanied with help file in PDF. Plugins are available free under an MIT license in the form of source code as well as compiled DLL and can be downloaded from https://signalplant.codeplex.com/ as part of SignalPlant software.

ACKNOWLEDGEMENT

Development was partially supported by the project GACR P103/11/0933.

REFERENCES

[1] PLESINGER, Filip, Juraj JURCO, Josef HALAMEK a Pavel JURAK. *SignalPlant*. Brno, Czech Republic: Institute of Scientific Instruments of CAS. 2015

[2] SCHNEIDER, R, A BAUER, P BARTHEL a G SCHMIDT. libRASCH - a Programming Framework for Transparent Access to Physiological Signals. In: *Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE* [online]. 2004, s. 3254–3257. Dostupné z: doi:10.1109/IEMBS.2004.1403915

[3] *LabChart*. Dunedin,New Zealand: ADInstruments. 2013

[4] ICHIMARU, Y. a G. B. MOODY. Development of the polysomnographic database on CD-ROM. *Psychiatry and Clinical Neurosciences* [online]. 1999, roč. 53, č. 2, s. 175–177. ISSN 13231316. Dostupné z: doi:10.1046/j.1440-1819.1999.00527.x

[5] KARLEN, W, M TURNER, E COOKE, G DUMONT a J M ANSERMINO. CapnoBase: Signal database and tools to collect, share and annotate respiratory signals. In: *Annual Meeting of the Society for Technology in Anesthesia (STA)*. 2010.

[6] KARLEN, W, S RAMAN, J M ANSERMINO a G A DUMONT. Multiparameter respiratory rate estimation from the photoplethysmogram. *IEEE transactions on bio-medical engineering*. 2013, roč. 60, č. 7, s. 1946–1953.

Bakalářské projekty

Automatizace a silnoproudá elektrotechnika

BELT FILLING LINE

František Majchrák

Bachelor Degree Programme (2), FEEC BUT

E-mail: xmajch01@stud.feec.vutbr.cz

Andrej Šimurda

Supervised by: Radek Štohl

E-mail: stohl@feec.vutbr.cz

Abstract: The goal was to create a mechatronic product, including sensory part, drives and control. The product is used to fill chosen drink flavors into the cup and transport it along the filling line. Management is realized through the Siemens PLC S-7 200.

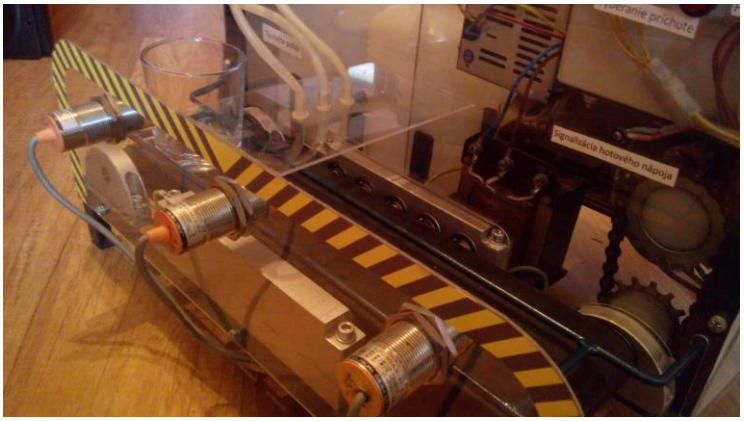
Keywords: conveyor, PLC, sensor, mixing drink

1. ÚVOD

Nápojový automat ako školský model slúži pre plnenie nápojov do pohárov podľa príchute, ktorú si zvolíme. Pohyb pohára je zabezpečený pomocou pásového dopravníku. Kvapalina je dopravovaná z nádob do pohára zubovými čerpadlami. Celý model je riadený programovateľným logickým automatom Siemens S7-200.

2. DOPRAVA POHÁRA

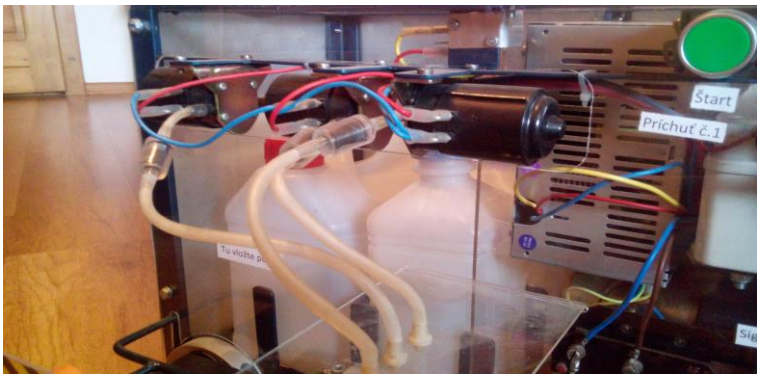
Pre dopravu pohára sme si zvolili dopravníkový pás. Pohár sa dopravuje medzi troma bodmi. Prvý bod je vloženie pohára. Pás následne dopraví pohár pod plniace dýzy. Po dokončení plnenia je pohár dopravený do miesta odkiaľ je možné pohár vziať. Jednotlivé miesta, na ktorých pohár zastavuje, sú snímané kapacitnými snímačmi.



Obrázok 1: Doprava pohára po dopravníkovom páse

3. DOPRAVA KVAPALINY DO POHÁRA

Kvapalina je dopravovaná do pohára pomocou zubových čerpadiel. Keďže sme nemali snímače prietoku, experimentálne sme určili dobu plnenia. Aby sa plnenie spresnilo na hadičky vedúce kvapalinu od čerpadiel k plniacim dýzam sme umiestnili spätné ventily, čím sme zaručili, že sa kvapalina nevráti do nádoby ale zostane v hadičkách. Tým sa plnenie samozrejme aj urýchli.



Obrázok 2: Plniaci mechanizmus

4. OVLÁDANIE MODELU

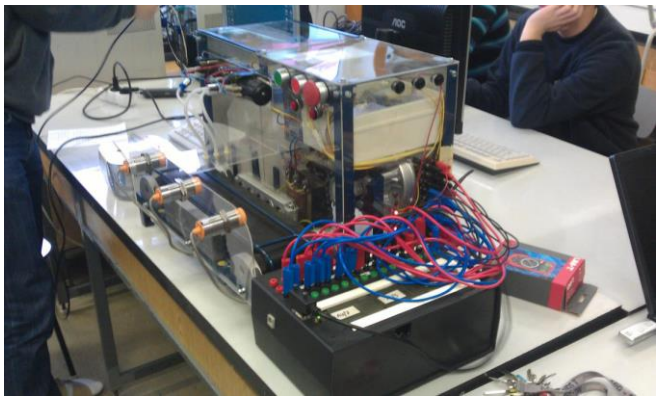
Na modeli je umiestnených osem tlačidiel. Sú to tlačidlá „štart“, „stop“, „centrál stop“, „voľba príchute č.1“, „voľba príchute č.2“, „spúšťanie čerpadla č.1“, „spúšťanie čerpadla č.2“ a „spúšťanie čerpadla č.3“. Tlačidlo „centrál stop“ slúži na úplné odpojenie modelu od elektrickej energie. Po stačení tlačidla pre voľbu príchute č.1 alebo č.2 môžeme stlačiť tlačidlo „štart“ čím odštartujeme celý proces plnenia. Tlačidlom „stop“ kedykoľvek zastavíme chod programu a stlačením tlačidla „štart“ ho obnovíme. Tlačidlá pre spúšťanie čerpadiel sú na modeli hlavne preto aby sa pri zapínaní modelu dala do hadičiek naťahať kvapalina.



Obrázok 3: Ovládacie prvky

5. RIADIACI SYSTÉM

Pre riadenie modelu sme si zvolili programovateľný logický automat od firmy Siemens S-7 200. Model je však navrhnutý tak, aby bolo možné použiť akýkoľvek riadiaci systém a to vďaka tomu, že sú všetky vstupy aj výstupy potrebné pre riadenie vyvedené na bočnej stene modelu. Pre riadenie je potrebné aby mal riadiaci systém k dispozícii 14 vstupov a 6 výstupov.



Obrázok 4: Pripojenie riadiaceho systému k modelu

6. SENZORY A INICIALIZÁCIA

Okrem spomenutých kapacitných snímačov sa v modeli nachádzajú ešte snímače, ktoré oznamujú nízku hladinu v nádobách a otvorené dvierka, ktorými sa vkladajú nádoby do modelu. Nízka hladina v nádobách je snímaná mechanickými spínačmi, ktoré sa pri malej hmotnosti kvapaliny v jednotlivých nádobách rozopnú. V nádobách však ostane dostatok kvapaliny pre dokončenie chodu programu. Pootvorenie dvierok je takisto snímané mechanickým spínačom, ak sa zopne, preruší sa proces plnenia. Po zatvorení dvierok alebo odstránení problému a stlačení tlačidla „štart“ bude proces plnenia pokračovať. Na prednej stene modelu sa nachádza led dióda, ktorá oznamuje správne prebehnutie programu a možnosť vyzdvihnúť si pohár s nápojom. Na bočnej stene sa zasa nachádzajú led diódy, ktoré oznamujú čo spôsobilo prerušenie chodu programu.



Obrázok 5: Pripojenie riadiaceho systému a inicializácia

7. ZÁVER

Pri vytváraní modelu sme sa snažili riešiť problémy spojené s plnením kvapalín do nádob. V práci by som chcel pokračovať aj naďalej a model by som rozšíril napríklad o inicializáciu a ovládanie pomocou dotykového panelu a prípadne o ďalšie snímače, ktoré by spresnili dávkovanie kvapalín a umožňovali by plnenie kvapalín do nádob rôznych tvarov. Takisto by som chcel riadiť model novším riadiacim systémom.

REFERENCIE

[1] Siemens: Systémový manuál - Programovateľný automat S7-200, 06/2004, A5E00307987-01

[2] Bukovinský, I.: Automatizačná technika: Úvod do činnosti a programovania logického automatu SIEMENS S7-200, Kysucké Nové Mesto, Spojená škola, 2011

[3] Vávra, P. a kolektív: Strojnícke tabuľky pre SPŠ Strojnícke, Praha, 1983

ELECTRONIC PARTS OF SMALL PHOTOVOLTAIC POWER PLANT

Dušan Benda

Bachelor Degree Programme (3), FEEC BUT

E-mail: xbenda10@stud.feec.vutbr.cz

Supervised by: Zdeněk Bradáč

E-mail: bradac@feec.vutbr.cz

Abstract: The aim of this project is to propose a comprehensive system for the management, control and monitoring of individual components of photovoltaic power plants. This article describes the various methods of approaches to solving each described parts including an outline of the issue. The output of work are functional and tested equipment.

Keywords: switching power supply, photovoltaic power plants, power management, LT8705, LT8490, LTC2946, high efficiency, off-grid

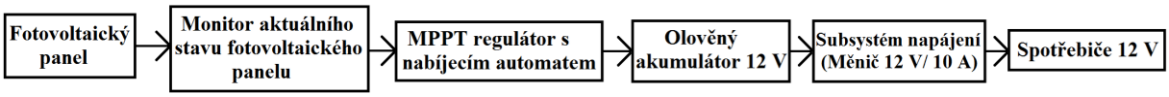
1. ÚVOD

Fotovoltaické elektrárny se staly nedílnou součástí našeho každodenního života. Za poslední desetiletí zažily boom převážně elektrárny velkých výkonů, typicky v řádech stovek kilowattů až jednotek megawattů špičkového výkonu. Budoucnost však bude patřit zřejmě malým ostrovním elektrárnám, řádově do 10 kW špičkového výkonu, takzvaným off-grid systémům, které nedistribuují vyrobenou energii do sítě, ale veškerá energie je spotřebována přímo lokálním výrobcem.

Tento projekt se zabývá návrhem a realizací elektronických obvodů určených pro malé fotovoltaické systémy s ohledem na zařazení jednotlivých podsystémů (nabíjecí automaty, napěťové regulátory, monitorovací systémy) do již vytvořeného systému řízení inteligentního domu.

2. KONCEPCE OSTROVNÍ FOTOVOLTAICKÉ ELEKTRÁRNY

Na následujícím blokovém schématu (Obrázek 2.1) jsou uvedeny základní elektronické subsystémy malé fotovoltaické elektrárny, které jsou popsány v jednotlivých kapitolách tohoto článku.



Obrázek 2.1: Blokové schéma zapojení elektronických součástí FV elektrárny.

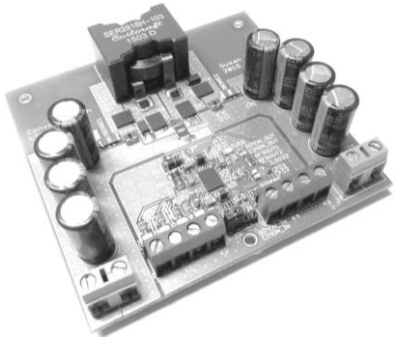
3. SUBSYSTÉM NAPÁJENÍ SPOTŘEBIČŮ

Při plném nabíjení akumulátoru dochází k situaci, kdy se napětí akumulátoru zvýší nad jeho nominální hodnotu. V případě 12 V olověných akumulátorů může svorkové napětí při plném nabití dosahovat až 14,6 V. Popisovaný subsystém napájení spotřebičů tento problém odstraňuje, protože vstupní napětí měniče se může pohybovat v intervalu od 8 V až do 20 V. Výstupní napětí je v tomto intervalu regulováno na hodnotě 12 V. Výstupní proud měniče je až 10 A. Měnič tedy slouží jako interface mezi akumulátorem a spotřebičem.

Pro řízení měniče byl zvolen jednoúčelový kontrolér LT8705 [1] od společnosti Linear Technology, který v sobě integruje veškeré potřebné periferie, tj. budiče výkonových tranzistorů, diferenciální zesilovače zesilující napěťové úbytky na snímacích rezistorech pro měření vstupního a výstup-

ního proudu, signalizační výstupy a regulační smyčky na výstupní napětí, výstupní proud, vstupní napětí a vstupní proud. Pokud jsou regulační smyčky aktivovány, tak mezi nimi kontrolér dle situace automaticky přepíná. Veškeré nastavení (výstupní napětí, maximální výstupní proud…) je provedeno pomocí rezistorových děličů, které jsou případně doplněny kondenzátory tvořící s rezistory RC články. Vstupní napětí kontroléru je možné volit v intervalu 2,8 V až 80 V a výstupní napětí v intervalu 1,3 V až 80 V. Dle intervalu vstupního napětí je kontrolér schopen pracovat s nižším (režim boost), vyšší (režim buck), případně stejným (režim buck/boost) vstupním napětím vzhledem k výstupnímu napětí.

Při návrhu byl nejprve vytvořen model měniče v simulačním prostředí LTspice. Z hlediska simulace byly nejdůležitější grafické výstupy nasnímané na významných pinech. Například pin VC reprezentující výstup chybového zesilovače a výstupy jednotlivých driverů pro výkonové tranzistory (piny TG1, BG1, TG2 a BG2). Při oživování zkonstruovaného měniče (osazená deska Obrázek 3.1) byly reálné výstupy kontroléru porovnávány s nasimulovanými průběhy modelu. Na měniči bylo provedeno celkem devět měření pro různá vstupní napětí (9 V, 13 V a 20 V) a pracovní režimy (DCM, FCM a BURST [1]). Nejvyšší účinnosti (97,2 %) měnič dosahuje při vstupním napětí 13 V, tedy v buck oblasti a pracovním režimu FCM. Naopak nejmenší účinnost měnič vykazuje v boost oblasti, kde se ale stále účinnost pohybuje nad hranicí 93 % při výstupním proudu 9 A. S ohledem na svorkové napětí olověných akumulátorů se bude účinnost měniče pohybovat nad hranicí 96 %.



Obrázek 3.1: Osazená deska měniče s obvodem LT8705.

4. MPPT REGULÁTOR S NABÍJECÍM AUTOMATEM

Nedílnou součástí každého fotovoltaického systému je regulátor výkonu vyrobené energie z fotovoltaického panelu (dále jen FV panelu). Úkolem tohoto regulátoru je měnit výstupní napětí, respektive proud FV panelu dle intenzity osvětlení panelu tak, aby byl výstupní výkon panelu vždy maximální. Regulátory využívající sledování maximálního bodu výkonu FV panelu se označují MPPT (Maximum Power Point Tracking) [2]. Algoritmů, respektive jejich modifikací využívající metodu MPPT regulace je nespočet. Nejrozšířenějšími a nejjednoduššími MPPT algoritmy jsou Constant voltage, Perturb&Observe a Incremental conductance. Stěžejní pro algoritmus Constant voltage je měření a změna aktuálního napětí FV panelu tak, aby napětí panelu dosahovalo vždy předem pevně nastavené referenční hodnoty. Vzhledem k této interpretaci však prakticky není možné dlouhodobě udržet skutečný bod maximální výkonu, a to i s ohledem na teplotní závislost bodu maximálního výkonu FV panelu.

Algoritmy Perturb&Observe a Incremental conductance využívají pro určení bodu maximálního výkonu první derivaci výkonu podle napětí, respektive proudu. Oba algoritmy využívají přírůstků napětí, případně proudu pro dosažení bodu maximálního výkonu. Perturb&Observe využívá znaménka první derivace pro určení směru přírůstků, tj. porovnává se aktuální a předcházející hodnota výstupního výkonu FV panelu a na tomto základě se rozhoduje, zda se bude k výstupnímu napětí FV panelu přírůstek přičítat, respektive odečítat. Pokud se přírůstek napětí (proudu) dynamicky nemění, tak opět není možné dosáhnout bodu maximální výkonu FV panelu, ale dojde k situaci, kdy je výstupní napětí FV panelu neustále rozmítáno kolem bodu maximálního výkonu. Naopak algoritmus Incremental conductance porovnává první derivaci výkonu podle napětí, respektive

proudu vzhledem k 0 a na základě toho určuje, zda se bude přírůstek přičítat, odečítat, případně nebude v aktuálním kroku reflektován. Metoda Incremental conductance využívá skutečnosti, kdy první derivace výkonu podle napětí (proudu) je nulová právě v lokálním maximu, tedy v bodě maximálního výkonu FV panelu.

Navrhovaný regulátor řízený kontrolérem LT8490 (vycházející z obvodu LT8705) využívá pro dosažení bodu maximálního výkonu nemodifikovanou metodu Perturb&Observe. Výkonová část kontroléru LT8490 včetně nastavení jednotlivých regulačních smyček je naprosto totožná jako v případě LT8705. Nabíjecí automat integrovaný v LT8490 disponuje nabíjecími algoritmy pro olověné a Li-Ion akumulátory. Vzhledem k tomu, že elektrárna, pro kterou byl regulátor navrhován, disponuje 12 V olověnými akumulátory, tak byla při nastavování periferií nabíjecího automatu regulátoru uvažována právě varianta s olověným akumulátorem. Nabíjecí proud regulátoru je až 20 A. Kontrolér LT8490 navíc disponuje jednosměrným (vysílač) UART kanálem, takže je z něj možné vyčítat v jakém stavu se právě nabíjecí automat nachází.

5. MONITOROVÁNÍ AKTUÁLNÍHO STAVU FOTOVOLATICKÉHO POLE

Pro řízení fotovoltaické elektrárny je nutné monitorovat aktuální stav fotovoltaických panelů, tj. jejich svorkové napětí, výstupní proud, případně i vyrobenou energii během časového intervalu, což realizuje popisovaný elektronický subsystém.

Pro měření všech popsaných veličin byl vybrán kontrolér LTC2946 [3], který využívá 12 bitového AD převodníku. Fotovoltaické pole čítá deset panelů, a proto byl tento obvod použit na navrhované desce plošných spojů desetkrát. Naměřená data se z obvodů LTC2946 vyčítají prostřednictvím sběrnice I^2C pomocí kontroléru STM32F103RBT6. ARM kontrolér STM32F103RBT6 dále disponuje sériovou sběrnicí CAN, která slouží pro páteřní komunikaci mezi dalšími subsystémy inteligentního domu.

Program naprogramovaný v kontroléru STM32F103RBT6 vykonává po inicializaci neustále následující činnost: Cyklicky jsou po sběrnici I^2C vyčítány hodnoty měřených veličin z jednotlivých obvodů LTC2946 (napětí, proud, výkon), zabaleny do paketu a na požadavek vystaveny na sběrnici CAN. Případně lze dle požadavku obdrženého po sběrnici CAN od nadřazeného kontroléru měnit limitní hodnoty, tj. minima a maxima měřených veličin, při kterých dochází k vystavení varovné informace na sběrnici CAN.

6. ZÁVĚR

Jednotlivé popsané elektronické subsystémy jsou plně funkční a v budoucnu k tomuto celku přibydou další části určené například pro monitorování vodovodních toků v domě, monitorování teploty jednotlivých částí domu doplněné regulací vytápění, případně další elektronické systémy. Díky páteřní komunikační sběrnici CAN je možné tuto síť dle potřeby neustále rozšiřovat.

REFERENCE

[1] LINEAR TECHNOLOGY. Datasheet LT8705 [online]. 2013. [cit. 2016-2-19]. Dostupný z: http://cds.linear.com/docs/en/datasheet/8705ff.pdf.

[2] MORALES, David Sanz. MPPT Algorithms for Photovoltaic Applications [online]. 2010. [cit. 2016-2-19]. Dostupný z: http://lib.tkk.fi/Dipl/2010/urn100399.pdf

[3] LINEAR TECHNOLOGY. Datasheet LTC2946 [online]. 2014. [cit. 2016-2-19]. Dostupný z: http://cds.linear.com/docs/en/datasheet/2946fa.pdf.

PROPULSION SYSTEM FOR ELECTRIC SCOOTER

Lukáš Doležal

Bachelor Degree Programme (3), FEEC BUT

E-mail: xdolez62@stud.feec.vutbr.cz

Supervised by: Libor Veselý

E-mail: veselyl@feec.vutbr.cz

Abstract: The paper is focused on a design of optimal propulsion system for folding portable scooter. The BLDC motor is used as a propulsion. The paper deals with the design of electronics for sensor control. Two control methods are compared: six-step commutation and vector control. Attention is given to efficiency of both methods and consecutive selection of the best of the two.

Keywords: EEICT, BLDC, PMSM, Vector Control, six-step, MC56f82748, Electric Scooter

1. ÚVOD

Na trhu se v poslední době velmi rozšířil sortiment elektrokol s přídavným motorem. Stejně tak jako klasické jízdní kolo, je možno o elektromotor doplnit i koloběžku. Výhodou je v tomto případě skladnost koloběžky a relativně nízká hmotnost. V této práci bude proveden návrh pohonné jednotky pro upravenou skládací koloběžku. Měnič bude navržen s ohledem na dosažení maximální účinnosti pohonu. Ovládání koloběžky bude možné pomocí otočné páky plynu a uživatel bude o provozu elektronických systémů informován prostřednictvím řídicí jednotky umístěné na řidítkách. Cílem práce je vyvinout koloběžku pro používání k dopravě ve městě.

2. MOTOR

V dané aplikaci je kritickým parametrem hmotnost celého pohonu. Proto byl použit vysokootáčkový modelářský EC motor s převodovkou 1:5. Váha motoru je mnohem nižší než v případě použití klasického nízkootáčkového motoru v trnu kola. Převodovka je tvořena ozubeným řemenem, kde hnací ozubené kolo je připevněno na hřídeli motoru a hnané ozubené kolo je součástí zadního kola koloběžky. Konstrukce byla navržena autorem a speciální díly vyrobeny na zakázku.

3. PŘESTAVBA KOLOBĚŽKY

Pro přestavbu byla vybrána skládací koloběžka SPARTAN. Jedná se o lehkou koloběžku konstruovanou s hliníku. Úpravy spočívaly v odstranění zadní části konstrukce, zajišťující upevnění kola, a nahrazení konstrukcí novou. Ta sestává z hliníkových pásovin přivařených na původní konstrukci. Pod ložem motoru vznikl prostor, který byl využit pro umístění měniče. Baterie byly umístěny pod podlážku koloběžky a jsou zakryty pouzdrem z hliníkového plechu.

4. ELEKTRONICKÁ ČÁST

Pohonná jednotka sestává ze dvou částí: řídicí jednotky a výkonového měniče. Navzájem jsou propojeny robustní sériovou sběrnicí RS485. Komunikace probíhá na vlastním protokolu se zabezpečením pomocí kontrolního součtu.

4.1 ŘÍDICÍ JEDNOTKA

Je navržena pro zobrazování provozních veličin. Mezi sledované veličiny patří: teplota měniče, fázový proud měniče, napětí baterie, aktuální rychlost a ujetá vzdálenost. Jednotka je dále vybavena slotem pro SD kartu a hodinami reálného času pro statistický záznam dat. Požadovaný moment motoru je nastavován pomocí otočného ovladače plynu připojeného do jednotky. Jako zobrazovací prvek byl použit LCD display 2x16 znaků . Ovládání je realizováno pomocí dvou tlačítek. Jednotku řídí procesor ATxmega128D4. Byl naprogramován v jazyce C ve vývojovém prostření Atmel Studio.

4.2 MĚNIČ

Pro koloběžku byl v rámci semestrální bakalářské práce vyvinut měnič. Z důvodu konstrukčních nedostatků a použitého procesoru o malém výpočetním výkonu, bylo přikročeno ke konstrukci nového měniče. Jádrem měniče je procesor MC56F82748. Patří do rodiny procesorů firmy NXP s jádrem DSP56800E. Sdružuje v sobě funkce signálového procesoru a klasického MCU. Tato kombinace je zvláště vhodná pro řízení motorů . Pro debugování byl využit software Free Master. Kód byl napsán ve vývojovém prostředí Codewarior.

Měnič je topologicky řešen jako trojfázový střídač [1]. Každá větev střídače je tvořena horním a dolním spínačem. Jeden spínač je realizován pomocí dvou paralelně spojených tranzistorů NMOS. Budič je realizován obvodem UCC27200, který zajišťuje plovoucí napětí pro buzení tranzistorů horního spínače. Návrh měniče je proveden s požadavkem maximální účinnosti. Z tohoto důvodu byly jako spínací prvky vybrány NMOS tranzistory s odporem R_{DSON} 1, 5$m\Omega$. Proudově je měnič dimenzován na 50A. Deska plošných spojů byla navržena v programu Eagle a vyrobena na zakázku. Tloušťka vodivé vrstvy je 90μm.

4.3 NÁVRH DESKY PLOŠNÉHO SPOJE

Deska plošných spojů byla navržena s ohledem na rozměry prostoru v koloběžce. Geometricky je deska řešena tak, aby byly maximálně potlačeny parazitní indukčnosti, na kterých se při vypínacích dějích indukují napěťové špičky. Důraz je kladen na co nejmenší délku spojů a blokování spínačů keramickými kondenzátory. Nad hlavní deskou měniče je umístěna napájecí kondenzátorová deska. Paralelně je zde zapojeno 19 elektrolytických kondenzátorů . Kondenzátory kompenzují vliv indukčnosti přívodních kabelů baterie. Veškeré proudové špičky a vysokofrekvenční proudy měničem jsou potom hrazeny z těchto kondenzátorů a nedochází ke zbytečnému zatěžování a ohřevu baterie. Spínací tranzistory jsou v pouzdrech D2PAK. Všechny součástky jsou spolu s tranzistory osazeny na vrchní straně a spodní strana desky je určena pro přitlačení na chladič. Teplo z tranzistorů je odvedeno pomocí množství procínovaných prokovů a přes teplo vodivou izolační podložku odvedeno do hliníkového plechu, na kterém je měnič připevněn.

4.4 SIX-STEP ŘÍZENÍ

Metoda sixstep je jednoduchou metodou řízení. V podstatě se jedná o nahrazení mechanického komutátoru. Na jednotlivá vinutí motoru je podle aktuální pozice rotoru připojováno napájecí napětí. V této konkrétní realizaci je poloha rotoru zjištěna pomocí magnetického encoderu AM4096. V jednotlivých komutačních krocích je proud motorem udržován na konstantní hodnotu diskrétním PS regulátorem. Regulační smyčka pracuje s periodou vzorkování 50μs a měření regulovaného proudu je realizováno pomocí snímačů s hallovou sondou. Návrh konstant regulátoru byl proveden metodou frekvenčních charakteristik. Mechanický výkon motoru je potom dán rovnicí [3]:

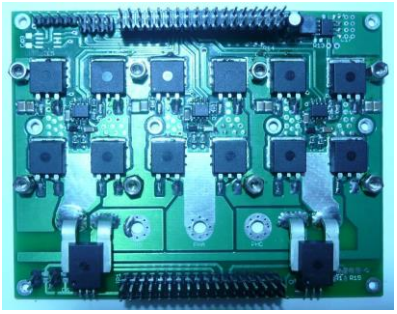
$$P_{el} = \sum_{n=0}^{\infty} U_{i,n} \cdot I_{f,n} \cos \Psi_n \qquad [W] \tag{1}$$

Přičemž n je řád harmonické, $U_{i,n}$ je efektivní hodnota n-té harmonické indukovaného napětí, $I_{i,n}$ je efektivní hodnota n-té harmonické fázového proudu a Ψ_n je vzájemný fázový posuv $I_{i,n}$ a $U_{i,n}$. Pro

maximální činný výkon je tedy nutné, dosáhnout stejného tvaru napájecího proudu a indukovaného napětí. Měřením bylo zjištěno, že tvar indukovaného napětí použitého BLDC motoru je téměř sinu sový. Vyšší harmonické napájecího proudu jsou přeměněny na teplo.

4.5 VEKTOROVÉ ŘÍZENÍ

Nejvyšší moment při daném statorovém proudu získáme při zátěžném úhlu motoru 90° elektrických. Toho lze dosáhnou použitím vektorového řízení. Motor se z hlediska řízení chová jako cize buzený stejnosměrný motor. Měnič pracuje s harmonickými proudy. To má za následek zvýšení ztrát na měniči, ale snížení tepelných ztrát na motoru. Cílem práce je porovnat tyto dvě metody řízení a stanovit, která je v dané konfiguraci efektivnější.



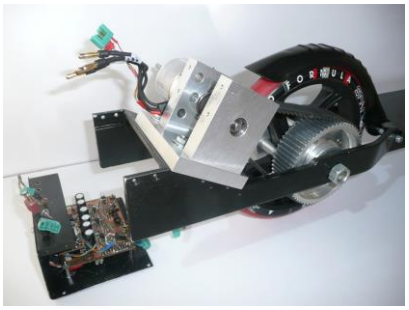
Obrázek 1: Hlavní deska měniče



Obrázek 2: Sestavený měnič



Obrázek 3: Řídicí jednotka



Obrázek 4: Upravená zadní část

5. ZÁVĚR

V době psaní příspěvku probíhalo programování měniče a testování hardware. Koloběžka s řídicí jednotkou již byla vyhotovena a připravena k použití. V práci bude dále provedeno měření a na základě jeho výsledků zvolen optimální algoritmus pro řízení pohonu koloběžky. Dále budou testovány tepelné vlastnosti navrženého chlazení. V neposlední řadě bude proveden test jízdních vlastností.

REFERENCE

[1] Patočka, M.: Vybrané statě z výkonové elektroniky, Brno, 2005

[2] Prokop, L.,Grasblum, P.: 3-Phase PM Synchronous Motor Vector Control, AN1931, 1/2005

[3] Vorel, P.: Synchronní stroje s permanentními magnety, Vyd. 1, Brno, Akademické nakladatelství CERM, 2005, ISBN 80-7204-417-6

ROBOCAR THE AUTONOMOUS DRIVING

Miroslav Hnát

Bachelor Degree Programme (3), VUT

E-mail: xhnatm00@stud.fit.vutbr.cz

Supervised by: Petr Petyovský

E-mail: petyovsky@feec.vutbr.cz

Abstract: Ve své práci jsem se zabýval požadavky kladenými na automatické řízení. Rozvrhnul jsem reálné řešení a vybíral vhodný SW nástroj pro simulaci a realizaci dané problematiky. Dále jsem řešil model pro simulaci řízení pohonu, brzdění a natáčení kol ve volném prostranství. Definoval jsem možné chybové stavy a navrhl nutné ochrany. Další důležitou částí byla identifikace parametrů reálného automobilu a ověření tak správnosti modelu pomocí simulace. Následovalo zhodnocení případných odchylek od modelu, což mělo za cíl ověřit reakce reálného systému. Celé řešení jsem pak implementoval do jednotlivých komponent a srovnal jednotlivé komponenty s výsledky simulace. Na závěr jsem zhodnotil dosažené výsledky a navrhl další možná vylepšení.

Keywords: Autonomní řízení, optimální trajektorie

1. ÚVOD

Ve své bakalářské práci se zabývám simulací, kompletní realizací a implementací autonomního řízení na klasické, sériově vyráběné vozidlo. Má práce obsahuje teoretickou i praktickou část. V teoretické části se zabývám specifickými požadavky pro optimální trajektorii jízdy vozu, výběrem vhodných komponentů pro reálnou implementaci. Hlavním prvkem jsou data od snímačů, po jejímž zpracování se bude PC ve voze rozhodovat, jak v dané situaci zareagovat.. Tato data jsou sbírána a následně analyzována pro odladění chyb. Google dnes již pracuje na konceptu plně autonomních vozů a společnost má několik prototypů, které jsou nasazeny do silničního provozu. V mé úloze pracuji s modelem pro simulaci řízení pohonu a simuluji navržené řešení v simulátoru. Zabývám se tedy rozdíly v modelu pro simulaci a reálného, sériově vyráběného vozu. Dále se zde zabývám sběrem dat od jednotlivých snímačů připojených na vozidle a simulací zvolené trajektorie. V teoretické části promýšlím možné chybové stavy a navrhuji nutné ochrany. Teorii a navržené metody pak vyzkouším v praxi a implementuji do vozu. Na závěr prezentuji dosažené výsledky a navrhuji další možná vylepšení.

2. POŽADAVKY AUTONOMNÍHO ŘÍZENÍ

Potkat se dnes v logistice s autonomními vozy, není nic zvláštního. V určitém rozsahu se již řadu let využívají ve vyspělých, skladovacích a výrobních provozech, kde zajišťují plně nebo částečně automatizované přesouvání zboží, dílů nebo materiálu. V dnešní velmi husté dopravě, především ve městech, se řízení automobilu stává velmi citelnou ztrátou času. Pro některé řidiče, kteří chtějí trávit čas za volantem jinou činností např prací, čtením emailů, věnováním se rodině, sledováním novinek či filmů, by autonomní řízení mohlo do budoucna být nepostradatelnou součástí. Autonomním řízením se automobilový průmysl zabývá již několik let. Velké automobilky (Volvo, Audi, Mercedes..) investují peníze do vývoje autonomního řízení. Například automobilka Volvo pracuje na konceptu, kdy si obsluha automobilu zvolí, kolik procent jízdy má odřídit samotný automobil a poté dle tohoto parametru, PC v automobilu, které dostává signály od GPS, radaru, lidaru, určí, kdy je bezpečné předat řízení od obsluhy řidiče k automobilu. Po ujeté vzdálenosti, kterou si v procentech zvolil řidič, automobil vyzve řidiče k převzetí ovládání. Pokud tak řidič

neučiní do určité doby, auto vyhodnotí tuto situaci jako „chybový stav“ a přejede na bezpečné místo, kde zastaví a vyčká na reakci od řidiče.

Z policejních statistik dopravních nehod vychází, že za minulý rok (2015) policie zaznamenala 195 851 nehod. Když se na toto číslo podíváme blíže, zjistíme, že nejčastější hlavní příčinou nehod řidičů motorových vozidel je nesprávný způsob jízdy konkrétně 49 807(63,7% z počtu těchto nehod).

Mezi pět nejčastějších příčin nehod řidičů motorových vozidel patří:

Pořadí	Pět nejčastějších příčin dopravních nehod	Počet nehod
1	řidič se plně nevěnuje řízení	**15 311**
2	nedodržení bezpečné vzdálenosti za vozidlem	**7 514**
3	nesprávné otáčení vozidla nebo couvání	**7 199**
4	jiný druh nesprávného způsobu jízdy	**7 097**
5	nepřizpůsobení rychlosti stavu vozovky	**5 682**

Nejdůležitější součástí pro autonomní řízení je snímání okolí kolem vozu. Snímání prostoru před automobilem je zajištěno lidarem, který snímá okolí v úhlu 270°, pro snímání okolí jsou na voze umístěny radary a kamery. Další, řekl bych možná ještě důležitější částí autonomního řízení, je vyhodnocení a reakce všech předvídatelných i nepředvídatelných situací na silnici. Křižovatky, semafory, chodce, cyklisty, auta v okolí, jízdní pruhy, nečekané překážky, to vše a mnohem více, musí počítač, který je v tomto případě umístěný v kufru automobilu, vzít v potaz, vyhodnotit a předat správný pokyn. Toto vyhodnocení musí počítač předat automobilu v co nejkratším čase, nejlépe „real-time“. Dlouhá odezva předání informace může vést ke špatnému rozhodnutí nebo dokonce k havárii vozu.

PC je myšleno jako klasický desktopový počítač s procesorem s výkonnější GPGPU pro výpočty a celkově pro zpracování obrazu. S vozem je PC propojeno sběrnicí CAN (**C**ontroller **A**rea **N**etwork), pomocí níž komunikujeme s jednotlivými komponenty. CAN sběrnici využíváme i pro komunikaci s jednotlivými snímači např. s radarem a drobnými snímači. Dále lidar a kamera jsou propojeny přes klasický síťový protokol, kvůli velkému toku dat. Pro představu, minutový záznam ze všech senzorů a kamery zabírá cca 350Mb.

Ovládaní pedálů brzdy, je řešeno mechanicky, pomocí servomotorů a táhel. Ovládání pedálu plynu bude řešeno elektronicky přes spínač a potenciometr. K tomuto řešení jsme přistoupili, jelikož není úplně jednoduché, reverzním inženýrstvím prolomit zabezpečení a ovládání automobilu. Automobilky, které jsme postupně oslovovali pro získání vozu a také pro bližší informace komunikace jednotlivých komponentů po sběrnici CAN, nejsou příliš shovívavé. Což je pochopitelné, z pohledu jejich investování do vývoje tímto směrem. Nejschůdnější volbou byla automobilka Hyundai, která nám částečně vyšla vstříc. Konkrétně model i40 s automatickou převodovkou[obr.1].



Obrázek 1: Roboauto ARTIN Hyundai i40

3. CO K TOMU POUŽÍVÁME?

Dnes rozšířené technologie jako je třeba ROS[3] či Docker[4] jsou nezbytnou součástí pro náš vývoj na autonomním řízení. Na našem PC běží OS Linux Ubuntu.

3.1. ROS

ROS je v současnosti nejpoužívanější platforma pro vývoj robotického softwaru, který využívá široké spektrum vývojářů od amatérů přes komerční využívání až po špičkové univerzity jako MIT apod. Je psán v jazyce C++ a architektura je založena na objektovém principu, kdy si uživatel doslova vyskládá svou aplikaci pomocí modulů. Ve své podstatě je ROS sada open-source softwarových nástrojů, knihoven a driverů vytvořených za účelem zrychlení a zjednodušení vývoje robotických aplikací. V ROS je k dispozici mnoho základních prvků, které jsou při tvorbě robotických aplikací potřeba, od ovladačů pro různé typy senzorů a hardwaru přes knihovny na zpracování obrazu, kalibraci senzorů, řešení kinematických úloh, plánování trajektorií či mapování prostoru, až po vizualizaci, ukládání dat a networkingu.

Velkou výhodou je, že jednotlivé moduly na sebe nejsou závislé, což v praxi znamená, že když by náhodou jeden modul přestal fungovat, nezničí to celé fungování vozu. Navíc ROS nahrává data z jednotlivých jízd, což nám následně pomáhá k analýze.

Jeden z modulů ROS je i simulátor, ve kterém si každý vývojář z týmu na svém stroji může odsimulovat chování modelu (setrvačnost, váhu, simulace překážek atd.).

3.2. DOCKER

Docker se řadí mezi takzvané softwarové kontejnery. Docker je skutečně často přirovnáván k tradičním přepravním kontejnerům, které je možné standardizovat a poté přepravovat různým způsobem. Do takového SW kontejneru je, zjednodušeně řečeno, možné vložit jednotlivé aplikace včetně nastavení a dalších úprav. Aplikace pak žije v tomto kontejneru a je jí možné jednoduše přenášet a nasazovat kamkoliv je potřeba. Kam vytvořený kontejner doručíme, tam bude fungovat. Vývojář může takto vytvořený kontejner hned přesunout do produkčního prostředí. Nasazení je rychlé a správa jednodušší.

4. ZÁVĚR

Autonomní řízení do budoucna může ušetřit spoustu času mnoha lidem, kteří se při řízení mohou věnovat pro nich potřebným činnostem. Vytížení lidé se tak budou moci věnovat práci i během jízdy, což jim ušetří mnoho času a nervů se zmatky a nedorozuměními v dopravě. Především v logistice můžeme najít několik uplatnění, ať už k převozu klasických materiálů či převozu nadměrných nákladů, které probíhají převážně v noci. I z tohoto důvodu je doprava bezpečnější a nehrozí žádné mikrospánky či únava řidiče. Myslím si, že v autonomním řízení je budoucnost a to i z hlediska bezpečnosti, velmi významný pokrok, kdy ze statistik bylo zřejmé, že nejčastější nehody jsou způsobeny nedostatečným věnováním se řízení.

REFERENCE

[1] 2010 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, Anchorage. Conference proceedings. Piscataway: IEEE, 2010. ISBN 9781424450404.

[2] 2011 IEEE INTELLIGENT VEHICLES SYMPOSIUM (IV), Baden-Baden. 2011 IEEE Intelligent Vehicles Symposium (IV 2011): Baden-Baden, Germany 5-9 June 2011. Piscataway, NJ: IEEE, 2011. ISBN 9781457708909.

[3] Docker. *Docker*. [online]. 15.3.2016 [cit. 2016-03-15]. Dostupné z: *https://www.docker.com/what-docker*

[4] ROS. *ROS*. [online]. 15.3.2016 [cit. 2016-03-15]. Dostupné z: *http://www.ros.org/core-components/*

BATTERY ELECTRIC VEHICLE CHARGING AND OPERATING COSTS ANALYSIS

Tomáš Zajac

Bachelor Degree Programme (3.), FEEC BUT

E-mail: xzajac02@stud.feec.vutbr.cz

Supervised by: Lukáš Radil

E-mail: radil@feec.vutbr.cz

Abstract: This paper deals with issues of battery electric vehicle charging. Key mechanisms are defined and the possibilities of charging station designs are presented. Active charging stations, their advantages and impacts on grid are discussed along with option of partial decoupling of the load from the grid, achieved by the application of energy storage elements – buffer accumulation. Furthermore, calculation methodology for operational cost analysis is proposed, programmed and the overall cost of the investment is discussed.

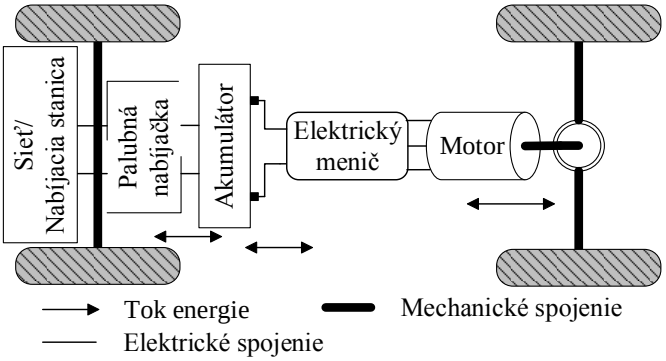
Keywords: electric vehicle, battery, charging, charging station, operating costs, analysis

1. ÚVOD

Elektromobilita je celosvetovo rastúci trend**.** Kvalitne navrhnutá a rozvinutá nabíjacia infraštruktúra je základnou podmienkou pre úspešné rozšírenie komerčného používania elektrických vozidiel (EV). Prieskum trhu uskutočnený v roku 2010 medzi iným ukázal, že až 81 % respondentov nie je schopných definovať, kde a ako by svoje EV nabíjali. Preto je nutné jasne kategorizovať aspekty nabíjacieho procesu. S kúpou EV je okrem technických aspektov úzko spojená aj otázka výhodnosti investície.

2. ELEKTROMOBILY

Pojmom elektromobil je možné vo všeobecnosti označiť všetky vozidlá, ktoré úplne alebo čiastočne využívajú premenu elektrickej energie na mechanickú, ktorú následne využívajú na svoj pohyb. Medzi takéto vozidlá patria okrem čisto elektrických vozidiel s akumulátorom (BEV – Battery Electric Vehicle) aj hybridy. Tie sa delia na pripojiteľné hybridné (PHEV – Plug-In Hybrid Electric Vehicle) a hybridné (HEV - Hybrid Electric Vehicle) vozidlá. Z pohľadu nabíjania a nabíjacej infraštruktúry sú zaujímavé hlavne BEV, keďže kapacita ich akumulátora je výrazne vyššia ako u hybridných vozidiel. Na Obrázku 1 je znázornené zjednodušené usporiadanie BEV [1].



Obrázok 1: Architektúra BEV [1].

1.1. AKUMULÁTORY

V prípade čisto elektrických vozidiel je technologické riešenie uskladnenia energie definujúcim parametrom ich funkčnosti. Výber akumulátora ovplyvňuje parametre vozidla ako napríklad dojazd, hmotnosť a nabíjací čas. Základné parametre uvažované pri porovnávaní akumulátorov sú menovitá kapacita, merný výkon a merná energia na jednotku hmotnosti (objemu). Z pohľadu prevádzky je najvýznamnejším parametrom úroveň nabitia (SOC – state of charge). Je definovaná ako percentuálny podiel okamžitej kapacity k menovitej kapacite batérie. Životnosť akumulátora je možné sledovať dvoma parametrami – stredná doba života udáva životnosť pri prevádzke v menovitých podmienkach a SOH (state of health) popisuje degradáciu menovitej kapacity batérie v čase. Prehľad najpoužívanejších akumulátorov v EV je uvedený v Tabuľke 1 [2].

Typ	NiMH	$LiCoO_2$	$LiMn_2O_4$
Anóda	kovové hydridy	amorfný uhlík	
Katóda	$Ni(OH)_2$	$LiCoO_2$	$LiMn_2O_4$
Kapacita článku (Ah)	6,5	4,4	
Merná energia ($Wh.kg^{-1}$)	63 - 75	90-140	160
Merný výkon ($W.kg^{-1}$)	220	760	1300
Počet cyklov (-)	600 - 1200	3000 - 20000	
Samovybíjanie ($\%.mes^{-1}$)	30	5-10	5-10

Tabuľka 1: Prehľad akumulátorových článkov [2].

3. NABÍJANIE A NABÍJACIE STANICE

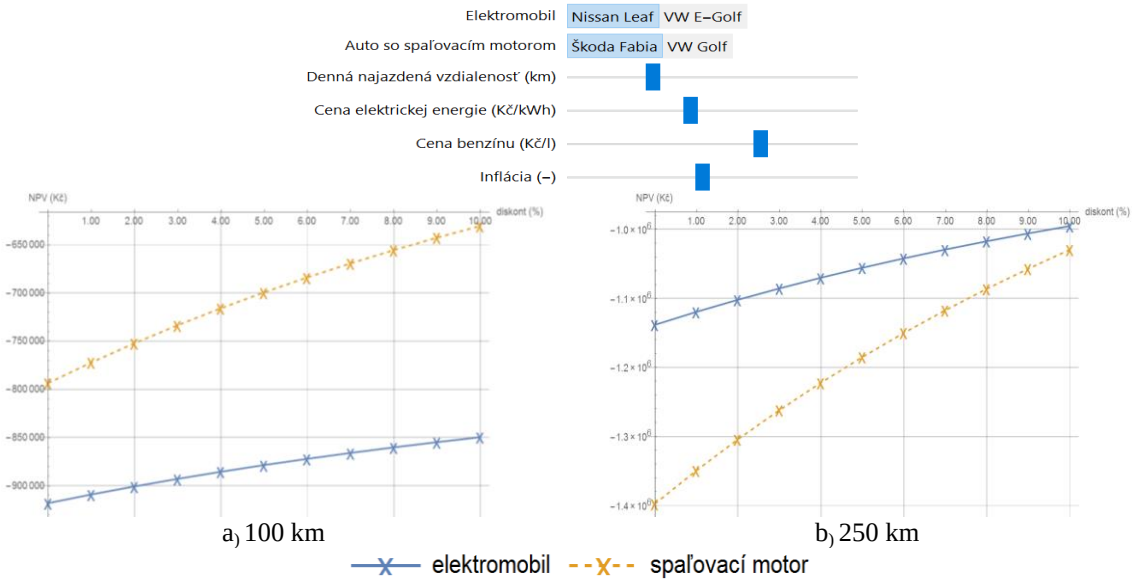
Nabíjanie je možné realizovať v aj domácnostiach na úrovni nízkeho napätia (NN), no z dôvodu zefektívnenia a skrátenia nabíjacieho procesu sú budované nabíjacie stanice (NS). Tie sa dajú kategorizovať na základe nabíjacieho príkonu, času nabíjania (Tabuľka 2), aplikovaných konektorov a štandardov a interakcie vozidlo-stanica-distribučná sústava. Striedavé NS najrozšírenejšie a ponúkajú nabíjanie príkonmi do 22 kW, za použitia konektorov ako Mennekes alebo SAE (USA). Jednosmerné NS sú určené na rýchlonabíjanie poskytujú vďaka AC/DC konverzii priamo v stanici vyššie nabíjacie príkony. Celosvetovo najrozšírenejší štandard pre DC nabíjanie je CHADeMO. Z pohľadu interakcie EV s nadradenou sústavou sa dajú definovať tri prípady – tok energie len do vozidla (V0G), kontrolovaný tok energie do vozidla na základe požiadaviek siete (V1G) a obojsmerný tok energie (V2G) [3].

Ako najpraktickejšie riešenie sa v súčasnosti ponúka koncept aktívnych nabíjacích staníc. Jedná sa o AC a DC, poprípade kombinované nabíjanie z jedného zdroja za súčasnej možnosti odľahčenia siete pomocou vnútornej akumulácie v stanici – tzv. buffer akumulátora, čo umožňuje využitie lacnej elektrickej energie počas nočných hodín a následné odľahčenie siete počas hodín špičkových. Stanice tohto druhu taktiež podporujú V2G systém a efektívnu implementáciu obnoviteľných zdrojov do nabíjacieho procesu [4].

4. POROVNANIE PREVÁDZKOVÝCH NÁKLADOV

Z pohľadu užívateľa je zásadným aspektom kúpy nielen otázka nabíjania, ale tiež výhodnosť investície do elektromobilu. Jednou z metód na jej stanovenie je výpočet čistej súčasnej hodnoty (NPV), ktorá vhodne odzrkadľuje vývoj investície v čase a prezentuje reálnu výšku nákladov v závislosti na diskonte. Za účelom porovnania bol vytvorený program v prostredí Wolfram Mathematica. Výpočet je realizovaný na základe nasledujúcich vstupných parametrov – náklady na pohonné hmoty sú 31,78 Kč.l-1, cena za kWh elektrickej energie je 4,55 Kč (nabíjanie na úrovni NN), medziročná inflácia 2 %, pri uvažovaní nákladov na údržbu auta so spaľovacím motorom 10 000 Kč, elektromobilu 7 000 Kč ročne a 8-ročnej životnosti oboch vozidiel. Program je možné doplniť o ďalšie EV či auto so spaľovacím motorom, všetky vstupné veličiny sú prispôsobiteľné na základe

individuálnych potrieb. Cena elektrickej energie môže predstavovať aj cenu za nabíjanie komerčnou nabíjacou stanicou, čím sú pokryté všetky možnosti prevádzky EV. Príkladom výstupu z programu je Obrázok 2, ktorý predstavuje porovnanie Nissanu Leaf a Škody Fabia 1,2 TSI, a to pre dve možné denné najazdené vzdialenosti – 100 km a 250 km.



Obrázok 2: Ukážka porovnania NPV v závislosti na diskonte.

5. ZÁVER

Pri súčasných požiadavkách distribučnej sústavy a náraste počtu elektrických vozidiel sa ako optimálna možnosť realizácie nabíjacej infraštruktúry ponúka koncept aktívnych nabíjacích staníc. Pomocou buffer akumulátora prispievajú k stabilizácii siete a implementáciou obnoviteľných zdrojov zas ku zníženiu environmentálneho dopadu výroby energie. Z porovnania prevádzkových nákladov (Obrázok 2) vyplýva, že pri 100 km najazdených denne je investícia do EV nevýhodnejšia. So zvyšujúcou sa najazdenou vzdialenosťou rozdiel v investičných nákladoch klesá. Pri vzdialenosti 250 km denne, ktorá presahuje dojazd väčšiny EV na jedno nabitie, je investícia výhodnejšia. Tu sa však do popredia dostáva otázka dostupnosti nabíjacej infraštruktúry. Odhliadnuc od investičných nákladov, EV majú široké využitie, a to najmä v mestách, kde dochádza k silnému lokálnemu znečisteniu.

REFERENCE

[1] KUMAR, Lalit; GUPTA, Krishna Kumar a JAIN, Shailendra. 2013. Architecture and configuration of electrified vehicles: A review. In: *2013 IEEE International Symposium on Industrial Electronics* [online]. IEEE, [cit. 2015-10-17]. DOI: 10.1109/ISIE.2013.6563825. ISBN 978-1-4673-5193-5.

[2] LINDEN, David a Thomas B REDDY. 2002. *Handbook of batteries*. 3rd ed. New York: McGraw-Hill, ISBN 00-713-5978-8.

[3] FALVO, Maria Carmen, Danilo SBORDONE, I. Safak BAYRAM a Michael DEVETSIKIOTIS. 2014. EV charging stations and modes: International standards. *2014 International Symposium on Power Electronics, Electrical Drives, Automation and Motion* [online]. IEEE, 1134-1139 [cit. 2016-03-08]. DOI: 10.1109/SPEEDAM.2014.6872107. ISBN 978-1-4799-4749-2.

[4] SIMONIK, Petr; HAVEL, Ales; HROMJAK, Michal a CHLEBIS, Petr. 2012. *Active Charging Stations for Electric Vehicles Charging* [online]. [cit. 2015-10-30]. ISSN 1559-9450.

MULTIPURPOSE ELECTROPLATING POWER SUPPLY

Michal Pinďák

Bachelor Degree Programme (3), FEEC BUT

E-mail: xpinda02@stud.feec.vutbr.cz

Supervised by: Petr Procházka

E-mail: prochazkap@feec.vutbr.cz

Abstract: The aim of this project is to design a four-quadrant bridge converter which could be used even for a galvanizing bath power supply. Nowadays the electroplating is one of the most used methods for surface treatment to all sorts of metal components, especially in electrotechnic industry. The article shows all basic structual parts of device, their parameters and the purpose of their use. The structure of an inverter control is implemented by program called LabVIEW.

Keywords: electroplating; four-quadrant bridge converter; myRIO; power supply; LabVIEW

1. ÚVOD

V dnešní době již nelze nalézt průmyslovou aplikaci, jež by nezahrnovala užití kovových materiálů a slitin. Toto nemalé zastoupení kovů v průmyslu však klade velké nároky na jejich vlastnosti, mezi které patří zejména odolnost vůči povětrnostním vlivům jako koroze a oxidace. Z hlediska elektrotechnického průmyslu je naopak podstatná kvalitní elektrická vodivost. Těchto vlastností lze dosáhnout použitím tzv. ušlechtilých kovů jako měď, zlato, stříbro, platina apod. Za nevýhodu těchto materiálů lze považovat jejich vysokou cenu.

Jelikož je mnohdy dostačující pouze povrchová odolnost výrobku vůči korozi, nebo pouze jeho elektricky vodivější kontaktní styk, lze v tomto případě využít cenově lukrativní metody galvanického pokovení. Výhoda této metody spočívá zejména ve výrazně nižší spotřebě ušlechtilého materiálu. V praxi pak lze použít jako nosný základ levnou ocel a na její povrch nanést tenkou vrstvu ušlechtilého materiálu. Tato metoda navíc umožňuje nanášení několika ušlechtilých materiálů ve více vrstvách.

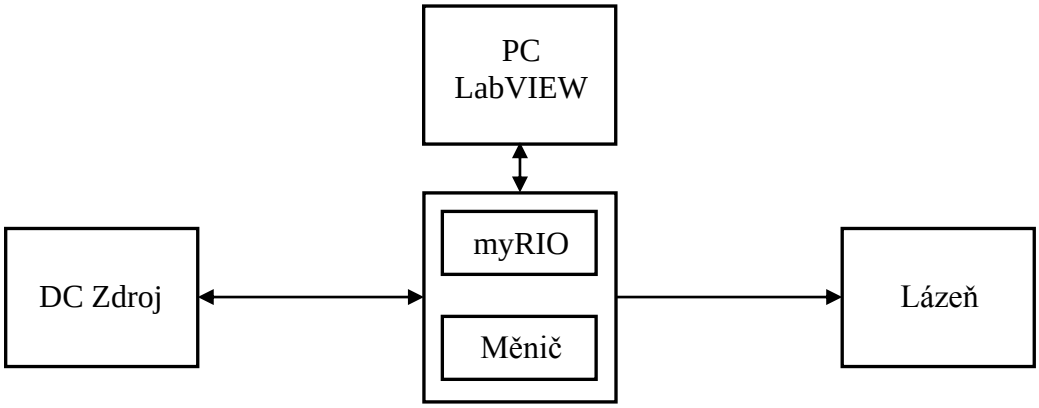
2. PRINCIP GALVANICKÉHO POKOVENÍ

Využívá se zde fyzikálně-chemického procesu zvaného elektrolýza. Ta spočívá v ponoření elektrod (katody a anody) do vhodného elektrolytického roztoku a následném připojení stejnosměrného napětí. Toto napětí vytvoří elektromagnetické pole mezi elektrodami a následně vzniklý elektrický proud vyvolá usměrněný pohyb iontů v roztoku. Kladné ionty jsou přitahovány směrem ke katodě a ionty záporné naopak směrem k anodě. Principiálně zde tedy dojde oproti kovovému vodiči k

přenosu látek.

3. POŽADAVKY NA ZAŘÍZENÍ

Zdroj pro galvanizační lázeň bude složen ze dvou základních částí. První část bude tvořena stejnosměrnými napájecími zdroji XP Power o výkonu 3 kW v paralelním chodu. Tyto zdroje jsou přímo dodány od zadavatele projektu. Druhou část zařízení tvoří univerzálním čtyřkvadrantový měnič včetně řídícího rozhraní myRIO. Každá z těchto dvou částí bude umístěna v samostatné montážní skříni typu „Rack". Tato dělená konfigurace posléze umožňuje víceúčelové využití samotného měniče. Tento měnič bude dle požadavků navrhován se značnou výkonovou rezervou.



Obrázek 1: Blokové schéma zdroje galvanizační lázně

4. ZÁKLADNÍ ČÁSTI ZAŘÍZENÍ

4.1. NAPÁJECÍ ZDROJ XP POWER HDS3000PS60

Jedná se o spínaný programovatelný AC/DC zdroj firmy XP Power o jmenovitém výstupním výkonu 3 kW s účinností až 93 %. Výstupní napětí a proud jsou regulovatelné od 0 do 105 % jmenovitého výkonu. Tato regulace bude v našem případě realizována pomocí analogového výstupního portu zařízení myRIO a umožní regulovat amplitudu spínaného napětí měniče. Taktéž bude možné omezit proud měničem na požadovanou hodnotu.

Parametry zdroje : U_{VSTUP} = 90 - 264 VAC, $U_{VÝSTUP}$ = 0 - 60 VDC, $I_{VÝSTUP}$ = 0 - 50 ADC [1]

4.2. ROZHRANÍ NI MYRIO - 1900

Zařízeni myRIO tvoří rozhraní mezi řídícím počítačem a samotným měničem. Na řídícím počítači jsou pomocí programu LabVIEW generovány požadované řídící signály, které jsou následně myRIEm zpracovány a adresovány na požadované hradlo IGBT. Pomocí myRIA je taktéž možné přenášet informace z měniče do počítače.

V našem případě bude využito digitálních výstupů myRIA pro řízení hradel IGBT a analogových výstupů pro řízení napájecího zdroje XP Power. Digitální vstupy budou naopak využity pro detekci havarijních stavů měniče a vstupy analogové pro měření napětí a proudu. Snímání vysokých napětí a proudů bude realizováno pomocí proudových a napěťových čidel firmy LEM.



Obrázek 2: Zařízení NI myRIO - 1900 [2]

4.3. MĚNIČ

Základ našeho měniče bude tvořen 2 výkonovými polomůstky IGBT od firmy POWEREX. Tyto polomůstky zvládají v DC režimu až 300 A při závěrném napětí 1200 V. Každý z dílčích tranzisto-

rů navíc obsahuje potřebnou nulovou diodu a termistor NTC pro nutnou kontrolu teploty. Ačkoli nám dle požadavků původně dostačovalo závěrné napětí pouze 600 V, volba 1200 V tranzistorů byla učiněna z důvodu menších tepelných odporů pouzdra. Tím je možné měnič snáze uchladit.

Maximální frekvence použitých tranzistorů v našem případě není kritická, jelikož bude čtyřkvadrantový měnič primárně pracovat v tzv. DC režimu. Nejedná se tedy o standartní řízení PWM. Z hlediska galvanizačních procesů je nutné v libovolném okamžiku měnit polaritu napětí na výstupu měniče a co nejpřesněji dodržet čas práce při dané polaritě. Z toho důvodu je zvolen elektronicky ovládaný čtyřkvadrantový měnič, u něhož lze dle požadavků zadavatele nastavit přesný algoritmus práce. Zařízení je však značně výkonově naddimenzované a zvládne frekvence i v řádu kHz.

Jako budiče jsou zvoleny hotové moduly od firmy SEMIKRON. Ty zajišťují požadované galvanické oddělení, ochranu výkonových tranzistorů před zkratem, dead time apod. Jelikož si však výstupní řídící napětí myRIA a vstupní řídící napětí budiče nejsou rovny, je nutné užít tzv. emitorových sledovačů.



Obrázek 3: Polomůstek CM300DX-24S1 (vlevo) [3], Budič Skyper 32 R (vpravo) [4]

5. ZÁVĚR

Úkolem tohoto článku je přiblížení základních konstrukčních prvků realizovaného univerzálního zdroje pro galvanizační lázeň. Tento projekt jsem si vybral z důvodu velkého zájmu o oblast výkonových spínaných zdrojů, měničů a jejich řídicích systémů. Prozatím byl proveden test řízení dodaných spínaných zdrojů a návrh desky DPS. Do této desky bude přímo zapájena řídící část výkonových tranzistorů IGBT a budou zde nachystány konektory pro nacvaknutí budičů. Dále zde budou taktéž osazeny bezindukční polypropylenové a elektrolytické filtrační kondenzátory.

Velké množství výkonových prvků (stykače, přídavné napájecí zdroje apod.) bude zvoleno s ohledem na rozměry až při samotné zástavbě měniče do skříní „Rack".

REFERENCE

[1] XP POWER. *HDS3000*. [online]. 29.9.2014 [cit. 2016-03-10]. Dostupné z: http://www.farnell.com/datasheets/1858130.pdf

[2] NATIONAL INSTRUMENTS. *myRIO - 1900*. [online]. 13.8.2013 [cit. 2016-03-10]. Dostupné z: http://www.ni.com/pdf/manuals/376047a.pdf

[3] POWEREX. *CM300DX-24S1*. [online]. 30.12.2013 [cit. 2016-03-10]. Dostupné z: http://www.farnell.com/datasheets/1794525.pdf

[4] SEMIKRON. *SKYPER 32 R*. [online]. 26.05.2009 [cit. 2016-02-19]. Dostupné z: http://www.farnell.com/datasheets/482336.pdf

THREE-PHASE INVERTER FOR HIGH SPEED INDUCTION MOTOR

Petr Veselý

Bachelor Degree Programme (3), FEEC BUT

E-mail:veselly.petr@seznam.cz

Supervised by: Petr Procházka

E-mail:prochazkap@feec.vutbr.cz

Abstract: This work deals with designing a power part of a three-phase inverter for induction motor. It contains choosing of some important parts such as IGBT modules or drivers, which depends on parameters, that are mentioned or calculated. Also a schema is included. General principle and brief theory is described in this paper too.

Keywords: IGBT, Three-phase inverter, induction motor

1. ÚVOD

Frekvenční měniče, jakožto nástroj k nejefektivnější regulaci velmi používaných asynchronních motorů, zažívají velký rozmach. Ten je podporován nejen onou potřebností a využitím frekvenčních měničů, ale i úspěšným vývojem polovodičových součástek, které jsou v těchto zařízeních stěžejní. Jejich vývoj především v oblasti zkracování vypínacích a zapínacích časů umožňuje dosažení stále vyšších spínacích. Tento fakt tedy vede i k rozvoji vysokootáčkových motorů, které jsou pak těmito měniči napájeny. Návrh takového silového obvodu měniče pro vysokootáčkový asynchronní motor o výkonu 50 kW je níže popsán.

2. PRINCIP FUNKCE

Frekvenční měnič slouží mimo jiné k napájení asynchronních motorů, u nichž je nejefektivnějším způsobem regulace otáček pomocí změny frekvence první harmonické napájecího napětí.

Silová část měniče se skládá z několika základních bloků. Na vstupu měniče je usměrňovač. V našem případě, kdy realizujeme měnič pro vysokootáčkový asynchronní motor o výkonu 50 kW, se musí jednat o třífázový usměrňovač. Za usměrňovačem se nachází LC-filtr, nebo sběrací kondenzátor, který funguje jako výkonový špičkový detektor. Tento LC-filtr, nebo sběrací kondenzátor, nám zajistí minimální zvlnění napětí ve stejnosměrném meziobvodu měniče. K tomuto meziobvodu jsou již připojeny tři větve střídače, z jejichž středu jsou vyvedeny tři fáze pro napájení asynchronního motoru. Díky vhodnému řídicímu algoritmu jsme schopni na těchto výstupních fází generovat pulsní napětí s proměnou střídou a frekvencí. Díky tomuto faktu jsme schopni měnit frekvenci a velikost první harmonické výstupního napětí.

3. KONSTRUKCE

3.1. PARAMETRY

Měnič je třífázový a je napájen ze sítě 230/400 V. Jelikož je jeho nominální výkon 50 kW, můžeme určit efektivní hodnotu fázového proudu podle vzorce:

$$I_{f,ef} = \frac{P_m}{3 * U_f * \eta_m * cos\varphi} = \frac{50\,000}{3 \cdot 230 \cdot 0{,}8 \cdot 0{,}8} = 113{,}22\ A \qquad [2\ ]\ (3.1)$$

Na tuto hodnotu je potřeba dimenzovat silové prvky. Také je potřeba součástky dimenzovat napěťově. Jmenovitou maximální hodnotu napětí na meziobvodu nám určuje vzorec:

$$U_d = U_{f,ef} \cdot \sqrt{2} \cdot \sqrt{3} = 230 \cdot \sqrt{2} \cdot \sqrt{3} = 563{,}4\ V \tag{3.2}$$

V důsledku konečné kapacity kondenzátoru bude střední hodnota napětí o něco nižší, my však musíme počítat s překmity na tranzistorech, jenž jsou způsobené parazitní indukcí a s nestálostí sítě. Proto silové prvky dimenzujeme na 1,2 kV. U sinusové pulsně šířkové modulace musí být splněno, že kmitočet první harmonické výstupního napětí musí být řádově nižší než frekvence spínání tranzistorů. Jelikož zařízení musí být schopno generovat napětí o frekvenci až 2 kHz, což je nestandartně vysoký kmitočet, musí být frekvence spínání tranzistorů také vyšší, než je běžné. Hodnota frekvence spínání byla zvolena 30 kHz, což nám splní podmínku řádového odstupu, a zároveň spínací ztráty ještě nejsou při vhodném spínání příliš velké. Z toho plyne hlavní kritérium pro výběr tranzistorů. Nejenže budou muset být dimenzovány na vysoké napětí a proud, ale také je potřeba je vybrat s ohledem na co nejkratší doby spínání, protože spínací ztráty jsou počtu sepnutí úměrné. Měnič bude sloužit pro napájení laboratorních a funkčních vzorků vyvíjených asynchronních motorů, proto musí být značně předimenzován.

3.2. SEMIKRON 400GB 125D

Pro spínání budou použity 3 výkonové moduly Semikron 400GB125D, každý pro jednu fázi. Každý modul obsahuje dva tranzistory (jedná se o horní a dolní spínač) a k nim dvě nulové diody, které slouží k vedení proudu po vypnutí tranzistoru.

Tato část měniče bude vykazovat největší ztráty, které se při sečtení spínacích ztrát a ztrát vedením dají vyčíslit na hodnotu 2133 W, což odpovídá přibližně 4 % výkonu. Takový ztrátový výkon bude pravděpodobně nutné chladit aktivně.

3.3. SKYPER R32

Tranzistory se řídí pomocí napěťových signálů 15 V a -8 V. Ty jsou přivedeny z řídící desky, musí ale být galvanicky odděleny. K tomuto účelu slouží budič, který zároveň poskytuje tranzistorům základní ochrany. Jako budič bude použit Semikron skyper 32 R, který umožňuje buzení obou tranzistorů ve větvi (horního i dolního). Mezi základní ochrany patří podpěťová ochrana a zkratová saturační ochrana. Také zajišťuje generaci potřebné ochranné doby a potlačuje krátké rušivé pulzy.

3.4. NAPÁJENÍ

Pro napájení je použit modul s neřízeným šestipulsním usměrňovačem a posléze LC-filtr s elektrolytickými a bezindukčními kondenzátory k potlačení parazitní indukce. Indukčnost tlumivky zjistíme pomocí vzorce 3.3

$$L = 0{,}00904 \cdot \frac{U_a}{\Delta I_z \cdot 2 \cdot \pi \cdot f} = 0{,}00904 \cdot \frac{565}{23{,}94 \cdot 2 \cdot \pi \cdot 50} = 0{,}679\ mH \tag{[1] (3.3)}$$

Pro kompletní realizaci LC-filtru je nutné navrhnout ještě kapacitu kondenzátorů, kterou volíme tak, aby se s tlumivkou nedostali do rezonance. Použijeme k tomu následující vzorec:

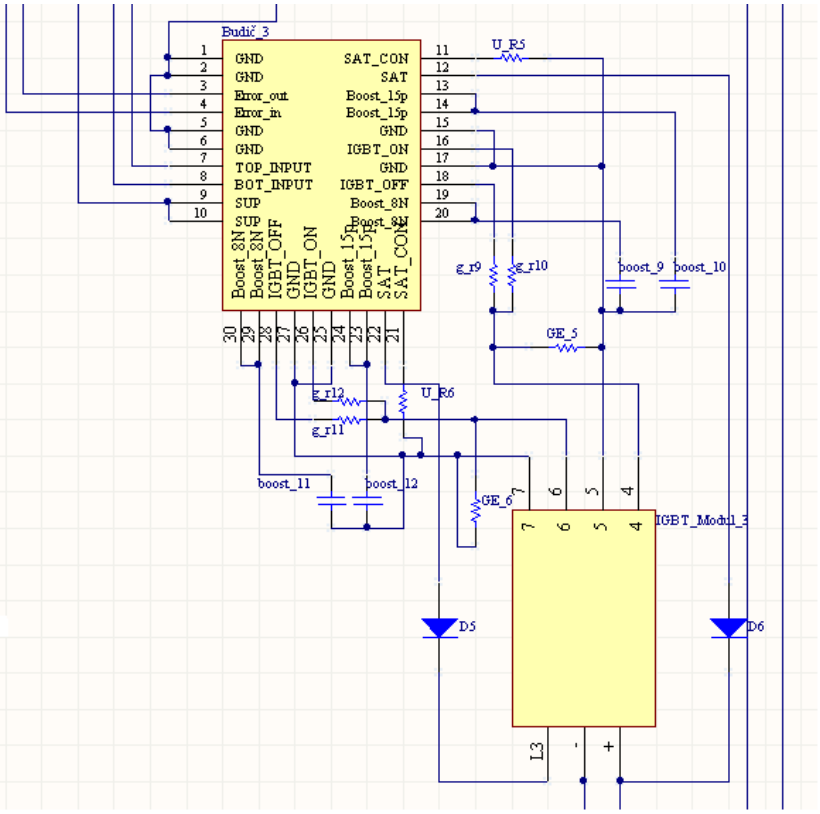
$$C = \frac{1}{4 \cdot \pi^2 \cdot f_{mez} \cdot L} = \frac{1}{4 \cdot \pi^2 \cdot 100} = 3{,}73\ mF \tag{[1] (3.4)}$$

3.5. ČIDLA PROUDU

Pro zpětnou vazbu v regulaci proudu je třeba snímat proud. Bude tak činěno pomocí dvou halových sond, každá pro jednu fázi. Proud ve třetí fázi lze dopočítat. Výstupem ze sond je napětí - jedná se tedy o nekompenzované sondy.

4. NÁVRH

Na obrázku 4.1 můžeme vidět detail návrhu schématu silové části. Jedná se o přivedení řídících signálů z řízení do jednoho z modulů přes budič. Je zde také řešeno napájení, signalizace chyb a saturační ochrana. Veškerá komunikace mezi silovou a řídící částí, která má svou vlastní desku plošných spojů, bude probíhat pomocí jednoho konektoru. Na této části se pracuje a do prezentace bude hotová v podobě desky plošných spojů.



Obrázek 4.1: Detail schématu – propojení budiče a výkonového modulu

5. ZÁVĚR

V této práci byly popsány základní mezníky dosavadního návrhu frekvenčního měniče. Na základě vhodných parametrů byl mimo jiné zvolen spínací IGBT modul Semikron 400GB125D a jeho budič Semikron Skyper R32. Na schématu je vidět jejich propojení, zároveň s dalšími funkcemi silového obvodu. Cílem práce je frekvenční měnič realizovat a ověřit teoretické předpoklady v praxi.

REFERENCE

[1] PATOČKA, Miroslav. *Výkonová elektronika BVEL: 1. část - usměrňovače, střídavé měniče napětí*. Brno, 2010.

[2] PATOČKA, Miroslav. *Vybrané statě z výkonové elektroniky Svazek II: Pulsní měniče bez transformátoru*. Brno, 2005.

[3] PATOČKA, Miroslav. *Matematický nástroj pro analýzu PWM signálů, jeho užití ve střídavých pohonech* [online]. Brno, 2000 [cit. 2015-12-15]. ISBN 80-214-1646-7. Dostupné z: http://www.vutium.vutbr.cz/tituly/pdf/ukazka/80-214-1646-7.pdf. Teze Habilitační práce. Vysoké učení technické v Brně.

Bakalářské projekty

Biomedicínské inženýrství

A TECHNIQUE FOR PARAMETER ESTIMATION OF EQUIVALENT-CIRCUIT MODELS

Filip Mivalt

Bachelor Degree Programme (2), FEEC BUT

E-mail: xmival00@stud.feec.vutbr.cz

Supervised by: Petr Sedlak

E-mail: sedlakp@feec.vutbr.cz

Abstract: The paper describes a simple algorithm for parameter estimation of equivalent-circuit models on the basis of comparison calculated and measured electrical impedance. The algorithm searches in multidimensional space in the vincinity of expected results. Despite the simple calculations, the algorithm easily outstrip number of high-order mathematical and algorithmic methods, such as non-linear least squares, neural networks etc. The algorithm is easily realizable and outputs are highly accurate in comparison to experimental data.

Keywords: Mathematical models, Models of electrical components, Estimation, Approximation

1. ÚVOD

Matematické modely slouží k popisu reálných jevů či soustav. Míra složitosti modelu je přímo úměrná míře komplexnosti popisu daného jevu. Modely se využívají k popisu a analýze elektronických součástek, biologických funkcí, mechanických systému a dalších. U modelů elektronických součástek a některých typů biochemických senzorů je základním úkolem po úspěšném navržení ideálního obvodu zejména odhad jednotlivých parametrů. Tenhle úkol se stává náročným, pokud odhadujete větší počet součástek stejné série, ale jiných parametrů. Například křemenné mikrovážky o různých tloušťkách aktivní vrstvy. Z těchto důvodů jsem navrhl realizačně jednoduchý algoritmus, který řeší výše zmíněný problém.

Pro realizačně jednoduchý princip cyklické optimalizace nejsou potřeba komplikované numerické metody. Tenhle algoritmus je tedy relativně snadné implementovat v různých programovacích jazycích.

2. ZÁKLADNÍ NÁHRADNÍ MODELY REÁLNÝCH ELEKTRONICKÝCH SOUČÁSTEK

Náhradní ekvivalentní model popisuje chování skutečné elektronické součástky, kde jednotlivé komponenty ideálního obvodu charakterizují vlastnosti modelovaného objektu. Např. odpor a indukčnost kontaktů, odpor dielektrika nebo u biosenzorů kapacita lipidové dvojvrstvy. Jednotlivé modely se používají pro simulaci frekvenčních charakteristik, přechodových dějů, popřípadě jiných charakteristických vlastností.

Popisovaný algoritmus jsem vyvinul pro snadnější odhad parametrů ekvivalentních náhradních modelů. Tyto modely popisují elektrické charakteristiky křemenných mikrovážek o různé geometrii a materiálu aktivní vrstvy. Stejný postup jsem úspěšně aplikoval pro náhradní modely kondenzátorů od různých výrobců.

3. APROXIMAČNÍ PREFÁZE

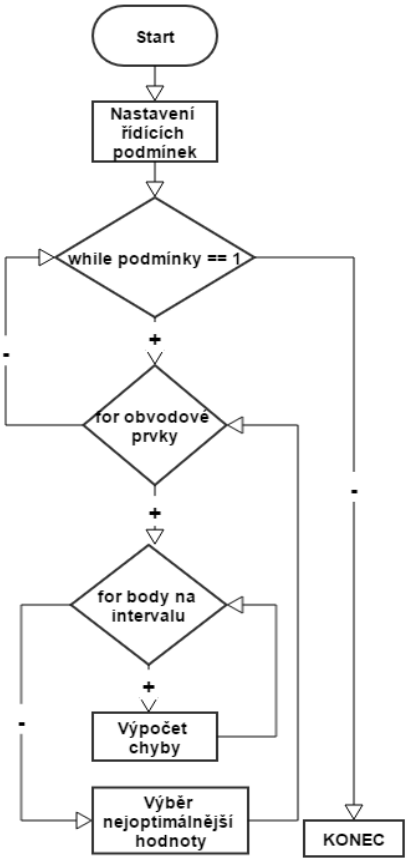
Před samotným použitím algoritmu je potřeba provést rozbor řešené úlohy. Musíme určit požadovanou přesnost modelu a zvážit reálnost této představy. Je potřeba zohlednit zašumění, popřípadě

jiné znehodnocení modelované křivky. Dále musíme rozhodnout, zdali chceme modelovat parazitní rezonance, které se v naměřených datech mohou vyskytovat. Je třeba věnovat zvláštní pozornost správnému nastavení kontrolních parametrů algoritmu. Zejména pokud si přejeme zmíněné parazitní složky vynechat, nebo naopak modelovat se zvýšenou přesností. Neméně důležitým krokem je metodika stanovení prvotního odhadu. Obecně platí, že čím je odhad přesnější, tím větší efektivity můžeme algoritmem dosáhnout. V případě, že nejsme schopni provést dostatečně přesný odhad, potřebujeme znát alespoň řád a přibližné poměry jednotlivých parametrů. Tahle varianta je výpočetně náročnější a také mnohdy dosahuje nižší přesnosti.

Pro dosažení lepších výsledků lze algoritmus modifikovat, avšak s novými modifikacemi přibývá kontrolních parametrů. Jejich správné nastavení pak může být problematické.

4. ALGORITMUS

Základní princip algoritmu spočívá v postupném procházení Euklidovského multidimenzionálního prostoru zvlášť pro každou dimenzi. V každé iteraci se snažíme přiblížit globálnímu minimu chybové funkce. Chybovou funkci jsem definoval jako průměr kvadrátů relativní chyby ve všech bodech měřených a modelových dat. To znamená, že hodnoty jednotlivých koeficientů posouváme ve směru gradientu záporného násobku této chybové funkce. Při samotné implementaci to znamená prohledávání předem definovaného intervalu v okolí aktuálních hodnot koeficientů se zvoleným vzorkováním. Vzorkování intervalu volím lineární. Jinou variantou může být vzorkování kvazilogaritmické, kdy se kolem středu prohledávaného intervalu vzorkuje s vyšší hustotou. Tato hustota klesá se vzrůstající vzdáleností od středu intervalu. Druhá zmíněná varianta je riziková. Zejména pokud je nutno zachovat poměry mezi jednotlivými obvodovými prvky.



Obrázek 1: Vývojový diagram algoritmu

5. VSTUPNÍ PARAMETRY, MODIFIKACE ALGORITMU

Řídícími parametry algoritmu jsou: hodnota maximálního počtu iterací, maximální počet iterací bez zlepšení, maximální přípustná relativní chyba a maximální přípustná průměrná relativní chyba mezi modelem a naměřenými hodnotami. Popřípadě další parametry plynoucí z případných modifikací algoritmu.

Pokud jsou měřená data znehodnocena šumem nebo se na nich vyskytují nechtěné rušivé elementy, je nutno před začátkem modelování tyto elementy odstranit. Avšak pokud se ale jedná o stacionární stochastický proces se zanedbatelnou amplitudou, je možno tenhle šum zanedbat. Nejčastěji se náhradní obvody sestrojují v logaritmickém měřítku a se zvyšujícími se frekvencemi roste citlivost systému, která se projevuje zvyšujícím se výkonem šumu. V předchozí úvaze si tedy musím vystačit s předpokladem, že tenhle proces je stacionární, alespoň na určitém intervalu. Délku tohoto intervalu lze definovat jako nejmenší možnou délku, kdy se střední hodnota amplitudy šumu na tomto intervalu rovná nule, nebo se alespoň k nule blíží.

$$RMSE = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{X_i - Y_i}{X_i}\right)^2 \qquad (1)$$

(1) Kvadrát relativní průměrné chyby; X – naměřená data, Y – model, N – počet vzorků

Pro každý bod na prohledávaném intervalu musíme spočítat průměr kvadrátu relativní chyby. V případě nutnosti můžeme kvadrát relativní chyby zaměnit za jinou metodu výpočtu chyby. Obecně výpočet volím podle způsobu, kterým váhuji různé části aproximované křivky. Váhování je důležitý aspekt, kterému je třeba věnovat velkou pozornost. Správné váhování je jedním z největších klíčů k úspěchu u komplikovanějších obvodů. Zvláště pokud modelujeme parazitní části křivky s nenulovou hodnotou druhé derivace. Například parazitní rezonance. Správné rozdělení vah může být rizikové a je tenká hranice mezi správným a přílišným váhováním.

6. VÝSLEDKY

Předloženým algoritmem jsem řešil odhad ekvivaletních náhradních modelů pro popis elektrických charakteristik křemenných mikrovážek. Stejný algoritmus jsem použil pro náhradní modely kondenzátorů od různých výrobců. Algoritmus je testovaný na náhradních modelech až o 28 obvodových prvcích. V naprosté většině jsem dosáhl maximální chyby menší než 10%. Ovšem průměrná chyba byla podstatně menší. Pohybovala se od 2 - 6%. Tato hodnota závisím na míře znehodnocení měřených dat.

Při porovnávání předloženého algoritmu s algoritmy komplexními jako například neuronové sítě bych jako hlavní výhodu zdůraznil relativní realizační nenáročnost. Další důvod proč zvolit právě tenhle algoritmus je mnohem vyšší přesnost odhadovaného náhradního modelu oproti použití běžných funkcí nabízených v Matlabu. Dále vyzdvihuji zejména možnost jednoduché modifikace algoritmu a tím zvýšení přesnosti výsledků, což u předdefinovaných funkcí navzdory jejich mnohým možnostem a nastavením není možné.

REFERENCE

[1] SCHREIBER, Rob. MATLAB. *Scholarpedia* [online]. 2007, 2(7), 2929 [cit. 2016-03-09]. DOI: 10.4249/scholarpedia.2929. Dostupné z http://www.scholarpedia.org/article/MATLAB

[2] VITÁSEK, Emil a E VITÁSEK. *Numerické metody*. 1. vyd. Praha: SNTL, 1987.

ACTIVE FOREARM PROSTHESIS

Martin Vařečka

Bachelor Degree Programme (3), FEEC BUT

E-mail: xvarec04@stud.feec.vutbr.cz

Supervised by: Jaroslav Balogh

E-mail: balogh@feec.vutbr.cz

Abstract: The aim of this work is to design active forearm prosthesis. This prosthesis scans myoelectric signals from upper limb, especially from forearm. Signals are processed by analog filters. Muscle contraction is evaluated by the Arduino, which also controls the servomotors. Motor unit is consisted of two motors. It is therefore prosthesis with two degrees of freedom.

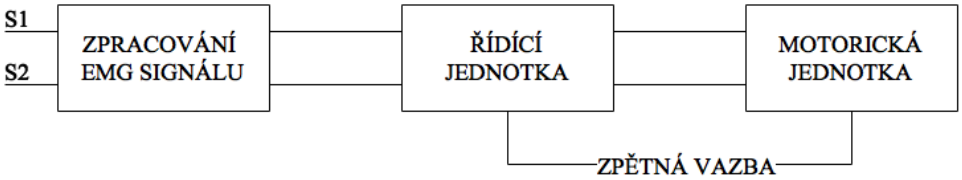
Keywords: forearm prosthesis, electromyography, signal processing, instrumentation amplifier, arduino, servo-motor, motor unit

1. ÚVOD

Protetika je obor Ortopedie zabývající se léčbou pacientů pomocí protetických pomůcek. Ty nahrazují část těla, která přišla o svou fyziologickou funkci. Tuto část těla je obvykle nutné amputovat ze života ohrožujících důvodů, jako jsou například traumata, infekce či tumory. Protézy lze rozdělit na aktivní a pasivní. V tomto příspěvku je popsán návrh aktivní protézy, která k vykonání pohybu využívá vnější sílu – bateriově napájené servomotory. Ovládání protézy zajišťuje vývojová platforma Arduino, jenž vzorkuje filtrovaný, zesílený a usměrněný elektromyografický (EMG) signál ze dvou svalových skupin. V řešení je využito analogové předzpracování signálů.

2. HARDWAROVÝ A SOFTWAROVÝ NÁVRH AKTIVNÍ PROTÉZY

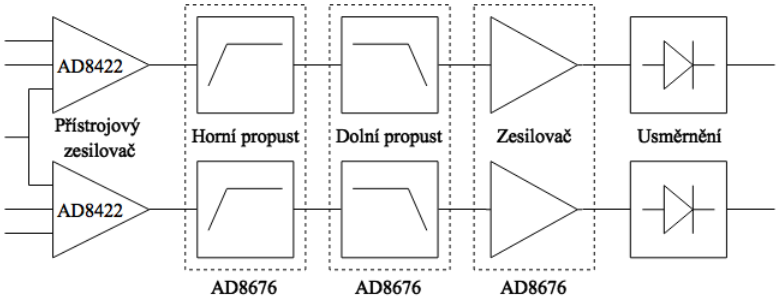
Schéma na **Obrázek 1** zobrazuje koncepci základního návrhu zařízení pro zpracování a vyhodnocení EMG signálů snímaných ze dvou antagonistických svalových skupin. Na každou svalovou skupinu připadají dvě elektrody měřící a jedna referenční. Referenční elektrody jsou nahrazeny jednou společnou elektrodou. Tyto elektrody jsou připojeny na blok pro zpracování EMG signálu. Vznikají dva signály (kanály), jeden pro každou svalovou skupinu. Pro snímání jsou využity povrchové nepolarizovatelné argentchloridové (Ag-AgCl) elektrody. Funkce jednotlivých bloků je popsána dále.



Obrázek 1: Základní blokové schéma zařízení

2.1. BLOK PRO ZPRACOVÁNÍ ELEKTROMYOGRAFICKÉHO SIGNÁLU

Tento blok je sestaven z pěti částí pro každý kanál. Blokové schéma obou kanálů je zobrazeno na **Obrázek 2.** Elektrody umístěné na ruce (či pahýlu) jsou přivedeny na vstupní zesilovač. Pro tento účel byl zvolen přístrojový zesilovač AD8422 od firmy Analog Devices. Jedná se o precizní přístrojový zesilovač, jenž vznikl jako třetí generace průmyslového standardu AD620. Nejdůležitějším parametrem přístrojových zesilovačů je diskriminační činitel CMRR.



Obrázek 2: Blokové schéma bloku pro zpracování EMG signálu

Pro zesílení $G = 100$ je CMRR přístrojového zesilovače AD8422 rovno 126 dB oproti přístrojovému zesilovači AD620, kde je hodnota rovna 110 dB. Tento rozdíl poukazuje na vývoj integrovaných obvodů, jenž zajistí lepší poměr signálu a šumu o 16 dB při stejném zesílení. Přístrojový zesilovač AD8422 je lepší ve všech ohledech. Za zmínku stojí klidový proud 0,33 mA oproti 1,3 mA (AD620). Zesílení přístrojového zesilovače AD8422 je závislé na hodnotě odporu připojeného na R_G piny. Zesílení vypočítáme dle rovnice (1). [1]

$$G = 1 + \frac{19800}{R_G} \tag{1}$$

Pro získání užitečného signálu je nutné provést předzpracování, tedy odstranění nežádoucích frekvenčních pásem. Toho je docíleno zařazením filtru typu horní a dolní propust do série. Frekvenční rozsah EMG signálu při povrchovém snímání se pohybuje v rozmezí od 20 do 500 Hz. Z těchto požadavků vychází i návrh analogových aktivních filtrů. První aktivní filtr typu horní propust má mezní frekvenci na 20 Hz a zesílení 2, druhý aktivní filtr typu dolní propust má mezní frekvenci na 500 Hz a zesílení 2. Oba filtry jsou Butterworthova typu a to zejména pro jeho plochou amplitudovou charakteristiku v propustném pásmu. Pro tento systém jsou dostačující filtry druhého řádu. [2]

Zesílení signálu je provedeno pomocí operačního zesilovače v neinvertujícím zapojení. Ve zpětné vazbě se nachází trimr, který slouží pro nastavení zesílení. Každý kanál je možné zesílit samostatně. Celkové zesílení signálu je dáno součinem zesílení všech předchozích prvků a tohoto neinvertujícího zesilovače. Pokud zesílení přístrojového zesilovače nastavíme na hodnotu 100, pak je zesílení předchozích prvků rovno 400. Zesílení neinvertujícího zesilovače je možné volit v rozmezí 1 až 26, to nám tedy dává celkové zesílení v rozmezí 400 až 10 400.

Aktivní prvek pro konstrukci analogových filtrů a neinvertujícího zesilovače je precizní operační zesilovač AD8676. Celkem je pro jeden kanál použito 4 operačních zesilovačů (přístrojový vstupní, HP, DP a výstupní) a jednotka má 2 kanály. V konstrukci je využito faktu, že veškeré OZ jsou vždy 2x v jednom pouzdře.

Usměrnění signálu je velice důležité pro fungování celého systému. Vychází z požadavků analogově-digitálního (AD) převodníku vývojové desky Arduino. Tento AD převodník není schopen zpracovávat negativní výchylky signálu. Z tohoto důvodu je nutné signál před vzorkováním dvoucestně usměrnit, což si vzhledem k povaze EMG signálu můžeme dovolit. Pro dvoucestné usměrnění využijeme Grätzův můstek. Při výběru diod je nutné dbát na co nejnižší prahové napětí, které otevře diodu v propustném směru. Těmto požadavkům nejlépe vyhovuje Shottkyho dioda. Prahové napětí této diody je cca 0,4 V, což je dostačující, jelikož signál zesilujeme na jednotky voltů. Takto usměrněný signál je přiváděn na analogový vstup Arduina.

2.2. ŘÍDÍCÍ JEDNOTKA

Řídící jednotka, má za úkol vyhodnotit již předzpracovaný signál. K tomuto účelu nám poslouží vývojová deska Arduino. Pro naše účely je nejlepší použít verzi MINI či NANO a to zejména pro jejich kompaktní velikost. Požadavky vývojové desky na napájení z externího zdroje je 7-12 V. Na vývojové desce se nachází stabilizátor napětí na hodnotu 3,3 V a 5 V. Stabilizované napětí

5 V používáme jako zdroj napájení pro blok zpracovávající EMG signál. Jelikož přístrojové a operační zesilovače požadují symetrické napájení, je nutné napětí +5 V invertovat na -5 V. Pro tento účel použijeme integrovaný obvod ADM660, jehož vlastnost je invertovat vstupní napětí v rozsahu +1,5 V až +7 V na -1,5 V až – 7 V s 94% účinností.

Jak již bylo zmíněno výše, usměrněné signály jsou přiváděny na analogové vstupy, kde jsou vzorkovány 10-bitovým AD převodníkem. Rozsah AD převodníku je od 0 V do A_{ref}. Výchozí rozsah (bez přivedení napětí na pin A_{ref}) je od 0 V do 5 V. Při výchozím rozsahu je rozlišení AD převodníku 4,9 mV. Vzorkované signály jsou dále filtrovány mediánovým filtrem, který tvoří křivky signálů, ze kterých je patrná aktivita svalstva. Mediánový filtr je realizován softwarově řídící jednotkou v jazyce C++.

Aktuální hodnoty iterace křivek jsou porovnávány s prahovými hodnotami. Výsledkem logické operace je 0 – false či 1 – true. Jsou tedy možné 4 kombinace: **0-0, 1-0, 0-1, 1-1** (kanál 1 – kanál 2). Tyto kombinace odpovídají událostem: **rozevření čelistí, rotace vpravo, rotace vlevo, sevření čelistí**. Analogie pro **druhou** kombinaci (1-0): „Hodnota křivky kanálu 1 je **vyšší než práh (1)**, hodnota křivky kanálu 2 je **nižší než práh (0)**, proveď **rotaci** terminální pomůcky **vpravo**."

2.3. MOTORICKÁ JEDNOTKA A KONSTRUKCE PROTÉZY

Pro tuto práci byla zvolena čelisťová terminální pomůcka vhodná k pracovnímu použití. U tohoto druhu terminální pomůcky lze za výhody považovat její odolnost, sílu úchopu a možnost precizního „pinzetového úchopu" mezi distální úchopové plošky. Nevýhodou je její nepřirozený vzhled. Motorická jednotka je tvořena dvěma servomotory. Ovládání servomotorů zajišťuje řídící jednotka pomocí PWM. První servomotor provádí sevření a rozevření čelisti, servomotor druhý rotaci vpravo a vlevo. Použité servomotory musí být dostatečně silné a rychlé. V projektu jsou použity voděodolné servomotory se sílou 12 kg/cm. Síla servomotorů je dostatečná při použité páce o délce 3 cm.

Veškeré výše zmíněné komponenty, včetně bateriového zdroje jsou umístěné v exoskeletu konstruovaného ze stavebnice Merkur a to zejména pro možnost snadné modifikace pří vývoji. Čelisťová terminální pomůcka je zhotovena zvlášť z hliníku. K té jsou připevněny servomotory. Rotace probíhá mezi čelistí a servomotorem. Proximální část servomotoru je pevně spojena s exoskeletem. Práce počítá s návrhem 3D modelu, resp. definováním požadavků na finální konstrukci exoskeletu protézy ze standardních materiálů (lehké slitiny/kompozity).

3. ZÁVĚR

Cílem práce je návrh aktivní myoelektrické protézy horní končetiny. Signály jsou snímány ze dvou svalových skupin se společnou referenční elektrodou. Zpracování signálů je řešeno hardwarově. Elektrody jsou přivedeny na vstupy přístrojových zesilovačů, dále je pomocí analogových aktivních filtrů získáno užitečné pásmo signálů v rozsahu 20 až 500 Hz. Signály jsou dále zesíleny a dvoucestně usměrněny. Ze vzorkovaných signálů jsou vytvořeny mediánovou filtrací křivky, jejichž hodnoty každé iterace jsou porovnávány s nastavenými prahovými hodnotami. Při překročení prahu je vyvolána odpovídající událost (sevření - rozevření čelisti, rotace čelisti vpravo - vlevo). Vylepšení je možné dosáhnout zapojením senzoru síly (ang. fsr – force-sensing resistor), který mění odpor v závislosti na síle, jenž jej stlačuje, ve zpětné vazbě. Tento senzor zabrání destrukci křehkých věcí. Motorická jednotka využívá dvou servomotorů – ovládání čelisti a rotace. Jedná se tedy o protézu s dvěma stupni volnosti. Použité servomotory mají zvýšenou odolnost proti vodě.

REFERENCE

[1] ANALOG DEVICES. AD8422. Data sheet. [online]. 2015 [cit. 2016-01-02]. Dostupné z: <http://www.analog.com/media/en/technical-documentation/data-sheets/AD8422.pdf>

[2] KOLÁŘ, R.: Lékařská diagnostická technika, Elektronické skriptum, Brno: Vysoké učení technické v Brně, 2014.

ACTIVE PROSTHETIC HAND

Michal Svoboda
Bachelor Degree Programme (3), FEEC BUT
E-mail: xvobo88@stud.feec.vutbr.cz

Supervised by: Jaroslav Balogh
E-mail: balogh@feec.vutbr.cz

Abstract: The aim of this thesis is to propose and produce prototype of active hand prosthesis. Prototype is based on myoelectric prosthesis principle, therefore it is meant to be controlled by EMG signals. In this specific case signals will be scanned from healthy arm. Arduino UNO development kit was chosen as the main controller. For measuring EMG muscle activity Shield EKG/EMG Olimax was selected.

Keywords: EMG, electromyogram, 3D printers, prosthetic hand, Arduino UNO

1 ÚVOD

Vzrůstající zájem o adrenalinové sporty, práce s těžkou technikou či s výbušninami, dopravní nehody, válečná zranění a mnoho dalších událostí, mohou vést k závažnému poranění s trvalými následky. I přes zlepšující se bezpečnostní prvky se stále setkáváme s lidmi, kteří museli podstoupit amputaci končetiny, nebo se narodili s vrozenou vadou těla. Pro tyto osoby se pak běžné denní činnosti stávají úkony nad jejich síly. Cílem této práce je vytvořit prototyp aktivní protézy ruky, respektive náhrady předloktí, zápěstí a prstů. Protézu bude ovládat nositel gesty své zdravé ruky. Myšlenka je tedy taková, že protéza bude zrcadlit ruku zdravou.

2 AKTIVNÍ PROTÉZA RUKY

Protéza je pomůcka nahrazující ztracenou nebo nevyvinutou část těla a její funkce. Věda zabývající se vývojem a výrobou se nazývá Protetika. Vývoj nových technologií a materiálů používaných při výrobě protéz se v posledním desetiletí urychlil v důsledku vylepšení vyrobních metod a používaných materiálů. Vede to k spolehlivějšímu a komfortnějšímu používání protetických pomůcek uživateli. Aktivní protetické náhrady lze rozdělit na protézy ovládané silou postiženého (tahové protézy), dále protézy ovládané silou zevní (myoelektrické protézy) a hybridní protézy. Neaktivní protézou je například kosmetická protéza, která nabízí minimální funkční možnosti a je využívána hlavně pro její, od originálu téměř nerozeznatelný, kosmetický vzhled.[4]

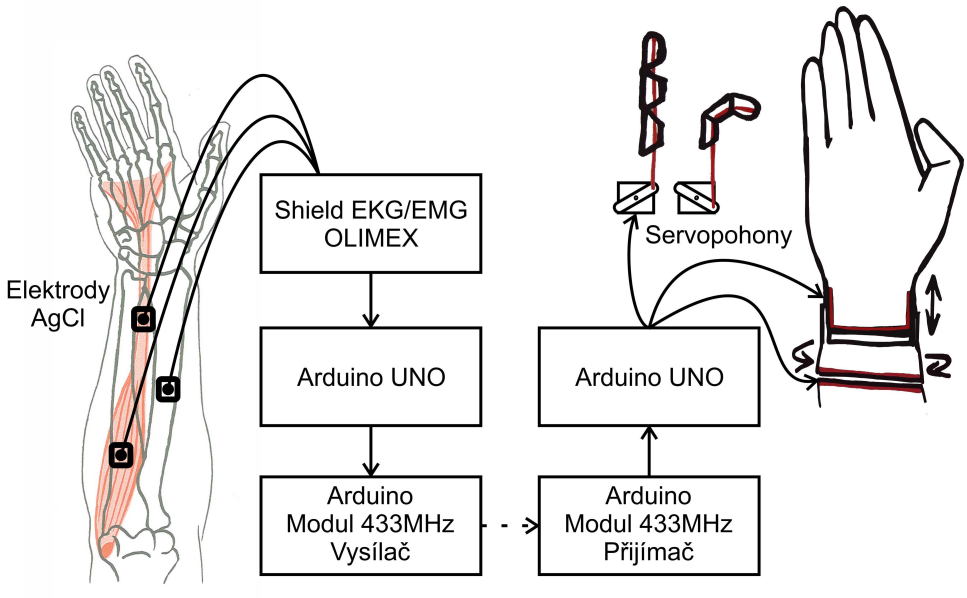
Tahové protézy kladou důraz především na funkčnost ruky. Obvykle bývá u těchto protéz náhrada ve tvaru kleští, kde pohyb těchto kleští je řízen jinou části těla (paží, ramenem, zády nebo hrudníkem) za pomoci upínacího postroje. Používání tahových protéz je fyzicky vysilující a naučit se s protézou pracovat je zpočátku velice náročné. Také lze provádět úkony pouze mezi úrovní úst a úrovní pasu. Celý systém zatěžuje nepostiženou stranu pacienta, jelikož pohyby pro ovládání je potřeba vykonávat v nefyziologických vzorcích. To vede k hrubšímu a méně plynulému uchopovacímu pohybu. Výhodami tahových protéz jsou menší váha oproti myolektronickým protézám, nižší pořizovací náklady, jednodušší údržba a servisní úkony. Tahové protézy jsou všeobecně odolnější na mráz, vlhkost, nečistoty a nárazy.[5]

Mioelektronické protézy obsahují jeden či více motorů, které umožňují oproti jiným protézám silnější a přesnější úchop. Kontrakce svalů pahýlu ruky vyvolává elektrické napětí, jež je velice malé a rušené.

Elektrody naměřené napětí předají řídící jednotce protézy, ta signál vyhodnotí a předá na motory, které pohybují rukou, zápěstím, popřípadě i loktem. V případě že je dobrý stav svalstva v pahýlu po amputačním zákroku a pacient se tyto svaly naučí ovládat, pak mu protéza může pomoci zvládat většinu běžných pohybů a úkonů. Myoelektrické protézy mají výrazně lepší estetický vzhled než tahové protézy, ale ani tak nedosahují dokonalosti protéz kosmetických. Nevýhodou je jejich značná hmotnost a vysoká pořizovací cena. S tím jde ruku v ruce větší poruchovost a nákladný servis. Většina dnešních myoelektrických protéz má nižší odolnost proti mrazu, vlhkosti a nárazům. Na druhou stranu dovolují pacientům vykonávat pohyb i nad úrovní úst, pod úrovní pasu a po stranách mimo středovou osu uživatele. Používání těchto protéz také braní atrofii svalstva amputačního pahýlu. Výsledkem této práce bude prototyp vycházející z principů myoelektrických protéz.[5]

2.1 Návrh prototypu

Svaly amputačního svalu bez řádného cvičení časem atrofují. Z tohoto se bude prototyp skládat ze dvou částí. První část bude sloužit ke snímání EMG aktivity svalů zdravé ruky, která se následně zpracuje do signálů, pomocí kterých se bude ovládat druhá část prototypu, samotná protéza. Návrh struktury systému je zobrazen na obrázku č. 1. Větší část dílů konstrukce protézy bude pro jednodušší a levnější výrobu vytištěna pomocí 3D tiskárny.



Obrázek 1: Blokové schéma systému pro snímání EMG signálu a aktivní protézy ruky.

2.2 Řídící jednotka Arduino UNO

Vyhodnocovací a řídící jednotku v tomto projektu zastává open-source počítačová platforma Arduino UNO. Bude využita nejprve ve snímací soupravě k vyhodnocení zpracovaných EMG signálu. Dále bude využita v aktivní protéze pro řízení servomotorů podle instrukcí, které přijme ze snímací soupravy. Vývojový kit Arduino UNO disponuje 14 digitálními vstupně výstupními piny, z toho lze 6 využít jako PWM výstupy, 6 analogovými vstupy s 10 bitovým ADC, 2 sériovými linkami a SPI. Mikrokontrolér v tomto vývojovém kitu je osmibitový AVR procesor Atmel ATmega328P s 32 kB flash pamětí, 1024B EEPROM a maximální pracovní frekvencí 20MHz.[2]

K programování Arduina se využívá vývojové prostředí ArduinoIDE a lze programovat v jazyce C nebo C++ s pomocí komplexních knihoven Wiring. Součástí desky kitu je i USB, sloužící ke komunikaci s PC.[1]

Moduly slouží pro jednosměrnou bezdrátovou komunikaci dvou vývojových kitů Arduino. Vysílací a přijímací moduly pracují na frekvenci 433,92 MHz. Rychlost vysílače je až 4KB/s.[3]

2.3 MĚŘENÍ EMG AKTIVITY

Pro snímaní elektrické aktivity svalů slouží elektromyograf (EMG). Ze získaných údajů jsme schopni zjistit zda sval je stažen, nebo se zrovna stahuje/roztahuje. Elektromyograf k neinvaznímu snímání EMG signálu využívá povrchových elektrod, kdy v tomto projektu jsou typu AgCl. Tento EMG signál se pohybuje ve frekvenčním pásmu 50-500Hz a hlavní složka signálů se nachází v pásmu 50-150 Hz. Pro naši potřebu se použije Shield pro snímání EKG/EMG pro vývojový kit Arduino od firmy Olimex.

2.4 POHYBOVÉ ÚSTROJÍ

V případě zápěstí musíme uvažovat o třech stupních volnosti. Protéza bude využívat pouze dva stupně volnosti, a to imitaci rotace ruky kolem osy předloktí a pohybu zápěstí nahoru a dolů. Oba tyto pohyby budou zajištěny pomocí servomotorů. Pohyb jednotlivých prstů bude prováděn za pomoci tahového ustrojí, které bude taženo silou serva a řízený mikroprocesorem. Návrat do roztažené polohy bude řešen systémem pružin nebo mechanismem kladek. U servomotoru bude potřeba kontrolovat odezvu na uchopený předmět, aby nedošlo k jeho rozdrcení či poškození protézy.

3 ZÁVĚR

Zařízení je v době publikace ve fázi vývoje prototypu na nepájivém poli. V současné době probíhá zkušební měření a zpracování EMG signálů. Předmětem další práce je zkompletování, vytvoření elektronické části a sady instrukcí pro řízení protézy vycházejících z naměřené aktivity.

REFERENCE

[1] Arduino: *Arduino board UNO*.[online], [2016], Citováno 8. března 2015.
`https://www.arduino.cc/en/Main/ArduinoBoardUno`

[2] Atmel Corporation: *ATmega328P*. [online], [2016], Citováno 8. března 2015.
`http://www.atmel.com/devices/atmega328p.aspx`

[3] Icstation.com: *433Mhz RF Transmitter And Receiver Kit*. [online], [2005-2016], Citováno 8. března 2015.
`http://www.seeedstudio.com/wiki/433Mhz_RF_link_kit`

[4] Brozmanová, B. a kolektiv: Ortopedická protetika, Osveta, 1990, ISBN 80-217-0133-1

[5] Paigerová, M.: Srovnání jednotlivých typů protéz horních končetin. *Ortopedická protetika*. [online], [1999], Citováno 8. března 2015.
`http://www.ortotikaprotetika.cz/oldweb/Wc1dfa89a7ea17.htm`

MICROCONTROLLER AUDIOSTIMULATION FOR BRAIN FMRI EXAMINATIONS

Marika Sobotková
Bachelor (3), FEEC BUT
E-mail: xsobot22@stud.feec.vutbr.cz

Supervised by: Jiří Sekora
E-mail: sekora@feec.vutbr.cz

Abstract: During fMRI acquisition of task related data the variability of time lag between the audio stimulus demand and its actual playback causes a huge problem for precise data analysis. The lag is probably arised because of using standard PC for whole stimulation management, where is not possible to achieve absolute control of stimuli timing. The developed device for audio stimulation based on Arduino Due microcontroller solves that problem. The control of the audio stimuli playback is realized by external trigger. All together allows to avoid variable and high lag of audio stimulation.

Keywords: audio stimulation, functional magnetic resonance, microcontroller, Arduino

1 ÚVOD

V neurovědním výzkumu se při zkoumání lidského mozku měří buď tzv. klidová data, nebo úkolová data, na kterých se sleduje reakce mozku na určitý podnět. Podněty mohou být např. vizuální nebo akustické.[1]

Při použití konvenčních nástrojů pro tvorbu stimulačního protokolu je člověk omezen tím, že jsou tato řešení závislá na použití počítače s určitým operačním systémem (dále jen OS) a stimulační program pak prostřednictvím OS přistupuje k jednotlivým periferiím, zde např. ke zvukové kartě. Vzniká tak situace, kdy výzkumník nemá přímou kontrolu nad časováním daného stimulu, protože čas od požadavku na spuštění daného audio podnětu a jeho skutečná reprodukce je velmi variabilní a v podstatě nepredikovatelný. Tím vzniká nedefinovatelné zpoždění v řádu jednotek až desítek milisekund. To např. znemožňuje výzkum evokovaných potenciálů [2] z mozku při audiostimulaci. Je třeba proto co nejvíce „obejít“ software počítače a generovat podněty maximálně hardwarovou cestou. Vzniká tím požadavek na tvorbu zařízení pro přehrávání stimulačních audiozáznamů v reálném čase a v požadované Hi-Fi kvalitě, které se maximálně vyčlení z vlivu OS a časově variabilních procesů, např. dekomprese zvukového záznamu.

2 ZAŘÍZENÍ PRO AUDIOSTIMULACI

Počítač, na kterém probíhá záznam měřených hodnot, bude nadále využíván pouze jako jakýsi nástroj pro značkování, kdy má být podnět generován, a řízení. Požadavek na spuštění daného podnětu vyšle stimulační software přes USB port do zařízení, které se na základě přijatého kódu postará o reprodukci vybraného audio stimulu. USB zaručuje odeslání i přijetí dat ve stejném pořadí, v jakém byla odeslána, a s konstantním zpožděním. Tímto způsobem je dodrženo optimální časování jednotlivých stimulů a správné vkládání stimulačních značek do záznamu, což umožňuje mnohem přesnější analýzu dat.

Vlastní zvukové záznamy jsou uloženy na paměťové kartě v tomto samostatném zařízení, takže nedochází k přenosu jiných dat mezi počítačem a zařízením, než jednoduchého krátkého protokolu, který

definuje, jaký záznam se má přehrát, případně zda má být provedena jiná funkce. Formát uložených skladeb je v „raw“ formátu, aby bylo odstraněno možné neměřitelné zpoždění vzniklé dekompresí zvukového souboru.

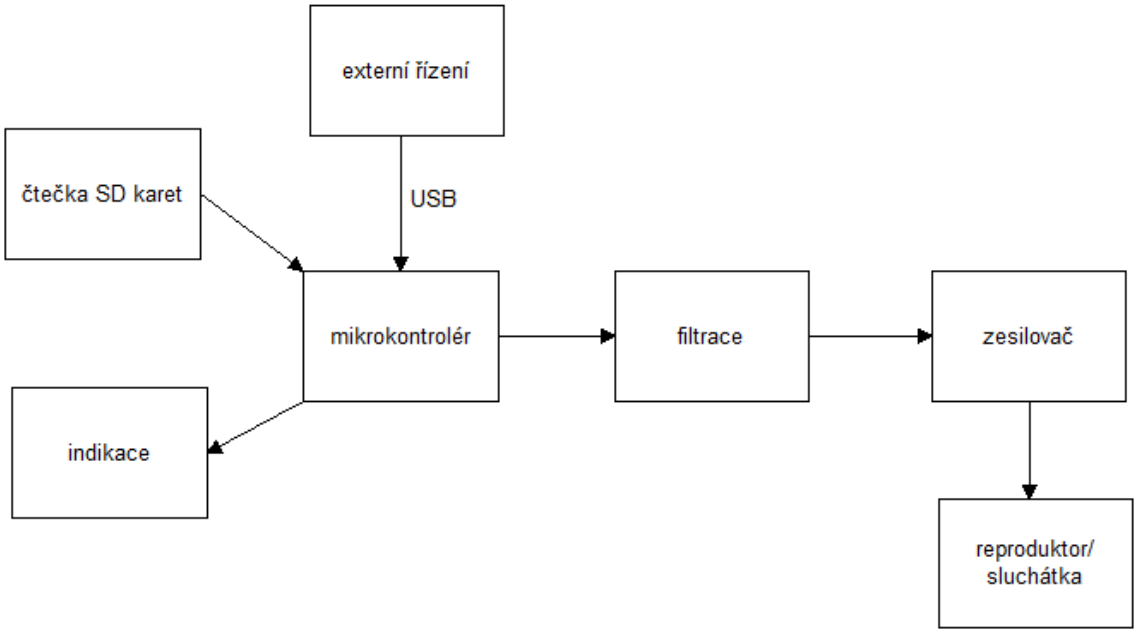
2.1 Zařízení

Samotné zařízení je postaveno na mikrokontroléru open-source platformy Arduino, konkrétně modelové řadě Due. Mikroprocesor použitý v řadě Due jako jediný z rodiny Arduino disponuje dostatečným výkonem pro přehrávání audio záznamu v požadované kvalitě, tedy $f_{vz} = 44100$Hz, 16 bitů a stereo. Zpracovaný záznam je převeden jako audio signál pomocí výstupů PWM, kdy je dolní propustí vyhlazen na analogový zvukový záznam.

Zvukový záznam na výstupu je dále zpracován audio-zesilovačem, zde konkrétně LM386, který slouží k impedančnímu přizpůsobení připojené zátěži — sluchátka nebo reproduktor, dále k nastavení požadovaného zesílení. Z důvodu požadovaného stereo výstupu budou použity dva zesilovače.

Zvukové stimulační nahrávky, které jsou uloženy na SD kartě, jsou načítány pomocí připojené čtečky SD karet. Komunikace mezi modulem čtečky a mikrokontrolérem probíhá po integrované sběrnici SPI [3], tedy opět na ryze hardwarové úrovni.

Navržené zařízení, včetně vazby na řídicí počítač (resp. osobní počítač pro záznam dat) je vyobrazeno v následujícím blokovém schématu.



Obrázek 1: Blokové schéma navrženého zařízení pro audiostimulaci

2.2 Řízení

Předpokladem pro návrh řídící části je, že při práci se zvukovým záznamem jsou využívány 4 základní funkce — přehrávání PLAY, pozastavení přehrávání PAUSE, pokračování v přehrávání CONTINUE a ukončení přehrávání STOP. Tyto funkce byly definovány jako požadavky vyšetřujícího pracoviště. Řízení je uskutečňováno pomocí 16bitového slova, které je rozděleno na 3 části. První částí je první bit, tzv. inicializační bit, který má vždy hodnotu „logická jedna“. Druhou částí je druhý a třetí

bit, které určují danou funkci. Pro přehrání skladby je použita dvojice bitů „11“, pro pozastavení přehrávání „10“, pro pokračování přehrávání „01“ a pro zastavení přehrávání „00“. Třetí a poslední částí je 13 nižších bitů. Touto sérií bitů je již přímo určeno číslo skladby, která má být přehrána. Představuje tak binární hodnotu, která koresponduje s číslem skladby na paměťovém mediu (samozřejmě pro přiblížení uživateli jsou skladby uloženy v dekadické soustavě).

2.3 SESTAVENÍ PROTOKOLU

Princip sestavení, resp. čtení protokolu je postaven na následujícím algoritmu. Namísto bitových operací (vyčleňování a posunování části bitového slova) byla volena jednoduchá matematická operace postavena na porovnávání a odečítání čísel. Pro každou z výše uvedených funkcí existuje určité základní 16bitové slovo. Pro zjednodušení je nyní převedeno do hexadecimální soustavy. Uživatel pomocí svého programu vygeneruje požadavek, který sestává z kódu skladby a funkce. Toto zadané číslo se porovná se základními hodnotami, definovanými pro PLAY, PAUSE, CONTINUE a STOP a skladbu číslo 0. Porovnáním hodnoty s hodnotou funkce dojde k přiřazení správné funkce a číslo skladby získáme odečtením základní hodnoty pro danou funkci od zadaného čísla. Jelikož je vyžadována pouze jednosměrná komunikace mezi počítačem a zařízením, bude přehrávání skladby, pozastavení a zastavení přehrávání doprovázeno indikací pomocí LCD, aby měl uživatel přehled o správné funkci zařízení.

Součástí řídicího softwaru je také inicializace SD karty, kontrola typu formátu skladeb, kontrola existence vybrané skladby. V případě chyby v některé výše jmenované části dojde k zobrazení konkrétní chybové hlášky na LCD.

3 ZÁVĚR

Podle požadavků pro zlepšení reprodukovatelnosti a správnosti dat při vyšetření funkční magnetickou rezonancí doprovázené stimulací zvukovými podněty bylo navrženo zařízení pro audiostimulaci. Zařízení je založeno na mikrokontroléru Arduino Due, což umožňuje eliminovat vliv PC v rámci celé přenosové cesty, konkrétně je tedy eliminováno nedefinovatelné zpoždění. Počítač je potom využíván pouze jako akviziční zařízení pro data generovaná z fMRI a ve vztahu k audiostimulaci pouze řídí okamžik spuštění audio stimulu. Kompatibilita s komunikačním rozhraním PC je zajištěna pomocí USB. Navržené zařízení umožňuje uživateli přehrávání audio stimulů s malým, a navíc konstantním zpožděním od požadavku na jejich spuštění, což umožňuje efektivně zaznamenávat značky v záznamu fMRI právě v místech, kde uživatel bude potřebovat audio stimul spustit.

REFERENCE

[1] Chlebus, Pavel, Mikl, Michal, Brázdil, Milan, Krupa, Petr: Funkční magnetická rezonance: Úvod do problematiky. [pdf]. Neurologie pro praxi. 2008, 2005(3/2005): 6. Dostupné z URL: <http://fmri.mchmi.com/articles/chlebus_prehled.pdf>.

[2] Bareš, Martin: Kognitivní evokované potenciály. [pdf]., Brno 2011 Dostupné z URL: <http://www.prolekare.cz/pdf?id=36052>.

[3] SPI library. Arduino [online]. Dostupné z URL: <https://www.arduino.cc/en/Reference/SPI>.

MOTION DETECTION OF SMALL OBJECTS

Svätopluk Blažej

Bachelor Degree Programme (3), FEEC BUT

E-mail: xblaze28@stud.feec.vutbr.cz

Supervised by: Petr Marcoň

E-mail: marcon@feec.vutbr.cz

Abstract: The paper describes a problem of detection small moving objects. A size of these objects is only one or few pixels. Relevant mathematical apparatus which are usually used for processing of dynamic pictures are included. There was proposed experimental module for standalone locator for real time operation. This module is consisted from two main parts. First part is used for imaging of scene, especially small moving objects. We use four camera systems. The Second part comprises a development board based on Raspberry PI 2. We use GNU/Linux system on the board. There were also tested methods useful for the essence of this project. On behalf of the tests was developed an algorithm to detect small particle in video sequence. C and python language routines were build including essence algorithms. This algorithm consists of: images are captured one by one from all 4 cameras 640 £ 480; RGB to Gray transformation; For each image are all shapes detected; unnecessary shapes are removed; Images are filtered with additional filters; Noise free images are then compared for movement detection. Images with movement are stored on SD card. Additional information: Devices captures and stores all flying objects. Data are accessible through LAN port. We use OpenCv for image analysis. Device is used for sky movement detection of small objects, birds. After detection, it can create acoustic signal to scare them away. This minimizes bird impact in Agriculture and it prevents bird-plane collisions on airports. This paper describes in more detail not only the entire detection algorithm, but also the results of the test algorithm.

Keywords: Raspberry Pi, motion detection, image processing, flying object detection, OpenCv

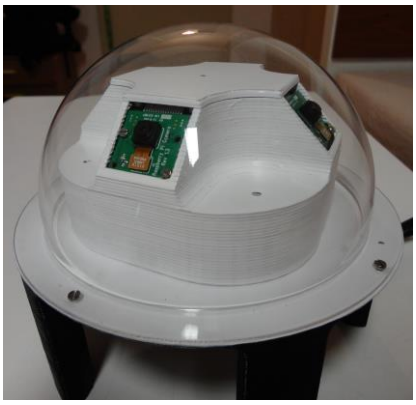
1. ÚVOD

Problematika detekcie pohybujúcich cieľov je v modernej dobe veľmi atraktívne odvetvie. Jej využitie je všestranné, od zapínania svetla po detekciu a rozoznávanie áut a ľudí, určenú na bezpečnostné účely. Doposiaľ sa detekovali objekty od určitej veľkosti, v mojej práci poukazujem na možnosť detekcie malých objektov, ktorých veľkosť je obmedzená len rozlíšením kamery - optickou sústavou. [1]

2. NAVRHNUTÉ RIEŠENIE

Ako projekt bolo navrhnuté zariadenie určené na detekciu pohybujúcich sa cieľov, zariadenie je určené pre detekciu vtáctva na oblohe, ktoré má ale využitie na detekciu akéhokoľvek pohybujúceho sa objektu. Pre prácu s obrazom bolo vybrané Raspberry Pi 2 ako lacná a komerčne dostupná platforma zaoberajúca sa analýzou obrazu. Zariadenie je napájané pomocou mikro-USB portu, je určené pre prácu v teréne, tým pádom je batéria nevyhnutným zdrojom energie. Momentálne je použitá batéria s kapacitou 12 Ah. Celková výdrž batérie sa momentálne pohybuje v rozmedzí 8-12 hodín. Spotreba je rovná 1-2 W v závislosti na vyťažení procesoru. Analýza obrazu je pomerne výpočetne náročná pri HD rozlíšení. Vysoké rozlíšenie je nutné práve pre veľkosť detekovaných objektov. Maximálna rýchlosť snímkovania dosahuje až 120 fps. Zariadenie bolo navrhnuté na chod za všetkých vonkajších podmienok. Celkový design a rozloženie kamier je zobrazené na obr. 1. Celý ovládací program je napísaný v jazyku Python [2] za použitia knižnice

OpenCv [3]. Model bol navrhnutý v aplikácii SolidWorks a vytlačený na 3D tlačiarni na ústave UTEE, ktorý bol zafarbený bielou farbou pre minimalizáciu ohrevu slnečnými lúčmi.

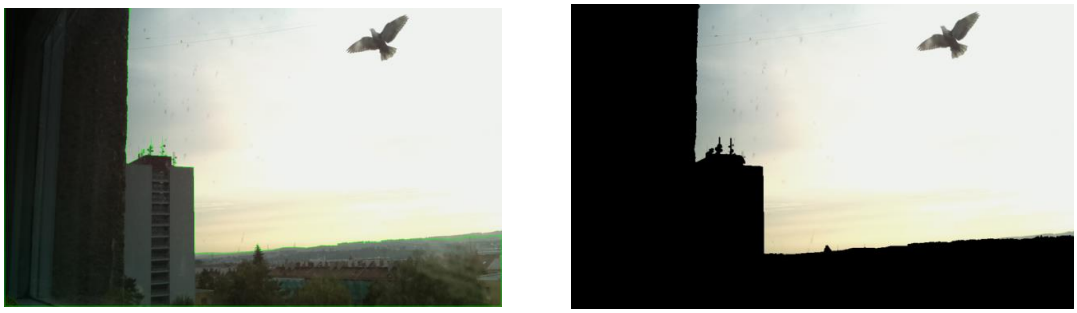


Obrázok 1: Samotné zariadenie

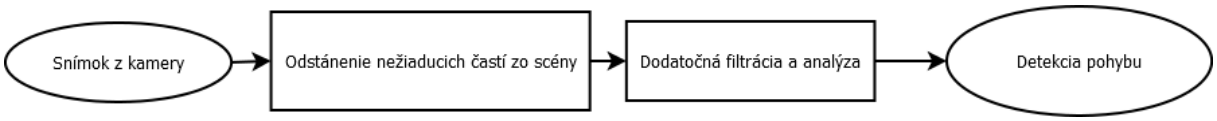
3. DETEKCIA A ODSTRÁNENIE NEŽIADUCICH ČASTÍ OBRAZU

Základnou potrebou pre správnu detekciu malých pohybujúcich sa objektov je odstránenie všetkých rušivých zložiek obrazu. Hrubé odstránenie rušivých častí, ako sú stromy, budovy a všetko ostatné, mimo plochy určenej na detekciu, je zrealizované ich detekciou a zmazaním.

Plocha a nepotrebné časti obrazu majú rozdielne farebné spektrum. Spektrum užitočnej časti je automaticky získavané počas chodu programu, a porovnávané s celým obrázkom. Na základe ich rozličnosti je zvolený prahovací filter, ktorý detekuje ich rozhranie, ktoré je znázornené vľavo na obr. 2. Po filtrácii, sú všetky detekované nepotrebné časti odstránené. Výstupný obrázok po hrubej filtrácii je znázornený na obr. 2 vpravo. Táto filtrácia sa prevedie vždy po zapnutí zariadenia, z čoho vyplýva nutnosť pevného umiestnenia kamier prípadne resetovanie zariadenia po každom premiestňovaní. Celý postup detekcie je zobrazený na obr. 3.



Obrázok 2: Snímok pred a po odstránení nežiaducich častí



Obrázok 3: Vývojový diagram

4. DODATOČNÁ FILTRÁCIA A DETEKCIA POHYBU

Náhodný šum je filtrovaný mediánovým filtrom. Po odstránení nežiaduceho šumu je pre každú kameru porovnávaný predchádzajúci a nasledujúci snímok. Pohyb je detekovaný zmenou rozloženia matice pixlov na nasledujúcom obraze. Rozdiel sa vyhodnocuje absolútnou diferenciou

oboch snímkov, ktorá vytvorí výsledný snímok, kde je tým svetlejší stupeň šedi, čím je rozdiel pixlov na danom mieste väčší.

Výstupný snímok po porovnaní, je prahovaný na bielu a čiernu zložku. Čiernej zložke odpovedá hodnota 0 a bielej zložke hodnota 255. Celkový počet rozdielnych pixlov je spočítaný a porovnaný s citlivosťou zariadenia, ktorá je štandardne volená na 5 pixlov. Jej hodnota je nastaviteľná podľa rozmeru detekovaných objektov. Výsledný obraz, ktorého príklad je znázornený na obr. 4, je po detekcii zaznamenaný a uchovaný na pamäti zariadenia.



Obrázok 4: Výsledný snímok zariadenia

5. ZÁVĚR

Zariadenie je určené pre prácu vo vonkajšom prostredí, pre pasívnu optickú lokalizáciu schopné detekovať pohybujúce sa objekty na oblohe zo všetkých scén, mimo priestoru pod zariadením. Uhol záberu kamier je dostatočne široký, na zachytenie rýchlo letiaceho objektu. Zariadenie je určené pre vonkajšie prostredie. V budúcnosti budú na zariadenie umiestnené fotorezistory so spínacím obvodom, pre automatické vypínanie zariadenia počas noci a znovu zapnutie v skorých ranných hodinách. Pre momentálne použitie, detekciu vtákov, nie je potrebná ich detekcia počas nočných hodín. Možným doplnením bude i senzor vlhkosti pre prípadné vypínanie počas dažďa.

Nevyhnutnými úpravami pred terénnym použitím, je umiestnenie tesnenia medzi podstavec a kryt zariadenia, umiestnenie látky, ktorá pohlcuje vlhkosť, pre minimalizáciu zarosenia.

REFERENCE

[1] *Basic motion detection and tracking with Python and OpenCV* [online]. [cit. 2016-03-29]. Dostupné z: http://www.pyimagesearch.com/2015/05/25/basic-motion-detection-and-tracking- with-python-and-opencv/

[2] *Python Guide Documentation* [online]., 1-113 [cit. 2016-03-14]. Dostupné z: https://media.readthedocs.org/pdf/python-guide/latest/python-guide.pdf

[3] *OpenCV 2.4.12.0 documentation: OpenCV Tutorials* [online]. , 1-451 [cit. 2016-03-14]. Dostupné z: http://docs.opencv.org/2.4/opencv_tutorials.pdf

MUSIC VOLUME ADJUSTMENT BASED ON DISTANCE DETECTION

Stanislava Serečunová

Bachelor Degree Programme (3), FEEC BUT

E-mail: xserec01@stud.feec.vutbr.cz

Supervised by: Oto Janoušek

E-mail: janouseko@feec.vutbr.cz

Abstract: The work is focused on user distance calculation from camera in real time by means of evaluating parameters of detected face. According to the distance of user from the camera sound level is adjusted, so the user can perceive the music on the same level. KLT method is chosen for the detection of face.

Keywords: face detection in real time, Viola-Jones, CAMShift, KLT, distance, volume

1. ÚVOD

Cieľom práce je zautomatizovať funkciu ovládania hlasitosti zvuku na základe vzdialenosti užívateľa od kamery. Hlasitosť sa pri zmene vzdialenosti adaptuje tak, aby užívateľ počul zvuk stále rovnakou intenzitou. Snahou počítačového videnia je porozumieť trojrozmernej scéne tak, aby počítač dokázal napodobniť určité úlohy, ktoré za normálnych okolností vykonáva človek.

2. NÁVRH REALIZÁCIE

2.1. METÓDA DETEKCIE A SLEDOVANIA TVÁRE POMOCOU HRANIČNÝCH BODOV

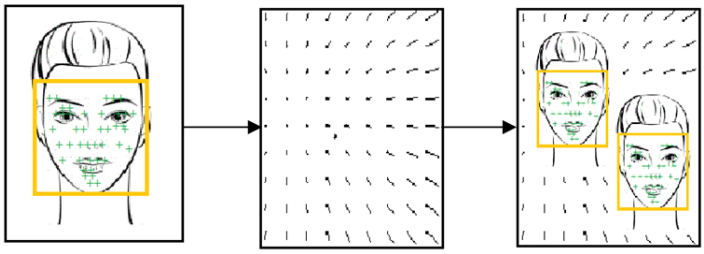
Základom metódy detekcie a sledovania tváre je detektor Viola-Jones. Detektor sa skladá z troch základných častí: vytvorenie integrálneho obrazu, určenie Haar-ových príznakov (Haar features) pomocou Haar-ových vlniek (Haar-like features), klasifikácia príznakov pomocou algoritmu AdaBoost (Adaptive Boosting). [1] Používa metódu kĺzavého okna, ktoré vyberie určitý región obrazu a ten sa vyhodnocuje porovnávaním výskytu príznakov, t.j. svetlých a tmavých obdĺžnikových oblastí, ako môžete vidieť na Obr. 1.



Obrázok 1: Detekcia tváre (vľavo Haar-ov príznak, uprostred lokalizácia príznaku na tvári, vpravo detegovaná tvár)

Metóda Kanade-Lucas-Tomasi označovaná ako KLT, využíva pre detegovanie a sledovanie tváre detekčný algoritmus Viola-Jones. Pri KLT metóde sú snímky snímané v okamžikoch za sebou a úzko spolu súvisia. Ako je vidieť na Obr. 2 metóda sleduje rad hraničných bodov na základe výpočtu optického toku, ktorý dokáže sledovať pohyb pixelov vo videosekvencii a na základe toho zistiť rýchlosť pohybu za časový interval. To sa využíva na zistenie presunu objektu, resp. aký po-

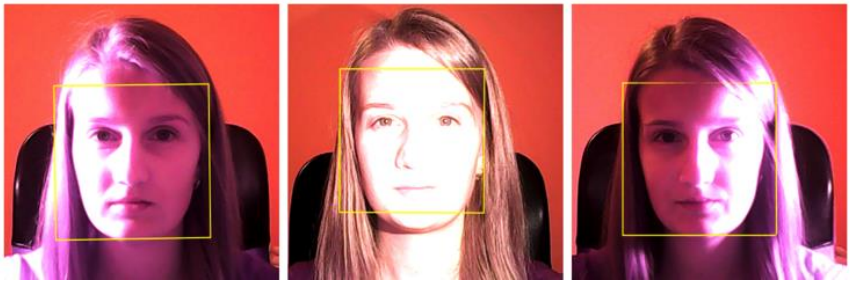
hyb objekt vykonáva. Sledovanie jedného pixelu nemôže byť použité, pretože je nepravdepodobné, že má výrazný jas vo vzťahu s pixelom predchádzajúceho snímku. Práve preto detektor využíva na sledovanie celé okno pixelov. To vedie k použitiu bohatého modelu s veľkým počtom parametrov, hoci malé okno s dvomi parametrami je v tomto prípade dostačujúce. [2]



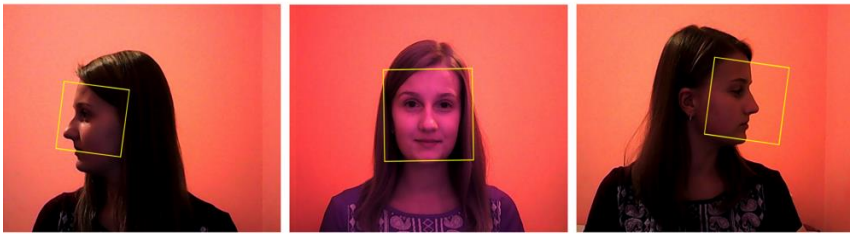
Obrázok 2: Princíp metódy KLT

Pri presvetlení scény metóda KLT detegovala tvár vo videosekvencii spoľahlivo (viď Obr.3) do rovinného útvaru, ktorý prispôsobuje výšku strany na základe jej šírky, teda chová sa ako štvorec. Práve toto zistenie je vhodné na realizáciu ďalších bodov práce.

Pri zmene otočenia hlavy detegovaný štvorec zachovával svoj tvar (viď Obr. 4). Pri otočení hlavy o 90° menil svoj obsah približne o 20%, no pri miernom natočení hlavy menil svoj obsah len približne o 0,06%.



Obrázok 3: (a) presvetlenie zľava, (b) frontálne presvetlenie, (c) presvetlenie sprava



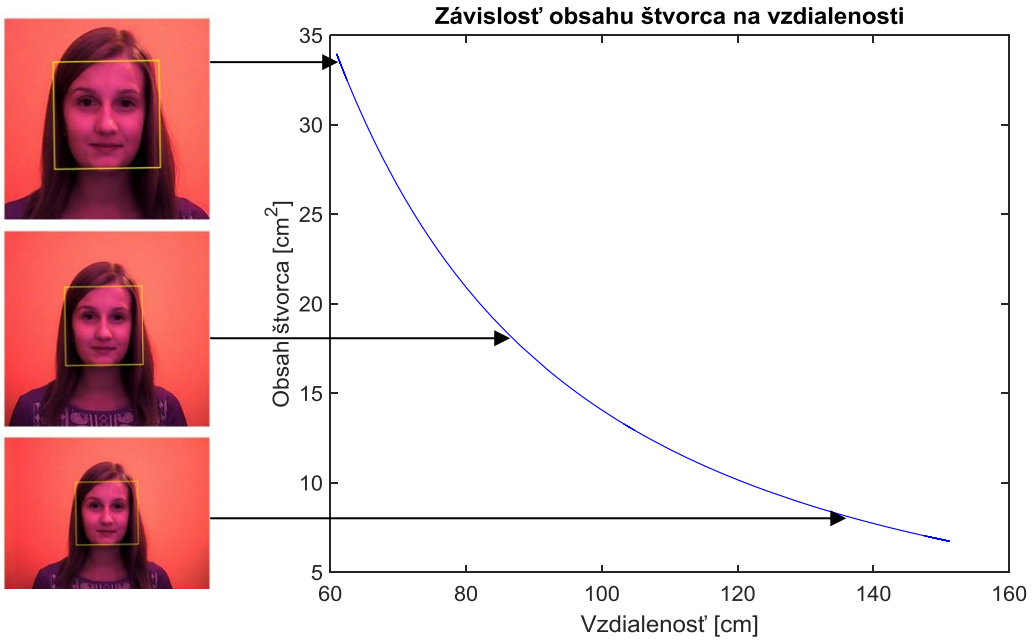
Obrázok 4: Detekcia tváre pri natočeniach tváre o 90°

2.2. VÝPOČET VZDIALENOSTI

Prenosová funkcia obsahu štvorca na vzdialenosti, na základe ktorej sa mení obsah štvorca, bola zistená pomocou testu. V miestnosti boli umiestnené detekčné značky, ktoré určovali presnú vzdialenosť sledovaného objektu od kamery. Objekt sa vzďaľoval o určitú vzdialenosť, pričom bola pri každej vzdialenosti určená veľkosť obsahu štvorca v danom momente. Na základe takto nameraných údajov bolo možné určiť funkciu závislosti, pri ktorej obsah štvorca so vzdialenosťou od zdroja exponenciálne klesá (viď Obr. 5). Funkciu je možné definovať ako

$$y = aS^x, \qquad (1)$$

kde rovnica každej reálnej hodnote *x* priraďuje práve jednu hodnotu *y*, ktorá vyjadruje vzdialenosť detegovanej tváre od kamery, *a* je reálna konštanta, $a > 1$, a $a \neq 1$, *S* je obsah detegovaného štvorca.



Obrázok 5: Závislosť obsahu detegovaného štvorca na vzdialenosti vo videosekvencii

2.3. REGULÁCIA HLASITOSTI

Pri návrhu spôsobu regulácie hlasitosti hudby podľa vzdialenosti poslucháča tak, aby poslucháč vnímal hudbu ustavične rovnako, bol pre zjednodušenie situácie zanedbaný vplyv odrazových vĺn v miestnosti. Ďalej bolo predpokladané, že zdroj zvuku je bodový a šíri sa izotropne. Intenzita zvuku bodového zdroja tak klesá so štvorcom vzdialenosti polomeru *r* od zdroja. [3] Na základe Weber-Fechnerovho zákona bolo určené, že hlasitosť, resp. intenzita zvuku, bude logaritmicky klesať so štvorcom vzdialenosti *r* od zdroja zvuku.

3. ZÁVER

Ušné implantáty sú neustále v technickom vývoji. Príkladom je aktívny stredoušný implantát, ktorý oproti bežným slúchadlám prevádza okolité zvukové signály na mechanické vibrácie. Navyše tieto nové implantáty používajú k regulácii hlasitosti bezdrôtový diaľkový ovládač, ktorý pacient môže použiť napríklad pri počúvaní televízora. Zautomatizovanie regulácie hlasitosti by mohlo byť v budúcnosti aplikované práve do takéhoto zariadenia, aby obmedzilo nadbytočné používanie ovládača a zefektívnilo tak použitie ušného implantátu pre pacienta.

REFERENCIE

[1] VIOLA, Paul; JONES, Michael. Rapid object detection using a boosted cascade of simple features. In: Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. IEEE, 2001. p. I-511-I-518 vol. 1.

[2] TOMASI, Carlo; KANADE, Takeo. Detection and tracking of point features. Pittsburgh: School of Computer Science, Carnegie Mellon Univ., 1991.

[3] HALLIDAY, David, RESNICK, Robert; WALKER, Jearl. Fyzika: vysokoškolská učebnice obecné fyziky. Vyd. 1. Brno: VUTIUM, 2003. ISBN 80-214-1868-0.

LED PIPETTING INDICATOR

Michal Demjan

Bachelor Degree Programme (2), FEEC BUT

E-mail: xdemja02@stud.feec.vutbr.cz

Supervised by: Pavel Škarvada

E-mail: skarvada@feec.vutbr.cz

Abstract: The main purpose of this work is making work for people in laboratories easier and more comfortable with limited budget. Project solved this problem and showed the affordable way how to improve laboratory environment and performances of laboratory workers. Designed and produced pipetting indicator, described on this work, is based on 8bit microcontroller. Indicator has ability to indicate up to 96 positions.

Keywords: pipetting, indication

1. ÚVOD

Z praxe je známe, že jedným z najväčších problémov pri manuálnom pipetovaní je zapamätanie si do ktorej skúmavky sa aplikoval roztok ako posledný. Pri nepozornosti takto môže ročne dôjsť aj k niekoľko tisícovým škodám. Je teda užitočné sa touto problematikou zaoberať, no zároveň treba zohľadniť finančné možnosti inštitúcií zastrešujúce pracoviská, v ktorých je táto problematika aktuálna.

Aj keď mikroskúmavkové dosky majú typizované rozmery ich okrajové časti si výrobcovia často individualizujú viď Obrázek 1. V bežných pracoviskách sa používajú hlavne 48(6x8) a 96(12x8) jamkové. Zariadenie by teda malo spĺňať univerzálnosť v spomenutých smeroch.

Existujú tiež komerčné riešenia tohto problému [1] [2] [3]. Vzhľadom na ich cenu, potrebu špecifického tabletu a malú intenzitu vyžarovaného žiarenia obrazovkou tabletu je ich používanie mizivé, či veľmi neprehľadné.



Obrázek 1: Mikroskúmavkové dosky s rôznymi okrajovými časťami [4].

2. POPIS NAVRHNUTÉHO RIEŠENIA

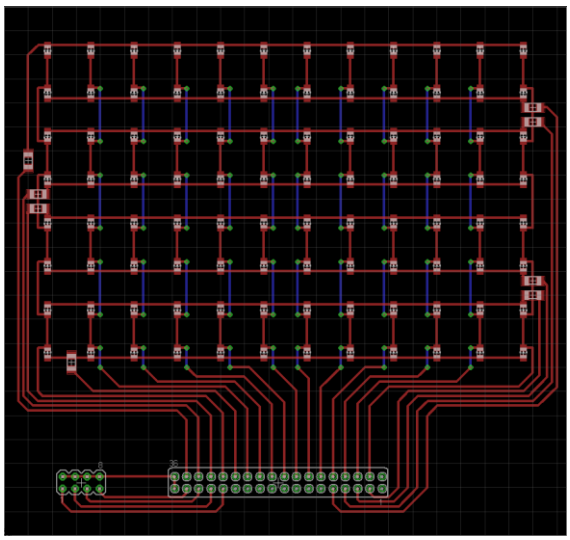
Zariadenie používa pre indikáciu maticu SMD LED diód, ktorá je osadená na doske plošných spojov. Jednotlivé LED diódy osvetľujú len jednu, jej pridelenú skúmavku. Teda nedochádza k osvetľovaniu okolitých skúmaviek.

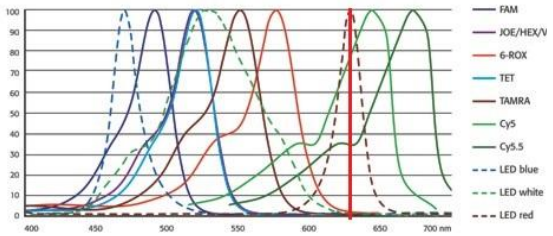
Na ovládanie LED matice zariadenie využíva platformu Arduino s čipom ATMega16U2, vďaka ktorej sa náklady držia v rozumných číslach. Vzhľadom k veľkému počtu ovládaných diód nie je s využitím bežných kontrolórov možné každú diódu ovládať zvlášť. Preto boli umiestnené do matice 12x8, ku ktorej ovládaniu postačí 20 výstupných liniek. Pri vhodnom návrhu nie je nutné použiť zosilňovacie tranzistory a celková spotreba (pri zapnutí všetkých LED) môže byť s rezervou pokrytá z výstupných portov platformy, z USB alebo z 9 V konektora, čo výrazne zjednodušuje konštrukciu.

Zapínanie jednotlivých LED funguje na základe XY sústavy, kedy sa na riadok privedie logická 1 a na stĺpec logická 0, čoho výsledkom je zasvietenie jednej LED diódy. V danom momente svieti vždy iba jedna LED dióda. Viac k danej problematike je možné nájsť v [5]. Na vyvolanie dojmu kontinuálneho žiarenia musí byť spínacia frekvencia vyššia ako 24 Hz a aby sa predišlo únave očí, tak sa používa frekvencia vyššia ako 100 Hz. Spínacia frekvencia 1 KHz bezpečne spĺňa tieto podmienky a zamedzuje vzniku zvýšenej únavy používaním tohto zariadenia. Vlnová dĺžka žiarenia vyžarovaná použitými diódami je 625 nm, čo nespôsobí rozpad u väčšiny bežne používaných fluorescenčných farieb viď Obrázek 2.

Hardware-ové voliče s platformou komunikujú a umožňujú spínanie LED diód buď v horizontálnom alebo vertikálnom smere. Zapnutie LED diódy iniciuje hardware-ové tlačidlo v podobe nožného pedálu. V prípade omylu je tu prítomné aj tlačidlo pre krok späť.

Povrch dosky plošných spojov bol ošetrený ochranným lakom odolným voči anorganickým kyselinám. Lak zamedzuje skratu a zjednodušuje čistenie v prípade kontaminácie.



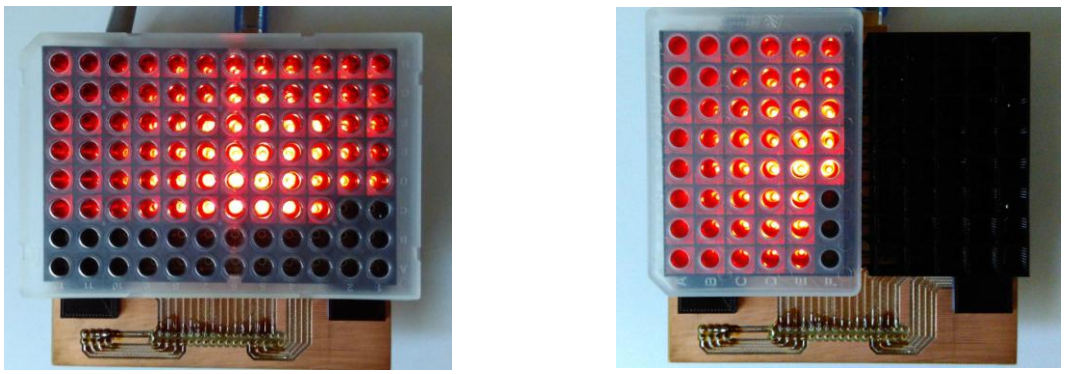


Obrázek 2: Navrhnutá doska plošných spojov s maticou SMD LED diód a absorbancia jednotlivých používaných farieb s vyznačenou vlnovou dĺžkou použitej LED [6].

3. DOSIAHNUTÉ VÝSLEDKY

Zhotovené zariadenie spĺňa všetky požiadavky pre správne fungovanie. Zariadenie takisto spĺňa podmienku nízkych finančných nákladov a jeho zhotovenie nevyžaduje žiadne náročné operácie, čo len prispieva k jeho prípadnému reálnemu využitiu v praxi.

Jediným problémom je samotné žiarenie, ktoré môže aktivovať niektoré svetlocitlivé farby. Jedným z riešení je naprogramovať inverzný mód, kde by tlačidlo iniciovalo vypnutie a nie zapnutie jednotlivých LED.



Obrázek 3: Spustené zariadenie v režimoch horizontálnej a zvislej iniciácie zapnutia LED diód.

4. ZÁVER

Tento projekt predstavil problém s ktorým sa stretávajú pracovníci v laboratóriách a úspešne ho vyriešil. Poskytol riešenie, ktoré je schopné sa uplatniť aj v praxi a zjednodušiť tak prácu mnoho ľuďom. Bola ukázaná možnosť ovládania 96 LED v maticovom režime a aplikácia pre využitie pri pipetovaní. Použité riešenie je jednoduché, lacné a poskytuje dostatočnú variabilitu pre prípadnú integráciu do vyšších systémov, prípadne má potenciál pre rozšírenie o automatickú detekciu pipety.

POĎAKOVANIE

Poďakovanie patrí MUDr. Michalovi Jurajdovi, Ph.D. a Ing. Vratislavovi Harabišovi, Ph.D. za konzultácie a spoluprácu.

REFERENCIE

[1] TRACKAMAN™. GILSON®. [vid.7.3.2016]. Dostupné z: http://www.gilson.com/en/Pipette/Products/45.265/Default.aspx#.Vt2lx7x3A2s

[2] PLATR SMART PIPETTING. Biosistematika®. [vid.25.3.2016]. Dostupné z: http://biosistemika.com/platr/

[3] Pipette Guide 96. BFrontier. [vid.25.3.2016]. Dostupné z: https://itunes.apple.com/us/app/pipette-guide-96/id548236701?mt=8

[4] Doska PCR Mikroskúmaviek, PP. In: Helago® [online]. ©2016 HELAGO-SK s.r.o [vid.7.3.2016]. Dostupné z:

http://www.helago-sk.sk/files/thumbs/mod_eshop/produkty/full/4675.1033825454.jpg

[5] Zbyšek Voda & tým HW Kitchen. PRŮVODCE SVĚTEM ARDUINA. Bučovice: Nakladatelství Martin Střiž, 2015. ISBN 978-1-4200-8201-2.

[6] TOptical Excitation by Led technology. In: Biometra [online]. ©2014-2016 Biometra GmbH [vid.7.3.2016]. Dostupné z: http://www.biometra.de/files/themes/biometra_standard/bilder/toptical_anregung.jpg

Bakalářské projekty

Mikroelektronika a komunikační technologie

AN EXTERNAL USER INTERFACE WITH USB

Michal Sládeček

Bachelor Degree Programme (3), FEEC BUT

E-mail: xslade18@stud.feec.vutbr.cz

Supervised by: Edita Hejátková

E-mail: hejatka@feec.vutbr.cz

Abstract: This article deals with problems of developing user interface which can communicate with host system via USB. The article includes comparation of existing alternatives and description of USB class used for communication between external system. The main part of work is choosing hardware parts for device and definition of software structure. The aim is to design a schema and build a hardware prototype.

Keywords: User interface, USB, hardware, display, joystick, Python

1. ÚVOD

Uživatelské rozhraní má umožnit uživateli ovládání systému, a poskytovat mu zpětnou vazbu pomocí zobrazovacích a akustických prvků. Hlavním důvodem vývoje tohoto zařízení je především vysoká cena a malá flexibilita již hotových komerčně prodávaných modulů. Další motivací je také fakt, že neexistuje žádné vhodné řešení problému ve veřejné (open source) doméně. Primární využití zařízení je určeno jako prvek pro automatizaci výroby. Operátor pracující ve výrobě, z displeje vyčítá instrukce, které má provádět a pomocí ovládacích prvků vrací zpětnou vazbu hostitelskému systému. Za hostitelský systém se dá považovat jakékoliv zařízení s operačním systémem (počítač, minipočítač, tablet apod.). Podmínkou je jednoduchá a levná výroba, která by měla umožňovat jak ruční, tak plně automatizovanou montáž.

2. DOSTUPNÁ KOMERČNÍ ŘEŠENÍ

Matrix Orbital patří mezi nejpoužívanější moduly uživatelského rozhraní. Obsahem modulu je displej většinou o rozlišení 192x64 px. Cena modulu se pohybuje nad hranicí 100 USD. [1] **Crystalfontz**, je firma vyrábějící podobné displeje jako jsou ty od Matrix Orbital, fungují na podobném principu. Výhodou zařízení je skutečnost, že poskytují dodatečné vstupně/výstupní piny. Dominantou modulu je displej s rozlišením 244x68 px. Cena se pohybuje v okolí 85 USD. [2] **PicoLCD**, je displej distribuovaný firmou Mini-Box.com. Displej je určen jak pro operační systémy Windows, tak Linux. Modul však funguje trochu jinak, než předchozí dva zástupci. Větší část řídící logiky je totiž realizována na straně hostitelského systému ve formě knihoven. To vyžaduje větší nároky na hostitelský systém. Dané řešení však výrazně zvyšuje flexibilitu použití modulu a zároveň snižuje komplexitu hardware. Modul je dodáván bez krabičky za cenu 60 USD. [3]

3. NÁVRH VLASTNÍHO ŘEŠENÍ

Aby bylo navrhované zařízení co nejefektivněji využitelné a jeho cena nepřesahovala komerčně vyráběné moduly, je uživatelské rozhraní navrženo tak, aby byla veškerá logika přenesena na hostitelský systém.

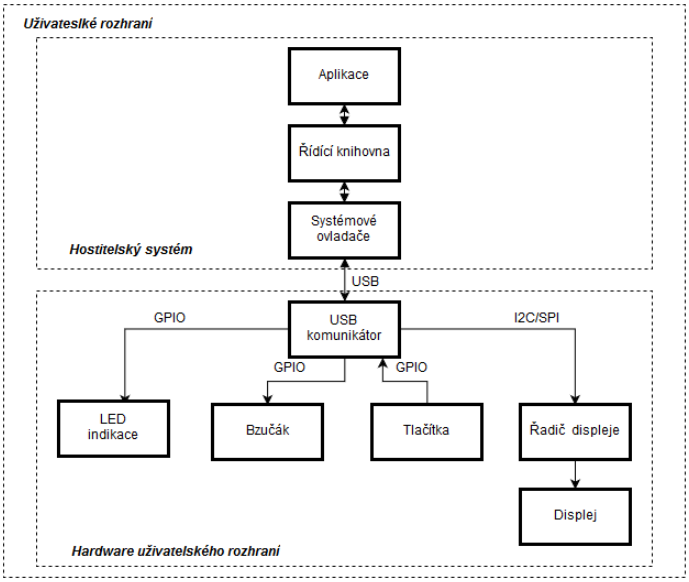
3.1. KOMUNIKACE PŘES USB

Komunikace přes USB probíhá vždy v rámci protokolu definovaného pro určitou třídu zařízení. Díky těmto třídám je pak možné zařízení v hostitelském systému rozpoznat a dále s ním pracovat.

Nejvhodnější třídou, pro komunikaci uživatelského zařízení s hostitelským systémem je třída **HID**, která je využívána například pro klávesnice, myší a podobná zařízení. Výhodou třídy je podpora většiny operačních systémů, které nepotřebují pro tyto zařízení na míru vytvořený ovládací program. Nevýhodou je však vcelku malá propustnost dat, která je však pro tento typ zařízení dostačující. [4]

3.2. ARCHITEKTURA ZAŘÍZENÍ

Výsledná architektura zařízení je znázorněna na obrázku č. 1. Logiku celého zařízení přebírá hostitelský systém (např. Raspberry Pi). A pomocí ovládací knihovny napsané v jazyce Python umožňuje jednoduché ovládání USB komunikátoru, který řídí jednoduchou elektroniku okolo něj.



Obrázek 1: Architektura výsledného zařízení

3.3. SOUČÁSTKOVÁ ZÁKLADNA

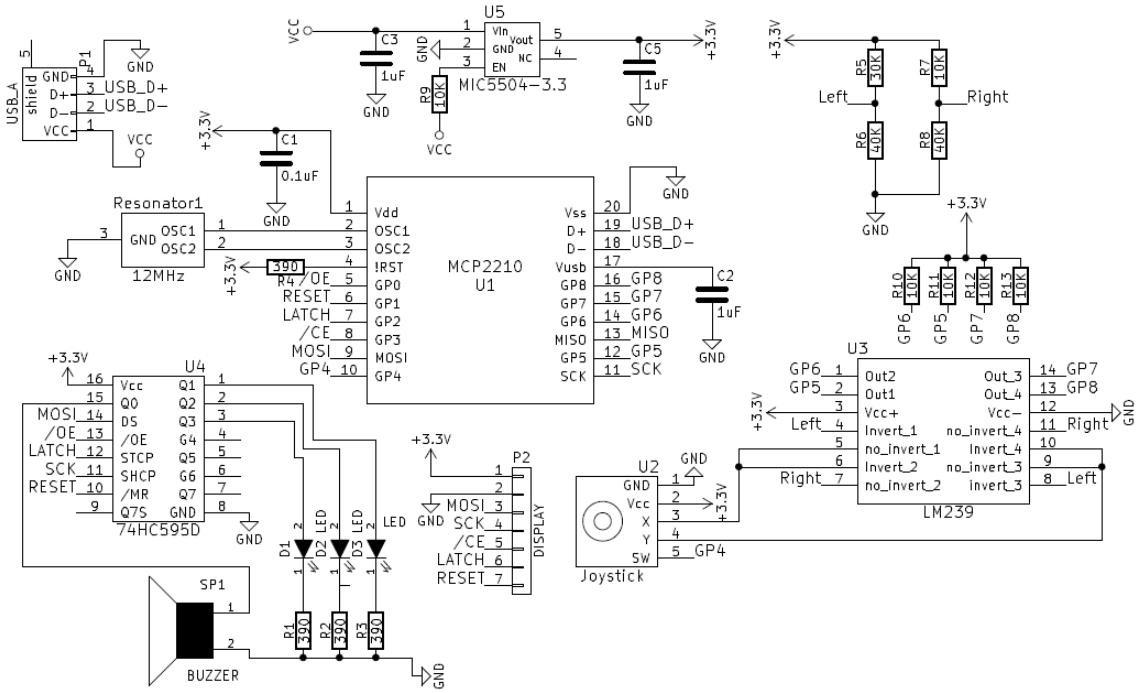
V tabulce č. 1 jsou znázorněny stěžejní součásti zařízení. Hlavním prvkem je obvod MCP2210 [5], který plně zvládá komunikaci přes USB HID class a spolu s devíti GPIO piny zajišťuje bezproblémovou implementaci navržené architektury. Jako zobrazovací prvek byl vybrán 4,3" displej s rozlišením 192x64 px disponující řadičem IST3020, který v porovnání kvality zobrazení spolu s pozorovacími úhly v závislosti na výstupní ceně jasně převyšoval ostatní displeje na trhu. Protože řídící prvek neobsahuje A/D převodník pro vzorkování ovládacího prvku (Joystisku), jsou v zapojení použity komparátory (LM239), které pro jednoduché vzorkování postačují. Posledním obvodem je posuvný registr 74HC595, který rozšiřuje omezený počet výstupních pinů MCP2210 a tím umožňuje připojit další především vizuální prvky.

Popis	Označení	Cena
USB-to-SPI Master converter + GPIO	MCP2210 [5]	2,10 USD
Regulátor napětí	MIC5504-3.3	0,24 USD
Monochromatický displej	4,3" 192x64 Display Graphic (IST3020)	9,15 USD
Mechanický ovládací prvek	3D analog joystick	1,10 USD
Komparátory pro řízení joysticku	LM239	0,13 USD
Posuvný registr pro navýšení výstupů	74HC595	0,37 USD

Tabulka 1: Součástková základna

4. SCHÉMA ZAPOJENÍ

Výsledkem vytvořené architektury a vybraných součástek je schematický návrh. Protože napájecí větev LCD displeje vyžaduje 3,3 V, je do celkového návrhu implementován regulátor napětí MIC5504 s maximálním výstupním proudem 300 mA.



Obrázek 2: Schéma zapojení uživatelského rozhraní.

5. ZÁVĚR

Práce popisuje návrh zařízení sloužícího jako uživatelské rozhraní systémů vyžadující zpětnou vazbu od uživatele. Jsou porovnány komerčně prodávané výrobky a v závislosti na požadavcích vytvořena architektura vlastního zařízení. S tím souvisí také výběr součástek, z kterého lze odvodit výsledná cena, která spolu s deskou plošných spojů vychází kolem 30 USD. Celková cena se zvedne na 40 USD, pokud vezmeme v úvahu také zapouzdření z ohýbaných plechů. Systém je ovládán pomocí knihovny napsané v programovacím jazyce Python. Výsledná aplikace pracující s ovládací knihovnou je následně spuštěna v minipočítači Raspberry Pi.

REFERENCE

[1] Matrix Orbital Corporation [online]. Calgary, Alberta, CA, 2015 [cit. 2015-12-13]. Dostupné z: https://www.matrixorbital.com/

[2] CRYSTALFONTZ AMERICA INCORPORATED. Crystalfontz [online]. 2015 [cit. 2015-12-13]. Dostupné z: https://www.crystalfontz.com/product/cfa835tfk-display-graphic-lcd-244x68

[3] PicoLCD 256x64 Sideshow. Mini-box.com: Pico size computing [online]. 2015 [cit. 2015-12-13]. Dostupné z: http://www.mini-box.com/picoLCD-256x64-Sideshow-CDROM-Bay

[4] USB org. [online]. 2015 [cit. 2015-12-13]. Dostupné z: http://www.usb.org/developers/docs/devclass_docs/

[5] Microchip Technology Inc. In: MCP2210 datasheet [online]. 2011 [cit. 2015-12-13]. Dostupné z: http://ww1.microchip.com/downloads/en/DeviceDoc/22288A.pdf

LASER TARGET FOR SIMULATING SHOOTING SPORT

Martin Poloch

Bachelor Degree Programme (3), FEEC BUT

E-mail: xpoloc00@stud.feec.vutbr.cz

Supervised by: Martin Friedl

E-mail: friedl@feec.vutbr.cz

Abstract: This work deals with complete design of a system for simulating shooting sport. Key aspects of the design are similarity with shooting from ordinary mechanical weapons, multifunctionality and low costs. The core of the simulation is a laser transmitter in the weapon and a photosensitive receiver in the target. The final version of the system is controlled by a group of microcontrollers and provides several types of feedback. For example lighting the place that has been hit with LED, sound signals, displaying shooting results, or evaluating them with a computer.

Keywords: LASER, microcontroller, ATmega8, LED, photosensor

1. ÚVOD

Sportovní střelba a její formy mohou být velmi nákladnou záležitostí. Například pro biatlon se využívají zbraně, které běžné stojí 50 000 – 100 000 Kč. Cena za vystřílené náboje za jeden trénink či závod se může pohybovat v rozmezí 500 – 1000 kč. Klasické mechanické zbraně jsou navíc také nebezpečné a neekologické (likvidace vystřelených nábojů, zplodiny doprovázející výstřel, prostředky na čištění zbraně atd.) . Cílem tohoto projektu bylo tyto nedostatky odstranit a navrhnout tedy systém, který ze střelby vytvoří bezpečný, ekologický, levnější, a tedy celkově přístupnější sport. Návrh byl ověřen na dvou prototypech nazvaných Gabriela a Rachel.

2. „STŘELBA LASEREM“

Jako vhodné řešení takovéhoto zadání byl zvolen systém zbraně vysílající laserový paprsek a elektronického terče pro zachycení a vyhodnocení „střelby“. Jako vysílač laserové „střely“ byly použity diody z běžně dostupných laserových ukazovátek s výkonem do 5 mW. Výběr světlocitlivého senzoru už byl obtížnější. Ze všech zvážených a otestovaných možností vyšlo nejlépe použití kapacity PN přechodu diody zapojené v závěrném směru.

2.1. SNÍMÁNÍ SVĚTLA POMOCÍ KAPACITY LED

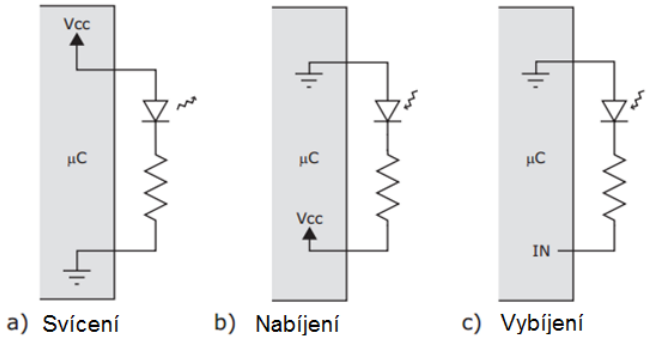
Dioda zapojená v závěrném směru se nabije a při osvícení světlem určité vlnové délky se dioda vybíjí. Rychlost vybíjení je úměrná intenzitě dopadajícího světla. Průzkumem a testováním tohoto efektu bylo zjištěno, že LED je vybíjena prakticky pouze světlem o nižší vlnové délce, než je vlnová délka světla diodou emitovaného při klasickém zapojení. Nejlépe pak detekují záření o stejné vlnové délce (nebo o několik nm menší) jako je vlnová délka emitovaného záření při normálním zapojení. To tedy znamená, že například zelené světlo (vlnová délka okolo 540 nm) bude špatně vybíjet červené LED (vlnová délka okolo 680 nm), velmi dobře zelené diody, a modré LED (vlnová délka okolo 450 nm) prakticky vůbec. [1]

U všech podobných zařízení je těžké odfiltrovat světlo z laseru od běžného denního světla. Na přímém slunci pak zařízení často chybují nebo nefungují vůbec. Proto je takováto filtrace vlnových délek pro tento projekt velmi prospěšná. Pro splnění zadání bakalářské práce byly vybrány LED emitující na vlnové délce 415 nm. Tato vlnová délka je na hranici viditelného světla (modrá barva)

a UV záření. Odfiltruje nám tedy nejsilnější složku světelného spektra – zelené záření a zároveň umožňuje „střelbu“ z běžně dostupného laseru s nejnižší vlnovou délkou 405 nm.

2.2. ŘÍZENÍ SNÍMÁNÍ MIKROKONTROLÉREM

Pro použití LED jako světelného senzoru je ji třeba nabít a poté sledovat, jak rychle se vybila. K řízení tohoto cyklu v tomto zařízení slouží mikrokontroléry. Každá LED je zapojena na 2 digitální piny mikrokontroléru (obr. 1) a proto není problém ji nabít. Na anodu se přivede kladné napětí – log. 1 a katoda se ponechá v log. 0. Tento proces trvá pouze několik µs. Poté se pin zapojený na anodu převede do vstupního režimu a zkoumá se, jak rychle se na pinu objeví log. 0 (dioda je vybita). Výhodou tohoto zapojení je, že ze stejné diody se může snímat a vhodným přepnutím pinů s ní i svítit. V tomto projektu je využito periodického střídání fáze svícení a fáze snímání vysokou rychlostí, takže lidské oko není schopno tyto fáze rozlišit, a je tedy možné snímat i ze zdánlivě svítící diody. Dioda přitom svítí jen krátký úsek z celkové periody, takže snímání zůstává spolehlivé.



Obrázek 1: Průběh řízení chování LED z mikrokontroléru [1]

Pro vytvoření laserového terče je třeba vytvořit pole desítek až stovek LED. Každá LED tedy nemůže být zapojena na 2 piny mikrokontroléru. Bylo tedy použito zapojení se společnou katodou, kdy katody všech diod jsou zapojeny na stejný pin mikrokontroléru. Úspěšně bylo také vytvořeno maticové zapojení, kde byly společné anody a katody, ale tím se několikanásobně snížila citlivost snímání a ta nakonec dostala přednost před úsporou mikrokontrolérů. Proto je v prototypu Gabriela použito k ovládání snímacích LED 13 mikrokontrolérů a v prototypu Rachel dokonce 25.

3. ZVOLENÝ MIKROKONTROLÉR

Hlavním mikrokontrolérem tohoto projektu se stal typ ATmega8 od společnosti Atmel. Kromě jednoduchosti programování byla hlavním kritériem co nejnižší cena na jeden vstupně/výstupní pin. ATmega jich má sice „pouze“ 20 (po odečtení pinu pro společnou katodu a jednoho pinu na komunikaci), ale jeho cena na 1 pin byla stále nejnižší z uvažovaných mikrokontrolérů.

Kromě mikrokontrolérů ovládajících snímací LED jsou v obou prototypech ještě další k ovládání segmentových displejů zobrazující informace o střelbě, mikrokontroléry ve zbraních a hlavní mikrokontrolér sbírající informace ze všech ostatních. Všechny tyto mikrokontroléry jsou opět ATmega8, kromě hlavního mikrokontroléru v prototypu Rachel, který byl natolik složitý, že požadavky na paměť převyšovaly možnosti tohoto čipu, a bylo tedy sáhnuto k mikrokontroléru s větší pamětí – ATmega16. Všechny tyto mikrokontroléry spolu komunikují pomocí jednodrátové komunikace (kvůli úspoře pinů) založené na protokolu, který byl vytvořen speciálně pro tento projekt.

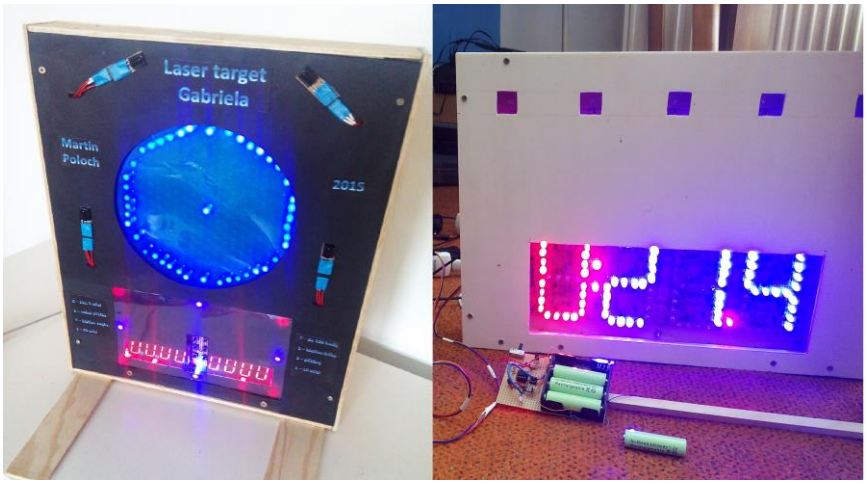
4. KOMUNIKACE ZBRANĚ A TERČE

Zbraň při stisku spouště vyšle laserový signál, který zasáhne jednu z LED. Terč to zaznamená a patřičně se podle toho zachová. Běžně se ale stává, že výstřel terč nezasáhne. Ten ale informaci o výstřelu musí dostat, aby byl schopen vyhodnocovat i netrefení terče. V prototypu Gabriela byla

k tomuto účelu použita infračervená komunikace pomocí IR LED a IR přijímače TSOP 1738. Tato komunikace neměla dostatečné parametry, zejména dosah se nepodařilo zvětšit nad 20 metrů. Navíc byla rušena slunečním zářením, proto v prototypu Rachel byla použita komunikace pomocí nRF24 modulů. Řešení to bylo několikanásobně dražší, ale dosah komunikace se zvýšil spolehlivě nad 50 metrů. Díky těmto modulům navíc terč může posílat informace i do dalších zařízení (jako např. osobní počítač), čehož je v prototypu Rachel plně využito.

5. PROTOTYP GABRIELA

Tento starší model je zaměřen kromě sportovní i na zábavní střelbu. Obsahuje 10 prstenců vytvořených z 253 LED. Tyto prstence jsou vytvořeny pro obodování střel od 1 do 10 bodů. Terč má 8 unikátních režimů střelby např. 10 střel, biatlon, moderní pětiboj, rychlopalba a další. V závislosti na zvoleném režimu svítí krajní nebo jiný prstenec LED a prostřední dioda, jak je vidět na obrázku 2. Vždy se také rozsvítí aktuálně zasažená LED. Vyhodnocování dosažených výsledků probíhá v reálném čase na sedmi-segmentových displejích, ať už zakomponovaných do terče, nebo v externím modulu velkého displeje (obrázek 2 vpravo). Jedná se o časové údaje, celkový počet bodů, počet bodů poslední střely a počet zbývajících střel. Pro prodloužení zmíněné nespolehlivé IR komunikace byla vytvořena sada dalších modulů s nimiž již terč spolehlivě funguje na 20 metrů.



Obrázek 2: Terč prototypu Gabriela, modul velkého displeje a prodlužovací modul

6. PROTOTYP RACHEL

Prototyp Rachel byl navrhován se zaměřením na biatlon a obsahuje tedy 5 terčů se 6 prstenci LED. Multifunkčnost byla ale zachována a rozšířena a terč tak obsahuje dokonce 54 režimů střelby. Kromě spolehlivější komunikace se zbraní byl doplněn zvukovým modulem pro poskytování ještě lepší zpětné vazby střelci a počítačovou aplikací, která slouží nejen jako zaznamenávání údajů, ale také jako simulace závodění s virtuálními protivníky.

7. ZÁVĚR

Pomocí popsaných postupů a technologií byl úspěšně vytvořen ekologický a bezpečný systém simulující sportovní střelbu, jehož pořizovací cena je nižší v porovnání s existujícími systémy. Díky multifunkčnosti terčů má navíc projekt obrovský potenciál z hlediska komerční využitelnosti.

REFERENCE

[1] Very Low-Cost Sensing and Communication Using Bidirectional LEDs. *MITSUBISHI ELECTRIC RESEARCH LABORATORIES* [online]. [cit. 2015-12-12]. Dostupné z: http://www.merl.com/publications/docs/TR2003-35.pdf

PRINTED PLANAR INDUCTOR

Pavel Krejčí

Bachelor Degree Programme (3), FEEC BUT

E-mail: xkrejc49@stud.feec.vutbr.cz

Supervised by: Ivan Szendiuch

E-mail: szend@feec.vutbr.cz

Abstract: This work deals with issues of application of the thick film technology in the field of planar inductors. It presents the essential impact of parameters on designed coil. Subsequently it describes equation for calculation of inductance of single layer inductor. Finaly, exemples of designed coil are shown and some technological ways are proposed.

Keywords: planar inductor, printed circuits, thick film technology, non-conventional application of thick film, Greenhouse formula

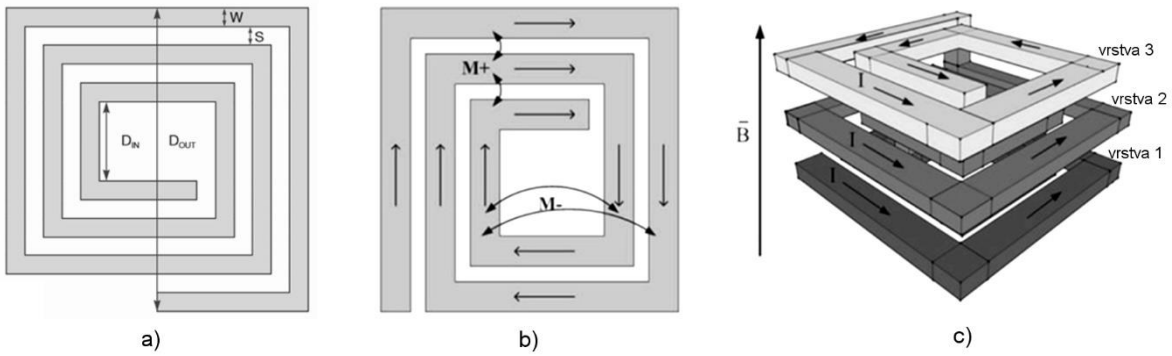
1. ÚVOD

Bezdrátový přenos elektrické energie pomocí indukčního či rezonančního principu si v posledních letech vydobývá čím dál tím větší pozornosti pro nesčetnou perspektivitu aplikací v nejrůznějších odvětvích (dobíjení mobilních telefonů, GPS navigací, biosenzorů). Tento článek pojednává o nekonvenčním způsobu vytváření planárních induktorů pomocí tlustovrstvové technologie na keramických a polykarbonátových flexibilních substrátech.

2. NÁVRH PLANÁRNÍHO INDUKTORU

Předmětem návrhu je tlustovrstvový induktor v řádech desítek μH s co nejnižší hodnotou parazitního sériového odporu a maximální účinností v oblasti frekvencí užívaných pro indukční přenos (desítky kHz až jednotky MHz). Cílem je zanalyzovat vztah mezi geometrickými parametry, hodnotou indukčnosti, parazitním sériovým odporem a činitelem jakosti. Dále bude posouzen vliv užitých past, substrátů a některých technologických námětů pro zdokonalení vlastností výsledné struktury.

2.1. PARAMETRY



Obrázek 2.1: a) geometrické parametry čtvercového induktoru **b)** pozitivní (M+) a negativní (M−) vzájemné indukčnosti závitů **c)** soustředný tok elektrického proudu ve vícevrstvé struktuře [1] [2]

Celková hodnota indukčnosti je dána součtem indukčnosti vlastní a vzájemné. Jak je patrno z obrázku 2.1b, blíže středu cívky začínají být vzájemné negativní indukčnosti mezi jednotlivými závity výraznější. Není vždy žádoucí zmenšovat vnitřní plochu, umožňuje-li to konkrétní aplikace.

- Počet závitů: S rostoucím počtem závitů roste i výsledná indukčnost. Zároveň však roste i hodnota parazitního odporu (pro totožnou pastu) a výsledný motiv zabírá větší plochu (je-li zachována šířka cesty w a rozestupy s).
- Šířka vodivé cesty: S rostoucí šířkou vodivé cesty a tedy se zmenšujícím se vnitřním průměrem, výsledná indukčnost klesá. Čím jsou vnitřní závity blíže sobě, tím je více převládající negativní vzájemná indukčnost M- (obrázek 2.1b). Nicméně současně s rostoucí šířkou vodivé cesty se snižují ztráty parazitního sériového odporu.
- Rozestupy vodivých cest: Čím jsou rozestupy větší, tím je výsledná indukčnost nižší, jelikož vnitřní průměr se zmenšuje též. Menší rozestupy vedou k vyšší hodnotě vzájemné pozitivní indukčnosti mezi vinutími, čímž se výsledná indukčnost cívky zvyšuje.

Z výše uvedeného můžeme vyvodit, že pro navýšení celkové indukčnosti planárního induktoru je možno [1] [2] [3]: zvýšit počet závitů, zvýšit velikost vnitřního průměru cívky nebo vytvořit vícevrstvé struktury a tím navýšit vzájemnou indukčnost mezi závity jednotlivých vrstev (obrázek 2.1c). První dvě možnosti v důsledku zabírají více plochy na substrátu. Poslední zmíněná umožňuje vyšší stupeň integrace. Možný výpočet vícevrstvých struktur je popsán v literatuře [1].

2.2. VÝPOČET INDUKČNOSTI

Předpokladem pro úspěšné nahrazování konvenčních způsobů vytváření cívek je dostatečná přesnost výpočtů indukčnosti navrhovaného planárního induktoru, jelikož je velmi náročné dostavovat vlastnosti po natisknutí struktury. Hodnotu indukčnosti planárního jednovrstvového induktoru L_{des} je možno vypočítat dle následujícího vzorce, známého jako „Greenhouse formula“ [1] [2] [3] [4] [5]:

$$L_{des} = \frac{\mu_0 \ N^2 \ D_{avg} \ C_1}{2} \left[\ln\left(\frac{C_2}{\rho}\right) + C_3 \ \rho + C_4 \ \rho^2 \right] \tag{1}$$

$$\rho = \frac{D_{\max} - D_{\min}}{D_{\max} + D_{\min}} \quad ; \quad D_{avg} = \frac{D_{\max} + D_{\min}}{2} \tag{2}$$

kde: N je počet závitů,
D_{max} a D_{min} jsou vnější/vnitřní průměr,
D_{avg} je tzv. střední hodnota geometrického průměru,
ρ je poměr volné plochy,
μ_0 je permeabilita vakua ($4\pi * 10^{-7}$ Hm^{-1}),
$C_1 - C_4$ jsou koeficienty dané tvarem induktoru (tabulka 2.1)

Tabulka 2.1: Koeficienty dané tvarem planárního induktoru [1] [4]

Tvar	C1	C2	C3	C4
Čtvercový	1,27	2,07	0,18	0,13
Hexagonální	1,09	2,23	0	0,17
Oktagonální	1,07	2,29	0	0,19
Kruhový	1	2,46	0	0,2

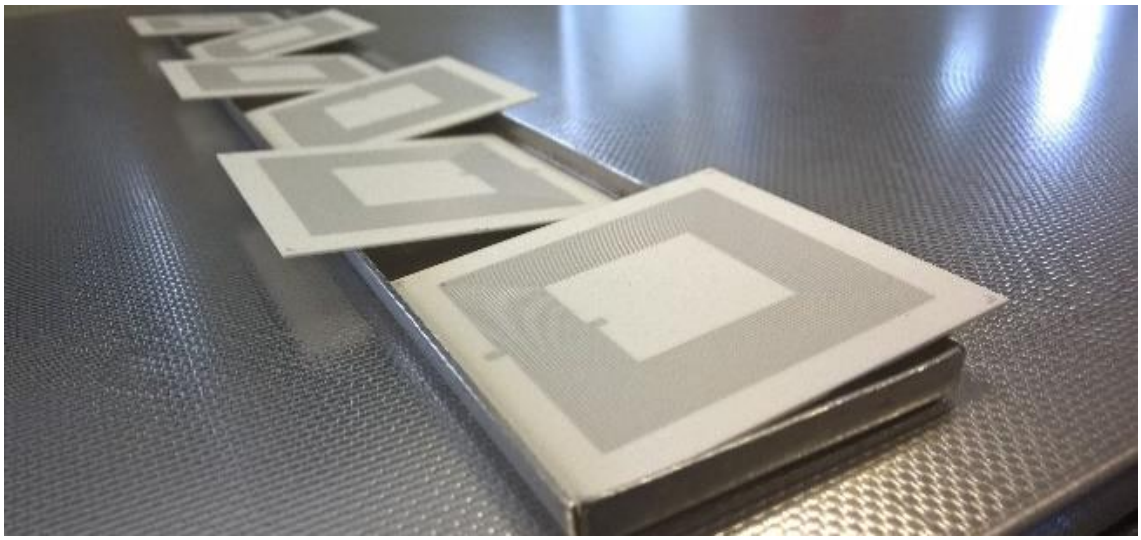
Přesnost výsledku uváděné rovnice klesá úměrně s rostoucími rozestupy mezi vodivými cestami. Maximální odchylka se pohybuje okolo 8% pro rozestupy třikrát větší, než je šířka vodivé cesty. [1] Nicméně je záhodno navrhovat induktory s rozestupy mezi vodivými cestami menšími, či rovnými šířce vodivé cesty. Větší rozestupy snižují magnetické vazby mezi závity a zabírají větší plochu.

2.3. REALIZACE

Rozměry motivu byly zvoleny s ohledem na běžné rozměry zadních krytů mobilních telefonů. Pro depozici sítotiskem na keramický substrát je použita pasta Heraeus C 4731 a síto s hodnotou 400 mesh, pro polykarbonátový substrát pasty ESL 1901-S a Dupont CB 115 Ag nanášené přes 230 mesh.

Tabulka 2.2: První návrhy parametrů planárních induktorů

Motiv	D_{out}	D_{in}	Závitů	W [µm]	S [µm]	L [µH] (teor.)
1	40,16	20,60	33	150	150	51,02
2	40,40	20,90	20	250	250	19,01



Obrázek 2.2: Realizované vzorky planárních induktorů

3. TECHNOLOGICKÉ NÁMĚTY

- Navýšením tloušťky natisknuté pasty lze snížit parazitní sériový odpor. Toho lze dosáhnout použitím síta s větším průměrem tkaniny nebo vícenásobným tiskem na již nanesenou a vysušenou vrstvu pasty, případně ponorem vytvořené struktury do adekvátní cínové lázně.
- Vytvoření vícevrstevného induktoru umožňuje miniaturizaci či volbu tvaru výsledné struktury při zachování požadované indukčnosti. Toho se dá značně využít v aplikacích, kde plocha je limitujícím faktorem nebo je nutno přizpůsobit tvar struktury konkrétní aplikaci.
- Přidáním feritového materiálu k induktoru lze rovněž navýšit hodnotu indukčnosti. Samotný feritový materiál pokrytý izolační dielektrickou vrstvou by mohl sloužit jako substrát, na který lze tisknout pastu.

4. ZÁVĚR

Článek poukazuje na dosud plně neobjevenou a neprozkoumanou oblast, která otevírá nové perspektivy pro vytváření planárních induktorů tlustovrstvovou technologií, což může vést k zdokonalování dobíjecích cívek či realizaci planárních transformátorů. Možným směrem zkoumání je tisk motivů pastou s vysokou vodivostí a vysokou rozlišovací schopností, což umožňuje navyšovat hustotu závitů. Je nutno však technologickými náměty dosáhnout snížení parazitního odporu. Dalším směrem je zvětšení šířky vodivé cesty pro docílení minimální hodnoty parazitního odporu. Zároveň se nabízí možnost tvořit vícevrstevné struktury, kde lze předpokládat dosažení hodnot indukčnosti srovnatelných s předchozím uvedeným způsobem přístupu.

REFERENCE

[1] Zhao, Jonsenser. *A new calculation for designing multilayer planar spiral inductors*. EDN Europe; Sep2010, Vol. 57 Issue 9, p7

[2] A. Islam, S. Islam and F. Tulip, *Design and Optimization of Printed Circuit Board Inductors for Wireless Power Transfer System.* Circuits and Systems, Vol. 4 No. 2, 2013, pp. 237-244

[3] Ostfeld, A. E. et al. *Screen printed passive components for flexible power electronics*. Scientific Reports. 5, 15959; doi: 10.1038/srep15959. 2015.

[4] Mohan, S. S., del Mar Hershenson, M., Boyd, S. P. & Lee, T. H, *Simple Accurate Expressions for Planar Spiral Inductances*. IEEE journal of solid-state circuits, vol.34, no. 10, 1999

[5] Olivo, Carrara, De Micheli. *A Study of Multi-Layer Spiral Inductors for Remote Powering of Implantable Sensors*. IEEE Transactions on Biomedical Circuits and Systems, vol.7, 2013

CONNECTION ESTABLISHMENT BETWEEN SDON CONTROLLER AND OF AGENTS WITHIN OXC

Tomáš Kloda
Bachelor Degree Programme (3), FEEC BUT
E-mail: xkloda00@stud.feec.vutbr.cz

Supervised by: Lukáš Kočí
E-mail: xkocil00@stud.feec.vutbr.cz

Abstract: This paper deals with SDON (Software Defined Optical Network) within OXC (Optical Crossconnect). SDON represents a new direction in the control of optical network. This direction concerns the management of optical network elements from one place through control plane and data plane. Outcome of this paper are simulations implemented in OMNeT++. These simulations show synchronization and communication during initialization of SDON with occurence of delay between SDON controller and OF (OpenFlow) agents.

Keywords: OF agent, OF protocol, OMNeT++, OXC, SDON, controller, Software Defined Optical Network

1 ÚVOD

V současné době je rozvoj řízení optických sítí jeden z nejdůležitějších trendů komunikačních technologií. Mnoho podniků a koncových uživatelů klade čím dál více požadavků, na které již stávající řízení optických sítí přestává být vhodné. Tato situace vede k tomu, aby současné optické sítě byly modernizovány. SDON představují nový směr v řízení optických sítí, který je založen na bázi tří-vrstvého modelu tradičních softwarově definovaných sítí SDN (Software Defined Network). SDN je realizováno na bázi open source, stejně, jako SDON, a není proto jasně definována výsledná podoba. V současné době již existují simulace vyšších vrstev SDN, avšak pouze v elektrické doméně [1]. Kvůli nekompatibilitě SDN pro přenos dat v optické doméně, je zavedena implementace tzv. rozšířeného záhlaví. Toto záhlaví rozšiřuje SDN o další parametry fyzické vrstvy, čímž je umožněno nasazení SDN v rámci optické domény [2].

2 NAVÁZÁNÍ SPOJENÍ MEZI SDON KONTROLÉREM A OF AGENTY

Navázání spojení SDON kontroléru s OF agenty je řešeno na základě implementace OF protokolu do navržené SDON. Jedná se o standard, který umožňuje komunikaci mezi datovou rovinou, ve které se nacházejí OF agenti, a oddělenou řídící rovinou, ve které se nachází SDON kontrolér. Způsob navázání SDON kontroléru a OF agentů přes OF protokol probíhá na základě výměny OF zpráv [3].

2.1 OFPT_FEATURES_REQUEST/REPLY

Jako první inicializuje spojení SDON kontrolér, který vyšle zprávu *OFPT_HELLO* obsahující záhlaví. SDON kontrolér v něm nastaví svou nejvyšší podporovanou verzi OF protokolu, typ přenášené zprávy, celkovou délku zprávy a jedinečný identifikátor xid, který usnadňuje párování mezi SDON kontrolérem a OF agenty. OF agenti tuto zprávu přijmou a odpoví na ni taktéž zprávou *OFPT_HELLO*, čímž dojde k její výměně a k nastavení správné verze OF protokolu, na které bude spojení probíhat [4].

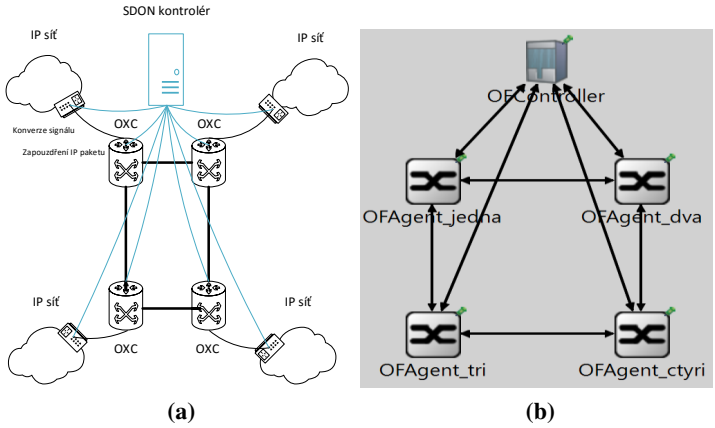
2.2 OFPT_FEATURES_REQUEST/REPLY

Po příjetí *OFPT_HELLO* zprávy, vyslanou OF agenty, si SDON kontrolér uloží její obsah do svých proměnných a následně vyšle zprávu *OFPT_FEATURES_REQUEST*. Tuto zprávu používá SDON kontrolér pro identifikaci a zjištění základních funkcí a schopností OF agentů. Po jejím obdržení vyšle OF agent asynchronní zprávu *OFPT_FEATURES_REPLY*, čímž odpovídá na požadavky zprávy *OFPT_FEATURES_REQUEST*. SDON kontrolér si obsah zprávy *OFPT_FEATURES_REPLY* opět uloží do svých proměnných [4].

2.3 OFPT_ECHO_REQUEST/REPLY

V dalším kroku je spojení mezi SDON kontrolérem a OF agenty udržováno na základě výměny zpráv *OFPT_ECHO_REQUEST* a *OFPT_FEATURES_REPLY*. Tyto zprávy slouží k výměně informací o zpoždění a šířce pásma [4].

Na obrázku 1 (a) se nachází navržená topologie SDON. Topologie je zaměřená na páteřní síť, která je složená ze čtyř OXC řízené SDON kontrolérem. Na každý OXC je napojena přístupová IP síť. IP sítě jsou v rámci topologie považovány za zdroje poskytující IP pakety pro přenos páteřní sítí do příslušné přístupové sítě. Přicházející paket z přístupové sítě je před vstupem do páteřní sítě zapouzdřen společně s parametry pro přenos skrz optickou síť [2]. Na obrázku 1 (b) se poté nachází propojení jednotlivých OF agentů s SDON kontrolérem v simulačním prostředí OMNeT++. Simulace je postavena na abstrakci fyzické vrstvy, kde OF agenti představují optické spoje OXC. Každý OF agent je s SDON kontrolérem spojen přes zpožďovací kanál.



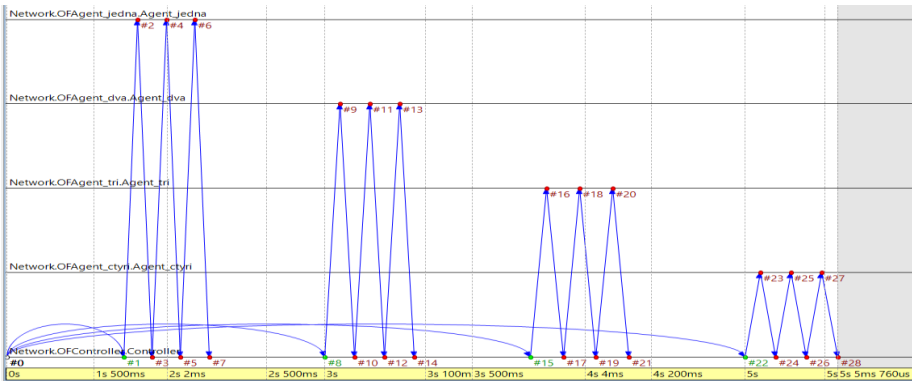
(a) **(b)**

Obrázek 1: Topologie SDON. a) Návrh topologie, b) Implementace topologie v OMNeT++.

3 VÝSLEDEK SIMULACE V OMNET++

Během vytváření topologie je počítáno s reálným zpožděním, které se vyskytuje již při navázání spojení OF agentů a SDON kontroléru. Zpoždění zahrnuje dobu zpracování příchozích OF zpráv SDON kontrolérem a OF agenty, avšak majoritní dobu zpoždění určuje vzdálenost OF agentů od SDON kontroléru. V tomto případě však OF v současné době neobsahuje žádný standardní mechanismus, který by byl schopen měřit zpoždění již při úvodním navázání spojení. V dostupných referencích se však vyskytuje několik lišících se experimentálních metod, které umožňují zpoždění měřit. V grafu na obrázku 2 lze vidět kompletní navázání spojení všech čtyř OF agentů s SDON kontrolérem, které proběhlo v maximálně stanoveném simulačním čase T = 5 s. Hodnoty zpoždění pro zpožďovací kanály jednotlivých OF agentů byly zvoleny na základě metody opakovaného měření pro maximálně 20 skoků (OF agentů) v rámci jednoho kanálu, kde prvnímu skoku odpovídá průměrný rozsah zpoždění od 0 ms do 2 ms [5]. Z uvedeného rozsahu má OFAgent_jedna definované zpoždění kanálu 0,65 ms,

OFAgent_dva 0,82 ms, OFAgent_tri 1,2 ms a OFAgent_ctyri 0,96 ms. Jelikož dojde v průběhu navázání spojení k výměně celkem šesti OF zpráv, je celková doba zpoždění navázání spojení pro jediného OF agenta 6 × X ms, kde X je definované zpoždění kanálu. V případě prvního OF agenta (OFAgent_jedna) je celková doba zpoždění v rámci navázání spojení 6 × 0,65 = 3,9 ms.



Obrázek 2: Výsledek simulace navázání spojení mezi SDON kontrolérem a OF agenty ze softwaru OMNeT++.

4 ZÁVĚR

Pomocí simulačního softwaru OMNeT++ bylo sestaveno spojení mezi SDON kontrolérem a OF agenty a odsimulován vznik zpoždění při navázání komunikace na základě reálných přenosových parametrů. Byla prokázána výměna inicializačních zpráv mezi SDON kontrolérem a jednotlivými OF agenty. Následně bylo prokázáno úspěšné navázání spojení všech OF agentů v rámci OXC.
Při realizaci bylo zahrnuto zpoždění, které vzniká majoritně vlivem rozdílné vzdálenosti mezi SDON kontrolérem a OF agenty. Provedená simulace tvoří předpoklad pro následnou datovou komunikaci v OXC. Předmětem dalšího výzkumu je rozšíření simulace o konfigurační zprávy a samotný datový přenos.

REFERENCE

[1] KLEIN, Dominik a Michael JARSCHEL. An OpenFlow extension for the OMNeT++ INET framework. In: *Proceedings of the 6th International ICST Conference on Simulation Tools and Techniques*. France: Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2013, s. 322-329. ISBN 978-1-4503-2464-9.

[2] CHANNEGOWDA, Mayur, Reza NEJABATI a Dimitra SIMEONIDOU. Software-Defined Optical Networks Technology and Infrastructure: Enabling Software-Defined Optical Network Operations [Invited]. In: *Journal of Optical Communications and Networking*. Bristol: Ori Gerstel and Patrick Iannone, 2013, s. 274-282. ISSN 1943-0620.

[3] HU, Fei. *Network innovation through OpenFlow and SDN: principles and design*. London: CRC Press, 2014. ISBN 978-1-4665-7210-2.

[4] OPEN NETWORKING FOUNDATION. *OpenFlow Switch Specification: Version 1.5.0 (Protocol version 0x06)* Dostupné z: https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-switch-v1.5.0.noipr.pdf

[5] PHEMIUS, Kevin a Mathieu BOUET. Monitoring latency with OpenFlow. In: *Proceedings of the 9th International Conference on Network and Service Management*. Zürich: International Federation for Information Processing, 2013, s. 122-125. ISBN 978-3-901882-53-1.

DESIGN OF OPTICAL WAVELENGTH SELECTIVE SWITCH

Vojtěch Sikora
Bachelor Degree Programme (3), FEEC BUT
E-mail: xsikor13@stud.feec.vutbr.cz

Supervised by: Petr Münster
E-mail: munster@feec.vutbr.cz

Abstract: Optical circuits switching is a rapidly growing area of photonic networks. With increasing requirements of photonic services only software defined networking (SDN) allows flexible and effective networks creation. The paper shortly introduces technologies of optical circuits switching, especially WSS (wavelength selective switching) for WDM (Wavelength Division Multiplexing) networks. Using specialized software VPIphotonics the WSS 1x4 for an interconnection between Prague, Brno, Ostrava and Vienna was designed.

Keywords: Optical network, Wavelength selective switch, Wavelength division multiplexing

1 ÚVOD

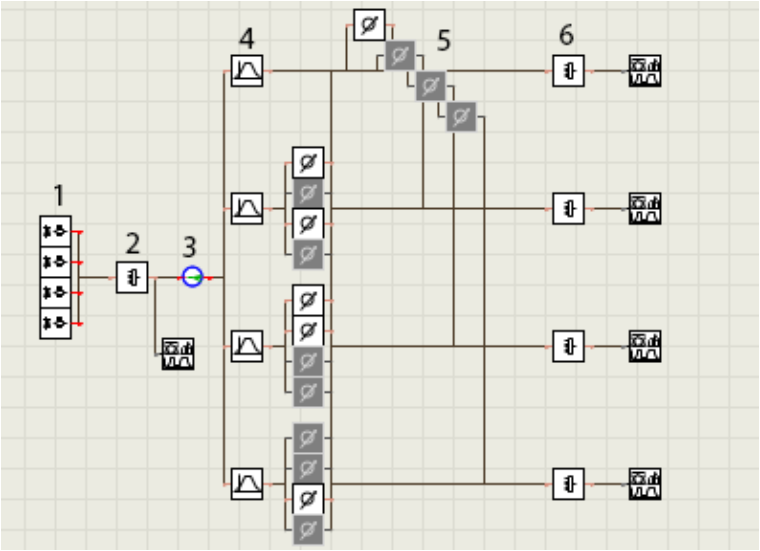
V posledních letech se v telekomunikačních sítích setkáváme s jevem exponenciálního nárůstu přenosu dat. V páteřních sítích jsou již jistou dobu standardem optické přenosy a v dnešní době se pracuje na technologiích, které dokáží zajistit větší pružnost těchto optických sítí. Jednou z cest tohoto výzkumu je nahrazování opticko-elektrických komponent, jako jsou směrovače a přepínače, za čistě optické zařízení. Vlnově selektivní přepínače jsou zařízení schopná měnit vlnové délky vysílaného WDM signálu v závislosti na stavu sítě, bez potřeby bezprostředního zásahu do zařízení obsluhou.

2 PŘEPÍNANÍ OPTICKÝCH SIGNÁLŮ

Přepínání optických okruhů na úrovni vlnových délek se v uzlech provádí pomocí optických multiplexorů, které umožňují z přenášeného signálu vyčlenit vlnovou délku, nebo ji do signálu přidat. Respektive, použitím optických přepínačů lze přepínat jednotlivé vstupní vlnové délky na kterýkoli výstupní port. V prvních generacích optických add-drop multiplexorů se musely signálové toky v uzlech a odbočkách do menších provozů přepojovat ručně. Obdobně, ručně, se signály musely připojovat do vstupních portů multiplexoru. Těmto manuálním zásahům bylo nutné se vyhnout z mnoha důvodů. Mezi hlavní patří zvyšování pružnosti sítě, možnost zavedení jisté automatizace a v neposlední řadě vzrůstající požadavky na celkové množství vlnových délek, které jsme schopni do vlákna multiplexovat. Tyto problémy efektivně řeší optický vlnově selektivní přepínač, používaný v konfigurovatelných optických add-drop multiplexerech.

3 VLNOVĚ SELEKTIVNÍ PŘEPÍNAČ

Vlnově selektivní přepínače se ve většině případů vyrábí v $1\times N$ formátu. Což znamená, že na vstup přichází WDM signál s určitým počtem kanálů (vlnových délek), které lze vydělit do N směrů. Způsobů přepínaní je mnoho. Hojně využívanou technologií je použití MEMS neboli mikroelektromechanických zrcadel. Jak už z názvu vyplývá, jedná se o mřížku, na které jsou umístěny mikrozrcadla, schopná natáčení a odrážení demultiplexovaných vlnových délek do různých výstupních portů. Neposlední v řadě jsou LCoS, použití tekutých krystalů. Podobně jako v LCD panelu pomocí natáčení tekutého krystalu se propouští vybrané vlnové délky, ostatní se blokují.



Obrázek 1: Vnitřní zapojení navrženého vlnově selektivního přepínače

3.1 Návrh vlnově selektivního přepínače

Můj vytvořený návrh je demultiplex-switch-multiplex typu 1×4. V návrhu se používá signál WDM se 4 vlnovými délkami s odstupem 100 GHz. Přesněji, vlnové délky v rozmezí od 193 THz do 193,3 THz. Na obrázku 1 je vnitřní zapojení přepínače. Jelikož použitý program VPIphotonics nezobrazuje vlnovody a uzly dostatečně viditelně, je první větev rozložena a ostatní jsou logicky zapojené podobně jako větev první. Všechny komponenty jsou ideální, tedy pokud nejsou definovány jinak, jejich útlum je 0 dB.

3.2 Popis funkce

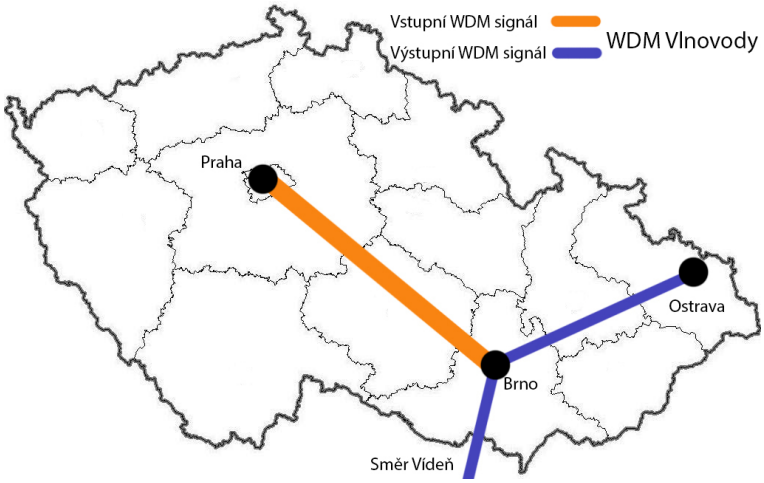
Budeme-li se pohybovat na obrázku 1 zleva, vidíme čtyři generátory výše zmíněných vlnových délek (1), které jsou pomocí multiplexoru (2) namultiplexovány do jednoho vlákna jako WDM signál. Následuje rozdělení WDM signálu (3) na 4 stejné signály totožné se signálem za multiplexorem (2). Na každém z optických filtrů (4) je ze signálu vydělena jedna vlnová délka, která je pomocí útlumových členů (5), které mají buď útlum 0 dB - propouští, nebo ∞ dB - blokují, podle potřeby odeslána na výchozí multiplexory (6). Za nimi již máme výstupní WDM signály.

4 Návrh použití

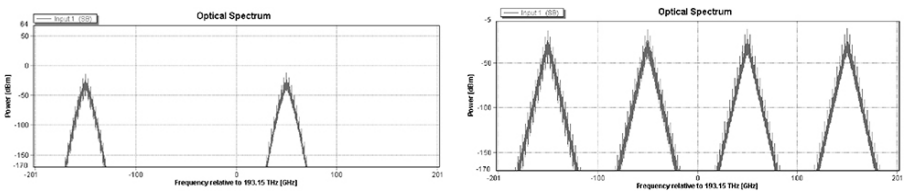
Pro ukázku použití přepínače jsem sestavil fiktivní síť mezi městy Praha, Brno, Ostrava a Vídeň. Vlnově selektivní přepínač je umístěn v Brně. Do vstupu přepínače je připojen vlnovod z Prahy, na výstupy jsou připojeny města Ostrava a Vídeň. To lze vidět na obrázku 2. Přepínač v Brně tedy může dle požadavků provozu sestavovat okruhy mezi městy na libovolných vlnových délkách. V případě, že by požadovaná vlnová délka již byla na zvolené trase zabraná jiným provozem, odešle přepínač informace na jiné volné vlnové délce. Na obrázku 3 vidíme sestavenou trasu mezi městy Praha - Brno - Ostrava. Přepínač v Brně vydělí 2 vlnové délky, které pokračují v cestě dál. Pokud by přišel jiný požadavek, lze vysílat na jiném kanále. To vše je možné ovládat vzdáleným přístupem, nemusí tedy nikdo bezprostředně u přepínače konfigurovat trasy manuálně.

Pokud bychom zapojili podobně další dva přepínače se vstupy ze zbylých měst a společně provázali, byl by již vytvořen jednoduchý model ROADM - rekonfigurovatelného add-drop multiplexoru. Má

dalších nespočetně mnoho výhod, mezi kterými jsou například vytvoření expresního průchodu uzlem, nebo vydělování a navazování vlnových délek pro místní provoz.



Obrázek 2: Fiktivní síť Praha - Brno - Ostrava - Vídeň



Obrázek 3: Ukázka výstupu a vstupu WSS, směr Praha - Brno - Ostrava

5 ZÁVĚR

Tato práce ve výsledku utváří čtenáři ucelený pohled na problematiku spínání optických okruhů pomocí vlnově selektivních přepínačů a jejich implementaci v optické sítí využívající WDM signály. Do budoucna je výzkum vlnově selektivních přepínačů kritickou záležitostí. Se vzrůstajícím přenosem dat je velmi neekonomické optické signály v přepínačích měnit na elektrické a opět odesílat v optické podobě, navíc se mohou při vyšších zátěžích vytvářet z těchto míst úzká hrdla zvyšující zpoždění celé sítě. Při shrnutí těchto poznatků dojdeme k výsledku, že implementace těchto zařízení je do budoucna obrovským přínosem jak ekonomickým, tak v oblasti rychlosti. Tyto dva parametry vždy byly, jsou a budou hlavním ukazatelem při tvorbě jakékoli sítě.

REFERENCE

[1] Filka, M.: Přenosová média 2012, ISBN 978-80-214-4444-7

[2] Kuchař, A. Přepínání a směrování optických signálů [online]. 2013, [cit. 10. 3. 2016]. Dostupné z URL: <http://www.action-m.com/ok2012/files/prednasky/kuchar-clanek.pdf>.

[3] Nagy, M. Wavelength selective switches for fiber optic telecommunications [online]. 2015, [cit. 10. 3. 2016]. Dostupné z URL: <http://www.photonics.com/Article.aspx?AID=27188>.

FIBER OPTIC MICHELSON INTERFEROMETER

David Dorazin

Bachelor Degree Programme (3), FEEC BUT

E-mail: xdoraz05@stud.feec.vutbr.cz

Supervised by: Milan Čučka

E-mail: xcucka00@stud.feec.vutbr.cz

Abstract: The aim of this paper is to theoretically desribe function of the Optic fiber Michelson interferometer and its parts with description of the most important issues which can occure due to bad design. Second part of this paper is dedicated to simulation of the optic fiber Michelson interefometer in professional simulation software VPIphotonics.

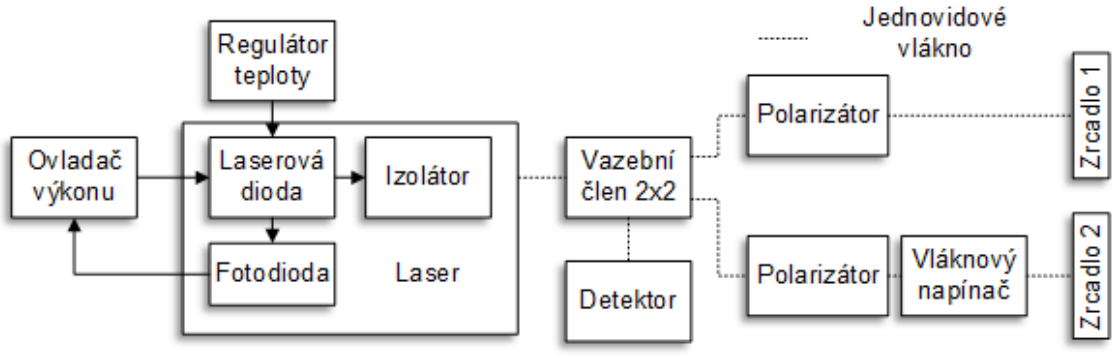
Keywords: Michelson interferometer, VPI photonics, distributed sensors, interference

1. ÚVOD

Distribuované optické vláknové senzory se vyvíjí od doby, kdy bylo vynalezeno samotné optické vlákno a v dnešní době se tyto senzory prosazují ve všech odvětvích měřící, řídící a automatizační techniky. Svými přednostmi ukazují v mnohých oblastech převahu nad konvenčními elektronickými senzory. Jedním z těchto senzorů je Michelsonův vláknový interferometr, který ke své činnosti využívá interferenci světla v optickém vlákně. Cílem této práce je teoretický rozbor a simulace funkce Michelsonova vláknového interferometru v programu VPIphotonics.

2. MICHELSONŮV VLÁKNOVÝ INTERFEROMETR

Interferometry jsou nejcitlivější vláknové senzory, které vyhodnocují změnu fáze v optickém vlákně. V našem případě MI pracuje na principu rozdělení světelného paprsku do dvou větví, kdy je jedna větev měřící a druhá referenční. Následně se vyhodnocují fázové rozdíly těchto dvou optických tras na výstupu vazebního členu. Světelné paprsky jsou vedeny jednovidovým optickým vláknem, kde jsou na konci obou větví odraženy reflexní plochou a interferují v optickém vazebním členu tzv. coupleru, z kterého je signál veden do detektoru [1].



Obrázek 1: Blokové schéma Michelsonova vláknového interferometru [1].

2.1. POPIS ZAŘÍZENÍ

Jako zdroj světelného záření se používá DFB laser, který je volen na telekomunikačních vlnových délkách a to 1310 nm nebo 1550 nm z důvodu kompatibility s optickým vláknem, což ovšem není podmínkou funkčnosti, může být použit i jiný laser s dostatečnou kvalitou výstupního paprsku.

DFB lasery mají relativně úzké spektrum emitované vlnové délky, dostatečnou koherentní délku, ale jelikož jsou stabilní lehce pod pokojovou teplotou, tak je zde použit regulátor teploty. Laser je opatřen optickým izolátorem, který slouží jako prevence před zpětným odrazem, čímž předchází nestabilitě a degradaci zdroje záření [2].

Optický vazební člen je zařízení užívané v optických systémech, sloužící k rozdělení vstupního optického signálu mezi více výstupů. V našem případě je použit vazební člen 2x2 s dělícím poměrem 50/50, což znamená, že je vstupní signál rozdělen do dvou větví s poloviční výkonem, protože v obou větvích potřebujeme navodit stejné podmínky. Druhý vstupní port slouží jako výstup interferovaného signálu [1].

Zde však nastává zásadní problém, když je laserový paprsek rozdělen a znovu spojen ve vazebním členu, souběžně putující signály optickým vláknem už nejsou schopné interferovat, protože odražený paprsek není schopen udržet původní stav polarizace. Jedním řešením je použití PM (polarization-maintaining) vlákna, které udržuje stejnou polarizaci v celém systému, ale z důvodu vysoké ceny je toto řešení často neatraktivní. Druhým a přijatelnějším řešením je použití Faradayova rotačního zrcadla, odrážející paprsek na obou koncích interferometru, které slouží jako fázové časované zrcadlo, které vytváří zpoždění 90°. Polarizátory se nepoužívají kvůli ztrátovosti signálu [1].

Napínač vlákna, může být použit ke kalibraci zařízení, využívá faktu, že křemíkové optické vlákno je elastické a může být nataženo až o 2% své délky [1]. Vláknové odrazové plochy na konci ramen interferometru a vazební člen by měli být kompatibilní s použitým laserem. Teoreticky jsou navrženy tak, že by měli pracovat na určité vlnové délce, ale v praxi, jsou levná vláknová zrcadla vyráběna z dialektrických materiálů a jsou širokopásmové. Například zrcadlo určené pro 1310 nm bude pravděpodobně jevit vysokou odrazivost i na vlnové délce 1550 nm [2].

Interferovaný signál může být detekován řadou elektronických součástek jako fotodioda, fotonásobič, fotovodivý detektor, CCD senzor nebo pyroelektrický detektor, v závislosti na vlastnostech měřeného signálu a použitém zdroji světelného záření. Fotodiody jsou nejpopulárnější a universální detektory, které pokrývají široký rozsah vlnových délek. Nicméně, propracované detekční zařízení předpokládají integraci detektoru v obvodu na zpracování signálu zahrnující jeho zesílení, Fourierovu transformaci signálu přijatelnou pro monitorování frekvenčních změn a k tomu převedení měřeného signálu pro přenos (vzorkování, kvantování, kódování), takto zpracovaný signál je dále odeslán do PC, kde je možno naměřené výsledky zpracovat podle druhu měření v matematickém prostředí jako je například Matlab [1].

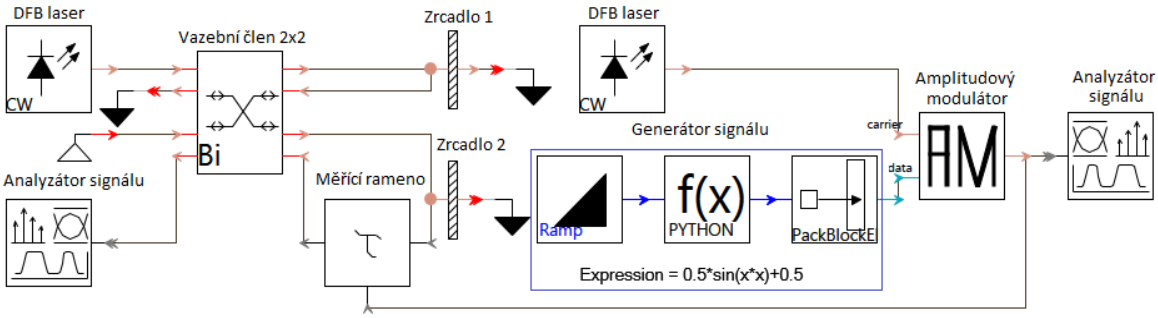
Michelsonův vláknový interferometr není schopen měřit absolutní hodnoty jako teplotu nebo vzdálenost, ale je velmi přesný při měření malých změn těchto veličin. Dále může být použit k měření vibrací [2].

3. SIMULACE

Simulace je prováděna v programu VPIphotonics, což je profesionální simulační software podporující aktivní či pasivní integrované fotonické obvody, aplikace optických vláken, optické přenosové systémy, síťové aplikace a jejich následnou analýzu.

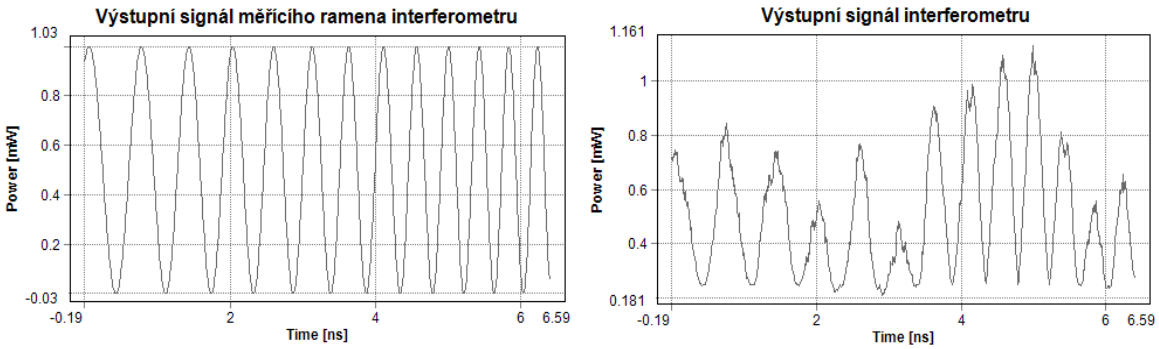
Jako zdroj světelného záření je použit DFB laser, který vysílá signál konstantní amplitudy a frekvence do obousměrného vazebního členu 2x2, který je nastaven na dělící poměr 0,5, což znamená, že vstupní signál je zde rozdělen do dvou ramen, kde v obou ramenech má poloviční výkon. Horní rameno, se zrcadlem 1 je rameno referenční, kde se rozdělený signál odrazí a putuje zpět do vazebního členu. Spodní rameno je měřící, sem signál putuje, odrazí se od zrcadla 2 a je ovlivňován měřenou veličinou, v tomto případě je simulováno působení vibrací nebo tahové síly sinusového průběhu na ideální optické vlákno. Toho je v simulačním programu docíleno tak, že je generován sinusový signál narůstající frekvence podle zadaného výrazu v generátoru složeného z tří bloků, kdy 1. blok inkrementuje hodnotu x, 2. blok počítá funkci sin a 3. blok převádí matematický signál na signál elektrický, který je pomocí amplitudového modulátoru převeden na optický signál. Takto převedený signál je přiveden do měřícího ramene interferometru a odtud zpět do vazebního členu, kde

signály z obou ramen interferují, a výsledný interferovaný signál je zobrazen pomocí analyzátoru signálu, který slouží jako detektor i zobrazení průběhu. Toto složité řešení je použito z toho důvodu, že Michelsonův vláknový interferometr je schopný měřit pouze změny měřené veličiny nikoliv hodnoty absolutní a tak je uměle vytvořen signál, který by vznikl působením vibrací či tahové síly.



Obrázek 2: Zapojení Michelsonova vláknového interferometru v simulačním prostředí.

Jelikož se jedná o simulaci a je použito ideálních prvků a ideálního přenosového média, nemusíme se zabývat problémy jako udržení stejné polarizace světla v ramenech, koherentní délky laseru nebo ztrát součástek nebo vedení. Trojúhelníkové značky slouží jako ošetření nepoužitých vstupů a výstupů.



Obrázek 3: Simulované průběhy signálů.

Na obrázku 3 vidíme na prvním grafu výstupní signál z měřícího ramena, který simuluje už dříve popsané vibrace či tahovou silu, které zvyšují svou frekvenci. Na druhém grafu je zobrazen výstupní interferovaný signál z obou ramen, kde jsou vidět maxima a minima, které odpovídají signálu z měřícího ramena.

4. ZÁVĚR

Práce se věnovala Michelsonovu vláknovému interferometru, který patří mezi distribuované optické senzory. Teoreticky jsem zde rozebral princip tohoto zařízení a probral součásti, z kterých je složeno a jejich důležité vlastnosti. Dále jsem interferometr navrhl a simuloval v programu VPIphotonics. Důležité však je, ověřit si dosažené výsledky simulace měřením na reálném zapojení.

REFERENCE

[1] Lucki, M., Bohac, L., Zeleny, R.: Fiber Optic and Free Space Michelson Interferometer – Principle and Practise, Optical Sensors – New Development and Practical applications, In-Tech, 2014, ISBN 978-953-51-1233-4

[2] Marret, D.: Construction of an All-Fiber-Optic Michelson Interferometer, 2012, dostupné z URL: http://www.conspiracyoflight.com/FO_MM/FOMM.html

NORMALIZED INTER-CLASS VARIANCE V PROUDOVÉ ANALÝZE

Vaněk Stanislav
Bachelor Degree Programme (3), FEEC BUT
E-mail: xvanek35@stud.feec.vutbr.cz

Supervised by: Martinásek Zdeněk
E-mail: martinasek@feec.vutbr.cz

Abstract: The internet of things leads to a massive interest in security of embedded devices. Nowadays, power analysis poses extremely effective and successful types of attacks to break confidential cryptographic algorithms such as AES (Advanced Encryption Standard), RSA (Rivest Shamir Adleman) and cryptographic devices such as smart cards. This paper deals with the method of localization of interesting markers in power analysis called NICV (Normalized Inter-Class Variance). The aim of the paper is to describe this method and its implementation.

Keywords: NICV, DPA, CPA, variance

1 ÚVOD

Masivní nasazení vestavěných systémů vede k velkému zájmu o jejich zabezpečení. Proudová analýza měří a zkoumá proudovou spotřebu kryptografického zařízení s cílem odhalit korelaci proudové spotřeby a senzitivní informace, kterou představuje většinou šifrovací klíč. Proudová analýza představuje v dnešní době efektivní a velice účinný způsob útoku na doposud bezpečné kryptografické primitiva jako AES, RSA a kryptografická zařízení jako čipové karty, jednočipové počítače atd.

Při proudové analýze představuje nejkritičtější část detekce zajímavých bodů, kterými je vyzařována požadovaná senzitivní informace. Je zřejmé, že detekce zajímavých bodů je důležitá také pro ochranu proti proudové analýze, protože dokážeme detekovat únik informace a následně jej eliminovat.

2 METODY DETEKCE ZAJÍMAVÝCH BODŮ

Pro detekci zajímavých bodů v proudové analýze se používají různé metody, nejznámější je DPA (Differnetial Power Analysis) založená na korelačním koeficientu. Dalšími metodami, které lze použít jsou např. metoda založená na rozdílu sumy čtverců (SOSD), nebo metoda založená rozdílu sumy čtverců s T-testem (SOST). Metoda NICV (Normalized Inter-Class Variance) není běžně implementována v prostředích zabývající se proudovou analýzou. Mým cílem bylo tuto metodu implementovat v prostředí Matlab a následně ověřit funkčnost na veřejně dostupných proudových průbězích ze soutěže DPA Contest.

3 NICV

NICV se dá volně přeložit jako analýza rozptylu. Oproti jiným metodám má útočník výhodu v tom, že nepotřebuje profilující zařízení, tedy pracuje přímo s průběhy, na které je aplikována proudová analýza (útok). Další výhodou je, že veškeré parametry potřebné k útoku jsou veřejné, a tudíž útočníkovi známé. Naopak nevýhodou této metody je, že je oproti jiným metodám méně přesná.

3.1 MATEMATICKÉ VYJÁDŘENÍ

Označme si jeden byte otevřeného nebo zašifrovaného textu X a proudovou spotřebu měřenou útočníkem Y. Pak vztah pro výpočet analýzy rozptylu mezi těmito veličinami bude následující:

$$NICV = \frac{(Var[E[Y|X|]])}{(Var[Y])}, \tag{1}$$

kde Var značí rozptyl a E průměr.

3.2 IMPLEMENTACE

Praktická realizace této metody byla provedena ve vývojovém prostředí MATLAB. Naměřené proudové průběhy a nástroje byly získány ze soutěže DPA Contest. Cílem této soutěže je poskytnout objektivní způsob porovnání metod proudové analýzy, především DPA útoku veřejnosti. Použité kryptografické zařízení je programovatelná čipová karta ATmega, která má v sobě naimplementována algoritmus AES-256 a k maskování tohoto algoritmu slouží metoda RSM (Rotating Sbox Masking). Implementovaný algoritmus je zobrazen na obr.1. Algoritmus funguje následovně, na začátku je vygenerován offset, který musí být udržen v tajnosti (hodnota 0 až 15). Podle této hodnoty je rotována sada šestnácti konstantních masek, která je veřejně známá. Následuje rotace řádků a poté mixování sloupců bytu otevřeného textu. Jako poslední je provedena korekce masky.

Implementace AES použitá pro DPAContestv4

Input : 16-bytes Plaintext X $[\![X_0, X_1 \cdots X_{15}]\!]$, // 16 bajtů otevřeného textu
Key, 15 16-bytes constants RoundKey$[r]$, $r \in [\![0, 14]\!]$ // Rozšiřování klíče
Output: 16-bytes Ciphertext X $[\![X_0, X_1 \cdots X_{15}]\!]$ // Šifrovaný text - 16 bajtů

Draw a random offset, uniformly in $[\![0, 15]\!]$ // Náhodně generovaný offset s náhodným rozdělením
$X = X \oplus \text{Mask}_{\text{offset}}$ // Maskování otevřeného textu

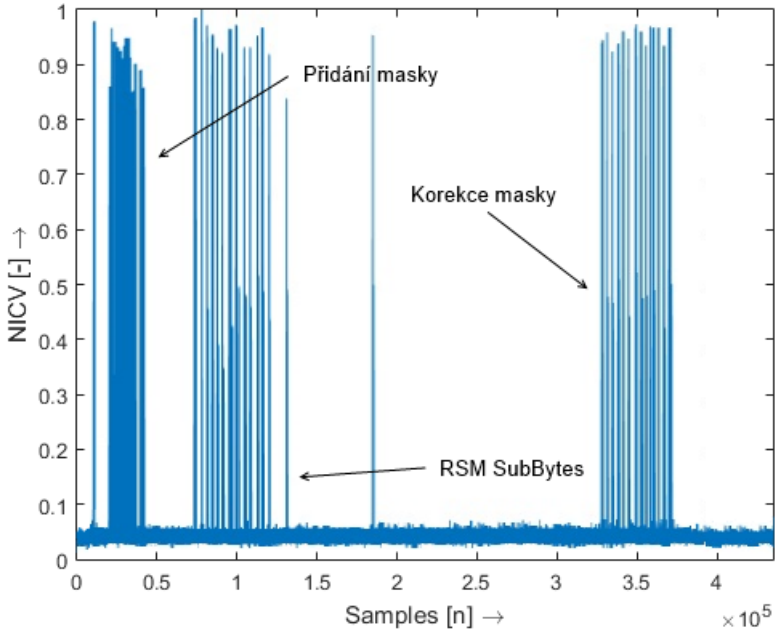
// Všechny rundy kromě poslední
for $r \in [\![0, 12]\!]$ **do**
  $X = X \oplus \text{RoundKey}[r]$
  for $i \in [\![0, 15]\!]$ **do** // Add round key
    $X_i = \text{MaskedSubBytes}_{\text{offset}+i+r}(X_i)$
  end
  $X = \text{ShiftRows}(X)$ // Rotace řádků
  $X = \text{MixColumns}(X)$ // Mixování sloupců
  $X = X \oplus \text{MaskCompensation}_{\text{offset}+1+r}$ // Korekce masky
end

// Poslední runda
$X = X \oplus \text{RoundKey}[13]$
for $i \in [\![0, 15]\!]$ **do**
  $X_i = \text{MaskedSubBytes}_{\text{offset}+13+r}(X_i)$
end
$X = \text{ShiftRows}(X)$
$X = X \oplus \text{RoundKey}[14]$

// Unmasking zašifrovaného textu
$X = X \oplus \text{MaskCompensationLastRound}_{\text{offset}+14}$

Obrázek 1: Algoritmus AES-256

Nejprve je třeba se rozhodnout, na co přesně chce útočník útočit. V tomto případě je to byte otevřeného textu po vykonání první rundy algoritmu AES, ze kterého by útočník mohl zachytit citlivé informace. Nejprve je pomocí NICV vypočítán rozptyl hypotézy proudové spotřeby. Následně je spočítán rozptyl pro měřenou proudovou spotřebu. Poté je provedena korelační proudová analýza (CPA). Výsledkem je matice koeficientů NICV, což je matice obálek korelačních koeficientů. Čím více se hodnota v matici koeficientů NICV blíží 1, tím větší je závislost mezi hypotetickou proudovou spotřebou a změřenou proudovou spotřebou. Naopak hodnoty blížící se k 0 značí velmi nízkou, téměř nulovou závislost.



Obrázek 2: NICV spočítané pro byte otevřeného textu

Na obr. 2 je vidět 48 bodů. Tyto body odpovídají bělení textu, přidání klíče, funkci SubBytes a korekci masky. Při úniku těchto dat ze zařízení může být jeden složený útok, na tento únik, dostačující pro obnovu klíče. Jestliže je známá sekvence masky, útočníkovi pak stačí na základě masky odhalit klíč a ten nahrát do svého zařízení (přístupové karty) a pomocí této karty se může vydávat za pravého majitele.

4 ZÁVĚR

Metoda analýzy rozptylu NICV je jedna z metod lokalizace zajímavých bodů u proudové analýzy. Nicméně jen samotné NICV k nalezení citlivých informací nestačí. Je nutné ještě provést korelační proudovou analýzu (CPA), díky které už lze vyčíst z průběhů citlivé informace. NICV funguje tedy jako tzv. předpříprava průběhů či zrychlení analýzy pro další zpracování útočníkem. Velmi často je analýza rozptylu používána právě spolu s CPA pro zrychlení analýzy.

REFERENCE

[1] MORADI, Amir, Sylvain GUILLEY a Annelie HEUSER. *Detecting Hidden Leakages* [online]. [cit. 2016-03-7]. Dostupné z: https://eprint.iacr.org/2013/842.pdf

[2] BHASIN, Shivam, Nicolas BRUNEAU, Jean-Luc DANGER a Zakaria NAJM. *Analysis and Improvements of the DPA Contest v4 Implementation*.

[3] BHASIN, Shivam, Jean-Luc DANGER a Zakaria NAJM. *NICV: Normalized Inter-Class Variance for Detection of Side-Channel Leakage** [online]. , 17 [cit. 2016-03-7]. Dostupné z: https://eprint.iacr.org/2013/717.pdf

Magisterské projekty

Elektronika a komunikace

ELECTRONICALLY RECONFIGURABLE FREQUENCY FILTERS

Adam Gajdos

Master Degree Programme (2), FEEC BUT

E-mail: xgajdo12@stud.feec.vutbr.cz

Supervised by: Jan Jerabek

E-mail: jerabekj@feec.vutbr.cz

Abstract: This article presents designed reconnection-less filter with electronically reconfigurable type and parameters of its transfer function. The Filter uses VDCC (Voltage Differencing Current Conveyor) and CA (Current Amplifier) as active elements. This paper includes also results of computer simulations of designed circuit.

Keywords: current amplifier, reconfigurable frequency filters, Voltage Differencing Current Conveyor

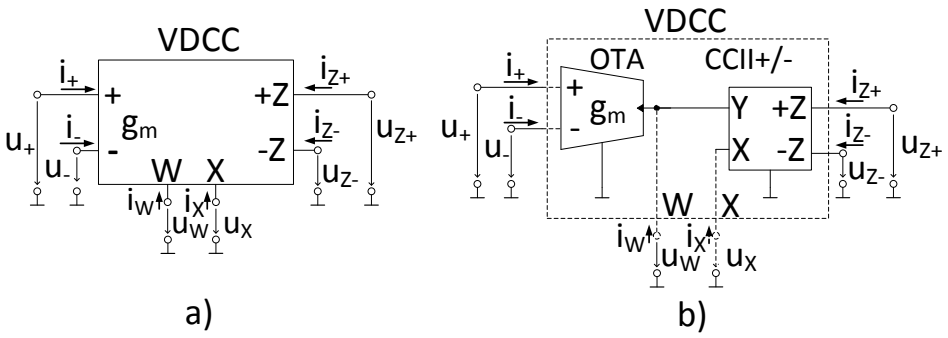
1. ÚVOD

Kmitočtové filtry se využívají zejména k odstranění, potlačení či zesílené některých složek spektra signálu [1]. Tyto filtry mají většinou jeden vstup a více výstupů [2], na kterých se nachází různé přenosové funkce nebo naopak mají více vstupů a jeden výstup [3], kdy při buzení do jednotlivých vstupů dostáváme různé přenosové funkce. V těchto případech je bohužel nutné vždy přepojit výstup resp. vstup, abychom mohli změnit přenosovou funkci obvodu. To ovšem není vždy možné nebo výhodné, například při implementaci na čipu.

Tato práce popisuje využití parametrů aktivních prvků, z nichž se skládá samotný filtrační obvod, pro změnu přenosové funkce bez nutnosti jakéhokoliv přepojení [4]. Obvod tedy bude mít pouze jeden vstup a výstup. Filtrační obvod umožňuje mimo změny přenosové funkce, také změnu mezního kmitočtu a činitele jakosti nezávisle na sobě.

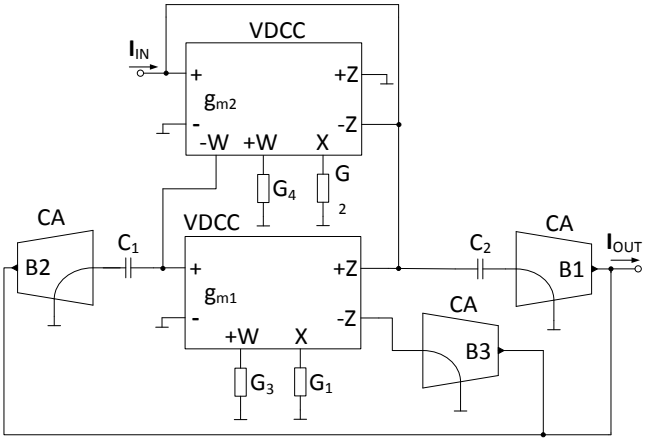
2. NÁVRH OBVODU

Navržený obvod využívá tři proudové zesilovače CA a 2 aktivní prvky VDCC. Řiditelným parametrem proudového zesilovače je zesílení *B*, kterým je vynásoben vstupní proud. Prvek VDCC je složen s operačního transkonduktančního zesilovače (OTA), ten může mít i více výstupů (BOTA, MOTA) a proudového konvejoru druhé generace (CCII). Řiditelným parametrem VDCC je tedy transkonduktance g_m. Vnitřní struktura VDCC je vidět na Obr. 1.



Obrázek 1: a) Schematická značka VDCC b) Vnitřní struktura VDCC

Schéma zapojení obvodu je na Obr. 2. Obvod realizuje přenosové funkce typu iDP, PP, iHP, iPZ, iFČ pouhou změnou zesílení *B* u jednotlivých proudových zesilovačů. Za pomocí proudových zesilovačů lze také realizovat dolní nebo horní propust s nastavením přenosu v nepropustného pásma.



Obrázek 2: Schéma zapojení filtrační struktury

Přenosová funkce obvodu z Obr. 2 je

$$K(p) = \frac{\boldsymbol{I}_{OUT}}{\boldsymbol{I}_{IN}} = \frac{-\boldsymbol{p}^2 B_1 C_1 C_2 G_3 G_4 + \boldsymbol{p} B_3 C_1 G_3 G_4 g_{m2} - B_2 G_1 G_4 g_{m1} g_{m2}}{\boldsymbol{p}^2 C_1 C_2 G_3 G_4 + \boldsymbol{p} C_1 G_2 G_3 g_{m2} + G_1 G_4 g_{m1} g_{m2}}. \quad (1)$$

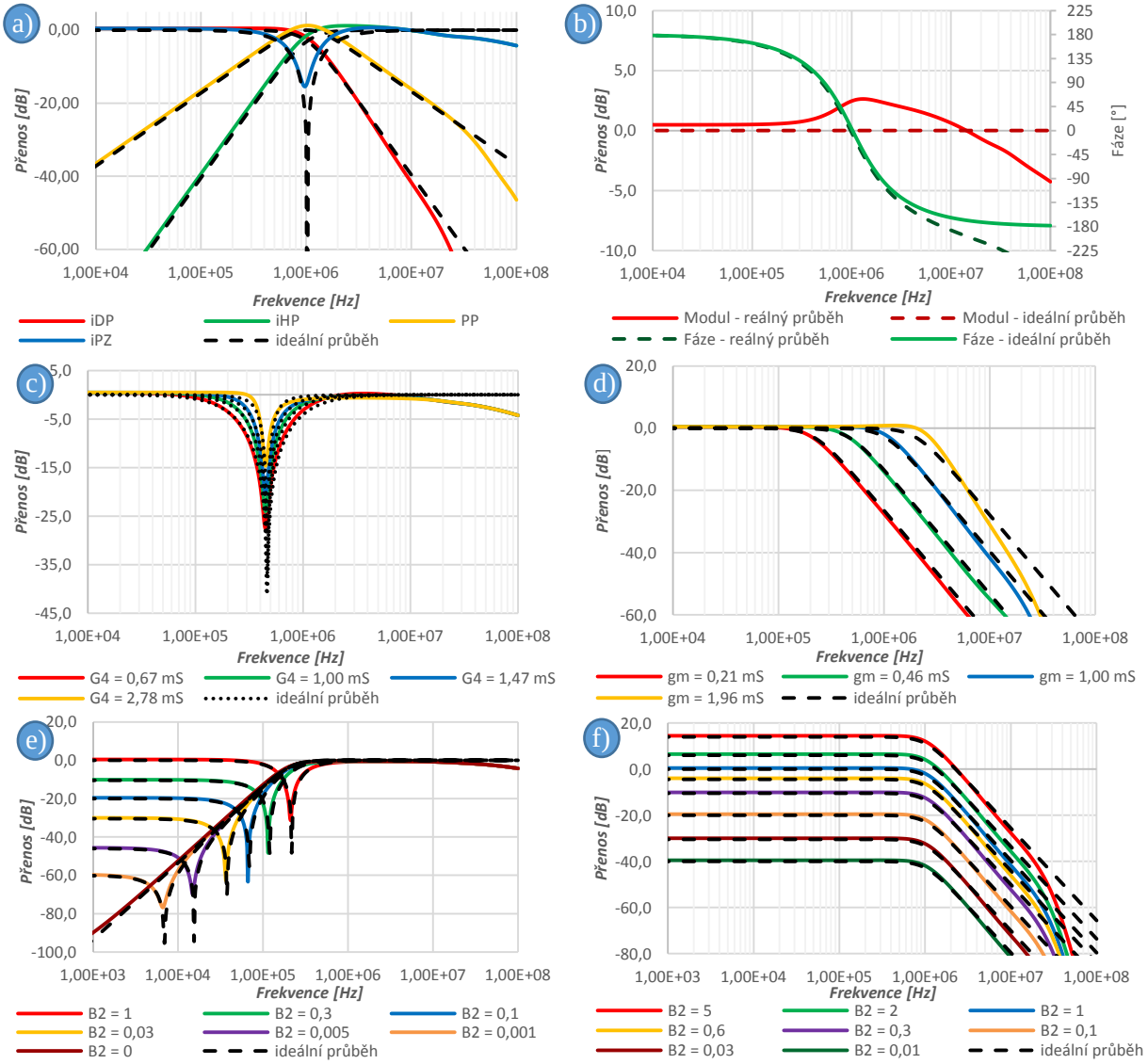
Jednotlivé přenosové funkce filtru jsou dány řízením zesílení B_1 až B_3 takto:

iHP - když $B_1 > 0$ a zároveň $B_2 = 0$, $B_3 = 0$. Zesílení přenosu v propustném pásmu je řiditelné za pomoci B_1; **PP** - když $B_3 > 0$ a zároveň $B_1 = 0$, $B_2 = 0$. Zesílení propustného pásma je řiditelné za pomoci B_3; **iDP** - když $B_2 > 0$ a zároveň $B_1 = 0$, $B_3 = 0$. Zesílení propustného pásma je řiditelné za pomoci B_2; **iPZ** - když $B_1 = B_2 > 0$ a zároveň $B_3 = 0$. Zesílení propustného pásma je řiditelné za pomoci $B_1 = B_2$; **iFČ** - když $B_1 = B_2 = B_3 > 0$. Zesílení je řiditelné za pomoci $B_1 = B_2 = B_3$; **iHP s nulou v nepropustném pásmu** - když $B_1 > B_2 > 0$ a zároveň $B_3 = 0$. Zesílení propustného pásma je řiditelné za pomoci B_1 a zesílení nepropustného pásma je dáno B_2; **iDP s nulou v nepropustném pásmu** - když $B_2 > B_1 > 0$ a zároveň $B_3 = 0$. Zesílení propustného pásma je řiditelné za pomoci B_2 a zesílení nepropustného pásma je dáno B_1.

Z koeficientů jmenovatele přenosové funkce jsou odvozeny vztahy pro úhlový kmitočet $\omega_0^2 = G_1 g_m^2 / C_1 C_2 G_3$ (kdy g_{m1}a g_{m2} musí být v rovnosti, aby byla možná změna mezního kmitočtu a zároveň nedocházelo ke změně činitele jakosti) a činitel jakosti $Q^2 = C_2 G_1 G_4^2 / C_1 G_2^2 G_3$. Z rovnic je jasné, že lze nezávisle na sobě řídit mezní kmitočet $f_m = \omega_0 / 2\pi$ a činitel jakosti Q, s vyjímkou PP, kdy změna činitele jakosti způsobí i změnu zesílení, což je vlastnost obvodu, která není žádoucí. Činitel jakosti lze ladit změnou vodivosti G_2nebo G_4. Z rovnic jsou vyjádřeny vztahy pro výpočet kapacit C_1 a C_2, kdy hodnoty zbylých prvků jsou voleny následovně $f_m = 1MHz$, $Q = 0{,}707$, vodivosti $G_1 = G_3 = 0{,}4mS$, $G_2 = G_4 = g_m = 1mS$. Hodnoty vypočtených kapacit jsou $C_1 = 225pF$ a $C_2 = 113pF$. Pro simulace byly využity modely universálního proudového konvejoru UCC jako stavební prvek VDCC a obvod EL2082 reprezentující proudový zesilovač CA. Výsledky simulací jsou na Obr. 3.

3. ZÁVĚR

Tento článek popisuje jedno ze zapojení filtračních obvodu mé diplomové práce. Jak je vidět, z výsledků simulací, obvod se ve většině případů blíží ideálním průběhům a má tedy i potenciál k realizaci a skutečnému měření. Velkou výhodou tohoto obvodu oproti klasickým filtrům je plynulý přechod mezi přenosovými funkcemi a možnost regulace zesílení v propustném i nepropustném pásmu.



Obrázek 3: Výsledky simulací navrženého obvodu z programu OrCAD PSpice: a) Výstupy simulací přenosových funkcí obvodu b) Výsledek simulace přenosové funkce iFČ c) Ladění činitele jakosti u přenosové funkce iPZ ($f_m = 450\ kHz$ z důvodu nižšího útlumu v nepropustném pásmu na vyšších kmitočtech) d) Ladění mezního kmitočtu u přenosové funkce iDP e) Ladění nuly nepropustného pásma u iHP ($f_m = 218\ kHz$) f) Ladění zesílení propustného pásma iDP ($f_m = 1{,}02\ MHz$)

REFERENCE

[1] DOSTÁL, T. Elektrické filtry. 1.vyd. Brno: FEI VUT, 2001. s. 96, ISBN 80-214-0877-4.

[2] JEŘÁBEK, J., VRBA, K. SIMO type low-input and high-output impedance current mode universal filter employing three universal current conveyors. AEU - International Journal of Electronics and Communications, roč. 64, 6. vyd., 2010, s. 588-593.

[3] MINAEI, S., YUCE, E., CICEKOGLU, O. Electronically tunable multi-input single-output voltage-mode filter. Proceedings of the 2005 European Conference on Circuit Theory and Design, roč. 3, 2005, s. 401-404, ISBN: 0-7803-9066-0.

[4] ŠOTNER, R., PETRŽELA, J., JEŘÁBEK, J., DOSTÁL, T. Reconnection-less OTA- based Biquad Filter with Electronically Reconfigurable Transfers. Elektronika Ir Elektrotechnika, roč. 21, 2015, č. 3, s. 33-37, ISSN: 1392-1215.

IDENTIFICATION OF SIGNIFICANT SPECTRAL COMPONENTS IN SPEECH SIGNAL IN STRESS

Egor Dulesov

Master Degree Programme (2), FEEC BUT

E-mail: xdules00@stud.feec.vutbr.cz

Supervised by: Jitka Poměnková

E-mail: pomenkaj@feec.vutbr.cz

Abstract: The purpose of this work is to compare the chosen methods of spectrum estimate and ways of identification of spectral components for realization to detect the stress in voice signal. Based on learning the special literature chooses the suitable methods of spectrum estimate. Determines an expected form and thresholds of voice signal in stress situation.

Keywords: Testing of significance, spectrum estimate, the periodogram, Bartlett, Welch, Burg, Yule-Walker,Matlab, speech, formants

1. ÚVOD

Lidskou řeč je možné považovat za nízkofrekvenční signál, který má svůj vlastní průběh a spektrální reprezentaci. Řečník se může nacházet v různém psychickém rozpoložení, jako je například normální (klidový) stav, nebo stres. O tom, v jakém stavu se člověk nachází, lze usuzovat podle toho, jak vypadá spektrální reprezentace jeho řečového signálu. Proto je třeba použít odhad spektra daného signálu a stanovit, jaké by měly být prahové hodnoty pro stresový a normální signál.

Cílem této práce je porovnat vybrané metody odhadu spektra a způsob identifikace významných složek spektra pro implementaci při detekci stresu v hlasovém signálu.

2. METODIKA

2.1. METODY ODHADU SPEKTRA SIGNÁLU

Existují dvě základní skupiny metod odhadu spektra signálů: neparametrické a parametrické metody. Neparametrické metody jsou metodami odhadu výkonového spektra, které provádějí odhad přímo z navzorkovaného signálu, přičemž odhady se používají pro konečný soubor dat. K této skupině metod patří odhad spektra pomocí periodogramu, nebo jeho modifikací. Mezi tyto modifikace můžeme zařadit například metody Welchova nebo Bartlettova periodogramu. Široce využívanou je také metoda Fourierovy transformace.

Parametrické metody předpokládají rozděleni algoritmu odhadu spektra na dva kroky, kdy jsou nejdřív odhadnuty parametry autoregresního modelu, který hypoteticky modeluje signál, a pak na základě těchto parametrů se provádí odhad spektra. K těmto patří metody Yulear-Walkerova, Burgova nebo Tomsonova metoda. Jejich hlavní výhodou je to, že často přinášejí lepší výsledky, než klasické neparametrické metody pro případ malých délek signálů, resp. časových řad. [1]

Vlastnosti tónové struktury řeči

Lidská řeč se přenáší ve formě akustických, tj. nízkofrekvenčních, vln. Energie akustického signálu má nejvýznamnější hodnoty v okolí určitých frekvencí. Tyto frekvence se nazývají formantové

frekvence, a oblasti koncentrace energie se nazývají formanty a označují se čísly, počínaje formantem s nejnižší frekvencí (F_1, F_2, …, F_n).

Základními jednotkami zvukové stránky řeči jsou hlásky, které je možné dále rozdělit do dvou skupin: samohlásky a souhlásky. Samohlásky jsou nejvíce charakterizovány právě formanty, přičemž jsou důležité první dva až tři. V této práci byly pro analýzu vybrány formanty samohlásek *a, á,* pro které jsou $F_1 = 700 — 1100$ *Hz*, $F_2 = 1100 — 1500$ *Hz*. [2]

2.2. ALGORITMUS DETEKCE STRESU V ŘEČOVÉM SIGNÁLU

V této práci byl navržen následující algoritmus pro detekci stresu v řeči pomocí programu Matlab:

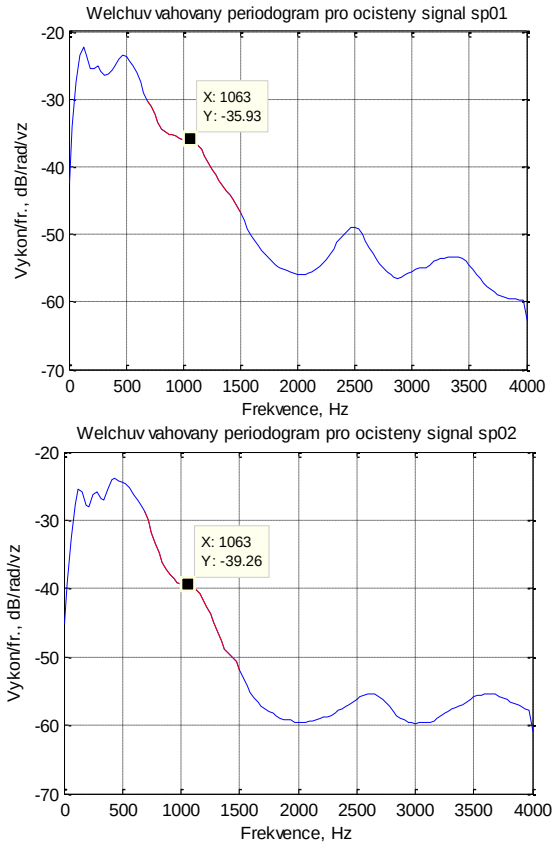
1. Načtení signálu s vhodnou vzorkovací frekvencí.
2. Rozdělení celého signálu do několika jednotlivých úseků - segmentů. To se provede tím způsobem, že je signál po načtení uvažován jako vektor a dělí se na segmenty pomocí transformace vektoru do matice, obsahující *m* řádků a *n* sloupců. Každý sloupec je pak samotný segment.
3. Hledaní formantů v každém jednotlivém úseku (v sloupci matice), které se provádí pomocí LPC (Linear predictive coding). [3]
4. Úprava původního signálu na nový „očištěný“, obsahující jen ty části původního signálu, ve kterých se vyskytují formanty (odstraněním těch sloupců matice, které neobsahují formanty) a jeho analýza.
5. Odhad spektra „očištěného“ signálu.
6. Výpočet statistických ukazatelů „očištěného“ signálu.
7. Stanovení prahových hodnot na základě vypočtených statistických ukazatelů, podle kterých se pak rozhoduje, zda je řečník ve stresové situaci, nebo není.

2.3. ZPRACOVÁNÍ VÝSLEDKŮ V MATLABU

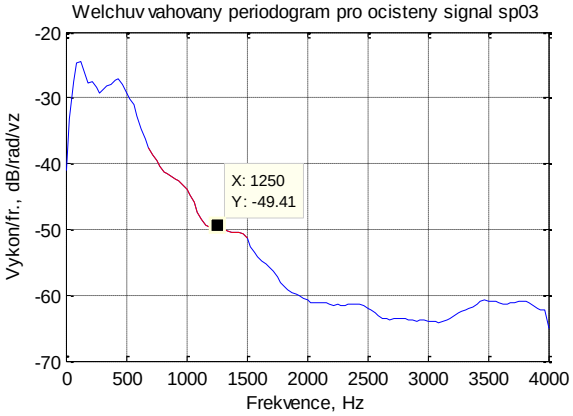
Pro analýzu spektra byla použita metoda Welchova periodogramu s váhováním Hammingovým oknem pro tři audio soubory *sp1*, *sp2* a *sp3* s maximální frekvencí *4000 Hz*. Z toho vyplývá, že vzorkovací kmitočet pro tyto signály je *fs = 8000 Hz*. Soubory *sp01* a *sp02* obsahují řeč ve stresovém stavu, soubor *sp03* řeč v normálním, nestresovém stavu. Výsledky odhadu spekter těchto signálů jsou uvedeny na obrázcích 1 a 2. Červenou barvou je označen interval pro formantové složky pro samohlásky *a, á*.

Jak je vidět na obrázcích 1 a 2, významné složky pro formant *a, á* se vyskytují na frekvencích *1063 Hz* pro stresové signály a na *1250 Hz* pro signál nestresový. V případě stresových signálů tyto složky mají amplitudy *-35,93 dB* a *-39,26 dB*, ale pro nestresový signál je tato amplituda pro stejný formant nižší, jen *-49,41 dB*.

Jak je uvedeno v části 2.2, formanty jsou oblastmi koncentrace energie akustického signálu, tj. energie formantové oblasti může být důležitým statistickým ukazatelem. V případě stresových signálů je energie, na základě výsledků vypočtených v programu Matlab, pro *sp01 -56,72 dB*, pro *sp02 -56,97 dB*. Pro nestresový signál *sp03* je tato hodnota nižší, pouze *-71,71 dB*.



Obrázek 1: Výsledný odhad spekter stresových signálů *sp01* a *sp02*



Obrázek 2: Výsledný odhad spektra nestresového signálu *sp03*

3. ZÁVĚR

V této práci byl navržen jeden z možných algoritmů detekce stresu v řečovém signálu. Hlavním principem je zde hledání a analýza spektrálních složek, které se vyskytují v oblasti formantů. Podle jejich hodnot výkonu a energie lze detektovat stres. Zpracování bylo zatím provedeno jen pro dva stresové a jeden nestresový signál a jen pro formanty samohlásek *a, á*. Podle dosažených výsledků lze uvést přibližně vypočítanou prahovou hodnotu hledané spektrální složky při stresovém stavu řečníka. Hodnota jejího výkonu se pohybuje v intervalu *-40…-35* dB. Z výsledků dále vyplývá, že energie oblasti formantu musí být pro stresový signál přibližně o *15 dB* výš.

REFERENCE

[1] PROAKIS, John G. Algorithms for statistical signal processing. Upper Saddle River, N.J.: Prentice Hall, c2002, xii, 564 p. ISBN 0130622192.

[2] PSUTKA, Josef. Mluvíme s počítačem česky. Vyd. 1. Praha: Academia, 2006, 746 s. Česká matice technická (Academia). ISBN 80-200-1309-1.

[3] http://www.mathworks.com/examples/signal/4522-formant-estimation-with-lpc-coefficients

IMPLEMENTATION ERROR CODING FOR NG-PON2 IN MATLAB

Martin Gallo

Master Degree Programme (2), FEEC BUT

E-mail: xgallo02@stud.feec.vutbr.cz

Supervised by: Tomáš Horváth

E-mail: xhorvath@feec.vutbr.cz

Abstract: The work deals with mathematical implementation of the newest NG-PON2 standard for PON networks from error correction coding point of view. The aim is to determine the effectiveness of this system in mathematical model for transmitting and receiving data via optical interface with help of artificial noise interface.

Keywords: NG-PON2, error correcion coding, matlab, BCH, RS

1. ÚVOD

Prvá myšlienka vytvorenia siete typu PON vznikla v 80. rokoch rokoch minulého storočia, keď sa začalo pracovať na zefektívnení prenášania objemov dát a takisto zakomponovať do neho práve roz- víjajúce sa optické riešenie. Takáto sieť prenáša dáta optickým prostredím medzi začínajúcim bodom pasívnej optickej siete OLT (Optical Line Terminal) a koncovým bodom nazývaným ONU (Optical Network Unit). Prvým počinom sa stal štandard B-PON (Broadband-PON), ktorý prestal vyhovo- vať s rozšírením internetu v 90. rokoch minulého storočia a cez rôzne medzi stupne sa dopracoval až na štandard G-PON, ktorý s jeho základnou architektúrou slúži v optických prístupových sieťi- ach dodnes. Tento článok pojednáva o jeho najnovšej verzii NG-PON2, ktorá ponúka až 40 Gbit/s (4×9.95328 Gbit/s) rýchlosť v režime donwstream, teda od jednotky OLT smerom k užívateľskej ONU. Jednou z techník, ktoré ponúka je zabezpečenie dátového toku pre potreby optického prostre- dia a práve tomuto sa venuje tento článok.

2. VRSTVOVÝ MODEL V NG-PON2

Z pohľadu referenčného sieťového modelu TCP/IP zastáva štandard vrstvu sieťového rozhrania, ktorá v tomto referenčnom modeli nie je bližšie definovaná. Technológia NG-PON2 teda poskytuje bližšie definovaný model a réžiu potrebnú na spoľahlivý prenos dát v optickom prostredí pri zachovaní požadovaných parametrov. Spomínanú pridanú hodnotu predstavuje Vrstva prenosovej konver- gencie (TC-TWDM), ktorú tvoria 3 subvrstvy:

- subvrstva adaptácie služieb
- rámcová subvrstva
- fyzicka adaptačná subvrstva.

Tieto 3 subvrstvy, ktoré sú definované v štandardizácii [1], tvoria jadro optickej siete založenej na NG-PON2. Každá má na starosti svoju úlohu v modeli a to definovaním polí, ktoré obsahujú mnohé identifikátory a logické prenosové kanály dôležité pre fungovanie optickej siete. Jednotky SDU, ktoré predstavujú dáta vyšších vrstiev si preberá najvyššia subvrstva adaptácie služieb a rozdelením toku, pridaním XGEM záhlaví, dáta posúva vrstva druhej v poradí. Tu sa XGEM rámce zoskupia do FS rámcov o veľkosti 135 432 bajtov, z ktorých malú časť zaberá premenlivé záhlavie a zápätie. Na zabezpečenie záhlaví týchto vrstviev sa používa HEC (Hybrid Error Correction)

zabezpečenie, ktoré sa skladá z cyklického samoopravného BCH kódu a jedného paritného bitu. Toto zabezpečenie je použité aj pre XGEM záhlavie na tretej subvrstve. Dáta následne pokračujú k poslednej fyzickej adaptačnej subvrstve, ktorá má na starosti pripravenie dátového toku na vyslanie optickým médiom. Rámce z druhej subvrstvy sú tu rozdelené na kódové slová dlhé 216 bajtov a a následne použité RS (Reed-Solomon) samoopravné blokové kódovanie z rodiny FEC (Forward Error Correction) typu typu RS(248,216). Slovo dlhé 216 bajtov je zabezpečené 32 bajtovou paritou. Takto zabezpečený dátový tok je pred vyslaním ešte scramblovaný, čo zabezpečí rozprestrenie zhlukových chýb.

2.1. HYBRIDNÝ SAMOOPRAVNÉ KÓDOVANIE HEC

Prvú časť kódu tvoria samoopravné BCH kódy využívajúce ku svojej prácu teóriu konečných (Galoisových telies) a to v režime BCH (63, 51, 12). Kód sa používa na zabezpečenie 51 bitov v záhlaví a pridanie 12bitového syndrómu. Je schopný opraviť 2 chybné bity. Posledný 13. bit HEC poľa je vyhradený pre párnu paritu, ak je nastavený na 1, počet jednotiek v poli je párne číslo.

2.2. REED-SOLOMONOVE SAMOOPRAVNÉ KÓDY

Tieto samoopravné FEC kódy pracujú s blokmi dát, v tomto prípade bajtami. Sú z rodiny BCH kódov a teda ich princíp bude podobný. V modeli NG-PON2 sa používajú na spodnej subvrstve v režime RS (248, 216) a teda 216 bajtov je doplnených o 32 bajtové protichybové zabezpečenie. Na prijímacej strane, podobne ako pri BCH kódoch je táto redundancia odstránená.Reed-Solomonov kód je schopný opraviť t chybných symbolov, pričom platí $2t = n - k$. Kvôli šetreniu šírky pásma je toto kódovanie pri nízkych hodnotách chybovosti v štandarde NG-PON2 vypínané.

3. PREVEDENIE ZABEZPEČ ENIA V PROGRAME MATLAB

Implementácia zabezpečenia TC-TWDM v programe matlab si vyžaduje prístup k problematike na úrovni bitov. Vrchné dve subvrstvy používajú na zabezpečenie jednotlivých svojich polí HEC kódova- nie, kdežto spodna subvrstva práve rozdelenie FS rámca druhej subvrstvy na 627 kódových slov nad ktorými vykoná zabezpečenie RS (248, 216) kódom. Pridaná redundancia teda zväčší kódové slovo zo 1728 bitov na 1984.

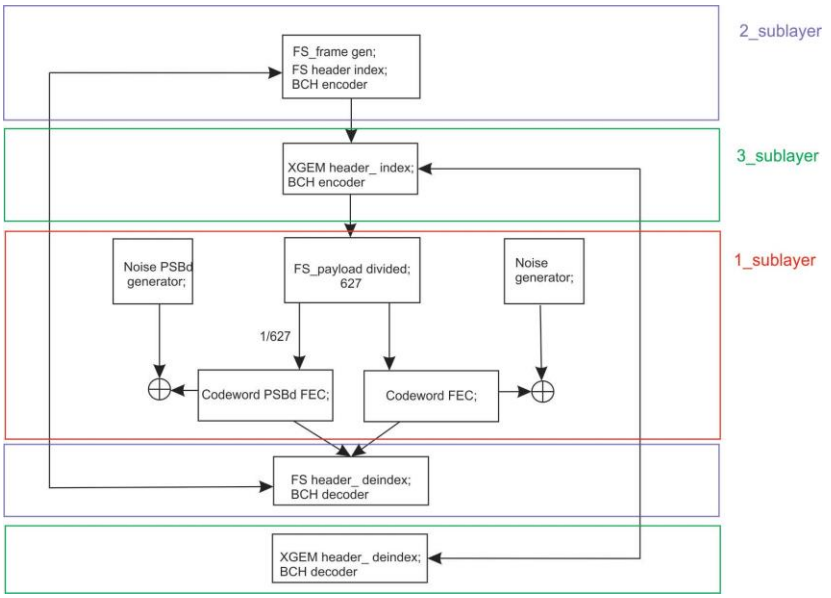
Keďže dátové jednotky tretej subvrstvy XGEM rámce majú premenlivú veľkosť, začiatok programu predstavuje generátor FS rámca o veľkosti 1 083 456 bitov s rovnakou pravdepodobnosťou núl a jednotiek $P\{1\} = P\{0\} = 0,5$, teda na subvrstve druhej.

```
FS_frame = round(rand(1,(1083456)));
```

Pre potreby sledovania chybovostí v jednotlivých poliach subvrstiev sú následne zaindexované polia druhej subvrstvy, ako napríklad pole identifikátora *Alloc_ID*, podľa [1]

```
Alloc_ID = FS_frame(1,32:45);
```

Táto indexácia bude neskôr použitá pri dekódovaní dát na strane prijímača. Zjednodušený diagram programu v matlabe je na obr. 1. Dáta postupne prechádzajú vrstvami, pričom každá zabezpečí vybraný blok dát. Je dôležité dodržať pozíciu jednotlivých polí vo vrstvách pre následné dekódovanie a určenie jednotlivých chybovostí. Zatiaľ čo BCH opravné kódy pracujú s bitmi, RS kódy naopak pra- cujú s blokmi dát, v konkrétnom príapde s h´lbkou 8, teda bajtami. Nutné je teda bitový tok premeniť na logický bajtový s rešpektovaním MSB a LSB bitov, následne prevedenie do desiatkovej sústavy a privedením do scramblera. Po výstupe zo scramblera dáta opäť premeniť na binárnu sústavu a z matici bajtov vyčítať jednotlivé bity zaradom do bitového toku. Takto pripravený bitový tok potom logickou funkciou OR sčítať s generátorom šumu, teda generátorom jednotiek určenej pravdepodob- nosti. Chybovosť je teda počítaná pre každé zo 627 kódových slov zvlášť, dáta sú ukladané do matice a posúvané vyššej subvrstve.



Obrázek 1: Zjednodušený vývojový diagram

Medzi jednotlivými subvrstvami je takisto nutné vyriešiť správne rozmery jednotlivých matíc, pri ich populácii a následnom porovnávaní. Podobne je dôležité dopraviť pôvodné vygenerované dáta až na koniec programu, pre zistenie celkovej chybovosti.

4. ZÁVER

Konkrétna implementácia zabezpečenia v matlabe je schopná zistiť chybovosť nielen celkovú, ale aj chybovosť v jednotlivých poliach subvrstiev. Toto je možné aj vďaka umelému matematickému šumovému prostrediu, ktoré je realizované dvomi generátormi a logickým súčtom OR na spodnej fyzickej subvrstve PHY. Keďže program matematického modelu zabezpečenia bol pôvodne písaný pre spoluprácu so simulačným nástrojom VPI photonics, neimplementuje schému PON siete, ale len jednoduché point-to-point zapojenie pre downstream. Z nameraných výsledkov pre model generáto- rového šumu vyplýva, že systém je schopný spoľahlivo korigovať chyby pri dodržaní stanovených hodnôt BER, do pravdepodobnosti šumového generátora $P\{1\} = 0, 012$.

5. REFERENCE

[1] Recommendation ITU T G.989.3. 40-Gigabit-capable passive optical networks (NG-PON2): Transmission Convergence Layer. 04/2013.[online] [cit. 2016-3-6]. Dostupné z URL: <http://www.itu.int/rec/T-REC-G.989.3/en>.

[2] Recommendation ITU T G.987.3. 10-Gigabit-capable passive optical networks (XG-PON): Transmission convergence (TC) layer specification 01/2014.[online] [cit. 2016-3-6]. Dostupné z URL: <http://www.itu.int/rec/T-REC-G.987.3/en>.

IMPROVEMENT OF BIT ERROR RATE IN FREE SPACE OPTICAL LINK

Marek Novák
Master Degree Programme (2016), DREL FEEC BUT
E-mail: xnovak0m@vutbr.cz

Supervised by: Lucie Hudcová
E-mail: hudcova@feec.vutbr.cz

Abstract: The article describes an inovative bit error rate reduction technique principle and its practical implementation. The design is implemented in an FPGA and can be combined with other more conventional BER reduction techniques. The presented approach benefits from properties of an optical channel which a general RF channel does not have.

Keywords: BER, FSO, FIFO, FPGA, Xilinx MIG

1 INTRODUCTION

One of the most important factors deteriorating the bit error rate in free space optics (FSO) is atmospheric turbulence. If the size of the "turbles" (turbulent cells) producing this turbulence is larger than the transmitted beam diameter, it can be completely deflected [1]. This effect called beam wander causes this way a complete signal loss. The duration of this effect is in the order of milliseconds. If a practical communication speed of 1.25 Gbps and more is considered, such a channel unavailability cannot be simply covered by interleaving and channel coding. Herein proposed technique solves this problem.

2 PROPOSED BER REDUCTION TECHNIQUE

As shown for example in [2], the principle of reciprocity applies in the classical optics. This fact is the root idea of our technique. In the considered communication setup, a full duplex fibre-less free space optical link is used. This allows each side of the link to analyse the incoming signal from the other side to estimate instantly whether the channel is available for transmission or not. This cannot be easily applied in radio frequency links, because of the multipath propagation which does not apply in FSO.

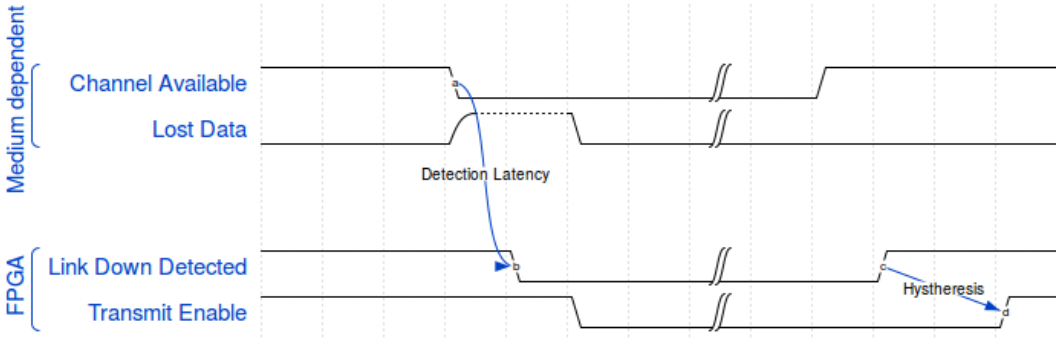


Figure 1: High level operation description

2.1 PRINCIPLE OF OPERATION

When an obstacle in form of a turbulent cell or a flying bird or other object appears in between the two communication terminals, it makes the channel temporarily unavailable. This can be detected either by a drop in receiving optical power and/or by observing integrity of data being received from the other side of the link. The last mentioned detection method can be used mainly when only one side of the link transmits data and the other one is idling and sending so-called comma symbols (which are also used to realize the symbol boundary bit alignment). When an unavailability is detected, the transmission is stopped and further data to be transmitted is put in a FIFO-based buffer. The latency depicted in Figure 1 is introduced to avoid intermittent drops in the communication. Multiple consequent errors must occur for link to be down.

Thanks to this approach, the optical communication terminal can safe some of the payload from being lost. There is however always a portion of the payload lost because of the detection latency as shown in Figure 1. This losses can be partly recovered by interleaving and channel coding or by forward error correction (FEC) mechanisms used at the application level by actual communicating endpoints. Furthermore, when the loss occurs in the middle of a frame, the whole frame is likely lost.

2.2 REQUIREMENTS

The proposed approach requires the communication speed to be fast enough, when the actual received data (comma symbols) is to be used in the channel unavailability detection. Another requirement is placed on the temporary payload storage buffer. It needs to be large enough so that it does not overflow, when the communication is not possible for a longer period of time (e.g. 200 ms). Of course, the required storage capacity grows with the communication speed.

For example a single comma symbol which consists of 10 bits is transmitted in 8 ns on a 1.25 Gbps link. On the same link, a 200 ms drop-out produces about 25 MB of data. In a real design, a DDR memory of 256 MB should be therefore sufficient.

3 PRACTICAL IMPLEMENTATION

Figure 2 shows an overall architecture of the proposed technique. The Loss-of-Signal (LOS) information is used both to stop reading out of the data from the FIFO and to switch the transmitter from transmission of payload to transmission of comma symbols.

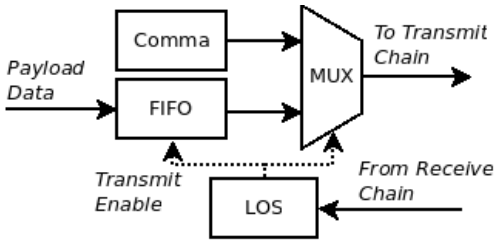


Figure 2: BER reduction technique block diagram

DDR FIFO

In order to overcome the problem with memory requirements described in 2.2, a DDR-based FIFO memory was designed. The implementation itself uses the memory interface generator [3] (MIG) IP core from Xilinx and interfaces this with “DDR Reader and Writer”. This module uses the half-duplex communication interface with DDR2 SODIMM module and implements a circular buffer in

it. Thanks to the bus width of 128 bits, the module has enough time to keep the "Input FIFO" empty and the "Output FIFO" full during non-interrupted data transmissions. 128 bit wide communication bus was selected, since it is the maximum bus width supported by the ML505 Evaluation Platform used in testing of the design and it gives enough time margin for the processing of input and output FIFOs.

Data frames are stored in the DDR FIFO in a microframe memory scheme. The scheme was developed to separate individual data frames. It enables the output module to replicate the Start-of-Frame (SOF) flag at the beginning and the End-of-Frame (EOF) flag at the end of each data frame. The scheme uses 128 bit data units called microframes. The first fifteen bytes of microframe contain data payload and the last byte is a control character. It indicates whether the SOF or EOF flags should be generated and the size of data payload in each microframe. The indication of the data size in the last microframe is crucial as the incoming frames are not 15 B aligned and 15 B alignment is needed for the microframe memory scheme.

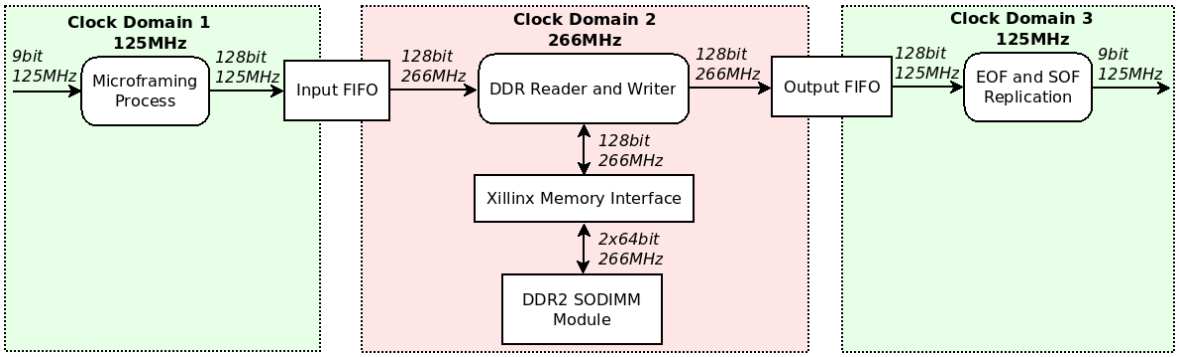


Figure 3: DDR FIFO internal data flow

4 CONSLUSION

This article briefly describes a simple, yet efficient technique to reduce the BER of an FSO link. Its performance is however to be evaluated in real environment, nevertheless the theory and current on-the-table testing shows, that this approach can cooperate with existing techniques to reduce the overall bit error rate. The source codes are available at a GIT repository: `https://gitlab.com/OK2NMZ/thesis-fso.git`

ACKNOWLEDGEMENT

The described research was performed in laboratories supported by the SIX project - the registration number CZ.1.05/2.1.00/03.0072.

REFERENCES

[1] WEYRAUCH, Thomas; VORONTSOV, Mikhail A. Free-space laser communications with adaptive optics: Atmospheric compensation experiments, p. 26, 2008.

[2] MANSURIPUR, MASUD. Reciprocity in Classical Linear Optics. Optics and Photonics News. 1998. DOI: 10.1364/OPN.9.7.000053. ISSN 10476938.

[3] XILINX. Memory Interface Solutions: UG86. 2010. Available from: `http://www.xilinx.com/support/documentation/ip_documentation/ug086.pdf`

MULTIMODE VHF RECEIVER

Vladimír Hradňanský

Master Degree Programme (2), FEEC BUT

E-mail: xhradn00@stud.feec.vutbr.cz

Supervised by Jiří Šebesta

E-mail: sebestaj@feec.vutbr.cz

Abstract The aim of this work is to design a microprocessor-controlled multimode VHF radio which can receive AM, FM and DAB. The design of the receiver circuits comes from the single-chip programmable defined radio Si4689 which covers all required modes. The control unit based on RTOS SC23 processor contains a firmware which ensures all needed function of the prototype. The firmware has been verified including testing of all parameters. A multi-mode FM radio, MP3 player and on-line radio stream player operated via touch green and remote control and supported web management is the final result of this work.

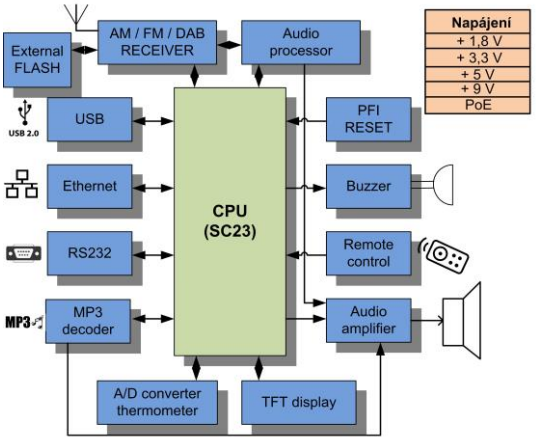
Keywords: DAB, SC23, BECK-IPC, AD7417, WEB-player, VHF receiver, Si4689, TFT display,FT801, DS1307, RS232, remote control, www, telnet.

1. ÚVOD

Cílem této práce bylo navrhnout mikroprocesorem řízený multimódový VKV rozhlasový přijímač s možností příjmu AM, FM a DAB. Na základě analýzy možností řešení kombinovaného FM/DAB přijímače s využitím moderní součástkové základny byl vybrán jako jádro multimodového přijímače obvod Si4689, který reprezentuje tzv. programově definované rádio. Pro tento obvod byla vytvořena kostra (knihovna) obslužného firmware a zkonstruován prototyp přijímače, následně pak ověřeny jeho funkce a parametry. Výsledkem práce je multimódový VKV rozhlasový přijímač rozšířený o MP3 přehrávač a přehrávač webových rádiových streamů. Přijímač je ovládán přes dotykovou obrazovku a dálkový ovladač s možností webové správy.

2. BLOKOVÁ STRUKTURA VÝROBKU

Blokové schéma přijímače je uvedeno na obrázku 1. Jádrem přijímače je již zmíněný programově definovatelný přijímač Si4689 [1]. Řídící procesor SC23 [2] s vlastním RTOS systé-

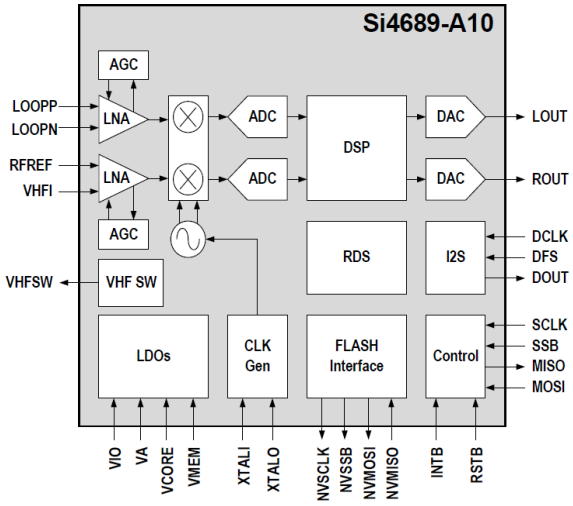


Obrázek 1: Blokové schéma realizovaného přijímače

mem byl zvolen díky implementovanému ethernetovému a USB rozhraní pro webovou správu, ovládání přijímače a možnosti přehrávání mp3 streamu a souborů. Uživatel přijímače jistě rád ocení možnost audio korekcí díky vestavěnému audio procesoru. Dále je tu možnost dálkového ovládání, barevný TFT 4.3 kapacitně dotykový display a RTC obvod pravidelně synchronizovaný s NTP serverem a mnoho dalších funkcí. Komunikace uvnitř výrobku probíhá na jednoduché I2C sběrnici s výjimkou mp3 dekodéru, který díky větším datovým tokům vyžaduje SPI sběrnici.

2.1. Čip přijímače Si4689

Výkonným jádrem systému je jednočipový multimódový CMOS přijímač AM, FM a DAB v pouzdře QFN48 Si4689-A10 od firmy Silicon Labs [1]. Tento obvod zajišťuje podporu nejen analogového vysílání AM v rozsahu 520 až 1750 kHz. Dále i celosvětový rozsah kmitočtů HD FM pro příjem od 76 do 108 MHz a integrovaný dekodér DAB a DAB+ a DBM podporující III TV pásmo v rozsahu 168 až 240 MHz. Podpora všech módů přijímače je realizována díky internímu firmwaru pro DSP jádro pro dané módy příjmu, který se volí výběrem adresy v externí flash paměti. Napájení je dvojího typu, a to 1,8V a 3,3V. Komunikace s řídicím procesorem probíhá pomocí I2C sběrnice. Blokové schéma obvodu přijímače je zobrazeno na obrázku 2.



Obrázek 2: Bloková struktura čipu přijímače Si4689 [1]

2.2. Praktická realizace

Zapouzdření elektroniky je provedeno v estetické krabičce pro přijímač s vestavěnými reproduktory. Přijímač je realizován na dvou deskách plošných spojů obsahujících převážně součástky SMD technologie, kdy nejmenší pasivní prvky jsou velikosti 0603 a pouzdro čipu QFN48.



a b

Obrázek 3: Pohled na realizovaný multimódový přijímač a) zepředu, b) zezadu



Obrázek 4: Vnitřní uspořádání přijímače

3. ZÁVĚR

Multimódový VKV přijímač je navržen za pomocí moderní součástkové základny. Přijímač by měl umět přijímat klasické analogové rádiové vysílání v amplitudové modulaci od 520 do 1750 kHz a frekvenční modulaci od 88 do 108 MHz. Přijímač realizuje i dekódování dodatkových informací (tzv. RDS), které budou v další části práce zobrazeny na dotykovém a barevném TFT displeji. Frekvenční rozsah navrženého digitálního rozhlasového vysílání (DAB+) je v rozsahu 168 až 240 MHz. Multimedialita přijímače je v navržené schopnosti přijímat rozhlasové vysílání na třech různých modulacích, přijímat a dekódovat datový audio stream z internetu a přehrávat hudební soubory na USB klíčence. Přijímač byl navrhnut tak, aby uměl komunikovat s uživatelem více způsoby. Jako hlavní komunikační prostředek je zvolen barevný dotykový 4.3 palce velký TFT displej [3], na kterém budou zobrazovány provozní informace. Dalším komunikačním prvkem jsou www stránky komunikující po ethernetovém rozhraní. Na www stránkách budou zobrazovány veškeré dostupné informace dynamicky reagující na daný provozní mód. IP adresu přijímače lze konfigurovat jako pevnou nebo dynamickou. Nepostradatelnou částí komunikace je dálkový ovladač komunikující na protokolu NEC. Potvrzení nebo výstraha může být zvýrazněna pomocí akustického piezoměniče.

Audio korekce (výšky, středy, hloubky) lze upravit pomocí audio procesoru, který jej realizuje jak pro signál z AM/FM/DAB přijímače, tak i pro MP3 dekodér. Datum a čas je realizovaný díky RTC obvodu, který je pravidelně synchronizován s NTP serverem. Funkce reagující na neočekávané vypnutí (PFI) umožní zálohovat data pro daný provozní mód a při opětovném spuštění tak přijímač uvést do předchozího stavu.

Napájení přijímače je realizováno pomocí externího adaptéru s výstupním napětím +5 V nebo pomocí vestavěného PoE měniče. Výstup lze přesměrovat na audio sluchátkový či na I2S datový stream z výstupu AM/FM/DAB přijímače pro další zpracování. Maximální výkon vestavěných 4 Ω reproduktorů je 3 W stereo, změřený výstupní výkon koncového stupně je 2,6 W do 4 Ω zátěže. Příkon přijímače je 2,9 W. Špičkové proudy, zaviněné pracovní třídou D u koncového stupně mohou být až 0,8 A.

REFERENCE

[1] Silicon Laboratories Inc. *Si4689-A10: SINGLE-CHIP, AM/FM/HD/DAB/DAB+ RADIO RECEIVER* [online]. 1. Austin: Silicon Laboratories Inc., 2014 [cit. 2016-03-07]. Dostupné z: https://www.silabs.com/Support%20Documents/TechnicalDocs/Si4689-A10-short.pdf

[2] SC23 Embedded Web Controller: SC23. *BECK-IPC* [online]. Německo: Beck IPC GmbH, 2016 [cit. 2016-03-07]. Dostupné z: http://www.beck-ipc.com/en/products/sc2x/sc23.asp

[3] *LCD TFT datasheet: RVT4.3B480272CFWC81* [online]. 1.4. Itálie: Riverdi, 2014 [cit. 2016-03-07]. Dostupné z: https://riverdi.com/datasheet/RVT4.3B480272CFWC81

NETWORK INTERFACE FOR MICROBLAZE

Lukáš Janík
Master Degree Programme , FEEC BUT
E-mail: xjanik12@vutbr.cz

Supervised by: Lucie Hudcová
E-mail: hudcova@feec.vutbr.cz

Abstract: This article deals with interfacing of Xilinx MicroBlaze (UB) softcore microprocessor to 1 Gb/s Ethernet data bus via custom simple Ethernet switch. This interface is implemented on an FPGA board and serves as the lower communication layer for administration of gigabit Ethernet-to-optical bridge.

Keywords: Free space optics, MicroBlaze, Ethernet, Block RAM

1. INTRODUCTION

Free space optical communication (FSO) brings many advantages compared to common radio frequency (RF) wireless links, however its reliability and availability is considerably dependent on atmospheric conditions [1]. Information about current atmospheric conditions and link state (bit error rate, loss of signal) should be gathered and monitored. The implementation of softcore microproces-

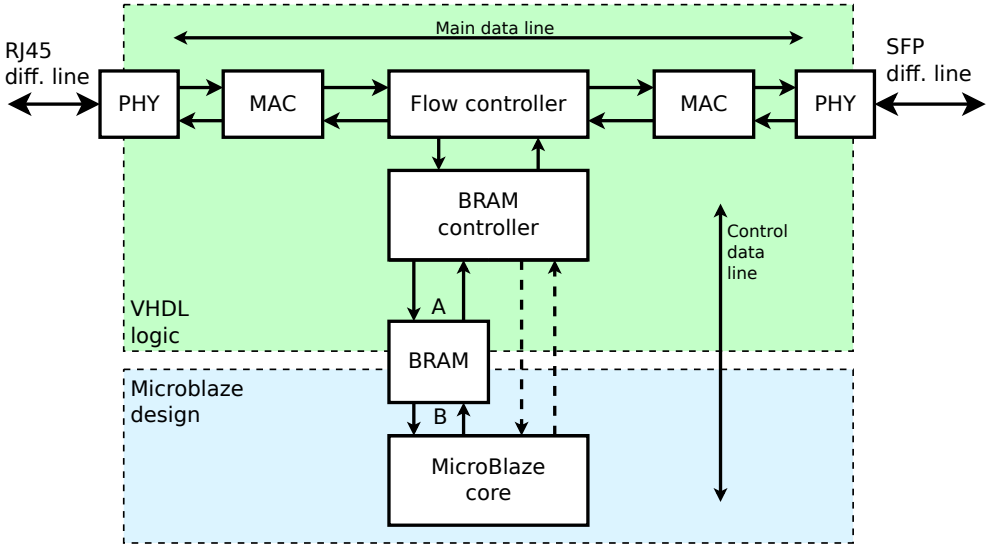


Figure 1: Discussed design block scheme. Green area contains user-logic written in VHDL, blue area contains Microblaze design blocks, designed in Xilinx platform studio (XPS).

sor on single chip together with Ethernet-to-optical bridge (main data line) presents effective solution for statistical processing and data gathering without the need for another external device or software.

There are many example implementations of Xilinx MicroBlaze Processor connected to Ethernet however they use own implementation of Media Access Controller (MAC) and use PHY chip exclusively, disallowing further data bridging to optical interface (additional MAC and gigabit Ethernet

switch would be needed for frame routing). Further presented design is based on unmanaged media bridge design and utilizes current MAC.

2. DESIGN OVERVIEW

An example of Ethernet-to-optical bridge allowing full duplex 1 Gb/s data rate can be seen in Figure 1. This design is implemented on Xilinx ML505 Evaluation Platform on single FPGA chip (Xilinx Virtex-5) and interfaced with onboard PHYs on metallic (1000Base-T) and optical (single-mode 2.5 Gb/s SFP optical transciever) side.

Main data line signals from metallic interface are processed by PHY chip and passed via SGMII interface to Xilinx Tri-Mode Ethernet Media Access Controller (TEMAC). TEMAC is connected to user logic via simple client interface [2] (signals DV, DIN, DOUT, GF, BF, ACK). Flow controller serves as a frame buffer between different clock regions, it drops corrupted frames and routes frames to and from MicroBlaze (Control data line). On SFP side, the MAC is significantly simplified and ensures frame synchronization, 8b/10b conversion and conversion between client interface and SGMII interface; none of other advanced MAC functions are needed.

2.1. FRAME ROUTING

Frame routing is being done in flow controller, Figure 2. Data from Ethernet interface is written to ETH-SFP FIFO; first six bytes of each frame containing destination MAC address are compared with MAC address register. If both addresses are equal, signal is sent to ETH-SFP read-out state machine, signalizing that current frame should be written to Microblaze FIFO (ETH-UB FIFO). The MAC address register is controlled from the Microblaze, allowing the change of MAC address from the software.

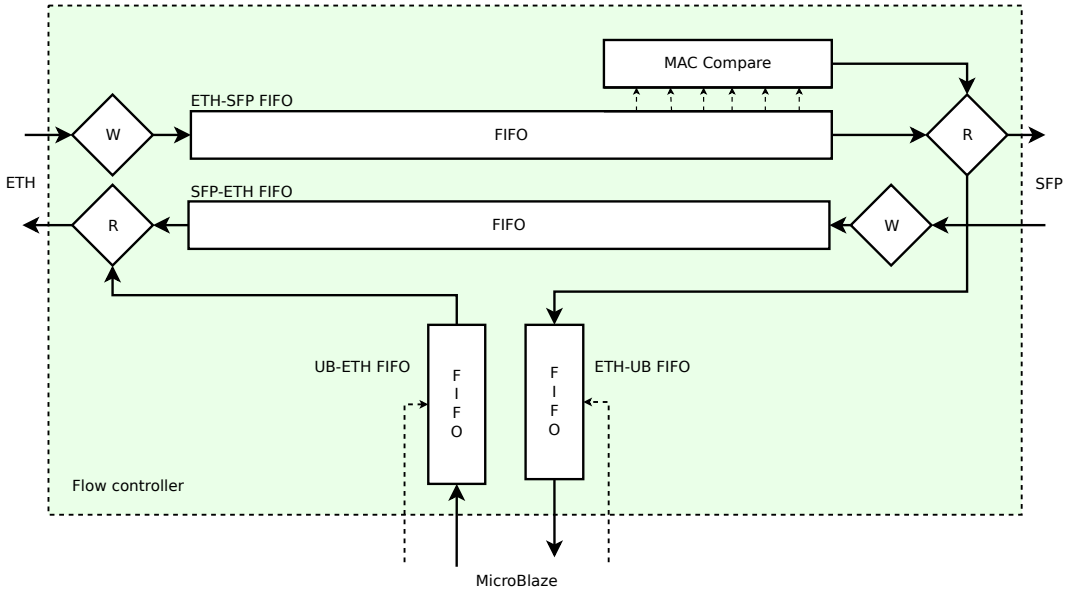


Figure 2: Flow controller FIFOs with MAC compare unit and read-out (R) and write-in (W) finite state machines

In the Microblaze uplink, the UB-ETH FIFO is emptied by SFP-ETH read-out state machine. The priority of main data line is ensured by first polling the SFP-ETH FIFO, data from Microblaze is sent only when SFP-ETH FIFO is empty. This behaviour is sufficient for common situations where interframe gaps occur, however, in the case when very long data stream comes from SFP interface,

frames from MicroBlaze would not be processed until the stream ends. For this purpose, simple 1:1 aggregation can simply be applied.

2.2. BRAM CONTROLLER

Since Ethernet frames vary in size, they cannot be read directly from FIFO memory and additional RAM must be added, allowing random access to the frame (single frame storage is enough). With maximum ethernet frame size of 1530 B, the 2 KB RAM for each direction is sufficient. On Virtex-5 devices, there are many small block RAM (BRAM) cells spread over the FPGA structure, those BRAM cells can be joined to greater ones (up to 1024 kB) and have two identical interfaces allowing read and write operations on both sides (even simultaneously).

The BRAM controller creates the interface between Flow controller FIFOs and BRAM block. This block is a part of Microblaze design, with its interface A connected to VHDL design and interface B connected via internal MicroBlaze memory controller and peripheral bus (PLB) to MicroBlaze core. First data is loaded to BRAM as soon as it is present in the ETH-UB FIFO. Data ready signal is sent to MicroBlaze, invoking RX interrupt. Next data is loaded to BRAM right after ACK from MicroBlaze. Opposite direction works similarly - uplink data are loaded to BRAM and data ready signal is sent to BRAM controller. Another frame can be loaded as soon as ACK from BRAM controller arrives.

Address space of Microblaze memory controller and thus address space of BRAM block is a part of monolithic address space of Microblaze processor, allowing direct software access to the frame.

3. CONCLUSION

Discussed design allows simple Ethernet frame switching between the Ethernet interface and the MicroBlaze microprocessor at layer of TEMAC client interface. The new control data line was added to the design without the need for additional MAC controller and without affecting the main data line throughput. At MicroBlaze, frames are read or written by accessing memory (e.g. by pointer dereference in C), which could be further interfaced to LwIP (lightweight IP stack implementation) library providing common socket API and few socket-based applications (HTTP server). With MicroBlaze directly connected to network, internal data (coding scheme, signal loss, BER, ...) can be sent directly via socket to user client application running on remote PC or viewed as a static web page using web browser.

This design is a part of opensource project available at:
`https://gitlab.com/OK2NMZ/thesis-fso`

ACKNOWLEDGEMENT

The described research was performed in laboratories supported by the SIX project - the registration number CZ.1.05/2.1.00/03.0072.

REFERENCES

[1] BOUCHET, Olivier. *Free-space optics: propagation and communication*. New- port Beach, CA: ISTE, 2006, 219 p. ISBN 9781905209026.

[2] XILINX. *Virtex-5 FPGA Embedded Tri-Mode Ethernet MAC User Guide* [online]. [cit. 2016-03-03]. Available at: `http://www.xilinx.com/support/documentation/user_guides/ug194.pdf`

PRINCIPLE OF THE NEW ADJUSTED ARCHITECTURE OF R-2R D/A CONVERTER

Zuzana Polešáková
Master Degree Programme (year 2), FEEC BUT
E-mail: xpoles03@stud.feec.vutbr.cz

Supervised by: Jan Jeřábek
E-mail: jerabekj@feec.vutbr.cz

Abstract: Original R-2R D/A Converter architecture has a shortcoming – while resolution is bigger than 7 bits, switching of MSBs causes a significant non-linearity error, which may even cause DAC to be non-monotonous. A possible solution to this issue is shown in this paper: dividing MSBs into subbits of a lower weight. Analog circuitry and digital driving is published in this paper.

Keywords: Analog Integrated Circuits, D/A Conversion, DNL, INL, R-2R DAC, Small Chip Area

1 INTRODUCTION

A design of an R2R D/A Converter (R2R stands for R-2R resistor ladder) with DNL (Differential Non-linearity Error) and INL (Integral Non-linearity Error) lower than 1 LSB (Least Significant Bit) and small area consumption is published in this paper. Well-known R2R D/A Converter has a disadvantage: when MSB (Most Significant Bit) is switched on and all the lower bits are switched off (0111→1000), DNL is the highest in the whole range. The height of the step may surpass 1 LSB significantly, especially if number of bits is high. If the analog value of the code word 1000 is higher than analog value of 1001, D/A converter is no longer monotonous [2].

Solution of this problem is described in this paper. D/A converter is designed in ON Semiconductor technology I4T (45 V, 180 nm). It is optimized for automotive applications. Technical parameters are 9 bits, $V_{DD} = 3.3$ V, $V_{REF} = 1.6$ V.

2 ARCHITECTURE OF THE R2R D/A CONVERTER

Well-known R2R D/A Converter is based on binary weighing of the bits by resistor ladder [1]. LSB of 9-bit R2R DAC has a weight of $\frac{V_{REF}}{512}$, while MSB $\frac{V_{REF}}{2}$. Switching of MSB causes huge non-linearity error. Mismatch error of this particular resistor has a huge influence on the precision of the whole converter. Solution is suggested in this paper: to avoid switching of such a significant weight in one step.

Experimentally it was proved that switching of the weight bigger than $\frac{V_{REF}}{8}$ is inconvenient [2]. 2 MSB's were divided into 6 subbits of weight $\frac{V_{REF}}{8}$, $(\frac{1}{2}+\frac{1}{4})V_{REF} = \frac{6}{8}V_{REF}$. 3rd MSB, which is originally $\frac{V_{REF}}{8}$, has different digital driving compared to the original DAC. In this paper, they are referred to as 7 subbits even though one of them is technically a bit.

7 subbits have the same weight – they are weighted linearly. All 3 MSB's in total have the weight of $(\frac{1}{2}+\frac{1}{4}+\frac{1}{8})V_{REF} = \frac{7}{8}V_{REF}$. Each subbit has the weight of $\frac{V_{REF}}{8}$. Weight of 1st MSB is equivalent to 4 subbits, weight of 2nd MSB to 2 subbits and the weight of 3rd MSB to 1 subbit.

Details of switching subbits instead of original bits is shown in Table 1. There are two options

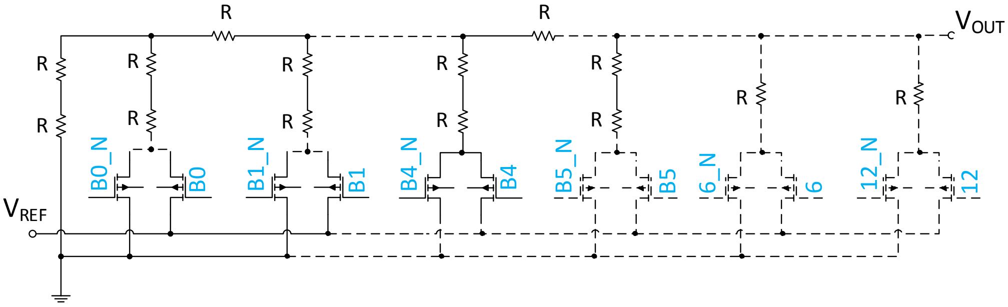


Figure 1: 9-bit R2R DAC. Gate labels stand for digital driving.

presented in this paper. DNL of the crucial steps is diminished via alternative switching. Circuitry is presented in Figure 2.

3 DIGITAL DRIVING

Original Bits			**Thermometric Driving**							**Advanced Driving**						
B8	B7	B6	12	11	10	9	8	7	6	12	11	10	9	8	7	6
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1
0	1	0	0	0	0	0	0	1	1	0	0	0	0	1	1	0
0	1	1	0	0	0	0	1	1	1	0	0	0	0	1	1	1
1	0	0	0	0	0	1	1	1	1	0	0	1	1	1	1	0
1	0	1	0	0	1	1	1	1	1	0	0	1	1	1	1	1
1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Table 1: Thermometric and Advanced Driving: Truth table of 3 MSB's and 7 subbits.

Truth table shown in table 1 is transformed into digital circuits driving the subbits. Thermometric driving requires 6 logic gates (shown in figure 2). Advanced driving requires only 2 logic gates, however, it increases DNL and INL (shown in Figure 3) as bigger weights are switched - two subbits at the same time. DNL and INL can be diminished if the area of the resistors is bigger. However, if thermometric driving is used, precision requirements are met while area consumption in total is smaller.

R2R DAC with both binary and linear weighing is described by equation 1

$$V_{OUT} = \frac{V_{REF}}{m+1}(sb_0 + sb_1 + ... + sb_{m-1} + sb_m) + \frac{V_{REF}}{2}b_{n-1} + \frac{V_{REF}}{4}b_{n-2} + ... + \frac{V_{REF}}{2^{(n-1)}}b_1 + \frac{V_{REF}}{2^n}b_0. \tag{1}$$

m stays for number of subbits weighted linearly. n stays for number of bits weighted binary. sb_m stays for the value of the subbit, 1 or 0. Weights are added decimaly ($1+1+1=3$). b_n stays for the value of a bit.

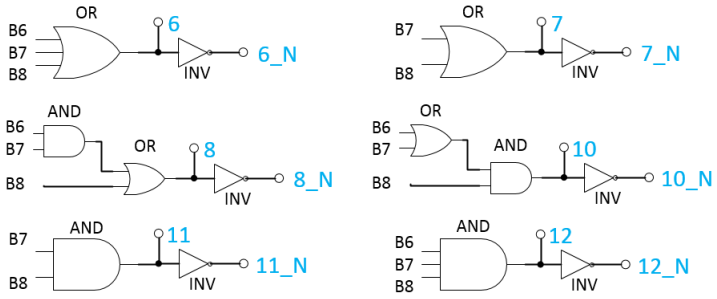


Figure 2: Digital driving circuits switching the subbits on and off – thermometric driving.

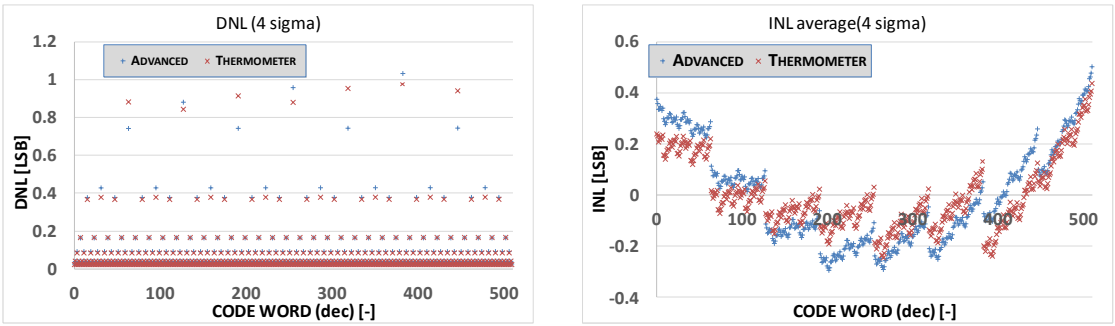


Figure 3: DNL and INL of Advanced and Thermometric solution.

4 CONCLUSION

Alternative architecture of R2R DAC was introduced in this paper. Two options of digital driving were simulated in Cadence Virtuoso 6. MonteCarlo mismatch simulation was performed (100 runs, Cadence Virtuoso), results are shown in Figure 3. It is obvious, that only switching two subbits at the same time makes DNL and INL bigger. If the MSB was switched as in original architecture, DNL and INL would significantly increase.

The area of the resistor bank on the chip is approximately 8 times smaller while using alternative architecture with 7 linearly weighted subbits and thermometric driving compared to the standard one and two times smaller compared to the one using advanced driving.

ACKNOWLEDGEMENT

This paper was created thanks to the support of company ON Semiconductor, which provided the author with sufficient software and helped by advice of the experienced workers.

REFERENCES

[1] Maloberti, F.: Data Converters. Springer 2007, e-book, ISBN-10 0-387-32486-0.

[2] Internal documents of the company ON Semiconductor.

PSEUDO-DIFFERENTIAL FILTER USING DDCCS

Ondřej Sládok

Master Degree Programme (2nd year), FEEC BUT

E-mail: xslado00@stud.feec.vutbr.cz

Supervised by: Jaroslav Koton

E-mail: koton@feec.vutbr.cz

Abstract: This article describes a pseudo-differential second-order filter, using three differential difference current conveyors (*DDCCs*) and five passive elements. The proposed structure is able to realize all types of frequency responses. The circuit has a high input impedance, low active and passive sensitivity and high common-mode rejection ratio (*CMRR*).

Keywords: Pseudo-differential filter, frequency filter, conveyors, common-mode rejection ratio, low-pass, high-pass, band-pass, band-reject, all-pass.

1. ÚVOD

Při návrhu analogových obvodů jsou v oblasti funkčních bloků často řešeny kmitočtové filtry, neboť své uplatnění nachází ve většině elektronických zařízeních. Svoji funkci zastávají např. při filtrování signálu na symetrické vedení, kde oddělují datový signál od telefonního. Jelikož v současné době je při návrhu funkčních bloků kladen důraz na nízké napájecí napětí a příkon, roste zájem o navrhování plně diferenčních kmitočtových filtrů, které v porovnání s nediferenčními kmitočtovými filtry vykazují velké procento potlačení souhlasného signálu (*CMRR*), jsou charakteristické nižším harmonickým zkreslením signálu a efektivně potlačují změny napájecího napětí (*PSRR*).

2. PLNĚ DIFERENČNÍ VS. PSEUDODIFERENČNÍ FILTRY

Obecně platí, že při analýze diferenčních obvodů pracujících v napěťovém režimu se předpokládají tyto vztahy [1]:

$$U_{1\mathrm{d}} = u_{1+} - u_{1-}\ ,\ U_{2\mathrm{d}} = u_{2+} - u_{2-}\ ,\ U_{1\mathrm{c}} = \frac{u_{1+} + u_{1}}{2}\ , \tag{1}$$

kde, $U_{1\mathrm{d}}$, $U_{2\mathrm{d}}$ a $U_{1\mathrm{c}}$ označují diferenční vstupní napětí, diferenční výstupní napětí a souhlasné vstupní napětí. Signál $U_{1\mathrm{d}}$ je rozdíl mezi dvěma vstupními signály u_{1+} a u_{1-}, zatímco $U_{1\mathrm{c}}$ souhlasný vstupní signál je průměr dvou vstupních signálů u_{1+} a u_{1-}.

Pokud budeme brát v úvahu rovnice (1), potom diferenční výstupní napětí $U_{2\mathrm{d}}$ můžeme definovat takto:

$$U_{2\mathrm{d}} = U_{1\mathrm{d}} \cdot A_{\mathrm{dm}} + U_{1\mathrm{c}} \cdot A_{\mathrm{cm}}\ , \tag{2}$$

kde, A_{dm} a A_{cm} je diferenční a souhlasné zesílení signálu, dle pořadí. Při vyhodnocování potlačení souhlasného signálu se od diferenčního zesílení odečítá zesílení souhlasné. Poměr zadrženého souhlasného signálu (common-mode rejection ratio) je pak dán vztahem:

$$CMRR = A_{\mathrm{dm}} - A_{\mathrm{cm}}\ .\ \left[\mathrm{dB}\right] \tag{3}$$

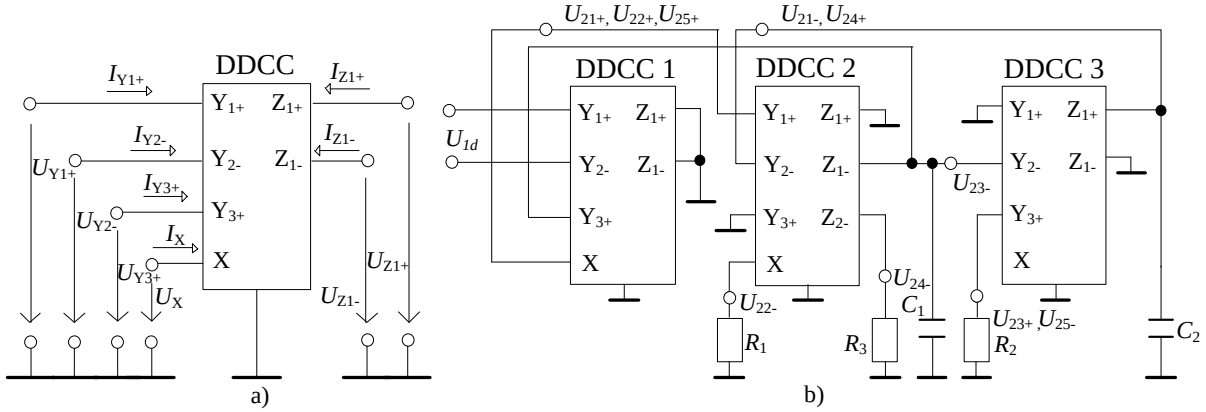
Parametr *CMRR* určuje, jak je aktivní prvek či nějaká diferenční struktura schopna potlačovat nežádoucí vstupní signál společný pro oba vstupy, vzhledem k požadovanému diferenčnímu signálu.

Nicméně, jak je zřejmé z matematického popisu (1) jsou k analýze takových obvodů používány pouze vstupní a výstupní signály. Lze tak popisovat a navrhovat tzv. pseudodiferenční filtry, které

mají diferenční vstupní a výstupní napěťové brány, kdy však samotné obvodové řešení filtru je nediferenční [2]. Tyto typy filtrů zajišťují dostatečně velké potlačení souhlasné složky podobně jako plně diferenční struktury, ale oproti nim je obvodové řešení pseudodiferenčního filtru méně složité.

3. NAVRŽENÝ PSEUDODIFERENČNÍ FILTR

Pro návrh pseudodiferenčního filtru byl použit aktivní prvek *DDCC* [3], jehož schematická značka je uvedena na Obr. 1a). Vztahy mezi branovými proudy a napětími je: $U_{\mathrm{X}} = U_{\mathrm{Y1}} - U_{\mathrm{Y2}} + U_{\mathrm{Y3}}$, $I_{\mathrm{Z1+}} = -I_{\mathrm{Z1-}} = I_{\mathrm{X}}$, $I_{\mathrm{Y1}} = I_{\mathrm{Y2}} = I_{\mathrm{Y3}} = 0$. Navržený pseudodiferenční filtr druhého řádu je na Obr. 1b), který využívá tři *DDCC* a pět uzemněných pasivních prvků.



Obrázek 1: a) Schematická značka aktivního prvku *DDCC*, b) navržený pseudodiferenční filtr

Navržený filtr realizuje všechny typy základních kmitočtových odezev. Uvažujeme-li ideální aktivní prvky, tak pro horní propust jednotlivá výstupní napětí jsou:

$$U_{21+} = \frac{p^2 C_1 C_2 R_1 R_2 + 1}{p^2 C_1 C_2 R_1 R_2 + p C_2 R_2 + 1} \cdot U_{1\mathrm{d}} \quad U_{21-} = \frac{1}{p^2 C_1 C_2 R_1 R_2 + p C_2 R_2 + 1} \cdot U_{1\mathrm{d}}, \qquad (4a,b)$$

potom pro diferenční výstupní napětí $U_{21\mathrm{d}}$ s využitím (1) platí:

$$U_{21\mathrm{d}} = U_{21+} - U_{21-} = \frac{p^2 C_1 C_2 R_1 R_2}{p^2 C_1 C_2 R_1 R_2 + p C_2 R_2 + 1} \cdot U_{1\mathrm{d}} \qquad (4c)$$

Ostatní odezvy kmitočtového filtru ve srovnání s horní propustí se upravují obdobně, tudíž pro ně budou vypsány jenom diferenční výstupní napětí:

$$U_{22\mathrm{d}} = \frac{1}{p^2 C_1 C_2 R_1 R_2 + p C_2 R_2 + 1} \cdot U_{1\mathrm{d}} \quad U_{23\mathrm{d}} = \frac{2 \cdot p C_2 R_2}{p^2 C_1 C_2 R_1 R_2 + p C_2 R_2 + 1} \cdot U_{1\mathrm{d}}, \qquad (4d,e)$$

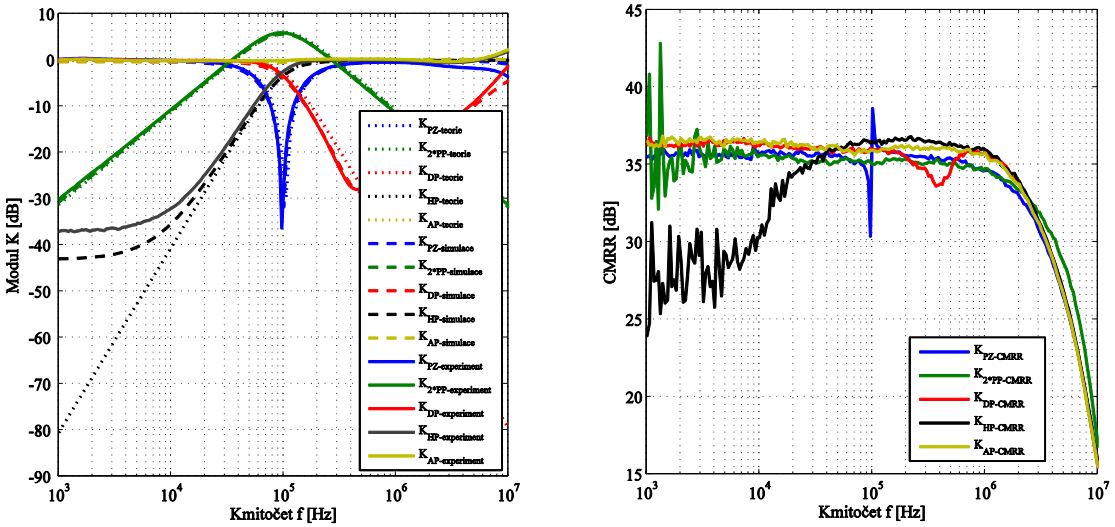
$$U_{24\mathrm{d}} = \frac{p^2 C_1 C_2 R_1 R_2 + 1}{p^2 C_1 C_2 R_1 R_2 + p C_2 R_2 + 1} \cdot U_{1\mathrm{d}} \quad U_{25\mathrm{d}} = \frac{p^2 C_1 C_2 R_1 R_2 - p C_2 R_2 + 1}{p^2 C_1 C_2 R_1 R_2 + p C_2 R_2 + 1} \cdot U_{1\mathrm{d}}. \qquad (4f,g)$$

Obvodové řešení z Obr. 1b) je možné použít k realizaci horní (4c), dolní (4d), pásmové propusti (4e), pásmové zádrže (4f) u které se předpokládá $R_1 = R_3$ a fázovacího článku (4g).

4. SIMULACE A EXPERIMENTÁLNÍ MĚŘENÍ

Vybrané diferenční přenosové funkce získané simulacemi jsou zobrazeny na Obr. 2a). Kmitočtový filtr byl navržen pro charakteristický kmitočet f_0 = 100 kHz, činitel jakosti Q = 0,707. Hodnoty pasivních prvků jsou $C_1 = C_2$ = 1 nF, $R_1 = R_3$ = 1100 Ω a R_2 = 2200 Ω Aktivní prvek *DDCC* byl pro

simulace a experimentální měření realizován pomocí UCC-N1B [4]. Obr. 2b) znázorňuje schopnost pseudodiferenčního filtru potlačit souhlasný signál.



a) b)

Obrázek 2: a) Diferenční přenosové funkce, b) potlačení souhlasného signálu

K výsledkům experimentálního měření pseudodiferenčního filtru uvedených na Obr. 2a) můžeme říci, že v porovnání simulací s teorií se průběhy téměř shodují. Na Obr. 2b) je pak uveden parametr *CMRR*, který pro všechny typy uvažovaných přenosových funkcí dosahuje hodnoty cca 36 dB, což pro řadu aplikací je možné považovat za dostatečné.

5. ZÁVĚR

V tomto článku je prezentován pseudodiferenční filtr, který se především vyznačuje vysokou vstupní impedancí a vysokým potlačením souhlasného signálu. Jak simulační tak experimentální měření se shodují s teoretickými analýzami. Vyhodnocením CMRR bylo ukázáno, že tyto struktury kmitočtových filtrů dosahují dostatečného potlačení souhlasného signálu a pseudodiferenční filtry tak lze považovat za vhodné funkční bloky pro zpracování analogových signálů.

PODĚKOVÁNÍ

Výzkum popsaný v tomto článku byl podpořen projektem č. LO1401 Národního programu udržitelnosti a projektem č. 16-11460Y Grantové agentury České republiky. Pro výzkum byla použita infrastruktura Centra SIX.

REFERENCE

[1] Tietze, U., Schenk, Ch., Gramm, E.: Electronic Circuits: Handbook for Design and Apllications, Springer, Berlin, 2008, s 1543.

[2] Horng, J., W., Wu, C., M., Herencsar, N.: Fully differential first-order allpass filters using a DDCC, Indian J. Engineeringand Materials Sciences, roč. 21, 2014, s. 345-350.

[3] Chiu, W., Y., Horng, J., W.: Voltage-Mode Highpass, Bandpass, Lowpass and Notch Biquadratic Filters Using Single DDCC, Radioengineering, roč. 21, 2012, s. 297-303.

[4] Datasheet UCC-N1B - Universal Current Conveyor (UCC) and Second-Generation Current Conveyor (CCII+/-), Brno University of Technolohy, On Semiconductor Ltd., Rev. 1, 2012

WIRELESS CONTROL OF EMBEDDED M2M DEVICES USING UNLICENSED ISM BAND

Martin Štůsek

Master Degree Programme (2), FEEC BUT

E-mail: xstuse01@stud.feec.vutbr.cz

Supervised by: Pavel Mašek

E-mail: masekpavel@feec.vutbr.cz

Abstract: This research work follows the idea of Machine-to-Machine (M2M) communication designing the radio controlled devices using the ISM radio band of 433 MHz. Implemented solution consists of transmitter unit (Raspberry Pi) and sets of receiver units based on Arduino Uno platform. To offer better user experience, simple management interface on the side of transmitter is imple- mented.

Keywords: Arduino Uno, ISM 433 MHz, Machine-to-Machine (M2M), OSGi, Raspberry Pi

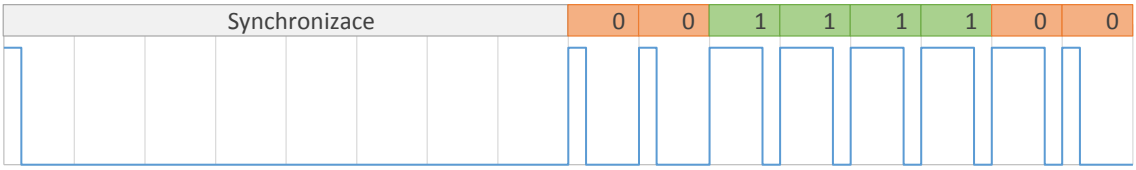
1. ÚVOD

Tato práce se zabývá bezdrátovým ovládáním elektronických zařízení s využitím jednodeskového počítače Raspberry Pi a 8-bitového mikrokontroleru Arduino Uno. Raspberry Pi je použito jako server s vysílací částí zajišťující interakci s uživatelem. Mikrokontroler Arduino slouží jako klient pro přijímání dat umožňující ovládání elektrických zařízení. K bezdrátové komunikaci je využito vysílačů/přijímačůpracujících na frekvenci 433 MHz.

2. IMPLEMENTACE BEZDRÁTOVÉHO PROTOKOL

Jelikož na frekvenci 433 MHz komunikují zejména zařízení s nízkým výpočetním výkonem tzv. embedded zařízení, je při komunikaci mezi zařízeními využito jednoduchého klíčování typu OOK (On- Off Keying). Komunikace v pásmu ISM (Industrial, Scientific and Medical) na 433 MHz je bezlicenční, a proto je využívána velkým množstvím zařízení. Z důvodu možného rušení souběžně komunikujícími zařízeními je ale nutné každý povel vyslat opakovaně. Nejmenší doporučený počet opakování je tři, z důvodu velkého počtu komunikujících zařízení není ovšem tato hodnota dostačující – proto je v této implementaci každý kód odeslán desetkrát.

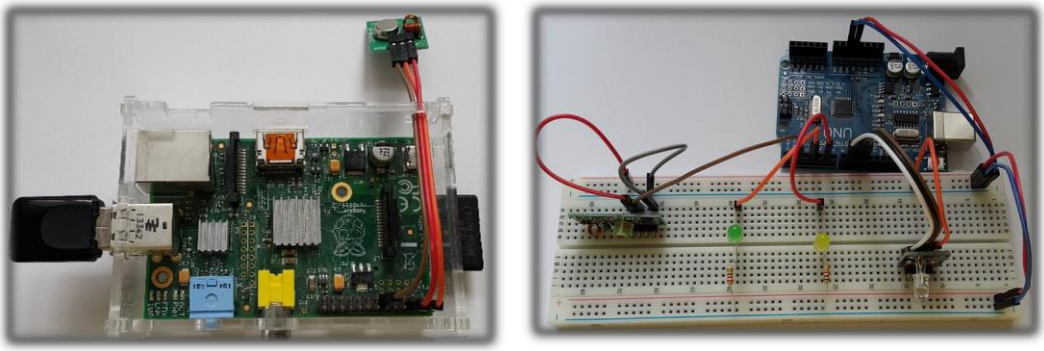
Implementovaný protokol vychází z kódování využívaného v čipech EV1527, kdy jsou využity pulzy o délce 300 až 500 μs. Jak lze vidět na Obr. 1, přenos je zahájen jedním pulzem vysoké úrovně a 31 pulzy nízké úrovně. Po synchronizační sekvenci jsou přenášena již samotná data. `Log 1` je vyjádřena třemi pulzy vysoké úrovně a jedním pulzem nízké úrovně. Naopak `log 0` je definována jedním pulzem vysoké úrovně a třemi pulzy nízké úrovně. Pro adresaci zařízení je využito protokolu s délkou 8 bitů– lze tedy adresovat až 255 zařízení (adresu 0 nelze využít). Pro přenos samotných příkazůje využit 32bitový protokol umožňující nastavení každé složky RGB LED v 255 úrovních [1].



Obrázek 1: Průběh 8 bitové zprávy při přenosu čísla „62“.

3. VYSÍLACÍ Č ÁST

Jako zařízení vysílající příkazy je využito Raspberry Pi model B gen. 1, které je s vysílačem propojeno skrze GPIO (General-purpose input/output), viz Obr. 2a. Vysílací jednotka také implementuje aplikaci zajišťující správnou interpretaci a převod příkazůz uživatelských vstupů. Dále je vytvořen webový server poskytující uživatelské rozhraní pro ovládání připojených zařízení.



(a) Vysílací část. (b) Přijímací část.

Obrázek 2: Zapojení zařízení pro testování.

3.1. BACKEND

Celý sytém je napsán v jazyce Java s využitím OSGi (Open Service Gateway Initiative) frameworku Knopflerfish. Standard OSGi definuje sadu doporučení pro dynamický systém balíčků(bundles) v jazyce Java. Tento systém umožňuje instalovat, spouštět či zastavovat jednotlivé balíčky bez nutnosti restartu programu. Základní filozofií OSGi je tvořit jednoduché aplikace, které poskytují jednu, přesně definovanou, službu například pro HTTP (Hypertext Transfer Protocol) komunikaci.

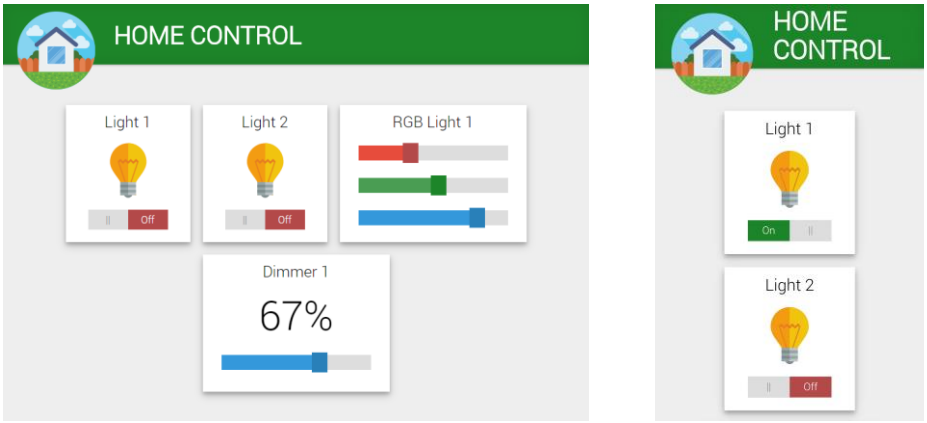
Základními prvky architektury OSGi jsou tzv. `bundles`, což je sada tříd rozšířená o speciální soubor `manifest` – poskytuje dodatečné informace pro OSGi class loader. OSGi definuje také sadu služeb, která umožňuje komunikaci mezi jednotlivými balíčky, v této práci je to zejména vrstva služeb (`service layer`) a `OSGi Event Admin` [2].

Výsledná aplikace je rozdělena do tří samostatných balíčků:

- **ItemsRegistry** – definuje datovou strukturu položek a udržuje v paměti seznam všech položek s jejich aktuálními hodnotami. Tento balíček musí být spuštěn jako první, v opačném případě není možné spustit zbylé balíčky, které jsou na něm závislé. Jako registr položek slouží objekt typu `ConcurentHashMap`, který poskytuje bezpečný přístup k položkám i ve vícevláknových aplikacích.
- **Transmitter** – stará se o zpracování přijatých příkazůa samotné vyslání povelů. Komunikace s GPIO portem je realizována pomocí knihovny Pi4J [3], která umožňuje přímí přístup na GPIO rozhraní z prostředí Java.
- **Web API** – zajišťuje komunikaci s webovým uživatelským rozhraním. Ke komunikaci je využito protokolu WebSocket, který je základní součástí jazyka HTML5 (HyperText Markup Language). Na základě vlastností protokolu WebSocket je komunikace mezi serverem a uživatelem velmi rychlá bez viditelného zpoždění, které je patrné u technologií využívající pooling. Data jsou následně přenášena jako objekty typu JSON (JavaScript Object Notation).

3.2. UŽIVATELSKÉ ROZHRANÍ

Aplikaci lze ovládat skrze vytvořené uživatelské rozhraní, viz Obr. 3. Prvek `Light` slouží jako spínač se dvěma stavy. `Dimmer` umožňuje nastavit výstupní hodnotu v rozsahu 0 až 255 (vhodné pro stmívače) a `RGB` umožňuje ovládat jednotlivé složky RGB diody.



Obrázek 3: Uživatelské rozhraní aplikace.

4. PŘÍJÍMACÍ ČÁST

Jak bylo zmíněno v úvodní části, jako přijímací část slouží mikrokontroler Arduino Uno s rádiovým přijímačem, viz Obr 2b. Příjem dat je signalizován externím přerušením, z tohoto důvodu je nutné datový pin přijímače připojit k pinu 2 nebo 3 – u mikroprocesor Atmega 328p mohou pouze piny 2 a 3 vyvolat externí přerušení.

Maximální délka přenesených dat je z důvodu využití datového typu `unsigned long` pro příkazy omezena na 32 bitů. Není také možné přenést nulová data, tzn. v sekvenci dat se musí vyskytovat alespoňjeden bit logické hodnoty 1. Nulová data by totiž mohla být snadno zaměněna se šumem.

Zpracování příkazůse skládá ze dvou částí, nejprve jsou vyslána data obsahující adresu zařízení a v další fázi jsou přenesena data s příkazem. Pokud jsou přijata data pro zařízení typu `switch` je na výstup digitálního pinu zapsána logická hodnota 0 nebo 1, dle typu příkazu. Pro zařízení typu `dimmer` či `RGB`, jsou použity analogové výstupní piny v režimu PWM (Pulse Width Modulation).

5. ZÁVĚR

V této práci byl proveden popis charakteristiky bezdrátového ovládání zařízení a jednoduchého protokolu pro bezdrátovou komunikaci. Dále byla popsána implementace vysílací části postavené na Raspberry Pi a přijímací části využívající zařízení Arduino Uno.

Důraz byl kladen na (i) popis vysílací části, která obsahuje logiku pro zpracování a vyslání příkazů, (ii) rozhraní pro komunikaci skrze protokol WebSocket a (iii) vytvořený webový server s uživatelským rozhraním.

REFERENCE

[1] KnowHow_LineCoding. *GitHub* [online]. 2016 [cit. 2016-03-02]. Dostupné z: https://github.com/sui77/rc-switch/wiki/KnowHow_LineCoding

[2] OSGITM ALLIANCE. *OSGi: The Dynamic Module System for Java* [online]. 2015 [cit. 2015-11-22]. Dostupné z: https://www.osgi.org

[3] *The Pi4J Project: Connecting Java to the Raspberry Pi* [online]. 2015 [cit. 2016-03-02]. Dostupné z: http://pi4j.com/

DESIGNING A DIRECTIONAL ACOUSTIC TRANSMITTER

David Hladký

Master Degree Programme (5), FEEC BUT

E-mail: xhladk08@stud.feec.vutbr.cz

Supervised by: Jan Mikulka

E-mail: mikulka@feec.vutbr.cz

Abstract In this paper the parametric transmission of sound through an amplitude-modulated ultrasonic wave, using automatic demodulation in a non-linear environment is discussed. Methods for the processing of the transmitted signal are presented; these techniques employ modified amplitude modulation to suppress the environmental effects, which otherwise distort the signal. A special emphasis is placed on the mathematical instruments describing the parametric sound beam transmission in a non-linear medium. Within our research, we fabricated a parametric speaker consisting of 50 ultrasonic acoustic transducers with the resonant frequency of 40 kHz. The functionality of the acoustic source was verified via measurement of the spatial radiation characteristics and through determination of the total harmonic distortion (THD). In order to control the parametric speaker, a digital element comprising a sampler, a transducer of the input transmitted signal, and a microcontroller to preprocess this signal and to facilitate the amplitude modulation of the ultrasonic carrier wave is designed. The applied algorithms were analyzed to establish their computational intensity; such an analysis enabled us to select a microprocessor basis suitable for the implementation of the signal processing algorithms.

Keywords: Parametric acoustic array, parametric loudspeaker, amplitude modulation, STM32F411

1. ÚVOD

V lineární akustice, lze jen s velkými obtížemi realizovat zvukové zařízení, které dokáže směrově vysílat zvukovou vlnu. Aby bylo možné vyslat v určitém směru zvukovou vlnu, je nutné využít jevů vyskytujících se v nelineární akustice, zejména jevu, který se nazývá parametrické akustické pole (PAP). Pomocí PAP je možné vysílat zvukovou vlnu v úzkém svazku do určitého místa v místnosti anebo určité skupině osob. Tento fenomén se využívá v zařízeních pro přenos parametrického (směrového) zvuku. Tato zařízení nalézají uplatnění především ve veřejných prostorách, jako jsou galerie, odchodní centra a muzea, kde je potřeba předat zvukovou informaci pouze určité osobě či skupině osob. Výhodou těchto zařízení je, že dokáží přenášet zvuk s vysokou směrovostí na rozdíl od klasické reproduktorové soustavy. Značnou nevýhodou zmíněných zařízení je jejich celkové harmonické zkreslení (THD) vysílané akustické vlny a tudíž nutnost předzpracovat vstupní signál. Při konstrukci těchto zařízení je potřeba použít amplitudový modulátor a zařízení pro přenos parametrického zvuku tzv. parametrický reproduktor, který se skládá ze značného množství ultrazvukových akustických měničů. Cílem tohoto článku je popsat základní princip parametrické akustiky a technik předzpracování signálu pomocí amplitudové modulace a předložit konkrétní návrh řešení zařízení pro parametrický přenos zvuku.

2. PRINCIP PARAMETRICKÉHO AKUSTICKÉHO POLE A PARAMETRICKÉ AKUSTIKY

Prvním, kdo teoreticky popsal princip parametrické akustiky, byl v roce 1963 P. J. Westervelt [1]. Svým teoretickým popisem objasnil, jakým způsobem dochází ke vzniku slyšitelné zvukové složky ze dvou ultrazvukových vln o úhlovém kmitočtu ω_1 a ω_2, přičemž $\omega_1 > \omega_2$. Jestliže se

z intenzivního ultrazvukového zdroje vygenerují dvě ultrazvukové vlny o úhlovém kmitočtu ω_1, ω_2 a jsou následně vyslány skrze nelineární prostření (např. vzduch), dochází k jejich vzájemné interakci a vznikají sekundární vlny, které tvoří součet a rozdíl původních dvou vln. Výsledkem nelineární interakce je rozdílová složka, odpovídající rozdílu ω_1 - ω_2, která se síří ve stejném směru jako původní vlny a zachová si jejich směrovost (čím má zvukový svazek vyšší frekvenci, tím více je užší), což je výhodou ve srovnání s lineární akustikou. Výše zmíněný jev se nazývá parametrické akustické pole.

Z teoretického popisu PAP vyšel H. O. Berktay [2] a provedl jednoduchý pokus, kdy z ultrazvukového zdroje vyslal amplitudově modulovaný signál. Výsledkem tohoto experimentu bylo, že autodemodulovaný signál $p(\tau)$ v jisté vzdálenosti od zdroje ultrazvukového vlnění odpovídá přibližnému řešení Berktayově rovnice (1). Tedy, že autodemodulovaný signál je roven druhé derivaci kvadrátu obálkové funkce $E(\tau)$ amplitudově modulovaného signálu (AM). V rovnici (1) je symbol β koeficient nelinearity prostředí (pro vzduch β = 1,2), d značí průměr apertury zdroje ultrazvuku, r je vzdálenost od zdroje ultrazvuku, p_0 odpovídá akustickému tlaku nosného signálu generované z ultrazvukového zdroje, c_0 označuje rychlost šíření zvuku, α_0 vyjadřuje absorpční koeficient nosného signálu na dané frekvenci. Hustotu prostředí označuje symbol ρ_0, τ odpovídá retardovanému času.

$$p(\tau) \approx \frac{\beta p_0^2 \mathrm{d}^2}{16\rho_0 c_0^4 r \alpha_0} \frac{d^2}{d^2\tau^2} E^2(\tau). \tag{1}$$

3. TECHNIKY AM PRO PŘEDZPRACOVÁNÍ SIGNÁLU

Z Berktayovy rovnice (1) je patrné, že výsledný autodemodulovaný signál neodpovídá původnímu signálu a je značně zkreslený v důsledku druhé derivace a kvadrátu obálkové funkce $E(\tau)$ AM, výše popsaný jev je způsoben nelinearitou prostředí. Hlavním úkolem technik pro předzpracování vstupního signálu je upravit obálkovou funkci $E(\tau)$ tak, aby bylo dosaženo co možná nejmenšího zkreslení THD výsledného autodemodulovaného signálu. S ohledem na dosažené výsledky, především malá výpočetní náročnost a nízké THD, byly vybrány pro ověření funkce zařízení dvě amplitudové modulace a to jednostranná AM bez potlačené nosné (SSB-C) viz rovnice (2) a pokročilá AM [3] [4] podle rovnice (3). Každá z uvedených technik AM upravuje jiným způsobem modulační obálku $E(\tau)$. Obě dvě modulace vycházejí z kvadraturní AM. Symbol m v rovnici (2) a (3) značí hloubku modulace, signál $g(\tau)$ označuje vstupní signál, ω_c je úhlový kmitočet nosného signálu. V rovnici (2) vystupuje signál $\hat{g}(\tau)$, který odpovídá vstupnímu signálu $g(\tau)$ po Hilbertově transformaci.

$$u_{\mathrm{AM}}(\tau) = m\left[g(\tau)\cos(\omega_c\tau) \pm \hat{g}(\tau)\sin(\omega_c\tau)\right] + \cos(\omega_c\tau), \tag{2}$$

$$u_{\mathrm{AM}}(\tau) = \left[1 - \frac{1}{2}m^2 g^2(\tau) - \frac{1}{8}m^4 g^4(\tau)\right]\sin(\omega_c\tau) + \left[1 + mg(\tau)\right]\cos(\omega_c\tau). \tag{3}$$

4. ZAŘÍZENÍ PRO PŘENOS PARAMETRICKÉHO ZVUKU

Z předložených poznatků bylo zkonstruováno zařízení pro přenos parametrického zvuku. Zařízení se skládá ze tří základních částí. První část tvoří digitální amplitudový modulátor, v němž jsou implementovány výše uvedené amplitudové modulace, které je možné pomocí přepínačem přepnout, v závislosti na měření vlivu a zkreslení dané AM. Hilbertova transformace v rovnici (2), byla implementována pomocí FIR filtru typu III. Základní části digitálního modulátor tvoří A/D převodník typu MCP3201, dále pak mikrokontrolér řady STM32F411 a výstup je zpětně rekonstruován pomocí D/A převodníku typu MCP4291. Druhá část se sestává z výkonového audio zesilovače s TDA7293, který pracuje ve třídě AB. Posledním stupněm je parametrický reproduktor, který se

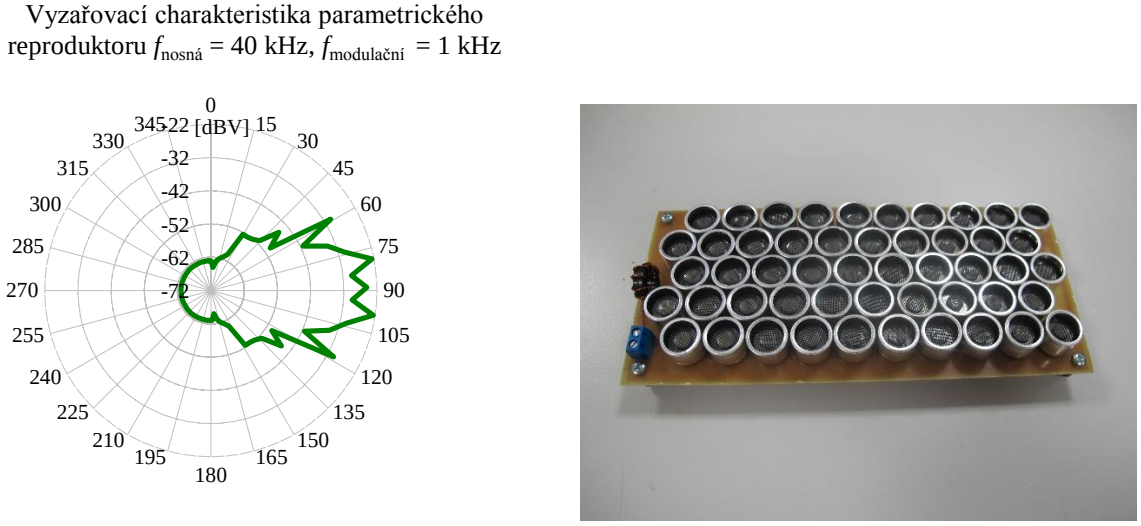
sestává z matice 50 ultrazvukových měničů typu BPU-1640TOAH12, jenž jsou poskládány těsně vedle do matice 5 × 10. Celkové blokové schéma je uvedeno, viz obrázek 1.



Obrázek 1: Blokové schéma digitálního amplitudového modulátoru

5. MĚŘENÍ VYZAŘOVACÍ CHARAKTERISTIKY PARAMETRICKÉHO REPRODUKTORU

Vyzařovací charakteristika parametrického reproduktoru, viz obrázek 2**,** byla změřena pomocí digitálního osciloskopu Agilent DSOX3034A, který byl použit jako spektrální analyzátor, ze kterého se odečítal útlum přijaté modulační složky vyzářeného zvuku z parametrického reproduktoru. Parametrický reproduktor byl umístěn ve vzdálenosti 1,5 m od elektretového mikrofonu typu PMOF-6027PN. Signál z mikrofonu byl zesílen pomocí dvou kaskádně řazených operačních zesilovačů typu OPA2604, celkové zesílení signálu bylo 61,5 dBV. Zesílený signál byl následně veden do výše zmíněného osciloskopu. Ke generování signálu pro parametrický reproduktor byl použit generátor Agilent 33511B, který měl v sobě implementovanou pokročilé AM s nosným kmitočtem 40 kHz a modulačním kmitočtem 1 kHz. Měření probíhalo v bez odrazové komoře na Ústavu teoretické a experimentální elektrotechniky.



Obrázek 2: Vyzařovací charakteristika (vlevo) a fyzická realizace parametrického reproduktoru (vpravo)

6. ZÁVĚR

V předloženém článku byl prezentován teoretický popis parametrického akustického pole a parametrické akustiky, dále pak techniky AM pro předzpracovaní vstupního signálu pro parametrický reproduktor a konkrétní řešení digitálního AM modulátoru v podobě blokového schématu. Poslední část článku uvádí praktické měření vyzařovacích charakteristiky parametrického reproduktoru. Při bližším pohledu na vyzařovací charakteristiku, viz obrázek 2**,** jsou patrné výběžky na úhlech 60° a 120°, tyto výběžky jsou pravděpodobně způsobeny odrazy od stěn komory. Další měření budou provedena v otevřeném prostoru, aby se potlačil výše zmíněný jev.

REFERENCE

[1] WESTERVELT, Peter J. *Parametric Acoustic Array*. B.m.: Department of Physics, Brown Univesity, Providence, Rhode Island. 1965

[2] WOON-SENG, Gan, Yang JUN a Kamakura TOMOO. A Review of Parametric Acoustic Array in Air. *Applied Acoustics*. 2012, roč. 12, s. 1211–1219. ISSN 0003-682X.

[3] WOON-SENG, Gan, Ee-Leng TAN a Sen M. KUO. *Audio projection* [online]. 2010. Dostupné z: http://eeeweba.ntu.edu.sg/DSPLab/audiobeam/docs/sp.pdf

[4] TAN, Ee-Leng, Woon-Seng GAN, PeiFeng JI a Jun YANG. *Distortion Analysis and Reduction for the Parametric Array* [online]. 2010. Dostupné z: http://eeeweba.ntu.edu.sg/dsplab/audiobeam/Publication/2008_C3.pdf

COMPARISON OF MULTI-OBJECTIVE OPTIMIZATION METHODS

Martin Marek
Master Degree Programme (2), FEEC BUT
E-mail: xmarek54@stud.feec.vutbr.cz

Supervised by: Petr Kadlec
E-mail: kadlecp@feec.vutbr.cz

Abstract: This paper deals with comparison of multi-objective optimization methods. Basic properties of multi-objective optimization are explained here. Algorithms NSGA-II, MOPSO and GDE3 are briefly introduced and compared using performance metrics on several test functions.

Keywords: Multi-objective, optimization, NSGA-II, MOPSO, GDE3

1 INTRODUCTION

Multi-objective optimization (MOOP) is a process of finding set of optimal solutions (known as Pareto-optimal solutions) when objective functions of solved problem are conflicting. Pareto-optimal solutions express the trade-off between individual objectives. The decision whether one solution is better than the other or not is made using fitness functions values.

In the multi-objective optimization, there are two goals: To find a set of solutions as close as possible to the Pareto-optimal front. To find a set of solutions as diverse on the Pareto-optimal front as possible. First goal is identical with single-objective optimization task. Second one is specific for multi-objective optimization only.

2 OPTIMIZATION METHODS

In this section NSGA-II [1], MOPSO [2] and GDE3 [3] optimization methods are briefly described.

2.1 NSGA-II

Elitist Non-dominated Sorting Genetic Algorithm [1] uses non-dominated sorting [4] on combined parent and offspring population of $2N$ members. In a selection process, only N members are selected to form next parent population.

After selection process, new parent population is submited to crossover and mutation operations to create new offspring population. Whole process is repeated for given number of iterations.

2.2 MOPSO

Multi-objective Particle Swarm Optimization [2] simulates the movement of swarm of bees. At the beginning, agent's positions and velocities are randomly assigned. Then, non-dominated solutions are inserted to an external archive. For each agent personal best (*pbest*) and global best (*gbest*) positions [2] are assigned. In the next step, velocity vectors are computed to find a new agent's positions.

New positions can be out of decision variable's ranges, thus boundary conditions has to be applied. After this step, content of external archive and agent's personal and global best positions are updated. Whole process is repeated for given number of iterations.

2.3 GDE3

Third Version of Generalized Differential Evolution [3] creates trial vector $\vec{u}$ from vector of decision variables $\vec{x}$ according to equation (1).

$$u_{j,i,g} = x_{j,r_3,g} + F \cdot (x_{j,r_1,g} - x_{j,r_2,g}), \tag{1}$$

where j denotes j-th decision variable, i denotes i-th agent, g denotes g-th generation, r_1, r_2, r_3 are randomly chosen agent's indexes and F is scaling factor [3].

Afterwards $\vec{u}_{i,g}$ or $\vec{x}_{i,g}$ or both are selected to create new vector of decision variables $\vec{x}_{i,g+1}$ based on dominance relation. If both vectors are selected, overall number of solutions is larger then N, thus crowding comparison [1] has to be used to select most promising solutions.

3 COMPARATIVE STUDY

Efficiency of the above described methods was tested using several test functions. Comparison between optimization algorithms was based on Generational Distance (GD) [4], Spread (Δ) [4] and Hypervolume metrics [4]. Another considered parameter was computational complexity of algorithms.

Test functions used for algorithm comparison are: Schaffer's study (SCH), Fonseca's and Fleming's study (FON), Kursawe's study (KUR), Poloni's study (POL), Zitler-Deb-Thiele's: ZDT1, ZDT2, ZDT3, ZDT4 and ZDT6 [4][1].

3.1 SIMULATION RESULTS

Table 1 contains average values from 20 simulation runs. Values of computational time, generational distance, spread and hypervolume were taken after 200 iterations of populations with 200 individuals for each test function. It is not possible to localize true Pareto-optimal front for KUR and POL testing functions. Thus, generational distance cannot be calculated.

The algorithms were for purposes of our comparative study set as follows:

NSGA-II settings:	Probability of crossover - $P_C = 1$,
	Probability of mutation - $P_M = 0.5$,
	Size of mating pool (tournament selection) - $\|MP\| = 2$,
	Binary precision - $BP = 20$.
MOPSO settings:	Cognitive learning factor - $C_1 = 1.5$,
	Social learning factor $C_2 = 1.5$,
	Inertia weight - linearly decreasing from $W = 0.9$ to $W = 0.5$.
GDE3 settings:	Scaling factor - $F = 0.2$,
	Probability of crossover - $CR = 0.4$.

NSGA-II's computational time is quite stable for every test function. Small time variances between individual problems are caused by different number of decision variables. In MOPSO, if small number of non-dominated solutions is found, non-dominated sorting is performed on smaller population (in NSGA-II strictly $2N$) and diversity maintanance is not required which results in smaller computational effort. GDE3's computational time instability has similar reason as MOPSO's, but it is less apparent.

Generational distance of MOPSO and GDE3 for ZDT4 test function is quite large. Both algorithm converged to local Pareto-optimal front, which is main difficulty with ZDT4 test function.

Hypervolume values are diverse for each test function due to different sizes of objective space.

	SCH	FON	KUR	POL	ZDT1	ZDT2	ZDT3	ZDT4	ZDT6
Computational time (s)									
NSGA-II	14.551	14.202	14.33	14.28	16.553	15.918	16.599	14.193	14.03
MOPSO	3.521	2.803	2.134	4.753	1.629	0.8195	1.395	0.334	0.387
GDE3	11.737	10.13	9.666	8.979	6.006	5.663	6.045	6.211	6.347
Generational Distance (-)									
NSGA-II	1.65E-5	6.96E-4	-	-	1.15E-2	0.267	3.4E-3	2	0.251
MOPSO	2.43E-5	2.10E-3	-	-	1.24E-2	0.238	1.13E-2	7.64	4.8
GDE3	1.73E-5	3.14E-4	-	-	2.33E-2	0.176	1.08E-2	26.7	0.484
Spread (-)									
NSGA-II	0.139	0.110	0.191	0.109	0.513	0.588	0.577	0.662	0.704
MOPSO	0.084	0.300	0.849	0.184	1.439	1.187	1.438	1.042	1.099
GDE3	0.119	0.100	0.185	0.119	0.368	0.792	0.484	0.782	0.755
Hypervolume (-)									
NSGA-II	12.639	0.297	25.84	359.5	0.649	0.308	0.778	1.280	0.204
MOPSO	13.307	0.298	26.78	377.9	0.284	0.094	0.501	1.4E-5	0.906
GDE3	13.303	0.303	26.13	357.3	0.674	0.128	0.789	0.016	0.205

Table 1: Simulation results

4 CONCLUSION

Resulting values from simulation shows that although MOPSO is the fastest algorithm it is outperformed by other two algorithms in almost every cases and for complex test functions (ZDT) shows poor diversity maintanance. Algorithms GDE3 and NSGA-II show similar results for easier test function such as SCH, FON, KUR and POL, but results from more complex test functions (ZDT family) are in favor of NSGA-II. NSGA-II converges systematically to true Pareto-optimal front in wide spread of solutions (ZDT test functions need more time to scour vast decision variable space), but GDE3 can converge closer to true Pareto-optimal front if satisfactory number of individuals and iterations is used (SCH, FON and KUR test functions are sufficiently searched through with comparative study settings).

ACKNOWLEDGEMENT

This work has been supported by the Technology Agency of the Czech Republic, grant TA 04010457 "Tools for synthesis of antennas and sensors".

REFERENCES

[1] Deb, K., Pratap, A., Agarwal, S.: A fast and elitist multiobjective genetic algorithm: NSGA-II. *Evolutionary Computation, IEEE Transactions on*, 2002, 6.2: 182-197.

[2] Reyes-Sierra, M., Coello, C. C.: Multi-objective particle swarm optimizers: A survey of the state-of-the-art. *International journal of computational intelligence research*, 2006, 2.3: 287-308.

[3] Kukkonen, S., Lampinen, J.: GDE3: The third evolution step of generalized differential evolution. In: *Evolutionary Computation, 2005. The 2005 IEEE Congress on*. IEEE, 2005. p. 443-450.

[4] Deb, K.: Multi-objective optimization using evolutionary algorithms. John Wiley & Sons, 2001.

Magisterské projekty

Zpracování signálů, obrazu a dat

ALGORITHM FOR DETECTION OF POSITIVE AND NEGATIVE TEXT

David Musil
Master Degree Programme (2), FEEC BUT
E-mail: xmusil49@stud.feec.vutbr.cz

Supervised by: Lukáš Povoda
E-mail: xpovod00@phd.feec.vutbr.cz

Abstract: In the present, obtaining and sorting knowledge from data produced by various sources requires significant effort which is not ensured easily by a human, meaning machine processing is taking place. Purpose of this work was to create a system capable of positive and negative emotion detection from text along with evaluation of its performance. System allows training with use of large amount of data (known as Big Data), exploiting Spark library. Classificator model was created with use of Support Vector Machines. Highest achieved accuracy is 78,05% for Czech, 79,73% for German and 91,88% for English.

Keywords: artificial intelligence, Big Data, emotion detection, text-mining

1 ÚVOD

Dolování znalostí, zejména emocí z textových dat, je poměrně rozsáhlým oborem, těšícím se v současné době značné popularitě. Emoce mají značný dopad na lidské vztahy v mnoha aspektech každodenního života. V moderní společnosti nejen mezilidská komunikace, ale i komunikace mezi člověkem a přístrojem, tíhne k přesunu do světa počítačů a internetu. Množství textových dat drasticky stoupá s nárůstem využití současných trendů v komunikačních technologiích.

Cílem této práce je vytvoření systému, který je schopen identifikovat pozitivní a negativní emoce v jemu předloženém textu, a sice takovým způsobem, aby byla zajištěna možnost jeho trénování pomocí velkých objemů dat (Big Data, přiblíženo v článku [1]), konkrétně s využitím knihovny Spark. Jedná se o víceúčelový clusterový výpočetní framework pod open source licencí vhodný pro zpracování velkých objemů dat.

První kapitola článku představuje hlavní přínos práce a základní poznatky o problematice, druhá kapitola zahrnuje aktuální situaci v oblasti zpracování textových dat a uvádí některé podobné práce. Třetí kapitola objasňuje proces zpracování textu včetně popisu algoritmu umělé inteligence. Čtvrtá kapitola uvádí systémem dosažené výsledky, pátá kapitola nabízí rekapitulaci práce a její zhodnocení.

2 SOUVISEJÍCÍ PRÁCE

Přístup k úkolu dolování znalostí za účelem rozpoznání emocí v textu se u různých prací může značně lišit. Známou možností je vyhledávání pomocí klíčových slov, jež použili například autoři článku [2], kde se nejvyšší dosažená úspěšnost pohybovala okolo 75%. Takovýto systém detekuje emoce za pomoci seznamů slov roztříděných do specifických kategorií. Další přístup využívá umělé inteligence jako kupříkladu v článku [3], kde nejúspěšnější ze série experimentů dosáhl na úspěšnost 86% pomocí algoritmu podpůrných vektorů. V experimentu [4] autoři provádí analýzu sentimentu v datech s využitím algoritmů Naive Bayes a SVM, přičemž maximální úspěšnost se pohybuje v rozmezí 89% až 94%. V neposlední řadě existují hybridní techniky, kde jsou předešle zmíněné přístupy kombinovány. V nedávné době bylo provedeno množství výzkumů na datech z populární sociální sítě Twitter,

kupříkladu článek [5], kde je pro analýzu emocí v příspěvcích využito několik různých typů klasifikátoru včetně Naive Bayes, SVM, k-nejbližších sousedů a rozhodovacích stromů. Za zmínku dále stojí zrychlování samotných výpočtů jako u experimentu [6], kde byl výpočet urychlen algoritmem k-nejbližších sousedů a kombinací více grafických karet až 750-násobně oproti samotnému procesoru.

3 POPIS FUNKCE SYSTÉMU

Pro projekt byla použita databáze textů, ze které jsou čerpána data jak pro trénování, tak pro testování. Jedná se o komentáře uživatelů z internetových serverů, odrážející jejich subjektivní názor na konkrétní produkt nebo dílo, jde tedy o nestrukturovaný text. Každý z komentářů je celistvý text opatřený označením, zda v něm převažuje pozitivní či negativní emoce. Je možno zde nalézt vzorky ve třech jazycích, a sice v češtině, němčině a angličtině. Další rozdělení dat je podle určení pro trénování či testování. Polovina vzorků v každém jazyce je určena pro trénování, druhá pak pro testování. Pro každý z jazyků obsahuje databáze 12 000 vstupů, celkově je tedy k dispozici 36 000 záznamů. Toto množství dat je pouze zlomkem objemů, pro které má být možno systém využívat.

První částí procesu je import dat pro trénování klasifikátoru. Texty obsahují subjektivní názor uživatelů se všemi jejich náležitostmi včetně gramatických chyb, překlepů, apod. Následuje segmentace textu neboli rozčlenění vět na tokeny, sloužící jako základní nositele informace, v tomto případě jednotlivá slova věty. Vstupní text je převeden na tokeny tak, že je využito znaků oddělujících slova jako mezery, tečky, číslice a ostatní znaky, jenž nejsou písmeny, načež je veškerý nyní již rozdělený text převeden na písmena malé abecedy. Tímto krokem úprava textu pro další práci končí, není tedy aplikován lemmatizátor či stemmer, seznam nejčastějších zkratek, filtr stop slov, či slovník pro eliminaci překlepů. Tyto skutečnosti je nutno brát v úvahu při posuzování naměřených výsledků.

Aby bylo možné text dále využít, je nutné jej upravit do takové podoby, aby mohl být zpracován algoritmem strojového učení. Text je tedy reprezentován za pomoci vektorového modelu, což je zajištěno prostřednictvím výpočtu hodnoty TF-IDF [3]. Jde o metodiku pro hodnocení relevance jednotlivých slov v textu. Umožňuje slovům přiřazovat váhu na základě převrácené četnosti slova v dokumentech, pokud je tedy velký počet dokumentů ze sbírky, ve kterých se dané slovo vyskytuje, jeho důležitost se snižuje. Pro určení váhy jednotlivých složek vektoru je nejprve vypočítána četnost výskytu tokenu v dokumentu, následně je vypočtena důležitost tokenu v rámci celého korpusu, načež vynásobením těchto dvou hodnot dostáváme celkovou výslednou váhu tokenu.

V oblasti zpracování textu je zřejmě nejpoužívanější technikou algoritmus podpůrných vektorů (SVM) založený na konceptu hledání rozhodovacích rovin, který je pro klasifikátor využit. Jak již bylo zmíněno dříve, systém je postaven na projektu Spark, jež tvoří sjednocený systém určený ke zpracování dat v různých podobách. K dispozici je několik knihoven Sparku, z nichž pro tuto práci je významnou knihovna MLlib určená pro strojové učení, která obsahuje SVM algoritmus zde využitý pro vytvoření modelu klasifikátoru. Jedná se o SVM s lineárním jádrem, jehož parametry byly nastaveny podle předešle zjištěné konfigurace vhodné pro texty. Volání funkce pro vytvoření modelu obsahuje parametr počtu iterací, označující kolikrát má proces učení proběhnout. Počet iterací je podstatným prvkem se značným vlivem na úspěšnost klasifikace, jak bude dále ukázáno v prezentaci výsledků.

4 VÝSLEDKY

Úspěšnost natrénovaného modelu pro klasifikaci zobrazená v Tab. 1 je zhodnocena ověřením správnosti rozhodování modelu za pomoci testovacích dat. Jedná se o poměr počtu správně zařazených dokumentů do dané kategorie vzhledem k počtu veškerých dokumentů zařazených do dané kategorie (včetně nesprávně zařazených). Zvažované počty iterací v každém jazyku jsou v rozsahu 1 až 10 000.

Tab. 1 zobrazuje dosaženou přesnost natrénovaného modelu pro každý z jazyků při různých počtech iterací, kde úspěšnost je vyjádřena procentuálně. Lze vyčíst, že adekvátní počet iterací se pohybuje

	jazyk		
iterace	čeština	němčina	angličtina
1	67,13	67,00	73,00
5	75,92	70,33	79,00
10	78,00	78,28	91,05
100	78,05	79,73	91,88
1000	77,02	79,05	91,45
10000	75,47	77,78	88,05

Tabulka 1: Úspěšnost modelu pro jednotlivé jazyky a počty iterací

kolem hodnoty 100 iterací, a sice pro všechny 3 jazyky. Celkově nejnižší přesnosti bylo dosaženo u českého jazyka. Možným důvodem je fakt, že čeština je odlišnější od zbylých dvou jazyků více, než jsou tyto jazyky odlišné navzájem od sebe, navíc jde o bohatější jazyk se složitější větnou skladbou. Dále mohla k nižší úspěšnosti přispět skutečnost, že vzorky sice pochází z české databáze, zlomek z nich však byl vytvořen ve slovenštině. Nejvyšší přesnosti je při stejných počtech iterací jako u ostatních jazyků dosaženo pro anglický jazyk, a sice více než 90%. Přesnost se pak pro český a německý jazyk pohybovala na společné hodnotě, konkrétně téměř u hranice 80%.

5 ZÁVĚR

Článek popisuje vytvořený systém pro detekci pozitivní či negativní emoce v textu v českém, německém a anglickém jazyce. Texty použité pro trénovací množinu pocházejí z internetových diskuzí a obsahují označení převažující emoce. Z prezentace výsledků je zřejmé, že natrénovaný model byl při adekvátním počtu iterací schopen správně zařadit přibližně 3 ze 4 vzorků v českém jazyce, podobně jako u němčiny. Pro stejný počet iterací dovede model správně klasifikovat dokonce přibližně 9 z 10 předložených vzorků v anglickém jazyce. Přesnost lze považovat za relativně uspokojivou, přihlédneme-li k faktu, že na vstupním textu byla provedena pouze tokenizace, podoba slov tedy byla zachována v jejich původním tvaru, ve které byla autorem příspěvku vytvořena.

REFERENCE

[1] POWER, D.J. Using 'Big Data' for analytics and decision support. *Journal of Decision Systems, 23:2, s. 222-228, 2014*. DOI: 10.1080/12460125.2014.888848.

[2] CHUANG, Z.J., WU, C.H. Multi-modal emotion recognition from speech and text. In *The Association for Computational Linguistics and Chinese Language Processing*, vol. 9, s. 45-62, 2004.

[3] DE SILVA, J., HADDELA, P.S. A term weighting method for identifying emotions from text content. *Industrial and Information Systems (ICIIS), 2013 8th IEEE International Conference on Industrial and Information systems*. ISBN 978-1-4799-0908-7.

[4] TRIPATHY, A., AGRAWAL, A., RATH, S.K. Classification of Sentimental Reviews Using Machine Learning Techniques In *Procedia Computer Science*, vol. 57, 2015, pp. 821–829.

[5] LIMA, A.C.E.S, DE CASTRO, L.N., CORCHADO, J.M. A polarity analysis framework for Twitter messages, In *Applied Mathematics and Computation*, vol. 270, 2015, pp. 756–767.

[6] MAŠEK, J., BURGET, R., KARÁSEK, J., UHER, V., DUTTA, M.K. Multi-GPU Implementation of k-Nearest Neighbor Algorithm. In *2014 37th International Conference on Telecommunications and Signal Processing (TSP)*. Berlin, Germany: 2014. s. 652-655. ISBN: 978-80-2144983-1.

INVISIBLE IMAGE WATERMARKING USING FREQUENCY DOMAIN FEATURE POINTS

Dávid Pacura

Master Degree Programme (2), FEEC BUT

E-mail: xpacur00@stud.feec.vutbr.cz

Supervised by: Petr Petyovský

E-mail: petyovsky@feec.vutbr.cz

Abstract: For robust invisible watermarking scheme for images, it is crucial to withstand both malicious and common image manipulations (e.g. JPEG compression, pixel jittering or cropping). In this paper, local image features are used. For this, a Maximally Stable Extremal Regions (MSER) detector, which is fast and invariant to affine transformations, is used. The watermark itself is embedded multiple times in circular areas across image. This is achieved by manipulation of magnitude coefficients in frequency domain. During detection, watermark is recovered using majority criterion.

Keywords: Invisible image watermarking, robust image watermarks, image protection, discrete Fourier transform (DFT), peak signal to noise ratio (PSNR)

1. INTRODUCTION

Because a considerable portion of economics resources is dedicated to the creation of intellectual property, a creator may expect some measure of compensation for the resources expended in the process of creation [1]. With rise of the Internet, the traditional forms of intellectual property protection are no longer sufficient – the high amount of available multimedia data makes it easy to violate these intellectual rights. One of possible solutions is invisible watermarking of digital signals. It consist of embedding of additional information (watermark) into carrier signal (cover image) in such way, that it is imperceptible to the human observer (stego image), but easy to detect by computer algorithm [2]. The proposed scheme consist of following steps: feature point's selection, local characteristic regions (LCRs) creation, discrete Fourier transformation (DFT) and watermark W generation, which are common for both embedding and detection part. The embedding continues with data embedding itself, inverse DFT and replacing of original LCRs with their watermarked counterparts (see **Figure 1:** Flowchart of the watermark embedding process.Figure 1). Detection part continues only with recovering of watermark in the Fourier space.

2. FEATURE POINTS SELECTION

For a detection of feature points, Maximally Stable Extremal Regions (MSER) [3] detector was chosen. It provides stability under various attacks such as affine transformations, perspective changes, noise reduction and addition or illumination changes.

The final output of the MSER feature detector is a list of regions R, where each region is described as a list of points P contained in this region. These regions are then ordered ascendingly by its size and removal of overlapping regions is performed. Then, new circular Local Characteristic Regions (LCRs) are created producing a list L. LCRs are of fixed diameter with centers identical to those of detected feature points and contains data from corresponding locations in original image.

3. WATERMARK EMBEDDING

Prior to the data embedding itself, a transformation from spatial domain to frequency domain is performed. Firstly, circular LCRs are zero padded: An image of dimensions $d \times d$ with zero values is created, where d is the diameter of the LCR circle. Then LCR data are copied into it and transformed image T is created using discrete Fourier transform (DFT).

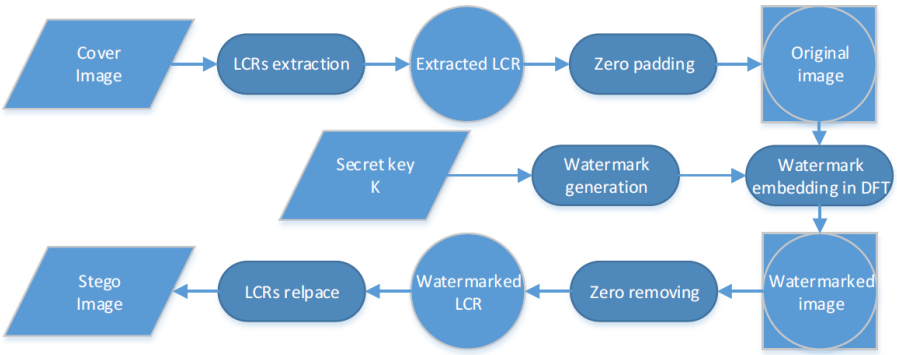


Figure 1: Flowchart of the watermark embedding process.

The watermark W itself (generated from secret key K) is embedded into magnitude part T_m of transformed image. In order to make it imperceptible, a middle frequencies are selected [4]. This is achieved by selection of annular area between two circles with radius R_1 and R_2. The $R_1 = \left\lceil \frac{R}{4} \right\rceil$ and $R_2 = R_1 + 32 \leq \left\lceil \frac{7R}{8} \right\rceil$. Then $C(r_i), i = 1, \ldots, 32$ is the set of all homocentric circles around the zero frequency in the selected area. The watermark is then embedded into this area: one bit per circle. For each starting point, another paired point 90° apart is generated by negating the y coordinate [4]:

$$\{(x_i{}^p, y_i{}^p), (x_i{}^p, -y_i{}^p)\}, i = 1, \ldots, 32; p = 1, \ldots, n \tag{1}$$

where i is the i-th bit of watermark, p is the current pair and n is the count of pairs in the circle c.

Then, the embedding itself is achieved by satisfying following condition in each circle:

$$\begin{aligned} M(x_i{}^p, y_i{}^p) - M(x_i{}^p, -y_i{}^p) &\geq \alpha \quad if\ W_i = +1 \\ M(x_i{}^p, y_i{}^p) - M(x_i{}^p, -y_i{}^p) &\leq \alpha \quad if\ W_i = -1 \end{aligned} \tag{2}$$

where α is the watermark strength and M is the magnitude of the coefficient at the location $T_m(x_i, y_i)$.

The parameter α is determined automatically for each LCR from its local properties in spatial domain in order to ensure the same imperceptibility of watermark across all LCRs:

$$\alpha = k \times sum(M) \times \frac{b}{mean(LCR)} \tag{3}$$

where b is the maximum possible value of single image channel pixel (255 for 8 bit images).

Finally, the region is transformed into spatial domain using inverse DFT. Then, zero padding removal is performed prior to the replacement of the watermarked regions with the ones in original image.

4. WATERMARK DETECTION

The detection process is similar to embedding process: The first three steps (feature points selection, LCRs generation and DFT) are all the same. The difference comes after the angular area is calculated. For each circle $c \in C$, a set B_i containing extracted i-th bit of watermark $\widehat{W}$ is created [4]:

$$\widehat{W}_i^{\,p} = \begin{cases} +1 & M(x_i, y_i) - M(x_i, -y_i) \geq 0 \\ -1 & M(x_i, y_i) - M(x_i, -y_i) < 0 \end{cases} \quad (1)$$

where M is the magnitude of the coefficient at the location $T_m(x_i, y_i)$ and $\widehat{W}_i^{\,p}$ is the extracted bit.

The real watermark values are computed by usage of majority criterion:

$$d(x) = \frac{count(x,B)}{n} \quad (2)$$

where $count(x, L_i)$ is the count of value x in the set B_i and n is the count of pairs in the circle c.

The final value of the extracted watermark bit $\widehat{W}_i$ is then computed as:

$$\widehat{W}_i = \begin{cases} +1 & d(+1) \geq d(-1) \\ -1 & d(+1) < d(-1) \end{cases} \quad (3)$$

Finally, an image can be declared as watermarked, if $\widehat{W} = W$ is present in at least two LCRs.

5. RESULTS

The algorithm was tested on two testing sets of images. The first set contained images rendered using computer graphics by both rasterization and raytracing. They were stored in PNG format with resolution of $2{,}560 \times 1{,}600\ px$. The second testing set consisted of photographs taken by both DSLRs and compact cameras. They were stored in JPEG format and contained low amount of noise. Their resolution ranged from 5 to 16 MPx. Both sets contained N = 20 RGB images with 8 bits per channel.

The 32 bit pseudorandom watermark was inserted into all three channels of each image. The total number of watermarked areas ranged from 48 to 1128 per image, with median value of 159 for first testing set and 543 for second testing set. The PSNR ranged from 40 dB to 56 dB with average of 48 dB. Overview of performed attacks and average results (n – count of images declared as watermarked and % of LCRs declared as watermarked) for each testing set can be found in Table 1.

Attack type	Image set 1 (N = 20)		Image set 2 (N = 20)	
	n	% of detected	n	% of detected
median 3x3, JPEG Q = 50 %	19	10	18	9
brightness & contrast +/- 40 %	20	30	20	30
crop 50 %, sharpen 5x5	20	33	20	39
equalized histograms	20	40	20	36
noise deviation = 15, shear 3 %, ration 20°	18	11	17	7
conversion to grayscale, skip 10 % of pixels	20	48	20	50

Table 1: Robustness test results of proposed algorithm over various attacks for two testing sets.

6. CONCLUSION

The proposed method showed excellent imperceptibility with average PSNR = 48 dB. However, some watermarks can still be visible with proper zoom. This issue can be solved by adjusting watermark strength in color channels according to human visual system sensitivity.

The robustness benchmark against common image modifications like JPEG compression and illumination changes showed great resistance – watermark was detected in all of the attacked images and in high percentage of embedded LCRs (see Table 1). Also, the robustness against malicious attacks was very good, especially against addition of Gaussian noise and rows and columns skipping. However, weaknesses of this watermarking scheme are affine transformations like rotation and shearing, where robustness is limited. The problematic is also scaling operation, where proposed algorithm shows good results only in scale changes up to 3 %. To address this issue, further research is required.

REFERENCES

[1] M. Arnold, M. Schmucker and S. D. Wolthusen, Techniques and applications of digital watermarking and content protection, Boston: Artech House, 2003, p. 274.

[2] J. Seitz, Digital watermarking for digital media: Information Science Pub, 2005, p. 262.

[3] J. Matas, O. Chum, M. Urban and T. Pajdla, "Robust Wide Baseline Stereo from Maximally Stable Extremal Regions" British Machine Vision Conference, pp. 384-393, 2002.

[4] X. Wang, J. Wu and P. Niu, "A New Digital Image Watermarking Algorithm Resilient to Desynchronization Attacks" IEEE Transactions on Information Forensics and Security, vol. 2, no. 4, pp. 655-663, December 2007.

ORDER ANALYSIS

Lukáš Honc
Master Degree Programme (5), FEEC BUT
E-mail: xhoncl00@stud.feec.vutbr.cz

Supervised by: Petr Beneš
E-mail: benesp@feec.vutbr.cz

Abstract: This article deals with order analysis. The main purpose is design and realization of own library for online order analysis as a toolkit for LabVIEW. Library contains methods for order analysis including basic functions for processing vibration and tacho signal and functions for tach-less analysis of vibration signal.

Keywords: order analysis, order tracking, tach-less analysis, vibrodiagnostics, condition based maintenance

1 ÚVOD

Při vibrační diagnostice rotačních strojů je žádoucí oddělit deterministické signály (např. od zubů převodovky) od těch náhodných (šum, vibrace z okolí). Tyto deterministické signály jsou přesnými násobky frekvence otáčení, nejvhodnějším řešením je tedy vztáhnout tyto signály k otáčkám stroje [1]. Pro tyto účely byly vyvinuty metody řádové analýzy, které odstraňují fluktuace v rychlosti a které se využívají pro analýzu strojů pracujících při proměnlivých otáčkách (tzv. rozběh nebo doběh). Řádem (angl. order, zkr. ord) rozumíme spektrální složku signálu, jejíž frekvence je pevně spjatá se základní fundamentální frekvencí stroje [2].

Cílem této práce byla implementace vlastní knihovny funkcí jako toolkitu do vývojového prostředí LabVIEW a jednotlivé metody optimalizovat a ověřit jejich vlastnosti na reálných a simulovaných signálech.

2 ŘÁDOVÁ ANALÝZA

Rychlostní profil: Pro získání rychlostního profilu se běžně používá signálu z tachosondy. Protože je dnes žádané extrahovat tuto informaci přímo ze signálu vibrací, byl vedle klasické metody naimplementován Bayesův odhad fundamentální frekvence ω_0.

Metoda Bayesova odhadu je založena na předpokladu, že signál vibrací d $[\mathrm{m.s}^{-1}]$ je tvořen konkrétními řády (k-násobky základní frekvence ω_0 $[\mathrm{rad.s}^{-1}]$), viz. rovnice (1), kterou lze přepsat do maticové podoby (2). Šum e $[\mathrm{m.s}^{-1}]$ je Gaussovský s rozptylem σ^2 [3].

$$d(t) = a_0 + \left[\sum_{k=1}^{K} a_k \cos(\omega_k t) + b_k \sin(\omega_k t)\right] + e(t) \tag{1}$$

$$d = Gb + e \tag{2}$$

Rychlostní profil základní frekvence je vypočten z L překrývajících se záznamů o délce N vzorků podle rovnic (3) až (6) za předpokladu, že se signál vibrací skládá z K frekvenčních složek [3][4]. V grafech 1 a 2 je možné vidět porovnání jednotlivých metod pro získání rychlostního profilu z dat

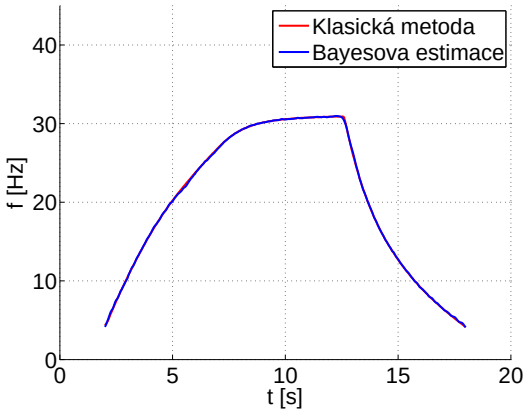
naměřených na počítačovém ventilátoru se čtyřmi vinutími (Pro výpočet byl použit pouze 4-násobek fundamentální frekvence s rozptylem rovným 0,125 Hz a okno o délce 800 vzorků, signál byl vzorkován frekvencí 3,2 kHz). Pokud opomeneme začátek a konec záznamu z důvodu délky okna a nízkých otáček, je relativní chyba určení fundamentální frekvence, až na ostré změny v trendu rychlosti a rezonanci na 22Hz (čas přibližně 5,6 s), menší než 1%. Při testování na reálných a simulovaných signálech se ukázalo problematické právě křižování jednotlivých řádů s konstantními a rezonančními frekvencemi, které můžou v daném okamžiku svou energií v signálu převažovat. Tomu je možné zamezit použitím krátkého okna. Jako další možné řešení se zde nabízí použít apriorní znalosti o vlastnostech rotoru a jednotlivé rezonanční frekvence penalizovat.

$$p(\omega_0|d) \propto p(d|\omega_0)p(\omega_0) \tag{3}$$

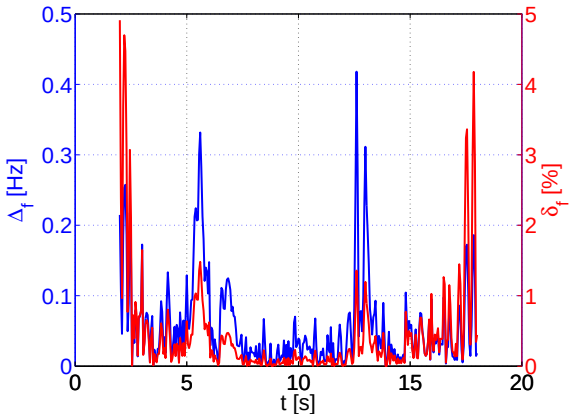
$$p(d|\omega_0) \propto \frac{(d^T d - d^T G(G^T G)^{-1} G^T d)^{-(N-2K-1)/2}}{\sqrt{|G^T G|}} \tag{4}$$

$$p(\omega_0) = \frac{1}{\sqrt{(2\pi\sigma^2)}} \exp -\frac{1}{2\sigma^2}(\omega_0 - \mu)^2 \tag{5}$$

$$\omega_0^{(i)} = \left\{ arg\ max_{\omega_0^{(i)}}\ p(\omega_0^{(i)}|d^{(i)}) \right\} \quad \text{kde} \quad i = 0, 1, ..., L \tag{6}$$



Obrázek 1: Rychlostní profil získaný klasickou metodou a pomocí Bayesových vzorců

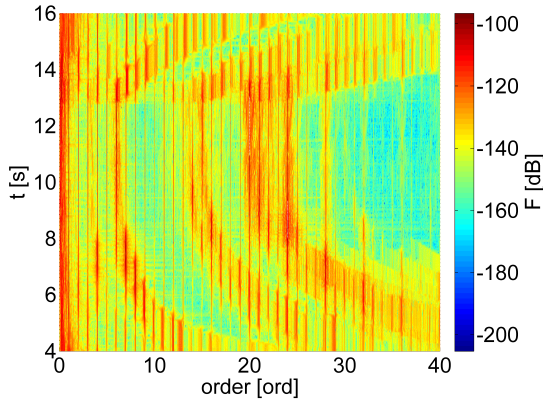


Obrázek 2: Chyba určení fundamentální rychlosti pomocí Bayesových vzorců

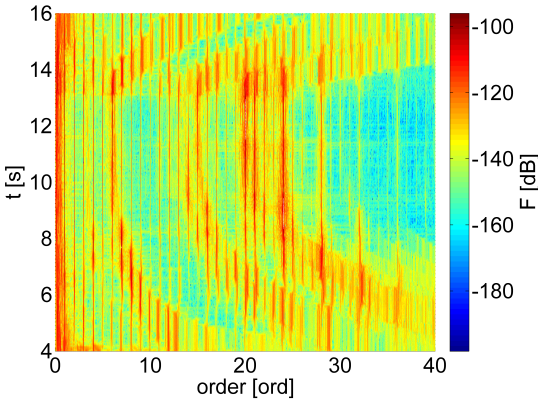
Integrace: Mechanické vibrace mají zhruba rovnoměrné frekvenční spektrum z hlediska rychlosti [2]. Rychlost vibrací je tedy vhodná pro vyhodnocení limitů z důvodu vazby na poškození stroje, ale i jako vstup pro odhad fundamentální frekvence pomocí Bayesových vzorců, kdy má každý řád stejnou váhu. Jelikož jsou nejčastěji používanými senzory ve vibrodiagnostice akcelerometry, byly naimplementovány dvě metody integrace: integrace v časové oblasti (TDI) a integrace ve frekvenční oblasti (FDI). Experiment s interferometrem měřicím přímo rychlost vibrací a akcelerometrem, jehož výstup byl na rychlost integrován, ukázal, že nejpřesnější metodou je FDI. Při TDI se ukázal jako stěžejní parametr typ horní propusti pro odfiltrování nízkých frekvencí, která způsobuje fázový posun jednotlivých spektrálních složek. Tento problém při FDI není, jelikož jsou nízké frekvence jednoduše nahrazeny nulou, nevýhodou je naopak vyšší výpočetní náročnost a navzájem nenavazující záznamy při online analýze.

Převzorkování: Signál je nutné s využitím rychlostního profilu převzorkovat konstantním počtem vzorků na periodu. Experiment ukázal, že pro interpolaci okamžité rychlosti i mezi pulzy z tachosondy je vhodné použít delší trend rychlosti v závislosti na dynamice daného rotoru. Při online analýze vede předchozí věta na zpracování signálu jako navzájem překrývajících se bloků. Výsledný rychlostní profil je integrován na fázi, mezi každými dvěma pulzy normalizován na interval $^{2\pi}/_{n}$, kde n je počet pulzů za otáčku, a interpolován na ekvidistantní úhlové rozložení nových vzorků. Při Bayesově odhadu otáčkové frekvence není fáze po integraci normalizována, ale úhel roven k-násobku 2π je využit jako normalizační faktor při převzorkování.

Řádové spektrum: Pro výpočet řádového spektra je použit algoritmus krátkodobé Fourierovy transformace s plovoucím oknem o délce n period. Tím je docíleno řádového rozlišení $df = \frac{1}{n}$ [ord]. Na obrázcích 3 a 4 lze vidět, jaký vliv má na řádovou analýzu použitý rychlostní profil.



Obrázek 3: Řádové spektrum vypočtené pomocí klasického rychlostního profilu (celkový výpočetní čas 39 ms)



Obrázek 4: Řádové spektrum vypočtené pomocí odhadovaného rychlostního profilu (celkový výpočetní čas 2,03 s)

3 ZÁVĚR

Podařilo se naimplementovat a z velké části otestovat knihovnu pro řádovou analýzu signálu jako toolkit pro vývojové prostředí LabVIEW. Výsledné funkce je možné použít jak pro offline tak online analýzu signálu a většinu z nich se rovněž podařilo paralelizovat a optimalizovat tak pro více-jádrové procesory. Předmětem dalšího vývoje je zpřesnění metody Bayesova odhadu otáčkové frekvence.

REFERENCE

[1] Tůma, J.: Zpracování signálů získaných z mechanických systémů užitím FFT. Štramberk, 1997, ISBN 80-901936-1-7.

[2] Randall, R. B.: Vibration Based Condition Monitoring. Chichester, West Sussex, PO19 8SQ, United Kingdom: John Wiley & Sons, 2011, první vydání, ISBN 978-0-470-74785-8.

[3] Pedersen, T., Herlufsen, H., Hansen, H.: Order Tracking in Vibro-acoustic Measurements: A Novel Approach Eliminating the Tacho Probe. SAE Technical Paper 2005-01-2266, 2005, DOI:10.4271/2005-01-2266.

[4] Gondo, H. K.: Classical & Bayesian Spectral and Tracking Analysis. 2007. *Master's thesis*. Technical University of Denmark.

PERSON FOLLOWING BY SPOT LIGHT USING CAMERA

David Rajnoch

Master Degree Programme (2), FEEC BUT

E-mail: xrajno07@stud.feec.vutbr.cz

Supervised by: Libor Boleček

E-mail: bolecek@feec.vutbr.cz

Abstract: This paper describes design of person tracking system connected to spot light fixture. System is able to track actor or narrator walking on stage in theatre. System is designed to be easy to use and configure. Several issues connected with such tracking is mentioned in text below.

Keywords: Object tracking, Computer Vision, C++, OpenCV, Spot Light

1. ÚVOD

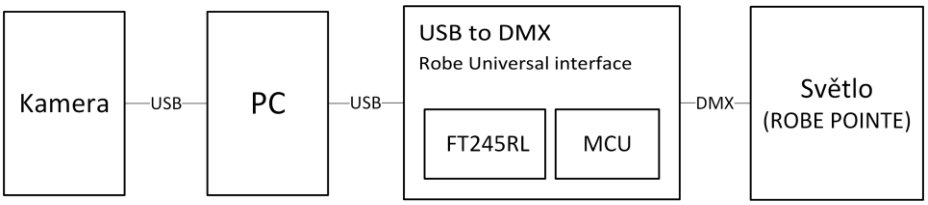
Dnešní digitální světla obsahují několik procesorů, přes desítku motorů a jejich funkcionalita otevírá stále větší možnosti uměleckému vyjádření. Proč ale nejsme schopni masově automatizovat tak jednoduchý úkol, jako je sledování osoby na pódiu. Především protože se pohybujeme v odvětví živého vystupování, ve kterém není místo pro chyby. Sledovací systém tak může uspět pouze při splnění náročných kritérií a musí se dostat do podvědomí jako spolehlivý. Snahou tohoto projektu je vytvořit spolehlivý jednoduše nastavitelný, téměř autonomní systém dané funkcionality implementovatelný třeba jako vlastnost světla.

2. ROZBOR SITUACE

Největším problémem popsané detekce je časově proměnné světelné prostředí scény. Může také dojít k situaci, že použité světlo je jediným zdrojem světla. Je proto důležité definovat podmínky, při kterých bude světlo testováno a pro které bude program přizpůsoben. Další komplikací je nerovnoměrné rozložení intenzity světla ve světelném kuželu, při odlišných zoom faktorech a vysoká intenzita oproti zbytku obrazu. Je potřeba použít algoritmus pro korekci expozice z lokálně maskovaného obrazu. Snaha o minimální odezvu naráží na konečnou rychlost pohybu světla ale největší měrou přispívá zpoždění při snímání obrazu a vnitřní buffer kamery. Je vhodné vyčítat snímky z kamery nezávisle na hlavní smyčce programu a zajistit tak nejaktuálnější snímek. I přes toto opatření program musí počítat se snímkem několik desítek milisekund zpožděným a patřičně na to reagovat.

3. REALIZACE

Základní blokové schéma navrženého systému zobrazuje obrázek 1. Kamera je pomocí USB rozhraní připojena k PC. PC zpracovává obrazová data a výsledná poloha je odeslána přes USB – DMX převodník do světla. Kamera je umístěna na světle rovnoběžně s vyzařovaným paprskem a pohybuje se ve stejném směru jak světlo. Externí kamera by musela být přesně umístěna případně kalibrována na aktuální polohu, což přináší komplikace při instalaci, proto je použití kamery upevněné na světle základní předpoklad pro jednoduchost systému.



Obrázek 1: Blokové schéma aplikace.

4. POPIS PROGRAMU

Stavový popis aplikace zobrazuje obrázek 2 (vlevo). Během *Inicializace* je inicializována kamera a převodník, provedeno první nastavení expozice a načteny nastavení ze souboru.

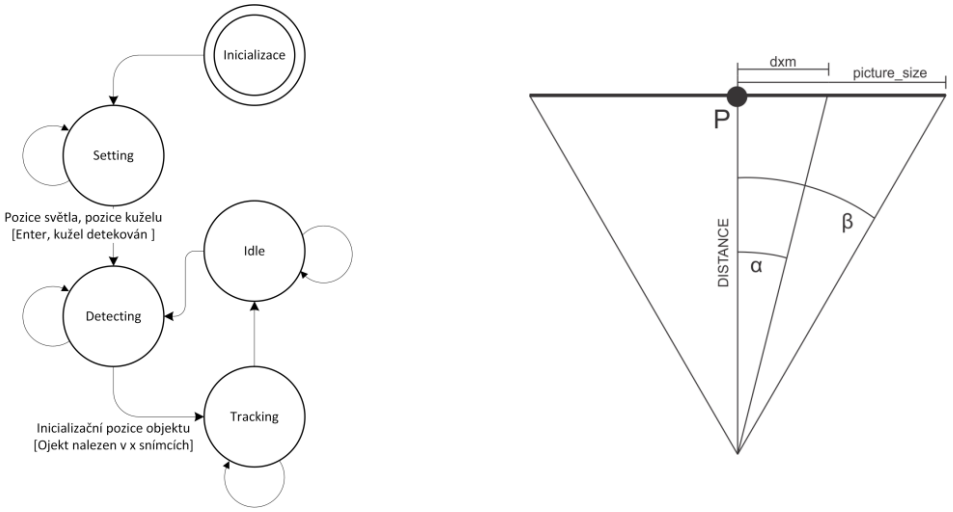
Setting umožní nastavit světlo pomocí klávesnice na výchozí pozici. Během tohoto stavu musí proběhnout kalibrace světelného kuželu, která je nezbytná pro velké rozdíly ve vzálenosti k podiu a při testování (na krytí nedostatku v uchycení kamery). Výsledkem je hodnota středové souřadnice kuželu proti které je kompenzován pohyb sledovaného člověka.

Ve stavu *Detecting* je nalezen pohyb pomocí diference dvou po sobě jdoucích snímků. Tato metoda slouží k prvotní inicializaci osoby. Předpokládá se nepohybující se světlo natočené směrem k pódiu a pouze jeden detekovatelný, pohybující se člověk. Detekcí pohybu v x snímcích se přechází do stavu *Tracking.*

Stav *Tracking* obdrží souřadnici nalezené osoby a dojde k pohybu světla na danou souřadnici. Dále se pravidělně aktualizuje poloha světla na základě informací o pohybu osoby. Při ztrátě dochází k přechodu do stavu *Idle.*

V *Idle* je podle parametrů vyhodnoceno, jestli dojde k opětovnému přechodu do stavu *Detecting* nebo *Tracking* a s jakými parametry, případně program zůstává ve stavu *Idle*. Jako rozšíření je možné implementovat změnu parametrů svělného paprsku tak, aby například pomohl kameře osvítit a znovu detekovat člověka.

Takto lze popsat základní přechody mezi stavy. V plánu je vytvořit módy, odpovídající různým funkcionalitám. Momentálně je jimplementován mód *MODERATOR*, který navíc detekuje, jestli osoba opustila oblast ve které byla inicializována. Pokud ano a vrátí se do oblasti blízké oblasti inicializace, dojde k ukončení sledování.



Obrázek 2: Stavový diagram (vlevo), Geometrie polohy (vpravo).

4.1. GUI

Snahou je vyvinout téměř autonomní systém, proto program disponuje pouze jednoduchý uživatelským rozhraním, které akceptuje vstup z klávesnice pro ladění. Je možné měnit expozici, polohu, spustit záznam, sloužící jako log, nastavit světlo do předvolených poloh, vyvolat histogram, nebo dílčí data některých funkcí.

4.2. DOPLŇUJÍCÍ SYSTÉMOVÉ INFORMACE

Výchozí rozlišení snímaného obrazu je 1024x768, program však v časově náročných funkcích zpracovává pouze část snímku, která je důležitá. Snímková rychlost kamery je až 120 snímků za sekundu. Systém pracuje okolo rychlosti 15 snímků za sekundu. Program je vyvíjen v C++ s použitím knihovny OpenCV v aktuální verzi 3.1 a volně šířitelné implementace TLD.

Použitím TLD trackeru (track-learn-detect), který implementuje reinicializaci pomocí naučeného modelu, je možné minimálně v případě plošně osvětlené scény jednoduše znovu inicializovat člověka, který například zašel za překážku. Své výhody také nese možnost uložení a vyvolání modelu.

Pro přepočet mezi obrazovými body a reálným prostorem je použitý rovinný geometrický model z obrázku 2 (vpravo). Jeho vstupem je vzdálenost k osobě, úhel záběru kamery, počet kroků na rozsah a pohybový rozsah světla.

5. ZÁVĚR

Tento projekt je stále ve stádiu vývoje. Je potřeba dokončit implementaci TLD trackeru, automatické expozice, podmínek reinicializace a testovat jej v reálných podmínkách divadla. Je velmi pravděpodobné, že ve fázi testování dojde k významným úpravám základních částí (inicializujícího detektoru a trackeru). Jsem nicméně přesvědčen, že je možné dospět k funkční realizaci, která by při dlouhodobém vylepšování dokázala nahradit roli osvětlovače. Velmi vhodné by bylo implementovat učení na úrovni prostoru, které by posílilo spolehlivost systému v kritických situacích. Možnost zaznamenat video jako formu logu nese do budoucna potenciál v ladění kritických situací a možnosti lepšího pochopení nároků na systém.

REFERENCE

[1] OpenCV: OpenCV modules. OpenCV: 3.0.0-dev, Open Source Computer Vision [online]. 2015 [cit. 2015-11-08]. Dostupné z: http://docs.opencv.org/master/

[2] C. GONZALEZ, Rafael a Richard EUGENE WOODS. Digital Image Processing. New Jersey: Pearson/Prentice Hall, 2008. ISBN 013168728X, 9780131687288. Dostupné z: https://books.google.cz/books?id=8uGOnjRGEzoC

[3] IAN T. YOUNG, Ian T.Jan J. Fundamentals of image processing. version 2.2. Delft: TU Delft, Faculty of Applied Physics, Pattern Recognition Group, 1995. ISBN 90-756-9101-7.

PLAGIARISM DETECTION IN PROGRAM CODES USING MAPPING TECHNIQUE

Jakub Kašpar

Master Degree Program (2), FEEC BUT

E-mail: xkaspa36@stud.feec.vutbr.cz

Supervised by: Martin Vítek

E-mail: vitek@feec.vutbr.cz

Abstract: The aim of this paper is to introduce the problem of plagiarism and propose a method for plagiarism detection in program codes. In the first part of this paper the basic definition of plagiarism is described. Further in the paper the principle of preprocessing and localization process for signs of plagiarism is introduced. The last part of this paper presents an algorithm for comparison of the detected signs to get the best results possible. The detector was tested on student projects from the BTBIO study program.

Keywords: plagiarism, detection, preprocessing, signs, mapping

1. INTRODUCTION

One of the most serious problems in the academic field, mostly among students, is plagiarism. There are two types of plagiarism. The first one is plagiarism of scientific papers, publications, books, project or even ideas. The second type is plagiarism of program codes. Plagiarism basically means using someone else's work as your own, without acknowledging the original author. In this work a method for plagiarism detection in program codes is introduced. The proposed method does not declare exactly what is and what is not a plagiarism, but alerts about a probability of two program codes being similar or even the same. Focus on full automatization of the program is also a high priority, so that it could process a large set of data in the shortest time possible. The whole algorithm consists of two main parts, each of them being further divided into smaller elements, as it is shown in the Figure 1.[1]

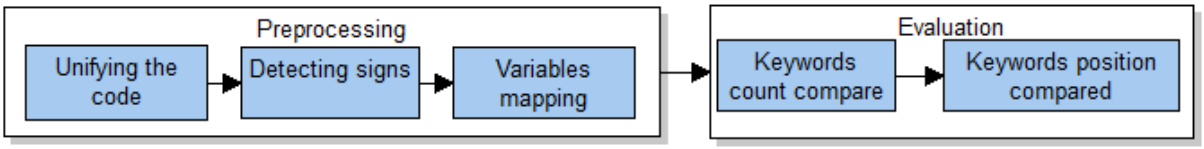


Figure 1: Block diagram

2. PREPROCESSING

Program codes come in a large variety of forms and different formats and often contain a lot of redundant information. Before the beginning of whole sign detection procedure, it is necessary to eliminate all the useless data, such as lines with code commentary, empty lines, and all the blank space in front of and behind the code. Setting all letters to lower case is also helpful for further processing. With these adjustments all processed codes are in a unified format. Without the preprocessing step, adding even minor format changes to the code would lead to inability to detect the similarities. [1]

2.1. SIGNS DETECTION

The next step before the final similarity evaluation of two compared codes is the detection of significant signs. These signs comprise of a large number of keywords typical for the used program environment. In case of the detector proposed in this paper the program environment is Matlab. The keywords were chosen from the most commonly used functions such as *for, while, if,* and also very specific functions, typical for signals and image processing and bioinformatics.

The proposed program scans separately all lines of the preprocessed code and searches for the keywords. Apart from the final count of detected words in the code, the exact position of these words is also recorded. This is very important for further evaluation, as even if the amount of keywords in two codes is the same, it is not necessarily a sign of plagiarism since those keywords can occur in unrelated positions in the code and be used to fulfill a different purpose. Important signs are also the size of the code in bytes and number of rows and characters.

The keywords are searched in different ways. The first group of keywords are the functions, which can be placed only at the beginning of a line e.g. *for* function. For the purpose of finding the keywords the standard *find* function is used. But in case of this kind of functions the position in the code must be also checked along with detection of a *space* sign behind the function. This procedure makes sure, that the detected keyword is actual function and not only part of some variable e.g. *forcount*. It is important that detection of these keywords is done before the space deletion process.

The second group consist of the functions which are usually the first function in a row and have parenthesis behind them e.g. *imshow*. The detection process is very similar to the one used in the first group with exception of the *space* sign detection. There are also keywords like *figure* which can be used in code with or without parenthesis. These words are searched for by combination of both previously described methods.

The third group of keywords are the functions, which can be situated anywhere in the code. The *find* function is again used for detection. Problematic are the functions, which are part of the name of some other function. Good example of this is the *fft* function which is part of the inverse *ifft* function. In order to prevent the false detection in these kind cases, the sign in front of the third group functions are checked and the right function is determined. In the detected keywords a check is carried out whether these keywords are not in the string format, which would mean that they are not used as functions. This process is done by detecting the number and positions of quotes in the row and determining whether or not the keyword is between the active pair.

The last keyword to detect is *end*. This step is done by simple string comparison, since there should not be any other signs on the same line after the preprocessing

2.2. VARIABLES DETECTION

One of the most common ways to hide plagiarism in program codes is to change the names of the used variables. In order to detect these changes, all variables in the code are found and their positions recorded. This process creates maps with coordinates of the variables in all the processed codes which are later compared to each other. Additionally a vector of distances between neighboring variables is created from these coordinates, which is later used for more general comparison.

The variables are detected with the use of information about positions of the first two groups of keywords. The declared variables have to be on every row apart from these positions.

3. EVALUATION

The final part of proposed method is evaluation of the detected signs. The signs from two different codes are compared together and weighted. Weights are not set on constant value, but are instead calculated for every code and sign independently. Every comparison of two signs can take the value from 0 to 1 and so do the weights. The following equations are used for calculation of probability of plagiarism (1, 2) and for the determination of weights (3, 4).

$$P = \frac{\sum_{i=1}^{N} P_i \cdot V_i}{\sum_{i=1}^{N} V_i} \cdot (1 - V_p) \quad (1)$$

$$P_i = \frac{A_i}{B_i} \quad (2)$$

$$V_i = P_{mi} \cdot \frac{4 \cdot P_{bi} + 5 \cdot P_{ri} + P_{ci}}{10} \quad (3)$$

$$V_p = \frac{C_{mi}}{C_i} \quad (4)$$

Here P_i is ratio of two compared signs, A_i and B_i are counts of i-th sign in compared codes, V_i is the weight for compared sign, P_m stands for comparison of the position maps of the sign and is further described below, P_{bi}, P_{ri} and P_{ci} are the ratios of the code sizes in bytes, numbers of rows and numbers of characters respectively. V_p is the weight for the whole comparison, C_{mi} is number of keywords used only in one of the two codes and C_i is number of all keywords in both codes. P represents the probability of plagiarism.

The comparison of the two position maps represented by variable P_m is done by use of the dynamic time warping (DTW) method on the both position vectors, which are used for the calculation of correlation coefficient to compare the similarity. The DTW method makes the shorter of the compared vectors the same length as is the length of the other vector by adding additional values to it, while preserving the initial similarity.

The detector was tested on a dataset consisting of school projects from the APRG course. 100 different codes were divided in to three groups by the topic of the project. All codes in groups were compared to each other and every pair with the higher comparison value than threshold was marked as suspicious from plagiarism. The results of testing are in the Table 1.

Statistics\Projects	Numeric Integrating	Genetics	Ciphers and games
SE	100 %	100 %	100 %
+P	85 %	100 %	77 %
Number of codes	40	21	39
Known plagiarism	4	0	7
Newly discovered plagiarism	7	0	3

Table 1: Results of testing of the proposed detector

Overall 1731 comparisons were made in the total time under 8 seconds. The main problem which caused the false detections were codes, that were made by the same person on top of which the problems were almost identical e.g. numeric integration with rectangle and trapeze method. The sensitivity (SE) results are based on the known cases of plagiarism in the database. Positive predictive value (+P) is based on both known and newly discovered cases of plagiarism.

4. CONCLUSION

The final result of this paper is novel software for detection of plagiarism in program codes. The software is based on similarity comparison of used functions and their positions. Results of the testing has SE 100 % and +P 87 % on average. One of the key advantages of the software are adaptive weights, which bring better results than constant weights. Furthermore the software is fully automatic therefore it is possible to evaluate large set of data in very short time.

REFERENCES

[1] ALSMADI, Izzat, Ikdam ALHAMI a Saif KAZAKZEH. Issues Related to the Detection of Source Code Plagiarism in Students Assignments. *International Journal of Software Engineering and Its Applications* (Vol.8, No.4 (2014), pp.23-34): [cit. 2015-12-30].

TRAINABLE IMAGE SEGMENTATION USING DEEP NEURAL NETWORKS

Martin Majtán
Master Degree Programme (2), FEEC BUT
E-mail: xmajta02@stud.feec.vutbr.cz

Supervised by: Pavol Harár
E-mail: harar@phd.feec.vutbr.cz

Abstract: This paper is focused on trainable segmentation of image with use of deep neural networks. In this paper, the principle of creating images from magnetic resonance, generating data with algorithm of sliding window, creating a data set used for training neural network and principal segmentation of image with neural network is described. In practical part the algorithm of sliding window is created for generating data from magnetic resonance images and created model of artificial neural network used for image segmentation. In the practical part was achieved accuracy of segmentation 64 %.

Keywords: neural network, deep learning, segmentation, MRI, DL4J, multiple sclerosis

1 ÚVOD

Roztrúsená skleróza je ochorenie centrálneho nervového systému ktoré postihuje autoimunitný systém, zrakové nervy, miechu a mozog. Pri ochorení imunitný systém napáda myelínový obal nervových vlákien. Myelínový obal ktorý obklopuje nervové vlákna je potrebný k správnemu šíreniu elektrických signálov. [3]

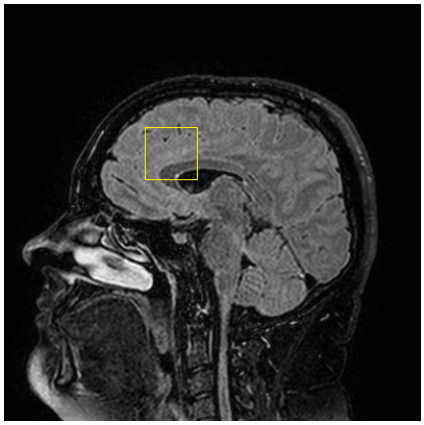
Táto práca sa venuje trénovateľnej segmentácii obrazu s použitím hlbokých neurónových sietí. Segmentácia obrazu je proces pri ktorom sa obraz rozdeľuje na segmenty tak že jednotlivé segmenty v obraze predstavujú reálne objekty. V danej práci sa trénovateľná segmentácia obrazu využíva na segmentáciu snímkov hlavy ľudí z magnetickej rezonancie a následnú detekciu roztrúsenej sklerózy.

2 PRÍPRAVA DÁT PRE NEURÓNOVÚ SIEŤ

Ako vstupné dáta pre proces segmentácie boli použité čiernobiele troj-rozmerné snímky hlavy z magnetickej rezonancie a k nim prislúchajúce masky, ktoré určujú kde v obraze sa nachádza roztrúsená skleróza. Masky majú rovnaké rozmery ako snímky. Na tom istom mieste kde sa v maske nachádza biela farba je v snímku vyznačená roztrúsená skleróza. Masky boli vytvorené manuálne lekárom. Pomocou snímkov a masiek boli vytvorené dva datasety ktoré slúžili na trénovanie a testovanie neurónovej siete.

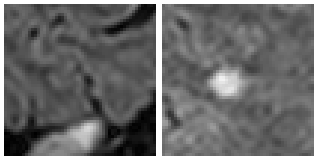
Pomocou algoritmu kĺzajúceho okna boli vytvorené podobrazy zo snímkov, ktoré sa použili na trénovanie a testovanie neurónovej siete. Algoritmus kĺzajúceho okna postupne posúva okno po jednotlivých pixeloch s určitým krokom a z pixelov ktoré sa aktuálne nachádzajú v okne vytvorí nový podobraz. Trieda do ktorej patrí aktuálny podobraz sa určí podľa toho akú farbu má pixel v maske, ktorý je na rovnakej pozícii ako stredový pixel aktuálneho podobrazu. Ak sa v maske nachádza biely pixel tak aktuálne okno patrí do triedy true (trieda true označuje okná kde sa nachádza roztrúsená skleróza), inak patrí do triedy false. Takýmto procesom sa vygenerovalo 8000 podobrazov pre trénovací a testovací dataset. Každý dataset bol rozdelený na triedu true a false, pre každú triedu bolo vygenerovaných 4000 podobrazov. Vygenerované podobrazy mali veľkosť 29×29 pixelov. Obrázok 1

zobrazuje snímok hlavy z magnetickej rezonancie, žltý štvorec predstavuje kĺzajúce okno použité na generovanie podobrazov pre neurónovú sieť. [1, 2, 5]



Obrázek 1: Obrázok zobrazuje snímok z magnetickej rezonancie a kĺzajúce okno (žltý štvorec).

Pretože vygenerované podobrazy rovnako ako originálne snímky sú čiernobiele a hodnoty pixelov sú v rozsahu od 0 do 255 je potrebné normalizovať hodnoty pixelov. Normalizácia vstupných dát pre neurónovú sieť je dôležitá pre efektívnejšie a rýchlejšie učenie neurónovej siete a minimalizáciu chýb. Normalizácia sa robila prepočítaním hodnôt pixelov z rozsahu od 0 do 255 do rozsahu od 0 do 1.



Obrázek 2: Vygenerované podobrazy pomocou algoritmu kĺzajúceho okna.

Obrázok 2 zobrazuje dva vygenerované podobrazy pred normalizáciou. Prvý podobraz patrí do triedy false, nezobrazuje roztrúsenú sklerózu. Druhý podobraz patrí do triedy true, zobrazuje roztrúsenú sklerózu. Po normalizácii podobrazov sa dvojrozmerné dáta previedli na vektor dát a pridalo sa k nim označenie do ktorej triedy jednotlivé podobrazy patria.

3 MODEL UMELEJ NEURÓNOVEJ SIETE

Na vytvorenie modelu umelej neurónovej siete bola použitá knižnica Deep Learning for Java (DL4J) [4]. V tejto knižnici bol vytvorený model umelej neurónovej siete s ôsmimi vrstvami. Pri trénovaní modelu neurónovej siete bol použitý krok učenia 0,000001, na učenie bolo použitých 8000 vygenerovaných okien pri 250 iteráciach a algoritmus spätného šírenia chýb. Aby sa zabránilo pretrénovaniu neurónovej siete, tak po každej trénovacej iterácii sa sieť otestovala na trénovacích dátach. Následne sa uložili aktuálne váhy a úspešnosť modelu. Po skončení trénovania sa vyhľadala najvyššia dosiahnutá úspešnosť modelu a použili sa váhy patriace k tejto úspešnosti. Tabuľka 1 zobrazuje vrstvy a parametre umelej neurónovej siete použitej pri segmentácii obrazu. Pri trénovaní modelu neurónovej siete sa dosiahla úspešnosť 96 %. [1]

Segmentácia obrazu pomocou natrénovanej siete sa robila rovnako ako generovanie okien. Vytvorilo sa okno ktoré sa normalizovalo a vložilo do natrénovanej siete. Výstupom siete boli dve pravdepodobnosti, kde každá určovala príslušnosť k jednej z dvoch tried. Podľa pravdepodobností sa určilo či

Vrstva neurónovej siete	Popis vrstvy neurónovej siete
Konvolučná vrstva 1	veľkosť filtru 5×5 pixelov, aktivačná funkcia typu Relu
Podvzorkovacia vrstva 1	veľkosť filtru 2×2 pixelov, výber pixelu s maximálnou hodnotou
Konvolučná vrstva 2	veľkosť filtru 2×2 pixelov, aktivačná funkcia typu Relu
Podvzorkovacia vrstva 2	veľkosť filtru 2×2 pixelov, výber pixelu s maximálnou hodnotou
Perceptronová vrstva 1	počet neurónov 500, aktivačná funkcia typu Relu
Perceptronová vrstva 2	počet neurónov 400, aktivačná funkcia typu Relu
Perceptronová vrstva 3	počet neurónov 300, aktivačná funkcia typu Relu
Výstupná vrstva	aktivačná funkcia typu Softmax

Tabulka 1: Tabuľka zobrazuje jednotlivé vrstvy a parametre umelej neurónovej siete.

pixel v strede podobrazu predstavuje alebo nepredstavuje roztrúsenú sklerózu. Výsledný obraz predstavoval masku k vstupnému snímku z magnetickej rezonancie. Biele pixely vo vytvorenej maske určovali roztrúsenú sklerózu označenú modelom neurónovej siete. Validácia natrénovanej neurónovej siete bola robená segmentáciou snímku. Dosiahnutá úspešnosť bola 64 % a pri validácii bolo použité nastavenie neurónovej siete z tabuľky 1. [1, 2, 5]

4 ZÁVER

Práca sa venuje trénovateľnej segmentácii obrazu s použitím hlbokých neurónových sietí. Trénovateľná segmentáci obrazu sa v práci využíva na segmentovanie snímkov hlavy z magnetickej rezonancie a následnú detekciu roztrúsenej sklerózy. V práci bol vytvorený algoritmus kĺzajúceho okna na generovanie podobrazov a model umelej neurónovej siete použitý na segmentovanie snímkov z magnetickej rezonancie. Pri trénovaní modelu neurónovej siete sa dosiahla úspešnosť 96 %. Pri validácii natrénovaného modelu neurónovej siete sa dosiahla úspešnosť 64 %. Takto natrénovaný model neurónovej siete nie je vhodný na detekciu roztrúsenej sklerózy kvôli nízkej úspešnosti pri segmentácii. Nízka úspešnosť je spôsobená slabým natrénovaním modelu neurónovej siete, malým počtom iterácií pri trénovaní a malým datasetom trénovacích dát. Ďalšie zvýšenie presnosti sa dosiahne použitím vyššieho počtu vygenerovaných podobrazov, ktoré budú použité na trénovanie umelej neurónovej siete a optimalizáciou jednotlivých parametrov neurónovej siete (počet vrstiev a neurónov).

REFERENCE

[1] NIELSEN, M.: Neural Networks and Deep Learning [online]. 2016 [cit. 2015-03-06]. Dostupné z: http://neuralnetworksanddeeplearning.com/

[2] KRIZHEVSKY, A., SUTSKEVER I., HINTON, G.: ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems. 2012, (25): 9. Dostupné z: http://www.cs.toronto.edu/%7Efritz/absps/imagenet.pdf

[3] Sclerosis multiplex: Ochorenie [online]. 2016 [cit. 2015-03-06]. Dostupné z: http://sclerosis-multiplex.sk/ochorenie/

[4] What is Deeplearning4j? [online]. 2016 [cit. 2015-03-06]. Dostupné z: http://deeplearning4j.org/

[5] CHEN, L., PAPANDREOU G., KOKKINOS I., MURPHY K., L. YUILLE A.: Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs [online]. 2014 [cit. 2015-03-27]. Dostupné z: http://arxiv.org/abs/1412.7062v3

Magisterské projekty

Automatizace a silnoproudá elektrotechnika

CLOSED-LOOP INTERFEROMETRIC FIBER-OPTIC GYROSCOPE WITH UNPOLARIZED LIGHT

Michal Skalský
Master Degree Programme (2), FEEC BUT
E-mail: xskals01@stud.feec.vutbr.cz

Supervised by: Zdeněk Havránek
E-mail: havranek@feec.vutbr.cz

Abstract: This article deals with utilizing special techniques to improve parameters of interferometric fiber-optic gyroscope (I-FOG) using unexpensive single-mode (SM) fiber and based on all-fiber components. Firstly, it was shown that polarizer must be used to compensate for depolarizing unperfections. Optimization of Er-doped superfluorescent fiber source (SFS) greatly increased SNR of the sensor. Finally, solution for closed-loop operation with slow piezoelectric phase modulator was found yielding significant extension of measurement range as well as its linearity.

Keywords: Closed-loop unpolarized fiber-optic gyroscope, all-fiber optical gyro, superfluorescent fiber source, polarizer, Lyot depolarizer, depolarized light modulation, piezoelectric phase modulator

1 ÚVOD

Optické gyroskopy jsou považovány za nejpřesnější snímače úhlové rychlosti a jsou tudíž využívány u systémů pro inerciální navigaci. Nacházejí uplatnění zejména v letectví, astronautice, vojenství či robotice. Vývoj optických vláknových gyroskopů (FOG) vedl ke snížení jejich vysoké ceny, avšak současná řešení jsou stále poměrně komplikovaná [1].

Dříve demonstrovaná sestava interferometrického optovláknového gyroskopu (I-FOG) [2] naznačila potenciál celovláknového uspořádání, měřicí rozsah i rozlišení však byly značne omezené. Cílem dalšího vývoje byla koncepce komplexního I-FOG ve zpětnovazebního uspořádání, umožňující značný měřicí rozsah s vysokým rozlišením, a to i při zachování nízkonákladových komponent.

2 PRINCIP INTERFEROMETRICKÉHO OPTOVLÁKNOVÉHO GYROSKOPU

I-FOG je založen na Sagnacovu jevu. Využívá interference svazků koherentního záření, které se šíří proti sobě optickým vláknem navinutým v několika kruhových závitech. Při rotaci mezi nimi vzniká fázový posuv, který je přímo úměrný úhlové rychlosti. Ten pak ovlivňuje míru interference, která je detekována jako intenzita výstupního záření I [W], daná

$$I = I_0(1 + \cos \phi_s), \tag{1}$$

kde I_0 [W] je střední hodnota intenzity a ϕ_s [rad] je fázový rozdíl svazků [1]. Převodní charakteristika je tedy nelineární a periodická, navíc nerozlišuje směr rotace.

3 ZAJIŠTĚNÍ RECIPROCITY

Nutnou podmínkou správné funkce I-FOG je reciprocita, tj. identita optických drah obou protichůdných svazků, a to i s ohledem na proměny jejich stavu polarizace [1].

Předchozí výzkum ukázal možnost realizace I-FOG v otevřené smyčce s použitím běžného jednovidového (SM) vlákna pracující s nepolarizovaným zářením [2]. Eliminace vlivu proměnné dvojlomnosti SM vlákna zde byla dosažena pomocí vláknového Lyotova depolarizátoru. I při tomto řešení však byla pozorována jistá zbytková nereciprocita znamenající nestabilitu výstupu. Ačkoli citlivost snímače byla poměrně vysoká (0,01 °/s), jeho absolutní přesnost byla více než o řád horší.

Přestože je pro zajištění stability nepolarizovaného řešení I-FOG použití Lyotova depolarizátoru teoreticky dostatečné [3], jeho vláknové provedení vykazuje některé parazitní vlivy. Především je to částečné prolínání polarizačních módů i ve vlákně zachovávajícím polarizaci (PM), kterým je tvořeno. Dále je nutné zdůraznit, že polarizace takto upraveného záření je pseudonáhodná a vychází pouze z omezené časové korelace záření širokopásmového zdroje.

Experimenty ukázaly, že výrazného zvýšení stability lze dosáhnout zařazením polarizátoru mezi vazební členy, jak ukazuje obr. 1. Přestože depolarizátor zajistí rozložení do všech polarizačních módů, polarizační filtr vybírá jen tu složku, která si z obou protibežných svazků odpovídá a má shodný fázový posuv. Zatímco u I-FOG s PM vláknem je polarizátor běžný [1], zde bylo zjištno, že součinnost polarizátoru a depolarizátoru umožňuje značně zlepšit parametry také při použití mnohem levnějšího SM vlákno.

4 ZDROJ NEPOLARIZOVANÉHO ZÁŘENÍ

Pro I-FOG se používají výhradně širokopásmové zdroje záření, zejména z důvodu potlačení vlivu Rayleighova rozptylu [1]. Z důvodu nižších nákladů a možnosti vysokého výkonu byl zvolen zdroj využívají stimulované spontánní emise erbiem dopovaného vlákna (SFS) [2]. Ukázalo se však, že významnou roli v účinnosti tohoto zdroje, tedy v transformaci budicího laserového záření o vlnové délce 980 nm na výstupní širokopásmové žáření s vlnovou délkou 1550 nm, hraje excitace atomů erbia. Přibližně šestinásobným zvýšením budicího výkonu tak došlo k šedesátinásobnému nárůstu výkonu výstupního.

Dále, přestože se jedná o jednoprůchodový SFS, jako nutné se ukázalo zařazení optického izolátoru, jak je patrné z obr. 1. Ten poté brání zpětnému vniku záření, které jinak způsobuje stimulaci převážně v opačném směru. Touto úpravou bylo dosaženo přibližně pětinásobného navýšení výkonu.

5 REALIZACE I-FOG V UZAVŘENÉ SMYČCE

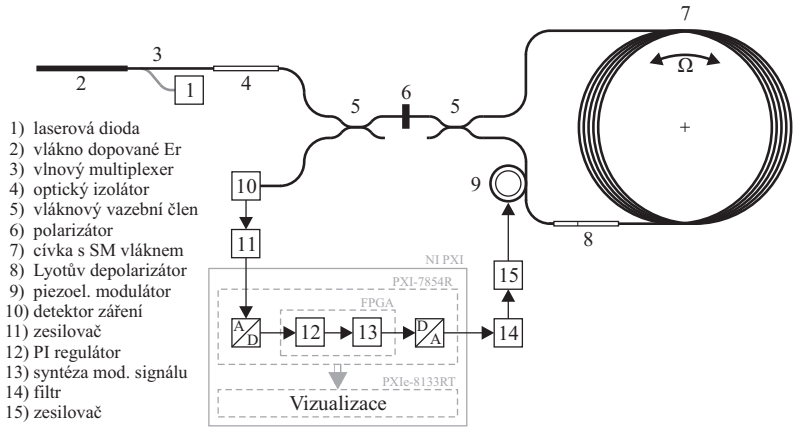
Zásadní nevýhoda předchozího řešení I-FOG v otevřené smyčce spočívala v analogovém zpracování signálu pomocí synchronní modulace-demodulace. Toto řešení však poskytuje pouze omezený měřicí rozsah a linearitu pouze v oblasti ±12 °/s [2].

Současné I-FOG využívají fázové modulace nejen pro rozmítání fáze pro zajištění maximální citlivosti a rozlišení směru, ale také pro kompenzaci nereciprokého Sagnacova fázového posuvu pomocí zpětné vazby. Detekovaný fázový posuv je pak trvale udržován v nule a generovaný akční zásah přímo odpovídá Sagnacovu posuvu, tedy měřené úhlové rychlosti [1].

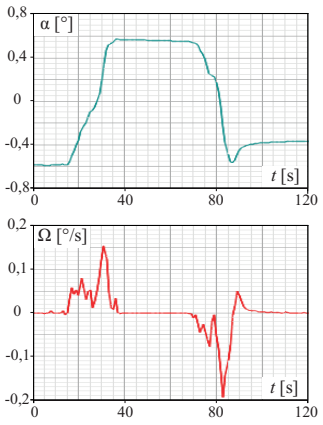
Standardní řešení I-FOG používá modulační schéma s velkou šířkou pásma, jaké lze uskutečnit pouze s použitím elektrooptického fázového modulátoru. Ten ovšem využívá dvojlomné struktury lithiumniobátu, lze proto použít pouze pro lineárně polarizované záření, tedy pro I-FOG s PM vláknem.

Pro I-FOG využívající nepolarizované záření a běžné SM vlákno je jediným řešením piezoelektrický fázový modulátor, vyznačující se úzkým použitelným kmitočtovým pásmem. Pro realizaci I-FOG v uzavřená smyčce s tímto modulátorem a s digitálním zpracováním signálu bylo navrženo speciální modulační schéma vyžadující vysoký stupeň synchronizace. Té bylo dosaženo implementací regulačního algoritmu pomocí FPGA modulu a PXI National Instruments v prostředí LabVIEW, viz obr. 1. Z regulačního hlediska se jedná o kompenzaci poruchy, která je způsobena úhlovou rychlostí.

Pro tento účel byl na základě simulace celého nelineárního systému navržen regulátor typu PI pomocí integrálního kritéria s ohledem na minimální chybu výstupní hodnoty I-FOG. Již první testování ukázalo nárůst rozsahu linearity snímače na ±300 °/s. Zároveň bylo dosaženo i zvýšení rozlišení, které je nyní lepší než rotace Země. Přitom stále nebylo využito maximálního optického výkonu ani rozsahu vstupního A/D převodníku, předpokládá se tedy ještě značné zvýšení SNR a tedy i rozlišení snímače. Očekávaná stabilita údaje o absolutní úhlové poloze získaná integrací úhlové rychlosti je v řádu jednotek stupňů na hodinu. Na obr. 2 je ukázka změřeného průběhu úhlové rychlosti a z ní dopočítané úhlové výchylky pomocí měřicího a vizualizačního programu.



Obrázek 1: Schéma sestaveného uspořádání I-FOG v uzavřené smyčce s nepolarizovaným zářením



Obrázek 2: Ukázka měření úhlové výchylky a rychlosti

6 ZÁVĚR

Podařilo se sestavit funkční prototyp I-FOG se zpětnovazební kompenzací fázového posuvu pomocí modulačního schématu navrženého speciálně pro celovláknový piezoelektrický fázový modulátor. Díky tomu došlo ke značnému rozšíření měřicího rozsahu i přesnosti snímače, a to při zachování výhod původního řešení, kterými bylo použití komponent z běžného SM vlákna a nepolarizovaného záření. Dále došlo k výrazné optimalizaci zdroje záření a odstranění zbytkové nereciprocity pomocí polarizátoru, což vedlo ke zlepšení rozlišení přibližně o řád. Byla tak ukázána a odzkoušena alternativní koncepce k současným komerčním I-FOG, avšak využívající pouze optovláknových prvků, SFS zdroje a běžného SM vlákna.

PODĚKOVÁNÍ

Tato publikace vznikla za podpory grantu číslo FEKT-S-14-2429 - „Výzkum nových řídicích metod, měřicích postupů a inteligentních prostředků v automatizaci“ financovaného z Interní grantové agentury Vysokého učení technického v Brně.

REFERENCE

[1] LEFÈVRE, H. C.: The Fiber-optic Gyroscope. 2nd ed. London: Artech House, © 2014. ISBN-13: 978-1-60807-695-6.

[2] SKALSKÝ, M.: Design of Depolarized Interferometric Fiber-Optic Gyroscope. In: Proceedings of the 20th Conference STUDENT EEICT 2014. Brno: VUT, 2014. ISBN 978-80-214-4922-0.

[3] KERSEY, A. D., et al.: Fiber Optic Gyroscopes with depolarized light. Inventors: A. D. Kersey, W. K. Burns a R. P. Moeller. United States Patent No. 5,319,440. 7. 6. 1994.

FIELD INSTRUMENT WITH HART INTERFACE

Jan Kunz
Master Degree Programme (2), FEEC BUT
E-mail: xkunzj00@stud.feec.vutbr.cz

Supervised by: Petr Beneš
E-mail: benesp@feec.vutbr.cz

Abstract: This paper describes design and development of of creating field instrument demonstration module. Module is capable of measuring multiple sensors such as thermocouples, RTD, or pressure sensor, analogue sensor simulation is also provided. Sensor's side is isolated from output, which is composed of 4 - 20 mA current loop and HART interface. Current loop also provides power supply and module can communicate via HART also when alarm current (3,2 mA) is set. Basic safety features like open wire detection, over and undervoltage protection are also implemented.

Keywords: current loop, HART, thermocouple, RTD, pressure sensor

1 ÚVOD

Senzory s proudovým výstupem jsou díky svým vlastnostem i dnes v praxi hojně rozšířené. S rozvojem nízkopříkonových součástek se objevují možnosti vytvářet stále sofistikovanější senzory napájené pouze proudovou smyčkou. K časté výbavě těchto senzorů patří HART komunikace s nadřízeným systémem, nebo galvanické oddělení snímačů. Zkombinovat obě tyto energeticky náročné možnosti a zachovat napájení pouze z proudové smyčky je velmi obtížné. Tvorbou právě takového modulu se tato práce zabývá.

Při napájení z proudové smyčky a možnosti komunikovat přes HART nesmí celková spotřeba modulu přesáhnout 2,7 mA. To je při použití galvanického oddělení, HART modemu a AD převodníku velmi obtížný úkol. Navíc galvanické oddělení nesmí svým šumem ovlivnit HART komunikaci a AD převodník musí být schopen napájet i snímače. Je potřeba pečlivý výběr jednotlivých komponentů, jejich optimální propojení a citlivé využívání řídícím programem. K AD převodníku bude navíc možné díky univerzálnímu obvodovému řešení připojit různé snímače, například termočlánek, platinový RTD snímač, nebo snímač tlaku.

Hlavním cílem této práce je vytvořit univerzální modul pro úpravu signálu z různých typů snímačů a s HART komunikací na proudové smyčce, ze které bude modul současně i napájen.

2 REALIZACE

Pro realizaci modulu musely být vyřešeny dva hlavní problémy. První je zajištění spotřeby celého modulu tak, aby mohl být modul napájen z proudové smyčky a zároveň nerušil HART komunikaci. Druhou je navržení univerzálních analogových obvodů pro připojení různých typů snímačů a vhodné zpracování jejich výstupů.

Pro dodržení maximálního proudového odběru byly vybrány integrované obvody kumulující do sebe co nejvíce užitečných funkcí, ale s co nejnižší spotřebou. Byl použit DA převodník AD5421 firmy Analog Devices. Jedná se o DA převodník, který je napájený z proudové smyčky a obsahuje dva nezávislé napěťové stabilizované výstupy, které byly nastaveny na 3,3 a 9 V. Jedno napětí (3,3 V)

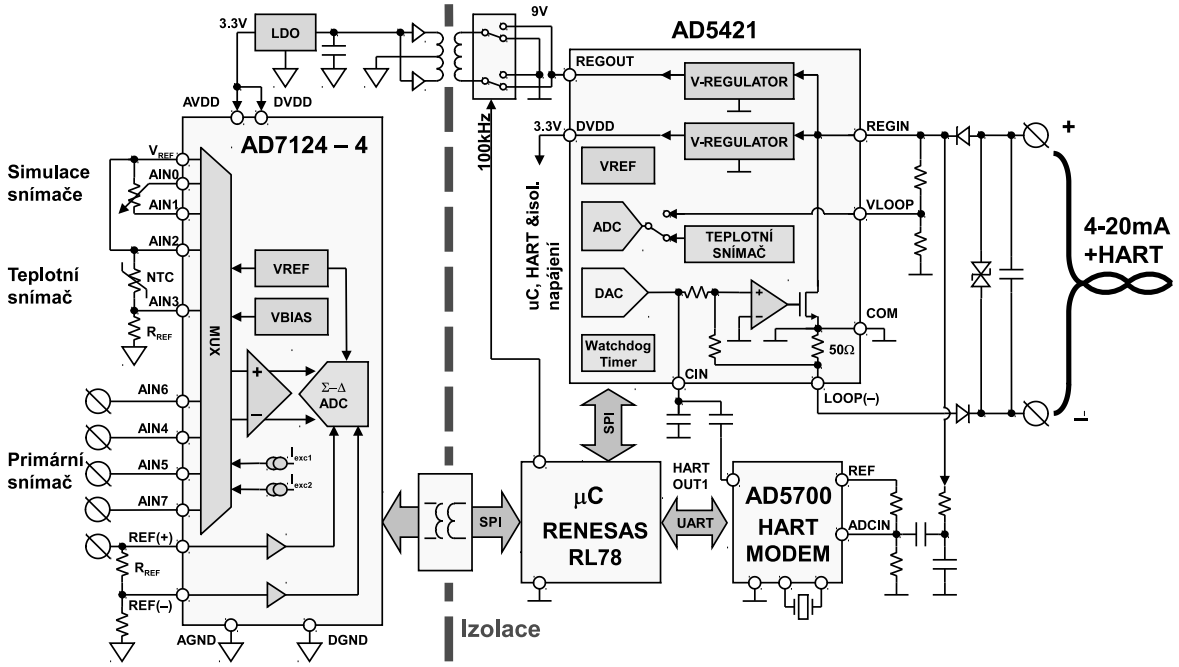
je použito pro napájení okolních obvodů a druhé (9 V) napájí přes transformátor (2:1) izolovanou senzorickou část modulu.

Velkým problémem je galvanické oddělení SPI komunikace s co nejnižší spotřebou. Spotřeba nejlepších opto-izolátorů se pohybuje nad 1 mA což je stále vysoká spotřeba pro tento modul. Proto musely být použity izolátory pracující na jiném principu. Byl vybrán izolátor ADuM1441 firmy Analog Devices, který pracuje na magnetickém principu a spotřebovává méně než 0,5 mA. Tato spotřeba je již dostatečně malá pro použití v tomto modulu.

Celkové řešení zapojení snímačů bylo vytvořeno tak, aby poskytovalo co nejvíce možností jejich zapojení. Platinový RTD snímač je možné měřit pomocí dvou-, tří-, i čtyřvodičové metody. Tenzometrický tlakový snímač (tvoří plný most) se dá excitovat proudově i napěťově. U každého měření je možné signál zesílit a dokonce i zjistit, zda není některý přívodní vodič přerušen a pokud ano, tak který. Je zde i možnost analogové simulace snímače, která dovoluje demonstrovat funkčnost modulu, i když snímač není připojen, nebo je problematické dosáhnout krajních poloh jeho rozsahu.

Z blokového schématu celého modulu (obr. 1) jde vidět základní funkčnost modulu. Modul je řízen 16-bitovým procesorem Renesas RL78, který ovládá periferie, zpracovává naměřená data a komunikuje s nadřízeným systémem. Vstupní signál jde přes přepěťovou ochranu do HART modemu (AD5700) a do DA převodníku. Příchozí HART zprávy jdou přes modem do procesoru, kde se zpracují a odchozí zprávy jdou přes modem do DA převodníku, kde se HART komunikace superponuje na proudovou smyčku. DA převodník reguluje proud procházející smyčkou, měří napětí ve smyčce (nebo teplotu) a současně z ní produkuje dvě napájecí napětí. Napětí 3,3 V napájí procesor, HART modem a přepínač. Druhé napájecí napětí (9 V) prochází přes galvanické oddělení, poté přes LDO regulátor a poskytuje napájecí napětí pro AD převodník (AD7124-4).

AD převodník měří signál primárního snímače, teplotu desky a analogovou simulaci primárního snímače. Primární snímač může být napájen dvěma proudovými zdroji v AD převodníku, nebo napájecím napětím. AD převodník také průběžně kontroluje zda je primární snímač správně připojen.



Obrázek 1: Blokové schéma modulu

3 PARAMETRY

Modul byl vytvořen, naprogramován a jeho parametry změřeny. Kromě celkové spotřeby se pozornost soustředila na přesnosti měření jednotlivých snímačů a šumové parametry. Spotřeba celého modulu při plné funkčnosti se pohybuje od 1,8 mA (u termočlánku) do 2,2 mA (tlakový snímač s napěťovou excitací). Tyto hodnoty jsou s velkou rezervou pod limitem, takže je možno dokonce uvažovat o přidání nějaké funkčnosti, například modulu pro bezdrátovou komunikaci.

Teplotu pomocí Pt100 RTD snímače teploty třídy A dle IEC 751 měří modul s přesností specifikovanou touto třídou v tří-, i čtyřvodičovém zapojení. Chyba měření je tedy mnohem menší než chyba snímače.[1] Z toho se dá se usuzovat, že je modul vhodný i pro měření kalibrovaných RTD snímačů, ovšem otestováno to nebylo.

Při měření termočlánkem je důležitá i přesnost měření teploty studeného spoje. Proto byla přesnost měření teploty termočlánkem rozdělena na dvě části. Při měření byl použit termočlánek typu T a teplota srovnávacího spoje byla měřena NTC termistorem. V první části byl srovnávací spoj držen na stálé teplotě (25 °C) a měřící spoj byl vystavován teplotám od 0 do 100 °C. Maximální chyba během tohoto měření nepřesáhla 0,5 °C. Ve druhé části byl měřící spoj v konstantní teplotě (25 °C) a byla měněna teplota srovnávacího spoje od 0 do 80 °C. I zde maximální chyba během měření nepřesáhla 0,5 °C. Chyba měření je zde způsobena hlavně chybami snímačů, jelikož termočlánek i NTC termistor mají chybu 1 °C (nebo 1% podle toho, která hodnota je větší). Teplotu srovnávacího spoje by šlo měřit s menší chybou, například platinovým RTD snímačem, ale měření teploty NTC termistorem se vzhledem k přesnosti termočlánku jeví jako dostatečné a levnější řešení. [2]

U měření tlaku byla ověřena funkčnost a proveden test na šum. Při obou způsobech napájení (napěťově a proudově) je efektivní rozlišení 19 (respektive 18) bitů, což je pro průmyslové použití dostatečné, jelikož nejpřesnější průmyslové snímače tlaku dosahují rozlišení 15 bitů.

4 ZÁVĚR

Byl vytvořen modul pro zpracování signálů ze snímačů s galvanickým oddělením od výstupu, který tvoří proudová smyčka a HART komunikace. Modul úspěšně prošel HART testy fyzické vrstvy. Maximální spotřeba modulu při plném provozu je 2,2 mA, což může být poskytnuto proudovou smyčkou a modul tak stačí připojit pouze dvěma dráty.

Modul je schopný měřit termočlánek, RTD snímač a snímač tlaku. Při testech těchto snímačů hrála vždy chyba daného snímače dominantní roli v celkové chybě. Navržený modul lze tedy rovnou použít pro běžná měření daných veličin.

Modul splňuje všechny požadavky, které na něj byly kladeny. Další vylepšení jsou možná v lepší stabilizaci izolovaného napájecího napětí a také je možno obohatit funkčnost modulu dalšími vymožeností, jelikož aktuální spotřeba leží 0,5 mA pod maximem.

REFERENCE

[1] KREIDL, Marcel. *Měření teploty: senzory a měřící obvody*. 1. vyd. Praha: BEN - technická literatura, 2005. Senzory neelektrických veličin. ISBN 80-7300-145-4.

[2] KESLER, Walt, James BRYANT a Walt JUNG. *Practical design techniques for sensor signal conditioning*. Norwood, Mass: Analog Devices, 1999, s. 7.2 -7.9. ISBN 0916550206.

[3] ANALOG DEVICES. AD5421: Data Sheet. rev. H.

[4] ANALOG DEVICES. AD5700: Data Sheet. rev. F.

[5] ANALOG DEVICES. AD7124-4: Data Sheet. rev. A.

SYSTEM FOR MEMS GYROSCOPE TESTING

Stanislav Hasík

Master Degree Programme (2), FEEC BUT

E-mail: xhasik02@stud.feec.vutbr.cz

Supervised by: Stanislav Klusáček

E-mail: klusacek@feec.vutbr.cz

Abstract: This paper presents fundamental information about an apparatus for tests on MEMS gyroscopes. Within the introductory section, the measurement system used to test MEMS gyroscopes at a Honeywell laboratory is characterized; the set-up comprises a Polytec MSA-500 analyzer, goniometers, and a controller. The related practical part then describes an application developed with the LabVIEW software to control the position of the goniometers. Importantly, the discussed system is able to fully control two goniometer stages and to align the MEMS device surface into an orthogonal position with respect to the Polytec MSA-500 measurement unit.

Keywords: MEMS gyroscope, Polytec MSA-500, Goniometer, Controller, Topography

1. ÚVOD

MEMS gyroskopy jsou Mikro-Elektro-Mechanické Snímače úhlové rychlosti vztažené k inerciální referenční soustavě a jde o snímače malých rozměrů a malé hmotnosti a nízkou spotřebou energie, což je jejich největší výhodou. MEMS gyroskopy jsou většinou založeny na měření Coriolisovy síly působící na pohybující se setrvačnou hmotu při současném působení úhlové rychlosti. Velikost Coriolisovy síly (F) je určena vztahem:

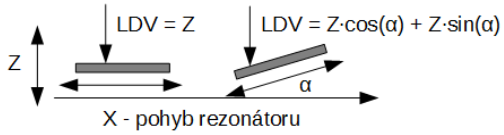
$$F = -2m\bar{\Omega} \times \bar{v} \tag{1}$$

Ve společnosti Honeywell International s.r.o. jsou vyvíjeny MEMS gyroskopy pro použití v inerciálních navigačních systémech pro letecký a vesmírný průmysl. Tyto senzory jsou tedy využívány jako snímače úhlového natočení získané integrací úhlové rychlosti měřené gyroskopem. Výsledná hodnota natočení je však při integraci zatěžována nepřesností měřené hodnoty a to především výstupní hodnotou při nulové úhlové rychlosti tzv. offsetem. Velikost chyby gyroskopu při nulové výchylce je rozhodující pro dlouhodobou integraci výstupu a tedy i pro přesnost inerciálního navigačního systému. Tato chyba offsetu je z velké části způsobena tím, že setrvačná hmota, elektrostaticky vybuzená do mechanické rezonance, nekmitá pouze v ose buzení, ale dochází k přenosu pohybu i do ostatních os a to zejména do osy, ve které dochází k měření Coriolisovy síly (viz **Obrázek 1**). Tímto jevem dochází k offsetu na výstupním signálu snímače. Tento přenos kmitání je z velké části způsoben nedokonalostmi výrobního procesu mechanické struktury gyroskopu, která u MEMS výrobních procesů dosahuje jednotek procent.

2. TESTOVÁNÍ MEMS GYROSKOPŮ

Jednou z nejpřesnějších metod jak testovat statické parametry gyroskopu (offset) je přímé měření výchylek setrvačné hmoty pomocí laserového vibrometru. Proto byl pro testování použit měřicí systém Polytec MSA-500, který obsahuje Laser Doppler Vibrometr (LDV) a dokáže měřit výchylku v kolmém směru s rozlišením 0,1 pm/√Hz. Aby bylo možné stanovit příčnou výchylku vybuzené setrvačné hmoty způsobenou nepřesnostmi při výrobním procesu (offset při nulovém natočení), je nutné nastavit klidovou polohu tak, aby rovina setrvačné hmoty byla dokonale kolmá k měřicí optice analyzátoru. Proto byl navržen komplexní měřicí řetězec, který mimo jiné

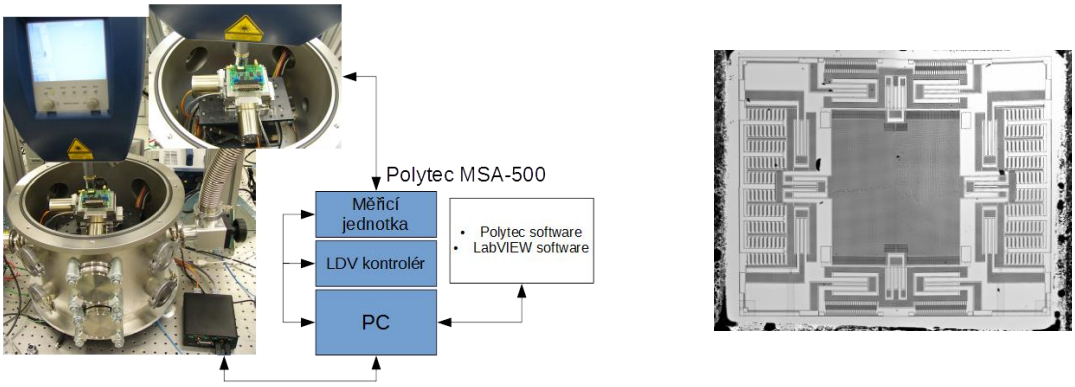
kompenzuje vliv náklonu struktury a tím umožňuje přesnější měření nedokonale pohybující se struktury gyroskopu, ve směru Z na obrázku 1, a určení mechanické složky offsetu.



Obrázek 1: Chyba měření způsobená náklonem

2.1. POPIS MĚŘICÍHO ŘETĚZCE

Pro dokonalé ustavení povrchu setrvačné hmoty do směru kolmého k měřicí optice LDV MSA-500 byla použita dvojice goniometrických plošin řízených prostřednictvím programu v prostředí NI LabVIEW. K získání referenčního údaje o náklonu zkoumaného povrchu se využívá další z vybavení měřícího analyzátoru MSA-500 a to topografické měření povrchu. Ukázka celé měřicí aparatury pro testování MEMS gyroskopů je na **Obrázek 2** a):



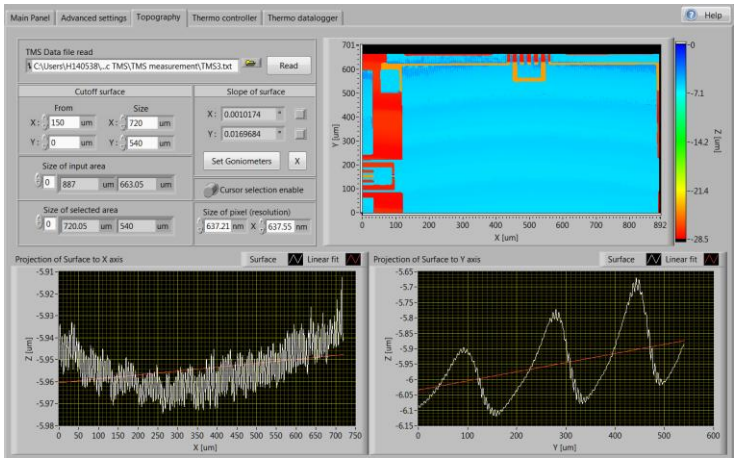
Obrázek 2: a) Měřicí aparatura pro testování gyroskopů; b) ukázka MEMS gyroskopu [1]

Příklad mechanické struktury MEMS gyroskopu je na **Obrázek 2** b). Jedná se o design vycházející z [1], jehož parametry (rezonanční frekvence, velikost kapacity řídicího a snímacího obvodu, atd.) byly upraveny tak, aby odpovídaly parametrům snímače vyvíjeného interně firmou Honeywell International s.r.o. Mechanická struktura gyroskopu je umístěna do LCC32 pouzdra, která společně s elektronikou tvoří gyroskopický systém. Tento sytém je poté upevněn na dvojici goniometrických plošin, které umožňují náklon systému v osách X a Y až o ±2,5° s rozlišením 10″. Celé toto zařízení je posléze umístěno do vakuové komory. Každý goniometr obsahuje pohon s krokovým motorem a dvojici koncových spínačů. Goniometry jsou řízeny kontrolérem, do kterého jsou po USB rozhraní posílány příkazy z prostředí LabVIEW. Kontrolér je zapojen do systému MSA-500 firmy Polytec, kde je spuštěn program pro obsluhu MSA-500, jak je naznačeno na **Obrázek 2** a).

2.2. POSTUP ZAROVNÁNÍ STRUKTURY

Měřený vzorek (gyroskop) je vložen do elektroniky a upevněn na povrch goniometrických plošin. Poté je pomocí software Polytec TMS změřena topografie povrchu části (setrvačná hmota) zkoumaného vzorku, kterou je potřeba zarovnat. Topografické měření povrchu je založeno na principu interference bílého světla Mirau interferometrem, jež je součástí optiky (**Obrázek 2** a). Ze software Polytec TMS je poté vyexportována matice hodnot (bodů), které udávají relativní výšku s přesností ±10 nm. Změřená matice bodů je načtena navrženou aplikací, která z hodnot opět sestaví topografickou mapu povrchu. Z těchto hodnot je pomocí průměrování získán průmět povrchu v osách X a Y. Tyto průměty jsou poté automaticky proloženy lineární závislostí a ze směrnice přímky je získán úhel náklonu zkoumaného vzorku v jednotlivých osách. Takto získaný úhel je použit pro vyrovnání (korekci) sklonu povrchu a tím zajištění ideální ortogonality vůči optice MSA-500. Chyba způsobená náklonem u tímto způsobem vyrovnané struktury je zanedbatelná a tuto strukturu je možné podrobit testování pomocí LDV. Přesnost topografického

měření umožňuje vyrovnat strukturu s přesností v jednotkách úhlových minut. Náklon struktury před vyrovnáním může být i v jednotkách stupňů, kvůli nepřesnostem při lepení struktury do pouzdra, umístění pouzdra do patice či samotné montáži zařízení do vakuové komory, která je nezbytná pro dosažení minimálního tlumení při pohybu struktury.



Obrázek 3: Vyrovnání struktury pomocí topografie: ukázka po vyrovnání

Před začátkem vyrovnání je doporučeno omezit zkoumanou oblast struktury (setrvačná hmota), protože velké změny výšky v oblastech pružin a nosníků mohou zhoršit přesnost vyrovnání. Před vyrovnáním byl sklon plochy v oblasti setrvačné hmoty cca 0,2° v ose X a 0,4° v ose Y. Po vyrovnání je v konečné iteraci (**Obrázek 3**), kdy byl cyklus měření a vyrovnání opakován, náklon roviny setrvačné hmoty asi 0,001°v ose X a 0,02° v ose Y. Zde se již dosáhne maximálního rozlišení topografie, kde dochází k jevu zvlnění naměřených dat vlivem vibrací a perforace struktury gyroskopu a také se projevuje vliv kmitání mechanického upevnění optické hlavy ve směru Y, viditelné na **Obrázek 2**.

3. ZÁVĚR

Byl navrhnut, vytvořen a ověřen měřicí systém a řídicí aplikace v prostředí LabVIEW, která je schopná plně řídit dvojici goniometrických plošin umístěných ve vakuové komoře. Vytvořený měřicí řetězec je schopen nastavit povrch zkoumané struktury do pozice kolmé vůči optice analyzátoru Polytec MSA-500 s přesností v rozmezí ±1.5′ za pomoci topografického měření. Na **Obrázek 3** je ukázka aplikace pro zpracování naměřených dat ze systému Polytec TMS, kde je možné jednoduchým způsobem provést korekci sklonu zkoumaného povrchu. Aplikace je rovněž schopna akceptovat změny v instalaci (orientaci) goniometrických plošin vůči měřicí jednotce MSA-500. Aplikace obsahuje další možnosti řízení goniometrických plošin pomocí úhlových souřadnic a dále algoritmy pro řízení Peltierových článků pro stabilizaci teploty měřeného zařízení.

PODĚKOVÁNÍ

Tento příspěvek vznikl za podpory společnosti Honeywell International s.r.o. divize Aerospace Advanced Technology Europe, za podpory grantu Technologické agentury České republiky číslo TE02000202 - "Pokročilé senzory a metody zpracování senzorových dat" a také grantu číslo FEKT-S-14-2429 - "Výzkum nových řídicích metod, měřicích postupů a inteligentních prostředků v automatizaci" financovaného z Interní grantové agentury Vysokého učení technického v Brně.

REFERENCE

[1] ALPER, Said Emre. MEMS gyroscopes for tactical-grade inertial measurement applications. 2005. PhD Thesis. MIDDLE EAST TECHNICAL UNIVERSITY. Dostupné také z: https://etd.lib.metu.edu.tr/upload/12606483/index.pdf

UNIVERSAL MEASUREMENT INTERFACE FOR DIGITAL AUDIO SIGNAL

Marek Gál

Master Degree Programme (2), FEEC BUT

E-mail: xgalma00@stud.feec.vutbr.cz

Supervised by: Petr Petyovský

E-mail: petyovsky@feec.vutbr.cz

Abstract: This article deals with a modification of existing project which is used as a helpful tool for tracking and measuring digital audio interface I2S. Modifications are based on practical one year experience when the device was tested and deals with new requirements for extension.

Keywords: Digital audio processing, I2S, USB, MCU, Codec, Isolator

1. ÚVOD

Cílem tohoto projektu je modifikace již existujícího zařízení [1], pracujícího jako specifická externí zvuková karta s běžným analogovým výstupem, doplněna také o digitální audio I2S rozhraním. Tato kombinace vytváří velmi užitečný nástroj pro testování a ladění zařízení, pracujících právě se zmíněným digitálním audio rozhraním. Příklad užití nalezneme například při sledování komunikace na sběrnici, případně lze signál injektovat do samotného I2S kanálu. Motivací na přepracování projektu bylo odhalení nedostatků a zcela nového potenciálu zvukové karty. Ta již také podstoupila přibližně roční testovací provoz.

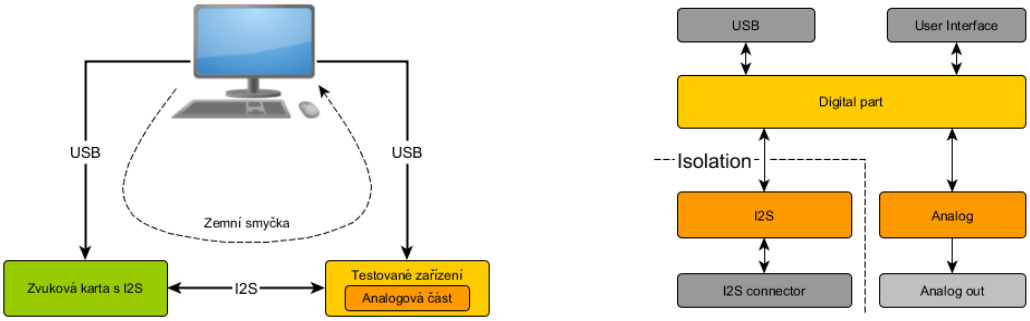
2. VÝCHOZÍ ZAŘÍZENÍ

V základu se projekt opírá o open source aplikaci nesoucí název SDR-Widget [2], která využívá rozhraní USB audio class 1 (AC1). Jedná se o standardizovaný přenos zvuk po sběrnici USB bez nutnosti instalovat ovladače. Projekt byl navržen pro 32 bitový mikrokontroleru (dále jen MCU) Atmel rodiny UC3 [3], který disponuje USB rozhraním a obsahuje také digitální audio sběrnici včetně režimu I2S. Firmware je postaven na real-time operačním systému FreeRTOS.

2.1. I2S ROZHRANÍ A JEJÍ

Jedná se o digitální stereo audio sběrnici, schopnou přenášet zvukový signál až s 32 bitovým rozlišením i na vysokých vzorkovacích frekvencích (například 96kHz). Linka obsahuje signál BCLK (bit clock) synchronizující jednotlivé přenášené bity, dále FSYNC (frame synchronization), který rozlišuje, zda jsou aktuální data platná pro levý nebo pravý kanál a dvojici datových linek RX a TX, po kterých jsou přenášena audio data. Z tohoto uspořádání plyne i topologie, která je typu master-slave a umožňuje připojení pouze dvojice navzájem komunikující zařízení. Toto rozhraní bylo v projektu zvoleno jako primární pro připojení externího testovaného zařízení především z důvodu své rozšířenosti.

Při nasazení zvukové karty vzniká velmi často modelová situace, kdy k jednomu počítači je prostřednictvím USB rozhraní připojena jak externí zvuková karta, tak i samotné testované zařízení (viz obrázek 1, vlevo). Tato konfigurace vede ke vzniku zemní smyčky, která může významným způsobem ovlivnit audio kvalitu analogové části testovaného zařízení. To vede ke zkreslení získaných informací. Aby se tomuto jevu zabránilo, je část připojení I2S rozhraní galvanicky oddělena. Izolace byla realizována za použití digitálních izolátorů firmy Silicon Labs, které umožňují přenos rychlých signálů vyskytujících se na I2S lince. Ty dosahují hodnot v řádu nejvýše desítek MHz.



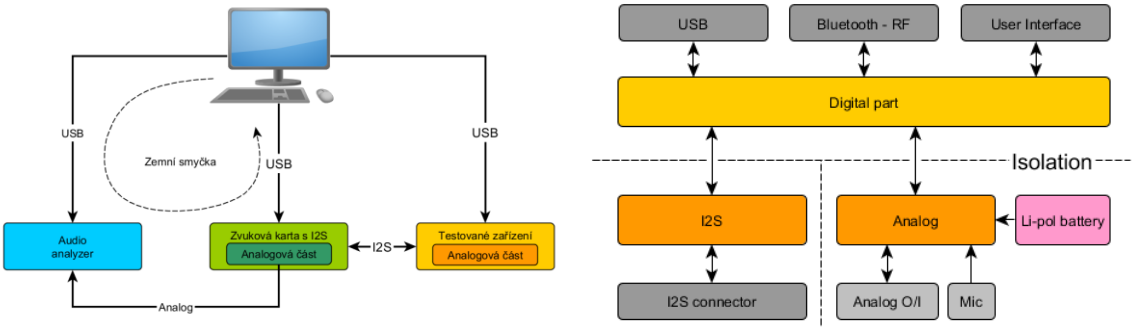
Obrázek 1: Vznik zemní smyčky (vlevo), izolace I2S rozhraní v původní variantě (vpravo)

2.2. POTÍŽE S PROVOZEM PŮVODNÍ VERZE

Přestože původní varianta zvukové karty naplnila své možnosti využití, vykazovala stálé nedostatky především v napájení izolované I2S části. Konektor bylo nutné napájet externím zdrojem ze samotného testovaného zařízení. Další potíže byly spojeny s analogovou částí zvukové karty, využívající kodek TLV320AIC33. Ten je napájen přímo z USB čímž je jeho zem galvanicky spojena s ovládacím počítačem. To opět vede ke vzniku zemní smyčky v případě, že chceme měřit analogový výstup ze zvukové karty (například pro ověření vlastností AD převodníku v testovaném zařízení). Situace je znázorněna na obrázku 2 vlevo.

3. NOVÉ ŘEŠENÍ

Hlavními body modifikace byla eliminace potíží popsaných v bodě 2.2. Nabízelo se také rozšířit stávající variantu o modul Bluetooth, který umožní testovat audioaplikace pracující právě na tomto standardu. Se začleněním tohoto prvku také vznikl požadavek na doplnění analogové části zvukové karty o vstupní část (linkový vstup a interní mikrofon). Tato funkcionalita zkomplikovala vnitřní propojení linky I2S, jelikož není možné, aby více než dvě zařízení současně bylo aktivně připojeno k lince. Vždy můžou komunikovat pouze dvě interní části zvukové karty. V původní verzi bylo spojeno MCU s I2S konektorem, přičemž kodek pouze sledoval jednu ze dvou datových linek. Nové řešení vyžaduje konfigurovatelné spojení mezi MCU, kodekem, připojitelným I2S rozhraním a modulem Bluetooth. Problém vyřešil návrh přepínače linky I2S, který je dále popsán v kapitole 3.1.



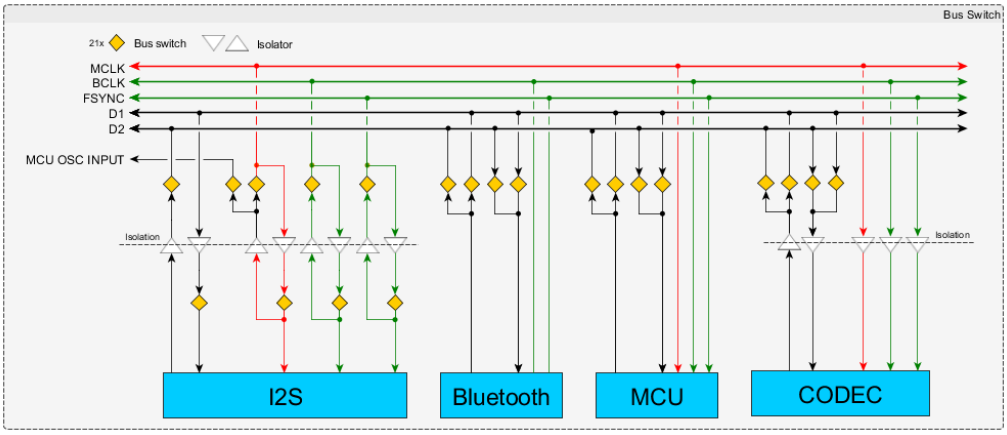
Obrázek 2: Problém zemní smyčky v neizolované analogové části (vlevo), nové vnitřní uspořádání (vpravo)

Pro zlepšení vlastností analogové části bylo rozhodnuto také o jejím galvanickém oddělení. Jako napájecí zdroj této části byl s ohledem na požadovaný nízký výstupní šum zvolen akumulátor. Nové vnitřní uspořádání se dvěma galvanicky oddělenými částmi zachycuje obrázek 2 vpravo.

Dalším požadavkem bylo zavedení zdroje hodin na vstup čítače MCU, což umožní měření frekvence na linkách I2S a uživateli usnadní konfiguraci zvukové karty na správnou přenosovou rychlost. Změny se týkají také firmwaru a jeho uživatelského rozhraní s ohledem na provedené modifikace.

3.1. PŘEPÍNAČ I2S SBĚRNICE

Přepínač signálů linky I2S s tak specifickými požadavky, jaké kladla aplikace v tomto projektu, bohužel není dostupný na trhu. Bylo tedy nutné navrhnout sofistikované řešení za použití třístavových jednobitových linkových driverů. Konkrétně bylo použito obvodů 74AHC125. Grafické znázornění celého propojení I2S linky je na obrázku 3. Figuruje zde také signál MCLK, který není nezbytně nutný pro funkci I2S rozhraní, ovšem některé zařízení jej vyžadují pro synchronizaci svého chodu. Jeho hodnota je obvykle udána v násobcích vzorkovací frekvence audiosignálu. Jedná se typicky o nejrychlejší hodinový signál, na což je třeba brát ohled při návrhu komponent.



Obrázek 3: blokové zapojení přepínače I2S sběrnice

4. ZÁVĚR

Návrh zařízení vychází z již existujícího projektu, který byl po dobu jednoho roku testován při vývoji aplikací využívající digitální audio I2S sběrnici. Byly identifikovány nedostatky spojené s jeho provozem a odhalen nový potenciál, který zahrnuje rozšíření o modul Bluetooth. Ten umožní testování aplikací pracujících s tímto standardem. Razantní modifikací prošla také analogová část zařízení, která byla doplněna o vstupní část v podobě vestavěného mikrofonu a linkového vstupu. Nově byla celá tato část také galvanicky oddělena od zbytku zařízení. Prvotní testy hardwaru již ukázali, že navržené modifikace skutečně splnili očekávání. V případě analogové části poklesl výstupní šum téměř od 10dB. Následující vývoj se bude především zaměřovat na úpravu obslužného firmwaru. Projekt by měl být po dokončení zveřejněn jako open-source aplikace.

REFERENCE

[1] STEJSKAL, M. Polymorfní USB – I^{2}S rozhraní. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2014. 81 s. Vedoucí diplomové práce: Ing. Martin Friedl.

[2] SDR-Widget [online], [citace 30.3.2016]. Dostupné na www: https://code.google.com/p/sdr-widget/

[3] ATMEL CORPORATION. *32-bit AVR Microcontroller: AT32UC3A3* [pdf]. ©2012, 1020 s. [cit. 2015-12-30].

IMPACT OF SURFACE-MOUNTED PERMANENT MAGNETS IMPERFECTIONS ON THE MACHINE

Adam Vícha

Master Degree Programme (2.), FEEC BUT

E-mail: xvicha02@stud.feec.vutbr.cz

Supervised by: Jan Bárta

E-mail: xbarta27@stud.feec.vutbr.cz

Abstract: Study sums up issues of mounting permanent magnets onto rotor surface of permanent magnet synchronous machine. Fully parametric model of rotor with surface mounted permanent magnet was created. Model was subjected to analysis of offset of permanent magnet, wrong magnetization and manufacturing width tolerance. Fourier transformation was performed with the results to separate the imperfections via cogging torque and induced voltages.

Keywords: magnet imperfections, cogging torque, induced voltage, FEM, synchronous machine

INTRODUCTION

Surface-mounted permanent magnet synchronous machines have a very wide spectrum of use. They are greatly popular due to high torque/volume density. But the design of permanent magnet machines has a few disadvantages. One of them is cogging torque, which is a result of interaction between magnetic field created by rotor permanent magnets and stator slotting. Cogging torque leads to undesirable vibrations and noise. Most of the motors are still produced by manufacturing and that can cause many different imperfections. Manufacturing imperfections highly increase the effect of cogging torque, so it‘s important to determine the impact of those imperfections. With knowing their impact it is possible to prevent them in design phase by applying certain measures.

1. HARMONIC COMPONENTS OF COGGING TORQUE

Cogging torque as a periodical phenomenon can be analyzed via Fourier transformation. Harmonic components of cogging torque depend on design, manufacturing and assembly tolerances. In order to characterize the impact of imperfections it is important to separate individual components from each other. Cogging torque harmonic components can be divided into native (NHC) and additional (AHC) harmonic components. [1]

$$T_{cog}(\alpha) = T_{NHC}(\alpha) + T_{AHC}(\alpha) \quad (1)$$

Native harmonic components are always present in the machine. Their origin is in the design of the machine. Additional harmonic components are caused by manufacturing and assembly imperfections of permanent magnets. Both components affect specific harmonic orders of cogging torque. Native harmonic components have orders: [1]

$$N_{NHC\,i} = LCM(Q, P) \cdot i \quad (2)$$

Where LCM stands for the least common multiple, Q is number of stator slots, P is number of magnetic poles and i is an integer.

Orders of additional harmonic components are:

$$N_{AHC\ rot\ i} = Q \cdot i \tag{3}$$

2. PARAMETRIC MODEL

A 2D parametric model was developed for this purpose. A synchronous machine with number of stator slots Q=36 and magnetic poles P=6 was chosen as a subject for analysis. It has four flat surface-mounted magnets per pole. The rotor model that was created is fully parametric. It allows unrestricted movement of permanent magnets and even different magnetization direction option.

This parametrization gives freedom to create the majority of possible setups of magnet imperfections. This study is focused on two main manufacturing imperfections that affect cogging torque the most.

The first setup is the offset of permanent magnet. This imperfection is caused by a worker using too much glue. Simulation via FEM is performed on one magnetic pole and magnet has only two possible states, 0mm and 1mm offset. That gives 2^4 possible combinations.

The second setup is false magnetization direction. This fault is a manufacturing tolerance. Simulation was performed on one half of the magnetic pole. Magnetization fault has three possible states, -10° direction from default, +10°, and the correct one - 0°.

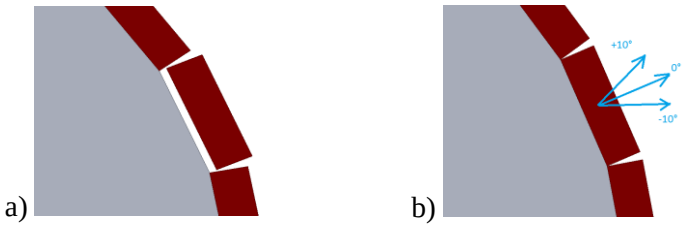


Figure 1: Simulated magnet imperfections a) offset b) false magnetization

3. SIMULATION RESULTS

A combination with the greatest influence from each simulation was chosen to demonstrate the result. Concerning the offset setup, an option when three magnets on one pole are affected by the fault and one of the middle magnets is mounted correctly. From the magnetization fault setup, the most influential combination is the one where magnetic flow diverges.

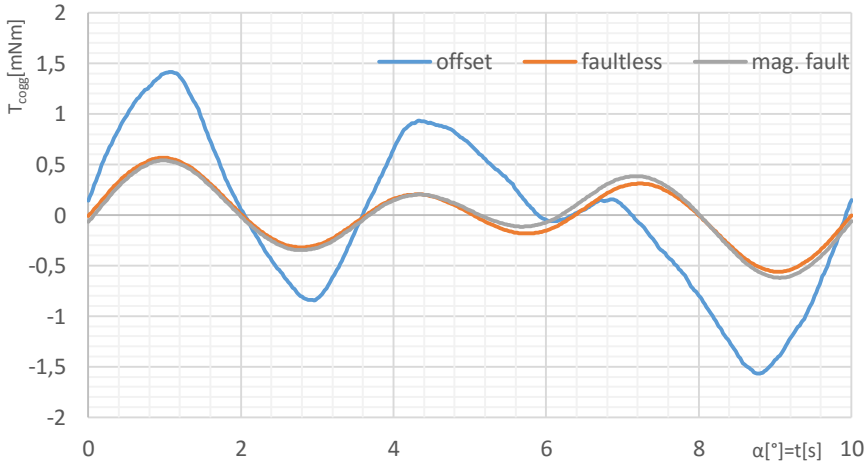


Figure 2: Cogging torque

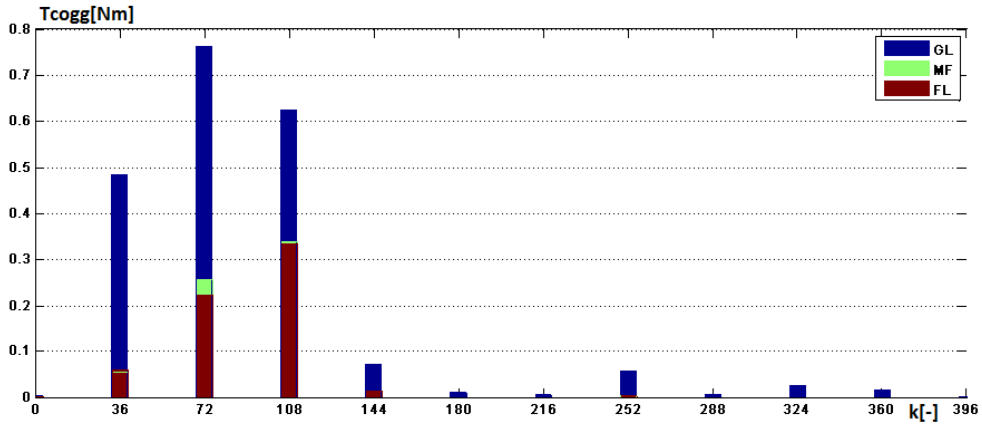


Figure 3: Cogging torque harmonic analysis

It is possible to see in Figure 2 and 3 that the offset imperfection (GL) affects cogging torque the most. Each harmonic order of cogging torque corresponds to equations in chapter 1. As for the design of the motor, native and additional harmonic components affect the same harmonic orders, because the least common multiple of number of stator slots and number of poles is the same as number of stator slots.

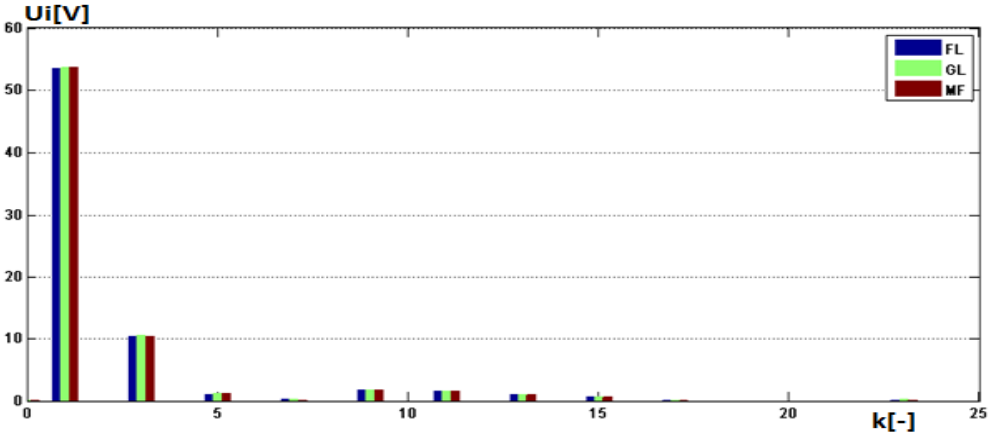


Figure 4: Harmonic analysis of induced voltage of phase A

In Figure 4 there is a harmonic analysis of induced voltages. The induced voltages are not affected as much as cogging torque. Closer examination shows that the offset imperfection slightly increases 1st, 3rd, and 5th harmonic. On the other hand, magnetization fault slightly increases 1st and 5th harmonic a slightly decreases 3rd and 7^{th}.

4. CONCLUSION

Simulations proved that surface-mounted permanent magnet imperfections have a significant affect on cogging torque. The biggest influence comes from the offset of permanent magnet. All simulation results were subjected to harmonic analysis. It is possible to see that each imperfection affects different harmonic orders of induced voltage. This could be the first step for detecting magnet imperfections via induced voltages.

REFERENCES

[1] Gasparin, L.; Cernigoj, A.; Markic, S.; Fiser, R., "Additional Cogging Torque Components in Permanent-Magnet Motors Due to Manufacturing Imperfections," *Magnetics, IEEE Transactions on* , vol.45, no.3, pp.1210,1213, March 2009

ROTOR SLOT OPTIMIZATION OF AN INDUCTION MOTOR

Vojtech Szekeres
Master Degree Programme (2.), FEEC BUT
E-mail: Vojtech.Szekeres@vutbr.cz

Supervised by: Jan Bárta
E-mail: Jan.Barta@phd.feec.vutbr.cz

Abstract: Proposed paper is dealing with optimization of an induction machine rotor slot by using artificial intelligence. Electrical machine under investigation is commonly manufactured induction motor with simple deep rotor bars. Goal is to achieve the highest possible value of efficiency and power factor. This paper is dealing with description of rotor parametric model and set-up of the optimization algorithm. Results of optimization are proposed as well.

Keywords: induction motor, rotor slot, optimization, SOMA, Ansys Maxwell, RMxprt

1 INTRODUCTION

In today's world, ecology and energy saving is one of the biggest topics. This trend continues in the field of electrical machines design as well. Much of electrical energy is consumed by electrical machines. Thanks to that, electrical machines represent huge potential for energy savings. Induction machines are one of the most used electrical machines nowadays. Optimization by using artificial intelligence is therefore very promising way how to increase induction machine's efficiency e.g. [1] and [2]. The one optimized is commonly manufactured induction machine with deep rotor bars.

2 OPTIMIZATION

Optimization is performed by Self-Organizing Migrating Algorithm (SOMA). It has been firstly published and described by Zelinka in [2]. This algorithm is fast and durable, comparable to genetic algorithms or other differential evolution algorithms and can be used to any optimization problem, written in any programming language. In this case, software MATLAB has been used. Optimization script was written in its own programming language. This programing language is based on Fortran, which is suitable for numerical and scientific computations. In most cases the MATLAB's script is perfectly compatible with GNU Octave, which is freely redistributable software. This script is written to cooperate with Ansys Maxwell tool RMxprt. RMxprt offers fast analytic computation of electrical machine that is accurate enough for proposed application.
Right definition of cost function is essential for optimization algorithm. Appropriate settings have major impact in terms of optimization results. In this work the cost function is used as follows:

$$f_{cost} = W_1 \left(\frac{\eta - \eta_w}{\eta_w} \right)^2 + W_2 \left(\frac{cos\varphi - cos\varphi_w}{cos\varphi_w} \right)^2 + W_3 \cdot smoothness \tag{1}$$

where η is calculated efficiency, η_w is wanted efficiency, cosφ is power factor, cosφ_w is wanted power factor. Coefficients determining priorities of optimization elements are marked as W. Optimization goals are set to be $\eta_w = 1$ and $cos\varphi_w = 1$, because there's no set value and the values wanted are high as possible. Weight coefficients are chosen as $W_1 = 2.5$, $W_2 = 2$ and $W_3 = 1$.
As seen, all elements in the formula (1) are squared. This means that computation is able to get to the wanted value from both sides - negative (minus) and/or positive (plus). Third element of this formula

is called smoothness. As the name itself implies, it is a function, that controlls the shape of slot's geometry in term of smooth transition between layers. This function compares neighbour layers by their width value. If these values are different, smoothness' value is risen by 1 then. Mathematic description of function is as follows:

$$\begin{aligned}
&\textbf{for} \quad i = 1 \quad \textbf{to} \quad i = n-1 \\
&\textbf{if} \quad b_i - (b_{i-1}) \geq 0 \quad \textbf{and} \quad b_{i+1} - (b_i) \geq 0 \quad \textbf{or} \quad b_i - (b_{i-1}) \leq 0 \quad \textbf{and} \quad b_{i+1} - (b_i) \leq 0 \\
&\textbf{then} \quad smoothness = smoothness + 0 \\
&\textbf{else} \quad smoothness = smoothness + 1
\end{aligned} \tag{2}$$

The more layers are in the parametrical model, the higher value may smoothness reach. Because efficiency and power factor are unable to reach value of 1, smoothness' values might cause bad optimization results.

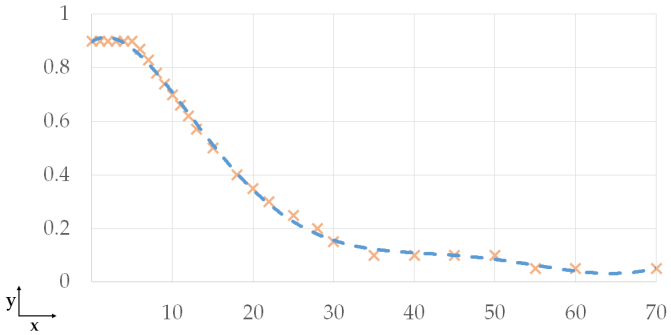


Figure 1: Custom function curve (x = smoothness, x = converted value)

That's why its value needs to be converted via a function, which values are in 0 to 1 range. To get a smooth shape of slot, smoothness' value needs to be small. First in mind is the exponential slope function e^{-x}, which has positive maximum in 0. Negative range is neglected, because smoothness can not achieve values lower than 0. Another solution is to define a custom function with chosen points and then create a curve by interpolation of these points. This gives the user a power to indirectly influence the result shape of slot or any optimized geometry. In this case the result function can be seen on fig.1 and its mathematical expresion is:

$$-2 \cdot 10^{-10} \cdot h^6 + 5 \cdot 10^{-8} \cdot h^5 - 6 \cdot 10^{-6} \cdot h^4 + 3 \cdot 10^{-4} \cdot h^3 - 0.0061 \cdot h^2 + 0.0189 \cdot h + 0.8979 \tag{3}$$

3 RESULTS

The optimization has been performed for following three variants:

A: Smoothness without any change. Its value is not controlled.

B: Smoothness as exponential function e^{-x}.

C: Smoothness as custom function, see equation 3.

The results were as follows.

parameter	marking	value				
		original machine (measuring)	original machine (simulation)	A	B	C
efficiency	η [-]	0.889	0.917	0.9175	0.916	0.919
power factor	cosφ [-]	0.83	0.77	0.86	0.87	0.85

Table 1: Optimization results (values)

Comparison of optimized shapes with original one is shown in Figure 2. As expected, because of the nature of exponential function is the slot geometry in case B very smooth. Steepness of the function in variant C is lower, so the smoothness is not pushed so hard to value of 0 and the geometry is much more varying.

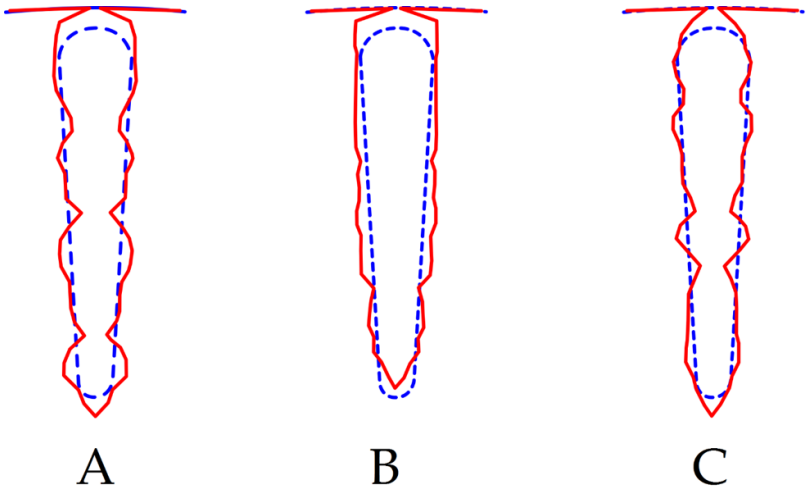


Figure 2: Optimization results (geometry)

By comparing observed values before and after optimization, it can be concluded that optimization lead to improvement of desired parameters. Highest efficiency shows case C, but best value of power factor has been reached in case B. It has the most smooth shape of slot as well, what makes it easier to manufacture. The one not mentioned, case A shows the second highest efficiency and power factor but the geometry is chaotic and it would be very hard to manufacture. Another important aspect of these geometries is mechanical strength.

4 CONCLUSION

Optimization is one of many aspects of improving efficiency and power factor and thus the ecology of an electrical machine itself. As shown, SOMA with little modifications and improvements is more than capable of beating this challenge not just in the field of optimizing rotor slot shape, but in improving any geometry in electrical machine. Only thing needed is a parametric model, which is easy to make in a software like Maxwell. That makes this kind of optimization available to any electric-machine-designing business. It can be concluded that parameters of electrical machine can be improved by using optimization aglorithms. Optimized electrical machine under investigation is going to be manufactured and measured. Presented paper can be easily used as a guidline for optimization of any induction machine rotor slot geometry.

REFERENCES

[1] CIBULKA, T. Možnosti optimalizace návrhu asynchronních strojů s využitím metod umělé inteligence. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2007. 95 s. Školitel disertační práce doc. Ing. Čestmír Ondrůšek, CSc

[2] ZELINKA, Ivan. Umělá inteligence v problémech globální optimalizace. 1. vyd. Praha: BEN - technická literatura, 2002, 189 s. ISBN 80-730-0069-5

[3] GALINDO, V.A., Xose M. LÓPEZ-FDEZ, J.A. DIAS PINTO a A. Paulo COIMBRA. Parametric study of rotor slot shape on a cage induction motor [online]. [cit. 2016-03-12].

DEVELOPMENT OF THE VERSATILE INDUSTRIAL MANIPULATOR

Michal Veteška

Master Degree Programme (2), FEEC BUT

E-mail: xvetes03@stud.feex.vutbr.cz

Supervised by: Jaromír Žák

E-mail: xzakja07@stud.feec.vutbr.cz

Abstract: This project deals with the universal industrianl robot. Design of mechanical parts, electronics and control system and a program for the microcontroller and the PC is described in this work. The industrial robot can be used in different environments. The precise operation of the robot is provided by a different sensors, mainly position sensor.

Keywords: industrial robot, microcontroller, position sensor

1. ÚVOD

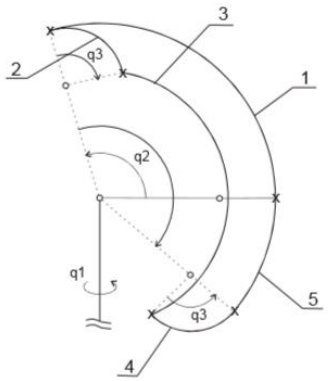
Robotické ruky alebo priemyselný roboti sú v súčasnosti veľmi rozšírenými robotmi. Ich využitie je hlavne v priemysle ale rovnako aj v lekárstve alebo v priestoroch, ktoré sú ťažko dostupné pre človeka. Základné koncepcie dosiahli v súčasnosti vysokej technickej úrovne a rovnako aj spoľahlivosti. Kvôli novým technickým odvetviam a lepšieho servisu sa však stále navrhujú nové koncepcie robotov manipulátorov.

Práca vznikla na podnet vytvorenia univerzálneho priemyselného robota, ktorý vie pracovať na rôznych pracovných miestach napríklad ako podávač, osadzovač apod. Pri jeho navrhovaní sa dbalo hlavne na jeho univerzálnosť a možnosť umiestniť ho na rôzne pracovné miesta s minimálnymi úpravami.

2. ROZBOR

2.1. MECHANIKA ROBOTA

Návrh mechanickej časti priemyselného robota bol odvodený z anregulárnej kinematickej štruktúry. Ide o štruktúru priemyselných robotov, ktorých pracovný priestor, teda priestor v ktorom sa robot môže pohybovať je daný obrysom plochy na seba nadväzujúcich oblúkov **Obrázok 1**. Tieto oblúky reprezentujú dosah jednotlivých ramien robota [1].



Obrázok 1: Prierez anregulárne kinematickej štruktúry priemyselného robota [1].

Mechanická konštrukcia je rozdelená do troch častí: spodnú časť, ktorú tvorí podstava a otočnica (pohyb q1), strednú časť, tvoriacu kolísku kde sú uložené motory a samotné rameno motora (pohyb q2). Poslednou častou je efektor robota (pohyb q3, viz. **Obrázok 1**). Vyrobenú a zostavenú konštrukciu zobrazuje **Obrázok 2**. Spodná časť je vyhotovená z nerezovej trubky, v ktorej je uchopené axialno-radiálne ložisko slúžiace na minimálne namáhanie motora. V ložisku je nalisovaná hliníková otočnica, na ktorej je postavené rameno robota. Na presné vystredenie a dodržaný konštantný dištanc medzi podstavou a otočnicou je použité ďalšie axiálne ložisko.

Stredná časť pozostáva z kolísky a ramien. V kolíske sú uložené štyri motory, ktoré zabezpečujú pohyb sústavy troch ramien. Týmto je zabezpečené odľahčenie ramien od váhy motorov. Krútiaci moment z motorov sa na ramená prenáša pomocou remeníc a remeňov. Tento spôsob prenosu momentu bol zvolený preto aby bola dosiahnutá maximálna pohyblivosť ramien. Pri iných spôsoboch prenosu, napríklad pomocou šróbovice by nebol dosiahnutý uhol otáčania ramien 180°. Kolíska aj ramená sú vyhotovené z nerezovej oceli. Celková efektívna dĺžka ramena bez efektora bola simuláciou stanovená na 760 mm.

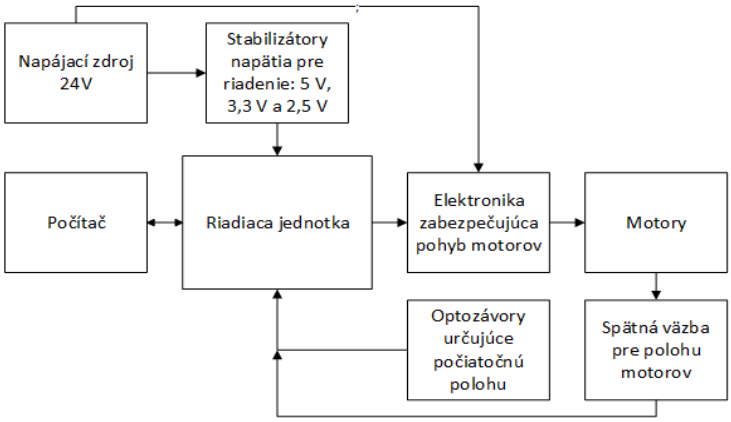
Efektor je uchytený na kužeľovom súkolí, ktoré zabezpečuje rotačný a súčasne aj zdvíhací pohyb. Kliešte efektora sa pohybujú pomocou šrobovice a šróbu, tie sú uchytené v kolíske. Na to aby kliešte ostali stále pod uhlom 90° je použitá sústava troch pák, simuláciou vyplynulo, že ich výstupný uhol musí byť vždy rovný 180°. Pokiaľ je táto podmienka splnená tak kliešte ostanú pri otváraní a zatváraní stále v kolmej polohe. Kliešte robota sú odnímateľné a vymeniteľné. Pre robota boli vyvinuté špeciálne kliešte, ktoré majú tendenciu sa prispôsobovať predmetu. Túto vlastnosť získali pomocou ich vyrobenia z pružného materiálu a vystuženia pevným materiálom. Obdobný koncept bol použitý napríklad v [2].



Obrázok 2: Mechanická konštrukcia robota bez koncového efektora

2.2. ELEKTRONIKA ROBOTA

Elektronika robota sa skladá z napájania, pohonu, riadenia pohybu motorov, spätnej väzby motorov, komunikácie medzi riadením a pc a elektroniku pre koordináciu robota v priestore. Základnú štruktúru zobrazuje **Obrázok 3**.



Obrázok 3: Principiálna schéma elektronickej časti robota.

Celá sústava je napájaná pomocou 24 V spínaného zdroja. Týmto napätím sú priamo napájané motory, pre riadenie muselo byť napätie stabilizované na referenčné hodnoty 5 V, 3,3 V a 2,5 V. Riadiacu jednotku tvorí procesor s jadrom ARM Cortex M3 a FPGA Cyclone IV. Procesor slúži na komunikáciu s počítačom a na nastavovanie polohy motorov zatiaľ čo hradlové pole spracováva výstupné dáta zo senzorov. Tato kombinácia bola zvolená pre optimálnu implementáciu komunikačného rozhrania so zachovanou rýchlou odozvou na detegovanú zmenu polohy ramien.

Na základe veľkosti a váhy ramena bola vypočítaná minimálna hodnota krútiaceho momentu motora a to tak, aby sa rameno zdvihlo z najhoršej polohy, ktorú reprezentuje vodorovná poloha ramena z podložkou. Hodnota krútiaceho momentu potrebného na zdvihnutie ramena je 16,5 Nm. Na základe toho boli vybrané hybridné krokové motory s prevodovkou, ktorých krútiaci moment je 34 Nm. Takto je zabezpečená dostatočná rezerva aby sa rameno vedelo zdvihnúť aj s nákladom o definované hmotnosti 500 g.

Pre spätnú väzbu polohy motorov a ramien slúžia polohové senzory, od spoločnosti InvenSense typ MPU-6050. Tieto senzory sú pripevnené ku každému ramenu robota. Ich úlohou je doladiť rameno do požadovanej polohy ak by došlo k prekĺznutiu kroku krokového motora. Tým je zabezpečená presnosť relatívneho pohybu. Pri každom spustený sa robot musí nastaviť absolútne do svojej počiatočnej polohy. Tento úkon je zabezpečený pomocou optozávor ktoré sú taktiež prichytené k ramenám robota.

2.3. PROGRAM ROBOTA

Program robota je rozdelený na dve hlavné časti a to na program pre riadenie elektroniky robota, spracúvanie signálov a zabezpečenie prenosu dát do počítača a na riadiaci program, ktorý je nainštalovaný v počítači.

Program v riadiacej jednotke má nasledovnú postupnosť. Pri zapnutí zariadenia MCU automaticky nastaví ramená do počiatočnej pozície. Následne očakáva dáta od programu v počítači, ktoré udajú uhol natočenia pre jednotlivé ramená. Tieto dáta sú predané pomocou paralelnej komunikácie do FPGA, ktoré začne zbierať údaje od polohových senzorov. Súčasne začne procesor krokovať motory. Po dokončení krokovania sú MCU vyžiadané dáta od FPGA, ktoré pošle koľko krokov je potrebných, aby bola dosiahnutá požadovaná poloha. Pokiaľ je dosiahnutá poloha FPGA pošle údaj o zhode, následne je odoslaný údaj do počítača o dosiahnutí požadovanej polohy a procesor čaká na ďalšie dáta. Týmto spôsobom využitia spolupráce FPGA s MCU dáva priestor MCU presne nastaviť súbežne všetky krokové motory, zatiaľ čo FPGA spracováva dáta od senzorov a upravuje ich tak aby mali čo najpresnejšiu hodnotu. Komunikácia medzi počítačom a procesorom je zabezpečená pomocou USB zbernice, v ktorej je nainštalovaný procesor ako virtuálny COM port.

Program v počítači je navrhnutý v prostredí windows form. Je v ňom možné nastavovať sekvencie, ktoré robot bude vykonávať. Tiež zabezpečuje ukladanie a načítanie sekvencii.

3. ZÁVER

Výsledkom tejto práce je navrhnutý a skonštruovaný univerzálny priemyselný robot, ktorého je možno zaviesť do rôznej prevádzky. Taktiež naprogramovaný riadiaci program pre počítač a pre MCU. Ďalej bola navrhnutá elektronika riadiaca krokové motory a zabezpečenie polohy robota v priestore pomocou senzorov. Práca taktiež popisuje teoretické znalosti z oblasti mechaniky priemyselných robotov a ich riadenia. Táto časť práce je však vzhľadom ku svojej rozsiahlosti v článku popísaná iba okrajovo.

REFERENCE

[1] ČERVEŇAK, A. *Hodnotenie pracovného priestoru robota*, 2005, s. 14.

[2] STOLL, Wilfried. *BionicTripod with FinGripper* [online]. [cit. 2016-03-07]. Dostupné z: https://www.festo.com/rep/en_corp/assets/pdf/Tripod_en.pdf

Magisterské projekty

Mikroelektronika a technologie

ANALYSIS OF APPLICATION FLUX AND SOLDER PASTE ON PCB FOR BGA COMPONENTS

Michal Toufar

Master Degree Programme (2), FEEC BUT

E-mail: xtoufa05@stud.feec.vutbr.cz

Supervised by: Radek Vala

E-mail: xvalar00@stud.feec.vutbr.cz

Abstract: This paper deals with rework of BGA components. The first part is focused on defects and errors in a solder joint. The main part chooses two methods for application flux and solders paste. The final part is focused on dipping and dispensing.

Keywords: BGA, flux, solder, repair

1. ÚVOD

Dosavadní trend vývoje v mikroelektronice a jejím odvětví zabývajícím se zapouzdřením polovodičových čipů má neustálé zvyšující se nároky dosáhnout co nejmenší zástavbové plochy na DPS při stále se zvyšujícím počtu kontaktů s ohledem na dosažení vysoké spolehlivosti se zachováním nízké ceny. Používané řešení s vývody po obvodu pouzdra (např. DIL) začalo být pro některé aplikace již nevhodné a z tohoto důvodu se začala více prosazovat pouzdra s vývody v maticovém uspořádání po celé jeho ploše, např. BGA (Ball Grid Array). Kdy bylo možné dosáhnout vysokého počtu připojovacích kontaktů na ploše pouzdra s pájkovými kulovými kontakty.

Z environmentálního pohledu je pro opravu zařízení v případě jeho poruchy spotřebováno mnohem méně surovin, energie a pitné vody než pro výrobu celého nového komponentu. Tímto se zamezí produkci toxických odpadů vznikajících při výrobě nového zařízení. Z tohoto důvodu lze efektivněji využívat obnovitelných a neobnovitelných zdrojů se zachováním zásob pro budoucí generace.

Z těchto mnoha důvodu je velice efektivní neustále vyvíjet, testovat a inovovat známé metody aplikace tavidla a pájecí pasty pro výměny BGA pouzder u opravovaných DPS (deskách plošných spojů) s následnou optimalizací procesu pro zvýšení jakosti pájených spojů, která zahrnuje i spolehlivost.

2. PROCES OPRAVY KOMPONENTU BGA

Problematika zahrnuje komplexní proces, kde jednotlivé operace postupují s návazností zahrnující dílčí kroky pro úspěšnou opravu. Operací předcházející aplikaci tavidla nebo pájecí pasty je odebrání vadného BGA pouzdra v opravárenské stanici s následným očištěním DPS od zbytku cínu a tavidla vhodným způsobem tak, aby nebyla poškozena pájecí maska nebo kontaktní plošky na DPS.

S přípravou na osazení nového BGA pouzdra je nezbytný výběr vhodné metody nanášení pájecí pasty nebo tavidla, aby se předcházelo defektům pájených spojů s následným vytvořením jakostního spoje. U špatné zvoleného dávkování může docházet tzv. „studeným spojům“ jde o špatné či vůbec elektricky nepropojené kontakty. Pouzdro může také tzv. „uplavat“ čímž je myšleno, že po sesouhlasení v opravárenské stanici BGA pouzdro s pájecími kontakty je následně usazeno na definovanou pozici u DPS. Avšak při procesu pájení přetavením se pouzdro následně posune s pájecími kontakty na vedlejší pozici pájecích plošek u DPS, než bylo původně usazeno. Lze zahrnout také důležité parametry, mezi které se řadí vhodně nastavený pájecí profil v opravárenské sta-

nici, ovlivnění procesu pájení přetavením s rychlou změnou okolního klimatického vlivu, způsobem odsávání vytvořených zplodin z naneseného tavidla a neposlední řadě také rychlost chlazení. Poté mohou nastat různé vady u pájených spojů nejčastěji vyskytující se v procesu oprav, které jsou popsány níže.

2.1. VOIDY

Jedná se o otvory nebo nekovové části v pájeném spoji. Obecně obsahují plyny a další látky z uvolněných tavidel, které se během pájení odpařují a nejsou schopny uniknout z roztaveného kovu před jeho zatuhnutím. Analýzu výskytu defektu lze vyhodnocovat rentgenovými snímky s následným zhodnocením dle normy IPC 610, zda se jedná o akceptovatelný nebo neakceptovatelný defekt. Snižujícím se průměrem pájecího kontaktu u aplikaci vystavující se cyklickému teplotnímu namáhání a vibracím je nutno tento defekt minimalizovat neboť tyto spoje jsou méně mechanicky odolné a jsou tedy náchylnější na prasknutí. Tento jev lze minimalizovat změnou množství a typem naneseného tavidla. V neposlední řadě je nutné také zařadit optimalizaci pájecího profilu.

2.2. ZKRATY

Mezi kořenové příčiny mohou být posunutá nepájivá maska již z výroby nebo při procesu čištění DPS vznikne defekt odloupnutím nepájivé masky s následným odhalením vodivé cesty. Následkem je snížení izolační mezery mezi pájenými ploškami. Kulové kontakty se následně roztékají po odhalené mědí, čímž vznikne elektrické spojení mezi sousedními pozicemi a deformace kulových kontaktů do atypických tvarů, které lze vyhodnotit optickou inspekcí nebo na rentgenových snímcích. Další příčinou vzniku zkratů může být nanesení nadměrného množství pájecí pasty na kontakty.

3. ZVOLENÉ METODY APLIKACE PRO PÁJECÍ PASTY A TAVIDLA

Byl proveden předvýběr vhodných metod dle zvolených kritérii pro danou problematiku. Z tohoto byly zvoleny dvě odlišné metody, u kterých probíhají testy a následně budou posouzeny na výskyt defektů a zhodnocení jakosti vytvořených pájených spojů. Dle předchozího testování na které je navázáno, byla vyhodnocena vhodná tavidla a aplikace je prováděna na pájecí kontakty pouzdra BGA. Testované metody se nazývají Dipping a Dispenzování.

3.1. DIPPING

Dipping je metoda namáčení nebo ponořování pájecího kontaktu. Vhodná je především pro pouzdra, která mají již předem osazené kulové kontakty. Pro metodu dipping jsou používány kovové přípravky s přesně vybroušenou rovinnou plochou pro aplikaci tavidla nebo pájecí pasty. Přípravek je vybaven drážkou pro aplikaci výšky naneseného media. Pro vytváření přesné výšky je používána aplikační stěrka.

Nejprve se nanese na okraj přípravku tavidlo nebo pájecí pasta a poté se aplikační stěrkou vytvoří rovinná vrstva. Následující operací je definované ponoření pájecích kontaktů BGA pouzdra dle nanesené výšky a po navrácení do výchozí polohy jsou viditelné otisky ve vytvořeném filmu a je naneseno medium na pájecích kontaktech. Tato metoda je vhodná pro vybrané typy pájecí pasty a tavidla.

3.2. DISPENZOVÁNÍ

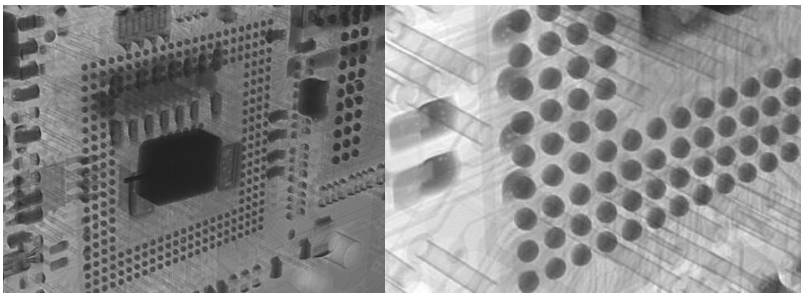
Dispenzer je zařízení ovládající přesné dávkování množství neboli dispenzování. Dle typu řízení lze rozdělit na dvě základní skupiny pístové a šroubové. U zařízení, které je řízené působícím tlakem po daný čas není dosahováno natolik přesného reprodukovatelného dávkování jak u typu řízeného Archimedovým šroubem. U zařízení dávkujícího přes přesnou šroubovici se nanesené množství definuje otočením šroubovice o daný úhel. Mezi hlavní výhody této metody lze zařadit snadné

řízení dávkovaného množství volbou vhodného průměru kapiláry. Jedná se o proces přesného reprodukovatelného dávkování pájecí pasty a tavidla na definovaná místa pouzdra BGA.

Dávkování je řízeno zařízením zvaným dispenzer a řízení pozice probíhá s automatickým strojovým naváděním s uloženým programem popisujícím trajektorii a dávkování. Tímto je metoda automatizovaná a lze minimalizovat vliv lidského faktoru ovlivňující nanesené množství při zachování reprodukovatelnosti procesu.

Ovlivňující faktor je typ použitého tavidla a jeho nanesené množství. Neboť překročení určitého množství dávkovaného tavidla může pouzdro při procesu pájení způsobit „zaplavání“ pouzdra nebo „studené spoje“ při nedostatečném naneseném množství.

Probíhající testy pro aplikaci tavidla na BGA komponenty mají pro tuto metodu pozitivní výsledky s minimálním výskytem chybovostí. Na obrázku 1 je patrné pouzdro zapájené v opravárenské stanici po aplikaci tavidla dispenzerem. Tavidlo je nanášeno přes kapiláru na pájecí kontakty BGA pouzdra dle nastavené trajektorie dávkování s následným vytvořením obrazce. Pájecí pasta je nanášena na DPS dle přednastavených souřadnic bodově na kontaktní plošky.



Obrázek 1: Zapájené pouzdro BGA po aplikaci tavidla metodou dispenzování (vpravo přiblížený pohled na zapájené kulové spoje)

4. ZÁVĚR

Jedná se pouze o nastínění problematiky výměny BGA pouzdra v procesu opravárenství se zaměřením na aplikaci tavidla a pájecí pasty. Byly zvoleny dvě hlavní metody, které jsou testovány. Metoda dispenzování je vhodná pro tisk různých geometrických útvarů nebo bodů i pro Fine Pitch rozteč. Metoda dipping je testována při různých nanesených tloušťkách po dobu ponoření a vliv naneseného množství pro výslednou jakost pájeného spoje. Obě metody jsou testovány pro různá tavidla a pájecí pasty a následným vyhodnocením na jakost pájeného spoje s následujícími optimalizacemi stávajících procesů v společnosti Sanmina.

PODĚKOVÁNÍ

Děkuji společnosti SANMINA Czech Republic s.r.o. a jejím zaměstnancům, zejména Ing. Radku Valovi za pomoc, rady a možnost zpracovávat práci v prostorách firmy.

REFERENCE

[1] JANÍČEK, Martin. ZVÝŠENÍ VÝTĚŽNOSTI BGA OPRAVÁRENSKÉHO PROCESU. Brno, 2015. Diplomová práce. Vysoké učení technické v Brně. Vedoucí práce Ing. Bc. Pavel Řihák.

[2] *SMT-INFO KONSORCIUM*. [online]. Brno: BRNO SMT-info, 2014 [cit. 2015-11-20]. ISBN 1211-6947. Dostupné z: http://www.smtinfo.cz/download/bulletiny/SMT_bulletin1410.pdf

[3] BOGATIN, E. *GRID ARRAY PACKAGING: BGA AND CSP*. Smithsonian: The Chip Collection [online]. [cit. 2015-10-10]. Dostupné z: http://smithsonianchips.si.edu/ice/cd/PKG_BK/CHAPT_10.PDF

ANALYSIS OF DEFECTS ON PCB USING X-RAY, 3D SW CERA AND MICRO-SECTIONS

Martin Vala

Master Degree Programme (2), FEEC BUT

E-mail: xvalam02@stud.feec.vutbr.cz

Supervised by: Pavel Řihák

E-mail: xrihak02@stud.feec.vutbr.cz

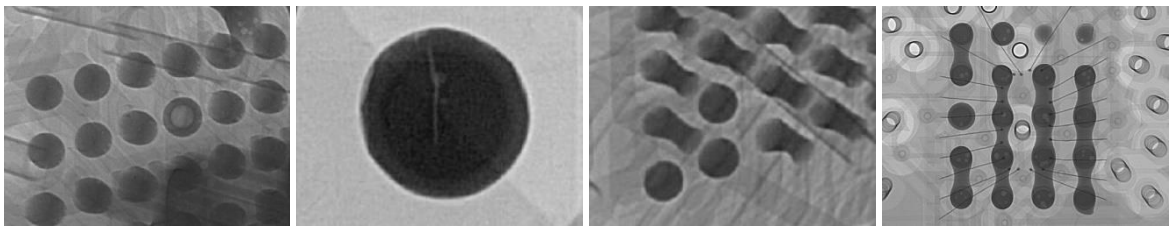
Abstract: This paper deals with detecting of defects on BGA (Ball Grid Array) components using X-ray. Defects are formed throw reflow process of BGA components during assembly, but also later due to mechanical and thermal stress. Therefore, there is an overview of defects and methods of diagnosis of BGA packages eg.: modern X-ray defect detection

Keywords: X-ray, Defect, Analysis, X-plane, Void

1. DEFEKTY NA DPS

Defekty vznikají nejčastěji při samotné výrobě zařízení (osazování nebo pájení součástek), ale mohou vznikat i později při používání výrobku. V průběhu životního cyklu je zařízení vystaveno vnějším vlivům okolí např. mechanické nebo tepelné namáhání. Toto namáhání může způsobit vznik nových defektů nebo rozšiřování těch, které vznikly již při výrobě.

Pájený spoj je jedním z kritických míst, kde je statisticky výskyt defektů nejčastější. Pájený spoj, zajišťuje mechanické, ale především elektrické propojení součástky s DPS (deskou plošných spojů). Zapájený spoj také zajišťuje odvod tepla ze součástky, tím se zajistí lepší chlazení čipu. S technologickým vývojem narůstal počet vývodů pouzder, což vedlo ke zmenšování pájených spojů a také umístění vývodů na spodní stranu součástek jak tomu je u BGA pouzder. U těchto pouzder jsou vývody realizovány malými kuličkami pájecí slitiny, které se při přetavení roztaví, smočí pájecí plošku a změní tvar z kuličky na tzv. „soudkovitý tvar“. Typickými defekty v těchto vývodech jsou různé deformace, dutiny (voids), prasklinky (cracks), otevřené spoje (open joints), nebo zkraty (shorting) tzv. můstky (bridging). Nejčastější defekty jsou zachyceny na obrázku 1.



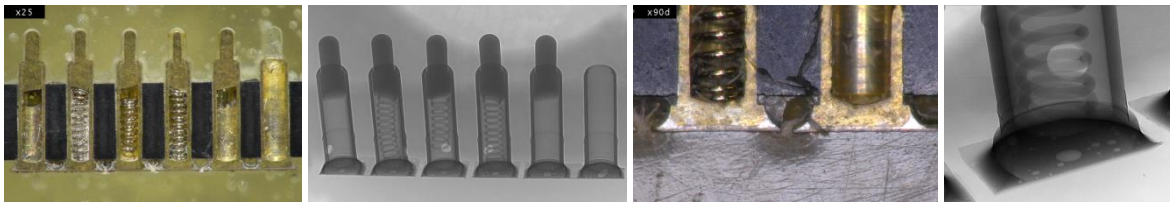
Obrázek 1: Ukázka defektů odhalených pomoci rentgenu NORDSON DAGE XD7600NT (dutina, prasklinka, otevřené spoje, zkraty/můstky)

2. ANALÝZA DEFEKTŮ

Umístěním vývodů na spodní stranu pouzdra se zkomplikovala kontrola kvality pájených spojů. První metodou, kterou jsme schopni kontrolovat pájené vývody BGA pouzder je ERSASCOPE. Jedná se o endoskop se speciální sondou, která snímá obraz pod úhlem 90°. Sondu umístíme mezi DPS a pouzdro, tímto způsobem lze zkontrolovat pouze první řadu kuliček. Nevýhodou je, že nelze odhalit chybějící kuličku v dalších řadách nebo defekty uvnitř jednotlivých vývodů.

Vlivem nedostatků první jmenované nedestruktivní metody se k analýze defektů začalo využívat rentgenové záření, které dokáže odhalit i vnitřní defekty pájených spojů. Rentgenová defektoskopie je založena na schopnosti RTG záření pronikat (prostupovat) zkoumaným předmětem. S průchodem záření přes předmět dochází k pohlcení části tohoto záření podle druhu materiálu, z kterého je předmět vyroben a jeho tloušťky. Materiály s vyšším protonovým číslem více pohlcují rentgenové záření. A k většímu útlumu záření dojde i při průchodu přes silnější vrstvy materiálu. Prošlé paprsky dopadají na detektor, díky kterému získáváme výsledný obraz zkoumaného předmětu. Obraz má světlejší a tmavší místa, které odpovídají množství rentgenového záření, které v daném místě prošlo předmětem. Z průběhu těchto změn absorbovaného záření určíme makroskopické defekty, jakými jsou neviditelné trhliny a dutiny, které bychom jiným způsobem nezjistili.

Část vnitřních defektů je obtížně zjistitelná i za použití rentgenu, proto je další používanou metodou mikrovýbrus. Tato metoda je časově náročnější a je destruktivní, což je její hlavní nevýhodou. Mikrovýbrus je však v některých případech nenahraditelný. Analyzovaná část DPS se musí vyříznout, zalít do fixační hmoty a po vytvrdnutí hmoty se takto připravený vzorek brousí a leští do doby, než se dostaneme do požadovaného místa, kde se nachází zkoumaný defekt. Příkladem je zhotovený mikrovýbrus „pogo“ konektoru na obrázku 2, kdy na RTG snímku není vidět prasklina.



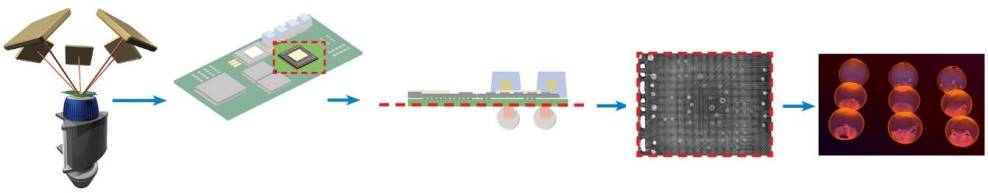
Obrázek 2: Mikrovýbrus a RTG snímek „pogo“ konektoru s prasklinou spoje

3. RENTGEN NORDSON DAGE

Společnost Sanmina má k dispozici rentgen od firmy NORDSON DAGE s modelovým označením XD7600NT. Uvnitř zařízení je radiační komora, ve které jsou tři nejdůležitější části rentgenu, kterými jsou zdroj RTG záření, pracovní stůl a detektor RTG záření. Statický zdroj rentgenového záření je umístěn ze spodní strany pohyblivého pracovního stolu rentgenu, na který se umisťují analyzované vzorky. Nad stolem je pohyblivý panelový CMOS detektor, který snímá množství prošlého záření a umožňuje přenášet obraz v reálném čase a ve vysokém rozlišení přímo k obsluze rentgenu. Celé zařízení je odstíněno a zkonstruováno tak, aby rentgenové záření neohrožovalo obsluhu ani nikoho jiného v okolí. Je využito mnoho ochranných prvků např. bezpečnostní dvířka apod.

3.1. METODA X-PLANE

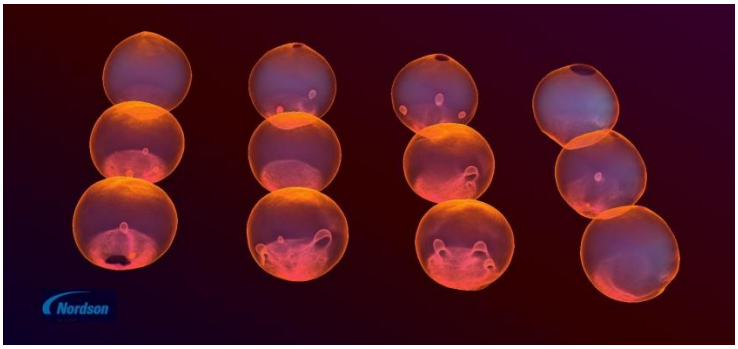
X-plane je metoda snímání určená pro sestavení 2D nebo 3D modelu vnitřních struktur součástek a pájených spojů BGA. Výsledným 2D výstupem je soubor ve formátu JPEG, nebo při použití pokročilého softwaru CERA i 3D model ve video formátu AVI. Vytvoření modelu je možné díky složení více snímků jednoho vzorku. Tato metoda vytvoří řezy diagnostikovaného vzorku. Jejich počet se dá nastavit od 36 do 720 snímků pod úhly 30° až 60°. Ze sestavených modelů lze určit polohu a velikost dutin a kontrolu jednotlivých vrstev DPS, pájeného spoje nebo čipu. Princip funkce metody X-plane je naznačen na obrázku 3. Snímač rotuje okolo vzorku v rozsahu celých 360° pod nastaveným úhlem a pořizuje množství snímků.



Obrázek 3: Princip metody X-plane

3.2. 3D REKONSTRUKCE DEFEKTŮ POMOCÍ SOFTWARU CERA

Software CERA vyvinula společnost DAGE ve spolupráci se společností SIEMENS a slouží k 3D rekonstrukci dat získaných metodou X-plane, která byla popsána v předešlém odstavci. Touto metodou se získávají raw data, která se používají jako vstupní informace pro software CERA. Ten umožňuje řadu základních funkcí nezbytnou pro vytvoření 3D modelu jako je např.: příprava raw dat, korekce špatných pixelů, korekce intenzity záření, další různé filtry, online rekonstrukce, vizualizace průběhu rekonstrukce a další pokročilé funkce. Pro názornou představu výsledný výstup softwaru CERA po zpracování dat získaných rentgenem XD7600NT je na obrázku 4.



Obrázek 4: 3D model vývodu BGA pouzdra s vnitřními defekty

4. ZÁVĚR

Hledání a oprava defektů je časově a finančně mnohdy náročná. Proto je nejlepší způsob defektům předcházet. To lze z velké části ovlivnit při výrobě, když se správně nastaví výrobní proces, budou použity kvalitní komponenty, materiály a velká pozornost bude věnována pájecímu procesu, který ovlivňuje přímo výslednou kvalitu pájeného spoje. Ovšem jak bylo zmíněno, defekty vznikají i během života výrobku, proto se jich nelze zcela vyvarovat a jejich diagnostika a oprava je zapotřebí. Vhodnou metodou je použití rentgenu, protože lze kontrolovat všechny řady vývodů, jejich tvar a hlavně vnitřní strukturu, v které se často vyskytují dutiny ovlivňující spolehlivost pájeného spoje. Rentgen se řadí mezi nedestruktivní diagnostiku, takže nedochází k poškození analyzovaného výrobku a jeho velkou výhodou je, že s jeho pomocí lze kontrolovat jak kovové tak i nekovové materiály a předměty. V případech, kdy základní RTG kontrola neodhalila defekt spoje byl zhotoven mikrovýbrus, který potvrdil, že některé vnitřní defekty rentgen nezobrazuje. Další práce se proto zaměří na použití 3D rekonstrukce RTG snímků, aby se dal vyvodit závěr jestli 3D modely defektů budou schopny nahradit destruktivní metodu zhotovování mikrovýbrusů.

ACKNOWLEDGEMENT

Tento příspěvek vznikl díky podpoře společnosti Sanmina-SCI Czech Republic s.r.o., která mi umožnila využívat nejmodernější zařízení této firmy. Především přístup k rentgenu XD7600NT napomohl ke zpracování mé diplomové práce a k vytvoření tohoto příspěvku.

REFERENCE

[1] SZENDIUCH, Ivan. Základy technologie mikroelektronických obvodů a systémů. Vyd. 1. Brno: VUTIUM, 2006, 379 s. ISBN 80-214-3292-6.

[2] AGE PRECISION INDUSTRIES LIMITED. NORDSON XD7600NT: uživatelská příručka. 2004.

[3] BERNARD, D. X-RAY inspection criteria and common defect analysis. 2006.

[4] ČEZ a.s, Rentgenové záření [online]. [cit. 2015-12-02] Dostupné z: <http://www.cez.cz/edee/content/microsites/rtg/histor.htm>

BOOTLOADER FOR SCI-TRACE

Lukáš Novák

Master Degree Programme (2.), FEEC BUT

E-mail: xnovak0b@vutbr.cz

Supervised by: Pavel Šteffan

E-mail: steffan@feec.vutbr.cz

Abstract: This paper analyze a design of bootloader for microcontrollers AVR and made proposal for implementation to the device Sci-Trace. Create design circuit for patch board and control board and implementation into the rack system.

Keywords: Avr, Atmega644P, Atmel, Bootloader, Rack, Schroff, Fuse bits, RS485, FTDI.

1. ÚVOD

Sci-Trace je zařízení sloužící pro analýzu prvků pro Středoevropský technologický institutu VUT v Brně (viz **Chyba! Nenalezen zdroj odkazů.**). Hlavní součástí zařízení jsou interakční komora, pulzní laser, spektrální analyzátor, vzduchotechnika a řídicí elektronika. U samotné analýzy prvků pomocí pulzního laseru je vypařena část vzorku v interakční komoře a spektrálním analyzátorem se zjišťuje samotné složení vzorku při různých tlacích, případně v různých plynech.

Již zmíněná elektronika slouží k řízení motorů, ovládání osvětlení, vzduchotechniky či závěrky pro kameru. Jednotlivé moduly musejí být připojeny k základní řídicí desce, která zajišťuje spojení s počítačem.



Obrázek 1: Fotografie zařízení Sci-Trace.

2. KONCEPT BOOTLOADERU

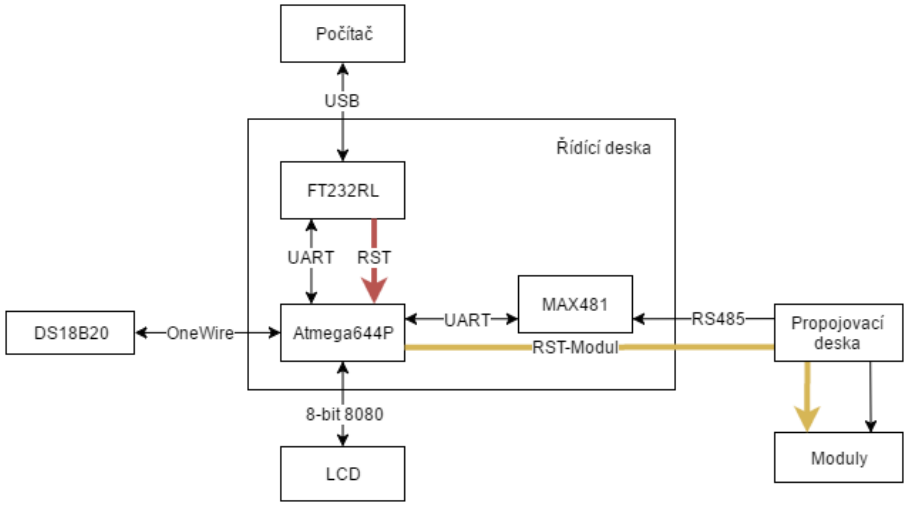
Koncept bootloader, zvaný též zavaděč, lze nalézt v různých mikrokontrolérech, případně složitějších čipech. Mikrokontroléry od firmy Atmel mají rozdělenou Flash paměť na dvě části: Aplikační část a bootloader část, kterou lze aktivovat pomocí tzv. programovacích pojistek (fuse bits). V továrním nastavení jsou pojistky pro zavádění bootloaderu neaktivní a tato část se chová jako klasická aplikační část. Při aktivaci příslušných pojistek nastává při restartu změna v programové adrese, kde nejprve dojde k načtení bootloaderu a až poté startuje aplikační část od nulové adresy. Adresa bootloaderu je pro každý typ procesoru specifická a to kvůli rozdílným velikostem programové paměti.

Mohlo by se zdát, že tento způsob je zcela zbytečný a lze část kódu bootloaderu pro mikrokontrolér implementovat na začátek aplikační části. Takovéto tvrzení lze považovat za správné, ale pouze pokud chce implementovat nějaký jednodušší kód, například blikání informačních LED diod při restartu. Úplně jiný případ nastává v mnohem složitějším kontextu, kde přes různé příkazy tato malá část kódu, dokáže přes komunikační sběrnici přijímat data a následně přepisovat těmito daty aplikační část ve Flash paměti, což je i pravý důvod proč byl koncept bootloaderu vynalezen. Lze tedy tvrdit, že přes poměrně malou část kódu dokáže mikrokontrolér přeprogramovat sám sebe. Pro aktualizaci firmware tedy nepotřebujeme jiné externí zařízení (programátor).

3. KONSTRUKČNÍ USPOŘÁDÁNÍ

Celé zařízení je konstrukčně navrženo do rackového systému od firmy Schroff [1], který je standardizován dle normy ČSN 60297-3-100. S ohledem na normu je vytvořena propojovací deska pro euro-karty a zajištěna rozšiřitelnost pro libovolný počet propojovacích desek. Samotná propojovací deska slouží k rozvedení napájecích a datových pinů mezi všemi moduly. Na zadní straně propojovací desky jsou zvlášť vyvedeny piny pro připojení řídicí desky, která je hlavní částí celé elektroniky.

Samotné moduly lze vkládat do propojovací desky v libovolném pořadí což činí zařízení univerzální. Nevýhoda této možnosti je v nutnosti umístění konektorů na přední panel, jehož šířka se tím zvětší. To způsobuje snížení celkového počtu modulů zapojitelných do racku.



Obrázek 2: Blokové schéma konstrukce.

4. ŘÍDICÍ DESKA

Řídicí deska (**Chyba! Nenalezen zdroj odkazů.**) využívá 8 bitový mikrokontrolér od firmy Atmel se dvěma periferiemi UART. Jedna periferie UART je připojena k FTDI čipu (FT232RL [3]), který komunikuje s počítačem. Druhá periferie UART je připojena přes převodník MAX481 [4], který jej propojí se sběrnicí RS485. Přes obě rozhraní se posílají jak funkční, tak programová data. Kromě základního zapojení převodníku FTDI je ještě vyveden jeden datový vodič z převodníku na vstup RESET mikrokontroléru, jehož funkcí je aktivace bootloader části. Zmíněný pin je zvýrazněn na blokovém schématu (obrázek 2) červeně.

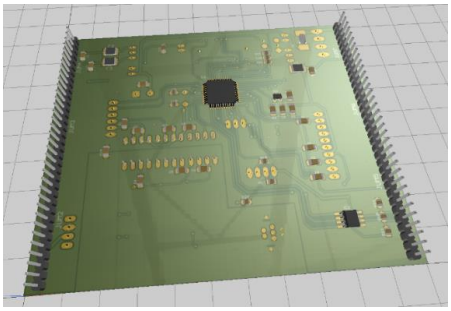
Sběrnice RS485 je dále rozvedena na propojovací desce ke všem modulům viz model na **Chyba! Nenalezen zdroj odkazů.**. Příkaz lze poslat přes počítač a řídicí elektroniku až k modulům.

Na blokovém schématu (obrázek 2), si lze povšimnout žlutě zvýrazněného datového pinu, který vede z pinu od mikrokontroléru z řídicí elektroniky, přes propojovací desku až k jednotlivým modulům. U modulů je pin připojen na reset mikrokontroléru a lze na dálku přeprogramovat nejenom samotný řídicí modul, ale i ostatní moduly, připojené přes sběrnici RS485. Protože datový paket o

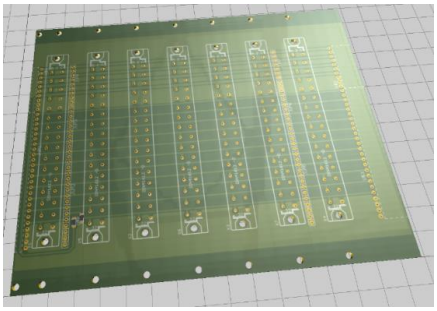
délce 9 bajtů v sobě zahrnuje adresu cílového modulu, lze všechny mikrokontroléry resetovat pomocí jednoho ovládacího pinu.

Na přední části racku je z řídicí desky vyveden informační displej pro sledování teploty uvnitř racku, lze tak kontrolovat přehřátí samotného zařízení. Dále se na displeji zobrazují aktuálně připojené moduly, jejich adresy, identifikační typy, kontrola napájecích větví, stav připojení k počítači a stav komunikační sběrnice RS485. Samotný displej komunikuje přes 8 bitový protokol 8080 [5].

Na řídicí desce je vyveden i sběrnice OneWire pro připojení teplotních čidel společnosti Dallas, pro již výše zmíněné měření vnitřní teploty racku.



Obrázek 3: Model řídicí desky.



Obrázek 4: Model propojovací desky.

5. ZÁVĚR

Článek popisuje návrh elektroniky pro zařízení Sci-Trace. Nejprve je popsáno samotné zařízení Sci-Trace. V dalším kroku je popsán systém funkce bootloaderu, kterým lze přeprogramovat elektroniku bez externích zařízení. Dále je popsán řídicí modul připojený k propojovacímu modulu, který propojuje datový tok mezi počítačem a moduly. Hlavní částí celého přístroje je řídicí modul, který zajišťuje ovládání uživatelsky výměnných modulů. Displej připojený k řídicí desce zobrazuje stav elektroniky, teplotu uvnitř racku a napájecí napětí. Pro představu je zde také uvedeno zjednodušené blokové schéma a modely propojovací a řídicí desky.

6. PODĚKOVÁNÍ

Tento příspěvek vznikl za podpory Ministerstva školství, mládeže a tělovýchovy České republiky pod záštitou projektu CEITEC 2020 (LQ1601) a projektem IGA „Nové typy elektronických obvodů a senzorů pro specifické aplikace, FEKT-S-14-2300“. Dále bych poděkoval konstrukčnímu a vědeckému týmu Laboratoře laserové spektroskopie FSI VUT a svému vedoucímu za podporu a cenné rady.

REFERENCE

[1] A. Comporation, "AVR109: Self Programming," Atmel, 2004. [Online]., [cit. 2016-03-30] Dostupné z: http://www.atmel.com/images/doc1644.pdf.

[2] ČSN, Mechanické konstrukce pro elektronická zařízení - Rozměry mechanických konstrukcí řady 482,6 mm (19 palců) - Část 3-100: Základní rozměry čelních panelů, skříní, stojanů, zásuvných jednotek a koster, 1.9.2009.

[3] FTDI, *Future Technology Device International*. [Online]., [cit. 2016-03-30] Dostupné z: http://www.ftdichip.com/Support/Documents/DataSheets/ICs/DS_FT232R.pdf.

[4] MAX481 Low-Power. *Maxim Integrated*. [Online]., [cit. 2016-03-30], Dostupné z: https://www.maximintegrated.com/en/products/interface/transceivers/MAX481.html.

[5] T6963c. *Buydisplay*. [Online]., [cit. 2016-03-30] Dostupné z: http://www.buydisplay.com/download/ic/T6963.pdf.

CALIBRATION OF THE FBG MEASURING SYSTEM BY LASER INTERFEROMETER

Michal Jelínek

Master (2) FEEC BUT

E-mail: xjelin36@stud.feec.vutbr.cz

Supervised by: Břetislav Mikel

E-mail: mikel@feec.vutbr.cz

Abstract: This work deals with calibration method of the FBG based sensors and the nuclear power plant containment shape monitoring system. The design and the implementation of this system with the optical filter thermal compensation is described. The measurement of length changes is compared with the Michelson interferometer. The control system was programmed in LabVIEW.

Keywords: Bragg grating, measuring of optical spectrum, fiber sensors, Michelson interferometer

1. ÚVOD

Měření optickými vlákny s Braggovými mřížkami je v posledních letech stále více využíváno. V aplikacích se liší metody měření, vyhodnocení a konstrukce senzorů. Využití těchto optických senzorů v oblasti jaderného průmyslu je dosud minimální. Souvisí to s mnoha speciálními požadavky na konstrukci senzorů i na vyhodnocovací jednotky.

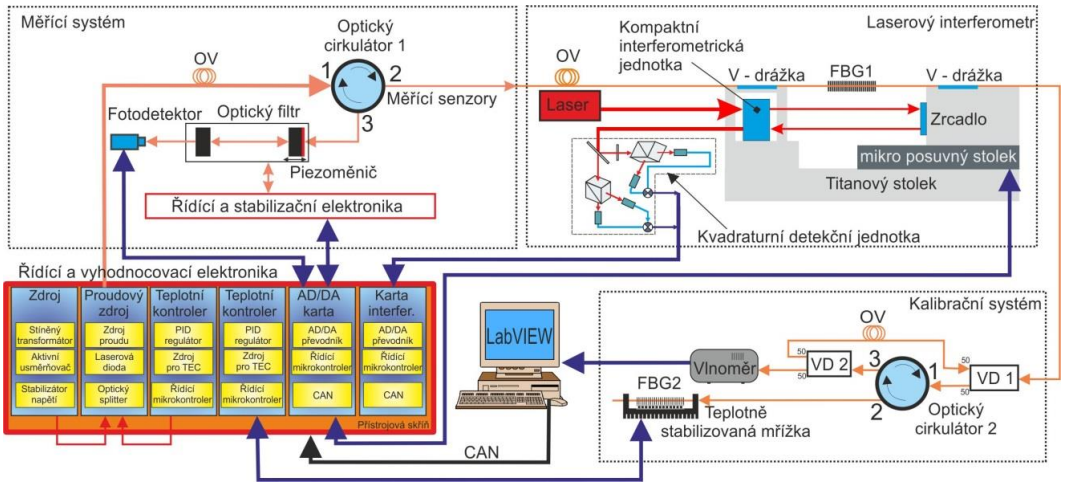
Navrhli jsme a připravili systém, který je v současnosti nainstalován na první srovnávací měření v jaderné elektrárně Temelín, kde měří tvarové změny kontejnmentu. Využité principy měření a laboratorní verze tohoto systému je zde prezentována. Pro návrh, kontrolu a výrobu optovláknových senzorů pro JE Temelín byl navržen a otestován systém na kalibraci linearity měření a ověření vláknových Braggových mřížek.

2. PRINCIP MĚŘENÍ

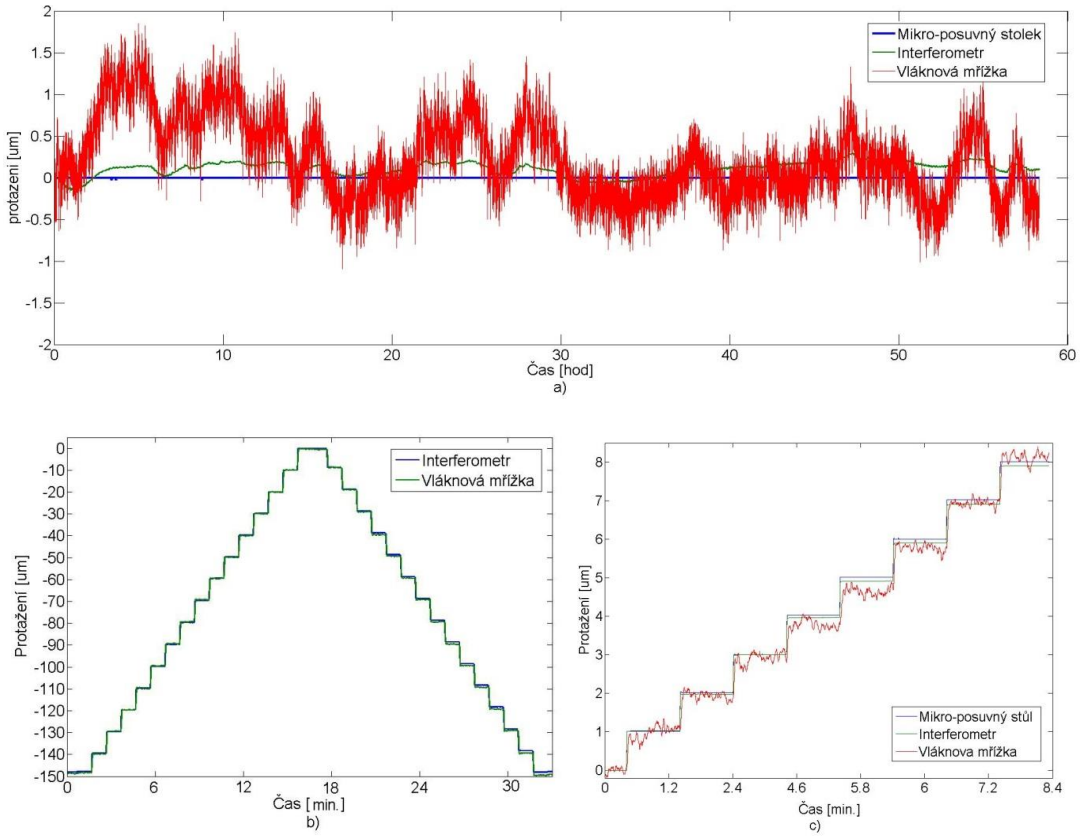
Princip zapojení celého systému pro měření délkových změn je vidět na obr. 1. Světlo ze širokospektrální LED diody CS5203A o optickém výkonu 20 mW je přivedeno na vláknovou Braggovu mřížku. Vlnová délka světla odpovídající periodě Braggovy mřížky se odrazí zpět do optického cirkulátoru a je nasměrována do optického spektrálního analyzátoru OSA [1]. Přesnost měření frekvence odraženého světla je velmi závislá mj. na změnách teploty optického filtru. Pro nastavení měřícího systému byla proto vytvořena kalibrační sestava, která v laboratorním provedení obsahuje i referenční měření prostřednictvím vlnoměru [2]. Pro kalibraci systému i vláknových mřížek byl celý systém rozšířen o laserový interferometr, který měří protažení senzorů ve speciální titanové kalibrační lavici. Na základě těchto rozšíření lze celou experimentální sestavu rozdělit do několika funkčních bloků: měřící systém, laserový interferometr, kalibrační systém a řídící elektronika. Celá experimentální sestava byla kompletně řízena z PC programem realizovaným v LabVIEW.

2.1. KALIBRACE

Přesnost měření s optickým filtrem je ovlivněna teplotními změnami filtru, nelinearitami elektroniky a hysterezí piezoměniče. Nejvíce přesnost měření ovlivňují teplotní změny elektroniky filtru a zejména pak teplotní nestability samotného filtru/rezonátoru. Kalibrační část proto slouží zejména k odstranění těchto teplotních nestabilit.



Obrázek 1: Schéma zapojení senzorového systému; VD - vláknový dělič, OV - optické vlákno; FBG - vláknová Braggova mřížka.



Obrázek: 2: Záznamy z laboratorního měření a) Záznam z 60 hodinového měření stability optického filtru; b) Časový graf zobrazující krokové délkové změny senzoru po 10 µm za 10 s; c) Detail záznamu měření krokových změn senzoru po 1 µm za dobu 10 s.

Spektrum světla procházející vláknovou Braggovou mřížkou FBG1 se využívá pro kalibraci a je proto přivedeno na druhou Braggovu mřížku FBG2, která je teplotně stabilizovaná. Její teplota je řízena teplotním kontrolorem s přesností ± 1mK. Odražená vlnová délka je detekována v optickém filtru. Aby bylo možné detekovat odražené spektrum obou mřížek v optickém filtru současně, je nutné jejich Braggovy vlnové délky vzájemně posunout tak, aby se nepřekrývaly v celém plánovaném rozsahu měření. Měřené spektrum z teplotně stabilizované mřížky je pak ovlivňováno pouze

nestabilitami optického filtru. Vzájemným srovnáním referenčního spektra FBG2 a měřeného spektra FBG1 je pak možné tyto nestability eliminovat a měřit přesněji.

2.2. LASEROVÝ INTERFEROMETR

Pro ověření linearity a přesnosti měření vláknových mřížek i navrhovaného měřícího systému byl navržen a realizován systém s laserovým interferometrem. Zapojení měřícího systému s interferometrem je na obrázku 1.

Měřící délková mřížka je přilepena k titanovému stolku. Na něm, na pevné části, je umístěna kompaktní interferometrická jednotka Michelsonova interferometru, která má implementováno referenční zrcadlo a další optické prvky. Na druhém, posuvném konci titanového stolku je měřící zrcadlo představující měřící rameno interferometru. Mezi touto posuvnou a pevnou částí je vlepeno optické vlákno s Braggovou mřížkou. Zdrojem světla pro Michelsonův interferometr je frekvenčně stabilizovaný Helium-Neonový laser s vlnovou délkou λ_l = 632,8 nm a optickým výkonem 5 mW. Při změně polohy pohyblivé části s měřícím zrcadlem se mění vzdálenost, kterou projde paprsek světla. Výsledná interference paprsku z referenčního a měřícího ramene je detekována kvadraturní detekční jednotkou. Změnou polohy pohyblivé části je natahováno nebo zkracováno i měřící vlákno s Braggovou mřížkou. Měřená změna délky je tak zaznamenána posunem spektra odrazu od Braggovy mřížky a současně detekční jednotkou interferometru. Celý měřící proces je řízen programem vytvořeným v LabVIEW, kde je zpracován signál z interferometru a ze spektrometru o protažení optického vlákna s Braggovou mřížkou. Na obrázku 2 jsou data z několika testovacích měření, kdy byla ověřována přesnost měření s vláknovou Braggovou mřížkou. Dlouhodobé měření stability optického filtru při implementované kalibrační metodě eliminace teplotních vlivů je vidět na obrázku 2a. V další fázi testů byla měřena odezva změny detekované délky senzoru na její reálné změně, která byla monitorována laserovým interferometrem. Délka byla měněna v rozsahu $\Delta p = < -150; 0 >$ s krokem k = 10 μm, viz obr 2b. Na obr 2c je pak detail záznamu měření při změně délky s krokem k = 1 um, kde je zobrazena i změna délky mikro-posuvného stolku. Směrodatná odchylka měření vláknovou Braggovou mřížkou je σ^2 = 0.438 μm pro 60 hodinový záznam měření.

3. ZÁVĚR

V článku je prezentován optický měřící systém pro měření tvarových změn kontejnmentu jaderné elektrárny Temelín prostřednictvím vláknových Braggových mřížek. Základní systém měření délkové roztažnosti byl rozšířen o část na odstranění teplotních nestabilit optického filtru a o část kalibrační. Pro průběžnou kalibraci systému je využívána teplotně stabilizovaná Braggova mřížka, která při každém měření reprezentuje měřící etalon. Přesnost systému byla ověřována s využitím Michelsonova interferometru, který měřil souběžně s vláknovou Braggovou mřížkou délkové změny senzoru. Směrodatná odchylka měření vláknovou Braggovou mřížkou při dlouhodobém – 60 hodinovém měření stability byla σ^2 = 0,438 μm.

PODĚKOVÁNÍ

Tento příspěvek vznikl za podpory institucionálního financování RVO:68081731, projektu MŠMT LO1212, projektu EK CZ.1.05/2.1.00/01.0017 a grantového projektu TAČR č. TA03010835.

REFERENCE

[1] JELÍNEK, Michal. *CALIBRATION SYSTEM OF FIBER BRAGG GRATINGS MEASUREMENT*. 1. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2015. ISBN 978-80-214-5148-3.

[2] JELÍNEK, Michal. *Návrh optovláknového senzorového systému pro měření změn délky.* Brno, 2014. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Vedoucí práce Ing. Břetislav Mikel, Ph.D.

DRIVER UTILIZING HDMI INTERFACE FOR MODULAR LED DISPLAYS

Tomáš Bartek
Master Degree Programme (2), FEEC BUT
E-mail: xbarte11@stud.feec.vutbr.cz

Supervised by: Vojtěch Dvořák
E-mail: xdvora99@stud.feec.vutbr.cz

Abstract: This paper presents VHDL implementation of complex system for controlling modular LED displays. In this project we mainly focus on utilizing HDMI input interface on FPGA which controls LED displays, but also on ensuring communication via MODBUS protocol, identification through DDC and thermal security of LED displays itself.

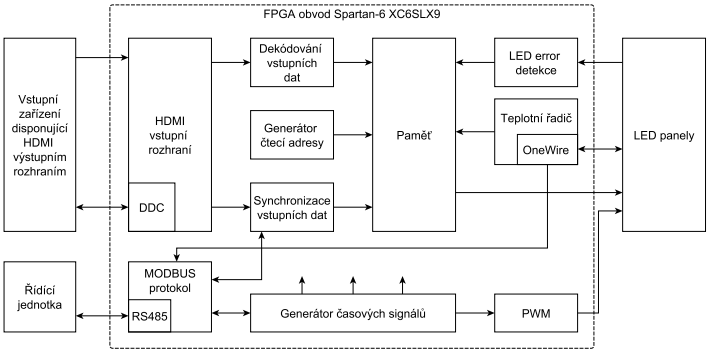
Keywords: Information panels, HDMI, LED, VHDL, FPGA

1 ÚVOD

Tento článek se věnuje oblasti informačních LED panelů. Konkrétně představuje návrh řadiče s FPGA obvodem, který zpracovává HDMI signál a na základě příchozího obrazu generuje řídicí signály pro LED moduly. Pro realizaci řadiče byl vybrán FPGA obvod od firmy XILINX, Inc. Spartan-6 XC6SLX9. Při realizaci bylo nutné implementovat i pokročilé funkce pro řízení průmyslových LED panelů. Jedná se o měření teploty a teplotní řízení, komunikaci s nadřazenou jednotkou a detekci vadných LED na panelu. HDMI rozhraní bylo vybráno kvůli své rozšířenosti a tedy univerzálnosti z pohledu zdroje video signálu.

2 BLOKOVÉ SCHÉMA

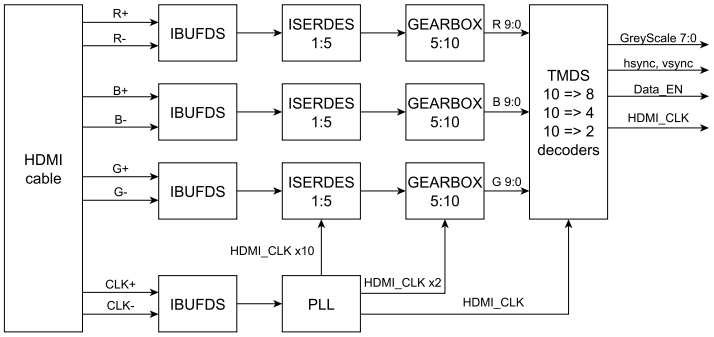
Zjednodušené blokové schéma je naznačeno na obrázku č. 1. Ze schématu vyplývá, že se jedná o komplexní návrh, který obsahuje kromě samotného HDMI rozhraní i možnost detekce nefunkčních LED, teplotní řadič a MODBUS protokol pro komunikaci FPGA s řídicí jednotkou. Použité LED moduly jsou monochromatické a jedná se o prototypy určené k zavedení na trh. Disponují vstupním rozhraním založeném na protokolu HUB08, který není standardizován a neexistuje k němu literatura. Detekce vadných LED je podporována osazenými budiči SCT2027 [1].



Obrázek 1: Zjednodušené blokové schéma implementace

3 HDMI ROZHRANÍ

HDMI rozhraní je založeno na velmi rychlé serializaci nejen obrazových dat. V FPGA obvodu je proto nutné zabezpečit zachycení těchto dat a jejich následné zpracování. Realizace příjmu dat v obvodu je znázorněna na obrázku 2.



Obrázek 2: Implementace vstupní HDMI komponenty

Použitý obvod Spartan-6 XC6SLX9 disponuje vstupními diferenciálními páry IBUFDS [6], které jsou použity pro příjem LVDS signálů. Další zpracování se liší podle toho, zda-li se jedná o obrazový nebo hodinový kanál. Signál z hodinového kanálu je předán do bloku fázového závěsu PLL [7], který generuje vstupní hodinový signál HDMI dat. Pro zpracování vstupních signálů je nutné generovat jednonásobek, dvojnásobek a desetinásobek vstupního hodinového signálu.
V případě obrazových kanálů je signál předáván do dalšího nativního bloku obvodu FPGA ISERDES2 [6]. Ten zabezpečuje zachycení velmi rychlého sériového datového toku. V obvodech Spartan-6 jsou přítomny ISERDES2 bloky s poměrem zachycení dat 1:4. Pro účely FPGA je ovšem nutné zachytávat vstupní data v poměru 1:5. V návrhu je toto zabezpečeno pomocí dvou ISERDES2 bloků pracujících v režimu Master a Slave.
Jakmile jsou data v 5bitových vektorech, je nutné je spojit do 10bitových TMDS slov, které je poté možné dekódovat na samotná obrazová data a synchronizační signály [2]. Spojení obstarává komponenta GearBox, která uchovává tři po sobě následující 5bitové vektory. V těchto 15bitových vektorech vyhledává speciální 10bitová synchronizační slova [2]. Jakmile tento vektor nalezne, je schopna generovat na svém výstupu platná 10bitová TMDS slova.
Posledním krokem zpracování vstupních dat je převod TMDS vektorů na platná obrazová data. Toho je dosaženo pomocí algoritmu popsaného v HDMI specifikacích [2]. Výstupem z HDMI komponenty však nejsou RGB signály, ale pouze jeden 8bitový kanál se škálou šedé. Převod je zabezpečen pouhým sečtením jednotlivých barevných složek obrazového bodu a vydělení číslem 4. Dělení je v FPGA obvodu realizováno odebráním dvou LSB. V FPGA obvodu je poté možné nastavit od jaké hodnoty šedé je daný obrazový bod rozsvícen.
Celá vstupní HDMI komponenta v FPGA obvodu využívá 190 LUT, 182 Slice Registers, 7 LUTRAM, 1 PLL a 3 BUFG bloky.

4 IDENTIFIKACE ZAŘÍZENÍ

Pro zabezpečení kompatibility navrhovaného systému s libovolným zařízením je nutné, aby FPGA obvod byl schopen se identifikovat pomocí DDC rozhraní [2]. V rámci tohoto komunikačního kanálu se FPGA obvod identifikuje jako zobrazovací zařízení, které je schopno přijímat obraz o libovolném rozlišení, maximálně však 1920x1080 bodů. Samotná komunikace je založena na standardu I^2C a data jsou přenášena ve formě bloku 128 16bitových slov. Přesná struktura identifikačních dat je definována pod pojmem EDID [3].

5 TEPLOTNÍ ŘÍZENÍ

V oblasti průmyslového použití informačních panelů je nutné počítat s velikou teplotní zátěží LED modulů. Každý jednotlivý modul obsahuje teplotní senzor DS18B20, se kterým je možno komunikovat pomocí protokolu 1-Wire[4]. FPGA obvod tedy detekuje všechny teplotní senzory a pravidelně v rozestupech 5 vteřin měří teplotu na LED modulech. Pokud některý dosáhne kritické teploty (to může být způsobeno například díky lokálnímu osvětlení sluncem nebo kvůli poruše ventilačního systému v LED panelu), vyhodnotí to FPGA jako chybovou situaci a vypne zobrazování na příslušném panelu. Díky tomuto opatření je možné zabránit nenávratnému poškození panelu. FPGA obvod následně ohlásí chybový stav nadřazené řídicí jednotce a dále kontroluje teplotu na vypnutém panelu. Pokud teplota poklesne o definovanou hodnotu, je panel opět zapnut.

6 KOMUNIKACE

Komunikace FPGA obvodu s nadřazenou jednotkou je zajištěna pomocí MODBUS protokolu [5], který využívá jako fyzickou vrstvu rozhraní RS485. Nadřazená jednotka může nastavovat celou řadu parametrů, například jas, výchozí pozici zobrazování, práh zobrazení pixelu a teplotní limity. Dále je schopna číst stavové informace o LED panelu. Jsou to například chybová hlášení, počet LED modulů, teploty modulů, počet LED.

7 ZÁVĚR

Funkčnost řadiče s FPGA obvodem je ověřena na testovacím LED panelu s rozlišení 160x32 obrazových bodů. Při této konfiguraci je v FPGA obvodu Spartan-6 XC6SLX9 využito 37% LUTs, 52% Slices a 25% BRAM16BWERs a maximální pracovní kmitočet 86 MHz (řadič pracuje na kmitočtu 50 MHz). Maximální kmitočet pro zpracování HDMI signálu je 164 MHz. Proudový odběr v této konfiguraci je 290 mA při 5 V. Teoretické maximální rozlišení, které je schopen jeden řadič řídit, je 313 tisíc pixelů, což odpovídá například rozlišení 640x480 px.

REFERENCE

[1] StarChips Technology. 16-bit Constant-Current LED Driver with Modeless Error Detection [online]. 03-06-2013 [cit. 2016-03-25]. Dostupné z: http://www.starchips.com.tw/pdf/datasheet/SCT2027V01_03.pdf

[2] HDMI Licensing, LLC. HDMI Specification Version 1.3a [online]. 10-11-2006 [cit. 2016-03-15]. Dostupné z: http://www.microprocessor.org/HDMISpecification13a.pdf

[3] Video Electronics Standards Association. Vesa Enhanced Extended Display Identification Data Standard, Release A, Rev. 2 [online]. 25-11-2006 [cit. 2016-03-15]. Dostupné z: ftp://ftp.cis.nctu.edu.tw/pub/csie/Software/X11/private/VeSaSpEcS/VESA_Document_Center_-Monitor_Interface/EEDIDrAr2.pdf

[4] DALLAS Semiconductor. Programmable Resolution 1-Wire Digital Thermometer [online]. [cit. 2016-03-24]. Dostupné z: https://datasheets.maximintegrated.com/en/ds/DS18B20.pdf

[5] Modbus-IDA. Modbus Application Protocol Specification v1.1b [online]. 28-12-2006 [cit. 2016-03-24]. Dostupné z: http://www.modbus.org/docs/Modbus_Application_Protocol_V1_1b.pdf

[6] XILINX, Inc. Spartan-6 FPGA SelectIO Resources [online]. 21-11-2015 [cit. 2016-03-25]. Dostupné z: http://www.xilinx.com/support/documentation/user_guides/ug381.pdf

[7] XILINX, Inc. Spartan-6 FPGA Clocking Resources [online]. 19-06-2015 [cit. 2016-03-25]. Dostupné z: http://www.xilinx.com/support/documentation/user_guides/ug382.pdf

EFFECT OF COMPACTION PRESSURE TO THE ELECTRO-CHEMICAL PROPERTIES OF THE ELECTRODES FOR LI-S ACCUMULATORS

Kamil Jaššo

Master Degree Programme (2), FEEC BUT

E-mail: xjasso00@stud.feec.vutbr.cz

Supervised by: Tomáš Kazda

E-mail: kazda@feec.vutbr.cz

Abstract: Over the last decade, demand for batteries increased a lot thanks to the wearable devices and electric cars. The current trend in all industries is reducing the size and weight of batteries and increasing their capacity. The demand for even more powerful batteries is often faced with theoretical and practical limits of the technology. Battery industry compared to other industries is just slightly behind and batteries became a limiting factor for technologies.The purpose of this experiment is to describe the impact of compaction pressure on the electrochemical properties of the positive electrode for lithium-sulfur batteries.

Keywords: Battery, Accumulator, Lithium-ion, Li-ion, Compaction pressure, Li-S, Lithium-sulfur batteries, Cyclic voltammetry, CV

1. ÚVOD

Od svojho vzniku v roku 1991 prešli lítium-iónové akumulátory veľkým vývojom. Na trhu sú v súčasnosti dostupné Li-ion batérie v širokej škále tvarov, veľkosti, kapacít a zložení. Rôzne druhy Li-ion akumulátorov sa líšia použitými materiálmi a majú rôzne vlastnosti. V súčasnosti prebieha mnoho výskumov s cieľom zlepšiť vlastnosti Li-ion akumulátorov. Tieto výskumy sú zamerané na všetky zložky týchto akumulátorov. Medzi najperspektívnejšie sa radia najmä výskumy aktívnych materiálov kladných elektród a to konkrétne síry a kyslíku. Výsledkom týchto výskumov sú batérie (Li-S a Li-O) s podstatne vyššou energetickou hustotou než akú dosahujú konvenčné Li-ion akumulátory. Ich vysoká energetická hustota ich robí príťažlivými najmä pre oblasti elektro-automobilového priemyslu. Dôležitou vlastnosťou elektród je porozita a vnútorná súdržnosť materiálov. Pri použití vysokého lisovacieho tlaku môže dôjsť k uzavretiu štruktúry a k zabráneniu priechodu elektrolytu. Naopak pri nelisovanej elektróde môže (najmä pri Li-S) dôjsť k rozpadu štruktúry.

2. LÍTIUM-SÍROVÝ AKUMULÁTOR (LI-S)

Na rozdiel od Li-ion akumulátorov lítium-sírový akumulátor nepracuje na princípe interkalačného procesu, ale je konverzným typom batérie a na jeho nabíjaní a vybíjaní sa podieľajú elektrochemické reakcie. [1][4]

Reakcia prebiehajúca na zápornej elektróde:

$$2Li \leftrightarrow 2Li^{+} + 2e^{-} \tag{1}$$

Reakcia prebiehajúca na kladnej elektróde:

$$\mathrm{S} + 2e^{-} \leftrightarrow \mathrm{S}^{2-} \tag{2}$$

Celková reakcia (→ vybíjanie, ← nabíjanie):

$$2Li + \mathrm{S} \leftrightarrow \mathrm{Li}_2S \tag{3}$$

Síra je vďaka svojej vysokej teoretickej kapacite (1675 mAh/g), energetickej hustote (2600 Wh/kg), veľkej dostupnosti a nízkej cene perspektívnym nástupcom súčasných konvenčných aktívnych elektródových materiálov používaných pre kladné elektródy Li-ion článkov. Nevýhodou síry je jej nízky potenciál voči Li/Li^+ a slabá elektrická vodivosť. V prírode sa síra vyskytuje v molekule obsahujúcej 8 atómov síry, čo znamená že príhodnejší zápis rovnice celkovej reakcie je:

Celková reakcia (→ vybíjanie, ← nabíjanie):

$$16\ Li + S_8 \leftrightarrow 8\ Li_2S \tag{4}$$

Li_2S sa ale pri vybíjaní formuje postupným prechodom z S_8 cez polysulfidy Li_2S_x ($1 \leq x \leq 8$) až k Li_2S. Niektoré z týchto medziproduktov (Li_2S_8 a Li_2S_4) sú rozpustné v používanom elektrolyte a usadzujú sa na povrchu lítiovej anódy, čím vo výsledku dochádza k strate aktívneho materiálu katódy, k pasivácii anódy a k rýchlemu poklesu kapacity behom cyklovania. Aby sa tomuto problému zabránilo, je nutné uzavrieť síru do elektricky vodivej vrstvy schopnej prepustiť ióny lítia. Perspektívnou metódou, ako sa tomuto problému vyhnúť, je uzavretie síry do pórov amorfného uhlíku, uhlíkových nanovlákien, či iných uhlíkových produktov alebo do vodivých polymérnych puzdier. Výhodou tejto modifikácie je tiež zvýšenie vodivosti aktívneho materiálu katódovej elektródy. Ďalším predmetom výskumov s cieľom zabrániť rozpúšťaniu polysulfidov je modifikácia použitých elektrolytov. [2][3][4]**Chyba! Nenalezen zdroj odkazů.**

Najväčšou nevýhodou lítium-sírových akumulátorov je najmä veľká objemová zmena pri nabíjaní a vybíjaní článku, čo spôsobuje na katóde mechanický stres a zhoršuje elektrický kontakt medzi uhlíkom a sírou. Tento objemový nárast môže spôsobiť rozpad aktívnej hmoty katódy. [2]

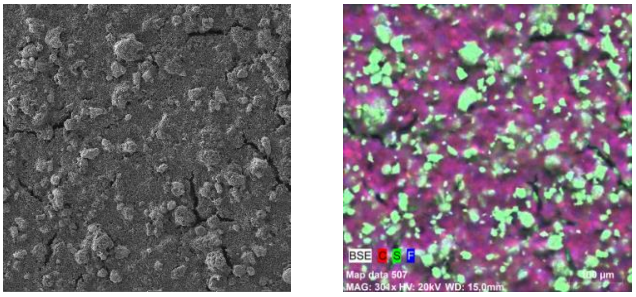
3. EXPERIMENT

Pre experiment bola vytvorená kladná elektróda s 200 µm vrstvou aktívneho materiálu v zložení: 60% Síra (Sigma-Aldrich ≥99,5%), 30 % uhlík (Super P) a 10 % spojivo (PVDF*).

Nelisovaná elektróda bola vložená do štandardnej meracej cely El-Cell©. Ako referenčná elektróda bolo použité lítium. Do meracej cely bol kvapnutý elektrolyt v zložení 0,25 mol/l $LiNO_3$, 0,75 mol/l LiTFSI (DME:DOL** v pomere 2:1).

(*PVDF - Polyvinyliden fluorid, **DME - 1,2-Dimethoxyethane, **DOL - 1,3-dioxolane).

3.1. ANALÝZA KLADNEJ ELEKTRÓDY



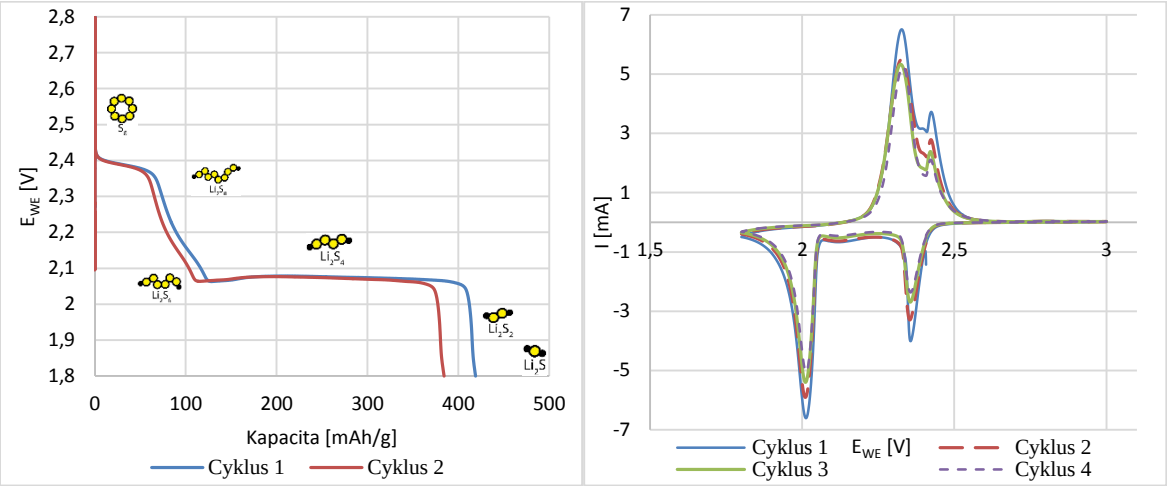
Obrázok 1: A) Zobrazenie pod SEM (Zorné pole: 415 µm) B) Rozloženie prvkov (C, S, F)

Pred vložením do meracej cely bola elektróda skúmaná pod SEM mikroskopom a bola spravená prvková analýza. Zobrazenie nelisovanej elektródy pod mikroskopom ukázalo nerovnosť povrchu a tvorbu mikropuklín, ktoré vznikli pri manipulácii s elektródou a pri sušení. Na prvkovej analýze bolo možné vidieť nerovnomerné rozloženie použitých chemických prvkov (zhluky síry).

3.2. VYBÍJANIE ČLÁNKU A CYKLICKÁ VOLTAMETRIA

Prebehli dva cykly. Nabíjací/vybíjací prúd bol nastavený na 0,2C. Potenciálové okno bolo nastavené v rozmedzí 1,8 – 2,8 V. V grafe je možné si všimnúť dve relatívne stabilné platá pri

2,38 V a 2,07 V, ktoré sú spôsobené tvorbou polysulfidov. Tieto vybíjacie charakteristiky, podobne ako aj priebeh cyklickej voltametrie zodpovedajú priebehom uvedeným v odbornej literatúre.



Obrázok 2: A) Vybíjacie charakteristiky cyklovania (0,2C) B) Cyklická voltametria (0,1 mV/s)

Pri CV bolo potenciálové okno nastavené v rozmedzí 1,8 – 3,0 V. Ako je z grafu 2. možné vidieť, pri skenovacej rýchlosti 0,1 mV/s je priebeh cyklickej voltametrie pri štyroch cykloch relatívne stabilný, ale s jasne viditeľnými kapacitnými stratami.

4. ZÁVER

Napriek tomu, že experiment nie je momentálne ešte ukončený, je možné z doposiaľ získaných dát povedať, že nelisovaná sírová elektróda počas cyklovaní výrazne stráca kapacitu vplyvom rozpúšťania polysulfidov. Vzhľadom k viditeľne malej súdržnosti a pevnosti elektródy sa očakáva jej rozpad pri ďalšom cyklovaní.

POĎAKOVANIE

Ďakujem Ing. Tomášovi Kazdovi, Ph.D. za účinnú metodickú, pedagogickú a odbornú pomoc a Ústavu elektrotechnologie UETE, FEKT VUT za poskytnuté priestory a vybavenie.

REFERENCIE

[1] YOO, Hyun Deog, Elena MARKEVICH, Gregory SALITRA, Daniel SHARON a Doron AURBACH. On the challenge of developing advanced technologies for electrochemical energy storage and conversion. Materials Today. 2014, 17(3): 110-121 .ISSN 1369-7021

[2] NITTA, Naoki, Feixiang WU, Jung Tae LEE a Gleb YUSHIN. Li-ion battery materials: present and future. Materials Today [online]. Elsevier Ltd, 0150n. l., 18(5): 252-264 [cit. 2015-12-05]. ISSN 1369-7021

[3] BRODD, R. Batteries for sustainability: selected entries from the Encyclopedia of sustainability science and technology. vi, 513 pages.

[4] FEDORKOVÁ, Andrea, et al. New Composite Cathode Materials for Li/S Batteries: A Review. Int. J. Electrochem. Sci, 2013, 8: 10308-10319.

[5] ČECH, Ondřej, Andrea FEDORKOVÁ a Marie SEDLAŘÍKOVÁ. Progresivní elektrodové materiály pro lithno-iontové akumulátory: Kladná elektroda na bázi elementární síry. Electroscope. Západočeská univerzita v Plzni. Fakulta elektrotechnická, 2013, 2013(4), 4. ISSN 1802-4564.

EFFECT OF PLASMA OXIDATION OF CARBON WORKING ELECTRODE ON ITS ELECTROCHEMICAL RESPONSE

Lukáš Koporec

Master Degree Programme (2), FEEC BUT

E-mail: xkopor00@stud.feec.vutbr.cz

Supervised by: Petra Majzlíková

E-mail: businova@feec.vutbr.cz

Abstract: Screen-printed electrochemical sensors are very perspective because they are much smaller than common electrodes used in electrochemistry, portable, ecological and can be very easily modified in order to increase their sensitivity and selectivity for specific analytes. In this study, plasma treatment was used to modify the carbon working electrode of a screen-printed three-electrodes system. Plasma treatment was performed using RF plasma discharge in argon/oxygen atmosphere. The plasma modified electrodes exhibited considerable enhancement in electrocatalytic activity towards studied electroactive compounds.

Keywords: Screen-printing, carbon electrode, plasma treatment, cyclic voltammetry, electrochemical sensor

1. ÚVOD

Hrubovrstvové tlačené elektrochemické senzory predstavujú perspektívu pre elektroanalytickú chémiu. Ich nespornou výhodou sú rozmery a tým možnosť ich ľahko prenášať a používať v podstate skoro kdekoľvek. V porovnaní s bežnými elektródami, vykazujú vyššiu vodivosť a nízky prúd pozadia kvôli ich poréznej štruktúre, ktorá vytvára veľkú plochu povrchu elektród.[1] Ďalšie výhody oproti klasickým elektródam sú výrobný proces, nízka cena a ekologickosť. Pri hrubovrstvových senzoroch môžeme namiešať pasty s rôznymi vlastnosťami podľa toho na aký účel bude senzor použitý. Tieto senzory sa potom po výrobe tiež môžu ešte dodatočne modifikovať. [2] Jednou z možností modifikácie je vystavenie povrchu elektród pôsobeniu plazmy. Plazma je kvázineutrálny plyn, ktorý obsahuje rovnaký počet kladne aj záporne nabitých častíc. Vystavením senzora pôsobeniu kyslíkovej plazmy dôjde k oxidácii povrchu elektródy a k zlepšeniu ich elektrochemických vlastností.

2. EXPERIMENTÁLNA ČASŤ

2.1. ÚPRAVA PRACOVNEJ ELEKTRÓDY V PLAZME

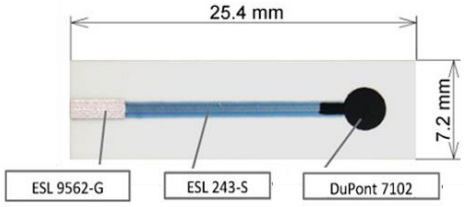
V tejto práci boli vzorky vystavené pôsobeniu plazmy vytvorenej pomocou argónu a kyslíka. Tieto dva plyny boli zmiešané v pomere 4:1 ($Ar:O_2$). Tohto pomeru bolo dosiahnutého pomocou prietokomerov, ktoré boli nastavené na 20 sccm a 5 sccm. Pracovná komora bola najprv odčerpaná na počiatočný tlak 3 Pa. Následne boli do nej napustené plyny a dosiahnutý tlak bol 92,60 Pa. Výkon plazmy bol nastavený na 25 W a čas jej pôsobenia na vzorky bol 5 minút.

2.2. ELEKTROCHEMICKÁ CHARAKTERIZÁCIA

Elektródy boli charakterizované metódou cyklickej voltampérometrie. Táto technika poskytuje informácie o prenose elektrónov a o umiestnení redoxných potenciálov. Pri použití pracovnej elektródy v nepremiešavanom roztoku sa používa trojuholníkové napätie a meria sa prúdová

odozva na toto napätie. Prúdové maxima a ich potenciál sú podľa Nicholsona a Shaina základnými parametrami cyklického voltampérogramu. [3]

Na charakterizáciu boli použité roztoky 2,5mM ferro-ferrikyanidu draselného v 0,1M chloride draselnom a 2,5mM dopamín v 0,01M kyseline chlorovodíkovej. Zvolené analyty tvoria redoxný pár a preto sú vhodné k charakterizácii elektród. Meranie bolo uskutočnené v trojelektródovom usporiadaní. Pracovná elektróda bola uhlíková, vyrobená pomocou technológie hrubých vrstiev (Obrázok 1) a upravená plazmou. Pasty použité pri výrobe elektródy boli od firmy ESL ElectroScience a DuPont. Pomocná platinová a referenčná argento-chloridová elektróda boli štandardné, dodávané firmou Metrohm.



Obrázok 1: Tlačená hrubovrstvová polymérna uhlíková elektróda, podľa [2].

Pracovná elektróda bola pred meraním aktivovaná v základnom roztoku 0,1M chloridu draselného s polarizačnou rýchlosťou 50 mV/s. Tento cyklus bol zopakovaný 10 krát. Následne bola elektróda premeraná v 2,5 mM ferro-ferrikyanide draselnom s polarizačnými rýchlosťami 500, 400, 300, 200, 100, 50, 25, 10 a 5 mV/s. Každá polarizačná rýchlosť prebehla len raz. Potenciálový rozsah bol nastavený od -1 do +1 V. Rovnaký postup bol použitý aj pri meraní v roztoku dopamínu. Len potenciálový rozsah bol nastavený od -0,5 do +1 V a základným roztokom bola 0,01 M kyselina chlorovodíková. Charakterizované boli neupravené (UT) a takisto aj plazmou upravené (PT) pracovné elektródy.

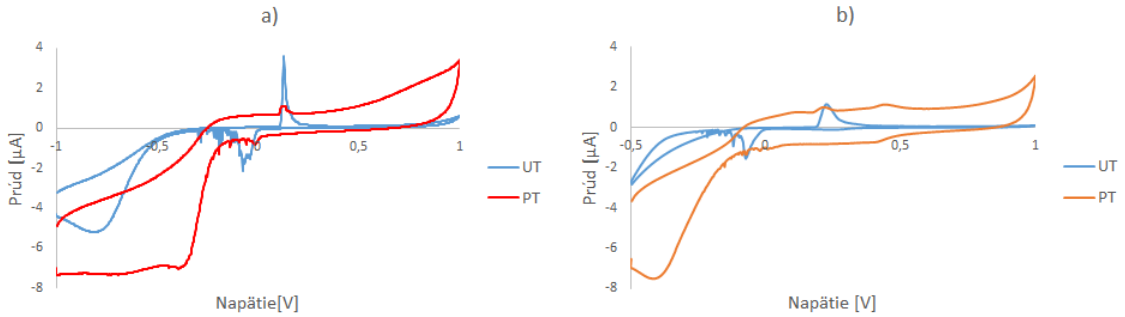
3. VÝSLEDKY A DISKUSIA

Dôležitým prvkom pri skúmaní elektród je sledovanie parametrov charakterizujúcich reverzibilné systémy. Rozdiel potenciálov oxidačného a redukčného píku (ΔE_p) sa u jednoelektrónových reverzibilných systémov rovná hodnote 59 mV a so zmenou rýchlosti polarizácie nedochádza k posunu potenciálov píkov [3]. Ani u jedného typu použitého senzora (UT/PT) nebolo dosiahnutého takéhoto výsledku (viď Tabuľka 1). Stanovené hodnoty ΔE_p boli vyššie a bolo pozorované, že s rastúcou rýchlosťou polarizácie sa ΔE_p zvyšuje. V prípade nemodifikovaných elektród boli hodnoty ΔE_p veľmi vysoké a systém je možné charakterizovať ako ireverzibilný. V prípade modifikovaných elektród došlo k výraznému zníženiu hodnôt ΔE_p a systém je možné charakterizovať ako kvazireverzibilný.

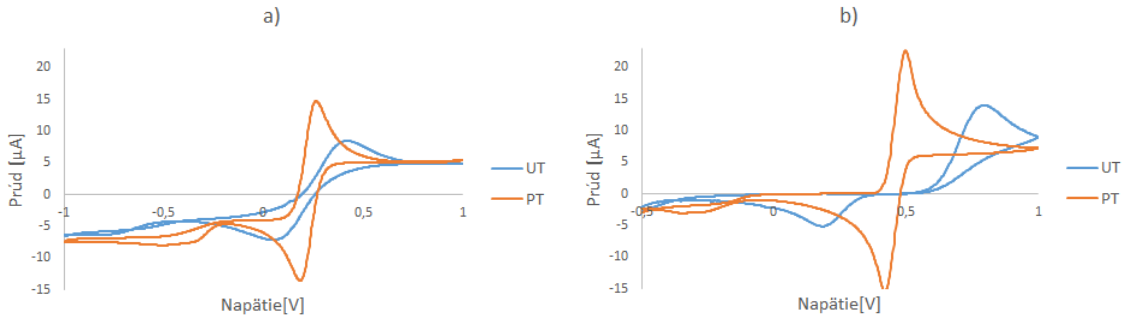
Polarizačná rýchlosť [mV/s]	Ferro-ferrikyanid draselný		Dopamín	
	ΔE_{pUT} [V]	ΔE_{pPT} [V]	ΔE_{pUT} [V]	ΔE_{pPT} [V]
5	0,374	0,081	0,613	0,078
10	0,459	0,081	0,647	0,927
25	0,557	0,088	0,698	0,120
50	0,669	0,095	0,740	0,149
100	0,681	0,107	0,784	0,183
200	0,779	0,127	0,837	0,232
300	0,781	0,140	0,867	0,266
400	0,818	0,151	0,894	0,293
500	0,842	0,159	0,918	0,317

Tabuľka 1: Rozdiel potenciálov oxidačného a redukčného píku

Plazma ďalej zlepšuje odozvu a znižuje negatívny vplyv zložiek pasty, z ktorých je senzor vyrobený (viď Obrázok 2 a **Chyba! Nenalezen zdroj odkazů.**). Pri neupravenej elektróde sa do reakcie zapája pravdepodobne aj pojivová časť pasty, ktorá spôsobuje píky navyše (Obrázok 2 UT). Úpravou elektródy ale dochádza k odstráneniu tohto pojiva z povrchu elektródy a tým sa zväčší elektroaktívny povrch čo vedie k minimalizovaniu týchto píkov (Obrázok 2 PT).



Obrázok 2: Porovnanie odozvy upravenej (PT) a neupravenej (UT) elektródy pre 5. cyklus v základnom roztoku a) 0,1M KCl a b) 0,01M HCl pri 50 mV/s.



Obrázok 3: Porovnanie odozvy upravenej (PT) a neupravenej (UT) elektródy vo a) ferro-ferrikyanide draselonm a b) dopamíne pre 5 mV/s.

4. ZÁVER

Merania v základných roztokoch KCl, HCl a taktiež vo ferro-ferrikyanide draselnom a dopamíne ukázali, že úprava elektród hrubovrstvových senzorov v plazme má pozitívny vplyv na výsledok. Dochádza k zvýšeniu odozvy elektród, zmenšuje sa vplyv ostatných zložiek pasty na prebiehajúcu elektrochemickú reakciu medzi elektródou a analytom a veľmi výrazne sa znižuje rozdiel potenciálov oxidačného a redukčného píku, čo ukazuje zlepšenie elektrokatalytickej aktivity elektród.

REFERENCIE

[1] WANG, S. C. et al. Enhancement of electrochemical properties of screen-printed carbon electrodes by oxygen plasma treatment. *Electrochimica Acta*, 2009, sv. 54, č. 21, s. 4937-4943. ISSN 0013-4686.

[2] MAJZLÍKOVÁ, P. et al. Comparison of different modifications of screen-printed working electrodes of electrochemical sensors using carbon nanotubes and plasma treatment. *physica status solidi (a)*, 2014, sv. 211, č. 12, s. 2756-2764. ISSN 1862-6319.

[3] WANG, J. *Analytical electrochemistry*. New Jersey: John Wiley & Sons, Inc., 2006. 262 s. ISBN 978-0-471-67879-3.

FPGA MODULE FOR FAST WAVEFORM GENERATION AND SYNCHRONOUS DATA ACQUISITION

Josef Eliáš
Master Degree Programme (2), FEEC BUT
E-mail: xelias10@vutbr.cz

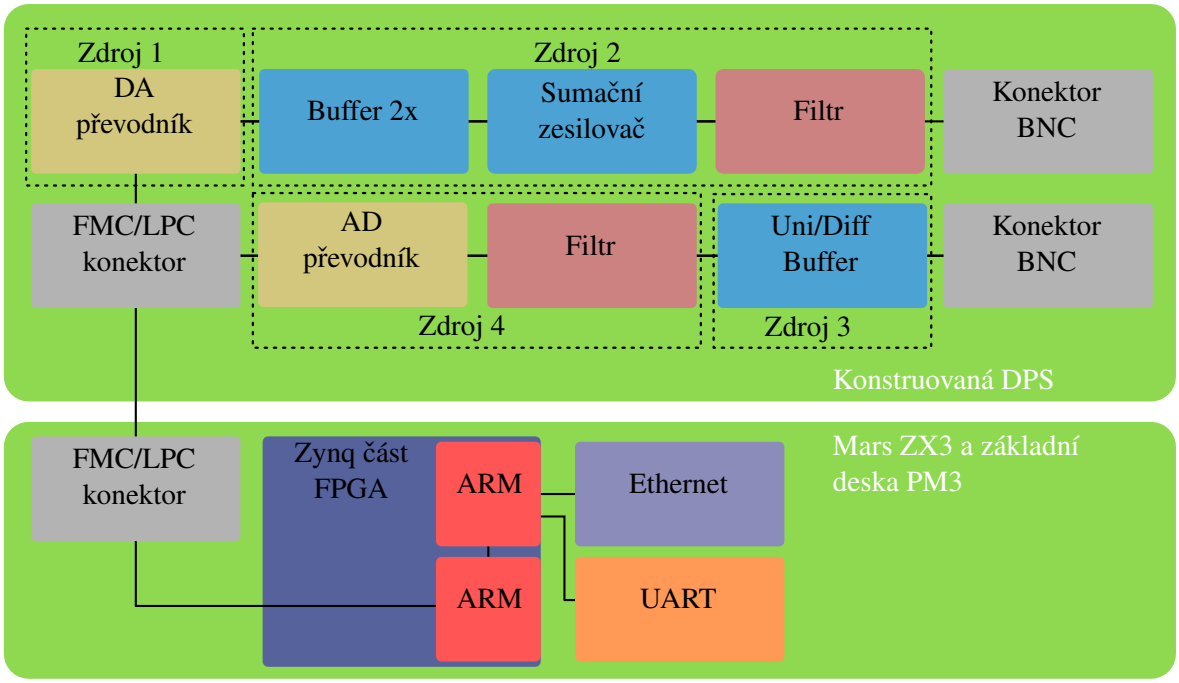
Supervised by: Milan Holík
E-mail: xholik02@vutbr.cz

Abstract: This article deals with system for fast waveform generation and synchronous data acquisition. System is used for evaluation of optical sensors by sweep through narrow band light source. The DAQ system use dual core ARM processor with build in FPGA array. We can achieve direct sharing memory between ARM core and AD/DA periphery through SoC architecture. Then we can optimally distributed core usage between processors and peripherals.

Keywords: waveform, data aquisition, FPGA, SoC, ARM, AXI

1 ÚVOD

V tomto článku bude popsána konstrukce zřízení pro vyhodnocování optických vláknových senzorů založených na principu změny vlnové délky v reakci na měřenou veličinu. Princip měření je založen na rychlém rozmítání úzkopásmového filtru při synchronním vzorkování signálu z fotodetektoru. Obě hlavní komponenty (úzkopásmový filtr a rychlý fotodetektor) jsou součástí firemního know-how, proto bude popsán pouze navrhovaný synchronní systém pro řízení obou komponent. Základem je obvod Zynq 7020, který nabízí dvě ARMová jádra ve struktuře programovatelného hardwaro-



Obrázek 1: Blokové schéma systému pro vyhodnocení optického signálu

vého hradlového pole. Procesor bude využit pro řízení a zpracování, zatímco hradlové pole bude sloužit pro ukládání dat do sdílené paměti. Navrhovaný obvod disponuje dvojicí převodníků,DA převodník pro generování řídicího signálu a AD pro vzorkování fotodetektoru. Oba převodníky pracují na frekvenci 80 MHz, proto byla zhotovena vývojová deska, optimalizovaná pro pracovní frekvence nad 100 MHz. Ilustrační blokové schéma je na obrázku 1. Zelené bloky znázorňují jednotlivé desky plošných spojů, vnitřní bloky jsou klíčové fyzické obvody, které budou popsány v následujících odstavcích.

2 ZPRACOVÁNÍ DAT A ŘÍZENÍ

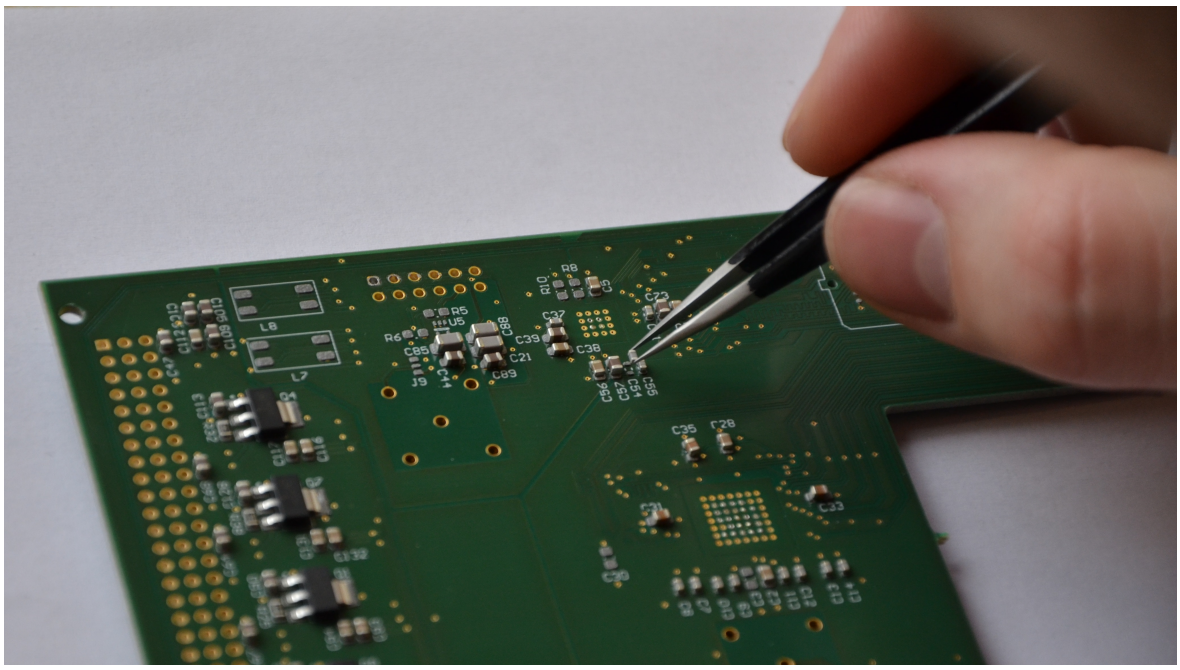
Zpracování dat probíhá na jednom ARMovém procesoru pomocí přístupu do sdílené paměti. Díky periferii, navržené v této práci jsou data nahrávána do sdílené paměti bez asistence procesoru, je možné data zpracovávat a vyhodnocovat měření v reálném čase na frekvenci 10 kHz. Druhé procesorové jádro slouží k provozu webserveru, pomocí kterého bude možné precizně nastavovat tvar generovaného signálu, spouštět nebo zastavovat veškeré komponenty systému a zobrazovat průměrné naměřené hodnoty.

3 GENEROVÁNÍ ŘÍDICÍHO SIGNÁLU

Periferie DA převodníku slouží k napěťovému řízení úzkopásmového optického filtru. Převodník generuje signál na dvou kanálech v rozlišení 14b. První kanál generuje řídicí funkci. Druhý kanál generuje offset pro kompenzaci nesymetrií řízeného filtru. Signál je impedančně přizpůsoben a sečten na trojici přesných nízkošumových precizních zesilovačů a přiveden na BNC konektor. Interně je generován synchronizační pulz, který spouští vzorkování na druhém převodníku.

4 SYNCHRONNÍ VZORKOVÁNÍ FOTODETEKTORU

Druhý převodník vzorkuje signál z dvojice fotodetektorů. Získaná data jsou pomocí speciální periferie navržené pro procesor ukládána do sdílené paměti opět bez pomoci procesoru. Použitý převodník pra-



Obrázek 2: Realizovaná deska pro AD/DA převodníky

cuje na frekvenci 80 MHz s rozlišením 16b a to s diferenciálním signálem. Signál fotodetektoru proto převádíme pomocí diferenciálního operačního zesilovače a přizpůsobujeme jej impedančně na 50 Ω. Aby nedocházelo k převzorkování signálu, jsou na diferenciálních vstupech použity RC články dle doporučení výrobce.

5 SDÍLENÁ PAMĚŤ

Pro generování signálu je třeba vystavovat každých 12,5 ns nový 14b vzorek s periodou 1 ms. Pro vzorkování dat z fotodetektorů je potřeba synchronně ukládat 2 kanály po 16b. Dohromady takový datový tok znamená uložení a přečtení 348 kb paměti každou milisekundu. Vyhovujícím blokem paměti, který je navíc přístupný jak z procesoru, tak z navrhované periferie, je paměť typu DDR3, osazená přímo na vývojovém kitu Enclustra ZX3 [1], [4] . Do té je možné přistupovat pomocí AXI sběrnice z několika míst [2], [3]. Pracovní frekvence 533 MHz je použitím sběrnice snížena na maximálních 150 MHz, přesto je tato frekvence při využití 32b datové sběrnice, dostupné zároveň v obou směrech, plně dostačující. Díky velikosti 512 Mb je možné data ukládat více period a počítat z hodnot exponenciální klouzavý průměr.

6 ZÁVĚR

Účelem konstrukce tohoto systému je nahrazení současného řešení, využívajícího složitý a nákladný průmyslový počítač spolu s řadou přídavných zařízení pro práci s periferiemi v reálném čase. Další nevýhodou současného systému je velikost a softwarová závislost. Navrhovaný systém bude schopen nahradit pracovní stanici a miniaturizovat tak vyhodnocovací jednotku do podoby několika málo Eurokaret a zvýší vyhodnocovací frekvenci systému. To přinese možnost konstrukce nových a propracovanějších vláknových senzorů a posun o řád rychlosti vyhodnocování získaného signálu z takových senzorů až na desítky kiloherz. Tématem pro další vývoj je zavedení linuxového systému na druhé ARMové jádro a usnadnit tak práci se systémem. Na tuto práci bude přímo navazovat výzkum zavádění zpětnovazební smyčky, která linearizuje výstupní signál na normálu. Tímto způsobem bude zjednodušen algoritmus výpočtu hodnot ze senzoru. Do zpracování dat díky zpětné vazbě nebudou zaváděny statistické chyby, což vylepší dlouhodobou stabilitu a přesnost

REFERENCE

[1] GLATTFELDER, CH. *MARS ZX3 SOC MODULE USER MANUAL Datasheet [online].* Enclustra, 2015 [cit. 2016-20-2]. Dostupné z: *http://download.enclustra.com/public_files/SoC_Modules/Mars_ZX3/Mars_ZX3_User_Manual.pdf*.

[2] JOHNSON, J. USING THE AXI DMA IN VIVADO course [online], 2014 [cit. 2016-26-2]. Dostupné z: *http://www.fpgadeveloper.com/2014/08/using-the-axi-dma-in-vivado.html*.

[3] SADRI, M. LESSON 1 : WHAT IS AXI? video [online], 2014 [cit. 2016-25-2]. Dostupné z: *http://www.googoolia.com/wp/2014/03/20/lesson-1-what-is-axi-part-1/*.

[4] ZIEGLER, S. *MARS PM3-R5 USER MANUAL Datasheet [online].* Enclustra, 2015 [cit. 2016-20-2]. Dostupné z: *http://download.enclustra.com/public_files/Base_Boards/Mars_PM3/Mars_PM3_User_Manual_R5.pdf*.

OPTIMIZATION OF SWITCHING REGULATED DC-DC CONVERTER DESIGN

Daniel Appel

Master Degree Programme 5, FEEC BUT

E-mail: xappel00@stud.feec.vutbr.cz

Supervised by: Edita Hejátková

E-mail: hejatka@feec.vutbr.cz

Abstract: The goal of this project was an optimization of existing DC-DC converter design with regards to its efficiency, minimalizing power loses and radiated electromagnetic inference. Design and optimization of converter circuit, PCB, components and also measurement automatization in Simple are discussed in this paper.

Keywords: DC-DC converter, full bridge, zero voltage transition, measurement automatization

1. ÚVOD

Jeden z hlavných trendov vo vývoji elektroniky je zmenšovanie rozmerov zariadení. Použitím spínaných zdrojov je možné zmenšovať veľkosť zariadení, zvyšovať ich účinnosť a znižovať ich cenu. Táto práca pojednáva o dokončení rozpracovaného projektu prototypu DC-DC konvertora firmy Meatest a jeho optimalizácií. V prototype je pre maximalizáciu účinnosti použitá topológia plného mostu s fázovým posunom. Po osadení, oživení a odmeraní pôvodného prototypu budú preskúmané všetky možnosti optimalizácie a bude vyrobená nová DPS konvertora. Optimalizácia DC-DC konvertora je zameraná predovšetkým na zvýšenie jeho účinnosti, zníženia výkonových strát na spínacích prvkoch a zníženie miery vyžarovaného rušenia. Konvertor je plánované po dokončení použiť spolu s modulom regulátora PFC ako spínaný zdroj. Predpokladané parametre konvertora sú výstupné regulovateľné napätie 1-12V, maximálny výstupný prúd 50A, pri vstupnom napätí 400V.

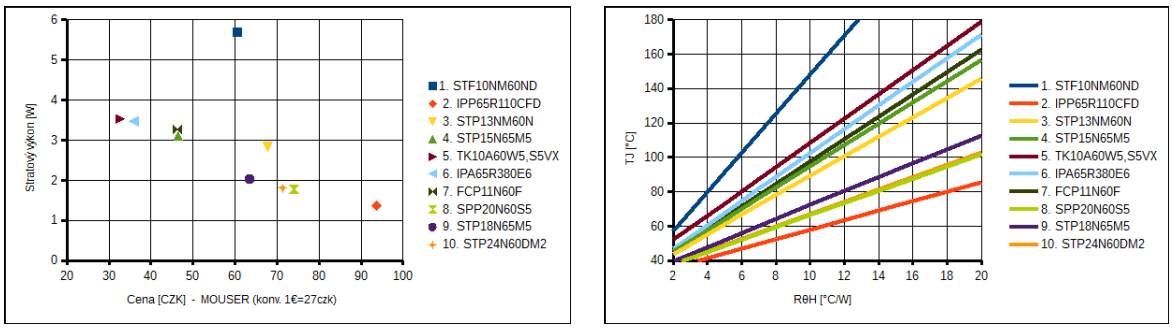
1.1. NÁVRH TRANSFORMÁTORA A CIEVOK

Potrebné parametre transformátora a cievok boli vypočítané podľa doporučení výrobcu riadiaceho obvodu od firmy Texas Instruments [1]. Návrh transformátora bol overený v programe PowerESim, konkrétne v nástroji Magnetic Builder, ktorý slúži na návrh transformátorov pre v práci použitú, a aj iné topológie konvertorov [2]. Prototyp transformátora bol ručne navinutí na teliesku ETD39 s použitím jadra s feromagnetickým materiálom triedy N97, izolačných materiálov firmy 3M a tienením medenou samolepiacou fóliou. Hotový transformátor je možné vidieť na obrázku 4.

1.2. OPTIMALIZÁCIA TRANZISTOROV

V topológií plného mostu (full bridge) s fázovým posunom je vhodné vyberať tranzistory s čo najnižšou hodnotou celkovej výstupnej kapacity C_{OSS}, keďže indukčnosti transformátora a cievky na primárnej strane musia byť dimenzované na uskladnenie dostatočnej energie pre ich nabíjanie a vybíjanie. Taktiež je vhodné vyberať tranzistory s čo najnižšou hodnotou odporu v zapnutom stave R_{DSON}, keďže stratový výkon generovaný na tranzistore v zapnutom stave je priamo úmerný tejto hodnote a nižšou hodnotou celkového náboja elektródy gate Q_G - náboj, ktorý je potrebné dodať na hradlo gate tranzistora pre úplne zapnutie tranzistora [3][4]. Vzhľadom na vyššie uvedené parametre bolo z aktuálnej ponuky vybraných dodávateľov vybraných niekoľko variant

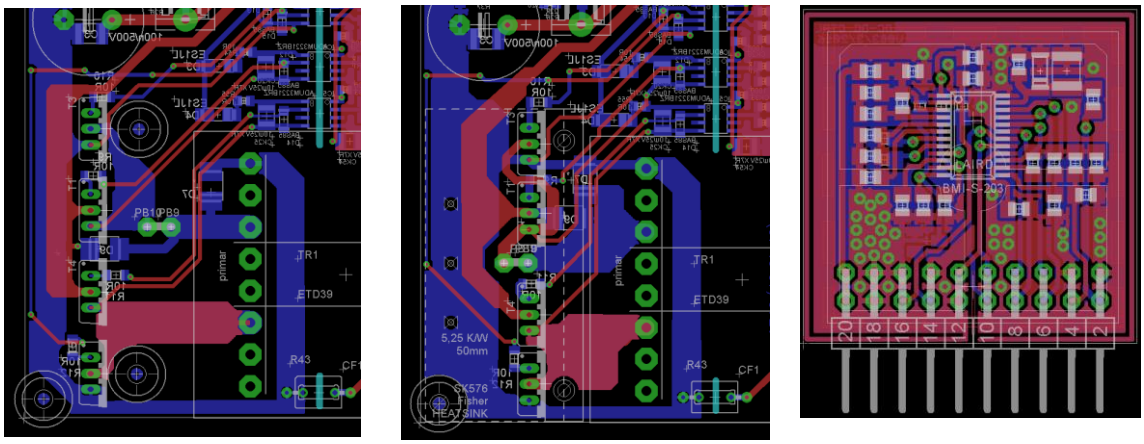
tranzistorov pre konvertor, pre ktoré boli spočítané ich stratové výkony, ktoré boli porovnané s ich cenou. Z výpočtov stratových výkonov boli následne spočítané potrebné tepelné odpory chladičov, teplota chladičov a súčiastok pri vybranom chladiči, výsledky boli vykreslené do grafov . Výsledkom optimalizácie je teoretické zníženie stratového výkonu o 16,6W pri náraste ceny o iba 6kč spolu pre všetky použité tranzistory. Taktiež vďaka nižšej hodnote kapacity C_{OSS} nových tranzistorov je možné použiť na primárnej strane transformátora cievku s menšou hodnotou indukčnosti.



Obrázok 1: Naľavo: porovnanie stratového výkonu a ceny vybraných tranzistorov (č.1 je pôvodný tranzistor). Napravo : Graf závislosti teploty tranzistora na tepelnom odpore chladiča.

1.3. OPTIMALIZÁCIA DPS

Vďaka výberu nových kondenzátorov na výstupe konvertora, bolo možné zmeniť zoskupenie súčiastok na pravej strane DPS (sekundárna strana transformátora). Tým bolo uvoľnené viac miesta v pravej dolnej časti DPS, ktorá bola použitá na umiestnenie chladiča Fisher SK489 s tepelným odporom o hodnote 5,75K/W oproti pôvodným malým individuálnym chladičom, ktoré by nedokázali spoľahlivo chladiť tranzistory pri väčšom prúde na výstupe konvertora. Na ľavej strane DPS bol jednoduchý hliníkový blok chladiča podobne nahradený chladičom Fisher SK576. Keďže v mieste styku s tranzistormi (základňa chladiča) je tento chladič podstatne tenší, tranzistory mohli byť posunuté bližšie k transformátoru, čím boli skrátené všetky cesty k nim a od nich vedúce na DPS ako je možné vidieť na obrázku 2. Dcérska DPS pre riadiaci obvod bola vytvorená v programe Eagle nanovo podľa doporučení [1] pre optimalizáciu najmä z hľadiska rušenia a EMC.



Obrázok 2: Úpravy DPS na primárnej strane transformátora, v strede upravená verzia. Napravo nanovo vytvorená DPS pre riadiaci obvod.

1.4. MERANIE A AUTOMATIZÁCIA MERANIA PARAMETROV KONVERTORA

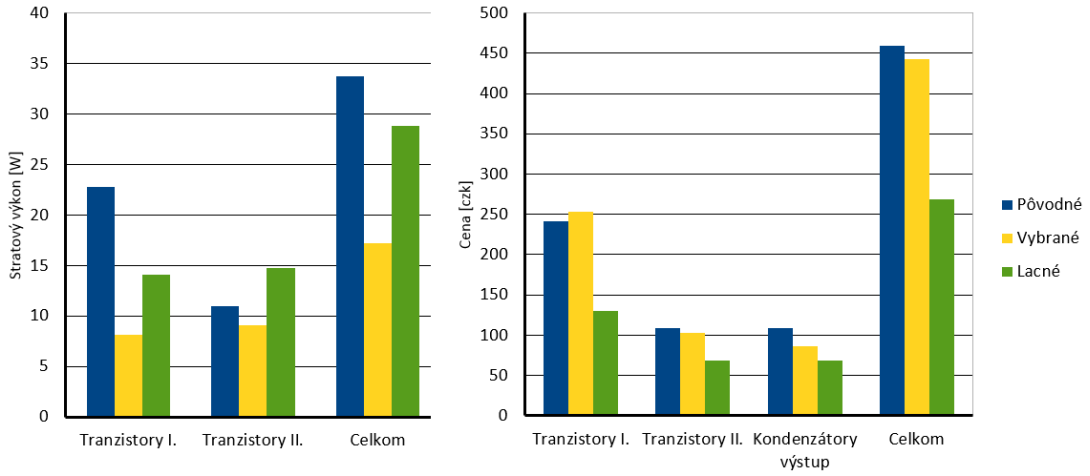
Po oživení konvertora pomocou programovateľného zdroja napätia na vstupe Chroma 61501, programovateľnej záťaže Chroma 63113 a osciloskopu bude vykonané merania a to najmä závislosť účinnosti na vstupnom napätí a výstupnom prúde, zvlnenie výstupného napätia, atď.

V programe Simple firmy Meatest bol naprogramovaný meriaci protokol, ktorý umožní automatizované meranie parametrov konvertora diaľkovým ovládaním meriacich prístrojov pomocou zbernice GPIB čo značne urýchli priebeh merania.

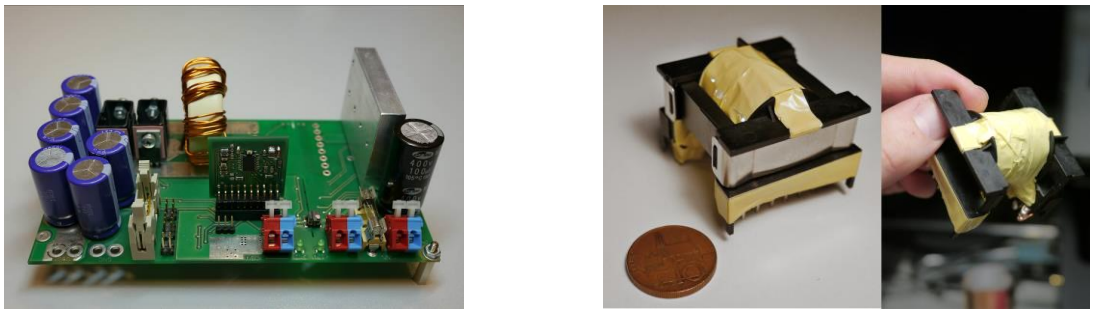
2. ZÁVER

Na základe výpočtov pri všetkých optimalizačných opatreniach boli stanovené teoretické hodnoty zníženia stratového výkonu prípadne zníženia ceny. Výsledky je možné vidieť v grafoch na obrázku 3, z ktorých vyplýva že použitím ideálnych prvkov je možné znížiť stratový výkon o 16,6W, prípadne použitím lacných prvkov je možné ušetriť 191kč pri miernom znížení stratového výkonu.

V rámci riešenia práce je momentálne finalizované osadenie konvertora a príprava pre oživenie a meranie jeho parametrov, na základe čoho budú preskúmané ďalšie možnosti optimalizácie, a neskôr vyrobená nová DPS zahŕňajúca všetky popísané optimalizácie konvertora. Aktuálny stav prototypu konvertora je možné vidieť na obrázku 4.



Obrázok 3: Výsledky optimalizácie.



Obrázok 4: Naľavo dokončený pôvodný prototyp DC-DC konvertora, napravo transformátor.

PODĚKOVÁNÍ

Ďakujem vedúcej diplomovej práce Ing. Edite Hejátkovej a Ing. Michalovi Sittovi z firmy Meatest za účinnú metodickú, pedagogickú a odbornú pomoc a cenné rady pri spracovaní práce.

REFERENCE

[1] UCC28950 Phase-Shifted Full-Bridge Controller With Synchronous Rectification. [online]. [cit. 2016-3-15]. Dostupné z: http://www.ti.com/lit/ds/symlink/ucc28950.pdf

[2] PowerESim [online]. [cit. 2016-03-15]. Dostupné z: http://www.poweresim.com/

[3] ŠTÁL, P., Výkonové audio zesilovače pracující ve třídě D základní principy a konstrukce zesilovače, Praha: BEN – technická literatura, 2008. 200s. ISBN 978-80 7300-230-5.

[4] Zero-Voltage Switching Full-Bridge Converter: Operation, Guidelines for MOSFET Selection [online]. [cit. 2016-03-26]. Dostupné z: http://www.vishay.com/docs/90936/an847.pdf

REACTIVE MAGNETRON SPUTTERING OF THE ALUMINIUM OXIDE COATING

Jan Dušek

Master Degree Programme (5), FEEC BUT

E-mail: xdusek19@stud.feec.vutbr.cz

Supervised by: Tomáš Fořt

E-mail: fortt@isibrno.cz

Abstract: This contribution deals with design, check and test of deposition parameters for new installed aluminium target. In addition of individual parts of deposition process and hysteresis of sputtering thin films of Al_2O_3 was verified. This made it possible to determine borders of reactive and metallic mode. The designed borders were created technological window for sputtering Al_2O_3 films. The sputtered thin films are tested on mechanical, optical properties and adhesion.

Keywords: Reactive magnetron sputtering, pulse sputtering, aluminium oxide coating

1. ÚVOD

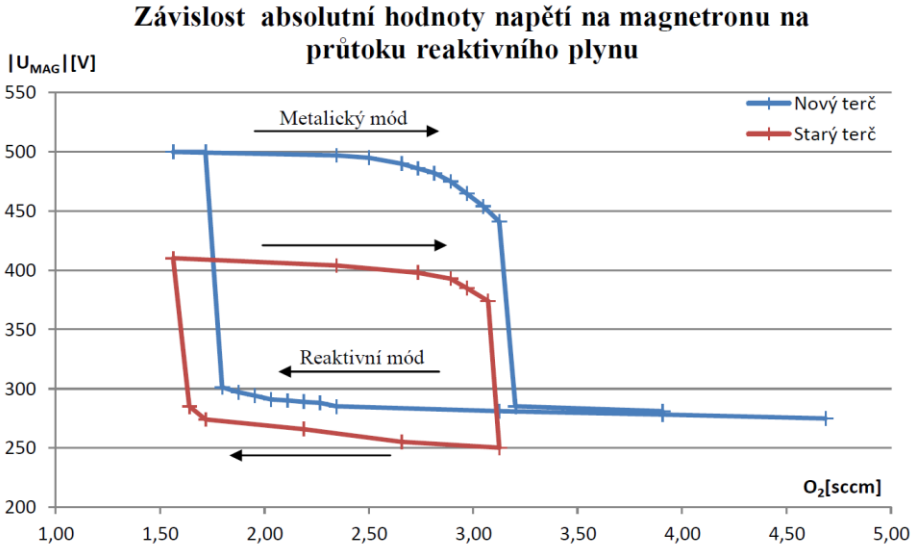
Reaktivní magnetronové naprašování je fyzikální a chemický proces, při kterém je do komory aparatury připouštěn pracovní a reaktivní plyn. Tento reaktivní plyn reaguje s materiálem terče a vytvoří na jeho povrchu sloučeninu požadovaného typu. Tato sloučenina je odprašována z povrchu terče na stolek se vzorky. Díky tomu je možné vytvářet unikátní vrstvy, které se v přírodě nevyskytují. Návrh, ověření a realizace naprašování vrstev Al_2O_3 byla provedena pomocí aparatury PLS160. Vrstva Al_2O_3 byla deponována z nového hliníkového terče za připouštění kyslíku. Dále byla zjištěna odpovídající oblast na hysterezní křivce, která odpovídá reaktivnímu módu. Kromě stechiometrických vrstev Al_2O_3 byl proveden i test pro depozici s nízkým průtokem kyslíku. Díky tomu byly vytvořeny vrstvy hliníku a nestechiometrického Al_2O_3. Vrstvy vytvořené v reaktivním módu byly podrobeny testům pro optické, mechanické vlastnosti a adhezi.

2. PULZNÍ REAKTIVNÍ MAGNETRONOVÉ NAPRAŠOVÁNÍ

V procesu magnetronového naprašování je nutné vyčerpat komoru aparatury na úroveň vakua. Poté je připouštěn pomocí průtokoměru pracovní plyn (argon). Tlak v komoře se udržuje od 0,1 Pa do 2 Pa. Pomocí tlakové špičky a záporného napětí na magnetronu je zapálen doutnavý výboj. Díky tomuto výboji začne docházet k ionizaci pracovního plynu. Kladné ionty jsou přitahovány k magnetronu, na kterém je připevněn terč. Elektrony jsou ovlivněny Lorentzovou silou tak, že pohybem po spirálách zvyšují pravděpodobnost ionizace. Po nárazu ionizované částice plynu do terče se uvolní atomy materiálu z povrchu terče. Vyražené atomy se deponují na stolku s vzorky a vnitřním povrchu aparatury. Pulzní reaktivní magnetronové naprašování rozšiřuje magnetronové naprašování. Hlavním rozdílem v procesu pulzního naprašování je připouštění reaktivního plynu (například: kyslík). Během depozice dochází k chemické reakci plynu na povrchu terče a vytvoření sloučenin, jako jsou například oxidy. Ty jsou odprašovány z povrchu terče. Pulzní naprašování se tedy používá zejména k vytváření nevodivých vrstev. Během doby jednoho pulzu na terči vzniká dielektrická vrstva, která je rozprašována. Poté dojde k otočení polarity pulzu a vybití náboje zachyceného v dielektrické vrstvě. Doba deponování vrstvy se tedy nastavuje pomocí střídy.

3. OVĚŘENÍ A STANOVENÍ HYSTEREZNÍ ZÁVISLOSTI

Pro realizaci vrstev Al_2O_3 byla navržena struktura procesu skládající se z čištění podložek, čištění terče, předprašování a naprašování. Jednotlivé kroky a parametry depozičního procesu byly ověřeny a zaznamenány. Pro čištění podložek byl použit vysokofrekvenční generátor Cesar, který byl připojen na stolek s výkonem 20 W a průtoku argonu 30 sccm po dobu 4 minut. Čištění terče bylo realizováno pomocí pulzního generátoru Pinnacle Plus připojeného k terči s výkonem 300W, frekvencí 150 kHz, dobou vybití terče 1 µs po dobu 2 minut za při 10 sccm argonu. Během předprašování bylo do aparatury kromě argonu připouštěno 12 sccm kyslíku, při kterém došlo k otrávení terče. Průtok kyslíku byl dále snížen až k hranici před skokem do metalického módu. V posledním kroku depozice byla otevřena clona a došlo k rozprašování vrstvy z terče. Během hledání optimální oblasti pro naprašování vrstev Al_2O_3 byla proměřena celá hysterezní křivka. Díky tomu bylo možné porovnat procesní parametry starého a nového terče (viz Obrázek 1).



Obrázek 1: Hysterezní závislost absolutní hodnoty napětí na magnetronu na průtoku reaktivního plynu

Z hysterezní křivky pro nový terč je patrný posun v ose napětí pro reaktivní mód a výrazný posun napětí pro metalický mód. Ze závislosti je dále patrný posun hysterezní křivky směrem k vyšším průtokům reaktivního plynu. Kromě zjištěného posunu hysterezní křivky byly vytvořeny vrstvy v reaktivním (Al_2O_3 stechiometrický) a kovovém módu (hliník a Al_2O_3 nestechiometrický).

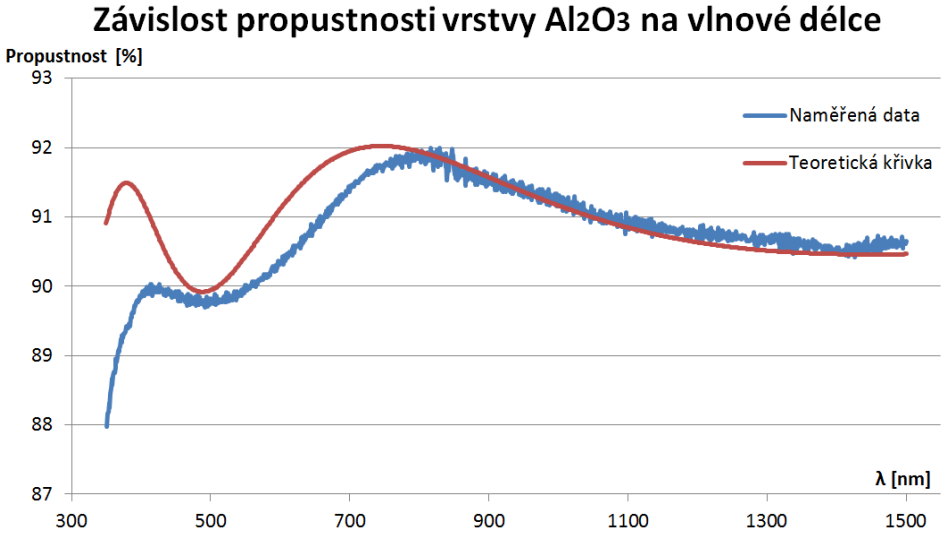
4. TECHNOLOGICKÉ OKNO NOVÉHO TERČE

Díky ověření hysterezní křivky v procesu reaktivního naprašování vrstev Al_2O_3 bylo možné předběžně určit oblast pro naprašování stechiometrických vrstev Al_2O_3. Tato oblast se nachází v části reaktivního módu velice blízko před skokem do kovového módu. Pro testování byly zvoleny 15 minutové depozice v rozmezí napětí 255 V – 295 V a průtoku kyslíku 1,95 sccm – 10 sccm. Během vyhodnocování výsledků bylo zjištěno, že se s rostoucí tloušťkou začne docházet k loupání vrstev při vysokém průtoku kyslíku. Pro vrstvy o tloušťce kolem 230 nm bylo možné použít rozmezí napětí na magnetronu 275 V – 295 V a průtoku kyslíku 1,95 – 4 sccm. Doba depozice se pohybovala v rozmezí 48 – 68 minut (depoziční rychlost od 3,38 nm/min do 4,8 nm/min).

5. TESTOVÁNÍ VRSTEV Al_2O_3

Pro testování stechiometrických vrstev Al_2O_3 byl zvolen test mechanických vlastností, zejména tvrdosti a modulu pružnosti. Dále adhezi měřenou pomocí scratchtestu. Tyto vlastnosti jsou testovány, ale výsledky zatím nejsou k dispozici. Z optických vlastností je testována propustnost ve

viditelném a infračerveném spektru. Na obrázku níže je možné vidět naměřenou a teoretickou charakteristiku propustnosti vrstvy Al_2O_3 v závislosti na vlnové délce. Naměřený průběh je pro depoziční podmínky napětí na magnetronu 294V a průtok kyslíku kolem 1,95 sccm. Propustnost (viz **Obrázek 2)** byla měřena pomocí spektrofotometru Varian Cary 5E.



Obrázek 2: Závislost propustnosti vrstvy Al_2O_3 na vlnové délce

Ze závislosti propustnosti na vlnové délce lze odečíst propustnost v infračerveném spektru, která se pohybuje kolem 91 %, ve viditelném spektru od 90 % do 92%. Dále je viditelný rozdíl v maximální propustnosti při vlnové délce: $\lambda = 750$ nm. Toto maximum odpovídá teoretické tloušťce vrstvy kolem 240 nm až 250 nm. Tloušťka měřená pomocí profilometru Talystep byla 233 nm.

6. ZÁVĚR

V uvedeném článku byla stručně popsána problematika pulzního reaktivního magnetronového naprašování. Dále byly navrženy a ověřeny parametry pro depozice tenkých vrstev Al_2O_3 spolu s hysterezní křivkou. Z této hysterezní závislosti bylo stanoveno technologické okno. V tomto technologickém okně byly provedeny depozice vrstev Al_2O_3 na vzorky skla a křemíku. Tyto vzorky byly dále podrobeny optickým testům, při kterých byla zjištěna propustnost kolem 91 % - 92 %. Kromě testu propustnosti byly provedeny také testy mechanických vlastností a adheze. Výsledky zbylých měření nejsou v současné době dostupné, protože probíhá jejich zpracování.

REFERENCES

[1] SPROUL, W. D. Reactive direct current magnetron sputtering of aluminum oxide coatings. *Journal of Vacuum Science*[online]. **13**(3), s. 1188- [cit. 2016-03-10]. DOI: 10.1116/1.579859. ISSN 07342101. Dostupné z: http://scitation.aip.org/content/avs/journal/jvsta/13/3/10.1116/1.579859

[2] KELLY, P. J. a R. D. ARNELL. Control of the structure and properties of aluminum oxide coatings deposited by pulsed magnetron sputtering. *Journal of Vacuum Science* [online]. 1999, **17**(3), s. 945- [cit. 2016-03-10]. DOI: 10.1116/1.581669. ISSN 07342101. Dostupné z: http://scitation.aip.org/content/avs/journal/jvsta/17/3/10.1116/1.581669

[3] HOUSKA, J., J. BLAZEK, J. REZEK a S. PROKSOVA. Overview of optical properties of Al2O3 films prepared by various techniques. *Thin Solid Films* [online]. 2012, **520**(16), s. 5405-5408 [cit. 2016-03-10]. DOI: 10.1016/j.tsf.2012.03.113. ISSN 00406090. Dostupné z: http://linkinghub.elsevier.com/retrieve/pii/S004060901200404X

SOLAR POWER INVERTER

Petr Gottwald

Master Degree Programme (2), FEEC BUT

E-mail: xgottw05@stud.feec.vutbr.cz

Supervised by: Michal Pavlík

E-mail: pavlik@feec.vutbr.cz

Abstract: Design of a solar power inverter with 1 kW output power capability is discussed in this paper. Interleaved boost topology used in the step-up stage as well as the output full-bridge stage are described. Control block utilizing Field-programmable-gate-array (FPGA) is described as well.

Keywords: FPGA, boost converter, full-bridge converter, MPPT

1. INTRODUCTION

Solar power is becoming more and more important in these days. While there are many MCU-based inverter designs reported in the literature, designs based on FPGAs are rather rare. FPGAs, compared to MCUs (or DSPs), may offer additional advantages like parallel signal processing capability and pin assignment versatility, which may ease the PCB design.

2. POWER CONVERSION PROCESS

Commercially available photovoltaic panels produce DC voltage ranging from few volts to few hundred volts. This DC voltage must be converted to an AC voltage in order to make the power available to the supply network. Furthermore, the power converter serves another purpose - it has to draw maximal power available from the panel. The input voltage and current are measured to determine if the power drawn from the panel is at the highest possible level. In other words, it does source and load impedance matching.

2.1. CONVERTER TOPOLOGY

Transformer-less boost converter was selected for the first (DC-DC) stage. It has some advantages compared to, for example, push-pull topology - the effects of parasitic inductances are not as severe as in transformer-based topologies, there is no "flux walking" problem, which leads to transformer core saturation. Full-bridge topology was selected as the output DC-AC stage. The input voltage is expected to be as low as 45 V. Assuming that the wasted power of the entire converter at maximum output power is 50 W (including losses in transistors, diodes, inductor windings and cores, capacitors), the total input power is 1050 W. Thus, the input current can be approximately calculated:

$$I = \frac{P_{OUT} + P_{LOSS}}{V_{IN(MIN)}} = \frac{1050}{45} = 23{,}3A \quad (1)$$

To decrease switching and conduction losses, the DC-DC stage was divided into four sections, having inductor currents shifted 90 degrees between each phase. This results in average current of 5,83 A through each phase. The output DC-AC stage is single - it does not handle as large current as the boost stage. Output sine voltage is generated by means of pulse width modulation. The switching frequency of both stages has been selected 20 kHz - to keep switching losses at reasonable level and to reduce EMI emissions in the frequency range of interest. PWM modulated signal of each full-bridge converter leg is filtered by LC filter consisting of 1 mH inductor and 0,0047 mF

capacitor, which means cut-off frequency of 2,3 kHz. The block diagram of the system is shown in Figure 1.

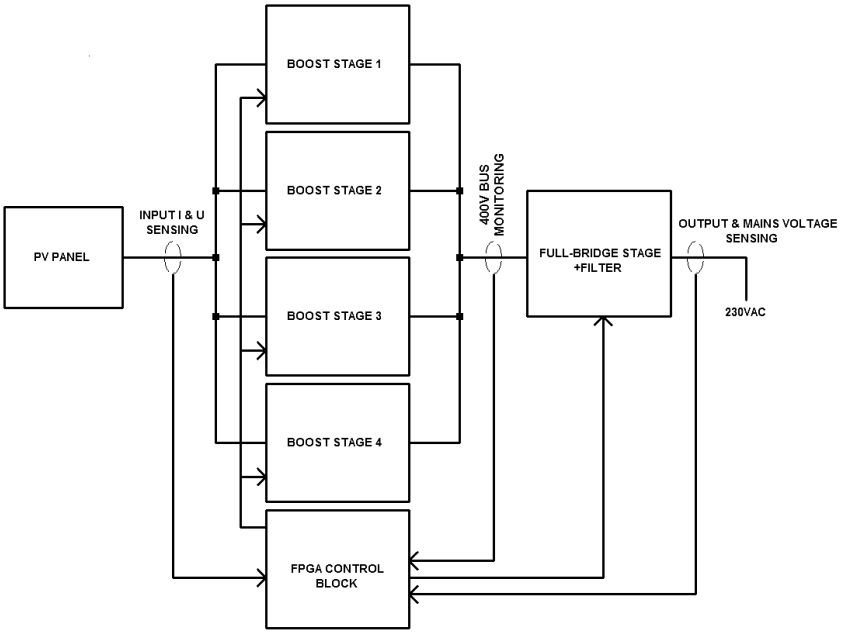


Figure 1: Solar power inverter simplified block diagram

FPGA XC3S200A from Xilinx, together with its power supply DC-DC converters, USB-UART bridge, reference voltage source, ADC converter and current/voltage sensing amplifiers, is placed on a separate six-layer PCB for the purpose of prototyping. Each boost converter phase has its separate double-sided PCB and full-bridge converter has its own double-sided PCB. This approach enables additional features, for example almost arbitrarily increase number of boost phases, or a 3-phase output voltage. Full-bridge stage block diagram is displayed in Figure 2.

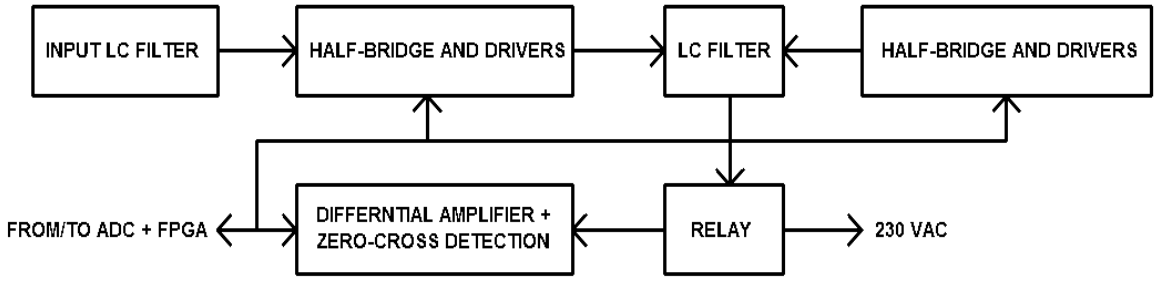


Figure 2: Full-bridge stage simplified block diagram

There are two possible modes of operation - either standalone or connected to the AC mains, in which case the generated sine wave voltage follows the mains voltage. For such case it is very important that the mains and inverter output are separated from each other, until the inverter output is in phase with mains voltage. This is accomplished by power relay. After startup, the phase of the generated voltage is increased, until the difference between mains and generated voltage is zero. Thereafter, the inverter output is connected to the mains and the mains voltage is replicated.

2.2. MAXIMAL POWER POINT CALCULATION (TRACKING)

Assuming constant load connected to the 400 VDC bus, the output voltage of the boost stage is given solely by input power. When lightly loaded, MPPT logic forces maximal duty cycle of the boost stage, in order to keep input power at the maximal level. MPPT logic is based on "Perturb and observe"[1] algorithm. The input current and voltage are multiplied to obtain the input pow-

er. Next a numerical differentiation is conducted to get information if the positive change of duty cycle has caused positive change of the input power. If not, it says that maximal power point has already been reached and the duty cycle is decreased. The duty cycle is decreased as long as the input power rises. After the input power has begun to drop, the duty cycle is increased again and so on.

2.3. SINE WAVE GENERATION

Sine wave samples are stored in the block memory of the FPGA. The principle of Direct digital synthesis was used for this purpose. Phase accumulator (frequency) resolution of 32-bits was selected. This resolution gives output voltage frequency step of approximately 0,0116 Hz (assuming FPGA clock of 50 MHz), which seems adequate. Sine sample resolution is 10-bit amplitude resolution and 10-bit phase resolution - only 10 upper bits of phase accumulator are used to address sine samples in the block memory. The rest of phase accumulator bits, which are not used for sample addressing, is truncated. Then, the sine samples are digitally attenuated (multiplied by factor of 0,85 - multiplied by 205 and 8 upper bits used as the result), which yields maximal amplitude of the generated signal of 330 V, when the inverter input is supplied from the boost stage output, which has nominal voltage of 400 V. The resulting gate drive signal is passed through anti-cross-conduction logic block, which has to avoid that both transistors of a full-bridge leg are switched-on simultaneously.

The resulting sine signal has distortion of approximately 0,7 % as shown in Figure 3, assuming that the most significant high-frequency spectral component (PWM carrier signal in this case) is separated from the 1st harmonic by 42 dB and other harmonics can be neglected as they are too low in magnitude. The converter has been simulated using LTspice IV circuit simulator.

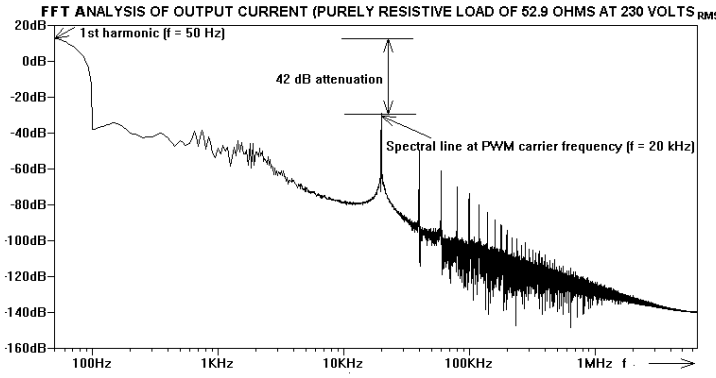


Figure 3: Output current distortion (vertical axis referred to 1 A_{RMS})

3. CONCLUSION

The solar power inverter was designed, simulated and briefly described. The PCBs were designed for the final realization of the project. Simulations have shown theoretical efficiency of 98 %. Nevertheless, the efficiency of the final inverter may vary, due to inherent difficulties of circuit simulations with respect to circuit parasitic elements, which are difficult to model.

REFERENCES

[1] MathWorks. 2010. MPPT Algorithm. [ONLINE] Available at: http://de.mathworks.com/discovery/mppt-algorithm.html. [Accessed 14 March 16].

STREAMLINING THE PROCESS OF CLEANING PCB AFTER REMOVING BGA

Tomáš Starčok
Master Degree Programme (2), FEEC BUT
E-mail: xstarc00@stud.feec.vutbr.cz

Supervised by: Radek Vala
E-mail: radek.vala@sanmina.com

Abstract: This paper deals with issues connected with solder residues cleaning process of modern BGA packages. First part is focused on contact cleaning method of solder residues. It describes main disadvantages and incurred defects of this method. Second part is focused on a new contactless cleaning method with special tool and then suggests variety of verification possibilities.

Keywords: BGA, solder, residues, rework

1 ÚVOD

Neustály rozvoj a miniaturizácia v oblasti mikroelektroniky viedla ku potrebe vytvoriť puzdro na čipové súčiastky, ktoré by vyhovovalo moderným trendom. Hlavnými požiadavkami bolo zmenšenie celkovej plochy puzdra a vyvedenie vývodov z puzdra čo najefektívnejším spôsobom a zároveň aby tieto vývody mohli byť spájkovateľné dostupnými technológiami. Puzdra s vývodmi po stranách prestávali vyhovovať a do popredia sa začali dostávať technológie všeobecne založené na povrchovej montáži a spájkovania pretavením. Preferovanými sa stali puzdra s maticovým usporiadaním vývodov vo forme spájkových guličiek, ktoré sú všeobecne známe pod označením Ball Grid Array [1]. Tieto puzdra sú v súčasnej dobe veľmi rozšírené a od prvých aplikácií sa dosť výrazne zmenšili rozstupy jednotlivých vývodov Na pracoviská zaoberajúce sa procesmi spojenými s týmito puzdrami sú často kladené oveľa vyššie nároky na technológie a skúsenosti. Aby finálne riešenie bolo schopné čo najviac konkurovať, je dôležité klásť dôraz hlavne na ekonomické aspekty každého kroku pri výrobe. Práve kvôli tomuto dôvodou je väčšina výrobkov navrhovaná ako neopraviteľná, nakoľko z ekonomického hľadiska nie je prínosné, aby boli realizovateľné opakovateľné opravy. Napriek tomuto existujú aj výrobky pri ktorých sú opravy montážnej zostavy výhodné. Problematika uvedených puzdier je pri takýchto opravách zastúpená v značnej miere. Ako už bolo spomenuté, pri tomto procese je potrebné uvažovať vplyv veľkého množstva rozličných faktorov [2].

2 KONTAKTNÁ METÓDA ODSTRAŇOVANIA SPÁJKY

Aktuálne najviac používanou metódou na odstraňovanie zvyškov spájky je kontaktná metóda. Jedná sa o manuálnu metódu a odstraňovanie zvyškov spájky je realizované prostredníctvom zaškolených operátorov. Pri čistení sa používa kontaktné spájkovanie pomocou ručnej spájkovačky, špeciálny spájkovací hrot, odsávací medený pásik, gélové tavidlo a isopropylalkohol. Pri tejto metóde je značný vplyv ľudského faktora na výslednú kvalitu očistenia. Pri tejto metóde čistenia je opakovane tepelne namáhaná lokálna časť DPS a zároveň môžu vznikať rôzne mechanické poškodenia na povrchu DPS.

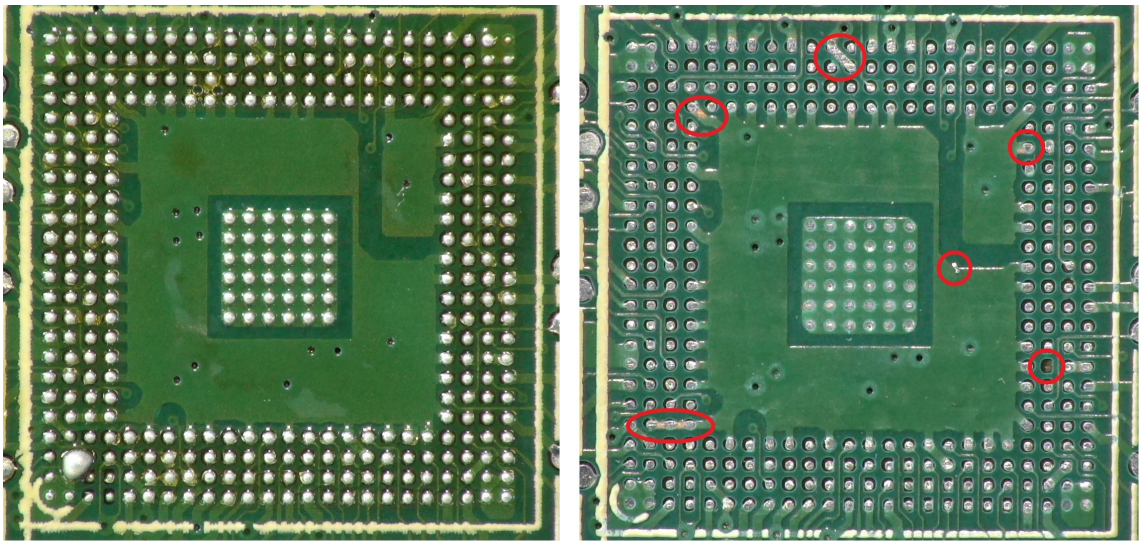
2.1 ANALÝZA VZORIEK

Optická kontrola vzoriek DPS po kontaktnom odstránení zvyškov spájky po odobratí BGA puzdra bola realizovaná prostredníctvom optického stereo mikroskopu a pre nafotenie vzoriek bol použitý

digitálny mikroskop s Full HD rozlíšením.

Stav vzorky po odobratí požadovaného BGA puzdra je zobrazený na obrázku 1. Oblasť pod BGA puzdrom obsahuje po odobratí značné množstvo prebytočnej spájky na spájkovacích plôškach. Túto prebytočnú spájku je nutné odstrániť pred osadením a spájkovaním nového BGA obvodu.

Po kontaktnom čistení táto vzorka obsahuje pomerne veľké množstvo poškodení na nespájkovacej maske, delaminované spájkovacie plôšky a malé množstvo zvyšnej spájky na nespájkovacej maske. Všetky uvedené defekty, ktoré vznikajú pri kontaktnom odstraňovaní spájky sú neakceptovateľné pre ďalšie kroky opravy DPS spojené s osadením a spájkovaním nového BGA obvodu. Výsledný stav jednej zo vzoriek je zobrazený na obrázku 1.



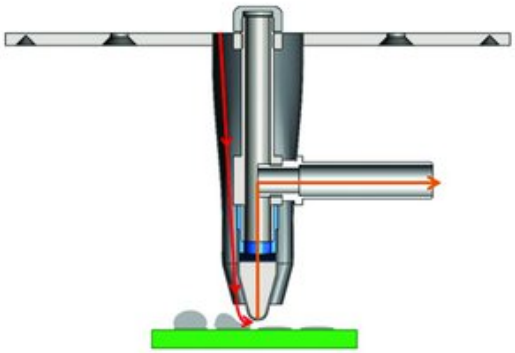
Obrázek 1: Ukážka vzorky, vľavo po odobratí BGA, vpravo po kontaktnom čistení

3 BEZKONTAKTNÁ METÓDA ODSTRAŇOVANIA SPÁJKY

Manuálne kontaktné čistenie zvyškov spájky nie je opakovateľný proces, ktorý prináša hlavne riziko poškodenia okolitých súčiastok, nespájkovacej masky, plošných vodičov alebo spájkovacích plôšiek. Tieto defekty môžu mať za následok poškodenú DPS, ktorá ďalej nemôže pokračovať v procese opravy. Momentálne existuje niekoľko riešení bezkontaktného čistenia, ktoré sa odlišujú technologickým spracovaním, a ako príklad jedného z nich, uvádzam riešenie od spoločnosti Finetech. Ďalšie spoločnosti, ktoré sa zaoberajú problematikou bezkontaktného čistenia sú napríklad Martin a Zevac. Spoločnosť Finetech vyvinula a ponúka riešenie cestou bezkontaktného odstránenia spájky v jedinom kroku, čo umožňuje bezpečné, opatrné a hlavne opakovateľné zaobchádzanie, a to i na veľmi malých miestach DPS s vysokou hustotou integrácie súčiastok. Po spustení procesu pretavenia je možné roztavenú spájku ľahko odstrániť z DPS pomocou výkonného vákua [3]. Nová generácia nástrojov ku odstaňovaniu spájky umožňuje bezkontaktné odstránenie zvyškov spájky bez narušenia spájkovacích plôšiek alebo nespájkovacej masky. Susedné súčiastky a DPS budú spoľahlivo chránené pred nežiaducim poškodením. Princíp bezkontaktného nástroja určeného na odstránenie spájky je zobrazený na obrázku 2.

3.1 PLÁNOVANÁ ANALÝZA VZORIEK

Analýza vzoriek bude realizovaná prostredníctom priamej alebo nepriamej metódy. Priamou metódou sa uvažuje precízne váženie vzoriek DPS pred a po odstránení zvyškov spájky. Rozdiel nameraných



Obrázek 2: Princíp bezkontaktného nástroja na odstránenie spájky od firmy Finetech [3]

hodnôt nám určí hmotnosť spájky, ktorá bola odstránená. Porovnaním rozdielov nameraných hodnôt medzi referenčným kontaktným a novým bezkontaktným spôsobom odstránenia spájky, sa vyhodnotí miera kvality nového procesu čistenia. Nepriame metódy vyhodnotenie sú napríklad 3D AOI, SPI alebo analýza povrchu profilometrom. Jednou z uvedených metód bude možné analyzovať a vyhodnotiť vzhľad povrchu spájkovavích plôšok po odstránení zvyškov spájky obidvomi uvedenými spôsobmi.

4 ZÁVER

Práca sa zaoberá praktickým porovnaním, analyzovaním a vyhodnotením dvoch rôznych metód odstraňovania zvyškov spájky na DPS po odobratí BGA puzdra. Cieľom je preskúmať možnosti automatického bezkontaktného čistenia DPS pomocou špeciálneho nástroja a zvoliť správnu metódu pre vyhodnotenie kvality tohoto nového procesu čistenia. Výhodou novej metódy je eliminovanie najčastejšie vznikajúcich defektov akými sú poškodenie nespájkovacej masky, delaminácia spájkovacej plôšky a zvyšky spájky na nespájkovacej maske. Všetky uvedené defekty na opravovaných DPS sú neakceptovateľné pre ďalší postup, akým je osadenie a spájkovanie nového BGA puzdra. Táto nová metóda čistenia zvyškov spájky sa javí byť perspektívna pre dosiahnutie ďalšieho zvýšenia výťažnosti procesu opráv DPS osadených BGA obvodmi.

POĎAKOVANIE

Tento príspevok vznikol za podpory spoločnosti Sanmina-SCI Czech Republic s.r.o., ktorá mi umožnila v ich priestoroch spracovávať moju diplomovú prácu.

REFERENCE

[1] Smithsonian Corporation. *10 Grid Array Packaging: BGA and CSP* [online]. 2009, [cit. 2015-10-28]. Dostupné z: `http://smithsonianchips.si.edu/ice/cd/PKG_BK/CHAPT_10.PDF`.

[2] SZENDIUCH, I. *Základy technologie mikroelektronických systémů obvodů systémů*. Vyd. 1. Brno: VUTIUM, 2006, 379 s. ISBN 80-214-3292-6.

[3] Finetech GmbH. *Residual Solder Removal* [online]. [cit. 2015-11-29]. Dostupné z: `http://www.finetechusa.com/rework/processes/residual-solder-removal.html`

Magisterské projekty

Biomedicínské inženýrství

ACTIVE CONTOURS INITIALIZATION FOR CELL TRACKING IN THE IMAGES FROM HOLOGRAPHIC MICROSCOPE

Tomáš Vičar
Master Degree Programme (2), FEEC BUT
E-mail: xvicar03@stud.feec.vutbr.cz

Supervised by: Radim Kolář
E-mail: kolarr@feec.vutbr.cz

Abstract: This paper describes the implementation of the segmentation method applied on images from holographic microscope. The method combines tresholding and active contour model. This segmentation method can be used on the first image of image sequence for creating contour initialization. Initialized contours, then can be used for cell tracking with active contours.

Keywords: holographic microscope, cell tracking, active contours

1 ÚVOD

Vysoké učení technické společně se společností TESCAN vyvinuly multimodální holografický mikroskop, který je vhodný pro dlouhodobé experimenty na živých buňkách. Experiment trvá typicky několik desítek hodin a během něj se nasnímá několik stovek obrazů, kdy každý obsahuje až několik buněk. Dostáváme velké množství obrazových dat, proto je vhodné jejich automatické zpracování. Zpracování těchto dat je založeno na segmentaci jednotlivých buněk, jejich sledování v sekvenci a následné extrakci vybraných parametrů (např. velikost či pozice).

Pro segmentaci byla zvolena metoda aktivních kontur, pro kterou je zásadní správná inicializace bodů kontury v prvním snímku sekvence. Tento článek popisuje metodu inicializace založenou na segmentaci jader, kolem kterých je vytvořena kontura, která je následně rozpínána na velikost buňky metodou aktivních kontur.

2 PRINCIP AKTIVNÍCH KONTUR

Aktivní kontury vycházejí z minimalizace energie v obraze umístěné kontury $\mathbf{v}(s)$. Energii této kontury lze obecně popsat jako

$$E_{kon} = \int E_{ext}(\mathbf{v}(s)) + E_{int}(\mathbf{v}(s))ds, \tag{1}$$

kde externí energie E_{ext} je dána pozicí kontury v obraze. Interní energie E_{int} je dána tvarem kontury a má udržet hladkost křivky.

Popisovaná metoda využívá parametrických aktivních kontur, kdy je křivka zadána svými body. Interní energie se skládá ze složek zajišťujících odolnost vůči protažení/stlačení a odolnost vůči ohybu. Tyto složky počítáme jako první a druhou derivaci po křivce. Při diskrétním výpočtu nahrazujeme tyto derivace pomocí diferencí. Interní síly lze počítat maticově, jak je popsáno v [1]. Minimalizace externí energie v základním modelu dosahujeme pohybem kontury proti směru gradientu, což zajišťuje konvergenci do lokálního energetického minima.

Pro sledování buněk je metoda aktivních kontur vhodná, neboť konturu z předchozího snímku můžeme použít pro inicializaci v následujícím snímku. Tím je do dalšího snímku přenesena informace o předpokládané pozici a tvaru hledané buňky.

3 IMPLEMENTACE METODY

Metoda byla implementována v programovacím prostředí MATLAB, přičemž využívá segmentaci buněčných jader pomocí prahování, vytvoření kontury v okolí jádra a její zvětšení s využitím modifikované metody aktivních kontur.

3.1 VYTVOŘENÍ KONTURY

Cílem prahování jader je dosáhnout binárního obrazu, kde každé jádro tvoří pouze jeden ucelený objekt. Lepších výsledků je možné dosáhnout rozmazáním prahovaného obrazu Gaussovým filtrem a vyčištěním binárního obrazu pomocí morfologických operací. Pro zajištění spolehlivosti je vhodná kontrola uživatelem s případným upravením prahové hodnoty. Převedení binárního obrazu na body kontury je provedeno použitím hraničních pixelů jako bodů kontury.

3.2 ROZPÍNÁNÍ KONTURY

Základní implementace neumožňuje výraznější rozpínání kontury, proto byl místo obyčejného gradientu použit tok gradientního vektoru popsaný v [2]. Metoda upravuje gradientní pole difuzí vysokých hodnot gradientu do oblastí s nízkým gradientem. Kontura je tak méně citlivá na inicializaci a je schopná se z velikosti jádra zvětšit na velikost buňky.

Jako externí energie byl použit rozmazaný obraz intenzity, čímž je zajištěna konvergence do okolí buňky (pozadí má vyšší intenzitu). Velké roztažení kontury může způsobovat chyby segmentace, proto byla vytvořena metoda pro zastavení kontury v oblastech pozadí. Pozadí lze identifikovat jako oblast, kde je malá hodnota rozdílu mezi aktuálním a následujícím snímkem. Zastavení kontury je zajištěno vynulováním gradientního vektoru v oblastech pozadí. Mezi dotýkajícími se buňkami se pozadí nenachází, ale vycházíme z předpokladu, že maximum intenzity je na hranici mezi buňkami.

Zpřesnění segmentace je následně dosaženo opětovnou úpravou kontury s použitím kombinace intenzity a gradientu jako externí energie. Gradient do výpočtu vnáší hranovou informaci a intenzita zajistí preferenci světlejších hranic buňky před hranicemi jádra. Preference hranic aktuální buňky před sousední buňkou je zajištěna diskriminací oblastí vně kontury snížením hodnot externí energie.

3.3 PŘEVZORKOVÁNÍ KONTURY

Velkým problémem metody aktivních kontur je nestabilita kontury - překrytí bodů, přílišné vzdálení bodů a vznik smyček. Tyto vlivy se projeví hlavně při rozpínání kontury nebo sumací při opětovném použití kontury pro sledování buněk. Existují metody pro hledání duplicitních a příliš vzdálených bodů, ty však nezajišťují úplnou stabilitu.

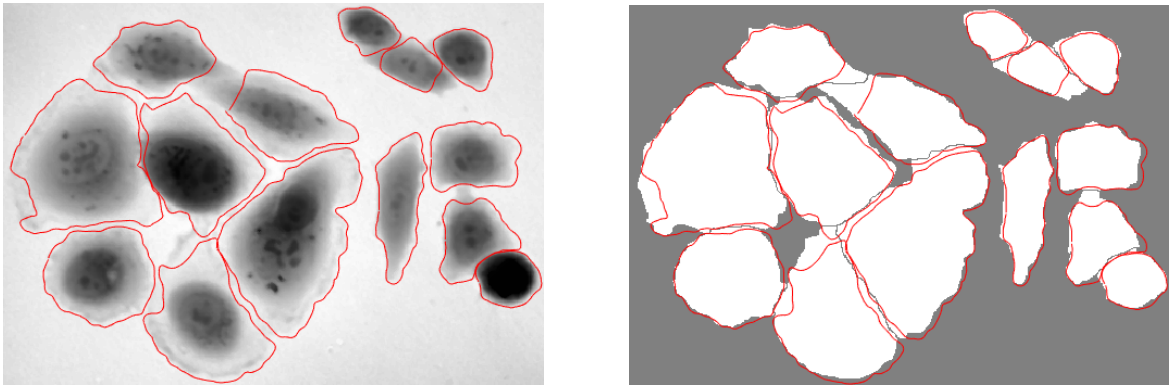
Pro opravu kontury byla vyvinuta metoda převzorkování. Chceme nalézt nové, stejně vzdálené body nacházející se na lineárním splinu spojujícím staré body kontury. Další správně vzdálený bod lze nalézt jako průsečík kružnice (vzdálenost) a přímky odpovídající části splinu. Průsečíky lze vypočítat z matematického popisu, a je nutné ověřit, zda se průsečík nachází na úsečce mezi starými body.

Využití převzorkování zajistí stabilitu kontury, ale změní její vlastnosti na hladší výslednou křivku. Tyto vlivy je možné regulovat použitím převzorkování vždy po určitém počtu iterací.

4 ZHODNOCENÍ DOSAŽENÝCH VÝSLEDKŮ

Popisovaná metoda v testovaných datech úspěšně inicializuje všechny buňky, čehož je dosaženo díky nastavení prahové hodnoty uživatelem. 100% spolehlivosti nemusí být dosaženo ve všech případech, neboť se buňky mohou překrývat. Takové buňky nelze prahováním rozdělit na jednotlivá jádra.

Důležitá je také přesnost výsledné segmentace. Ukázku segmentace ukazuje obrázek 1 na zvětšeném výřezu ze zorného pole 1. sekvence, kde můžeme vidět kontury zobrazené v původním obrazu a také porovnání s ručně segmentovanými daty. Tabulka 1 vyhodnocuje přesnost segmentace v porovnání s ručně segmentovanými daty. Hodnoty jsou vyjádřeny v poměru se segmentovanou plochou. Větší nepřesnost u falešně negativních pixelů (identifikace pixelů buněk jako pixelů pozadí) je způsobena slabými hranami mezi dotýkajícími se buňkami. Kontura se za slabou hranu nezachytí a dojde k jejímu smrštění až k okrajům jádra. Zbytek nepřesnosti u falešně negativních pixelů a nepřesnost falešně pozitivních pixelů jsou způsobeny malými chybami na okrajích buněk. To je způsobeno vlastnostmi kontury, která si udržuje hladkost a dochází tak k aproximaci malých výčnělků. Jelikož buňky mají malý kontrast, nelze ani ruční segmentaci považovat za 100% přesnou.



(a) Výsledek segmentace **(b)** Porovnání s ruční segmentací

Obrázek 1: Ukázka výsledné segmentace.

Číslo sekvence	Počet buněk	Průměrná velikost	Falešně pozitivní	Falešně negativní
1	25	4629 px	5,90 %	10,77 %
2	10	4831 px	6,82 %	8,33 %

Tabulka 1: Přesnost segmentace jako poměr segmentované plochy

5 ZÁVĚR

Cílem této práce bylo vytvoření iniciačních kontur pro sledování buněk s využitím modelu aktivních kontur. Popsaná a implementovaná metoda zajišťuje dostatečně přesnou inicializaci. Pro zajištění spolehlivosti je však nutná interakce s uživatelem, který nastaví prahovou hodnotu.

V budoucí práci bude tato metoda použita pro inicializaci sledování buněk. Pro úpravu kontury v dalším snímku je možné vhodně modifikovat implementaci aktivních kontur, která je použita při vytváření inicializační kontury.

REFERENCE

[1] KASS, Michael, Andrew WITKIN a Demetri TERZOPOULOS. Snakes: Active contour models. *International Journal of Computer Vision*. 1988, vol. 1, issue 4, s. 321-331. DOI: 10.1007/BF00133570.

[2] CHENYANG XU a J.L. PRINCE. Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*. vol. 7, issue 3, s. 359-369. DOI: 10.1109/83.661186.

ALIGNMENT-FREE METHODS FOR CLASSIFICATION OF METAGENOMIC DATA

Tereza Vaněčková

Master Degree Programme (2), FEEC BUT

E-mail: xvanec02@stud.feec.vutbr.cz

Supervised by: Helena Škutková

E-mail: skutkova@feec.vutbr.cz

Abstract: Metagenomics studies microbial communities by analyzing their genomic content directly sequenced from the environment. In this contribution, alignment-free methods based on word frequency will be introduced. It has been proven, that these methods are effective in processing of short metagenomic sequence reads produced by Next-Generation Sequencing technologies. To evaluate the potential of word frequency based methods, the *k*-mer analysis was applied on simulated dataset of metagenomic sequence reads with length of 600 nucleotides. Then the data were enrolled for a hierarchical cluster analysis. Results have shown that the proposed method is able to cluster genome fragments of the same taxa.

Keywords: metagenomics, alignment-free, nucleotide word frequency, hierarchical clustering

1. INTRODUCTION

With rapid development of sequencing technologies and applications, metagenomics is becoming an important approach for studying microbial communities in different enviroments and the human body. Metagenomic datasets consist of many raw reads that represent various parts of many individual genomes. [1]

The data volume generated by Next-Generation Sequencing (NGS) technologies is growing. Therefore, handling and processing such large datasets is becoming one of the major challenges in most metagenome research projects. [2] Alignment-based methods encounter difficulties in dealing with large datasets. To overcome this limitation many alignment-free methods have been proposed. Among them the method based on nucleotide word frequencies (*k*-mer analysis) may be the most developed alignment-free method. [3]

In this study the aim is to evaluate the ability of *k*-mer analysis approach to cluster reads of shotgun metagenomic data.

2. METHODOLOGY

Sequence signatures are normalized frequencies of *k*-mers in a sequence. Previous studies have shown that *k*-mer frequencies are similar across different regions of the same genome, but differ between genomes of different organisms. Therefore, they carry phylogenetic information. [3, 4]

2.1. NUCLEOTIDE WORD FREQUENCY

K-mer is a word that consists of *k* symbols from a finite alphabet of nucleotides A = {A, C, G, T}. The set W_k consists of all possible *k*-mers, also refers to variation with repetition (1). This set has *n* elements and the words are sorted alphabetically. [3]

$$W_k = \{w_{k,1}, w_{k,2}, \dots, w_{k,n}\} \qquad (1)$$

$$n = 4^k$$

Computationally, the counting of *k*-mers in a sequence of length *m* is usually performed by taking a sliding window of length *k*, that moves through the sequence from position 1 to *m* - *k* + 1. The overlapping of *k*-mers is allowed in this method. A sequence can be represented by *n*-dimensional vector c_k consisting of *k*-mers counts (2) [3]:

$$c_k = \left(c(w_{k,1}), c(w_{k,2}), \dots, c(w_{k,n})\right), \tag{2}$$

where *n* is the total number of possible *k*-mers. Vector of *k*-mer frequencies is counted as a relative number of each *k*-mer according to (3) [3]:

$$f_k = \left(f(w_{k,1}), f(w_{k,2}), \dots, f(w_{k,n})\right) \tag{3}$$
$$= \left(\frac{c(w_{k,1})}{m-k+1}, \frac{c(w_{k,2})}{m-k+1}, \dots, \frac{c(w_{k,n})}{m-k+1}\right).$$

2.2. HIERARCHICAL CLUSTERING

Hierarchical cluster analysis is widely used method for classification of features based on *k*-mer frequencies. [5, 6] It is a task of grouping a set of objects in such a way that objects in the same cluster are more similar to each other than to those in the other groups. Mutual similarity of sequences can be shown by creating a cluster tree or dendrogram. Ideally, branches of the dendrogram should contain a cluster of sequences of the same taxon. The hierarchical cluster analysis can be performed on a dataset by using following procedure [7]. First step is to find similarities or dissimilarities between every pair of objects (vectors of *k*-mer frequencies). Various distance metrics can be used with one of the most widespread Euclidean. Next, the objects are grouped into a binary hierarchical cluster tree with use of a linkage function (eg. nearest neighbor or furthest neighbor). The last step is to determine the number of clusters.

3. RESULTS

Analysis, as shown below in **Figure 1**, was performed on the simulated metagenomic dataset. This dataset consists of 60 genomic fragments (with length of 600 nucleotides) originating from 3 bacterial genomes of different bacterial phyla. For each of the fragments, the vector of *k*-mer frequencies (where *k*=7) was calculated. Based on the calculation of cophenetic correlation coefficient, the method of Spearman distance and furthest neighbor linkage function was determined as the most powerful method for clustering.

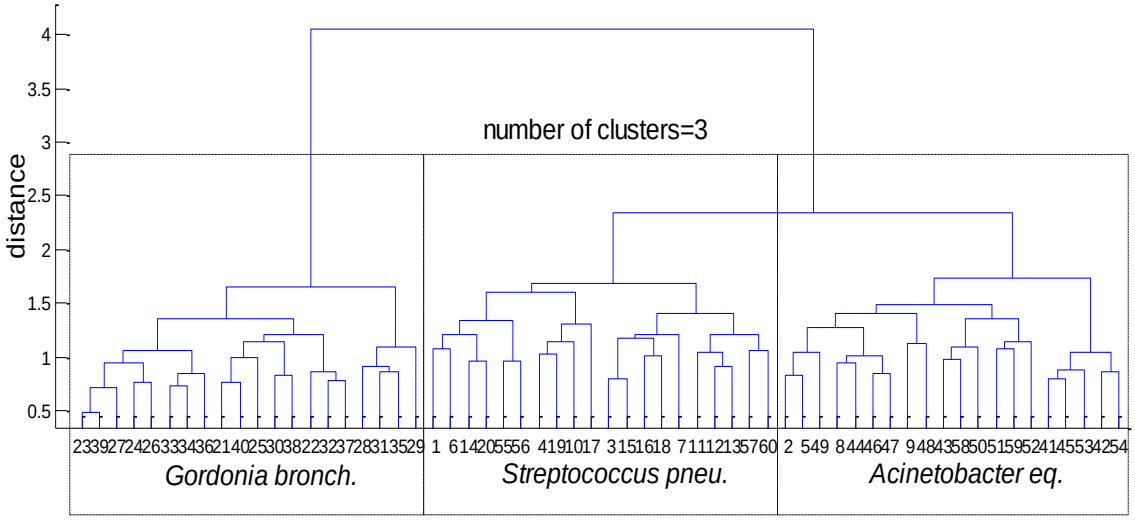


Figure 1: Analysis of 60 genomic fragments of 3 bacterial genomes. Rectangles represent clusters.

4. DISCUSSION AND CONCLUSION

To evaluate the potential of word frequency based methods for classification of metagenomic data, the *k*-mer analysis was applied on 60 simulated metagenomics sequence reads with length of 600 nucleotides. Then the data were enrolled for hierarchical cluster analysis.

As shown in the example (Figure 1), this method is able to separate sequences of taxonomically distant species, therefore to create clusters of genomic fragments originating from the same taxon. In this case, bacterial species of different bacterial phyla were used. The impact of *k*-mer length on method performance can be seen from the **Chyba! Nenalezen zdroj odkazů.** below. The number of misclassified fragments in analyzed dataset was 13 (78,3% success rate) when *k*=2. In contrast, when *k*=7, there was 8 misclassifications (86,6% success rate). Although there were found slightly better results with increasing length of *k*-mer, there is no point in counting *k*-mers longer than 7 in reads of length approximately 600 *bp* as no significant improvement in performance of the algorithm was found. The algorithm gives similar results to that publicated by Yang *et al*. [6], where rank orders of values of *k*-mer counts in a sequence are considered.

Length of *k*-mer	2	3	4	5	6	7
Success rate - this study [%]	78,3	76,6	76,6	80	78,3	**86,6**
Success rate - Yang *et al*. [%]	78,3	70	76,6	**86,6**	76,6	71,6

Table 1: Impact of *k*-mer length on success rate of the algorithms

The method proposed above has several unique features. First, this method is unsupervised, therefore it does not require any prior knowledge for the clustering or binning of the data. Next, no reference database is needed, as assessing similarity to existing genome sequence databases is very often limited by their incompleteness. Moreover, this method does not require sequence alignment, hence it is less computationally demanding.

However, further research regarding extracting sequence features based on *k*-mer frequencies and classification methods is needed. Limitation of this method may be lower accuracy in finding clusters in data that contain sequences of lower taxonomic rank, e.g. genus or family. One of the reasons can be high similarity of the *k*-mer frequency vectors between these sequences.

REFERENCES

[1] BASHIR, Y., S. PRADEEP SINGH and B. KUMAR KONWAR. Metagenomics: An Application Based Perspective. *Chinese Journal of Biology*. Hindawi Publishing Corporation, 2014. DOI: 10.1155/2014/146030.

[2] HODKINSON, B. P. and E. A. GRICE. Next-Generation Sequencing: A Review of Technologies and Tools for Wound Microbiome Research. *Advances in Wound Care*. 2015, 4(1), 50-58. DOI: 10.1089/wound.2014.0542.

[3] VINGA, S. and J. ALMEIDA. Alignment-free sequence comparison-a review. *Bioinformatics.* Oxford, UK. 2003, 19(4). ISSN 13674803.

[4] TEELING, H. *et.al*. Application of tetranucleotide frequencies for the assignment of genomic fragments. *Environmental Microbiology*. Oxford, UK: Blackwell Science Ltd, 2004, 6(9), 938-947. ISSN 14622912.

[5] JIANG, B. *et al*. Comparison of metagenomic samples using sequence signatures. *BMC Genomics*. London: BioMed Central, 2012, (13). DOI: 10.1186/1471-2164-13-730.

[6] YANG, X. and T. WANG. A novel statistical measure for sequence comparison on the basis of k-word counts. *Journal of Theoretical Biology*. 2013, (318), 91-100. DOI: 10.1016/j.jtbi.2012.10.035. ISSN 00225193.

[7] GENTLE, J. E., L. KAUFMAN and P. J. ROUSSEU. Finding Groups in Data: An Introduction to Cluster Analysis. *Biometrics*. 1991, 47(2), ISSN 0006341x.

ALIGNMENT-FREE VISUALIZATION OF METAGENOMIC DATA BY GENOMIC SIGNAL PROCESSING

Kristýna Kupková

Master Degree Programme (2), FEEC BUT

E-mail: xkupko00@stud.feec.vutbr.cz

Supervised by: Karel Sedlář

E-mail: sedlar@feec.vutbr.cz

Abstract: Alignment-free visualization of metagenomic data without prior knowledge about organism composition is one of the major issues in metagenomic bioinformatics. Such a visualization method must place fragments from one organism in close proximity to each other, be reproducible and with rapidly increasing amount of data also work ideally in linear time. A novel technique fulfilling all of these requirements is introduced in this paper. The method is based on transformation of genomic sequence into phase signal representation with extraction of three Hjorth's descriptors forming three dimensional space for visualization.

Keywords: metagenome; numerical representation; visualization; Hjorth's descriptors

1. INTRODUCTION

Next generation sequencing (NGS) with third generation sequencing (TGS) have allowed a rapid progress in metagenomics. However, the amount of data produced by both the NGS, and TGS is enormous which leads to the need for fast and efficient techniques which would allow interpretation of this data. Alignment-free visualization of metagenomic data without prior knowledge about the taxonomical composition is one of the major challenges. The state of the art is based on Barnes-Hut Stochastic Neighbor Embedding (BH-SNE) of oligonucleotide signatures with runtime of $O(n \log n)$ where n represents the number of fragments [1]. A new visualization method based on transformation of genomic sequences into signal representation followed by feature extraction is introduced in this article. The major advantage over the BH-SNE is not only the non-stationary character of this method which allows its reproducibility, but also the runtime of $O(n)$, rapidly accelerating the whole process of visualization as shown further.

2. MATERIALS AND METHODS

2.1. DATASET

In order to test the method effectiveness and compare it with the technique using BH-SNE algorithm, the same simulated dataset as in [1] was used. This dataset comprises of genomic fragments (three fragment lengths were tested: 500 nt, 1,000 nt, and 5,000 nt) of ten different taxa further introduced in Figure 1:. The whole genome sequences were obtained from National Center for Biotechnology Information (NCBI) GenBank database.

2.2. SIGNAL REPRESENTATION AND HJORTH'S DESCRIPTORS

The phase of a complex number has been chosen as a transformation technique for obtaining numerical sequences from genomic fragments. In case of this paper the method assigns the values $\{-3\pi/4, -\pi/4, \pi/4, 3\pi/4\}$ *rad* to the four nucleotides {C, T, A, G} respectively. This representation allows further use of signal processing tools, compared to the character representation, while keeping the biological information stored within the sequence [2].

After the sequence fragments are transformed into numerical signals a set of features consisting of the three Hjorth's descriptors (activity, mobility, complexity) is extracted from each of the signal. The Hjorth's descriptors, often used in e.g. EEG analysis, were selected for their ability to capture the features of non-stationary signals not only in signal time (position) domain but also in frequency domain of spectrum. The descriptors are computed according to equations (1)-(3), where σ_0^2, σ_1^2, and σ_2^2 are the variances of signal and its first and second derivatives respectively [3].

$$A = Activity = \sigma_0^2, \tag{1}$$

$$M = Mobility = \frac{\sigma_1}{\sigma_0}, \tag{2}$$

$$C = Complexity = \frac{\sigma_2 / \sigma_1}{\sigma_1 / \sigma_0} \tag{3}$$

3. RESULTS AND DISSCUSION

The Hjorth's descriptors from each fragment are placed into an imaginary three dimensional space as shown in Figure 1:. Here every point represents one 5000 nt long fragment related to the organism indicated by color in the legend. Except for the two *E. coli* organisms forming one cluster, each of the taxa forms its own cluster allowing a rather reliable visualization of the data.

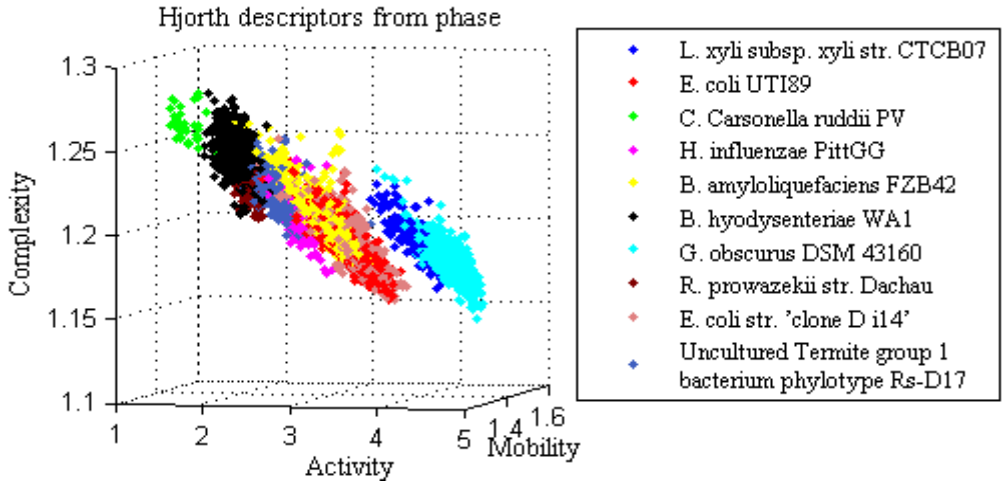


Figure 1: Visualization of Hjorth's descriptors from phase representation of genomic sequences

In order to quantify the effectivity of the introduced visualization technique, k-means clustering was performed. The algorithm was used on all the thee datasets differing in fragment length, specifically 500 nt, 1,000 nt and 5,000 nt. The final statistics for each organism in form of precision (*precision = TP/(TP+FP)*) and accuracy (*accuracy = (TP+TN)/(TP+FN+FP+FN)*) values can be observed in Table 1:. Here, we can see that in the majority of cases, both precision and accuracy values are being improved with the increasing fragment length. Compared to [1], where manual human augmented binning was used for classification, therefore the reached results are highly subjective, k-means algorithm used in this paper is an automated method providing objective results. Despite this fact the average precision and accuracy reach high values of 76.27%, and 92.78% respectively.

Another performance parameter, the runtime of the algorithm for different number of sequences, is presented in Table 2:. Here it is easy to observe that computational complexity of the method with Hjorth's descriptors is significantly lower compared to BH-SNE method and allows almost imme-

diate interpretation of metagenomic data. Furthermore, according to linear fitting with equation $t = 0.00039n + 13$, where t is runtime of Hjorth's descriptors and n represents the number of fragments, the computational complexity of Hjorth's descriptors is really linearly dependent on number of fragments. Whereas the BH-SNE based method is dependent on both number and length of fragments, as it is possible to see from Table 2:.

Table 1: Precision and accuracy values after applying k-means algorithm on Hjorth's descriptors obtained from phase representation of genomic fragments of variable lengths.

Organism	Precision (%)			Accuracy (%)		
	500 nt	1000 nt	5000 nt	500 nt	1000 nt	5000 nt
L. xyli	71.41	74.49	91.63	92.34	93.22	96.42
E. coli	44.24	47.63	42.17	77.74	80.26	75.80
Candidatus C.ruddii	92.97	95.31	92.31	95.53	94.29	99.90
H. influenzae	46.47	66.98	94.26	90.00	92.44	94.93
B. amyloliquefaciens	44.56	46.18	58.66	84.35	82.62	89.53
B. hyodysenteriae	50.15	49.24	70.46	90.87	90.35	96.19
G. obscurus	70.82	74.06	84.51	93.48	94.40	96.76
R. prowazekii	57.89	68.75	79.03	91.44	91.93	94.93
Termite group 1 bacterium	57.82	66.20	73.43	91.06	91.50	91.02

Table 2: Runtimes for the BH-SNE approach and for the Hjorth's descriptors approach implemented in MATLAB® using 2.4 GHz Intel® Core™ i5-2430M CPU.

Fragment length (nt)	No. of fragments	BH-SNE (s)	Hjorth's descriptors (s)
500	70,113	1,903.7	40.5
1,000	35,058	944.8	26.5
5,000	7,017	1,110.2	15.7

4. CONCLUSION

A new technique for visualization of metagenomic data without prior alignment to a reference database has been introduced in this paper. The method first transforms the short genomic fragments into numerical representation in form of phase of complex number and then three Hjorth's descriptors are extracted from each of the signals. The result is visualization of whole metagenomic dataset in three dimensional space where genomic fragments from the same organism cluster together. Unlike the previous technique, the introduced method is non-stochastic, therefore is reproducible. Furthermore the Hjorth's descriptors are computed from variances of signal and its first and second derivatives, which results in linear computational complexity allowing almost immediate analysis and on standard desktop or laptop computer.

REFERENCES

[1] Laczny, C. C., Pinel, N., Vlassis, N. and Wilmes, P., 2014, Alignment-free Visualization of Metagenomic Data by Nonlinear Dimension Reduction. Sci. Rep.. 2014. Vol. 4. DOI 10.1038/srep04516. Nature Publishing Group

[2] Cristea, P. D., 2003, Large scale features in DNA genomic signals. Signal Processing. 2003. Vol. 83, no. 4, p. 871-888. DOI 10.1016/s0165-1684(02)00477-2. Elsevier BV

[3] Oh, S.-H., Lee, Y.-R. and Kim, H.-N., 2014, A Novel EEG Feature Extraction Method Using Hjorth Parameter. *IJEEE*. 2014. P. 106-110. DOI 10.12720/ijeee.2.2.106-110. EJournal Publishing

ELECTROOCULOGRAPHY IN MARKETING

František Fajmon

Master Degree Programme (1), FEEC BUT

E-mail: xfajmo01@stud.feec.vutbr.cz

Supervised by: Oto Janoušek -

E-mail: janouseko@feec.vutbr.cz

Abstract: This article delas EOG analysis for marketing studies. Evaluation is performed using three methods. Data was recorded from 10 respondents. Proposed algorithm is realized in Matlab workspace.

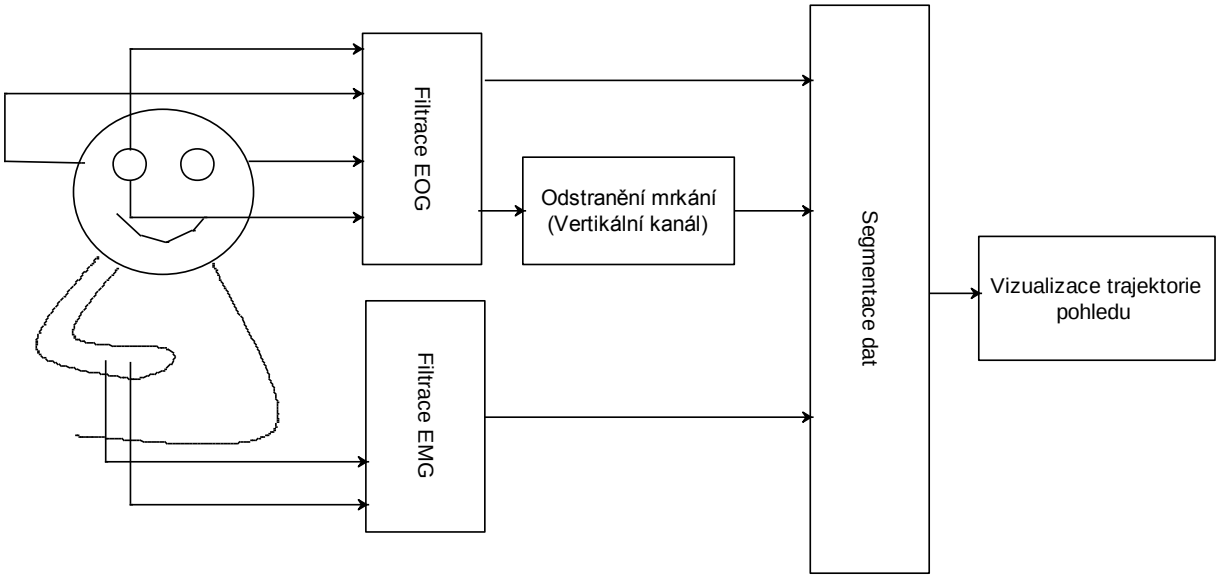
Keywords: EOG, marketing, blink detection

1. ÚVOD

Elektrookulografie je základní neinvazivní metoda pro měření corneo – retinálního potenciálu. Napěťová úroveň povrchovými elektrodami snímaného potenciálu se pohybuje kolem 1mV. Uplatnění této metody je možné nalézt v marketingu, kde je možné vyhodnocovat směr pohledu a pohybu oka po vizuální předloze a zobrazovat tak oblasti zájmu. V marketingových studiích je nutno znát kam a jak dlouho, popřípadě v jakém pořadí se zákazník na reklamní materiál dívá.

2. ZPRACOVÁNÍ SIGNÁLU A KALIBRACE

Zpracování signálu popisuje blokové schéma vyobrazené na obr. 1.

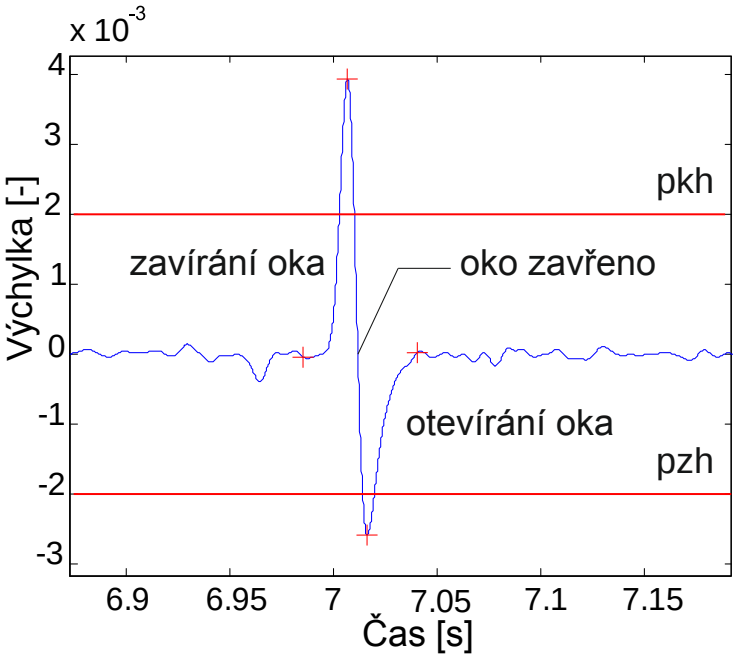


Obrázek 1: Blokové schéma zpracování signálu

Kanály horizontálního a vertikálního elektrookulogramu (EOG) zajišťují snímání polohy oka v rovině x a y. Kanál (elektromyografu) EMG slouží k synchronizaci EOG s jednotlivými snímky. EOG signály jsou filtrovány filtrem typu dolní propust s mezní frekvencí 30Hz. EMG signál zde slouží k synchronizaci EOG dat k jednotlivým snímkům, jehož výchylka stoupne v okamžiku, kdy respon-

dent stiskne šipku na klávesnici a posune tak pozorovanou obrazovou sekvenci o snímek dál. EMG signál je filtrován horní propustí o mezním kmitočtu 120Hz a následně derivován a umocněn. Tímto způsobem jsou získány špičky EMG, jejichž poloha koresponduje s posunem na následující snímek. Na obr. 2 je znázorněn algoritmus detekce mrknutí. Detekce mrknutí byla provedena pouze ve vertikálním kanálu. V signálu byla nalezena lokální maxima a minima, ze kterých byla vypočítána rychlost. Pokud rychlost odpovídala rychlosti mrknutí, algoritmus nalezl okrajové body artefaktu, které zaznamenal do paměti pro pozdější vymazání mrknutí v originálním signálu.

Kalibrace probíhala tím způsobem, že respondent musel sledovat fixační kříž, který se pohyboval po krajích monitoru. Z naměřených hodnot byla získána maxima a minima pro každý kanál a každou pozici. Následnou normalizací obrazových dat na tyto hodnoty bylo dosaženo kalibrace EOG na daný zobrazovací prvek.



Obrázek 2: Detekce mrknutí

2.1. ALGORITMUS HODNOCENÍ REKLAMNÍCH MATERIÁLŮ

Dle postupu v kap. 2 byly získány upravené EOG signály. Tyto signály jsou souřadnicemi pohledu oka po nazírané předloze. Aby bylo možné vyhodnocovat trajektorii pohybu po předloze, bylo nutné oba dva signály přepočítat na rozměr předlohy. Předlohy byly zobrazovány na monitoru s nativním rozlišením 1080 x 1920. Monitor byl vzhledem ke grafickému návrhu reklamních materiálů orientován na výšku. Pro zobrazení pomocí bodového grafu byla souřadnicová data „vzorkována“ 3 ms intervalem. Jednalo se pouze o to, aby výsledná trajektorie byla přehlednější. V případě kruhové grafu byla data přepočítána do matice o shodných rozměrech jako obraz, tím pádem jedna buňka matice reprezentuje jeden pixel v obraze. Pokud subjekt setrval na bodech zájmu delší dobu, byla inkrementována hodnota v matici. Následně byla použita konvoluční matice o rozměrech 5 x 5, střed matice odpovídal středu kruhu a hodnota matice jeho poloměru. Výběr středů probíhal na základě 3 ms intervalu. Pomocí výše uvedených algoritmů je možné sledovat trajektorii a časové setrvání pohledu oka od jednoho uživatele. Tyto metody nejsou vhodné pro srovnávací analýzy většího počtu respondentů. V případě teplotní mapy je možné hodnotit data i od více než jednoho respondenta, avšak pouze délku setrvání pohledu nikoli trajektorii. Výpočet teplotní mapy proběhl obdobně jako v případě kruhového grafu s tím rozdílem, že byla použita konvoluční maska 100 x 100 a nebyl použit 3 ms interval. Výsledek byl zobrazen jako pseudobarevná škála na šedotónový obraz reklamního materiálu. Porovnání metod je možné vidět na Obr.3.



Obrázek 3: Ukázka metod (vlevo bodový graf, uprostřed kruhový graf, vpravo teplotní mapa)

3. ZÁVĚR

V práci bylo pomocí vizualizačních metod vyhodnoceno deset dobrovolníků rovnoměrně rozdělených na muže a ženy. Ženy mají zájem o nasbírání co největšího množství informací o dané problematice, na rozdíl od mužů, kteří reklamu rychle zhodnotí pohledem a v místech zájmu, stráví více času. Pokud se v reklamním materiálu vyskytují slova, která upozorňují na slevu či akci, jsou vnímána více než slova *zdarma*. Tento jev může být způsoben tím, že zákazníci jsou k upoutávkám typu *zdarma* nedůvěřivý – zdarma nikdy nic není. Dalším společným znakem je fixování pohledů v místech postav a obličejů. Pokud se v reklamním materiálu objeví tvar naznačující šipku je její směr pohledem následován. Dalším společným bodem je poslední místo fixace. Jestliže reklamní materiál obsahuje na konci stránky kontaktní informace, prohlížení v těchto místech v 80 % končí. Pokud reklama neobsahuje na konci kontaktní informace, zákazníci se vracejí zpět na bod svého zájmu a opouštějí reklamní materiál. Doporučení pro atraktivní reklamu jsou následující. Reklama by měla mít atraktivní logo, který by ale nemělo příliš dominovat, kontrastní prvky jsou vítány. Každý výrobek i služby by měli být doplněny obrázkem, který bude dokonale ilustrovat danou problematiku. Pokud je potřeba zákazníka někam odlákat pohledem, může být použito šipky, obličeje a dalších prvků, které vizuálně odkazují na daný prvek. Kontaktní informace nebo doplňkové informace se hodí doplnit do „patičky“ reklamy.

REFERENCE (ANGLICKY = REFERENCES)

[1] CHENNAMMA, Chennamma a Xiaohui YUAN. *A SURVEY ON EYE-GAZE TRACKING TECHNIQUES* [online]. Indian Journal of Computer Science and Engineering, 2013 [cit. 2014-10-25]. ISBN 0976-5166. Dostupné z: http://arxiv.org/ftp/arxiv/papers/1312/1312.6410.pdf. Studie. University of North Texas, Denton, Texas, USA.

PULSE VISUALIZATION IN VIDEOSEQUENCE OF FACE

Pavel Bernátek

Master Degree Programme (2), FEEC BUT

E-mail: xberna11@stud.feec.vutbr.cz

Supervised by: Oto Janoušek

E-mail: janouseko@feec.vutbr.cz

Abstract: Eulerian video magnification method may be used for visualization of the pulse in the videosequence of face. This article describes its utilization with combination of Viola-Jones face detection algorithm and Kanade-Lucas-Tomasi algorithm for tracking faces in the videosequence.

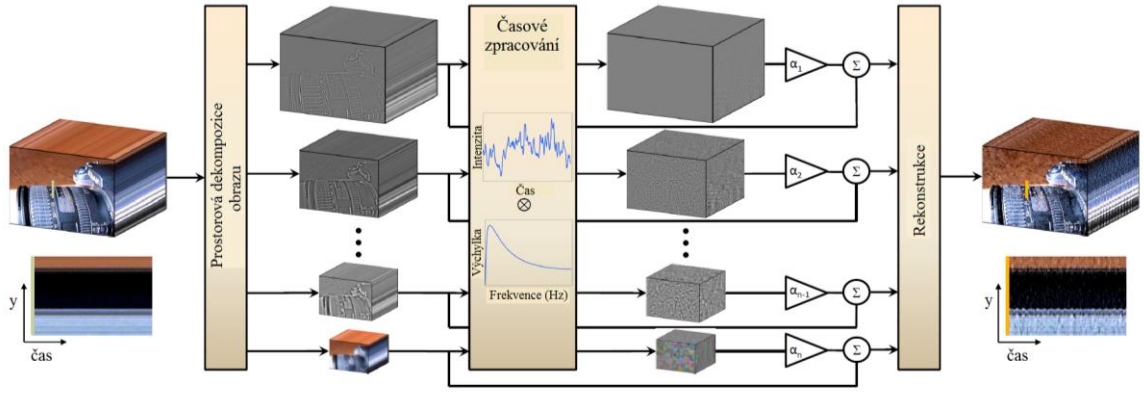
Keywords: Non-contact pulse measurement method, Eulerian video magnification, Viola-Jones algorithm, Kanade-Lucas-Tomasi algorithm, Matlab.

1. ÚVOD

Bezkontaktní detekce srdeční frekvence, která by mohla zvýšit komfort neinvazivního monitorování, spočívá v monitorování obličeje pacienta videokamerou. Získání hodnot srdečního tepu je podmíněno volbou vhodných metod pro analýzu videozáznamu.

2. EULERIAN VIDEO MAGNIFICATION

Metoda Eulerian video magnification se využívá k vizualizaci barevné variace scény a zvětšení jemných pohybů v obraze. Výsledný signál v dané prostorové frekvenci je pak zesílen, díky čemuž je pozorovatel schopen odhalit okem neviditelné informace ve videozáznamu.



Obrázek 1: Metoda Eulerian Video Magnification [1]

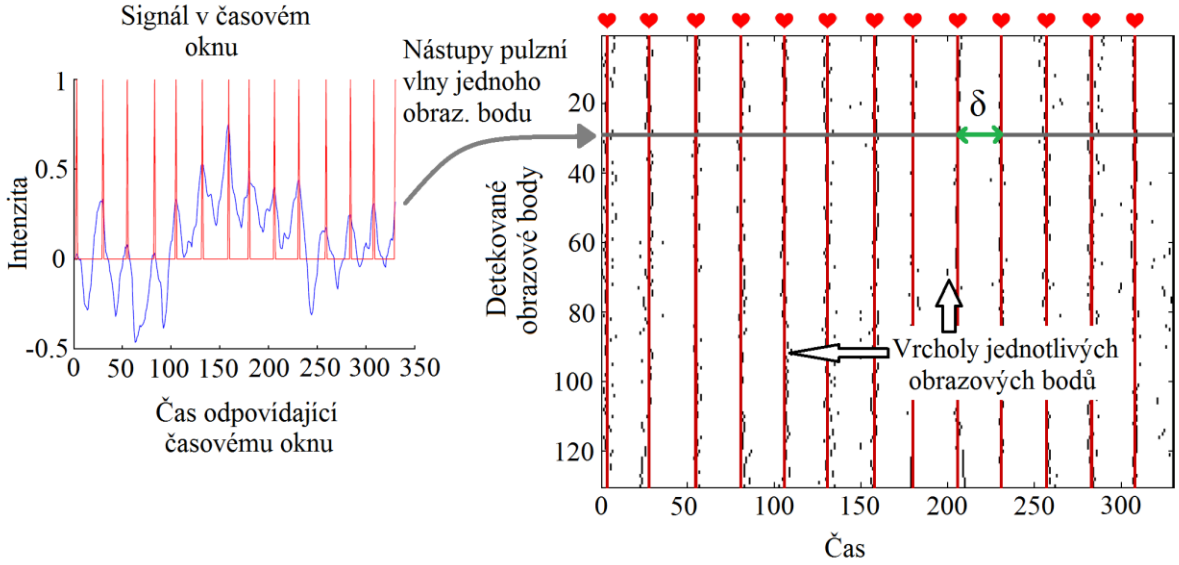
Metoda Eulerian video magnification zpracovává standardní videosekvenci, na níž aplikuje prostorovou filtraci, následovanou časovou filtrací. Systém nejprve rozkládá vstupní video na sekvenci snímků do různých prostorových frekvenčních pásem s využitím Gaussovy pyramidy. V této práci se signál rozkládá na 5 úrovní. Následná filtrace pásmovou propustí je jednotná pro všechny dekompozice, a pro všechny obrazové body v každé úrovni dekompozice. Hodnoty pásmové propusti aplikované v této práci mají hodnoty v rozmezí od 0.83 Hz do 1 Hz. Hodnoty pásmové propusti se však mohou lišit na základě aktuálního pulzu pacienta. Trend signálu je potřeba odstranit filtrací horní propustí. Poté se dle požadavků experimentátora násobí extrahovaný signál amplifikačním

faktorem *alfa*. Každé frekvenční pásmo může být individuálně zesíleno. V této práci se hodnota amplifikačního faktoru *alfa* pro frekvenční pásmo zájmu nastavovala na hodnotu 100. Šum, rušivé artefakty způsobené například pohybem obličeje či kamery a hodnoty prostorových frekvencí, ve kterých se kýžený signál určený k vizualizaci nenachází se pak násobí faktorem *alfa* menším jak jedna. Amplifikovaný signál se přičte k originálu a rekonstrukcí prostorové Gaussovy pyramidy se získá finální výstup. [1] Princip metody shrnuje obrázek 1.

2.1. EXTRAKCE SRDEČNÍHO PULZU

Cílem této práce je odhalit změny barvy kůže ve videu, které jsou obtížné nebo nemožné vidět pouhým okem a vizualizovat je. Základní přístup je zaměřen na obrazový bod a jeho barevné hodnoty měnící se v čase a zesílení kolísání barevné hodnoty v daném frekvenčním pásmu.

Výchylka signálu jednoho obrazového bodu může být závislá na hustotě periferních cév ve scéně. Popisovanou metodou lze zachytit pulz na více místech na tváři v nezpracované úrovni pyramidy. V úzkopásmovém signálu se najdou maxima signálu v rámci lokálního časového okna.



Obrázek 2: Detekce pulzních nástupů [1]

V levé části obrázku 2 je znázorněn úzkopásmový signál s jeho detekovanými pozicemi pulzních vrcholů z jednoho bodu obličeje ve videu v časovém okně. Velikost časového okna může být zvolena v závislosti na detekované frekvenci. Bylo zjištěno, že časové okno 10-15 snímků pro video o třiceti snímcích za sekundu dosahuje dobrých výsledků [1]. Pravá část obrázku 2 ilustruje jednotlivé pulzní nástupy zobrazené v časové ose. Každý řádek na ose obrazových bodů odpovídá jinému bodu na tváři. Černé obrazové body představují zjištěné polohy píků v každém bodě, a červené svislé čáry odpovídají mediánům pulzních nástupů srdce, které jsou použity pro výpočet srdeční frekvence. Extrémní odchylky větší než očekávaná chybovost (vyjádřená parametrem *delta* a znázorněná zelenou šipkou) lokálních pulzních nástupů v každém měřeném obrazovém bodě byly odstraněny. Lze si všimnout, že i když odhad pulzu v některých místech na obličeji může zanést nepřesnost do výsledků, odhad využívající kombinaci více bodů poskytuje robustní odhad pulzního signálu. [1]

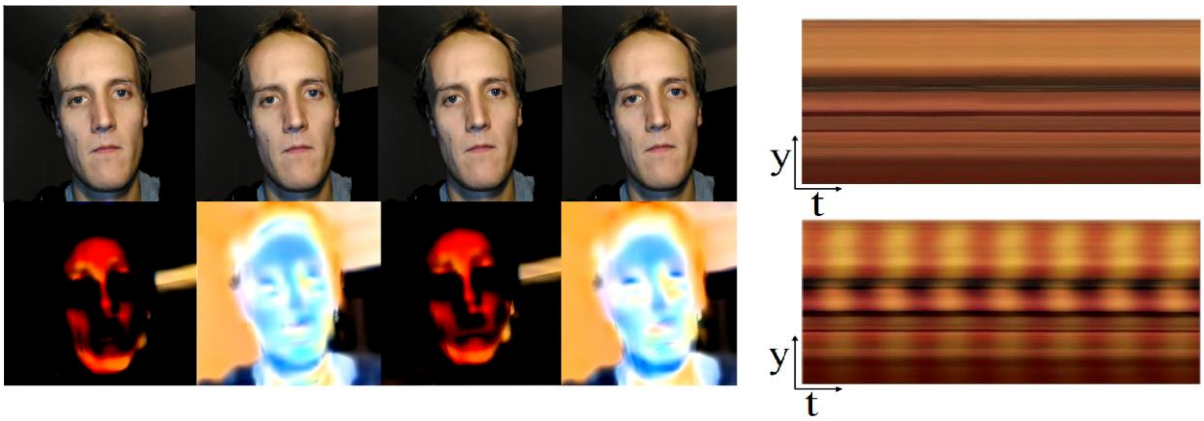
3. DETEKCE A SLEDOVÁNÍ OBLIČEJE

Pro robustnější analýzu videosekvence pomocí uvedené metody bylo nutné detekovat obličej figuranta. Detekce obličeje v obraze je realizována pomocí algoritmu Viola-Jones. Detektor pracuje s obrazy ve stupních šedi. Algoritmus tohoto detektoru vytvoří integrální obraz, následně něj aplikuje Haarovy příznaky a nakonec tyto příznaky sdružuje pomocí metody AdaBoost.

Za předpokladu, že lidská tvář byla v sekvenci obrazů lokalizována, pokračuje se ve sledování obličeje pomocí Kanade-Lucas-Tomasi algoritmu. Tato metoda je postavena na výběru znaků založených na texturách. Algoritmus počítá součet čtverců intenzity rozdílů mezi znaky v předchozím obrázku a znaky v aktuálním snímku. Posunutí specifických znaků je potom interpretováno jako posunutí obličeje. Algoritmus umožňuje kontinuální sledování obličeje. [2]

4. VÝSLEDEK

Videozáznam obličeje zpracovaný pomocí navržené metody barevně pulzoval. Levá horní část obrázku 3 znázorňuje časovou sekvenci snímků kamery snímající figuranta. V levé dolní části je identická sekvence snímků zpracovaná pomocí programu realizující navrženou metodu, která demonstruje pulzující obraz videosekvence podle srdeční frekvence figuranta. Podle hodnoty frekvence změny barevné složky v obraze lze konstatovat, že tato frekvence odpovídá srdeční tepové frekvenci. V pravé horní části obrázku 3 je znázorněn spektrogram nezpracované videosekvence snímající figuranta. V pravé dolní části obrázku 3 je spektrogram videosekvence zpracované pomocí programu, na kterém je vidět jasná pulzující složka ilustrující srdeční pulzaci figuranta.



Obrázek 3: Videozáznam obličeje zpracovaný metodou Eulerian video magnification

5. ZÁVĚR

V článku je popsána metoda Eulerian video magnification, která se využívá pro vizualizaci srdečního pulzu v obličeji. Vysvětlen zde byl princip zpracování videosekvence. Metoda je unikátní ve snímání více obrazových bodů naráz. Dosažené výsledky naznačují kvalitní detekci srdeční pulzace. Metoda může nalézt uplatnění při neinvazivním snímání tepové frekvence na lékařských pracovištích.

REFERENCE

[1] RUBENSTEIN, Michael. *Analysis and visualization of temporal variations in video*. Massachusetts Institute of Technology, 2014. Disertační práce. Massachusetts Institute of Technology. Vedoucí práce William T. Freeman.

[2] WAGENER, Dirk W. a Ben HERBST. *Face Tracking: An implementation of the Kanade-Lucas-Tomasi Tracking algorithm*. Matieland 7602, South Africa

THE SIMULATION OF EFFECT OF MR PULSE SEQUENCES ON THE MEASUREMENT OBJECT TEMPERATURE

Markéta Kosková

Master Degree Programme (2), FEEC BUT

E-mail: xkosko00@stud.feec.vutbr.cz

Supervised by: Lucie Grossová

E-mail: xgross03@ stud.feec.vutbr.cz

Abstract: This paper deals with a heating tissues problem of the measured object in magnetic resonance. This is the consequence of using MRI pulse sequences. At the beginning of this paper is introduction to this topic and the ways of gathering the data necessary for the further processing. In experimental part of this paper the program for object heating simulation in MRI is proposed. This should help with planning the sequences so the researcher does not have to wait for the object to cool down, and also create the better conditions for the experimental animals to survive the process.

Keywords: nuclear magnetic resonance, tissue heating, graphical user interface

1. ÚVOD

Nukleární magnetická rezonance je fyzikálně – chemická metoda založená na analýze magnetických momentů atomových jader. Jedná se o původně analytickou metodu, která byla rozvinuta v zobrazovací techniku používanou k tomografickému zobrazování vnitřních orgánů lidského těla [1]. Přestože je tato metoda z biologického hlediska považována za metodu bezpečnou, existují významná rizika s ní spojená. Tato práce se zabývá právě jedním z rizik a to zahříváním tkání v průběhu měření, které je způsobeno absorpcí energie radiofrekvenčního pole. Právě zmíněné pole vzniká jako důsledek aplikace pulzních sekvencí, pomocí kterých dochází k excitaci atomových jader v měřeném objemu tkáně a následně k zisku obrazové informace [2].

Celá tato práce vznikla primárně pro účely ÚPT AV ČR v Brně z důvodu kolísání teploty myši v průběhu experimentálního výzkumu. Cílem práce bylo provedení série měření pro otestování vlivu různých akvizicí pulzních sekvencí na teplotu měřeného fantomu a následná aplikace získaných poznatků na živý objekt. Data z proměřování fantomu byla dále využita pro vytvoření programu k simulaci zvyšování teploty po aplikaci vybraných akvizicí pulzních sekvencí při plánování MR experimentů.

2. ZÍSKÁVÁNÍ DAT

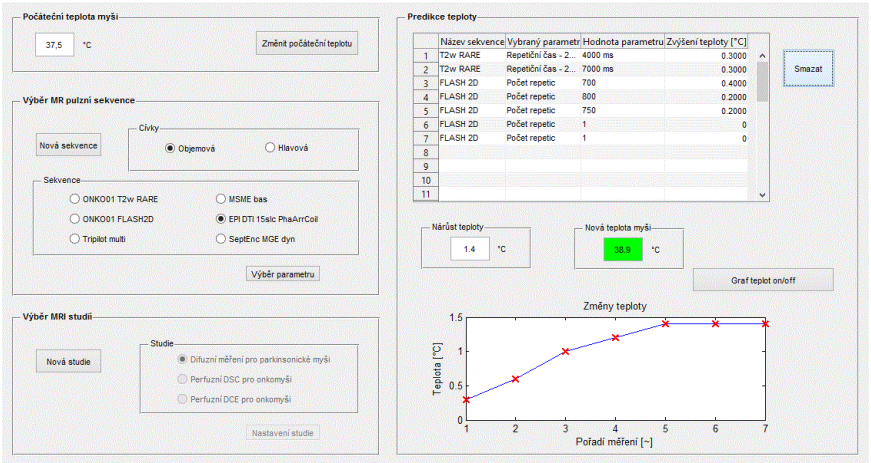
Aby mohl simulační program správně fungovat, bylo nutné nejprve získat potřebná data. Pro tyto účely byl tedy vytvořen testovací fantom, který se skládal z plastové lahvičky naplněné agarovou želatinou. Rozměry lahvičky byly voleny 10 cm na výšku a 3 cm na šířku, aby co nejvíce odpovídaly tělu myši, na které bylo prováděno následné srovnávání výsledků. Spolu s tím, že měl fantom napodobit myš rozměrem, musel také kvůli distribuci teploty co nejlépe simulovat vlastnosti živé tkáně. Toho bylo docíleno pomocí agarové želatiny. Pro zaznamenávání teploty fantomu v průběhu měření magnetickou rezonancí byl použit vláknově optický senzor Fiber Optic Temperature Sensor od firmy SA Instruments, Inc. s rektální sondou. Jedná se o teplotní sondu, která využívá fluorescenční technologie pro přesné a spolehlivé měření teploty v rozsahu 20 – 60 °C s přesností ± 0,2 °C a časovou odezvou 300 ms.

Celé měření proběhlo na 9,4T MR přístroji Bruker Biospec 94/30 USR na ÚTP AV ČR v Brně. Pro získání dat byly zvoleny čtyři pulzní sekvence: Rapid Acquisition with Relaxation Enhancement (RARE), Fast Low Angle Shot (FLASH), Echo Planar Imaging (EPI) a Multiple Gradient Echo (MGE). Vzhledem k tomu, že výsledky práce jsou pro účely ÚPT, vycházelo se z předem daného nastavení těchto sekvencí, jaké se na ÚPT běžně využívá k experimentálním studiím. Navíc se celé měření uskutečnilo dvakrát – pro hlavovou přijímací cívku a pro objemovou cívku.

Aby se projevil vždy samotný parametr (bez vlivu těch ostatních), zvětšovala či zmenšovala se vždy hodnota pouze jednoho konkrétního parametru a to tak, aby byla data získána vždy pro zvolený rozsah tohoto parametru. I když pro každou sekvenci byly zvoleny různé parametry, v této práci jsou prezentovány pouze ty, které jsou společné pro všechny sekvence. Jedná se tedy o délku repetičního času při jednom řezu i při 25 řezech fantomem, o velikost matice obrazu, o rozměry Field of view (FOV) a o počet řezů fantomem.

3. SIMULAČNÍ PROGRAM

Simulační program vychází z experimentálně naměřených dat na 9,4T MR přístroji Bruker Biospec 94/30 USR na ÚTP AV ČR v Brně. Základní premisou tohoto programu je, že nárůst teploty není závislý na počáteční teplotě. Tedy že při totožném nastavení pulzní sekvence a různé počáteční teplotě dojde vždy k přibližně stejnému ohřevu měřeného objektu. Princip programu spočívá v tom, že všechny MR sekvence a k nim změřené rozsahy parametrů jsou do programu uloženy spolu s korespondujícími reálně naměřenými nárůsty teploty. Tedy o kolik narostla teplota během experimentu, tak o tolik se zvýší teplota v programu u dané akvizice. Při plánování pomyslného experimentu potom program sečítá nárůsty teploty volených MR akvizicí a průběžná navýšení teploty pro přehlednost ukládá do grafu. Pro jednodušší používání bylo v programu MATLAB vytvořeno grafické uživatelské rozhraní, jehož úvodní stranu ukazuje **Obrázek 1**.

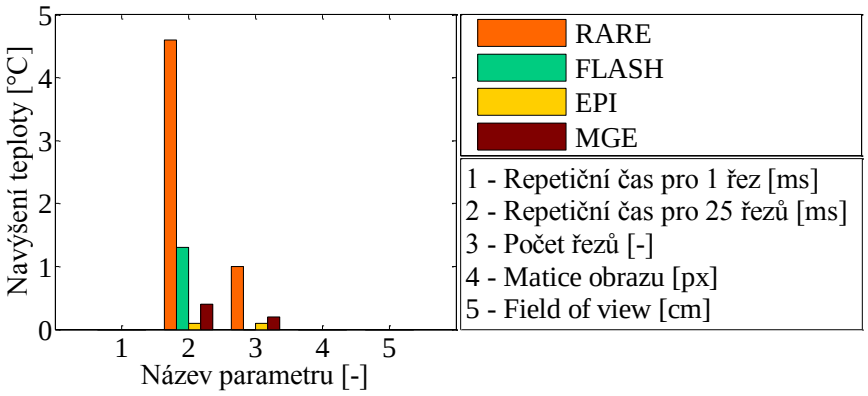


Obrázek 1: Grafické uživatelské rozhraní simulačního programu.

Po celou dobu simulace je možné změnit počáteční hodnotu teploty zkoumaného objektu (laboratorní myši). Tato změna je pak ihned implementována do celkových výpočtů. Podle výběru typu cívky a MR pulzní sekvence se následně volají jednotlivá grafická rozhraní s adekvátní databází naměřených dat. V těchto grafických rozhraních jsou dále zobrazeny jednotlivé parametry sekvence a jejich hodnoty. V závislosti na výběru jsou zvolené MR akvizice přidány do celkového experimentu a současně k nim naměřené hodnoty teploty vykresleny do přehledného grafu. V programu jsou zahrnuty i předpřipravené experimentální studie, které je možnost defaultně použít nebo je případně i upravovat dle potřeby. Veškerá vybraná data jsou ukládána do tabulky, kde jsou zobrazeny i jednotlivé popisné parametry vybraných dat. V průběhu celé simulace je implementován kontrolní algoritmus, který ve chvíli překročení kritické teploty myši upozorní na tuto situaci dialogovým oknem a barevnou indikací. Po celou dobu simulace je možnost kteroukoliv dříve vybranou hodnotu z tabulky odstranit.

4. VYHODNOCENÍ VÝSLEDKŮ

Vyhodnocení měření bylo provedeno spočítáním celkového nárůstu teploty při měření celého rozmezí jednoho zvoleného parametru. Příklad vyhodnocení výsledků je proveden z dat získaných při měření s hlavovou přijímací cívkou pro všechny parametry společné všem sekvencím, což ukazuje **Obrázek 2**. Jak je možné pozorovat, repetiční čas při jednom řezu, matice obrazu a FOV neměly na zvýšení teploty v žádném z měření vliv. Zbylé dva parametry, repetiční čas pro 25 řezů fantomem a počet řezů fantomem, naopak teplotu testovacího fantomu zvyšovaly. Podle teoretických předpokladů se potvrdilo, že největší vliv na zahřívání měřeného objektu měla sekvence RARE, která dokázala teplotu zvýšit o celých 4,6 °C.



Obrázek 2: Vyhodnocení naměřených dat s hlavovou přijímací cívkou.

Ověřování výsledků na laboratorní myši bylo provedeno dvakrát. Jednou byla myš v průběhu měření vyhřívána teplým vzduchem, v druhém případě cirkulující teplou kapalinou. V obou případech byly voleny takové akvizice pulzních sekvencí, které přispěly k největšímu nárůstu teploty během měření na fantomu. U sekvence RARE to byl repetiční čas při 25 řezech fantomem (4,6 °C), počet řezů (1 °C), tloušťka řezu (4,8 °C) a RARE faktor při 25 řezech fantomem (2,8 °C). Pro sekvenci FLASH to byl počet repetic (2,1 °C), sklápěcí úhel při 800 repeticích (1,9 °C) a opět repetiční čas při 25 řezech fantomem (1,3 °C). Při zahřívání myši teplým vzduchem docházelo k neustálému kolísání teploty myši, proto bylo provedeno druhé měření při zahřívání cirkulující kapalinou, kde se navyšování teploty myši přibližně rovnalo navyšování teploty fantomu.

5. ZÁVĚR

Tato práce prezentuje výsledky zaznamenávání teploty, kterých bylo dosaženo v průběhu MR měření na 9,4T MR přístroji Bruker Biospec 94/30 USR na ÚTP AV ČR v Brně. Výsledky jsou důležité především z biologického hlediska, protože díky nim bude experimentátor více zvažovat, jak za sebe lépe seřadit pulzní sekvence v průběhu experimentálních měření, aby se laboratorní zvíře nepřehřívalo a nemuselo se během měření ochlazovat. Navržený program tedy usnadní práci se seřazením sekvencí za sebe bez rizika tepelného ohrožení zvířete. Ze získaných výsledků je zřejmé, že ze sekvencí je pro zvíře nejnebezpečnější sekvence RARE, která se v experimentech běžně používá pro úvodní anatomické sekvence.

REFERENCE

[1] BROWN, Mark A a Richard C SEMELKA. *MRI: basic principles and applications*. 3rd ed. Hoboken, N.J.: Wiley-Liss, 2003, xiv, 265 p. ISBN 0471433101.Orság, F.: Vision für die Zukunft. Biometrie, Kreutztal, DE, b-Quadrat, 2004, s. 131-145, ISBN 3-933609-02-X

[2] FRANK G. SHELLOCK a JOHN V. CRUES. MR Procedures: Biologic Effects, Safety, and Patient Care1. *Radiology* [online]. 2004, **232**(3): 635-652 [cit. 2015-10-27]. DOI: 10.1148/radiol.2323030830.

Magisterské projekty

Komunikační technologie

DATA FLOW CONTROL IN AUTONOMOUS SYSTEMS

Marek Důbrava
Master Degree Programme (2), FEEC BUT
E-mail: Marek.Dubrava@vutbr.cz

Supervised by: Jaroslav Koton
E-mail: koton@feec.vutbr.cz

Abstract: The paper deals with the management of packet traffic over IP networks at the level of the autonomous system. A new real time mechanism for active network traffic control depending on the state of the network load is described. The new mechanism allows us to differentiate network traffic depending on the user and manage it according to the defined parameters.

Keywords: data flow, overload, trafic managment, QoS, shaping, HTB

1 ÚVOD

Telekomunikace jsou v současnosti nedílnou součástí společnosti. Požadavky na častější komunikaci s trendem přenášení multimédií kladou větší nároky na přenosové parametry sítě. Pokud síť takové požadavky není schopna zpracovat či obsloužit, dostává se do stavu přetížení. Tento stav sebou přináší až razantní zhoršení přenosových parametrů, které jsou neakceptovatelné [1]. Možným řešením je výstavba nových vysokorychlostních sítí, jejichž realizace je však nákladná. Proto jsou vyvíjeny metody, které mají za cíl využitím stávající infrastruktury zvýšit objem přenášených užitečných dat. V článku je představena nová metoda řízení datového provozu, která dále doplňuje skupinu již známých mechanizmů.

2 OCHRANY PROTI PŘETÍŽENÍ SÍTĚ

Stav přetížení není akceptovatelný, proto musí být síť realizována tak, aby se stavu přetížení předcházelo. Síťové prvky mohou podle různých normovaných metod informovat generátory provozu v síti o svém stavu zahlcení, aby omezily množství vysílaných dat. Mezi tyto metody patří Zpětný tlak, Tlumící paket, zahození paketu, ECN a další [1]. Hlavní nevýhodou těchto metod je omezená možnost řízení přetížení. Tyto metody nejsou schopny rozlišovat uživatele sítě a zacházet s provozem těchto uživatelů odlišně. Metody nedokážou zaručit parametry sítě tak, aby se účastníci sítě navzájem neovlivňovali. To znamená, že jeden uživatel dokáže při těchto metodách zhoršit parametry sítě ostatním uživatelům.

Na tvarování provozu a jeho rovnoměrné rozdělení jsou již představeny algoritmy jako Class Based Queueing (`CBQ`), Hierarchy Token Bucket (`HTB`), nebo další proprietární protokoly firem vyvíjející software pro síťové prvky. Tyto algoritmy mohou doplnit nedostatky popsány výše, ovšem při konfiguraci těchto algoritmů musí být předem definovaná přenosová rychlost omezovaného toku [2]. Ovšem existují případy, kdy přenosová rychlost spoje není známa, například při sdílení přenosové kapacity spoje, při jiné než stromové topologii sítě, při poruše části sítě, při zhoršení parametrů bezdrátového spoje z důvodů deště a při dalších událostech v síti. V těchto případech ztrácí algoritmy svou funkci.

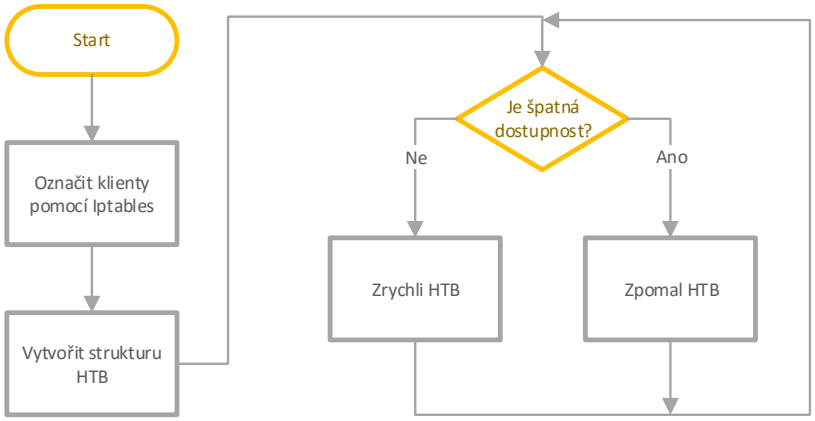
3 VLASTNÍ METODA ŘÍZENÍ PROVOZU V SÍTI

Do sítě jsou připojeny různé typy uživatelů, někteří mohou být omezeni, jiní naopak nikoliv. Hraniční směrovač dokáže s provozem jednotlivých klientů zacházet podle předem definovaných pravidel a i při vysoké agregaci sítě vyčlení kapacitu pro důležité uživatele. Pro správnou funkci algoritmu musí být síť brána jako určitá oblast obsahující přístupové body (POP - Point Of Presence), ke kterým jsou připojeni klienti. Na hranicích těchto oblastí bude provoz zpracováván a tvarován pomocí výkonných serverů, které dokážou provoz podrobně rozdělit podle uživatele a zacházet s ním tak, jak uživatel dle Servis Level Agreement (SLA) očekává.

Při návrhu nové metody byl kladen důraz na doplnění parametrů, které postrádají algoritmy uvedeny výše. Stavebním kamenem je program `HTB` dnes obsažený téměř ve všech jádrech linuxové distribuce. Nově popsaný algoritmus řeší nedostatky `HTB` a nastavuje rychlost přenosových tras dynamicky podle aktuálního stavu sítě, společně tak tvoří celek, který má schopnost udržet síť v nepřetíženém stavu za jakýchkoliv okolností. Z tohoto důvodu rozšiřuje působnost současných principů tvarování provozu i tam, kde nemohly být nasazeny, nebo kde bylo nasazení zbytečné. Celková síť s vylepšeními samozřejmě může obsahovat dříve popsané algoritmy, nebo další funkce jako je QoS (Quality of Service).

3.1 REALIZACE ALGORITMU

Program úzce komunikuje s jádrem systému, protože používá jeho funkce. Pro značení paketového provozu a rozdělení do skupin pomocí uživatelů byl použit program `Iptables`, který je součástí Linuxové distribuce. Pro zpracování skupin a tvarování provozu byl použit program `HTB`, který je



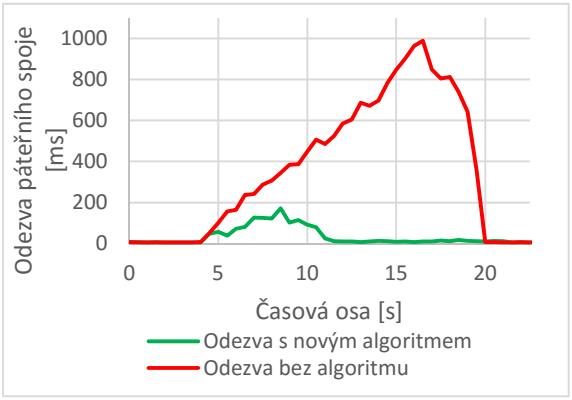
Obrázek 1: Vývojový diagram algoritmu

taktéž součástí jádra Linuxu a nastavuje se přes příkaz `tc`. Tento program dokáže omezovat provoz předem označených skupin, nebo rozdělovat šířku pásma rovnoměrně mezi uživatele sítě.

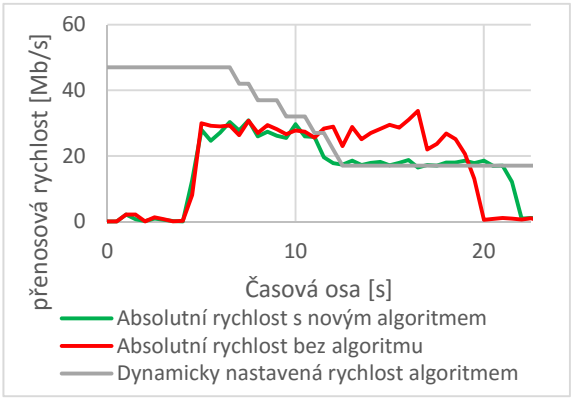
Algoritmus si při startu načte databázi známých uživatelů sítě a parametrů jejich služeb. Obdobně načte seznam míst POP a přiřadí k nim uživatele. Pomocí této topologie označí provoz procházející směrovačem a sestaví strukturu `HTB`, jak je uvedeno na Obr. 1. Právě přes strukturu `HTB` dokáže omezovat maximální rychlosti procházející sítí. Po celkovém sestavení přejde program do nekonečného cyklu, během kterého testuje dostupnost každého POP. Dostupnost je testována průměrně každých 100 ms, kdy tato rychlost testování se ukázala jako dostačující. V případě, že se síť začne přetěžovat, zvýší se odezva, nebo se začnou zahazovat pakety, jak je popsáno v kapitole 2. Tuto situaci algoritmus vyhodnotí jako blížící se stav zahlcení a může virtuálními frontami `HTB` snížit procházející provoz.

Tímto je zaručena dobrá dostupnost celé páteřní sítě v dané oblasti.

3.2 Výsledky reálného nasazení



Obrázek 2: Porovnání odezvy



Obrázek 3: Porovnání přenosové rychlosti

V grafu 2 je zobrazeno porovnání odezvy páteřního spoje bez a s ochranou novým algoritmem. Je patrné, že spoj nezahazuje přenášené data, ale uchovává si je pro pozdější odeslání, to však způsobí vzrůst odezvy sítě až na jednu sekundu. Při tak velké odezvě není možné na páteřním spoji provozovat služby v reálném čase, to znamená, že jeden uživatel měřící rychlost dokáže ovlivnit parametry sítě ostatním uživatelům. V grafu 3 jsou uvedeny přenosové rychlosti. Dva průběhy začínající a končící na nule zobrazují množství přenášených dat páteřním spojem. Nasazení nového algoritmu mělo za následek snížení přenosové rychlosti, pokles způsobil uvolnění kapacity páteřního spoje a snížení odezvy. Šedý průběh je výsledkem nově popsaného algoritmu, je nastavován v reálném čase podle aktuálního stavu sítě a pomocí programu `HTB` omezuje reálný provoz páteřním spojem.

4 Závěr

Nový algoritmus doplňuje důležité parametry pro efektivní řízení propustnosti v komunikačních a datových sítích. V současnosti je tento mechanismus pro potřeby ověření jeho činnosti a testování nasazen v ISP síti Durnet.cz [3]. Navržený mechanismus umožňuje správu datového toku klientů dle jejich tarifu. Nejdůležitější funkcí je však ochrana páteřních tras proti přetížení. Tato funkce dovoluje využít i spoje s relativně nízkou přenosovou kapacitou do maxima.

Reference

[1] KOTON. *Moderní síťové technologie.* Brno, 2014. Skripta. VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ.

[2] *HTB Home* [online]. 2001 [cit. 2016-03-30]. Dostupné z: http://luxik.cdi.cz/ devik/qos/htb/

[3] *Durnet.cz připojení k internetu* [online]. Dobrkovice, 2016 [cit. 2016-03-08]. Dostupné z: http://durnet.cz/

EMULATOR TRANSMISSION PARAMETERS OF DATA NETWORKS

Štěpán Grabovský
Master Degree Programme (2), FEEC BUT
E-mail: stepan.grabovsky@vutbr.cz

Supervised by: Václav Zeman
E-mail: zeman@feec.vutbr.cz

Abstract: This paper deals with a development of a network emulator which is part of a complex ICT tester. This work contains an analysis of concurrent systems for network emulation, both proprietary and open-source. Afterward describes a conception of the emulator and discusses chosen technologies for each part of the emulator software structure.

Keywords: network, emulator, transmission parameters, netem, dummynet

1 ÚVOD

V kontextu této práce je emulátorem přenosových parametrů datových sítí myšlen síťový prvek, který přenáší přenosové podmínky globálních sítí do lokálního prostředí. Tato technika umožňuje testování vyvíjených síťových aplikací a protokolů, ověřování funkčnosti již implementovaných síťových zařízení anebo studium chování datových sítí v podmínkách, které jsou charakteristické pro globální sítě. Emulátor je autonomní síťový prvek skládající se z dedikovaného hardwaru, především výkonných síťových karet a speciálního softwaru, který dokáže pracovat s datovým tokem počítačových sítí. Jelikož je tento projekt součástí vývoje komplexního ICT testeru, je v této práci řešena pouze softwarová část emulátoru, konkrétně rozbor současného stavu problematiky a návrh koncepce zařízení.

2 SOUČASNÉ SYSTÉMY A DOSTUPNÉ TECHNOLOGIE

V současnosti existuje velké množství emulátorů, a to buď v podobě komplexního produktu (veskrze komerční produkty), anebo pouze softwarový program, který je možné získat i jako tzv. open-source.

2.1 KOMERČNÍ ŘEŠENÍ

V úvodní části vývoje byla provedena analýza dostupných komerčních produktů, kterým by měl v této práci vyvíjený emulátor býti konkurencí. Analýza zpoplatněných systémů je z principu povahy těchto produktů obtížnější než u řešení s otevřeným kódem. Pro srovnání byly vybrány produkty 40,000/2 společnosti JAR Technologies, Ltd., dále Netropy 40G společnosti Apposite Technologies, Inc. a poslední Model 20 společnosti iTrinergy Inc.

Rozborem, který se opírá o veřejně dostupné materiály výrobců, bylo zjištěno, že vysoce-výkonové komerční emulátory dokáží pracovat v režimech směrovače i síťového mostu. Zvládají optické i metalické technologie až do rychlostí 40 Gbit/s a dokáží emulovat šířku pásma (bandwidth), konstantní i proměnné zpoždění, ztrátovost, chybovost a také záměnu či duplikaci datových jednotek. Pravděpodobnosti jednotlivých dějů je možné volit na základě definovaných modelů nebo pseudonáhodně. Komerční emulátory dále nabízí aplikaci filtrů na emulovaný provoz a to ve většině případů do čtvrté vrstvy RM ISO/OSI. Navíc jsou vybaveny pokročilým uživatelským rozhraním, jenž zajišťuje ob-

sluhu emulátoru a nabízí zpětnou vazbu uživateli, například pomocí grafů datového toku vykreslovaných v reálném čase.

2.2 OTEVŘENÉ TECHNOLOGIE

U open-source technologií byla pozornost zaměřena na systémy, které jsou kompatibilní s operačním systémem Linux. Po důkladné analýze bylo zjištěno, že všechna existující řešení umožňující emulaci jsou pouze nástroje, které zjednodušují ovládání nástrojů již dostupných v jádře operačního systému. V případě operačních systémů s jádrem Linux se jedná o nástroje balíčku `tc` (Traffic Control), který v linuxových operačních systémech obsluhuje řízení toku datových jednotek jádrem operačního systému. Balíček `tc` v sobě obsahuje tři nástroje: `qdisc` - (queueing discipline) pro nastavení řízení front, `class` - umožňuje tvorbu dědičné struktury řízení front, `filter` - umožňuje aplikovat řízení front pouze na zvolené datové toky. V kontextu emulace přenosových parametrů datových sítí je nejdůležitější řízení front, které dokáže tvarovat síťový datový tok dle přání uživatele.

Nastavení řízení front tak, aby splňovalo daný požadavek, například zpoždění datových jednotek, není triviální záležitostí. Pro zjednodušení nastavení řízení front je v linuxovém jádře nástroj `netem`. Tento nástroj používá například systém pro emulaci WANem [1].

Alternativním přístupem, v praxi také velmi používaným, je nástroj Dummynet. Tento nástroj je vyvíjený primárně pro operační systémy BSD, ovšem je dostupná i verze pro operační systémy s jádrem Linux. Systém Dummynet je postaven na nástroji `ipfw` [2], což je obdoba známého nástroje `iptables`.

První základní rozdíl obou technologií je v konceptu emulace. Zatímco u systému `netem` probíhá tvarování síťového toku pomocí hierarchického spojování řízení front a je realizováno na výstupu z protokolového zásobníku , u Dummynet je využíváno rour a front k abstrakci reálné sítě, kde se každá roura konfiguruje zvlášť a poté dochází k jejich následnému spojování dle aktuálních požadavků uživatele [2]. Další rozdíl mezi systémy lze nalézt v architektuře programů a stylu jejich vývoje. Nástroj `netem` je neustále vyvíjen širokou komunitou, a tak obsahuje mnoho přídavků a funkcionalit ovšem s nižší úrovní dokumentace a ne všechny přídavky jsou validovány. Naopak Dummynet je spravován užší skupinou vývojářů, kde každý nový přídavek je validován a dokumentován a jeho ovládání je snadné a intuitivní.

Na základě analýz publikovaných v [3] v roce 2011 a [4] v roce 2014 lze tvrdit, že nástroj `netem` je schopný dosahovat přesnějších výsledků při emulaci šířky pásma a konstantního i proměnného zpoždění. Navíc je na rozdíl od systému Dummynet schopen duplikace datových jednotek a emulace jejich poškození. Nástroj Dummynet však dokáže abstrakci síťové topologie.

3 STRUKTURA VYVÍJENÉHO EMULÁTORU

Strukturu softwarové části emulátoru lze rozdělit na tři části. První částí je jádro emulátoru, což jsou výše rozebírané nástroje. Druhou částí je API, které zprostředkovává komunikaci mezi uživatelským rozhraním a jádrem emulátoru. Poslední, částí je uživatelské rozhraní, které volá jednoduché služby API a zobrazuje uživateli aktuální stav nastavení emulátoru.

3.1 JÁDRO EMULÁTORU

Jako jádro emulátoru byly zvoleny nástroje balíčku `tc` společně s nástrojem `netem`. Hlavním důvodem této volby je vyšší přesnost emulace a více možností nastavení. Další výraznou výhodou `netem` je jeho snadná rozšiřitelnost a úprava zdrojového kódu. Nevýhodami této volby jsou především náročnější konfigurace, obsluha nástrojů a možnost abstrakce topologie. Tyto nedostatky jsou v této práci

odstraněny pomocí obsáhlého aplikačního rozhraní. Nepřítomnost programu Dummynet v operačním systému CentOS Linux nebylo považováno za nevýhodu. Dummynet lze pro tento typ operačního systému zkompilovat jako externí modul jádra, a poté nainstalovat.

3.2 APLIKAČNÍ ROZHRANÍ

Funkcí části aplikačního rozhraní je převádět jednoduché pokyny z uživatelského rozhraní na složitější příkazy, jejichž pomocí je jádro emulátoru řízeno. Obtížnost nastavení spočívá ve faktu, že je nutné kromě samotného řízení front vytvořit tomuto řízení adekvátního potomka v dědičné hierarchii a aplikovat jej daným filtrem. Pro nastavení i zdánlivě triviálního parametru (například zpoždění) je tedy nutné zavolat minimálně tři samostatné příkazy.

Jako API byla zvolena standardní webová služba. Tento přístup představuje nejobecnější příklad API, jehož hlavní výhodou je absolutní nezávislost klienta na technologii aplikačního rozhraní. Webová služba je realizována v programovacím jazyce Java a provozována na serverové technologii Apache Tomcat.

3.3 UŽIVATELSKÉ ROZHRANÍ

Jako primární uživatelské rozhraní bylo zadáno webové uživatelské rozhraní. Tento typ uživatelského rozhraní usnadňuje vzdálené ovládání bez nutnosti instalace programů, je nezávislé na operačním systému, jeho nastavení atp. Pro výstavbu webového uživatelského rozhraní byly zvoleny technologie, které pracují na straně klienta, tj. HTML, CSS a JavaScript. Hlavní výhodou tohoto přístupu je, že šetří výpočetní zdroje serveru a množství přenesených dat. Veškerá data jsou poslána klientovi pouze jednou a poté jsou již jen aktualizovány změny. Navíc se jedná o tradiční technologie, které mají širokou podporu pro zvolený operační systém a jsou velmi dobře zdokumentovány.

4 ZÁVĚR

V této práci byl popsán vývoj emulátoru přenosových parametrů datových sítí. Konkrétně byla pozornost zaměřena na rozbor schopností komerčních emulátorů a analýzu volně dostupných softwarových emulátorů. Dále byla popsána struktura vyvíjeného emulátoru a diskuse výhod zvolených technologií pro jednotlivé části emulátoru. Celý projekt je nyní ve stavu, kdy probíhá poslední etapa vývoje alfa verze a v nejbližší době bude započato testování softwarové části emulátoru.

REFERENCE

[1] TATA CONSULTANCY SERVICES (TCS). *WANem 2.0: Wide Area Network Emulator* [online]. 2008, 34 s. [cit.2016-01-03]. Dostupné z: http://sourceforge.net/projects/wanem/files/Documents/wanemulator-all-about-v2.0.pdf.

[2] CARBONE, Marta a Luigi RIZZO. *Dummynet Revisited* [online]. Universita‘ di Pisa, Dipartimento di Ingegneria dell’Informazione, 2009, 8 s. [cit. 2005-12-03]. Dostupné z: http://info.iet.unipi.it/ luigi/papers/20091201-dummynet.pdf.

[3] JURGELIONIS, Audrius, Jukka-Pekka LAULAJAINEN, Matti HIRVONEN a Alf Inge WANG. An Empirical Study of NetEm Network Emulation Functionalities. *2011 Proceedings of 20th ICCCN* [online]. IEEE, 2011, : 1-6 [cit. 2016-01-03]. ISBN 978-1-4577-0637-0.

[4] LUBKE, Robert, Peter BUSCHEL, Daniel SCHUSTER a Alexander SCHILL. Measuring accuracy and performance of network emulators. *2014 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)* [online]. IEEE, 2014, : 63-65 [cit. 2016-01-03]. ISBN 978-1-4799-4067-7.

MPT AND MPTCP MECHANISMS IN DATA NETWORKS

Petr Sejkora

Master Degree Programme (2), FEEC BUT

E-mail: xsejko00@stud.feec.vutbr.cz

Supervised by: Jaroslav Koton

E-mail: koton@feec.vutbr.cz

Abstract: This work deals with multipath transmission in the data networks to speed up transmission over networks with limited transmission speed. Work compares the characteristics, efficiency and response to changes in individual transmission paths. Work is specifically dedicated to MPT and MPTCP mechanisms.

Keywords: MPT, MPTCP, multipath data transport

1. ÚVOD

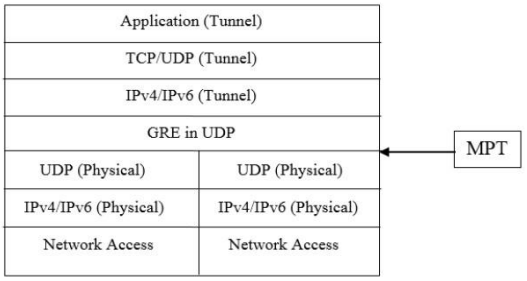
V současné době dochází k připojování nepřeberného množství zařízení k datovým sítím. S tím, jak se zařízení rozšiřují v počtu a schopnostech, se rozšiřuje i nabídka služeb, které jsou skrze zařízení poskytovány. Nároky na kapacitu a spolehlivost datových sítí se tedy zvyšují.

Vývoj probíhá i v rámci datových sítí, které tato zařízení propojují do vyšších celků. Na úrovni koncových zákazníků zpravidla jedna linka stačí. Na úrovni ISP, větších podniků a poskytovatelů služeb už ne. Architektura sítí musí mít také redundanci k zajištění provozu služby i v případě selhání některého prvku. Vznikají tak složitější celky s více cestami.

Pokud máme více cest, přichází na řadu otázka, jak je efektivně využít. Balancování zátěže se věnuje tato práce a klade si za cíl porovnat možnosti, efektivitu a chování vybraných mechanismů sloužících k tomuto účelu.

2. KNIHOVNA MPT

Knihovna vznikla jako projekt pod vedením dr. Béla Almási na Fakultě informatiky University of Debrecen v Maďarsku [1]. Klade si za cíl vytvořit možnost přenosu více cestami na úrovni síťové vrstvy. Využívá IETF RFC s názvem "GRE in UDP" [4], který je zatím ve formě návrhu a rozšiřuje jej o možnost přenosu dat více cestami. Pokud chce aplikace komunikovat s protistranou přes MPT, musí využít předem nakonfigurovaného tunelu s druhou stranou. Odešle požadavek na komunikaci s IP adresou dostupnou na druhém konci tunelu a předá nižší vrstvě. Změna nastává až na linkové vrstvě, kdy se rámec nepředává fyzické vrstvě, ale vstupuje do GRE tunelu. Jeho provoz zajišťuje právě MPT a rozděluje jej do více, až dvaceti, toků datagramů UDP. Ty jsou dále zpracovány standardním způsobem.

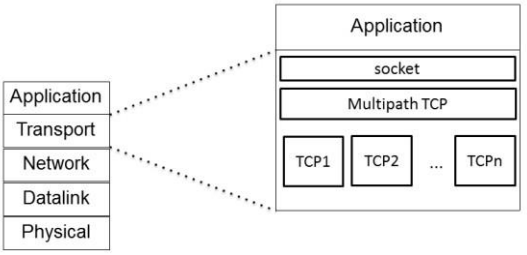


Obrázek 1: Architektura MPT [1]

Aplikace MPT serveru musí být spuštěna na obou stranách tunelu. Může se jednat o koncové stanice nebo o směrovače spojující celé sítě. Konfigurace nabízí obě možnosti. Provádí se formou zápisu dat do souborů, které si MPT server načte při spuštění. Po spuštění je možné posílat příkazy na vzdálený server pomocí klienta. Servery si mezi sebou posílají keepalive zprávy. Knihovna vytváří tunel s IP adresou na obou koncích a dokáže tedy přenášet jak UDP, tak TCP, případně IPv6 skrze IPv4. Vzhledem k tomu, že je tento tunel provozován na UDP protokolu, není zaručeno doručení odeslaných dat. Je tedy obtížné řízení toku dat cestami, které nejsou stejné nebo vykazují odchylky ve spolehlivosti, propustnosti nebo zpoždění. Nelze tedy očekávat bezproblémový přenos. Protokol umožňuje použití vyrovnávací paměti před výstupem z tunelu (pro seřazení paketů, pokud záleží na pořadí), ale když dojde ke ztrátě během přenosu, tak následky neřeší. V případě přenosu TCP se chybějící segment vyšle znovu nad úrovní tunelu, ale UDP datagram je ztracen. V konfiguraci je dále možné nastavit počet a váhu cest, sledování funkčnosti dané cesty nebo autentizaci řídících zpráv [1].

3. MULTIPATH TCP

Vývoj tohoto protokolu začal jako magisterská práce Dr. Sébastiena Barré na Université catholique de Louvain v Belgii [2]. Momentálně vývoj vede Dr. Christoph Paasch. MPTCP se snaží využívat přenosu více cestami pro přenos jednoho TCP spojení za účelem lepšího využití zdrojů, vyšší propustnosti a plynulejší reakce na výpadek [3]. Aplikace vytvoří požadavek na TCP spojení na danou IP adresu. Spojení se navazuje použitím 3-way handshake. Strana, která vytváří spojení, odešle paket s příznakem `SYN` a `MP_CAPABLE`, ve kterém je i šifrovací klíč na pozdější použití. Pokud má protistrana o spojení zájem odešle též `SYN+ACK+MP_CAPABLE` se svým klíčem. Navazující strana potvrdí `ACK` a tím je vytvořena relace MPTCP s prozatím jedním spoje- ním. Pokud má jedna ze stran více adres k dispozici, tak se pokusí navázat s protistranou další spo- jení zprávou s příznakem `MP_JOIN` a identifikátorem, ke které relaci MPTCP se chce připojit. Protokol tedy musí rozlišovat sekvenční čísla jednotlivých subrelací i nadřazené relace a spravovat rozdělení toků. Spojení se tedy navazuje na požádání a veškerá konfigurace se provádí dynamicky dle aktuálních podmínek v síti.



Obrázek 2: Architektura MPTCP [2]

Vývojáři si kladli za cíl schopnost reálného nasazení protokolu, aby mohl fungovat v reálném prostředí stávajících sítí a Internetu. Nasadit novou technologii na všechny prvky Internetu je nemožné a některé směrovače a firewally zahazují provoz, který nedokáží identifikovat. Jako jedinou možnost tedy shledali postavit protokol tak, aby fungoval na již existující architektuře, bez zásahu. Protokol je specifikován v RFC 6824 [3].

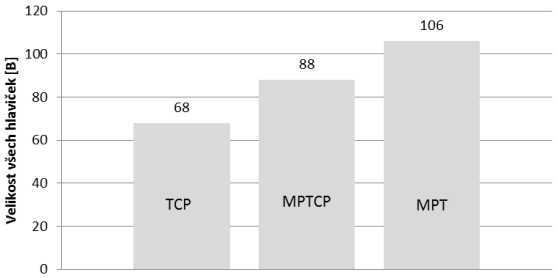
4. TESTOVACÍ SESTAVA

Při testování mechanismů TCP a MPTCP a pro jejich porovnání se základním protokolem TCP byla vyhodnocována jejich režie a schopnost využít více paralelních cest z pohledu celkové dosažené přenosové rychlosti na aplikační vrstvě a času potřebného k přenosu daného objemu dat. Měření probíhalo ve virtuálních strojích `VirtualBox` v OS `LUbuntu`. Hostitelský počítač měl dvě fyzická jádra. Každý stroj měl 1Gb/s síťovou kartu, přes kterou mohl komunikovat s druhým OS. Generování zátěže zajišťoval program `IPerf`, který na dané sestavě byl schopný vygenerovat

TCP provoz až 21,2 Gb/s. Omezení šířky pásma každé cesty na 10 Mb/s zajišťoval třetí počítač se dvěma síťovými kartami a programem `WANem`.

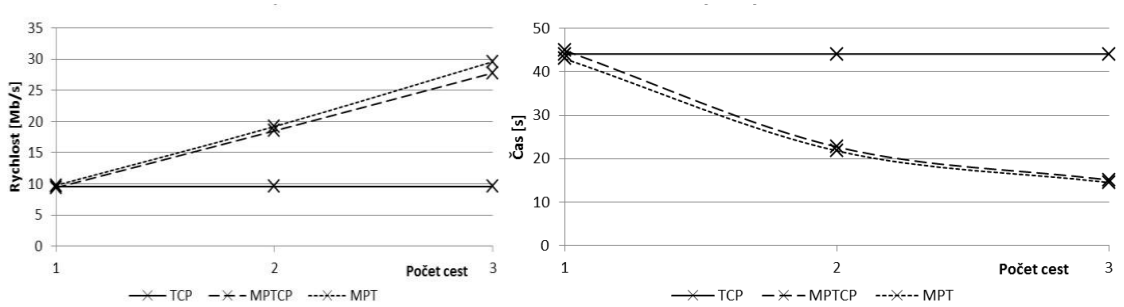
4.1. VELIKOST HLAVIČEK

Protokoly vnáší do přenosu kromě zpoždění a nároků na výkon také další data do hlaviček paketů. Problém nastává hlavně u MPT, kde v našem případě často docházelo k překročení MTU linky, fragmentaci a následné další režii. Proto je nutné věnovat pozornost konfiguraci MTU tunelu.



Obrázek 3: Režie protokolů

4.2. NAMĚŘENÉ VÝSLEDKY



Obrázek 4: (a) Rychlost přenosu (b) Čas potřebný pro přenos 50MB dat

Rychlost rostla s počtem linek téměř lineárně a linky byly vytěžovány na 100%. Režie samotných protokolů co do množství dat je zanedbatelná (viz. Obr. 3). V tomto ohledu se jejich nasazení určitě vyplatí. Jiná situace může nastat, pokud se začnou měnit vlastnosti jednotlivých linek. MPTCP bude na změny reagovat, MPT nikoli. Když byla rychlost linek limitována zahazováním nadbytečných paketů, propustnost protokolu MPT na třech linkách byla ještě nižší než samostatný tok TCP.

5. ZÁVĚR

V článku byly diskutovány mechanismy MPT a MPTCP, které dokáží využít více cest pro spojenímezi stanicemi.

Nutnou podmínkou provozování protokolu MPT je prvotní manuální konfigurace koncových stanic, kdy se tak předpokládá, že definovaná spojení jsou navázána staticky a protokol sám jen omezeně reaguje na dynamiku provozu na definovaných linkách.

Výhody protokolu MPT jsou ve schopnosti přenosu jak UDP, tak TCP spojení přes až 20 různých linek. Je možné ovládat server z druhé strany příkazovým kanálem. Protokol je možné použít i jako spoj mezi sítěmi.

Naopak protokol MPTCP je plně automatizovaný a reaguje dynamicky na změny parametrů sítě. Pracuje na spolehlivém protokolu TCP, který mu umožňuje okamžitě reagovat na změny. Spojení

se navazuje, když je potřeba (on demand) v závislosti na počtu a stavu linek a po uzavření přenášeného spojení se ukončuje.

REFERENCE

[1] B. Almási, The MPT-GRE Multipath Communication library, avaible from http://irh.inf.unideb.hu/~almasi/new/index.php/projektek/19-mpt-library

[2] C. Paasch, S. Barre, et al., Multipath TCP in the Linux Kernel, available from http://www.multipath-tcp.org

[3] RFC6824: TCP Extensions for Multipath Operation with Multiple Addresses, avaible from https://tools.ietf.org/html/rfc6824

[4] RFC "GRE in UDP Encapsulation", avaible from https://tools.ietf.org/html/draft-ietf-tsvwg-gre-in-udp-encap-11

SIMULATION OF STATIC GRANT ALLOCATION IN NG-PON2 NETWORKS

Patrik Cymorek
Master Degree Programme (1), FEEC BUT
E-mail: xcymor01@stud.feec.vutbr.cz

Supervised by: Tomáš Horváth
E-mail: horvath@feec.vutbr.cz

Abstract: This paper deals with issue of static grant allocation. The main goal of this paper is to present functionality of NG-PON2 networks and describe static form of grant allocation. The first part of the paper is focused on development of passive optical networks. Subsequently Next-Generation Passive Optical Networks and simulation in OPNET modeler are described. In modeler own designs of node and process models are used.

Keywords: NG-PON2, OPNET Modeler, Static allocation

1 ÚVOD

Již od dob 70. let minulého století je optické vlákno využíváno jako médium pro přenos dat. Od začátků užívání optických vláken uplynuly již desítky let, během kterých docházelo k nárůstu požadavků na připojení, způsobených rychlým vývojem společnosti.
V rámci vývoje optických sítí byly upřednostňovány sítě pasivního typu z důvodu minimalizace nákladů. Prvním standardem PON (Passive Optical Network) vydaným v roce 1998 společností ITU-T byl APON (ATM PON), následný vývoj zobrazuje tabulka 1 [2].

Tabulka 1: Vlastnosti standardů ITU-T.

	APON	BPON	GPON	XG-PON	NG-PON2
Sym. rychlost [Gbit/s]	0,155	0,622	2,5	2,5	10
Asym. rychlosti [Gbit/s] (Sestupný / Vzestupný směr)	0,622/0,155	1,244/0,155 1,244/0,622	2,5/1,2	10/2,5	10/2,5 2,5/2,5
Maximální počet stanic	32	32	64	128	256
Max. dosah [km]	20	20	20	40	60

Pro standard NG-PON2 jsou schváleny dvě finální specifikace, třetí je ve vývoji a je průběžně doplňován společenstvím FSAN. Druhým tvůrcem obdobných standardů je institut IEEE.
Bylo očekáváno, že následkem vývoje NG-PON2 vydá také institut IEEE standard, který se bude svými parametry podobat. Společnost ovšem vytvořila studijní skupinu, jejímž úkolem je vyvinout standard NG-EPON (Next Generation of Ethernet Passive Optical Networking), který by předstihl dobu a dokázal by uspokojit požadavky koncových zákazníků na několik let dopředu. Předpokládá se, že by bylo možné tohoto cíle dosáhnout použitím pokročilých modulačních technik. Již dnes jsou k dispozici publikace, prezentující vysokorychlostní přenosy na jedné vlnové délce. Studijní skupina uvažuje až 25 Gbit/s na jedné vlnové délce a zároveň je zvažována kombinace až 4 takových zdrojů signálu, tedy přenosová rychlost až 100 Gbit/s. V rámci zvažovaného standardu by teoreticky mohlo být využíváno koncových jednotek umožňujících příjem dvou i více vlnových délek, což by zvýšilo rychlost na 50 Gb/s či více. Vzhledem k rozpracovanosti standardu se zatím jedná spíše o vizi [1].

2 NG-PON2

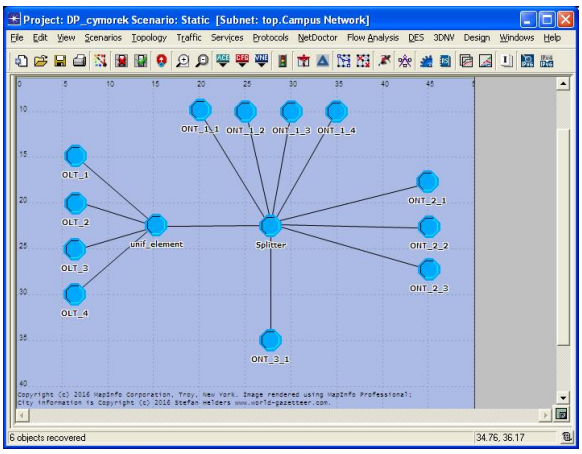
Nejnovější standard známý pod zkratkou ITU-T G.989 vydaný v roce 2015, ačkoli stále nedokončený, nese označení NG-PON2 (Next-Generation Passive Optical Network 2) [3].
Síť je tvořena nejdůležitějšími jednotkami:

- OLT (Optical Line Terminal) jednotkami na rozhraní přístupové a transportní sítě,
- ONT (Optical Network Terminal) jednotkami provádějícími převod optického signálu na elektrický na straně koncových uživatelů.

Provoz v síti se principiálně neliší od předešlých standardů. V sestupném směru jednotky OLT multiplexují XGEM rámce na optické vlákno využívající XGEM Port-ID jako klíč pro identifikaci, pro kterou koncovou jednotku je daný rámec vyslán. Ve splitterech dochází k vydělení příchozího optického signálu na všechny výchozí porty. Jednotky ONT filtrují pro ně určenou komunikaci na základě XGEM Port-ID, ostatní příchozí rámce jsou zahozeny. Multicast komunikace může být přijímána více jednotkami současně.
Pro vysílání ve směru vzestupném přiděluje jednotkám ONT časové intervaly pro vysílání jednotka OLT staticky nebo dynamicky. Jednotka OLT, které je směřována vzestupná komunikace je identifikována alokačním ID (Alloc-ID). Pro přiřazení rámců k různým logickým spojením využívají jednotky ONT jako identifikátor XGEM Port-ID.
Nejnovější standard, stejně jako jeho předchůdci, je zpětně kompatibilní. Společně s NG-PON2 je možno provozovat sítě standardu GPON a XG-PON na společném médiu díky rozdílným vlnovým délkám, čímž nedochází ke kolizím. Jednotky sítě se při užívání různých standardů ovšem liší. U standardu NG-PON2 se jedná především o jednotky OLT a ONT, které využívají unikátních prvků, a to aktivně laditelných laserů a filtrů.
Společnost Calix Inc. dodávající zařízení a software pro poskytovatele připojení ve svých laboratořích vytvořila model sítě, kdy byly na jediné optické vlákno multiplexovány přenosy ve standardech GPON, XG-PON a NG-PON2 a bylo v součtu sestupné a vzestupné komunikace dosaženo rychlosti 100 Gbit/s.

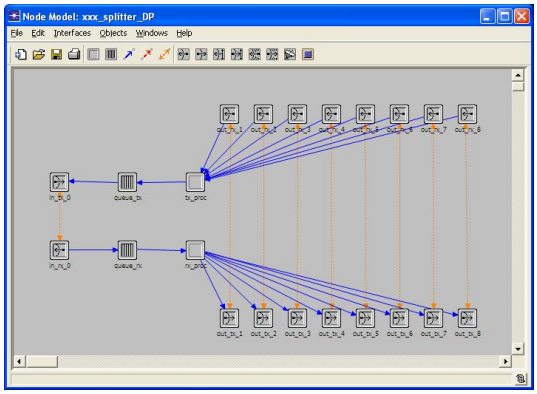
3 SIMULACE NG-PON2 V OPNET MODELERU

Pro simulaci statického přidělování grantů byl zvolen program OPNET Modeler v14.5. Navrženou topologii optické sítě NG-PON2 je možno vidět na obrázku 1.



Obrázek 1: Topologie sítě NG-PON2.

V rámci tohoto programu existuje mnoho implementovaných modelů uzlů a procesů, optická síť ovšem implementována není. Pro potřebu bylo tedy nutno vytvořit vlastní procesní a uzlové modely, které co nejvěrněji odpovídají skutečnému chování v síti standardu NG-PON2. Jako příklad je na obrázku 2 uveden uzlový model splitteru.



Obrázek 2: Navržený uzlový model splitteru.

Na výše uvedeném obrázku je možno vidět, že model se sestává ze vstupních a výstupních rozhraní, procesorů pro jednotlivé směry komunikace a front. Na rozdíl od reálných splitterů je nutno v modelu implementovat fronty, které vnáší do simulace zpoždění, se kterým je nutno dále pracovat. V kontrolním výpisu uvedeném níže, je možno vidět rozdíl mezi časy, kdy různé jednotky ONT zpracovaly rámce pro ně určené. Rozdíl je tvořen dobou, která je potřeba k vysílání paketů z front, jejichž výstupní rychlost je nastavena na 10 Gbit/s.

```
Packet in ONT_1_1 time: 0.000000512, grant start: 0.000002000, end: 0.125002000.
Packet in ONT_1_2 time: 0.000000614, grant start: 0.125003000, end: 0.250003000.
Packet in ONT_1_3 time: 0.000000717, grant start: 0.250004000, end: 0.375004000.
Packet in ONT_1_4 time: 0.000000819, grant start: 0.375005000, end: 0.500005000.
Packet in ONT_1_1 time: 0.500004512, grant start: 0.500005999, end: 0.625005999.
```

4 ZÁVĚR

V rámci této práce byl teoreticky představen standard NG-PON2 definován ITU-T, pro který byly následně v programu OPNET Modeler nezahrnujícím optické entity vytvořeny vlastní modely. Simulací bylo demonstrováno statické přidělování grantů v pasivní optické síti. Byla zobrazena část konzolového výpisu, včetně poukázání na rozdíly mezi reálnými a simulovanými prvky.

REFERENCE

[1] Knittle, C.: *Passive Optical Networking – for the Next Generation* [online]. 2015. Dostupné z URL: http://www.cablelabs.com/next-generation-passive-optical-networking/

[2] Cymorek, P.: *Simulace Triple Play služeb v sítích XG-PON v prostředí NS-3.* Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2015. 62 s

[3] ITU-T.: *Recommendation G.989.1* [online]. 2013. Dostupné z URL: https://www.itu.int/rec/T-REC-G.989.1/en

WAREHOUSE MODELING USING GRAPHICAL USER INTERFACE

Martin Rajnoha
Master Degree Programme (5), FEEC BUT
E-mail: xrajno05@stud.feec.vutbr.cz

Supervised by: Radim Burget
E-mail: burgetrm@feec.vutbr.cz

Abstract: This paper describes a new algorithm which enable efficient conversion of graphical representation of warehouse into graph theory representation and consequently accelerates estimation for route costs. The proposed algorithm computes route distances between any place in warehouses and does so significantly faster than traditional approaches. For this purpose an algorithm based on Breadth first search, image processing "skeletonization" and Dijkstra algorithm was proposed. Using the proposed algorithm it is possible to search routes in a warehouse effectively and fast using precomputed routing table. Searching time is approximately hundreds of microseconds using routing table and even it is independent on size of warehouse instead of using Dijkstra algorithm.

Keywords: warehouse, GUI, skeletonization, graph, searching, paths, routing, table

1. ÚVOD

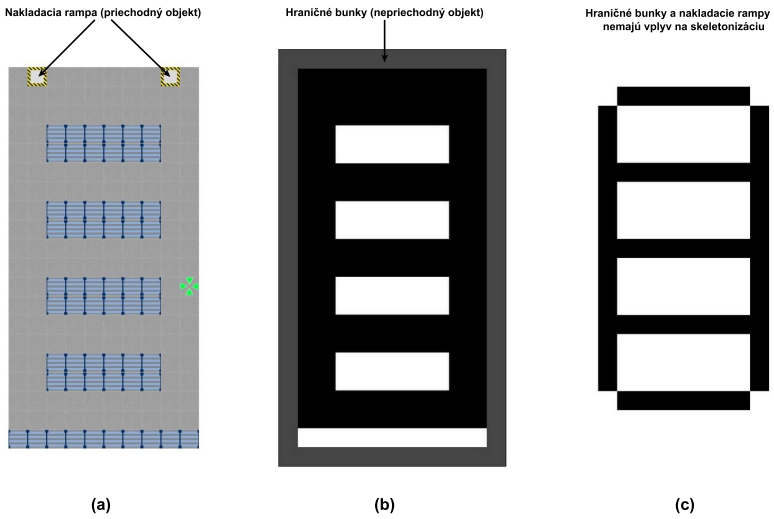
Sklad v podobe dátového modelu otvára mnohé možnosti nie len v oblasti skladového hospodárstva, ale predovšetkým plánovania a optimalizácie práce v sklade. Ako však vytvoriť dátový model, aký použiť a ako s ním pracovať aby sa dosiahli očakávané výsledky v oblasti časovej náročnosti optimalizačných algoritmov? Článok opisuje metódy prekonvertovania skladu do dátového modelu (grafu) a algoritmy na rýchle a efektívne vyhľadávanie ciest. Z nakresleného skladu sa vytvorí dvojrozmerné pole s priechodnými a nepriechodnými prvkami, ktoré sa ďalej prekonvertuje do obrázku binárneho typu. V ňom sa nájde kostra pomocou metódy skeletonizácie, v ktorej sa identifikujú miesta, kde sú smerovače, ktoré budú tvoriť vrcholy grafu a cesty, ktoré ich spájajú (budú hranami grafu). Pomocou vyhľadaných ciest a ich dvojíc smerovačov sa vytvorí graf. Na základe toho sa predpočíta pomocou Dijkstrovho algoritmu smerovacia tabuľka.

Hľavným prínosom práce je algoritmus na prevod skladu v grafickej podobe do dátového modelu (vytvorenie grafovej reprezentácie) a algoritmus, ktorý umožňuje vyhľadávať cesty v rádoch stoviek mikrosekúnd. To je dôležité pre implementáciu optimalizačných algoritmov, ktoré môžu využívať genetické algoritmy.

2. DÁTOVÁ ŠTRUKTÚRA SKLADU

Pole skladu je typu abstraktných buniek – abstraktná trieda buniek z ktorej dedia jednotlivé triedy na kreslenie skladu. Rozmery poľa sú odvodené z dĺžky a šírky skladu a hodnoty oddeľovača (určuje, po koľkých metroch sa sklad delí na bunky). Ďalej je nutné vytvoriť graf s čo najmenším počtom vrcholov. Sklady môžu byť rôznych tvarov a nemusia byť vždy nakreslené ideálne, preto vlastný algoritmus vyhľadávania vrcholov by bol zdĺhavý a nemusel byť vždy správny. Kvôli tomu sa využíva metóda skeletonizácie. Skeletonizácia je transformácia digitálneho obrazu do podmnožiny pôvodného obrazu, s cieľom znížiť množstvo dát, potrebného na reprezentáciu obrazu, alebo zjednodušiť tvar obrazu [3]. Pole sa prekonvertuje na obrázok binárneho typu, kde každá bunka poľa reprezentuje

jeden pixel v obrázku (musí naberať hodnoty len 0 a 1) [1]. Pomocou metódy iteratívneho stenšovania (iterative thinning) získame kostru – súbor tenkých čiar, oblúkov a kriviek (zvyčajne široká jeden pixel), ktoré sú spojené tak, aby zachovali pôvodné topologické a geometrické vlastnosti objektu [4]. Proces skeletonizácie je zobrazený na obr. 1.



Obrázok 1: Proces skeletonizácie, kde (a) zobrazuje nakreslený sklad, (b) znázorňuje sklad ako obrázok a (c) zobrazuje výslednú kostru vypočítanú skeletonizáciou.

V kostre sa pomocou ďalších algoritmov vyhľadajú križovatky, ktoré tvoria smerovače – vrcholy grafu. Následne sa prehľadajú jednotlivé vetvy kostry, tvoriace cesty – hrany grafu. Pixel, ktorý reprezentuje smerovač, musí ležať na kostre a taktiež jeho dva susedné pixely, ktoré sú navzájom kolmé. Vyhľadávanie ciest začína vždy od smerovača. Prehľadá sa jeho okolie a po narazení na pixel, ležiaci na kostre sa pomocou BFS (Breadth First Search) algoritmu vyhľadá celá cesta končiaca v druhom smerovači, prípadne v slepej uličke. Aby algoritmus pokračoval požadovanou cestou, je potrebné do množiny CLOSED pridať pixely v okolí počiatočného smerovača. Taktiež sa v každom kroku musí prehľadať celé okolie pixelu cesty (pixely kostry môžu následovať aj uhlopriečne a mohlo by sa naraziť na pixel inej cesty). Po skončení procesu prehľadávania sa do grafu pridajú cesty s dvojicou smerovačov – susedné vrcholy, spojené hranou. V mnohých prípadoch môžu medzi vrcholmi vzniknúť rovnobežné hrany. Po pridaní všetkých ciest, vzniká konečná podoba grafu – sieť skladu. [2]

3. VYHĽADÁVANIE CIEST V SKLADE

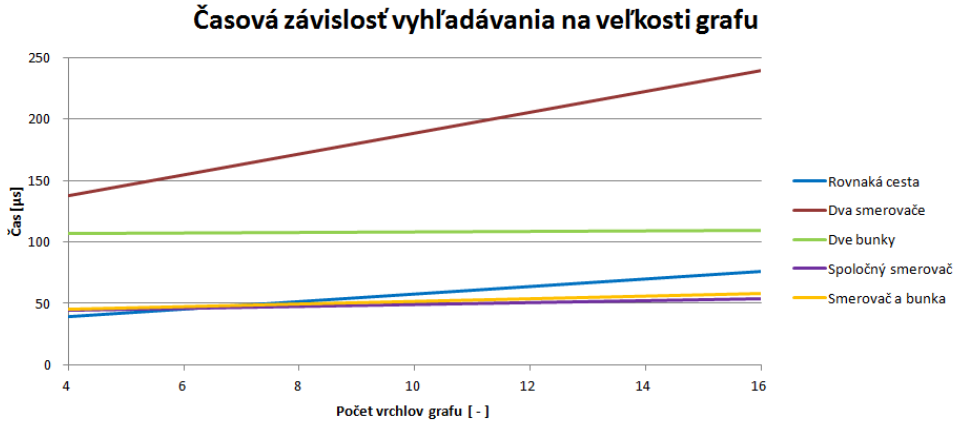
Vo vytvorenej štruktúre grafu, je už možné vyhľadávať, zatiaľ ale len medzi jeho vrcholmi. Pre potreby skladu je nutné umožniť vyhľadávanie z ktorejkoľvek pozície. Každej bunke sa preto ďalšími algoritmami priradia najbližšie smerovače a predpočítajú vzdialenosti k nim. Najbližší smerovač však nemusí vždy viesť k najlepšej ceste. Preto sú predpočítané obidva smerovače a prehľadávajú sa všetky kombinácie cez obidva smerovače štart a cieľ bunky a použije sa najkratšia cesta. Nie všetky cesty prechádzajú smerovačmi (štart a cieľ ležia na rovnakej ceste). Taktiež k nájdeniu cesty nemusí byť použitý graf (štart a cieľ majú spoločný smerovač – neexistuje kratšia cesta ako cez tento smerovač). Tieto prípady je nutné odlíšiť v algoritme pre vyhľadávanie ciest.

V ostatných prípadoch je nutné prehladať štruktúru grafu. Graf sa prehľadáva pomocou Dijkstrovho algoritmu, ktorý nájde najlepšie riešenie. Rýchlosť vyhľadávania však závisí na počte vrcholov grafu. Problém rieši smerovacia tabuľka, ktorá predpočíta cesty pre každú dvojicu smerovačov na základe

Dijkstrovho algoritmu. Cesty sú uložené v hešovacej mape. Pri vyhľadávaní ciest sa iba vyhľadá daný počiatočný smerovač a jeho najlepšia predpočítaná cesta k cieľovému smerovaču. Rýchlosť vyhľadávania ciest je tak výrazne lepšia (viď obr. 2).

4. VÝSLEDKY

V práci bola vytvorená aplikácia umožňujúca nakreslenie skladu, prevod do dátového modelu, identifikovanie smerovačov a ciest a následne vytvorenie grafovej štruktúry, v ktorej je možné rýchlo vyhľadávať cesty z ľubovoľných pozícií skladu. Najdôležitejším výsledkom je závislosť rýchlosti vyhľadávania ciest na veľkosti grafu (zvyčajne väčší sklad = väčší graf) zobrazená na obr. 2.



Obrázok 2: Časová závislosť vyhľadávania na veľkosti grafu. Závislosti sú preložené lineárne a sú takmer konštantné pre všetky typy vyhľadávania, najdôležitejšia závislosť je ale typu Dve bunky, čo je najviac pravdepodobná reálna situácia – vyhľadávanie z ľubovoľných pozícií v sklade.

5. ZÁVER

V práci boli vykonané testy a merania, ktoré dokazujú funkčnosť prevodu nakresleného skladu do grafu a rýchle vyhľadávanie ciest, ktoré sú popísané v sekcii výsledkov práce. Pri vyhľadávaní ciest nie je podstatná len samotná rýchlosť vyhľadávania (rádovo stovky mikrosekúnd), ale aj vplyv veľkosti grafu na čas vyhľadania cesty. Dôležitým prínosom práce je preto aj závislosť rýchlosti vyhľadávania na veľkosti grafu, ktorá je takmer konštantná, čo znamená, že veľkosť skladu neovplyvňuje rýchlosť vyhľadávania. Využitie týchto metód ako základ pre optimalizačné algoritmy, by mohol znamenať veľký posun v skladovom hospodárstve.

LITERATÚRA

[1] BURGER, W., BURGE M. J. *Digital Image Processing. An Algorithmic Introduction using Java. First Edition* 2008 ISBN 978-1-84628-379-6, e-ISBN 978-3-540-30941-3.

[2] BURGET, R. *Teoretická informatika* Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií, 2013. ISBN: 978-80-214-4897-1

[3] KLETTE, G. *Skeletons in Digital Image Processing* ITR Tamaki, University of Auckland Tamaki Campus, Building 731, Auckland, New Zealand. 29. 7. 2002

[4] VANAJAKSHI, B., SUJATHA B., KRISHNA S. R. *An Analysis of Thinning & Skeletonization for Shape Representation* International Journal of Computer Communication and Information System (IJCCIS) – Vol2. No1. ISSN: 0976–1349 7 – 12 2010

Doktorské projekty

Elektronika a komunikace

60 GHZ TIME DOMAIN CHANNEL SOUNDER

Josef Vychodil
Doctoral Degree Programme (3), FEEC BUT
E-mail: vychodil@phd.feec.vutbr.cz

Supervised by: Aleš Prokeš
E-mail: prokes@feec.vutbr.cz

Abstract: This paper describes a time domain channel sounder based on pseudorandom binary sequences. The channel sounder system consists of several off-the-shelf laboratory instruments connected together and controlled from PC. Frequency band of interest is the unlicensed millimeter wave band (MMW), 57–64 GHz. The system's architecture and its parameters are presented and discussed. Crucial feature of the proposed system is fast measurement speed enabling measurements of real scenarios with moving objects.

Keywords: MMW, millimeter wave band, channel sounding

1 INTRODUCTION

In modern communication, localization or radar systems, especially those working with huge bandwidths working for example in the MMW, allocated for unlicensed operation in the frequency band 57 to 64 GHz, perfect knowledge of the propagation channel parameters is advantageous. As these parameters get more precise, more advanced data processing methods and modulation techniques can be used to transfer the data, which lead to higher data rates and lower bit error ratios (BER). The enormous instantaneous bandwidths of several GHz in the MMW promise high data throughput and precise spatial localization.

The parameters of the radio channel may be obtained through channel sounding. There are several techniques and methods that can be utilized to perform those measurements. The measurements can be performed either in the frequency or in the time domain. Measurements in frequency domain are usually carried out using vector network analyzers (VNAs) which perform stepped frequency sweeping as it was shown for example in [1], while those in the time domain utilize transmitting and receiving wideband pulses or pulse sequences.

In [2], a precision time domain based sounding system utilizing custom silicon chips, whose development is very expensive, was presented. This paper presents a time domain channel sounder utilizing pseudorandom binary sequences (PRBS) built from several off-the-shelf laboratory instruments. The proposed system promises short measurement times and provides correlation gain over direct impulse based time domain methods. It is also simple to realize and replicate.

2 CHANNEL IMPULSE RESPONSE

A convenient way how to represent channel characteristics is to specify corresponding channel impulse response (CIR). It shows the channel output when the input is an ideal unit impulse. The general way to model CIR of a time invariant multipath wireless channel is to express it as a sum of attenuated, delayed and phase shifted replicas of the Dirac delta function:

$$h(t) = \sum_{n=0}^{L-1} \rho_n e^{j\Phi_n} \delta(t - \tau_n), \tag{1}$$

where L is the number of multipath components, δ is the Dirac delta function and τ_n, ρ_n and Φ_n is path delay, gain and phase shift of n-th component, respectively.

3 MEASUREMENTS USING VNA

Measurement in the frequency domain utilizing the vector network analyzer (VNA) is the conventional way to obtain the CIR. It has several advantages (e.g. great dynamic range) and also disadvantages – slow measurement speed and the fact that the receiver and the transmitter must be physically connected (synchronized) and located close to each other.

The VNA provides measurement of the channel transfer function $H(f)$ in terms of the s parameter, which represents the channel gain and phase shift as a function of frequency. We can convert between $H(f)$ and $h(t)$ utilizing fast Fourier transform (FFT) or more precisely inverse fast Fourier transform (IFFT):

$$\begin{aligned} H(f) &= \int_{-\infty}^{\infty} h(t) e^{-j2\pi ft} dt \\ &= \sum_{n=0}^{L-1} \rho_n e^{j\Phi_n} e^{-j2\pi f \tau_n}. \end{aligned} \tag{2}$$

A complete sweep over the frequency range of interest using VNA usually takes several tens of seconds, depending on the number of frequency steps. During this time, the channel must remain static. For this reason the VNA method cannot be used to measure either slowly changing channels (caused e.g. by people movement) or for measuring of fast changing channels (caused e.g. by vibrations). On the other hand, VNA offers an excelent dynamic range usually greater than 60 dB.

4 PROPOSED MEASUREMENT SYSTEM

There are several ways how to measure the CIR in the time domain. For example, it is possible to transmit Gaussian pulses (or its derivatives) and then recover the CIR by deconvolving the received signal utilizing the CLEAN algorithm. However, this work will be focused on the time domain channel sounding based on pulse sequences, rather than single pulse. This technique provides advantages as correlation gain and is not too challenging for the power amplifiers.

4.1 SYSTEM DESCRIPTION

Block diagram of the proposed channel sounding system is depicted in Fig. 1. The Anritsu Signal Quality Analyzer MP1800A serves as a baseband PRBS generator at data rates up to 12.5 Gbits/s with maximum RF output power of 13 dBm. Digital sampling oscilloscope Tektronix MSO72004C is utilized as a receiver. It provides 4 channels, 16 GHz bandwidth, 50 GS/s real time acquisition rate and 31.25 MS of data storage per channel. The Signal Quality Analyzer also provides 10 MHz reference and gating (triggering) signal to the oscilloscope. The power amplifier and low noise amplifier (LNA) are not depicted for the sake of simplicity. A PC can be used to control the instruments and interchange data to provide additional features (e.g. real-time and continuous channel sounding) but is not necessary for the data acquisition.

The Sivers IMA FC1003V/01 [3] up and down converter is used for frequency shifts between the baseband and the radio frequency. It is a direct conversion transceiver (sometimes called homodyne or zero–IF) with IF input bandwidth up to 5 GHz. The baseband transmit signal is feeded into I branch of the up converter. The Q branch is terminated by a 50 Ω terminator, which practically disables this input. This in ideal case leads to an Amplitude Modulation (AM) with symmetrical spectrum around carrier frequency $f_c = 59.6$ GHz. This RF signal is then propagated through the channel, which is

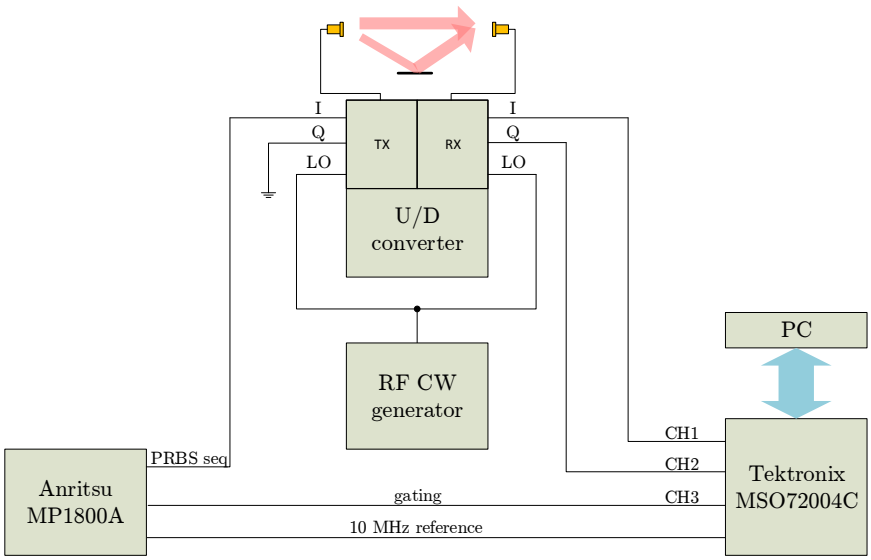


Figure 1: Proposed channel sounder architecture.

the object of measurement. Any type of antenna can be used, but typically an open ended WR15 waveguide is used. The received signal is then down converted and the output signal – complex, in the form of I and Q components – is passed to the Tektronix MSO72004C oscilloscope for sampling and further processing. To achieve perfect coherency, the Local Oscillator (LO) signal is generated in a RF generator, then it is split in the Wilkinson power divider, and fed into both up and down converter LO input.

A seamless repeating PRBS, concretely the m-sequence, was chosen as an excitation signal, mainly for its nearly perfect circular auto-correlation properties. The length of the PRBS is $N = 2^k - 1$ chips where k can in our case vary from 7 to 31. The parameter k determines the maximum observable CIR time span given by

$$T_{\max} = \frac{N}{f_{\text{chip}}} = \frac{2^k - 1}{f_{\text{chip}}}, \tag{3}$$

which is also the time period of the sequence. The parameter $T_{\max}$ is related to the maximum observable propagation distance by

$$D_{\max} = c \cdot T_{\max} = c \cdot \frac{N}{f_{\text{chip}}} = c \cdot \frac{2^k - 1}{f_{\text{chip}}} \tag{4}$$

where $c \approx 3 \times 10^8\,\text{m}\,\text{s}^{-1}$ is the speed of light. Considering $k = 11$ and $f_{\text{chip}} = 12.5$ GHz, the maximum propagation distance is $D_{\max} = 49.128$ m, which is sufficient for most of the short-range MMW channels.

Correlation gain, which improves the SNR at the receiver is given by

$$G_{\text{corr}} = 10 \cdot \log N = 10 \cdot \log(2^k - 1). \tag{5}$$

For our case of $k = 11$ the correlation gain is $G_{\text{corr}} = 33.11$ dB.

Main advantage of the proposed channel measurement system over the system based on VNA is the CIR measurement time. As stated in Sec. 3, it takes several tens of seconds to measure one single CIR using VNA. On the proposed system, the time to measure one single CIR equals the time length of the sequence given by (3). Considering $k = 11$ this leads to $T_{\max} = 0.163\,\mu\text{s}$. The reciprocal value,

$f_{\text{meas}} = \frac{1}{T_{\max}}$ denotes the maximum CIR measurements per second. For $k = 11$ this is theoretically 6.1 millions of measurements per second. The oscilloscope provides 31.25 MS of fast acquisition memory per channel, which equals to 0.625 ms at sampling rate of 50 GS/s, so only 3816 CIRs can be captured at once. Nevertheless using advanced triggering modes, the f_{meas} can be decreased to arbitrary value. It is also possible to continuously stream the real time acquired data to a PC at the rate $f_{\text{meas}} \approx 100$ Hz. In static channel measurement scenarios, averaging of the data can be utilized to further improve the SNR.

4.2 DATA PROCESSING

In order to obtain an estimate of CIR in the desired frequency band, which in our case is the MMW (∼55–65 GHz), the following steps are performed.

The transmitted PRBS signal $e(t)$ is periodic with period $T_{\max}$. Since we assume that the CIR is causal and has a finite length of $T_{\max}$ i.e.

$$h(t) = 0 \text{ for } t > T_{\max} \text{ and } t < 0, \tag{6}$$

the received signal

$$r(t) = e(t) * h(t), \tag{7}$$

where $*$ denotes the convolution operation, is also periodic with period $T_{\max}$.

Circular auto correlation function of the transmitted PRBS sequence $e(t)$ $R_e(t) = 1$ when t mod $T_{\max} = 0$, and $\frac{1}{N}$ otherwise. This property can be exploited to obtain an estimate of the CIR as

$$h(t) = r(t) * e(-t). \tag{8}$$

For faster calculation, the convolution operation in the time domain is realized as a multiplication operation in the frequency domain. The conversions from time to frequency domain and vice versa, are done via FFT and IFFT, respectively:

$$h(t) = \text{IFFT}\{\text{FFT}[r(t)] \cdot \text{FFT}[e(-t)]\}. \tag{9}$$

Please note that only the baseband transmitted signal $e(t)$ is real. The baseband received signal $r(t)$ and the CIR estimate $h(t)$ are generally complex valued.

After this, it is needed to filter out the spectral components, which are outside of the band of interest e.g. 55–65 GHz. This is done in the frequency domain – first, only spectral components, which are in the range of 55–65 GHz are taken and the rest is discarded. The remaining components are multiplied by Hann window function and transformed back to the time domain representation.

5 EXAMPLE

An example of a measurement performed by the proposed system is depicted in Fig. 2. It shows the magnitude of time-varying CIR captured in a moving car when the transmitter was held in hands, thus simulating a hand held device. The measurement rate was $f_{\text{meas}} = 1000$ Hz. It clearly shows the changes of the channel caused by slight movements of the car itself, the transmitter and the persons inside the car.

These data can be further processed to obtain channel parameters like RMS delay spread, delay Doppler spread, etc.

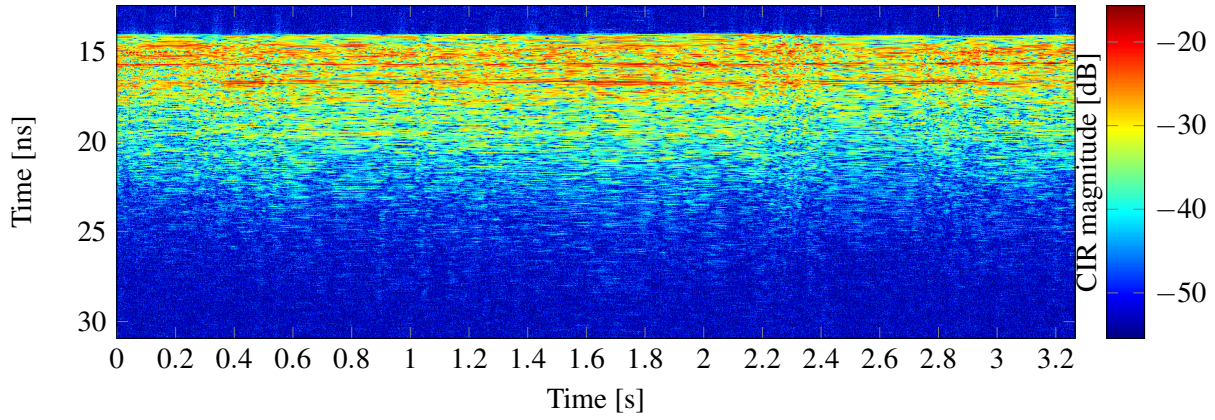


Figure 2: Example measurement – magnitude of time-varying CIR. The x axis is the "slow" time which corresponds to the measurement rate and the y axis is the "fast" time which corresponds to the travelling speed of the electromagnetic wave.

6 CONCLUSION

A 60 GHz channel sounder working in the time domain was presented in this paper. It gave a brief description of the instruments used and the principles and methods utilized in the system. An example measurement was shown and discussed. Main advantage over conventional channel sounding using the VNA is the possibility of very fast channel measurement.

The biggest disadvantage of this system is small dynamic range around 30 dB. This limitation arises mainly from the non-linear behavior of the devices in the measurement chain which causes spurious peaks in the CIR to appear. The noise contribution (which can be reduced for example by averaging) is not crucial regarding the dynamic range.

Future work will deal with improving the dynamic range of the system. There are also some other problems which arise from the real devices imperfections, for example the up and down converter IQ imbalance, which results in finite image signal attenuation.

ACKNOWLEDGEMENT

This work was supported by the Czech Science Foundation project No. 13- 38735S Research into wireless channels for intra-vehicle communication and positioning, and was financed by Czech Ministry of Education in frame of National Sustainability Program under grant LO1401. For research, infrastructure of the SIX Center was used.

REFERENCES

[1] M. Schack, M. Jacob, and T. Kurner, "Comparison of in-car UWB and 60 GHz channel measurements," in *Antennas and Propagation (EuCAP), 2010 Proceedings of the Fourth European Conference on*, April 2010, pp. 1–5.

[2] R. Zetik, M. Kmec, J. Sachs, and R. S. Thoma, "Real-time mimo channel sounder for emulation of distributed ultrawideband systems," *International Journal of Antennas and Propagation*, vol. 2014, p. 16, 2014. [Online]. Available: http://dx.doi.org/10.1155/2014/317683

[3] *FC1003V/01 V-band Converter*, Sivers IMA AB. [Online]. Available: http://siversima.com/wp-content/uploads/FC1003V01-Data-Sheet.pdf

ACCURACY COMPARISON OF SOME 2D NUMERICAL INVERSE LAPLACE TRANSFORM METHODS

Nawfal Al-Zubaidi R-Smith

Doctoral Degree Program 2, FEEC BUT

E-mail: xalzub00@stud.feec.vutbr.cz

Supervised by: Lubomir Brancik

E-mail: brancik@feec.vutbr.cz

Abstract: Numerical Laplace inversions are essential to use when the analytical manipulations of the Laplace transform tables are not possible to apply, and this becomes more and more difficult for applications with higher number of variables. In this paper we describe three methods for 2D numerical inverse Laplace transforms and analyse the methods as for their accuracy and calculation efficiency by implementing them in the Matlab environment. The relative and absolute errors for three different testing functions with previously known originals are presented in a comparative table for the selected points, and the 3D resulting plots for the inversion of a test function are displayed.

Keywords: Numerical inversion, two-dimensional Laplace transform, test functions, Matlab language, 3D plot.

NOMENCLATURE

- f is the time domain function
- F is the Laplace domain function
- t is time and x represents space (distance)
- s is the complex frequency variable
- c is distance from the imaginary axis
- Re is the real part
- Im is the imaginary part
- α_1 and α_2 are real numbers
- M is a positive constant
- T is the period of the Fourier series function, sec.
- t_{max} is the maximum time interval, sec.
- N is the number of terms; positive integer.
- K_{1i} , K_{2l} are the residues of approximating the exponential kernel
- z_{1i}, z_{2l} are the poles of approximating the exponential kernel
- *FFT* is fast Fourier transform
- T_i are sampling periods in the original time domain
- Ω_i is the frequency step
- E_r is the desired relative error

1. INTRODUCTION

Two-dimensional numerical inverse Laplace transforms (2D-NILTs) arise in electrical engineering related fields for applications that require solution of partial differential equations with two variables. Such applications are those defining transients in linear distributed-parameter systems, e.g. solving telegrapher equations, obtaining voltage, current distributions in the original time domain [1, 2], or more sophisticated such as that concerning slightly nonlinear systems described by second-order Voltera series expansion [3]. In this paper three inversion methods are described, these methods are based on using Fourier series techniques or the Padé approximation. These methods have proved to give good results and to be relatively universal for a wide range of applications [4]. Each method has different free parameters which could be optimized for higher accuracy results. The three methods in the next section are introduced in the following order, the first method is the

2D NILT based on Fourier series representation by Moorthy [4]. The second method is the 2D NILT by Singhal, Vlach and Vlach [5], and the third method is the accelerated 2D-FFT NILT by Brančik [6]. The numerical methods start with the main 2D inverse Laplace transform formula described as the two-fold Bromwich integral namely,

$$f(t,x) = -\frac{1}{4\pi^2}\int_{c_1-j\infty}^{c_1+j\infty}\int_{c_2-j\infty}^{c_2+j\infty} F(s_1,s_2)e^{s_1t+s_2x}ds_1ds_2, \quad (1)$$

where t and x are generally any independent variables, but due to our interest in engineering related applications, we consider them to be time and space variables respectively. The Laplace variable is defined as $s = c + j\omega$.

2. 2D NUMERICAL INVERSE LAPLACE TRANSFORM METHODS

2.1. 2D NILT BASED ON FOURIER SERIES REPRESENTATION

The method was developed by M. Moorthy in[4], it is an expansion of a 1D method based on a representation of the inverse transform by Fourier series, which was introduced by Dubner and Abate [7]. The inversion formula consists of a combination of two infinite sums and a double infinite sum and is listed as follows,

$$\tilde{f}(t,x) = e^{c_1t+c_2x}g(t,x), \quad (2)$$

such that,

$$\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} g^{jk}(t,x) = \frac{1}{2T^2}\left[\frac{1}{2}F1_{c_1,c_2} + \sum_{m=1}^{\infty}\left[Re\{F2_{c_1,c_2}\}Z1 - Im\{F2_{c_1,c_2}\}Z2\right] + \sum_{n=1}^{\infty}\left[Re\{F3_{c_1,c_2}\}Z3 - Im\{F3_{c_1,c_2}\}Z4\right] + \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left[Re\{F4_{c_1,c_2}\}Z5 + Re\{F5_{c_1,c_2}\}Z6 - Im\{F4_{c_1,c_2}\}Z7 - Im\{F5_{c_1,c_2}\}Z8\right]\right], \quad (3)$$

where the F annotations are functions with different approximations of s_1 and s_2 and the Z are different sinusoidal functions. For the detailed derivation and annotations the reader is advised to check reference [4].

For the derivation of the final formula the following assumptions were considered,

$$|f(t,x)| \le Me^{\alpha_1t+\alpha_2x}, \quad (4)$$

and the Laplace domain of the function $F(s_1,s_2)$ is analytic for $Re\{s_1\} > \alpha_1$ and $Re\{s_2\} > \alpha_2$.

To obtain an optimum choice of the parameters used, an error analysis done in [4] suggests that the choice of c_1 and c_2 to be larger than α_1 and α_2 respectively; for example by first choosing $c_1 > \alpha_1$, then [4],

$$c_2 = \alpha_2 - \frac{1}{2T}\ln\left(\frac{\varepsilon - e^{-2T(c_1-\alpha_1)}}{1-e^{-2T(c_1-\alpha_1)}}\right), \quad (5)$$

to choose the parameter T, after considering $t_{max} < 2T$, and after running different tests the author suggests the choice of $0.5t_{max} \le T \le 0.8t_{max}$ [4].

2.2. 2D NILT ROPOSED BY SINGHAL, J. VLACH, AND M. VLACH

In this 2D numerical inversion method the algorithm is an expansion of a 1D NILT based on the Zakian numerical scheme[5, 8]. The inversion technique uses the Padé approximation and the residual theorem. The formulae is described as follows [5],

$$\hat{f}(t,x) = \frac{1}{t*x}\sum_{i=1}^{M_1}\sum_{l=1}^{M_2} K_{1i}K_{2l}\, F\left(\frac{z_{1i}}{t_1}, \frac{z_{2l}}{t_2}\right), \tag{6}$$

z_{1i}, z_{2l} are the poles of the rational Padé approximation, shown as follows [5],

$$e^{s_1 t} \approx e^{s_2 x} \approx \xi N_k, M_k(z_k) = \frac{\sum_{i=0}^{N_k}(M_k+N_k-i)!\binom{N_k}{i}{z_k}^i}{\sum_{i=0}^{M_k}(-1)^i(M_k+N_k-i)!\binom{M_k}{i}{z_k}^i}, \tag{7}$$

where $z_{k=1} = s_1 t$, and $z_{k=2} = s_2 x$.

Tests performed on the method show that the method performs well especially for sufficiently smooth functions. In the tables and figures following in the paper the method is denoted as "Singhal-2D NILT".

2.3. ACCELERATED 2D FFT-BASED NILT

In this numerical procedure described in [6] the inversion is based on fast Fourier transform algorithms *FFT*. Basically, the method uses the sum of two-dimensional complex Fourier series calculated by *FFT* algorithms. Furthermore, the ε-algorithm of Wynn is combined into the method to further give a higher accuracy and improvement to the result. Following in the text the term FFT-ε-2D NILT will be used to refer to this method.

The basic inversion algorithm namely is [6],

$$\tilde{f}^{k_1,k_2} = C^{k_1,k_2}\{2Re[\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} F_{-n_1,-n_2}E_{-n_1,-n_2}^{k_1,k_2} + \sum_{n_1=0}^{\infty}(\sum_{n_2=0}^{\infty} F_{-n_1,n_2}E_{n_2}^{k_2})E_{-n_1}^{k_1} - \sum_{n_1=0}^{\infty} F_{-n_1,0}E_{-n_1}^{k_1} - \sum_{n_2=0}^{\infty} F_{0,-n_2}E_{-n_2}^{k_2}] + F_{0,0}\}, \tag{8}$$

where $\tilde{f}^{k_1,k_2} = \tilde{f}(k_1T_1, k_2T_2)$, is the discrete form of the approximate formula. The symbols used are given as follows [6],

$$F_{n_1,n_2} = F(c_1 + jn_1\Omega_1, c_2 + jn_2\Omega_2), \tag{9}$$

$$E_{n_1,n_2}^{k_1,k_2} = e^{jk_1T_1n_1\Omega_1 + jk_2T_2n_2\Omega_2} = E_{n_1}^{k_1}E_{n_2}^{k_2}, \tag{10}$$

$$C^{k_1,k_2} = \frac{\Omega_1\Omega_2}{4\pi^2}e^{c_1k_1T_1 + c_2k_2T_2} = C^{k_1}C^{k_2}, \tag{11}$$

where $\Omega_i = \frac{2\pi}{N_iT_i}$, $i = 1,2$, the discrete points given as $t_{k_i} = k_iT_i$, and $k_i = 0,1,\cdots,N_i - 1$.

The error analysis in [6], provides us with the ideal choice of c_i , that is to say by choosing E_r as the desired relative error then the choice of c_i can be given as,

$$c_i \approx \alpha_i - \frac{\Omega_i}{2\pi}\ln\frac{E_r}{2},\ i = 1,2. \tag{12}$$

The parameter N is introduced as a free parameter, such that the infinite sums in (8) are calculated up to $N-1$ terms; choosing $N = 512$ has shown to give good results, the terms above N are integrated into the ε-algorithm for higher accuracy. Generally, the ε-algorithm is considered as a non-linear algorithm used to speed up the convergence of the series and it is equivalent to a rational Padé approximation. The ε-algorithm is known to give good results when applied to power series; this is possible in the current case, since with some mathematical manipulations the complex Fourier series becomes as such a power series. In [6] the ε-algorithm diagram is shown with the detailed description of its application to the numerical method.

3. 2D NILT METHODS ACCURACY TESTS

After describing the 2D NILT methods and their inversion algorithms, here the absolute errors and relative errors for test functions with previously known originals are examined on the methods described in this paper.

The three methods were implemented by using the universal mathematical language Matlab and the test functions used here have been endorsed to be used as smooth test functions in both variables [4, 5]. The first test function in the Laplace domain with its known original is listed as follows,

$$F1(s_1, s_2) = \frac{1}{(s_1+1)(s_2+2)}, \tag{13}$$

$$f1(t, x) = e^{-1t-2x}, \tag{14}$$

The absolute errors and relative errors are calculated at the specific point $(t, x) = (1,1)$, the results are listed in an ascending order starting with the method of higher accuracy in Table 1, bellow.

2D NILT method	Absolute error	Relative error
FFT-ε-2D NILT	$2.9698 * 10^{-10}$	$5.9650 * 10^{-9}$
Singhal-2D NILT	$1.338 * 10^{-8}$	$2.687 * 10^{-7}$
Moorthy-2D NILT	$4.257 * 10^{-7}$	$8.55 * 10^{-6}$

Table 1: Accuracy tests for 2D NILT, function 1

The second function to be tested is described as,

$$F2(s_1, s_2) = \frac{1}{(s_1-1)\cdot(s_2-1)\cdot(s_1+s_2-1)}, \tag{15}$$

$$f2(t, x) = \begin{cases} e^x(e^t - 1), & x \geq t \\ e^t(e^x - 1), & x < t \end{cases}, \tag{16}$$

In Table 2, the absolute and relative errors are shown in an ascending order for the inversion methods tested on $f2(t, x)$ at $(t, x) = (1,1)$.

2D NILT method	Absolute error	Relative error
FFT_ε-2D NILT	$1.07288 * 10^{-3}$	$2.2970 * 10^{-4}$
Moorthy-2D NILT	$1.330 * 10^{-2}$	$2.85 * 10^{-3}$
Singhal-2D NILT	$5.320 * 10^{-2}$	$1.139 * 10^{-2}$

Table 2: Accuracy tests for 2D NILT, function 2

The results shown in the tables above have the FFT-ε-2DNILT method with the relatively highest accuracy among the three methods tested with an accuracy improvement in average of about 2 or-

ders. In Table 3, the different parameters used to implement the results of the methods are listed for each method respectively [4-6].

2D NILT Method	Parameters used to test the functions
FFT_ε-2D NILT	$\alpha_1 = 0, \alpha_2 = 0, E_r = 1e-7$. Same values for both functions.
Moorthy-2D NILT	For function1, $c_1 = 1, c_2 = 0, T = 3.5.$ For function2, $c_1 = 15, c_2 = 10.21, T = 1.0.$
Singhal-2D NILT	$M = 10, N = 8$. Same values for both functions used.

Table 3: Parameters used for different NILT methods

The results of inverting function 1, with the three 2D NILT methods are shown as 3D plot results in Figure 1. The 3D plots were determined through a grid of 256x256.

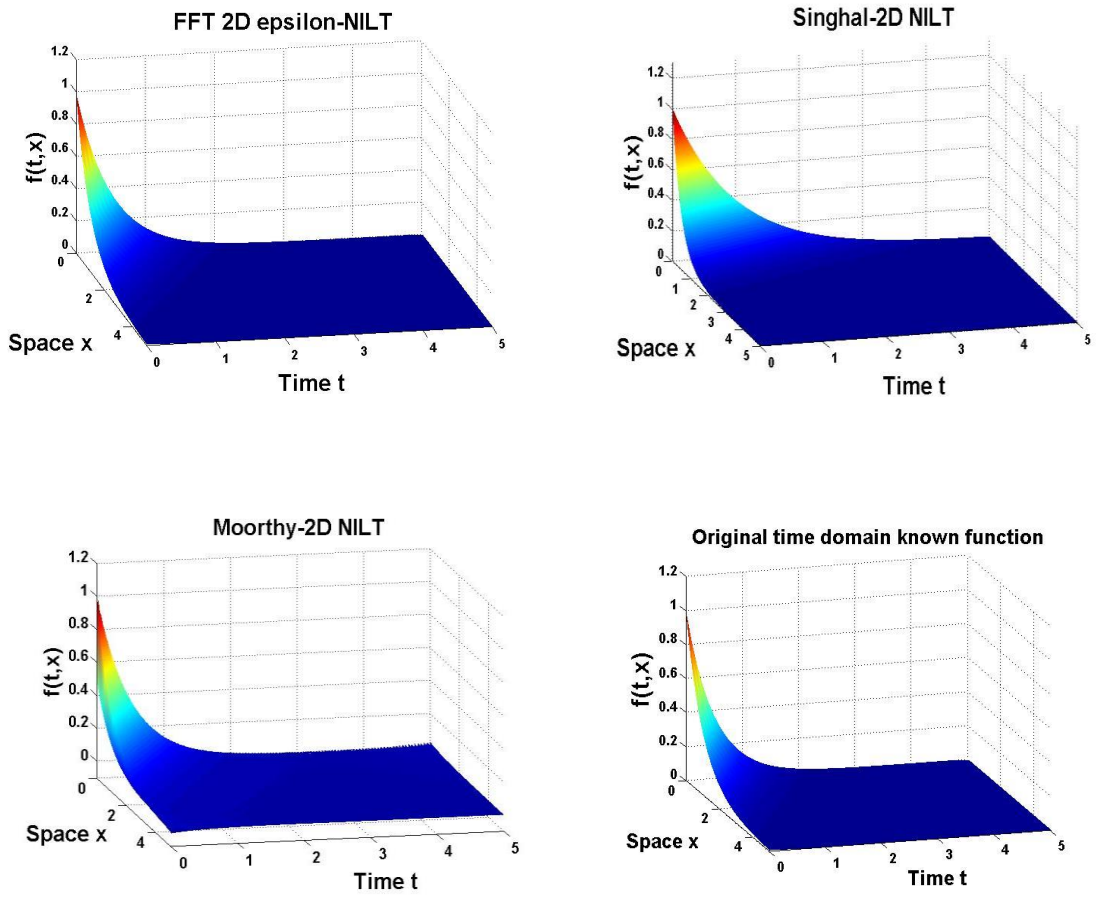


Figure 1: Results of the inverted function 1 with different NILTs and the known original

4. CONCLUSION

Three methods for 2D NILTs are described, implemented and compared in the paper. The NILT methods presented are devised based on using Fourier series techniques or Padé approximations which are classified as methods with good performance. The accuracy of the methods is emphasized by testing them with two functions that have previously known originals. We have programmed the methods and implemented them by using the universal mathematical language Matlab, and the results were shown in two tables and presented with the 3D plots of the inverted test function. The work is currently continued further by testing more methods with a larger number of functions of interest in electrical engineering along with the absolute error 3D plots and with further application of the methods on continuous space-time systems.

ACKNOWLEDGEMENT

Research disclosed in this paper was financially supported by the Czech Science Foundation under the grant 15-18288S; and for the research requirements, infrastructure of the SIX Center was used.

REFERENCES

[1] P. Gomez, L. Vergara, R. Nuricumbo-Guillen, and F. P. Espino-Cortes, "Two-Dimensional Definition of the Numerical Laplace Transform for Fast Computation of Transient Profiles Along Power Transmission Lines," *Power Delivery, IEEE Transactions on,* vol. 31, pp. 412-414, 2016.

[2] L. Brancik, "Improved method of numerical inversion of two-dimensional Laplace transforms for dynamical systems simulation," *Proceedings of ICECS 2002, The 9th IEEE International Conference on Electronics, Circuits and Systems,* Dubrovnik, vol.1, pp. 385-388, 2002.

[3] L. Brancik, "Utilization of NILTs in simulation of nonlinear systems described by Volterra series," *Przegląd Elektrotechniczny,* vol. 86, pp. 68-70, 2010.

[4] M. V. Moorthy, "Numerical inversion of two-dimensional Laplace transforms—Fourier series representation," *Applied Numerical Mathematics,* vol. 17, pp. 119-127, 1995.

[5] K. Singhal, J. Vlach, and M. Vlach, "Numerical inversion of multidimensional Laplace transform," *Proceedings of the IEEE,* vol. 63, pp. 1627-1628, 1975.

[6] L. Brancik, "An improvement of FFT-based numerical inversion of two-dimensional Laplace transforms by means of epsilon-algorithm," in *Circuits and Systems. Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium on*, vol.4, pp. 581-584, 2000.

[7] H. Dubner and J. Abate, "Numerical Inversion of Laplace Transforms by Relating Them to the Finite Fourier Cosine Transform," *J. ACM,* vol. 15, pp. 115-123, 1968.

[8] V. Zakian, "Numerical inversion of Laplace transform," *Electronics Letters,* vol. 5, pp. 120-121, 1969.

ADJUSTABLE FRACTIONAL-ORDER HIGH-PASS FILTER WITH TRANSCONDUCTANCE AMPLIFIERS

Jan Dvorak, Lukas Langhammer
Doctoral Degree Programme (1, 4), FEEC BUT
E-mail: dvorakjan@phd.feec.vutbr.cz

Supervised by: Jan Jerabek
E-mail: jerabekj@feec.vutbr.cz

Abstract: This paper deals with proposal of a fractional-order high-pass filter with electronically tunable pole frequency and its order. The filter is designed using operational transconductance amplifiers (OTAs), current follower (CF) and adjustable current amplifiers (ACAs). Function of proposed filter was verified using simulations in PSpice simulator. The circuit features were verified in simulations for three values of order of the filter and four values of cut-off frequency. In this paper, possibility to tune the cut-off frequency using the values of passive components is also mentioned.

Keywords: current mode, electronically tunable, fractional-order, high-pass filter

1 INTRODUCTION

Fractional-order circuits are currently of great interest, therefore the researchers aim their attention also to this matter [1]-[6]. Many designers are interested about design of the fractional-order filters [1]-[4]. Besides design of the fractional-order circuits, the research also deals with design of a fractional-order elements [5], [6]. Fractional-order circuits can be used for electrotechnics, accurate control technology and also for biology [5].

The filter structures are characterized by a transfer function and an order of the filter. The order of the filter determines the slope of attenuation of given transfer function. The slope of attenuation is desribed by $20 \cdot n$ dB/dec, where n is an integer order. The slope of attenuation of the fractional-order filter is described by: $20 \cdot (n + \alpha)$ dB/dec, where α is a number in range $0 < \alpha < 1$ and n is an integer order ($n \geq 1$) [2].

There are two basic methods how to propose a fractional-order filter. The first method is based on a design of a fractional-order passive element. This element can be created by the approximation using a passive RC network [5]. The second method is an approximation of the fractional-order transfer function by an integer-order function. The transfer function contains a fractional-order operator s^{α} which is approximated by an integer-order function [1]-[4]. For the realization of the approximated fractional-order transfer function, the FLF (Follow the Leader Feedback) [1], [4] topology is used, for example. The advantage of this method is the possibility to use available active elements with controllable parameters. Then a proposed fractional-order filter can have an electronic tunable order or a cut-off frequency [4], for instance.

This paper deals with the proposal of the Butterworth fractional-order high-pass filter operating in the current mode. Apart from the design of the fractional-order high-pass filter [1], [2], a fractional-order low-pass filter proposal is most frequently used [1]-[5]. The fractional-order circuits are often proposed in the voltage mode [3], however, fractional-order filters working in the current mode can be also proposed [1], [4].

2 PROPOSAL OF FRACTIONAL-ORDER FILTER

This chapter describes the design of fractional-order high-pass Butterworth filter having order (1+α). General equations and basic topology of the proposed filter are also included.

The basic transfer function [2] of the fractional-order high-pass filter can be given by:

$$K_{1+\alpha}^{\mathrm{HP}}(s) = \frac{k_1}{\frac{1}{s^{\alpha+1}} + \frac{k_2}{s^{\alpha}} + k_3}, \tag{1}$$

where $k_1 = 1$, $k_2 = 1.0683\alpha^2 + 0.161\alpha + 0.3324$ and $k_3 = 0.2917\alpha + 0.71216$. The coefficients k_1, k_2 and k_3 are used for shaping the pass-band region together with the desired fractional-order slope of attenuation of the stop-band [2]. The Butterworth approximation was selected to obtain a maximally flat pass-band response. Parameter s^{α} is fractional-order Laplacian operator which is described in [2], [3]. The following transfer function of fractional-order high-pass filter was created by using function in (1) and s^{α} approximation from [2]:

$$K_{1+\alpha}^{\mathrm{HP}}(s) \cong \frac{k_1}{a_0} \cdot \frac{s^3 a_0 + s^2 a_1 + s a_2}{s^3 + s^2 b_2 + s b_1 + b_0}, \tag{2}$$

where $a_0 = \alpha^2 + 3\alpha + 2$, $a_1 = 8 - 2\alpha^2$, $a_0 = \alpha^2 - 3\alpha + 2$, $b_2 = (a_1(k_3 + k_2) + a_2)/(a_2 k_3 + a_0 k_2)$, $b_1 = (a_0 k_3 + a_2 k_2 + a_1)/(a_2 k_3 + a_0 k_2)$ and $b_0 = a_0/(a_2 k_3 + a_0 k_2)$. The transfer function given in (2) was realized using FLF topology which is shown in Fig. 1. The circuit topology consists of three integrators, three feedback and three forward paths. The transfer function of this block topology is:

$$K(s) = \frac{I_{\mathrm{OUT}}}{I_{\mathrm{IN}}} = \frac{s^3 G_1 + s^2 \frac{G_2}{\tau_1} + s \frac{G_3}{\tau_1 \tau_2}}{s^3 + s^2 \frac{1}{\tau_1} + s \frac{1}{\tau_1 \tau_2} + \frac{1}{\tau_1 \tau_2 \tau_3}}. \tag{3}$$

Comparing particular polynomials from equations (2) and (3) can be determined G_x and τ_x parameters (It is obvious that parameter G_1 is equal $k_1 = 1$ for Butterworth approximation). Values of these parameters are important in order to calculate specific current gains and transconductances of the designed filter.

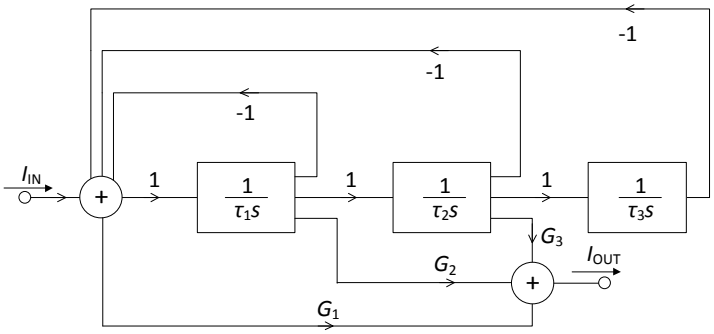


Figure 1: Block diagram of structure used for approximation of a fractional-order high-pass filter

The circuit structure of the proposed filter is shown in Fig. 2. This structure has been designed according to the schema which is illustrated in Fig. 1. The filter structure consists of three OTAs (Operational Transconductance Amplifiers), two ACAs (Adjustable Current Amplifiers), one auxiliary DO-CF (Dual–Output Current Follower) and three capacitors. OTA is described by following equation: $i_{\mathrm{OUT}+} = -i_{\mathrm{OUT}-} = g_m \cdot (u_{\mathrm{IN}+} - u_{\mathrm{IN}-})$, where parameter g_m is electronically controllable transconductance [7]. Current transfer of ACA is described by following relation: $i_{\mathrm{OUT}} = B \cdot i_{\mathrm{IN}}$, where B is current gain [8]. The last active element is DO–CF, which is described by equation: $i_{\mathrm{OUT}} = i_{\mathrm{IN}}$ [8].

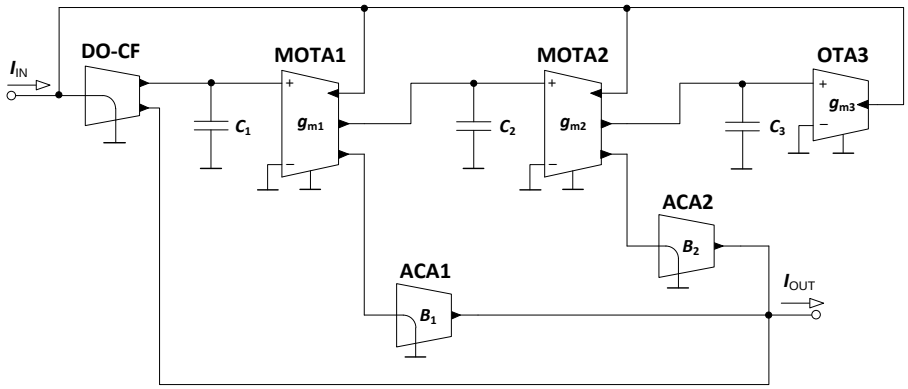


Figure 2: Circuit structure of the designed fractional-order high-pass filter

The transfer function of the circuit from Fig. 2 is:

$$K(s) = \frac{I_{\mathrm{OUT}}}{I_{\mathrm{IN}}} = \frac{s^3 + s^2 \frac{g_{\mathrm{m1}} B_1}{C_1} + s \frac{g_{\mathrm{m1}} g_{\mathrm{m2}} B_2}{C_1 C_2}}{s^3 + s^2 \frac{g_{\mathrm{m1}}}{C_1} + s \frac{g_{\mathrm{m1}} g_{\mathrm{m2}}}{C_1 C_2} + \frac{g_{\mathrm{m1}} g_{\mathrm{m2}} g_{\mathrm{m3}}}{C_1 C_2 C_3}}. \tag{4}$$

Values of the current gains, transconductances and passive elements can be calculated by comparing each polynomials in the equation (4) with polynomials in (3).

3 SIMULATION RESULTS

The proposed filter was simulated in PSpice. For simulation of the OTA and the DO-CF elements, a 3^{rd} level model of the UCC (Universal Current Conveyor) was used [9]. The ACA element was implemented by EL2082 chip [10]. Starting values of the simulations were chosen as follows: the cut-off frequency $f_0 = 100$ kHz, the order equals 1.5 and the capacitors $C_1 = C_2 = 470$ pF, $C_3 = 4.7$ nF. The order and the cut-off frequency of the proposed filter can be controlled electronically by changing values of the parameters of active elements. Adjustability of the order of the filter is independent to control of the cut-off frequency. The order is tuned by the current gains B_1, B_2 and the transconductances g_{m1}, g_{m2}, g_{m3}. The cut-off frequency can be tuned by changing of the values of transconductances g_{m1}, g_{m2} and g_{m3}, when keeping their ratio, in order maintaining the same order of the filter. Values of the parameters for three chosen orders (1.3, 1.5 and 1.7) are summarized in Table 1. The Table 2 summarizes required values of the parameters for selected cut-off frequencies (25, 50, 100 and 200 kHz).

Fig. 3 (a) illustrates the possibility to adjust order of the high-pass filter when the cut-off frequency was 100 kHz. It is obvious that simulations (solid lines) are almost identical with theoretical results (dashed lines). The slopes of attenuation of the theoretical results for selected orders are 26, 30 and 34 dB/dec. Values of the slope of attenuation obtained from simulations are 26.3, 31 and 35.4 dB/dec. The phase responses of the proposed high-pass filter for chosen orders are shown in Fig. 3 (b). Electronic tuning of the cut-off frequency for order 1.5 can be seen in Fig. 3 (c). The theoretical values (dashed lines) of the cut-off frequency are 25, 50, 100 and 200 kHz. Values of the cut-off frequency obtained from simulations (solid lines) are 28.2, 55.6, 106.4 and 208.4 kHz. The next possibility to adjust the cut-off frequency is by changing ratio of the values of capacitors. This can be seen in Fig. 3 (d). Selected values of cut-off frequencies are the same 25, 50, 100 and 200 kHz. To shift the cut-off frequency, values of all capacitors were multiplied by 4 for $f_0 = 25$ kHz and by 2 for $f_0 = 50$ kHz. To shift the cut-off frequency to $f_0 = 200$ kHz, values of the all capacitors were divided by 2. Values of the pole frequency obtained from simulation are: 27.7, 56, 106.4 and 209.9 kHz in this particular case. From magnitude and phase responses it can be seen that the results obtained from

simulations are in good agreement with theory. The obvious differences between the simulations and theory can be seen at high and low frequencies. This is mainly caused by parasitic properties and limited frequency range of the active elements (their models).

Table 1: Values of the parameters for three selected orders when $f_0 = 100$ kHz

Order	1.3	1.5	1.7
C_1 [pF]	470		
C_2 [pF]	470		
C_3 [nF]	4.7		
$1/g_{m1}$ [Ω]	894	1028	1158
$1/g_{m2}$ [Ω]	3377	3387	3439
$1/g_{m3}$ [Ω]	1051	840	768
B_1 [-]	0.693	0.607	0.515
B_2 [-]	0.116	0.069	0.033

Table 2: Values of the parameters for four values of cut-off frequency when order is 1.5

f_0 [kHz]	25	50	100	200
C_1 [pF]	470			
C_2 [pF]	470			
C_3 [nF]	4.7			
$1/g_{m1}$ [Ω]	4113	2056	1028	514
$1/g_{m2}$ [Ω]	13549	6774	3387	1694
$1/g_{m3}$ [Ω]	3361	1681	840	420
B_1 [-]	0.607			
B_2 [-]	0.069			

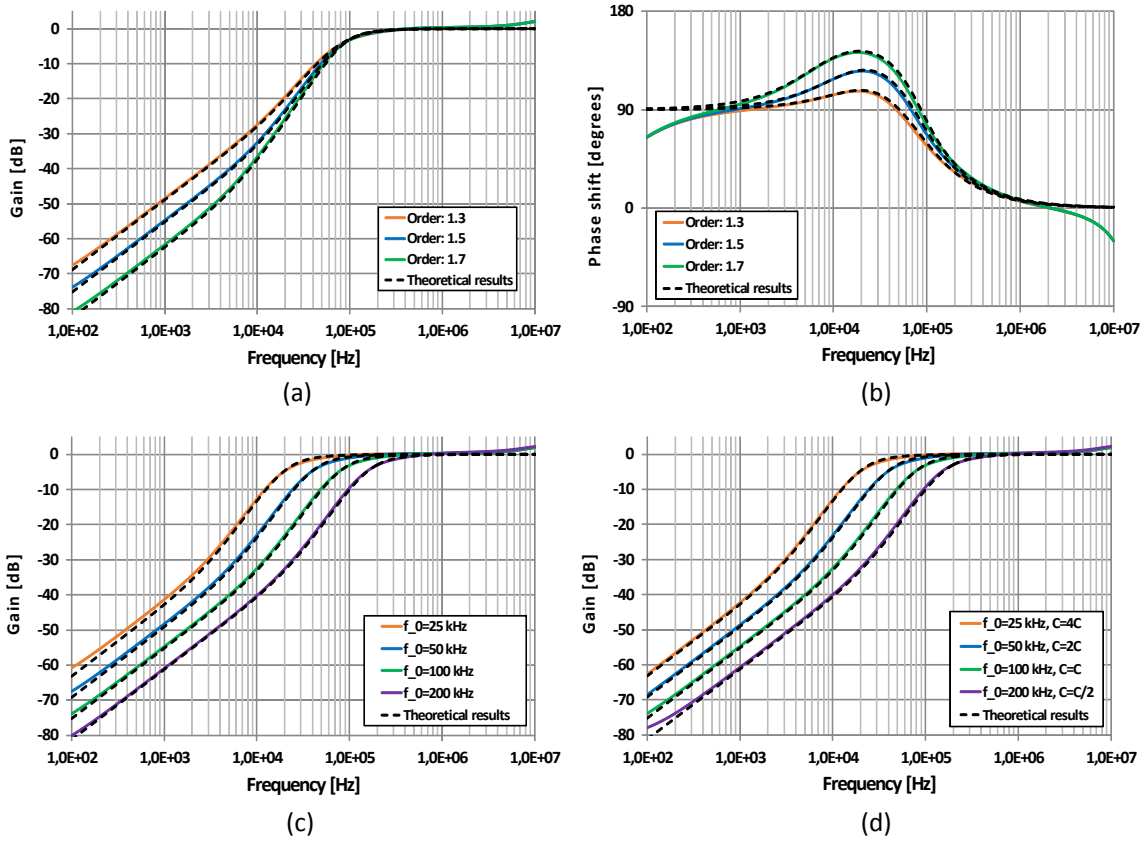


Figure 3: Simulation results of proposed fractional-order high-pass filter: (a) electronic control the order of filter by B_1, B_2, g_{m1}, g_{m2} and g_{m3} when $f_0 = 100$ kHz, (b) phase response for control the order of filter, (c) electronic control the cut-off frequency for order 1.5 by g_{m1}, g_{m2} and g_{m3}, (d) demonstration of possibility to control the cut-off frequency by changing values of capacitors

4 CONCLUSION

This paper describes the proposal of the fractional-order high-pass filter. The possibility of electronic control of the order and cut-off frequency is described. Control of the cut-off frequency by changing values of the capacitors is also mentioned. The function of fractional-order filter was verified using PSpice simulations. The obtained simulation results are compared with theory. It can be stated that the simulations results are very close to the theory for each selected value of the order and cut-off frequency of the proposed filter.

REFERENCES

[1] TSIRIMOKOU, G., C. LAOUDIAS and C. PSYCHALINOS. 2015. 0.5-V fractional-order companding filters. International Journal of Circuit Theory and Applications [online]. 43(9), 1105-1126 [cit. 2016-03-07]. DOI: 10.1002/cta.1995.

[2] FREEBORN, T. J., B. MAUNDY and A. S. ELWAKIL. 2010. Field programmable analogue array implementation of fractional step filters. IET Circuits, Devices [online]. 4(6), 514- [cit. 2015-12-28]. DOI: 10.1049/iet-cds.2010.0141.

[3] MAUNDY, B., A. S. ELWAKIL and T. J. FREEBORN. 2011. On the practical realization of higher-order filters with fractional stepping. Signal Processing [online]. 91(3), 484-491 [cit. 2015-12-28]. DOI: 10.1016/j.sigpro.2010.06.018.

[4] TSIRIMOKOU, G., C. PSYCHALINOS and A. S. ELWAKIL. 2015. Digitally programmed fractional-order Chebyshev filters realizations using current-mirrors. In: 2015 IEEE International Symposium on Circuits and Systems (ISCAS) [online]. IEEE, p. 2337-2340 [cit. 2016-03-07]. DOI: 10.1109/ISCAS.2015.7169152.

[5] ELWAKIL, A. 2010. Fractional-Order Circuits and Systems: An Emerging Interdisciplinary Research Area. IEEE Circuits and Systems Magazine [online]. 10(4), 40-50 [cit. 2016-03-07]. DOI: 10.1109/MCAS.2010.938637.

[6] SIVARAMA KRISHNA, M., S. DAS, K. BISWAS and B. GOSWAMI. 2011. Fabrication of a Fractional Order Capacitor With Desired Specifications: A Study on Process Identification and Characterization. IEEE Transactions on Electron Devices [online]. 58(11), 4067-4073 [cit. 2016-02-19]. DOI: 10.1109/TED.2011.2166763.

[7] JERABEK, J. and K. VRBA. 2010. Current-mode tunable and adjustable filter with digitally adjustable current amplifier and transconductance amplifiers. In: Proceedings of the European Conference of Circuits Technology and Devices (ECCTD'10) [online]. [cit. 2016-03-07]. ISBN 978-960-474-250-9.

[8] JERABEK, J., J. KOTON, R. SOTNER and K. VRBA. 2013. Adjustable band-pass filter with current active elements: two fully-differential and single-ended solutions. Analog Integrated Circuits and Signal Processing [online]. 74(1), 129-139 [cit. 2016-03-07]. DOI: 10.1007/s10470-012-9942-4.

[9] SPONAR, R. and K. VRBA. 2006. Measurements and Behavioral Modeling of Modern Conveyors. IJCSNS International Journal of Computer Science and Network Security [online]. 6(3A) [cit. 2016-03-07].

[10] EL2082 (Elantec) Current-mode multiplier (datasheet) [online]. 1996. [cit. 2016-03-07]. Available from: http://www.intersil.com/content/dam/Intersil/documents/el20/el2082.pdf

CURRENT-MODE MULTIFUNCTION FILTER UTILIZING LOSSES OF INDUCTOR SIMULATOR

Aslihan Kartci

Doctoral Degree Programme (1), FEEC BUT

E-mail: aslhankartc@gmail.com

Supervised by: Roman Sotner

E-mail: sotner@feec.vutbr.cz

Abstract: In this paper, a simple configuration for simulating the inductance using voltage differencing current conveyor (VDCC) as an active element has been presented. The proposed inductance simulator circuit uses only one VDCC, one current amplifier (CA), and one grounded capacitor. The equivalent value of the realized simulator can be tuned electronically through the transconductance parameter and current gain of the VDCC and current gain through the CA. The proposed circuit behavior together with its application is demonstrated using PSPICE simulations with commercially active devices. Functionality of the proposed circuit is verified through its application in the second-order circuit. The circuit that behaves as a filter, has one input and standard three output responses such as low-pass filter (LPF), band-pass filter (BPF), and high-pass filter (HPF).

Keywords: Multifunction filter, RLC circuit, 2nd-order circuit, voltage differencing current conveyor, VDCC, low-pass filter, band-pass filter, high-pass filter.

1. INTRODUCTION

Applications such as analog filters have benefited from the integration every module of a system in a single silicon. These implementations typically include capacitors that have also been part of circuit integration. Unfortunately, integrated inductors occupy significant space compared with other devices. For these reasons, designers have recently explored a wide diversity of techniques to magnify the inductive effect in order to reduce silicon area without compromising performance [1]. Some schemes offer very good accuracy, but limited multiplication factor, decreased effective bandwidth, increased power consumption, and silicon area; others have a large multiplication factor and effective bandwidth, although they have no accuracy and very limited dynamic range. In this brief, an inductance simulator with large output swing, large multiplication factor, high accuracy, and reduced silicon area is presented.

2. PROPOSED INDUCTOR SIMULATOR USING BENEFITS OF VDCC PARASITICS

The block diagram of a VDCC, ideally characterized by equations $I_Z = g_m(V_P - V_N)$, $V_Z = I_X R_X$, $I_{W_P} = I_X$, and $I_{W_N} = -I_X B_I$, is demonstrated in Figure 1, where P and N are input terminals, Z auxiliary terminal, X, W_P and W_N are output terminals. Except the X terminal, all of the terminals exhibit high impedance. In comparison to standard definition [2], the main advantage of this diagram is separated current transfer gain (B) control of the electronically controllable second-generation current conveyor (ECCII) from X to W_N of the output part of the VDCC.

For a complete analysis, it is important to take into account parasitics of active element shown in Figure 2. The parasitic resistances R_N, R_P and parasitic capacitances C_N, C_P appear between the high-impedance N (1 MΩ || 2.1 pF) and P (455 kΩ || 2.1 pF) input terminals of the VDCC and ground, respectively. The parasitic resistance R_Z and parasitic capacitance C_Z appear between high-impedance Z auxiliary terminal of the VDCC and ground, respectively, and their typical values are

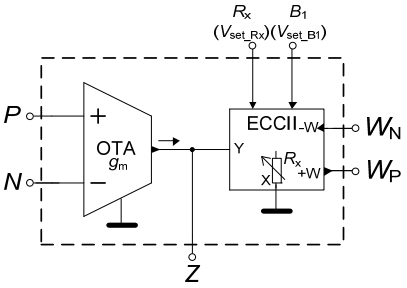


Figure 1: The block diagram of VDCC.

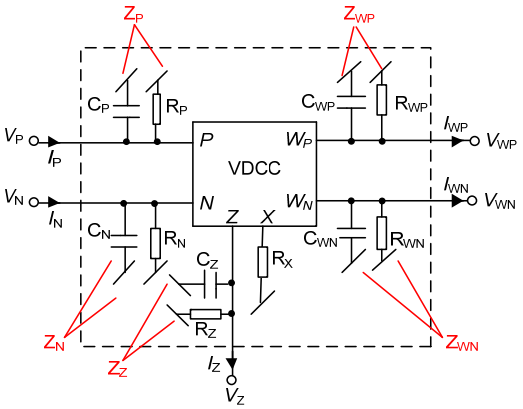


Figure 2: Model of the VDCC including parasitics elements.

48 kΩ || 5 pF. The parasitic resistance R_X appears at the low-impedance terminal X of the VDCC and its typical value is 95 Ω. The parasitic resistances and capacitances R_{W_P}, R_{W_N}, and C_{W_P}, C_{W_N} appear between the high-impedance W_P and W_N output terminals of the VDCC and ground, respectively, and their typical values are 1 MΩ || 5 pF.

2.1 Circuit Implementation with Commercially Available Active Devices

The proposed circuit for realizing a floating lossy electronically-controllable inductance simulator is shown in Figure 3. The circuit is constructed with single VDCC, one current amplifier, and one grounded capacitor. Its input impedance can be written as:

$$Z_{INP}(s) = R_{lossy} + sL_{eq} = \frac{1}{B_1}\left(\frac{B_2}{g_m} + s\frac{2CR_X}{g_m}\right) \tag{1}$$

Supposing $B_1 = 2$ instead of unity gain, equation (1) modifies to:

$$Z_{INP}(s) = \frac{B_2}{2g_m} + s\frac{CR_X}{g_m}. \tag{2}$$

Note that negative input of the OTA section is grounded therefore degradation resistor ($1/g_m$) in behavioral model in Figure 4 can be grounded without necessity to use diamond voltage buffer. The parasitic capacitance (C_{PWN} = 7.1 pF, caused by output impedance of EL2082 and input impedance of OPA860) connected to input terminal conjugated complex poles of external capacitances in frequency approximately about 8.3 MHz. Hence, equation (2) turns to the following expression:

$$Z_{INP}'(s) \cong \frac{B_2 + 2sCR_X}{g_m B_1 + sC_{PWN}B_2 + 2s^2 C_{PWN}CR_X}. \tag{3}$$

Simulations were performed using commercially available active devices such as current-mode two-quadrant multiplier EL2082 [3] (ECCII) with adjustable current gain (B_1) and one diamond transistor (OPA860 [4]) shown in Fig 4.

The supply voltages are selected as $V_{DD} = -V_{SS} = 5$ V. Note that the current from X terminal to output terminals (W_N, W_P) is divided equally as seen from general VDCC definition.

Equation for equivalent inductance with losses settings available from behavioral model has form:

$$L_{eq} = \frac{CR_X}{g_m} = \frac{CR_{Xinp}}{g_m(1-B_{Rx})} \cong \frac{CR_{Xinp}}{g_m(1-V_{SET\_Rx})}. \tag{4}$$

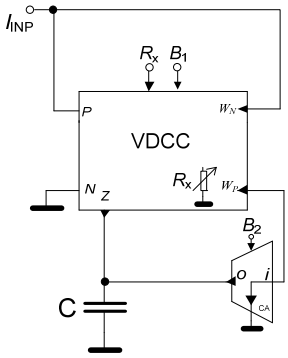


Figure 3: The proposed lossy inductance simulator.

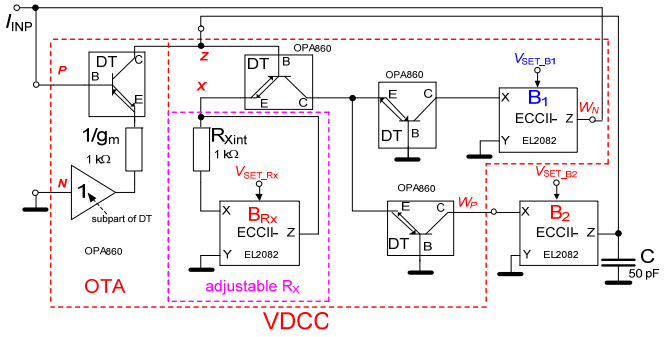


Figure 4: Circuit implementation with commercially available active devices.

Considering the effect of B_2 from short circuit impedance equation, the inductance simulator can be described as follows: (a) adjustable lossy inductance for $B_2 > 0$; (b) lossless inductance for $B_2 = 0$, and (c) adjustable inductance with negative losses for $B_2 < 0$. Descriptions (a) and (b) can be seen directly in equation (1), and (c) requires certain modification of ECCII. Initial value of lossy part is determined by $1/g_m \cong 1$ kΩ, while $B_2 = 1$. According to EL2082 datasheet, $B \cong V_{SET\_B}$. The adjustable R_X is the function of control voltage V_{SET_Rx}, which can be found as $R_X = R_{Xinp}/(1 - B_{Rx}) \cong R_{Xinp}/(1 - V_{SET\_Rx})$. Positive value of L_{eq} can be adjusted, while $1 > B_{Rx} \geq 0$ (i.e. $1 > V_{SET\_Rx} \geq 0$). Resistance of R_X can be set between 1 kΩ ($V_{SET\_Rx} = 0.1$ V) and 10 kΩ ($V_{SET\_Rx} = 0.9$ V).

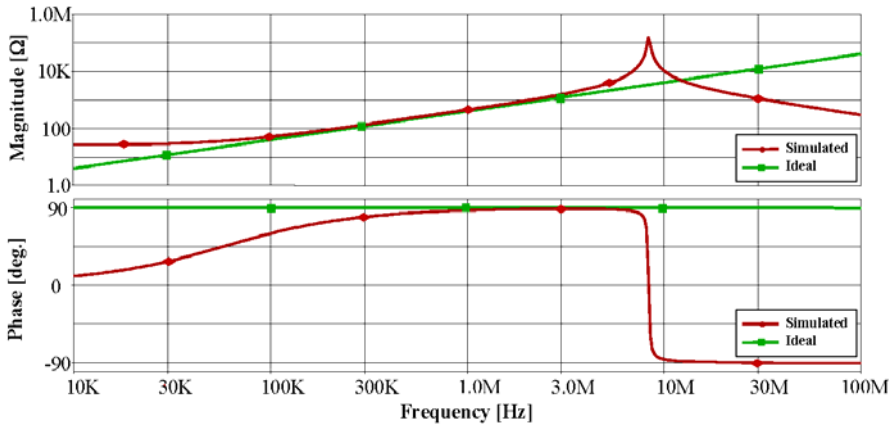


Figure 5: Magnitude and phase characteristics of the simulated inductance simulator using the behavioral model in Figure 4.

The proposed circuit shown in Figure 4 is simulated with the following passive element values: $C = 50$ pF, $1/g_m = R_{Xinp} = 1$ kΩ for minimum $V_{SET\_Rx} = 0$ V, which results in $L_{eq} = 50$ µH (70 µH simulated), and $1/g_m = R_{Xinp} = 10$ kΩ for maximum $V_{SET\_Rx} = 0.9$ V, which results in $L_{eq} = 500$ µH (531 µH simulated). In ideal case $B_2 = 0$ ($V_{SET\_B2} = 0$ V) and in maximum gain $B_2 = 3$ ($V_{SET\_B2} = 3.4$ V), the lossy resistance is found $R_{lossy} = 0$ Ω and $R_{lossy} = 1.5$ kΩ, respectively. The ideal and simulated impedance magnitudes and phase versus frequency for the proposed inductance simulator of Figure 4 are shown in Figure 5. Considering 5 deg. deviation in phase response with ideal phase response, it can be seen that the working frequency range of the proposed inductance simulator is from 300 kHz to 7.8 MHz.

The dependence of lossy part (R_{lossy}) by varying the B_2 (V_{SET_B2}) is shown in Figure 6. It is clear from the Figure 6 that the lossy part still exist even $B_2 = 0$ ($V_{SET\_B2} = 0$ V). The main reason is the grounded parasitic small-signal resistance R_Z in node Z (48 kΩ is the output resistance of DT OPA860), where the parasitic capacitance C_{PWN} connected. It effects on the equation (3) nu-

merator part as $B_2R_Z + 2R_X + 2sC_{PWN}R_XR_Z$. Negative value of B_2 is required to eliminate real part of the numerator from equation (3), but it is not possible with EL2082 used in our behavioral model.

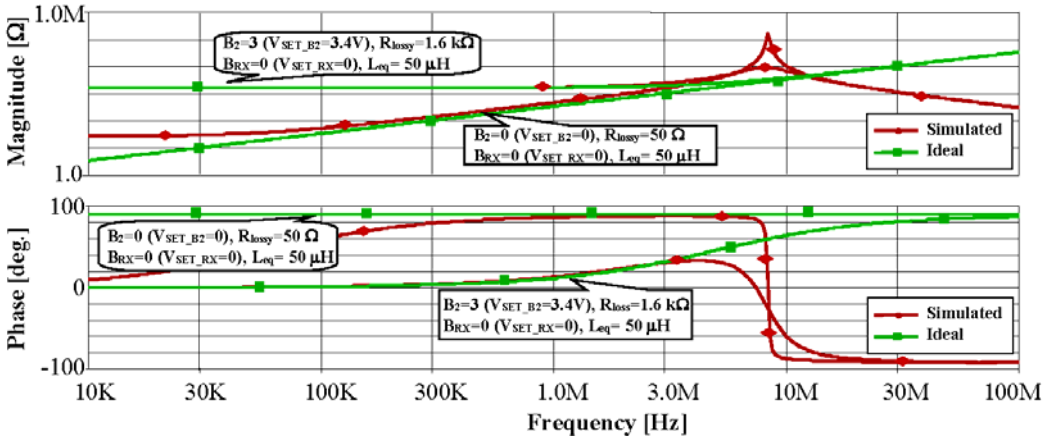


Figure 6: Change of losses in inductance simulator (Figure 4) for corner values of B_2: magnitude and phase responses.

3. REPLACEMENT OF INDUCTANCE SIMULATOR IN RLC CIRCUIT

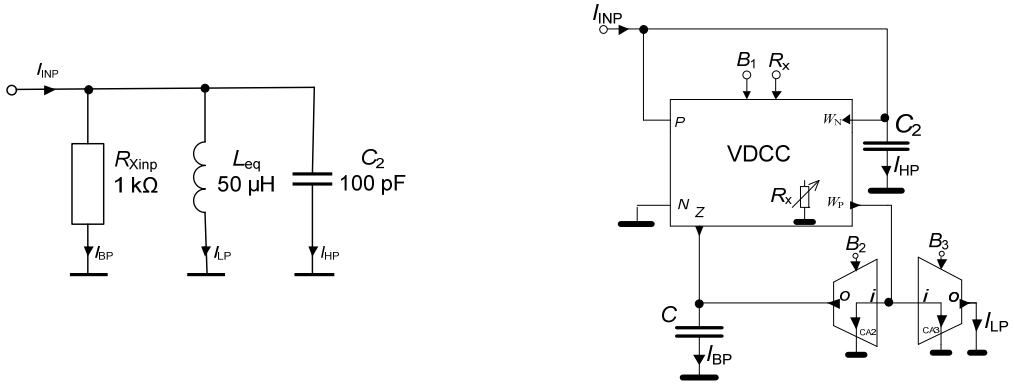
A simple passive filter can be obtained by using parallel and/or series R-L-C circuits. In addition, an active inductor circuit can be used instead of passive inductor in a filter circuit. As a result, an active filter can be constructed by making some modifications such as adding some passive components on active inductor circuits. Hence, the proposed lossy inductor of parallel R-L type can be easily used to make a multifunction 2nd-order filter providing electronically adjustable parameters as demonstrated in Figure 7, which is based on the formulation of a parallel RLC circuit formed by an inductor, a capacitor, and a resistor.

A routine circuit analysis of the below topology yields its characteristic equation as:

$$D(s) = 4s^2CC_2R_{xinp} + sC_2B_2 + 4g_m,\ K_{HP}(s) = -\frac{4s^2R_{xinp}CC_2}{D(s)},\ K_{LP}(s) = -\frac{g_mB_3}{D(s)},\ K_{BP}(s) = \frac{4sCR_{xinp}g_m}{D(s)}. \quad (5)$$

From the above characteristic equation, the natural frequency and quality factor (Q) are found to be:

$$\omega_0 = \sqrt{\frac{g_m}{CC_2R_{xinp}}} \text{ and } Q = \frac{1}{B_2}\sqrt{\frac{CR_{xinp}g_m}{C_2}}. \quad (6)$$



(a) (b)

Figure 7: (a) Model of multifunction filter using a simulated parallel RL type inductor, (b) the filter circuit using the proposed lossy inductor based on (a).

To verify the theoretical analysis, the behavior of the inductance simulator circuit from Figure 4 and RLC circuit used as a multifunctional filter shown in Figure 7 have been verified by PSPICE simulations using commercially available active devices. Current-mode multipliers EL2082 [3] as adjustable current amplifier (*B*) and one diamond transistor (OPA860 [4]) to build operational transconductance amplifiers (OTAs) are used for implementation of Figure 7 (a). Recommended supply voltage for the devices is ±5 V. Values of passive elements are selected as shown in Figure 7 (a). Initial setting of the filter was given by g_{m1} = 1 mS.

AC analysis of the proposed multifunction filter outputs such as low-pass, band-pass, and high-pass responses are given in Figure 8. The frequency response of the HPF output current through C_2, BPF output current through C, and LPF output current through output of additional current amplifier (CA_3) are obtained for the filter. These passive component values result in $Q \approx 2.08$ and $f_0 \approx 1.805$ MHz.

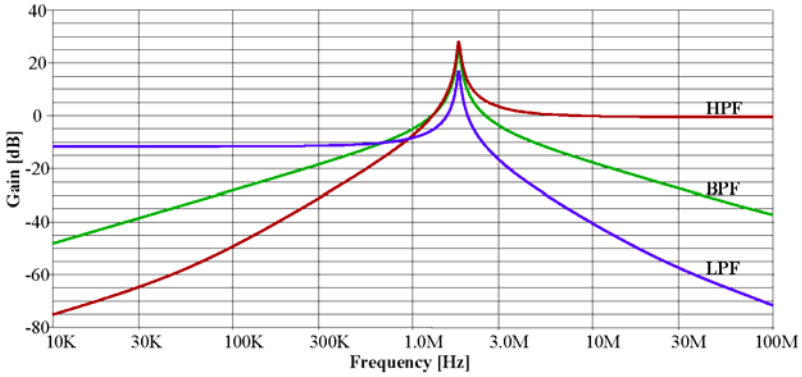


Figure 8: AC simulation results for the multifunction filter.

4. CONCLUSION

A new lossy inductor (parallel R-L type) using only a single VDCC along with only one capacitor and current amplifier has been presented. As an application, it is used to realize a 2nd-order circuit simultaneously providing BPF, LPF, and HPF responses by adding extra capacitor to the inductance simulator. The PSPICE simulations based on commercially available active devices to construct VDCC are included, which confirm the practical workability of the proposed lossy inductor and the multifunction filter circuit.

ACKNOWLEDGEMENT

This paper was supported by project no. FEKT-S-14-2281.

REFERENCES

[1] Kartci, A., Ayten, U. E., Herencsar, N., Sotner, R., Jerabek, J., Vrba, K.: Floating Capacitance Multiplier Simulator for Grounded RC Colpitts Oscillator Design. In Proceedings of the International Conference on Applied Electronics (AE), Pilsen, pp. 93-96, 2015.

[2] Biolek, D., Senani, R., Biolkova, V., Kolka, Z.: Active elements for analog signal processing: classification, review, and new proposals. Radioengineering, vol. 17, pp. 15-32, 2008.

[3] Intersil (Elantec). EL2082 CN Current-mode multiplier (datasheet), 1996, 14 p., accessible on www: http://www.intersil.com/data/fn/fn7152.pdf

[4] OPA860: Wide Bandwidth Operational Transconductance Amplifier and Buffer, Texas Instruments [online], 2005, last modified 8/2008. available at URL: http://www.ti.com/lit/ds/symlink/opa860.pdf

DIGITAL PREDISTORTION EXPERIMENTS WITH MATLAB-CONTROLLED SOFTWARE DEFINED RADIO

Martin Pospíšil
Doctoral Degree Programme (4), FEEC BUT
E-mail: xpospi29@stud.feec.vutbr.cz

Supervised by: Roman Maršálek, Tomáš Gotthans
E-mail: marsaler@feec.vutbr.cz

Abstract: This paper presents the experiments with digital predistortion for power amplifier linearization. It describes the architecture of used measurement setup based on the software defined radio USRP N200 with the focus on our custom MATLAB-based control of the USRP through UDP packets. Measurement results obtained for the 800 MHz power amplifier linearized by memory polynomial predistorter are presented, but the prepared framework to control of USRP in MATLAB can of course be used for experiments not only with such type of predistorter but for wide range of experiments where the USRP serves as the receiving/transmitting unit.

1. INTRODUCTION

Signals with non-constant envelope are often used in current wireless communication systems in order to fulfill demands for high data rates. In order to gain better Power Amplifier (PA) efficiency, the amplifiers are often used close to the saturation point. This leads to significant nonlinear distortion of the transmitted signals with undesirable effects on the increased bit error rate and spurious transmission into adjacent channels.

One of the solutions to combat the power amplifier effects is a digital predistortion (DPD) of PA, usually performed in the complex baseband domain. The first experiments with digital predistortion date back to 1980's [1]. DPD principle is based on the artificial distortion of the transmitted signal with a nonlinear function modeling the inverse of PA. The inverse function can be implemented in several ways differing in the complexity as well as in the performance. The simplest way is based on the Look Up Table [2] or polynomial function approximation. These two approaches cannot compensate for the memory effects in PA, significant especially for the wide-band input signals. In order to count for the memory effects, approaches like memory polynomials, Volterra series or generalized memory polynomials have been proposed. DPD is often adaptive - the inverse function adapts to PA changes caused by changing carrier frequency, power or temperature variations. DPD performance can be affected by RF impairments, e.g. IQ imbalances that can be corrected for, [3].

Currently chips for DPD [4] or specialized IP cores [5] are also available. Unfortunately, these solutions are not easy to use in the commercial software defined radios with small FPGA devices. To facilitate the experiments with DPD, it is desirable to be able to perform the algorithmic part in the simulation software (e.g. MATLAB) used in connection to available RF equipment.

2. PRINCIPLE OF ADAPTIVE DPD

The basic principle of adaptive DPD s is shown in left part of Figure 1. We consider the baseband DPD, where a transmitted baseband signal is upconverted to desired carrier frequency. A part of the transmitted signal is fed back to the baseband through a directional coupler to have a reference

signal for DPD adaptation. Such situation perfectly matches with a device available at a workplace - software defined radio USRP N200 populated with WBX front-end.

2.1. USED PREDISTORTER AND ADAPTATION METHOD

Several ways to adapt DPD block exists, the main two being a direct and an indirect method. In our example experiment presented below, the adaptation of DPD has been done using indirect-learning Recursive Least Squares (RLS) method, according to block schematic on right part of Fig. 1. The blocks of N data samples from the PA input and output are transferred from USRP to MATLAB, delay compensated [6] and DPD coefficients are computed from these blocks. Then the original data signal is processed by DPD and predistorted signal is uploaded to USRP for transmission.

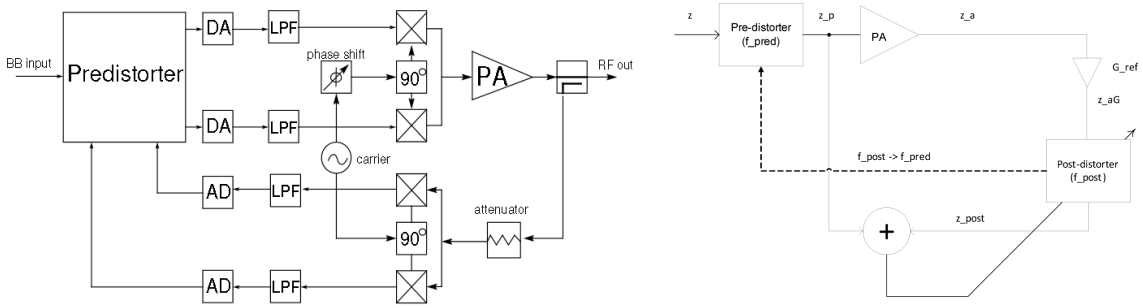


Figure 1: Digital predistortion principle (left) and indirect learning architecture (right)

From among a variety of predistorting functions, we have chosen Memory Polynomials with non-linearity order K and memory depth P due to its good compromise between the linearization performance and complexity. Given DPD coefficients f_{kp}, the expression for signal $z_p(n)$ at the output of DPD driven by the original data signal $z(n)$ at time instant n is then:

$$z_p(n) = \sum_{k=0}^{K} \sum_{p=0}^{P} f_{kp} z(n-p) \, |z(n-p)|^k \tag{1}$$

3. USRP AS A TRANSMITTING/RECEIVING DEVICE

The use of USRP from Ettus company (now National Instruments) is very popular among researchers in both academia and industry. In many cases, the USRP is used in connection with a host PC, and the more advanced data processing is done in PC using either GNU radio, GNU radio companion graphical environment, Simulink or LabView as in [7].

The main board of USRP N200 is equipped with the Xilinx Spartan 3A-DSP 1800 device and dual 14 bits A/D converters with 100 MSa./s. sampling rate. The USRP can be used in connection with various front-end modules ranging from DC to 6 GHz frequency bands. In the USRP, data from the antenna are received by the front-end module, converted to digital domain by the A/D converters and subsequently Digitally Down-Converted (DDC) to baseband. In the receiver, the user can access data from the front-end, input to the DDC, output of the DDC or baseband data. The transmitting chain is very similar. In our experiments we have worked with the final baseband data.

For the communication of the modules in the general USRP N2x0 design, an open source Wishbone bus is used. In this particular case the bus has 32-bit data width and 32-bit address space. There is one wishbone master in USRP design - the ZPU soft processor. The bus is configured to 16 slaves, but some of the slave positions are unused.

3.1. ARCHITECTURE CHANGES

In order to make use of USRP as a MATLAB-controlled device able to continuously transmit a vector of samples that drives the PA and to receive the PA output signal, some changes of the internal FPGA design had to be done. The overall architecture is shown on Fig. 2.

Two Block-RAMs (for transmitting-Tx and receiving-Rx path) are implemented in FPGA, both having a length of 4096 IQ samples. A multiplexer chooses between normal TX stream (i.e., from host PC) or data from TX BRAM.In this case, only TX BRAM is used as a transmission source. A counter is used to address port B on both BRAMs and is clocked with baseband sample rate. The counter operates in two modes: free-running or single pass. In the free-running mode, the signal from TX RAM is continuously repeated. It is used as normal output, for the measurements on the spectrum analyzer. The counter can be switched to other mode seamlessly, with no breaks in the transmitted signal. In the other mode, single pass, the counter stops at the end address. This ensures the received signal in the RX BRAM is one continuous block of data ready to be downloaded to MATLAB in a host PC.

On both BRAMs, the port A is connected to the Wishbone interface. This interface implements wishbone slave and it is connected to the normally unused position #4. Registers are mapped to the ZPU memory space starting with a base address of 0x5800. Data from BRAMs are accessible sample by sample only. To access the data, first the address register is written, then read or write operation is performed.

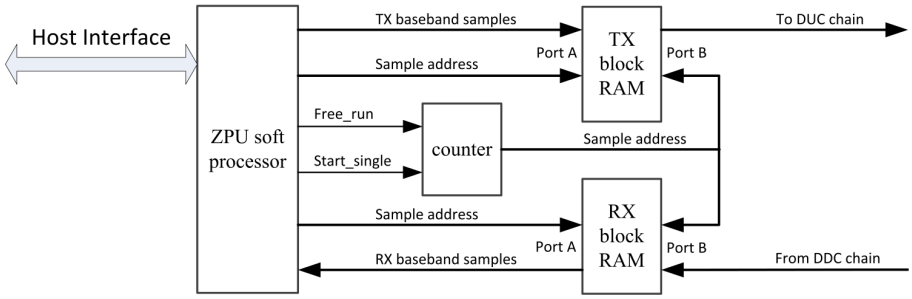


Figure 2: Block diagram of USRP N200 FPGA modifications that allow us to transmit and receive samples from MATLAB

3.2. MATLAB CONTROL OF USRP

Instead of using the standard communication with USRP using Simulink blocks, GNU radio or LabView software, we have concentrated on the USRP control from MATLAB using UDP packet operations in order to have the highest flexibility of control. For integration to MATLAB enviroment, UDP control library was written. It uses `java.net.DatagramSocket` and `java.net.DatagramPacket` classes for direct socket access.

In order to use the dynamic range of USRPs onboard A/D converters effectively and to be able to drive the power amplifier into the desired operational point, we needed to add a functionality to control the attenuators in the receiving and transmitting path. For direct control of the front-end attenuators from ZPU, a dedicated function needed to be written as this is normally only available from the standard UHD (Universal Hardware Driver) driver of USRP.

Three basic functions of our control library are thus now available, namely fill_bram, read_bram and control_set. Control_set function sets both attenuators and the control register. The data moving functions perform data type conversion, scaling and endian conversion. A new UDP listener is registered in the network stack on port 49180. The listen function implements a simple command protocol. First byte is the command number. Other bytes or packets carry data.

The stock UHD driver is still used to set-up the USRP. This is done by modified example application, written in C++. Data sent or received by streaming interface are discarded. UHD is used to set up sample rate (in our example setup set to 6,25 Msps) and carrier frequency (here 800 MHz).

4. MEASUREMENT SETUP AND RESULTS

During the experiment, USRP N200 was populated with WBX front-end module, whose output is amplified with the PA working in nonlinear region. At the PA output a directional coupler CP0603GN with -34dB coupling from AVX company has been used to get the attenuated part of PA output fed back to the USRP RX chain. The PA output was also monitored using a Rohde&Schwarz FSVR real-time spectrum and vector signal analyzer to get the output spectrum and to evaluate the adjacent channel protection ratio (ACPR) performance. The most important used components of WBX front-end and setup itself are summarized in Table 1.

The PA has been fed up with QPSK signal. The measured AM/AM characteristics are shown on Fig. 3. You can observe a dispersion of points in the characteristics caused possibly by the PA memory effects or incoherency of RX and TX oscillators. With the use of simple polynomial predistorter (memory depth P=0, nonlinearity order K=7), the ACPR improvement of around 12 dB's has been achieved as demonstrated on the spectral plots on right part of Fig. 3. Although such a performance is far from the theoretical values achieved in the pure simulations (20-30 dB) and marketing materials of DPD manufacturers, it has to be stated that till now we did not pay attention to correction of other RF imperfections such as IQ imbalances. As shown in Fig. 4, the temperature-dependent IQ imbalances are present in the used front-end and have to be compensated for to assure the full DPD functionality.

Transmit path		Receive path	
AD9777	Dual channel DAC	HMC174MS8	MMIC RF switch
ADL5385	Mixer	MGA62563	MMIC low-noise amplifier
HMC472LP4	Programmable attenuator	HMC472LP4	Programmable attenuator
GVA-84+	Amplifier	MGA82563	MMIC amplifier
HMC174MS8	MMIC RF switch	ADL5387	Mixer
Tesla 217.33	Power amplifier (to be linearized)	ADA4937-2	IQ baseband amplifiers
		ADS62P44	Dual channel ADC

Table 1: Active devices (chips) in DPD chain

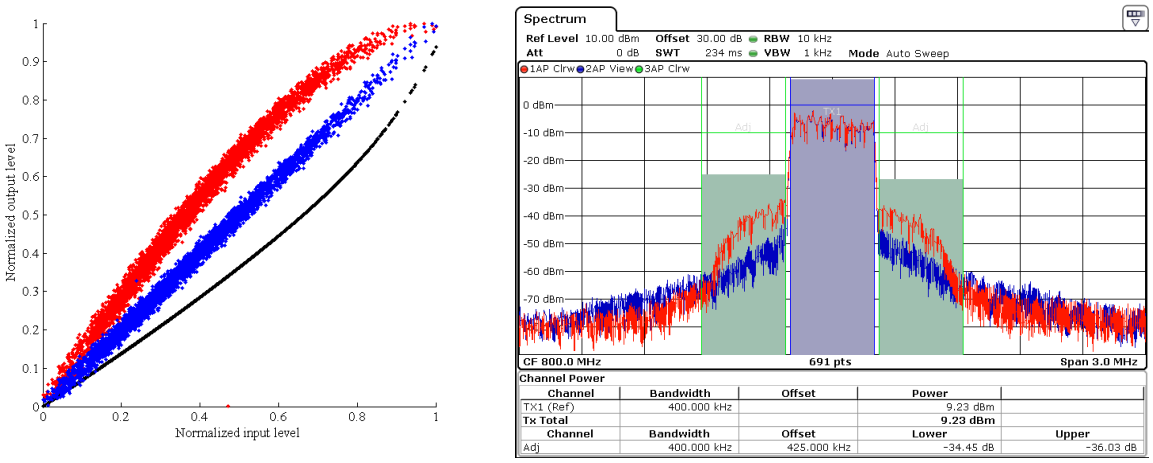


Figure 3: Left: AM/AM characteristics of PA (red), DPD (black) and model of DPD+PA (blue), Right: Measured spectra at PA output without (red) and with (blue) linearization

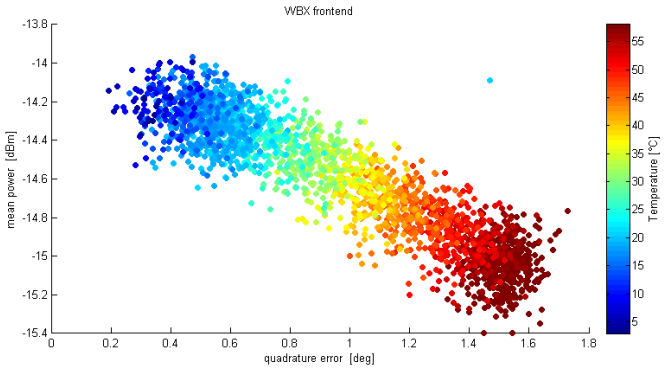


Figure 4: Transmit path temperature dependence.

5. CONCLUSION

The MATLAB-based control of USRP N200 software defined radio has been prepared. It allows to set the parameters such as Tx/Rx gain of this SDR, to transfer a signal created in MATLAB to internal memory of FPGA inside USRP for future playback, and to download received signal from FPGA internal memory back to MATLAB. Created MATLAB framework has been used to experiments with digital predistortion of power amplifiers based on memory polynomials. The improvement of up to 12 dB in term of ACPR has been achieved with the selected power ampliifier working at 800 MHz. The performance is significantly affected by the receiver impairments and further work would be to employ the corresponding compensation methods to further improve the predistorter's performance.

ACKNOWLEDGEMENT

The presented research has been supported by the BUT internal projects FEKT-S-14-2281 and FEKT-S-14-2177.

REFERENCES

[1] Nagata, Y., Linear amplification technique for digital mobile communications, In *IEEE 39th Vehicular Technology Conference*, 1989, vol.1, p. 159-164.

[2] Marsalek, R., Pospisil, M., Evaluation of digital predistortion using the USRP N200 software defined radio transceiver, In *Proceedings of NORCHIP 2014*, 2014, Tampere, Finland, p. 1-4.

[3] Antilla, L. et al., Recursive learning-based joint digital predistorter for power amplifier and I/Q modulator impairments, *International Journal of Microwave and Wireless Technologies*, 2010, 2(2), p. 173-182.

[4] Intersil, *Adaptive Predistortion Unsing the ISL5239*, Application Note, September 2002

[5] Xilinx, *LogiCORE IP Digital Pre-Distortion v5.0*, August 15, 2011.

[6] Gotthans, T.; Baudoin, G.; Mbaye, A., Influence of delay mismatch on digital predistortion for power amplifiers, in *20th International Conference Mixed Design of Integrated Circuits and Systems (MIXDES)*, 2013, p. 490-493.

[7] Dohl, J.; Fettweis, G., Evaluation of estimation and mitigation algorithms for nonlinearly distorted OFDM signals on a SDR platform, in *13th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2012 , p. 535-539.

DIRECTIVE DIELECTRIC RESONATOR ANTENNA

Michal Mrnka

Doctoral Degree Programme (3), FEEC BUT

E-mail: xmrnka01@stud.feec.vutbr.cz,

Supervised by: Zbyněk Raida

E-mail:raida@feec.vutbr.cz

Abstract: Relatively directive, single element, linearly polarized, dielectric resonator antenna for the ISM band applications (5.8 GHz) is presented in the paper. Its operation is based on the excitation of higher order modes inside a cylindrical dielectric resonator antenna of low relative permittivity material (ε_r= 6.15) fed by an aperture coupled microstrip line. The behavior of the mode properties inside the resonator is studied and geometry providing maximum gain is determined. Measurement results are given and compared with simulations results obtained in CST Microwave Studio in order to confirm the correct operation of the antenna.

Keywords: dielectric resonator antennas, directive antennas, dielectrics, higher order modes

1. INTRODUCTION

Dielectric resonator antennas (DRA) have become very popular radiating elements in microwave frequency bands [1]. The most important advantages of DRAs are high radiation efficiency, compact size, ease of excitation and a relatively large impedance bandwidth when compared to other resonant antenna elements, e.g. microstrip antennas.

A cylindrical resonator, operating with the low-order hybrid electromagnetic mode $HEM_{11\delta}$ placed above a large ground plane (theoretically infinite), is the most frequently used DRA configuration [1-3]. This mode generates a broadside radiation pattern with linear polarization and a gain of about 5 dBi. Several approaches have been suggested to increase the gain of the DRAs. Arraying of single element DRAs [2] is probably the most versatile method in which the gain value can be directly controlled by the number of elements in the array. Nevertheless, increased size, complexity and costs of the resultant antenna are the main disadvantages.

Altering a single element DRA can be used in cases, where medium gains up to around 10 dBi are sufficient. In general, two tactics to increase the gain of the single element DRA exist. First strategy is to use additional structures in close vicinity of the resonator operating in the low-order mode. These can represent surface mounted short horns [4], electromagnetic bandgap (EBG) structures [5] or superstrates [6].

The second strategy utilizes higher-order radiating modes in a single dielectric resonator. This approach has already been adopted in both rectangular and cylindrical DRAs. Petosa and Thirakoune [7] showed that a DRA based on higher-order $TE_{\delta 13}$ and $TE_{\delta 15}$ modes in a rectangular resonator can achieve gains of 8.2 dBi and10.2 dBi, respectively. The structure operating in $TE_{\delta 15}$ mode [7] re- quired a maximum resonator dimension of about $1.1\lambda_0$ when built from dielectric material with rel- ative permittivity ε_r = 10, where λ_0 is the free space wavelength. Guha et al. [8] managed to excite higher-order $HEM_{12\delta}$ mode in a cylindrical resonator by introducing an air-filled cavity in the ground plane below the resonator. This way, peak gain of about 10 dBi was achieved but only in a relatively narrow impedance bandwidth.

In this paper, we use excitation of a higher-order hybrid electromagnetic HEM_{133} mode in a single cylindrical dielectric resonator with partial excitation of the nearby HEM_{123} mode to accomplish

considerable gain enhancement. These modes require neither special feeding schemes nor ground plane modifications and can be excited similarly to the well-known $HEM_{11\delta}$ mode. The aperture coupling feed is selected and optimized to avoid excitation of unwanted TE/TM modes within the resonator. More detailed information on the topic can be found in [9].

The paper is organized as follows. Section II briefly describes the antenna concepts and the design process. The E-field distributions within the resonator are discussed. A simple parametric study and explanation of dimensions are given in Section III. Manufacturing of the prototype, final dimensions and a comparison between simulation and experimental results are summarized in Section IV. Finally, the paper is concluded by a short discussion in Section V.

2. CONCEPT OF THE PROPOSED ANTENNA

The antenna is composed of a single cylindrical dielectric resonator placed above the ground plane of a circular footprint. The resonator is excited through a rectangular slot in the ground plane of a microstrip line according to Fig. 1. An aperture coupling feeding mechanism was selected to minimize excitation of unwanted lower-order modes in the structure. Relative permittivity of the resonator in all simulations was 6.15. This value was simply selected due to availability of the material with the given ε_r. The substrate Arlon 25N with relative permittivity 3.38 was used in the feeding structure design.

Initial dimensions of the resonator (height *h* and diameter *d* according to Fig. 4) were found by the magnetic wall method using CST Eigenmode solver. All the walls were perfect magnetic conductors and the resonator dimensions were obtained so that the resonant frequency of both the target modes HEM_{133}, and HEM_{123} lied close to the desired frequency 5.8 GHz. Since the magnetic wall method does not take radiation losses of the resonator into account, considerably large inaccuracy is a result for low permittivity dielectrics. Nevertheless, the method provided a reasonable initial approximation of dimensions that needed to be tuned and optimized afterwards. Full-wave transient solver of CST Microwave Studio was used for this purpose.

Figure 2 depicts the simulated E-field and H-field distributions of the desired modes in several planes (see Fig. 1) of the resonator as obtained from the modal analysis and as excited within the designed DRA at target frequency 5.8 GHz. The comparison clearly shows the presence of the modes inside the DRA. Moreover, the modes are excited with quadrature phase shift which simplifies their identification. The E-field oscillates predominantly in the *x*-axis and thus the radiation is linearly polarized in the broadside direction (the direction along the *z*-axis).

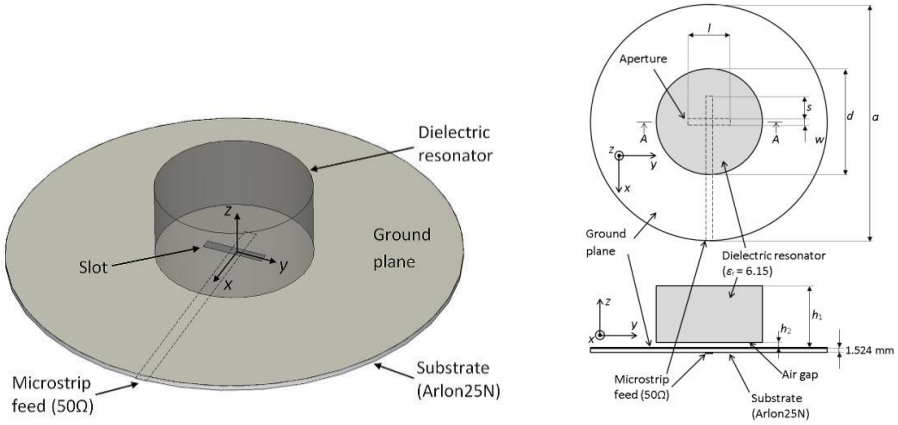


Figure 1: Geometry of the aperture-fed cylindrical (left) and its dimensions explained (right).

An aperture coupling excitation scheme is selected to excite both the desired modes in the DRA. This solution is preferred since the desired modes are of higher order and e.g. a probe feed would give rise to the excitation of unwanted TE/TM modes, deteriorating the near field distribution. The aperture is fed by a 50 Ω microstrip line placed on the other side of the ground plane according to

Fig. 4. By placing the aperture symmetrically below the DRA and tuning its dimensions, we can excite both the de- sired modes with a single feed. Correct values of parameters *l*, *w* and *s* have to be found. The length of the slot *l* is selected, so that the resonances of the slot are avoided. The optimum dimensions of the feed were found to be *l* = 17.73 mm, *w* = 3.49 mm with the stub length *s* = 10.94 mm.

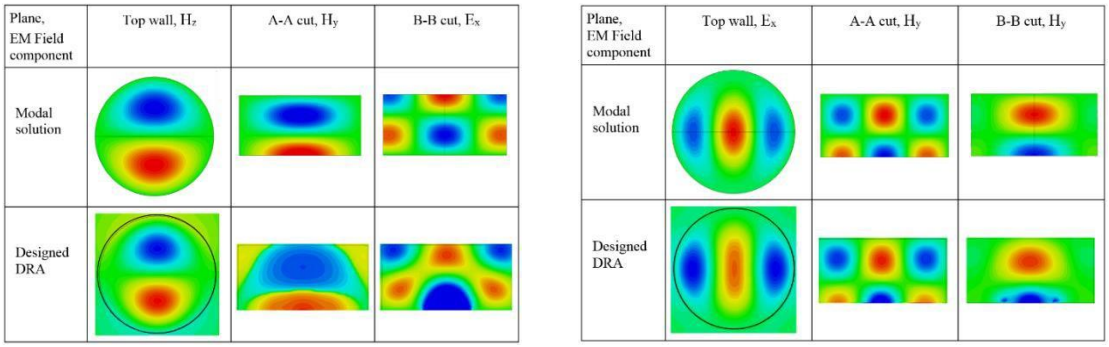


Figure 2: E-field and H-field distributions of the HEM_{133} mode (left) and HEM_{123} mode (right) as obtained by modal analysis and the same field components as excited in our cylindrical DRA at 5.8 GHz. The phase difference between the visualized E and H-field is 180°. For plane cuts see Fig. 1.

3. PARAMETRIC STUDY

The influence of two main design parameters of the designed DRA on antenna performance was studied. Namely the resonator diameter *d* and the ground plane diameter *a* were varied and the resonant frequency, impedance bandwidth and maximum gain were observed. Due to the restrictions of the antenna fabrication method, the height *h* must be an integral multiple of 1.575 mm only, which was the height of the Arlon substrate used for manufacturing (for details, see Section 4). The optimum height *h* was found to be 22.05 mm corresponding to 14 layers of the Arlon substrate. The desired resonant frequency was 5.8 GHz with the band of interest covering the frequency range 5.725-5.875 GHz. The band of interest corresponds to the relative bandwidth of 2.6 %. Throughout the parametric study, the dimensions of the feeding structure (i.e. the slot width *w*, its length *l* and the stub length *s*) were kept constant as well as the height *h* of the resonator.

The resonant frequency of the DRA and its impedance bandwidth were determined mostly by the dimensions of the resonator whereas its gain was in addition quite strongly influenced by the diameter of the ground plane. Fig. 3 shows the frequency response of the reflection coefficient for several *d* values (the reference impedance 50 Ω) with the ground plane size *a* = 132.5 mm. The height of the resonator *h* was fixed to the value 22.05 mm in all of the plots. Fig. 3 depicts the frequency response of reflection coefficient magnitude and frequency response of realized gain for several diameters of the resonator.

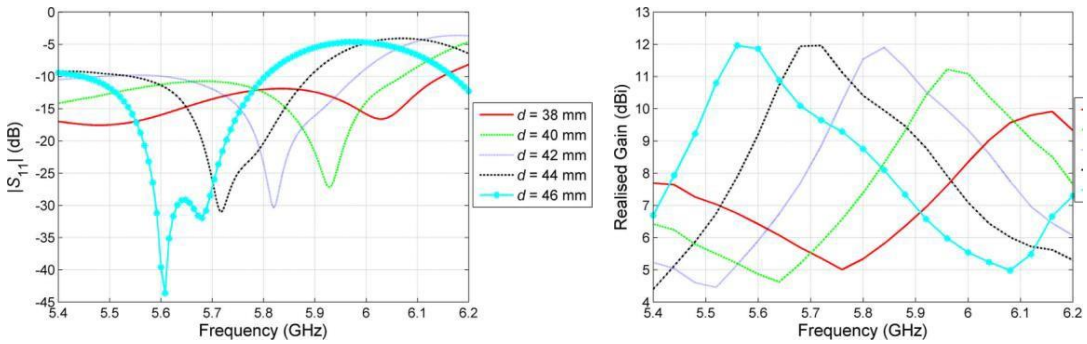


Figure 3: Reflection coefficient vs. frequency for several diameters of the resonator (left) and realized peak gain in the broadside direction vs. frequency for several diameters of the resonator (right).

4. PROTOTYPES AND RESULTS

The antenna was built by stacking up layers of a completely etched Arlon 600 substrate with $\varepsilon_r = 6.15$, loss tangent $\tan\delta = 0.0015$ and thickness 1.575 mm. The layers were held together by double sided duct tape of 65 μm thickness and relative permittivity 3, approximately. The resonant frequency had to be slightly tuned by altering the resonator diameter to compensate the effect of 14 layers of the duct tape in between the dielectric layers. The final dimensions of the resonator were h = 22.75 mm corresponding to 14 layers of Arlon 600 substrate interleaved with 14 layers of duct tape. The diameter of the resonator was 44.1 mm and the diameter of the ground plane a = 132.5 mm.

The comparison between the simulated and measured magnitude of the reflection coefficient is depicted in Fig. 4; reasonable agreement can be observed. Simulations were conducted in the transient solver CST Microwave Studio and measurement was done using a vector network analyzer.

Next, radiation patterns and gain of the designed antenna were measured in an anechoic chamber. First, the realized gain frequency response was measured in the broadside direction; a maximum gain of 11.59 dBi was obtained at the frequency 5.82 GHz (Fig. 4). Radiation patterns in two orthogonal principal planes (*xz*, *yz* in Fig. 1) were measured at the frequency of maximum gain 5.82 GHz for co- and cross-polarization components (Fig. 5). The co-polarization component corresponded to the *x* axis and the cross-polarization to the *y* axis according to Fig. 1.

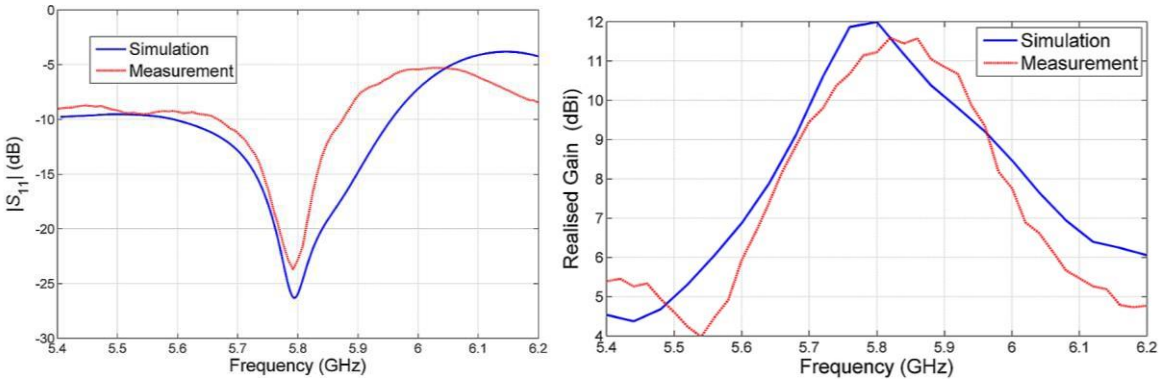


Figure 4: Comparison between simulated and measured magnitude of the reflection coefficient (left) and frequency response of the realized gain (right).

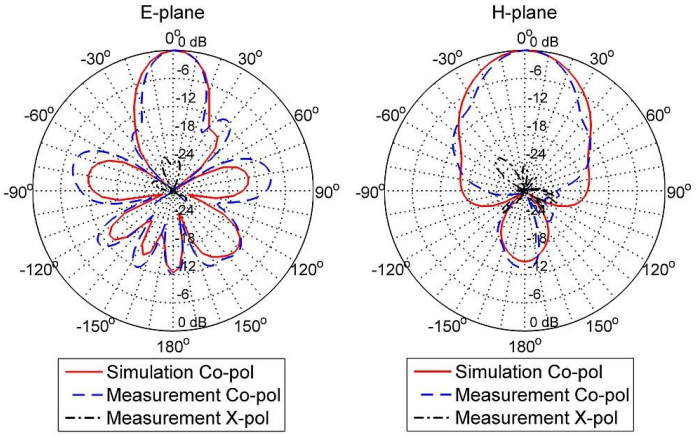


Figure 5: Simulated and measured radiation patterns at 5.82 GHz, E-plane corresponds to *xz* plane and H-plane to *yz* plane.

5. CONCLUSION

A concept of a dielectric resonator antenna based on the combination of higher-order modes in a simple cylindrical dielectric resonator was presented and experimentally verified. Good agreement between simulation and measurement was achieved. A directive radiation pattern with a high gain of 11.59 dBi was measured and impedance bandwidth was sufficient to cover the whole 5.8 GHz ISM band. Relatively high side lobe levels in the E-plane (8dB below the gain in the broadside direction) are the main disadvantage. The implementation of the antenna might be more convenient at higher frequencies where dimensions are less crucial. Since the structure operates with higher-order modes, the electrical size is increased when compared to the operation of low-order modes.

The proposed antenna concept might find application in all the areas of wireless communication where directive elements are required and where the bandwidth is not crucial, e.g. radio relay links. More complex shapes of the resonator might provide larger bandwidths and thus broaden the application range of this antenna element.

ACKNOWLEDGEMENT

Research described in this paper was financially supported by the Czech Science Foundation under grant no. P102/12/1274. Research is a part of the COST Action IC1301 supported by the grant LD14057 of the Czech Ministry of Education. The presented research was supported by the Internal Grant Agency of Brno University of Technology project no. FEKT-s-14-2483.For research, infrastructure of the SIX Center was used.

REFERENCES

[1] Long, S.A.; McAllister, M.; Liang Shen, "The resonant cylindrical dielectric cavity antenna," Antennas and Propagation, IEEE Transactions on , vol.31, no.3, pp.406,412, May 1983.

[2] Luk, K. M.; Leung, K. W. Dielectric Resonator Antennas. Research Studies Press Ltd. 2003.

[3] Kajfez, D.; Guillon, P. Dielectric Resonators2e. Noble Publishing Corporation Atlanta. 1998

[4] Nasimuddin; Esselle, K.P., "Antennas with dielectric resonators and surface mounted short horns for high gain and large bandwidth," Microwaves, Antennas & Propagation, IET , vol.1, no.3, pp.723,728, June 2007

[5] Denidni, T.A.; Coulibaly, Y.; Boutayeb, H., "Hybrid Dielectric Resonator Antenna With Circu- lar Mushroom-Like Structure for Gain Improvement," Antennas and Propagation, IEEE Trans- actions on , vol.57, no.4, pp.1043,1049, April 2009.

[6] Coulibaly, Y.; Nedil, M.; Ben Mabrouk, I.; Talbi, L.; Denidni, T.A., "High gain rectangular die- lectric resonator for broadband millimeter-waves underground communications," Electrical and Computer Engineering (CCECE), 2011 24th Canadian Conference on , vol., no., pp.001088,001091, 8-11 May 2011.

[7] Petosa, A.; Thirakoune, S., "Rectangular Dielectric Resonator Antennas With Enhanced Gain," Antennas and Propagation, IEEE Transactions on , vol.59, no.4, pp.1385,1389, April 2011

[8] Guha, D.; Banerjee, A.; Kumar, C.; Antar, Y.M.M., "New Technique to Excite Higher-Order Radiating Mode in a Cylindrical Dielectric Resonator Antenna," Antennas and Wireless Propa- gation Letters, IEEE , vol.13, no., pp.15,18, 2014

[9] M. Mrnka; Z. Raida, "Enhanced Gain Dielectric Resonator Antenna Based on the Combination of Higher Order Modes," in *IEEE Antennas and Wireless Propagation Letters* , vol.15, no., pp.710-713, 2016.

EDUCATIONAL MODEL OF SYNTHETIC APERTURE SONAR USING RANGE-DOPPLER ALGORITHM

Martin Bartoš

Doctoral Degree Programme (2), FEEC BUT

E-mail: xbarto85@stud.feec.vutbr.cz

Supervised by: Jiří Šebesta

E-mail: sebestaj@feec.vutbr.cz

Abstract: This paper deals with synthetic aperture imaging techniques in ultrasound spectrum. The paper provides description of Synthetic Aperture Sonar (SAS) experiment with low-cost components in laboratory environment. The Range-Doppler Algorithm for signal processing, which is one of the imaging algorithm, is also depicted in this paper. The theoretical resolution is compared with that of the experimental results.

Keywords: Experiment, SAS, RDA

1. INTRODUCTION

The technique known as synthetic aperture sonar (SAS)/radar (SAR) belongs to remote sensing systems, which can produce high-resolution images by moving a single sensor past a scene of interest along a track. SAS is based on SAR, which were developed and used earlier on spaceborn platforms. The basic difference between SAS and SAR is a type of wave (acoustic/electromagnetic) and from that reason speed of propagation in a given environment, carrier frequency, beam width or dimensions of transducer [1].

The fundamental principle of image reconstruction in the SAS/SAR system is based on post-processing of pulse responses which are digitized and stored in a memory. More often used SAS/SAR post-processing algorithms are range Doppler algorithm (RDA) [2], chirp scaling algorithm [2] and omega-K algorithm [2]. The RDA is presented in this paper. It has been selected for the multi-purpose SAR platform design due to its advantageous characteristics like a block processing efficiency.

Post-processing was realized in MATLAB. It was also used for defining the output signal which goes through a transducer. .

2. THE INSTRUMENTS

The main laboratory sonar parts and instruments characteristics are summarized here. The entire experiment system is based around a PC and external sound card and MATLAB, which depicts Figure 1. The sonar transmitted a LFM pulsed signal with a currier frequency 40 kHz with bandwidth 4 kHz. A propagation speed is approximately 340 m/s in air which causes quite a big limitation for defining Pulsed Repetition Frequency (PRF) – one of the most important SAS/SAR parameter.

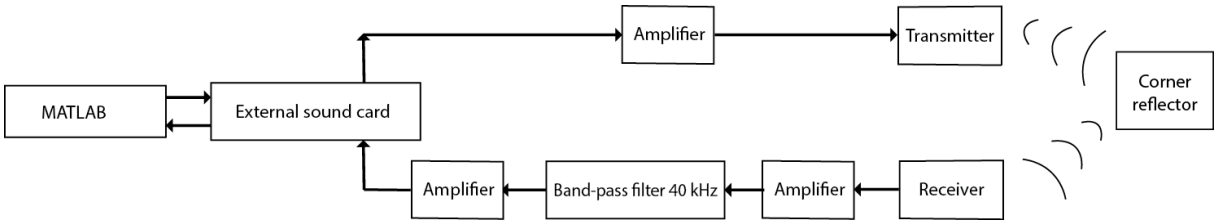


Figure 1: Block scheme

In MATLAB environment is defined LFM pulse with carrier frequency f_0 40 kHz, bandwidth B_0 4 kHz and pulse duration T_p 1 ns. MATLAB function `audioplayer` creates audio object for generating output signal from the sound card. Function `audiorecorder` solves inverse action for acquisition echoed signal. Both functions provide possibility to choose appropriate parameters such as given signal, sample rate or ID outputs and inputs.

The transmitter channel provides accurate amplification of the transmitted signal. It is composed of low noise JFET dual operational amplifiers TL072. Amplified signal is fed to transmitter transducer.

Echoed signal is received and modified in receiver channel. The band-pass filter with cutoff frequency 40 kHz is composed of two active low-pass filters and one band-pass filters. There are also amplifiers which increase Signal-Noise resolution. The attenuated signal by propagation and reflection is approximately amplified more than 1000 times.

For our experiment was used an external sound card Icon Cube because integrated sound card has not appropriate parameters. This card has 192 kHz sample rate, and its frequency response has been extended until 50 kHz by change of capacitors in output channel.

3. THE POST-PROCESSING

As was mentioned above, for the signal post-processing was chosen RDA, which was also created in MATLAB. Figure 2 illustrates the flowchart of the RDA process.

3.1. RANGE COMPRESSION

Compression of the echoed signal s_{rc} in the range direction is the first step of the RDA and this process can be described by:

$$s_{rc} = \text{IFFT}\{S_r(f_\tau, \eta)S_0(f_\tau\}, \tag{2}$$

where S_r is received signal in frequency domain and S0 is reference signal. Both signals are knocked under Inverse Fast Fourier Transform (IFFT). Increasing resolution in range can be gained by using of conventional windowing functions (Hann, Kaiser, etc) in addition to the rectangular envelope. The frequency matched filter is applied for each range line of received signal matrix.

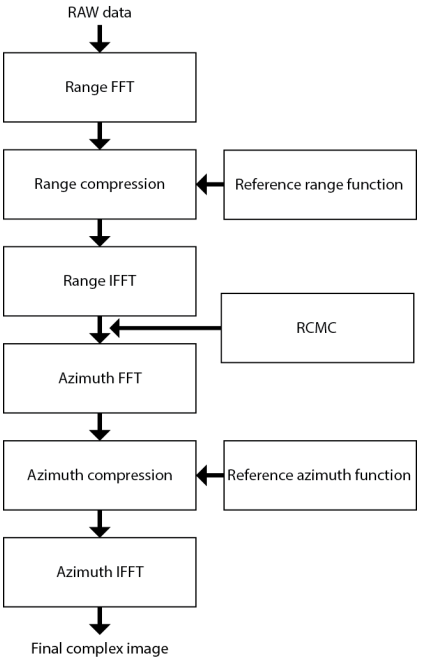


Figure 2: RDA diagram

3.2. RANGE CELL MIGRATION CORRECTION

Slant range $R_i(\eta)$ is changing due to movement of radar through synthetic aperture. This change causes additional Doppler shift. A hyperbolic form of instantaneous range (3) causes a phenomenon called range cell migration (RCM), which complicates the following signal processing.[2][3][4]

The contribution of RCM in frequency azimuth should be correct by:

$$\Delta R(f_\eta) = \frac{\lambda^2 R_0 f_\eta^2}{8V_r^2}, \tag{4}$$

where f_η is expressed as:

$$f_\eta = -K_a \eta, \tag{5}$$

and K_a is azimuth frequency rate, see (6). Note, the hyperbola (target trajectory) opens up with decreasing range in frequency domain, see (3)[2].

3.3. AZIMUTH COMPRESSION

After range compression and RCMC, the signal energy from target follows a trajectory in the two-dimensional matrix of echoed signal that depends on the changing range delay to the target as it passes through the antenna beam. In order to capture all signal energy from a point the target must be aligned in a signal range bin [2][3].

Azimuth compression is performed analogically to range compression. The difference is in the azimuth reference signal s_a:

$$s_a = exp\left\{-j\pi \frac{2V_g^2 \eta^2}{\lambda R_0}\right\} \tag{6}$$

The platform velocity V_g causes the (parasitic) phase modulation. Signal s_a has LFM characteristic and its frequency rate K_a is expressed as [3]:

$$K_a \approx \frac{2V_r^2}{\lambda R_0}. \tag{7}$$

4. THE EXPERIMENT AND RESULTS

The experiment was realized in a soundproofed laboratory. Transducers were placed approximately two meters from two reflectors. Transmitter and receiver were mounted on a rail on which were assessed align the track. The transmission was realized every 1 cm along the track (1 meter) to simulate accurate PRF.

Figure 3 depicts matrix of echoed signals. Peaks represent each reflection. There was a complication with synchronization of an acquisition time base, because CPU time is changing during each transmission of a signal. This latency causes CPU, memory, sound card, audio drivers and others. The latency can reach in order of tenths of second, which correspond to dozens of meters. For this reason was necessary to synchronize time base before RDA processing which results in final image of a scene.

To align all echoed signals was used threshold, because relatively low noise and high sample frequency allowed this approach. For this purpose can be also used built-in functions based on cross correlation, but since echoed signals are pretty different, I decided to use first mentioned approach.

Figure 4 gives better preview on received signal in range domain. There are three main peaks. First peak represents transmitted LFM pulse, second and third peak are caused by reflection from above mentioned targets. Figure 4 is composed of these signals. Due to MATLAB processing time was necessary to set acquisition time (recording time) to 1s, which is disproportionate to our scene and was also necessary to crop redundant part of the echoed matrix. The delay between main peaks corresponds to a distance between transmitter and targets.

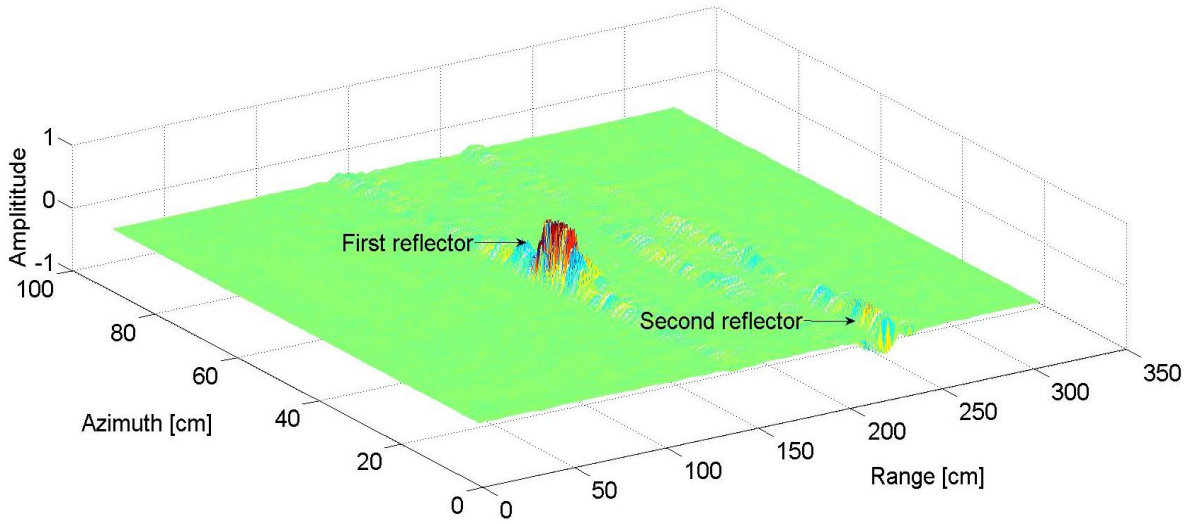


Figure 3: Echoed signals

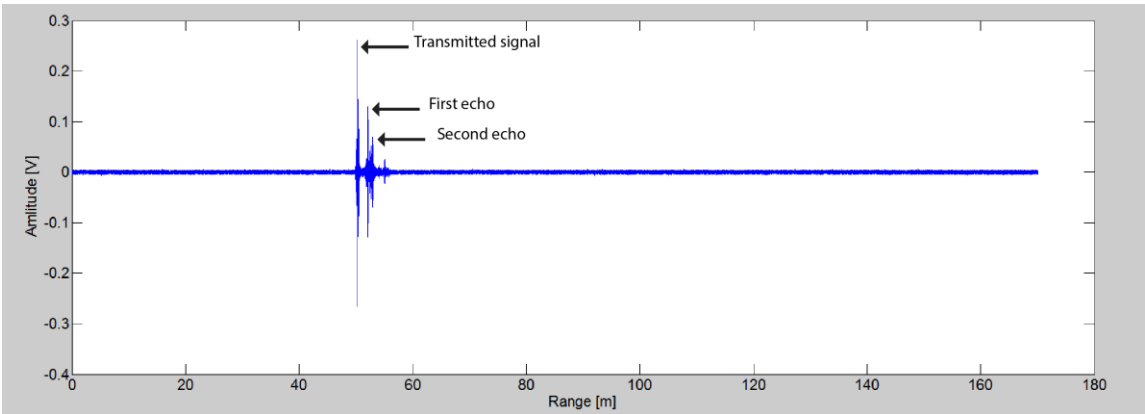


Figure 4: Received signal

5. CONCLUSION

Results show that the experiment confirms theoretical assumptions of the theme. The current phase of the experiment proves possibility to detect and resolve target in ultrasound spectrum. RDA was applied for the final image, which is discussed in third chapter. For the experiment, instruments used were described in second chapter. All measurements can be realized in Graphical Interface Unit in MATLAB, where a signal is generated and processed. Due to the difficult possibility of synchronization between generating and receiving time, input and output channel were operating simultaneously. Other task will carry out additional measurements, ideally outside in free space on bigger distances with metal scatters, which should ensure more transparent results. The final form of the laboratory model will be utilized for educational purposes.

ACKNOWLEDGEMENT

The presented research was performed in laboratories supported by the SIX project; no. CZ.1.05/2.1.00/03.0072, the operational program Research and Development for Innovation. Research described in this paper was also financed by Czech Ministry of Education in frame of National Sustainability Program under grant LO1401. Infrastructure of the SIX Center was used.

REFERENCES

[1] Wilkinson A. J., Mukhopadhyay P. K., Lewitton N., Inggs M. R. *Inverse Synthetic Imaging using a 40 kHz Ultrasonic Laboratory Sonar*. Grabouw, South Africa, 2004.

[2] Ouchi, Kazuo. *Recent Trend and Advance of Synthetic Aperture Radar with Selected Topics. Remote Sensing* [online]. 2013, DOI: 10.3390/rs5020716. ISSN 2072-4292

[3] Cumming, Ian G, Frank Hay-Chee Wong. *Digital processing of synthetic aperture radar data: algorithms and implementation*. Boston: Artech House, c2005, xxviii, 625 p. ISBN 15-805-3058-3.

[4] Shultz, Matthew. *Synthetic Aperture Radar Imaging Simulated in MATLAB*. San Luis Obispo, 2009. Master thesis. California Polytechnic State University.

H-PLANE SIW HOWRN ANTENNA FOR ON-BODY COMMUNICATION

Miroslav Cupal

Doctoral Degree Programme (1), FEEC BUT

E-mail: xcupal02@stud.feec.vutbr.cz

Supervised by: Zbyněk Raida

E-mail: raida@feec.vutbr.cz

Abstract: The substrate integrated waveguide (SIW) technology allows us to construct several types of low-profile antennas. One of these antennas is h-plane SIW horn antenna. However, this antenna shows constraints if thickness of the substrate is smaller than λ/10. A modified h-plane horn antenna to be used close to human body tissue is presented in this paper. Attention is paid to the decrease of the elevation angle of the main beam in E-plane. The antenna has been developed for wireless links of on body communication.

Keywords: Horn antenna, vertical polarization, low profile, substrate integrated waveguide.

1. INTRODUCTION

In this paper, we present substrate integrated h-plane horn antenna for point-to-point communication. The substrate integrated waveguide (SIW) technology is very popular in the field of microwave and millimeter-wave applications. The SIW is advantageous when integrating antennas and sub-circuits (filters, oscillators, amplifiers) into a single microwave substrate. Higher losses are a disadvantage of SIW technology [1], [2].

An h-plane horn antenna can be used for point-to-point communication since it is main lobe is relatively narrow in the H-plane. As a main problem, poor matching of the antenna can be considered if the substrate is thinner than $\lambda_0/10$. A wide main lobe in the E-plane and a lower gain compared to a pyramidal horn antenna are other disadvantages [3].

There are several ways of eliminating these disadvantages: In front of the aperture of the horn antenna, we can add non-planar dielectric elements, planar lenses or planar resonators. In the paper, we describe exploitation of planar resonators for improving impedance matching and bandwidth adjustment of the radiation pattern in the E-plane [4], [5]. In our case, the thickness of the substrate is smaller than $\lambda_0/20$. We have used the substrate Arlon CuClad 217 of thickness h = 1.524 mm and we have selected the center frequency f = 8 GHz. Since the substrate is very thin, there is a significant mismatch between the aperture of the horn and the air. This mismatch causes reflection at the interface and decreases the bandwidth. In order to increase the bandwidth of the antenna, a thicker substrate has to be used [3], [4], [5]. Choosing a higher substrate, we reach a better performance of the antenna, but we lose an advantage of a simple integration with other electrical circuits. The top and side views are in Fig. 1.

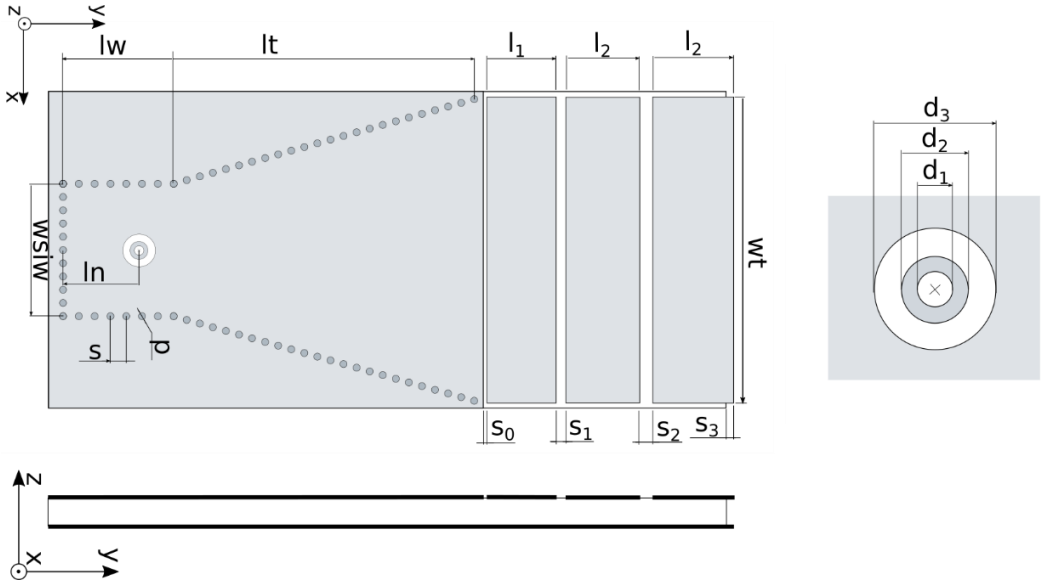


Figure 1: Designed antenna and feed structure.

2. H-PLANE HORN ANTENNA DESIGN

The design of an H-plane horn antenna is very similar to the design of a conventional pyramidal horn antenna. Designing the H-plane horn antenna, we can apply principles developed for one plane only. Since the antenna is fed by a substrate-integrated waveguide, we have to respect basic design rules for a single-mode excitation resulting in the condition

$$\lambda_0(2\varepsilon_r)^{1/2} < a < \lambda_0/\varepsilon_r, \tag{1}$$

where a is the width of the waveguide. The condition is valid for the height of the substrate h smaller than a. Other dimensions of the SIW and the horn antenna can be calculated using relations published in 0.

2.1 DESIGN METHODS

In 0 and 0, three methods of calculating basic dimensions of planar resonators were presented. We used method of coupled resonators.

In this case, we can reduce the 3D model of the antenna to a 2D model by extending the width of the antenna w_t (see **Chyba! Nenalezen zdroj odkazů.**) to infinity. Then, the structure becomes a series of strips, and each strip behaves like a parallel plate resonator. Resonators are coupled by gaps.

The resonant frequency of a resonator can be calculated as follows:

$$f_{r1} = \frac{c}{2l_{eq}\sqrt{\varepsilon_r}}. \tag{2}$$

Here, c is speed of light l_{eq} is an equivalent length of the resonator, and ε_r is the dielectric constant of the substrate used for implementing the resonator. The equivalent length l_{eq} is larger than the physical one l.

In our case, we are aimed to calculate the resonant frequency of a resonator capacitively coupled to a series of resonators. Therefore, we need calculate the capacitance of the coupling gap.

The resonant frequency of the coupled resonators can be calculated by:

$$f_{r2\pm} = \frac{f_{r1}}{\sqrt{1\pm k_2}}, \tag{3}$$

where k_2 is the coupling factor.

2.2 DESIGNED ANTENNA

The antenna was designed to operate close to human body tissue. The bottom side of the antenna was completely covered by a metal foil including the area below resonators in front of the horn. Feeding of the substrate-integrated waveguide is realized by the coaxial probe connected to the top layer of the antenna 0. This solution enables us to measure the antenna close to a phantom surface (0 mm, 5 mm). Dimensions of the feeding structure are depicted in Table 1.

Dimension	w_{siw}	w_t	l_w	l_t	l_n	s
Value [mm]	17.40	40.35	13.00	44.50	10.00	I.90
Dimension	p	s_0	s_1	s_2	s_3	l_1
Value [mm]	1.00	0.30	0.96	1.55	1.55	9.75
Dimension	l_2	d_1	d_2	d_3		
Value [mm]	11.35	1.30	2.50	4.50		

Table 1: Parameters of the designed antenna.

The designed antenna is depicted in Figure 1, and numerical values of all the parameters are shown in **Table 1**. The last resonator exceeds the substrate for s_3 = 1.55 mm. Thanks to this overlap, the resonator decreases the difference between the impedance of the last resonator and surrounding environment.

3. SIMULATIONS AND MEASURMENT RESULTS

The antenna was simulated in the CST Microwave Studio. Results of simulations of the antenna located in free space, close to a phantom (200 mm x 200 mm x 10 mm) and close to human body voxel model are presented in following paragraphs. Measurement results of the antenna placed in free space and close to an agar phantom are presented and compared with simulation results.

3.1 FREE SPACE

Frequency response of the reflection coefficient of the designed antenna operating in free space is depicted in Figure 1. The impedance bandwidth for S_{11} < -10 dB is BW = 1.01 GHz (12.6%). The antenna covers 63.8 % of UWB band-group 6 (7.392 to 8.976 GHz).

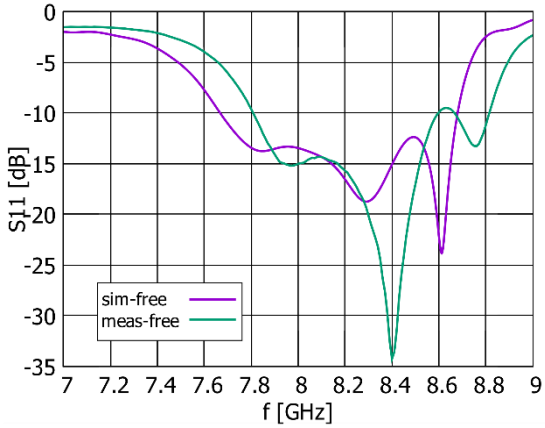


Figure 2: Simulated and measured results of the antenna located in free space.

In Figure 3, radiation patterns of the antenna in planes XY and YZ are depict. The simulated maximal gain of the antenna in free space is 10 dBi in the main lobe direction. The elevation angle in E-plane (YZ) is 40°. The measured gain is 4.12 dBi in H-plane (XY).

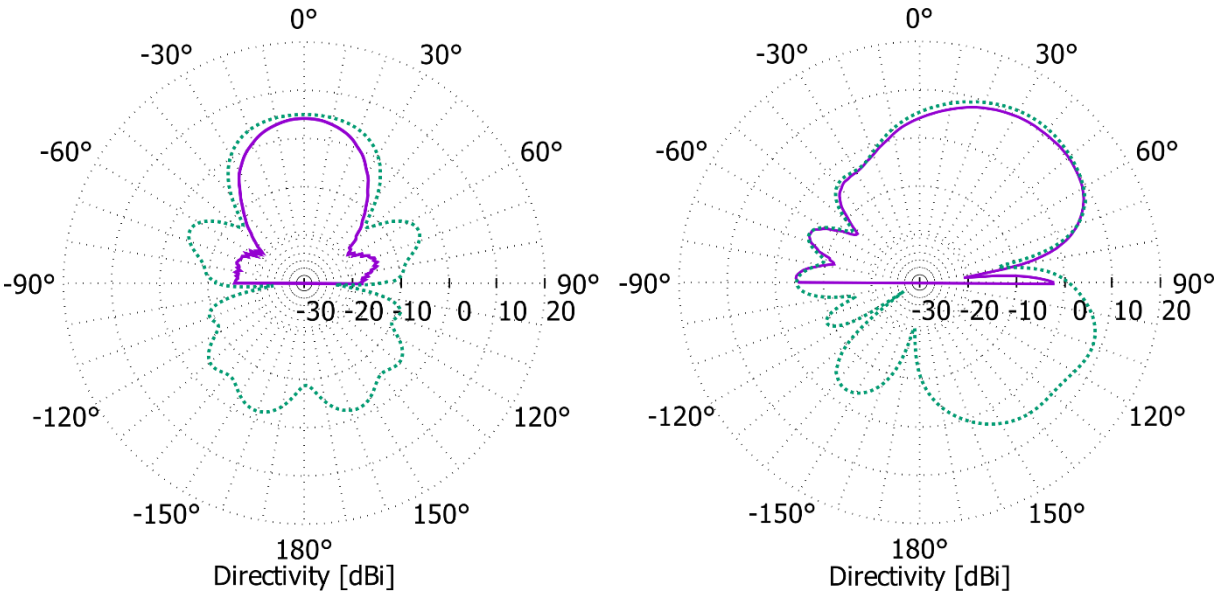


Figure 3: Simulated (dots) and measurement radiation patterns (H-plane left, E-plane right).

3.3 PHANTOM

Phantom was made by using agar gelatin and distilled water. The material parameters of the phantom are: relative permittivity ε_r = 48 and conductivity σ = 5 S/m. We measured reflection coefficient of the antenna 0 mm and 5 mm close to the agar phantom. Frequency responses of reflection coefficient at the input of the antenna are depicted in Figure 4. Photography of measurement is in the Figure 5.

Frequency responses of reflection coefficient at the input of the antenna were simulated and measured for two different distances between the antenna and the phantom (see Figure 4,5). The measured impedance bandwidth for S_{11} < -10 dB is BW_{0mm} = 760 MHz (9.50 %) and BW_{5mm} = 720 MHz (9.25 %).

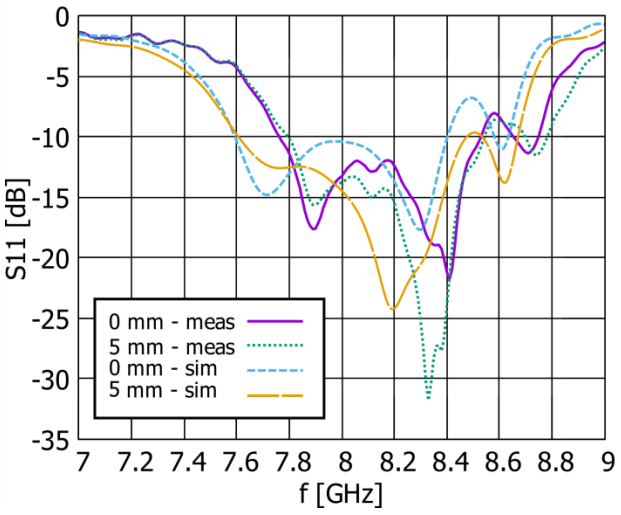


Figure 4: Reflection coefficient of the antenna close to phantom.

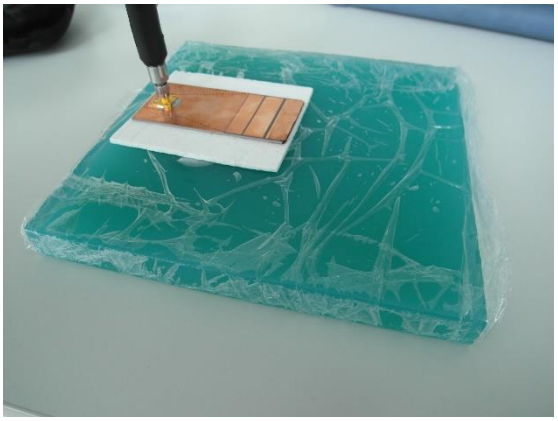


Figure 5: Fabricated antenna measured close to phantom.

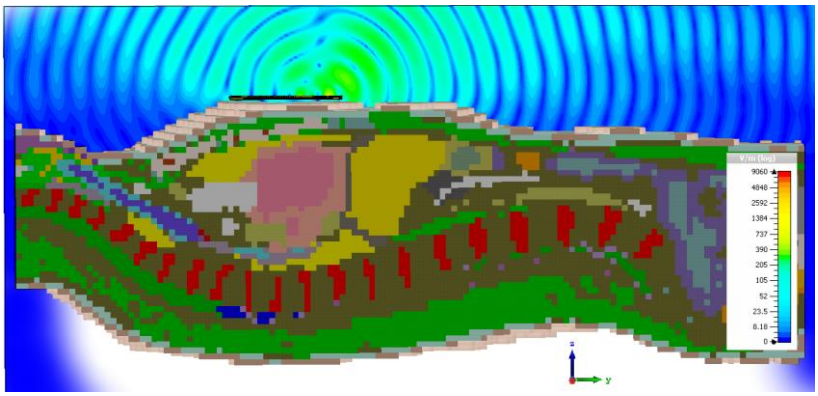


Figure 6: E-field density close to voxel model.

4. CONCLUSION

The H-plane SIW horn antenna with planar resonators was designed for the operation close to human body in the UWB band-group 6. The antenna has bandwidth 1010 MHz in free space and 760 MHz for the antenna placed 5 mm on the phantom. The measured antenna gain is 4.12 dBi in H-plane (xy) for free space cause. In the next phase, the antenna will be designed for textile substrate.

ACKNOWLEDGEMENT

The presented research was supported by the Internal Grant Agency of Brno University of Technology project no. FEKT-S-14-2483 and Czech Science Foundation by the grant 102/12/1274. Simulations and experiments in laboratories of the SIX Center were financed by the project LO1401 INWITE of the National Sustainability Program. Research is a part of the COST Action IC 1301 WiPE supported by the grant LD14057 of the Czech Ministry of Education.

REFERENCES

[1] Chen X., Hong W., Cui T., Chen, J., and Wu. K.: Substrate integrated waveguide (SIW) linear phase filter, *IEEE Microwave and Wireless Components Letters*, vol. 15, issue 11, pp.

[2] Giuppi F., Georgiadis, A., Collado, A., Bozzi, M., and Perregrini, L.: Tunable SIW cavity backed active antenna oscillator, *Electronics Letters*, vol. 46, issue 15, pp. 1053-, 2010.

[3] Yeh, C. I., Yang, D. H., Liu, T. H., Fu, J. S., Chin, K. S., Cheng, J. C., Chiu, and Kao, C. P.: MMIC compatibility study of SIW H-plane horn antenna, *2010 International Conference on Microwave and Millimeter Wave Technology*, pp. 933-936, 2010.

[4] Esquius-Morote, M., Fuchs, B., Zurcher, J-F., and Mosig. J. R.: A Printed Transition for Matching Improvement of SIW Horn Antennas, *IEEE Transactions on Antennas and Propagation*, vol. 61, issue 4, pp. 1923-1930, 2013.

[5] Esquius-Morote, M., Fuchs, B., Zurcher, J-F., and Mosig, J-R.: Novel Thin and Compact H-Plane SIW Horn Antenna, *IEEE Transactions on Antennas and Propagation*, vol. 61, issue 6, pp. 2911-2920, 2013.

[6] Balanis, C. A.: Antenna theory: analysis and design, 3rd ed. Hoboken: Wiley-Interscience, 2005. ISBN 978-0-471-66782-7.

[7] Mikulasek, T., Georgiadis, A., Collado, A., and Lacik, J.: Microstrip Patch Antenna Array Fed by Substrate Integrated Waveguide for Radar Applications,"IEEE Antennas and Wireless Propagation Letters, vol. 12, pp. 1287-1290, 2013.

[8] Lacik, J., and Mikulasek, T., Circular ring-slot antenna fed by SIW for WBAN applications, in Proc. 7th Eur. Conf. on Antennas and Propagation (EuCAP), Gothenburg, Sweden, Apr. 2013, pp. 213–216.

LINEARLY POLARIZED HMSIW U-SLOT ANTENNA FOR OFF-BODY COMMUNICATION

Patrik Hubka

Doctoral Degree Programme (3), FEEC BUT

E-mail: xhubka00@stud.feec.vutbr.cz

Supervised by: Jaroslav Lacik

E-mail: lacik@feec.vutbr.cz

Abstract: In this paper, a linearly polarized U-slot half-mode substrate integrated waveguide (HMSIW) antenna for off-body communication is presented. It is based on a combination of a U–shaped slot etched in the HMSIW top wall and two shorts placed in this slot. The antenna operates at 5.8 GHz. In order to verify performance of the proposed antenna placed in the proximity of a human body, a three-layer phantom of the human body – a chest model – is used. The antenna placed in proximity of the phantom of the human body achieves simulated impedance bandwidth of 2.02 % and the gain of 7.23 dBi.

Keywords: Half–mode substrate integrated waveguide, slot antenna, human body model

1. INTRODUCTION

In these days, a wearable antennas designed at Industrial, Scientific and Medical (ISM) bands (2.4, 5.8 and 61 GHz) attract a lot of attention. These antennas could be exploited for in-body, on-body or off-body communication [1]. Several antennas for off-body communication at 5.8 GHz were reported in [2]-[4]. The HMSIW technology [5] could be successfully exploited for the antenna design. In [2], a dual-band diamond-shape HMSIW textile antenna has been presented. The antenna is comprised by the two slots. The small one is for the tuning of the resonate frequency in the higher frequency band. Enlarging this small slot decreases the resonance frequency of the second-order mode and provides larger bandwidth in the higher frequency band. The measured impedance bandwidth of the antenna is 4.9 % and 5.1 % in 2.4 GHz and 5.8 GHz band, respectively. The measured gain of the antenna is 4.1 dBi and 5.8 dBi, in 2.4 GHz and 5.8 GHz band, respectively. In [3], a textile dual band-band patch antenna was presented. The antenna is comprised by an inner patch, operating at 5.8 GHz, surrounded by a parasitic rectangular ring element resonating at 2.45 GHz. In [3], the authors focus on crumbling of the antenna. It was demonstrated, that heavy crumbling has negative influences on operating frequency, minimum of the reflection coefficient and radiation patterns. A multilayered C-shape slot antenna has been introduced in [4]. The antenna is comprised by two patches, where the largest patch resonates at 2.45 GHz and the smaller patch resonates at 5.8 GHz. The antenna achieves impedance bandwidth of 3.88 % and 6.90 % and efficiency of 91 % and 95 % for 2.45 GHz and 5.8 GHz, respectively. Other antennas for off-body communication at 2.45 GHz and 60 GHz were reported in [6] and [7]. Antennas for on-body communication at 2.45 GHz, 5.8 GHz and 60 GHz were reported in [8]-[10].

In this paper, a half-mode substrate integrated waveguide U-slot antenna radiating linearly polarized wave designed for the 5.8 GHz ISM band for off-body communication is presented [11]. The proposed antenna is placed in proximity of human body model and its parameters are studied.

2. ANTENNA DESIGN

The antenna is depicted in Figure 1. The dielectric substrate of the length L, the width W, and the height h with the relative permittivity ε_r, and the loss tangent $\tan(\delta)$ is on its both sides covered by

metal sheets. The HMSIW is created by a row of vias and both its ends are shorted. It operates in the fundamental mode $TE_{0.5,0}$. The radiating U-slot is etched in the top wall of the HMSIW at the distance L_{viaX} from its short end in the x-direction, and from the row of vias at the distance L_{viaY} in the y-direction. The outer dimensions of the slot are L_{SLOT-X} and L_{SLOT-Y}. The width of the slot is W_{SLOT}. The slot is shorted by the strips of the width W_{SHORT1} and W_{SHORT1}. The antenna is equipped by a HMSIW to GCPW (grounded coplanar waveguide) transition.

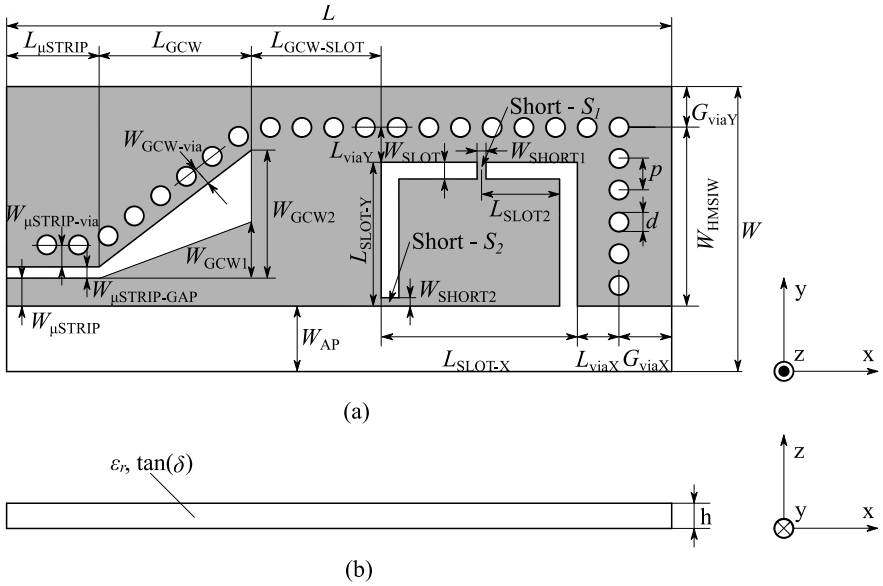


Figure 1: Top (a) and side (b) view of HMSIW U-shape slot antenna.

3. HUMAN BODY MODEL

The human body model used for simulations is composed from three layers – a skin, a fat, a muscle. From the viewpoint of the propagation of electromagnetic waves, all the described layers are lossy dielectrics of specific thicknesses and complex permittivity. The skin is a planar organ which plays the role of an external surface of organism and protects organism against surrounding environment. The thickness of the skin is between 0.4 mm and 4.0 mm. The fat is a reservoir of the energy and a heat insulator. The thickness of the fat layer is impossible to determine because everyone has unique anatomy. Muscles are tissues with elastic properties. Muscles are able to contract and relax. The layer of muscle has various thicknesses which depend on the type of muscle and the location on the body.

The configuration of the multilayer model (the phantom) of a chest is shown in Figure 2. The parameters of the phantom are summarized in Table 1 [12]. The phantom is a block of base 200 mm × 200 mm. The height of the phantom is given by the thicknesses of the layers representing the skin, the fat and the muscle. The distance d between the human body model and the antenna is 1 mm.

	ε_r	tan(δ)	σ (S/m)	thickness [mm]
Dry skin	35.114	0.32807	3.7170	2
Fat	4.955	0.18335	0.29313	10
Muscle	48.485	0.31715	4.9615	28

Table 1: Parameters of chest phantom.

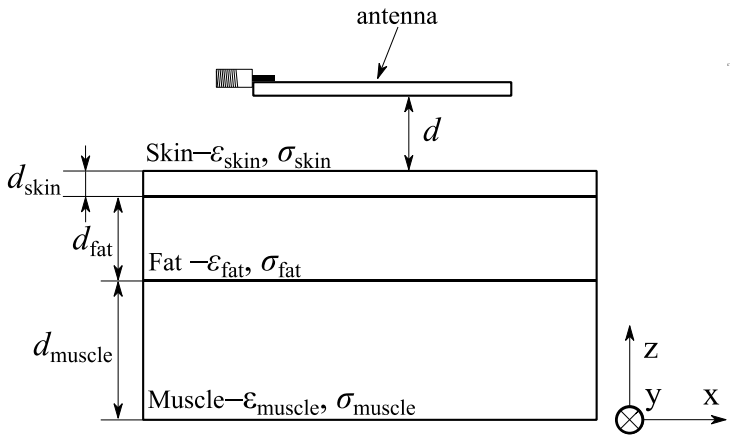


Figure 2: The chest phantom.

4. ANTENNA DESIGN AND SIMULATED RESULTS

The antenna was designed with the help of the time domain solver of CST Microwave Studio (CST MWS), where local optimization procedures Trust Region Framework and Nedler Mead Simplex Algorithm were used, for the operating frequency of 5.8 GHz on the dielectric substrate ARLON CuClad 217 with relative permittivity $\varepsilon_r = 2.17 \pm 0.02$, tangent loss $\tan(\delta) = 0.0009$, and height h = 1.524 mm. The resultant dimensions of the antenna are summarized in Table 1.

	[mm]		[mm]		[mm]
L	59.06	W	22.5	W_{AP}	3
$L_{\mu STRIP}$	8.76	$W_{\mu STRIP}$	2.34	W_{HMSIW}	16.98
L_{GCW}	8.77	$W_{\mu STRIP-GAP}$	1	G_{viaX}	9.57
$L_{GCW-SLOT}$	9.09	$W_{\mu STRIP-via}$	3.862	G_{viaY}	2.52
L_{SLOT-X}	20.95	W_{GCW1}	2.98	h	1.524
L_{SLOT-Y}	13.69	W_{GCW2}	9.99		
L_{SLOT2}	8.38	W_{SLOT}	1.687		
L_{viaX}	3.62	W_{SHORT1}	0.64		
L_{viaY}	3.29	W_{SHORT2}	0.64		

Table 2: Dimensions of designed antenna (Figure 1).

The distribution of the electric field in the substrate at the frequency 5.8 GHz is depicted in Figure 3. Thanks to the short S_1, the electric field at the slot is cutted into two halves and the linearly horizontally polarized wave in the far field is obtained.

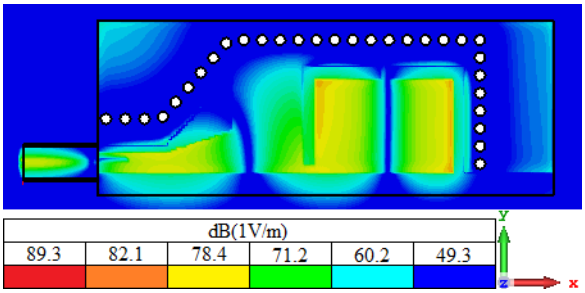


Figure 3: Distribution of electric filed intensity at 5.8 GHz in substrate (magnitude).

The simulated reflection coefficient of the antenna in the free-space and in the proximity of the chest phantom is depicted in Figure 4. The simulated reflection coefficient of the antenna achieves impedance bandwidth about 2.05 % and 2.02 % for the antenna placed in the free-space and in the proximity of human body, respectively. The minimum of the reflection coefficient of the antenna placed in proximity on human body is slightly shifted about 25 MHz to lower frequencies due to the influence of the phantom.

The normalized simulated radiation patterns in two orthogonal cutting planes in the free-space and in the proximity of the phantom at the frequency 5.8 GHz are depicted in Figure 5. It can be observed that the co-polarizations are in good agreement at the boresight direction in both cutting planes. The level of the cross-polarization at the boresight direction reaches higher level due to the influence of phantom in comparison to the cross-polarization of the antenna placed in the free-space. The simulated antenna peak gain is 6.49 dBi and 7.23 dBi for the antenna placed in the free-space and in the proximity of the phantom, respectively.

Simulated total efficiency at operating frequency f = 5.8 GHz of the antenna is 96 % and 82 % in the free-space and in proximity of the phantom of human body model, respectively.

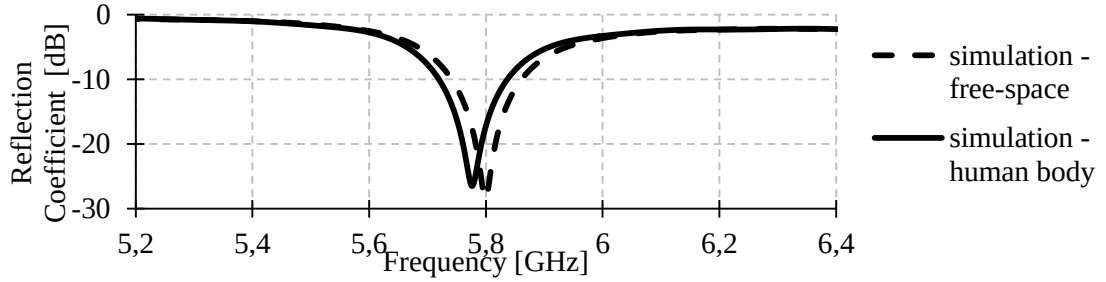
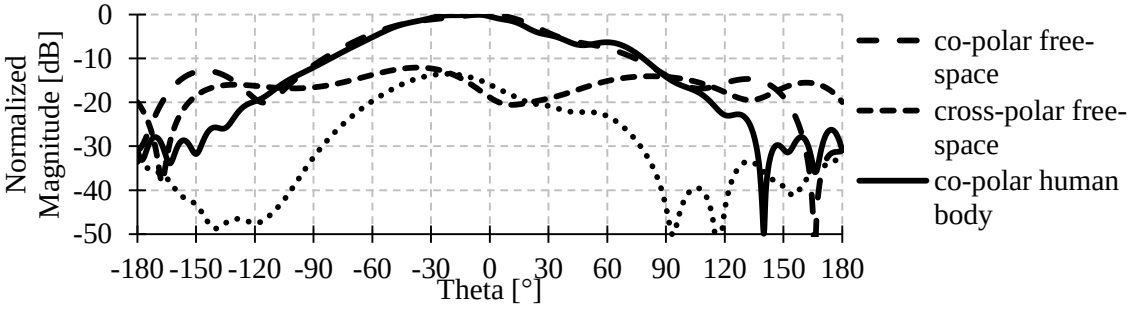
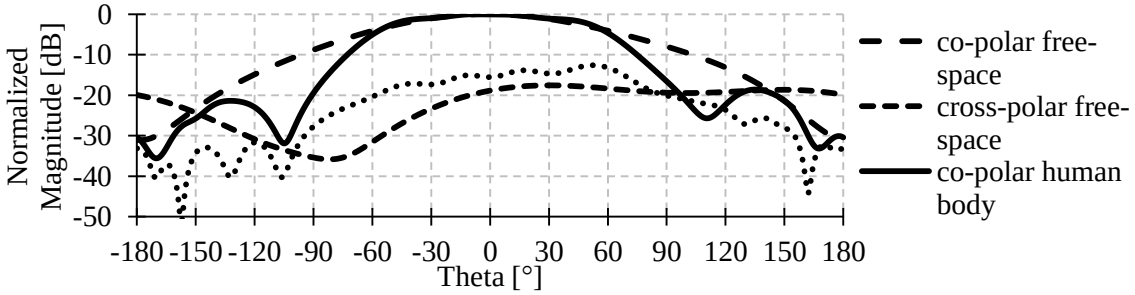


Figure 4: Simulated reflection coefficient.



(a)



(b)

Figure 5: Simulated radiation patterns of the antenna placed in free-space and in proximity of human body (phantom): (a) E-plane and (b) H-plane.

In comparison to antenna proposed in [2], the presented antenna has a higher gain, higher total efficiency and smaller size.

5. CONCLUSION

In this paper, a linearly polarized half-mode substrate integrated waveguide U-slot antenna operating at the frequency of 5.8 GHz in the proximity of a human body model (a chest phantom) has been proposed. The simulated results of the antenna placed in the proximity a human body achieves impedance bandwidth of 2.02 % and gain of 7.23 dBi. The antenna is suitable for off-body communication. The will be experimentally verified and the results will be presented at the conference.

ACKNOWLEDGEMENT

The presented research was supported by the Czech Grant Agency project no. P102/12/1274, by the Czech Ministry of Education in frame of National Sustainability Program under grant LO1401, and by the Internal Grant Agency of Brno University of Technology project no. FEKT-S-14-2483. The research is the part of the COST Action IC1301 which is financially supported by the grant of the Czech Ministry of Education no. LD14057. For research, infrastructure of the SIX Center was used.

REFERENCES

[1] P. S. Hall, "Antennas challenges for body centric communications," *in Proc. Int. Workshop Antenna Technol., Small Antennas Metamater. Appl.*, Mar. 2007, pp. 41–44.

[2] S. Agneessens, and H. Rogier, "Compact Half Diamond Dual-Band Textile HMSIW On-Body Antenna," *IEEE Transactions on Antennas and Propagation*, 2014, vol. 62, no. 5, pp. 2374–2381.

[3] Q. Bai, and R. Langley, "Crumpled textile antennas," *Electron. Lett.*, vol. 45, pp. 436–438, 2009.

[4] A. R. Chandran, and W. G. Scanlon, "Dual-band low profile antennas for body-centric communications," *International Workshop on Antenna Technology (iWAT)*, pp. 1 – 4, 2010.

[5] W. Hong, B. Liu, Y. Wang, Q. Lai, H. Tang, X. X. Yin, YIN, Y. D. Dong, Y. Zhang, and K. Wu, "Half Mode Substrate Integrated Waveguide: A New Guided Wave Structure for Microwave and Millimeter Wave Application," *Joint 31st International Conference on Infrared Millimeter Waves and 14th International Conference on Terahertz Electronics*, 2006, pp. 219.

[6] P. Salonen, and L. Humre, "A novel fabric WLAN antenna for wearable applications," *in Proceedings of the IEEE Antennas and Propagation Society International Symposium*, Jun. 2003, vol. 2, pp. 700–703.

[7] L. Panzin, and Y. Leviatan, "A compact 60-GHz tapered slot antenna printed on LCP substrate for WPAN applications," *IEEE Antennas and Wireless Propagation Letters*, vol. 9, pp. 272–275, 2010.

[8] I. Kahn, L. Yu, Y. I. Nechayev, and P. S. Hall, "Space and pattern diversity for on-body communication channels in an indoor environment at 2.45GHz," *in Proceedings of The Second European Conference on Antenna and Propagation (EUCAP 2007)*, Nov. 2007, pp. 1-6.

[9] I. Khan, and P.S. Hall, "Multiple Antenna Reception at 5.8 and 10 GHz for Body-Centric Wireless Communication Channels," *IEEE Transactions on Antennas and Propagation*, Vol. 57. No. 1, January 2009.

[10] S. S. Hsu, K. C. Wei, C. Y. Hsu, and H. R. Chuang, "A 60-GHz millimeter-wave CPW-fed Yagi antenna fabricated by using 0.18-μm CMOS technology," *IEEE Electron Devices Letters*, vol. 29, no. 6, pp. 625–627, Jun. 2008.

[11] P. Hubka, and J. Lacik, "Linearly polarized HMSIW U-slot antenna, " *In Proceedings of Conference on Microwave Techniques (COMITE)*, Pardubice (Czech Republic)," April 2015, pp 1-3.

[12] C. Gabriel, "Calculation of the Dielectric Properties of Body Tissues in the frequency range 10 Hz – 100 GHz. IFAC. [online]. 1997 [cit. 2016-06-03]," Available at: http:// http://niremf.ifac.cnr.it/tissprop/htmlclie/htmlclie.php.

MATRIX SOLUTION OF DIFFERENCE LINEAR SYSTEMS ARISING IN THEORY OF DIGITAL FILTERS

Kristýna Mencáková
Doctoral Degree Programme (1. ročník), FEEC BUT
E-mail: mencakova.k@fce.vutbr.cz

Supervised by: Josef Diblík
E-mail: diblik.j@fce.vutbr.cz

Abstract: The paper deals with utilisation of theory of linear difference equations to problems of digital filters modelling.

Keywords: Difference systems, digital filters, Putzer algorithm

1 INTRODUCTION

When linear digital filters with constant parameters are described by a mathematical model, systems of linear discrete difference equations are used as a basic tool of investigation (see, e.g., [1], [2], [3]). Linear difference equation of a higher-order with constant coefficients, when a system with a single input and a single output is considered, determines a relation among input, state and output quantities. In order to investigate such equations it is often more comfortable to transform them into systems of linear difference equations of the first order. As a result we get models of state descriptions. These systems are usually solved by Z-transform. The purpose of our contribution is to give a matrix solution utilising the theory of systems of discrete equations.

2 PRELIMINARY – SYSTEMS OF DISCRETE EQUATIONS

In the following text we use some formulas from the theory of systems of difference equations. Let us consider a nonhomogeneous system

$$y(n+1) = C(n)y(n) + g(n), \tag{1}$$

with an initial value

$$y(n_0) = y_0, \tag{2}$$

where $n \geq n_0$, $n_0 \in \mathbb{N}$, $C(n)$ is an $s \times s$ real matrix defined for $n \geq n_0$ and $g(n)$ is a real s-dimensional vector defined for $n \geq n_0$.

The solution of problem (1), (2) (see [4]) is given by the formula

$$y(n, n_0, y_0) = \Phi(n, n_0)y_0 + \sum_{k=n_0}^{n-1} \Phi(n, k+1)g(k) \tag{3}$$

where $\Phi(n, n_0)$ is an $s \times s$ fundamental matrix of the homogeneous system

$$y(n+1) = C(n)y(n),$$

such that $\Phi(n, n_0) = I$, I is the unit matrix, by definition

$$\sum_{k=n_0}^{n_0-1} \Phi(n, k+1)g(k) = \theta$$

and θ is an s-dimensional null vector.

When the fundamental matrix $\Phi(n,n_0)$ is specified by

$$\Phi(n,n_0)=\prod_{i=n_0}^{n-1}C(i),$$

then we turn formula (3) into

$$y(n,n_0,y_0)=\left(\prod_{i=n_0}^{n-1}C(i)\right)y_0+\sum_{k=n_0}^{n-1}\left(\prod_{i=k+1}^{n-1}C(i)\right)g(k),$$

where, by definition,

$$\prod_{i=n_0}^{n_0-1}C(i)=I,\quad \sum_{k=n_0}^{n_0-1}\left(\prod_{i=k+1}^{n-1}C(i)\right)g(k)=\theta.$$

When C is a constant matrix, the solution of the problem (1), (2) is given by

$$y(n,n_0,y_0)=C^{n-n_0}y_0+\sum_{k=n_0}^{n-1}C^{n-k-1}g(k). \tag{4}$$

3 DISCRETE SYSTEMS DESCRIBING DIGITAL FILTERS

A state model of a digital filter is described in [1] by partial state quantities ν_i, $i=1,\ldots,s$, by equations

$$\begin{aligned}
&\nu_1(n)=b_sy(n)-a_sx(n),\\
&\nu_2(n)=b_{s-1}y(n)+b_sy(n+1)-a_{s-1}x(n)-a_sx(n+1),\\
&\nu_3(n)=b_{s-2}y(n)+b_{s-1}y(n+1)+b_sy(n+2)-a_{s-2}x(n)-a_{s-1}x(n+1)-a_sx(n+2),\\
&\qquad\vdots\\
&\nu_s(n)=b_1y(n)+b_2y(n+1)+\ldots+b_sy(n+s-1)-a_1x(n)-a_2x(n+1)-a_3x(n+2)-\ldots\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\ldots-a_sx(n+s-1),
\end{aligned}$$

where $b_1,\ldots,b_s$ and $a_1,\ldots,a_s$ are constants, x is an input and y is an output.

In [1], assuming $b_s\neq 0$, is such a system transformed into

$$\begin{aligned}
\nu_1(n+1) &= \nu_2(n)-\frac{b_{s-1}}{b_s}\nu_1(n)+\left(a_{s-1}-b_{s-1}\frac{a_s}{b_s}\right)x(n),\\
\nu_2(n+1) &= \nu_3(n)-\frac{b_{s-2}}{b_s}\nu_1(n)+\left(a_{s-2}-b_{s-2}\frac{a_s}{b_s}\right)x(n),\\
&\vdots\\
\nu_{s-1}(n+1) &= \nu_s(n)-\frac{b_1}{b_s}\nu_1(n)+\left(a_1-b_1\frac{a_s}{b_s}\right)x(n),\\
\nu_s(n+1) &= -\frac{b_0}{b_s}\nu_1(n)+\left(a_0-b_0\frac{a_s}{b_s}\right)x(n).
\end{aligned}$$

Moreover the last system can be written in a matrix form as

$$\nu(n+1)=\mathbf{A}\nu(n+1)+\mathbf{B}x(n), \tag{5}$$

where

$$\nu(n+1) = \begin{pmatrix} \nu_s(n+1) \\ \nu_{s-1}(n+1) \\ \vdots \\ \nu_2(n+1) \\ \nu_1(n+1) \end{pmatrix}$$

and

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 & -\dfrac{b_0}{b_s} \\ 1 & 0 & 0 & \dots & 0 & 0 & -\dfrac{b_1}{b_s} \\ 0 & 1 & 0 & \dots & 0 & 0 & -\dfrac{b_2}{b_s} \\ & & & \vdots & & & \\ 0 & 0 & 0 & \dots & 1 & 0 & -\dfrac{b_{s-2}}{b_s} \\ 0 & 0 & 0 & \dots & 0 & 1 & -\dfrac{b_{s-1}}{b_s} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} a_0 - b_0\dfrac{a_s}{b_s} \\ a_1 - b_1\dfrac{a_s}{b_s} \\ a_2 - b_2\dfrac{a_s}{b_s} \\ \vdots \\ a_{a-2} - b_{s-2}\dfrac{a_s}{b_s} \\ a_{s-1} - b_{s-1}\dfrac{a_s}{b_s} \end{pmatrix}.$$

Let $n_0 = 0$. The matrix A is a constant matrix. Therefore we will use formula (4) to solve the initial problem (5), (6), where

$$\nu(0) = \nu_0. \tag{6}$$

In (4) we set

$$y(n, n_0, y_0) = \nu(n, 0, \nu_0), \quad C = \mathbf{A}, \quad g(k) = \mathbf{B}x(k).$$

Then the solution of problem (5), (6) is

$$\nu(n, 0, \nu_0) = \mathbf{A}^n \nu_0 + \sum_{k=0}^{n-1} \mathbf{A}^{n-k-1} \mathbf{B} x(k). \tag{7}$$

4 EXAMPLE

Consider a digital filter defined by the difference equation (see [1], p. 52)

$$y(n+2) - 1,25y(n+1) + 0,78125y(n) = x(n+2) - x(n).$$

The system that corresponds to (5) is a system of two equations ([1], p. 53)

$$\nu_2(n+1) = 0 \cdot \nu_2(n) - \frac{b_0}{b_2}\nu_1(n) + \left(a_0 - b_0\frac{a_2}{b_2} \right) x(n), \tag{8}$$

$$\nu_1(n+1) = 1 \cdot \nu_2(n) - \frac{b_1}{b_2}\nu_1(n) + \left(a_1 - b_1\frac{a_2}{b_2} \right) x(n), \tag{9}$$

$$\nu(0) = \nu_0, \tag{10}$$

where $b_2 = 1$, $b_1 = -1,25$, $b_0 = 0,78125$, $a_2 = 1$, $a_1 = 0$, $a_0 = -1$,

$$\mathbf{A} = \begin{pmatrix} 0 & -0,78125 \\ 1 & 1,25 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} -1,78125 \\ 1,25 \end{pmatrix}.$$

The solution of the initial problem (8) – (10) is given by formula (7). To compute powers $\mathbf{A}^n$ we use the Putzer algorithm (see, e.g., [4]). In accordance with this algorithm we get

$$\mathbf{A}^n = \lambda_1^n I + \left(\sum_{i=0}^{n-1} \lambda_2^{n-1-i} \lambda_1^i \right) \cdot (\mathbf{A} - \lambda_1 I), \tag{11}$$

where λ_1, λ_2 are the roots of the characteristic equation

$$\det(\mathbf{A} - \lambda I) = \lambda^2 - 1,25\lambda + 0,78125 = 0$$

and $\lambda_1 = 0,625 + 0,625\,j$, $\lambda_2 = 0,625 - 0,625\,j$.

By simplifying representation (11) we get

$$\begin{aligned} \mathbf{A}^n &= \lambda_1^n I + \left(\sum_{i=0}^{n-1} \lambda_2^{n-1-i} \lambda_1^i \right) (\mathbf{A} - \lambda_1 I) = \lambda_1^n I + \frac{\lambda_2^n - \lambda_1^n}{\lambda_2 - \lambda_1} (\mathbf{A} - \lambda_1 I) = \\ &= \frac{1}{\lambda_2 - \lambda_1} \left[\lambda_1 \lambda_2 (\lambda_1^{n-1} + \lambda_2^{n-1}) I - (\lambda_1^n - \lambda_2^n) \mathbf{A} \right]. \end{aligned} \tag{12}$$

We will eliminate complex numbers from (12). Expressing roots λ_1, λ_2 in an exponential form we have

$$\begin{aligned} \lambda_1 &= re^{j\varphi}, \\ \lambda_2 &= re^{-j\varphi}, \end{aligned}$$

where

$$\varphi = \frac{\pi}{4} \quad \text{and} \quad r = \sqrt{\lambda_1 \lambda_2} = \sqrt{0,78125} = 0,8838835.$$

Then, using the Euler formula,

$$\begin{aligned} \lambda_1^n - \lambda_2^n &= 2jr^n \sin(n\varphi), \\ \lambda_1^{n-1} - \lambda_2^{n-1} &= 2jr^{n-1} \sin((n-1)\varphi), \\ \lambda_1 - \lambda_2 &= 2jr \sin\varphi, \\ \lambda_1 \lambda_2 &= r^2 \end{aligned}$$

and (12) equals (after simple computations)

$$\mathbf{A}^n = \frac{r^{n-1}}{\sin\varphi} \left[(r\sin((n-1)\varphi)) I + (\sin(n\varphi)) \mathbf{A} \right],$$

or

$$\begin{aligned} \mathbf{A}^n \;=\; & (0,8838835)^{n-1} \cdot \sqrt{2} \cdot \\ & \cdot \begin{pmatrix} 0,8838835 \cdot \sin\left((n-1)\dfrac{\pi}{4} \right) & -0,78125 \cdot \sin\dfrac{n\pi}{4} \\ \sin\dfrac{n\pi}{4} & 0,8838835 \cdot \sin\left((n-1)\dfrac{\pi}{4} \right) + 1,25 \cdot \sin\dfrac{n\pi}{4} \end{pmatrix}. \end{aligned}$$

The solution of the initial problem (8) – (10) equals (in accordance with formula (7))

$$\begin{aligned} \nu(n,0,\nu_0) = \begin{pmatrix} \nu_2(n,0,\nu_0) \\ \nu_1(n,0,\nu_0) \end{pmatrix} = \frac{r^{n-1}}{\sin\varphi} \left[r\sin((n-1)\varphi) \cdot I + \sin n\varphi \cdot \mathbf{A} \right] \nu_0 + \\ + \sum_{k=0}^{n-1} \frac{r^{n-k-2}}{\sin\varphi} \left[r\sin((n-k-2)\varphi) \cdot I + \sin((n-k-1)\varphi) \cdot \mathbf{A} \right] \mathbf{B} x(k). \end{aligned}$$

5 CONCLUSION

In this article we demonstrated an application of the theory of difference linear equation to solving equations arising in the theory of digital filters.

ACKNOWLEDGEMENT

The author was supported by Grant FEKT-S-14-2200 of Faculty of Electrical Engineering and Communication, BUT.

REFERENCE

[1] Vích, R., Smékal, Z.: Číslicové filtry, Praha, Academia 2000, ISBN 80-200-0761-X.

[2] Proakis, J. G., Manolakis, D. G.: Digital Signal Processing, Fourth edition, New Jersey, Prentice Hall 2006, ISBN 978-0131873742.

[3] Oppenheim A. V., Schafer R. W., Buck J. R.: Disctere-Time Signal Processing, New Jersey, Prentice Hall 2009, ISBN 0-13-754920-2.

[4] Elaydi, S. N.: An Introduction to Difference Equations, Third edition, Springer 2005, ISBN 978-0-387-23059-7.

NEURAL MODEL OF TRANSMISSION CHANNEL IN 60 GHZ ISM BAND

Martin Kotol

Doctoral Degree (2), FEEC BUT

E-mail: xkotol00@stud.feec.vutbr.cz

Supervised by: Zbyněk Raida

E-mail:raida@feec.vutbr.cz

Abstract: In this paper, methodology of estimating parameters of a wireless transmission channel inside a car is presented. The work is focused on the utilization of artificial neural networks for channel modelling in the frequency range from 55 GHz to 65 GHz. Promising results have been reached by a feed-forward neural network and a radial basis function neural network. In order to train the networks, a wireless transmission was carefully measured in a testing car. Measured data were properly processed to be used both for training neural networks and validating neural models.

Keywords: Artificial neural network, feed-forward network, radial basis function network, parameter estimation, in-car wireless transmission.

1. INTRODUCTION

In the paper, we describe an efficient approach to the modeling of a transmission between two antennas which have been placed on an inner surface of a car body. The transmission has been influenced by a relative position of antennas. The ISM band from 55 GHz to 65 GHz [1] has been used for the experiment.

Numerical modeling of wireless in-car communication at 60 GHz is extremely time-consuming due to electrically large dimensions of in-car objects. We have therefore developed an effective approach to modeling which uses artificial neural networks (ANN). ANN is exploited for estimation of transmission between antennas.

Two types of neural networks have been selected for modeling – a feed-forward (FF) network and a radial basis function (RBF) network. Accurately created neural models have been compared to the measured signal transmission between the antennas to verify functionality.

2. ARTIFICIAL NEURAL NETWORK

ANN is a universal computational tool which is composed of simple units called neurons. Neurons are organized into layers which are interconnected and form a network. Input of each neuron in the layer M is connected to the outputs of all neurons in the previous layer $M-1$. Output of the neuron in the layer M is connected to inputs of all the neurons in the next layer $M+1$. Interconnections in between neurons are associated with different weights. Weights are set randomly during initialization. During training the network, weights are adjusted to minimize the error at the output of the ANN (the difference between the expected response and the actual one) [1].

Figure 1 shows a general ANN containing the input layer, the hidden layer and the output layer. Here, x_1 to x_n are signals at the input of input neurons, w_{11} to w_{mn} denote weights between neurons and f_1 to f_n are signals at the output of output neurons. The input layer distributes input signals among hidden neurons [3]. Neurons in the hidden layer transform signals using a specific activation function. The output layer forms output signals of ANN.

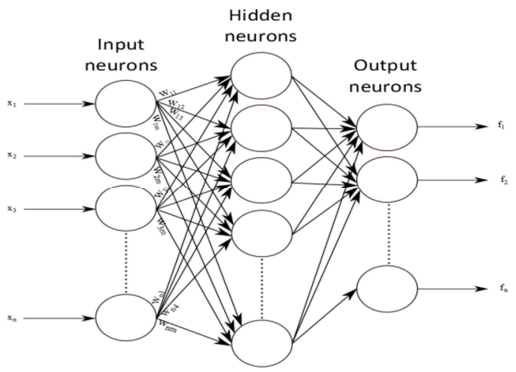


Figure 1: General diagram of the artificial neural network

For modeling in-car communication channels at 60 GHz, a feed-forward (FF) neural network and a radial basis function (RBF) neural network have been used.

In the FF ANN (Figure 2), information flows from inputs directly to outputs without any feedback. During training, weights between neurons are set to minimize the difference between expected outputs and actual ones. After training, the network is ready to estimate an appropriate output target from an arbitrary input pattern.

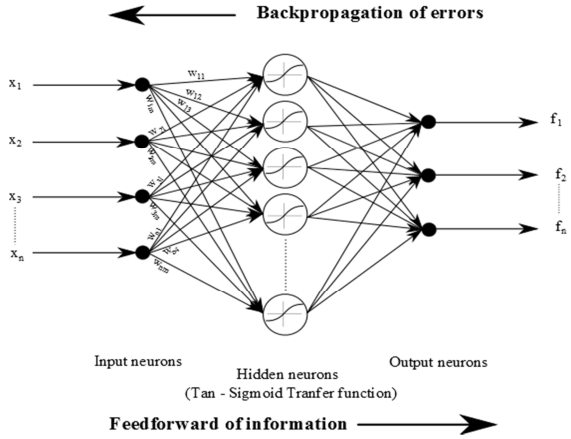


Figure 2: Structure of general feed forward neural network

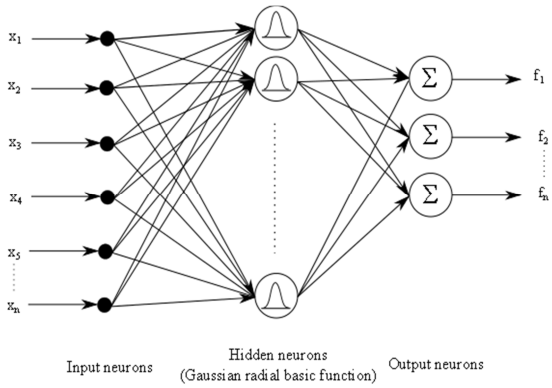


Figure 3: Structure of general radial basis function neural network

The RBF ANN (Figure 3) consists of the input layer, the hidden layer and the output layer. Neurons in the input layer distribute input signals among neurons in the hidden layer. Hidden neurons transform signals using the Gauss kernel function. Signals close to the center of the Gauss kernel function are amplified, and vice versa. Neurons in the output layer sum up signals at outputs of Gauss neurons and form the output pattern.

3. DESCRIPTION OF MEASUREMENTS

The transmission between the antennas was measured along the car roof. The transmission was measured for several positions of antennas in the frequency range from 55 GHz to 65 GHz with frequency step 10 MHz. Antennas were located on the inner surface of the roof, and the transmission was measured along the roof depending on the location of the antennas.

In the first configuration, the transmitting antenna was at a position Tx1 and the receiving antenna at the position Rx1. The position of the transmit antenna Tx1 was fixed, and the position of the receive antenna was changed to Rx2 and Rx3. For all positions, we measured transmissions. Similarly, transmissions were measured for positions Tx2 and Tx3 of the transmit antenna. The role of the transmit antenna and the receive antenna was played by an open end of the waveguide WR15. Positions of transmit antennas and receive ones are shown in Figure 4.

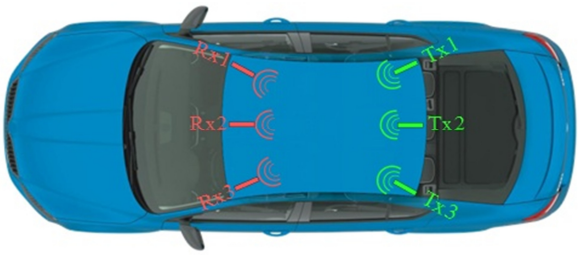


Figure 4: Location of transmit and receive antennas.

An example of measured frequency response of transmission coefficient is shown in Figure 5 for the configuration Tx3–Rx1 (red circles). The measured data were approximated by a low-order polynomial (blue line). Input patterns for training ANN were conceived as the triplets [frequency, position of transmit antenna, position of receive antenna]. Corresponding magnitude of the transmission coefficient was considered as the output target. Training patterns were created for frequencies from 55 GHz to 65 GHz with the step 150 MHz for all considered combinations of Tx and Rx antennas positions.

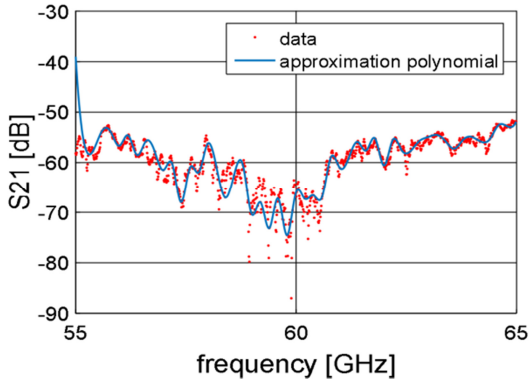


Figure 5: Measured frequency response of the transmission between antennas Tx3 and Rx1 (red). Polynomial approximation of frequency response (blue).

The training process is terminated if the relative training error is lower than 10^{-6}. Functionality of the trained NN is verified by input patterns differing from training ones.

Transmissions at testing frequencies between transmit antennas and receive ones, which were estimated by FF ANN, are shown in Figures 6 to 9. Transmissions estimated by RBF ANN are depicted in Figures 9 to 11. Red circles represent measured values and blue lines correspond to estimated ones. Figure 12 shows the relative error of estimation for FF ANN, and Figure 13 represents the relative error of estimation for RBF ANN.

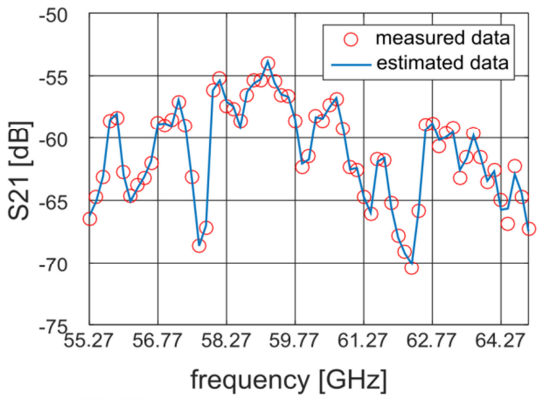


Figure 6: Transmission between Tx3 and Rx1. FF estimate (blue) versus measurement (red).

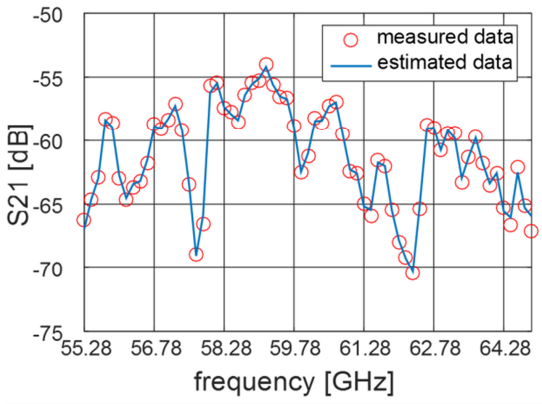


Figure 9: Transmission between Tx3 and Rx1. RBF estimate (blue) versus measurement (red).

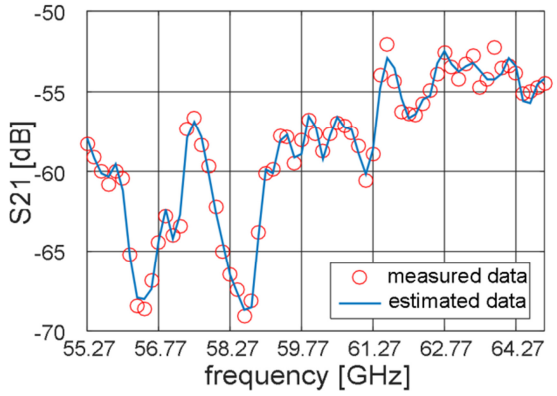


Figure 7: Transmission between Tx3 and Rx2. FF estimate (blue) versus measurement (red).

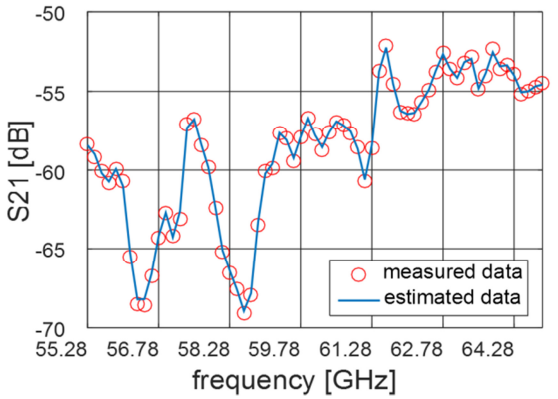


Figure 10: Transmission between Tx3 and Rx2. RBF estimate (blue) versus measurement (red).

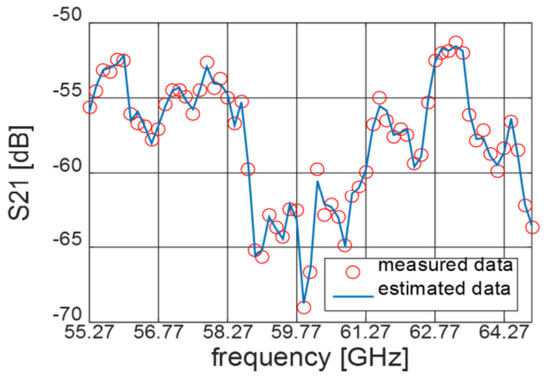


Figure 8: Transmission between Tx3 and Rx3. FF estimate (blue) versus measurement (red).

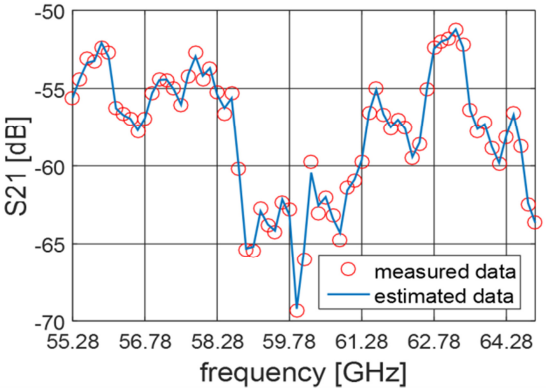


Figure 11: Transmission between Tx3 and Rx3. RBF estimate (blue) versus measurement (red).

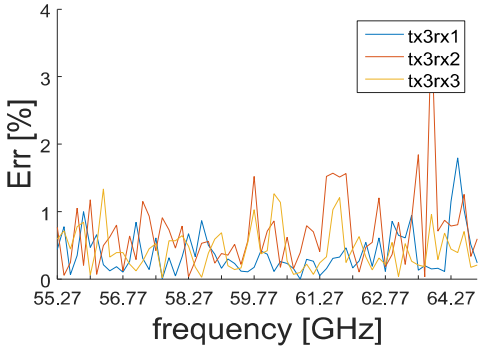


Figure 12: Relative error of FF estimation of transmission between antennas.

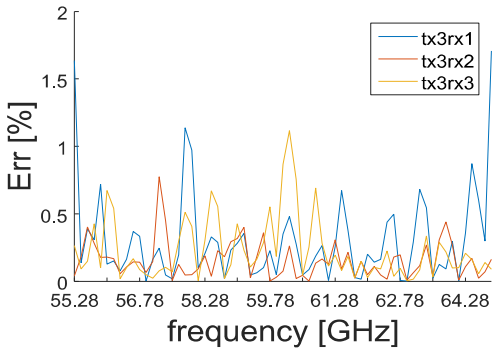


Figure 13: Relative error of RBF estimation of transmission between antennas.

4. CONCLUSION

The paper presents an exploitation of the feed-forward (FF) neural network and the radial basis function (RBF) neural network for modeling wireless transmission channels along the roof of a car. Networks were trained to estimate the transmission between the transmit antenna and the receive one depending on antenna location and operation frequency from the range 55 GHz to 65 GHz. The functionality of neural models was verified using testing patterns which differ from the training ones by the frequency. Accuracy of estimation is below 3% of the relative error.

ACKNOWLEDGEMENT

Research described in this paper was financially supported by the Czech Science Foundation under grant no. P102/12/1274. Research is a part of the COST Action IC1301 supported by the grant LD14057 of the Czech Ministry of Education. Measurements were performed at the SIX Research Center. The research was supported by the Internal Grant Agency of Brno University of Technology under the project no. FEKT-S-14-2483.

REFERENCES

[1] J. Blumenstein, et al, "In-vehicle mm-wave channel model and measurement," In Proceedings of Vehicular Technology Conference, fall 2014, Vancouver, Canada.

[2] D. Kriesel, "A Brief Introduction to Neural Networks," avail. at http://www.dkriesel.com, 2007.

[3] Jingwen Tian, Meijuan Gao, Fan Zhang, "Network Intrusion Detection Method Based on Radial Basic Function Neural Network," In Proceedings of E-Business and Information System Security, 2009, Wuhan, China.

[4] S. Haykin, "Neural Networks: A Comprehensive Foundation," 2/E, Prentice Hall, 1999.

[5] Z. Raida, "Modeling EM structures in the Neural Network Toolbox of MATLAB," IEEE Antennas and Propagation Magazine, 2003, vol. 44, no. 6, p. 46-56.

NUMERICAL ANALYSIS OF THE LIGHTNING ON SMALL AIRCRAFT

David Krutílek

Doctoral Degree Programme (3), FEEC BUT

E-mail: xkruti01@stud.feec.vutbr.cz

Supervised by: Zbyněk Raida

E-mail: raida@feec.vutbr.cz

Abstract: This paper deals with numerical analysis of lightning stroke effects on aircraft. Computations are implemented by the use of the time-domain solver (FIT). Attention is turned to a simple model of a composite nose part of the airplane EV-55. The analysis of the model itself is performed for two fundamentally different basic conceptions of a possible protection against direct effects of lightning.

Keywords: Numerical analysis, electromagnetic compatibility, (in)direct effects of lightning, lightning, aircraft

1. INTRODUCTION

Computational electromagnetism (CEM) is a branch established at the end of 1960's when the first applications for task solving in the field of electromagnetism with the aid of finite element method occurred. Retrospectively, computer technology was at rudimentary level and thus the computational costs were astronomical. As the computer technology was improving and spreading, the price of hardware was decreasing and new, more or less effective methods have been presented. These methods led to the fusion of computations and measurements and brought certain financial benefits for reduction of initial production costs. State of the art and CEM utilization in the aviation industry can be divided into two main groups based on the external effects affecting the aircraft during its common operation:

- Effects of fields with high intensity (high intensity radiated fields, HIRF)

  The main sources of HIRF can be radio and television transmission towers, radio-technical devices used at the airport, radars, etc.

- Influences of effects caused by lightning stroke

  The aircraft is in touch with the stroke of lightning discharge less often compared to the flight through HIRF. The lightning strike hits into a commonly used airplane once a year on average.

The skin of most airplanes is fabricated from duralumin, which is relatively well conductive and acts as a Faraday cage. This assures that the majority of the lightning current flows on the aircraft skin.

Nevertheless, carbon- and glass-fiber composites are increasingly applied in modern aircraft design to reach construction characteristics comparable to aluminum alloys but with a reduced weight. In a direct contrast to the excellent construction characteristics, composites have limited qualities to safely protect aircraft against indirect effects of lightning strokes (IEOLS), effects of high intensity radiated fields (HIRF) and direct effects of lightning stroke (DEOLS). This drawback is caused by a poor conductivity of composite materials. Therefore, composite materials are often completed by additional protective layers (protective conductive meshes, paint, etc.) to minimize this disadvantage.

Simulations of effects of lightning support better understanding of phenomena in the airplane structure when lightning strikes. By analyzing the results, we can identify weak points in an airplane from the perspective of lightning protection, and an adequate protection can be designed before the fabrication of the airplane prototype. That way, costs are reduced in the susceptibility testing and the certification to complete aircraft against direct lightning effects [1], [2].

A poor design of a lightning protection system (LPS) can cause a catastrophe. The worst crash of a big commercial airplane caused by lightning was recorded in 1963. The aircraft Boeing 707 was hit by lightning strike near Maryland, US, which resulted in explosion of a fuel tank. All 81 passengers on board died [3]. Thirteen years later, in 1976, another serious crash of the airplane took place, when Iranian Boeing 737 was approaching the airport in Madrid. That time, 17 passengers died due to the same fuel tank explosion [4]. Since that time, the aircraft protection against lightning was significantly improved. The development successfully continues thanks to new materials used in aviation.

2. MODEL UNDER TEST

The described analyses have been performed on a simplified, but still sufficiently realistic model of the nose part of an airplane EV-55 [5]. Analyses of the model of the nose part have been carried out for basic concepts of a possible protection against direct effects of lightning discharge. Main advantages and disadvantages of this concept are going to be further discussed through their different responses to the defined excitation signal. Nature of the excitation signal has to correspond to the standardized lightning discharge of the type A [6]. Since verification of calculated results by experimental measurements is relatively difficult, the main problem of lightning stroke was simulated using different solvers, and obtained outputs were compared.



Figure 1: Real plane EV-55 [5].

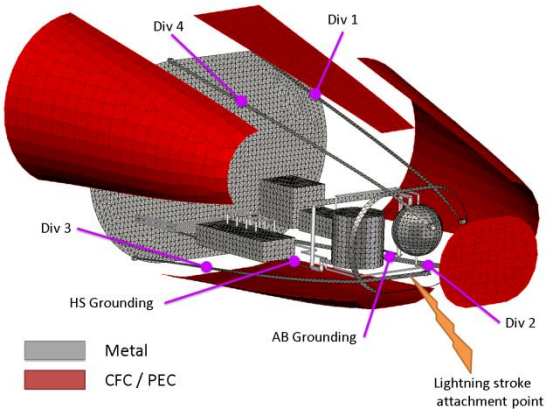


Figure 2: Morphed model of EV-55 nose.

2.1. CREATING OF NUMERICAL MODEL

When creating electro-magnetic model, compromises and simplifications make the model computable on one hand, but deviations from the real situation appear on the other hand. Compromises and simplifications of material inhomogeneities as well as imperfect fit parts are essential.

Inhomogeneities of the material are created during production of composite part already, and comprise various defects and overlaps that are caused by manufacturing technology. Samples of these defects are shown in Figure 3. These defects cannot be included into models, and have to be handled if the model is manufactured without these defects.



Figure 3: Inhomogeneity of the material (Cu mesh on composite parts).

CST Microwave Studio (CST MWS) is a program package equipped by a high-quality automated mesher. The time-domain solver of CST MWS uses the Finite Integration Technique (FIT) and a hexahedral discretization scheme with Fast Perfect Boundary Approximation (FPBA). CST MWS records input geometry, transforms it to a surface and creates a mesh according to prescribed requirements. Impossibility of additional mesh adjustments is a disadvantage of this solution.

2.2. LIGHTNING WAVEFORM

Each natural flash of lightning is unique, and its parameters are different. Corresponding testing signals are precisely defined by a standard ED-84 – Aircraft Lightning Environment and Related Test Waveforms [6]. When constructing, the worst case should be considered – idealized standard external current waveforms – Component A.

The behavior of structures after being hit by lightning stroke is analyzed in time domain mostly. Component A, which is shown in Figure 4, is used for excitation. The time domain is described by the following equation:

$$i(t) = I_0(e^{-\alpha t} - e^{-\beta t}) \tag{1}$$

Here, I_0 = 218.810 A, α = 11.364 s^{-1}, and β = 647. 265 s^{-1}.

Concerning numeric simulations, calculations in frequency domain (namely, low-frequency phenomena) might be worth. Results obtained in the time domain are processed by inverse Fourier transformation (IFT). The spectrum (see Figure 5) is defined for the frequency domain as well. The spectrum shows that only frequencies up to 100 kHz are used, which eliminates computations at higher frequencies. The source can be described as:

$$i(\omega) = I_{sim}(\omega) \cdot I_{source}(\omega) \tag{2}$$

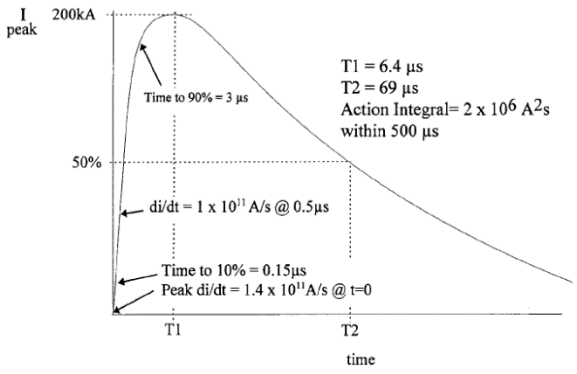


Figure. 4: Time domain waveform [6].

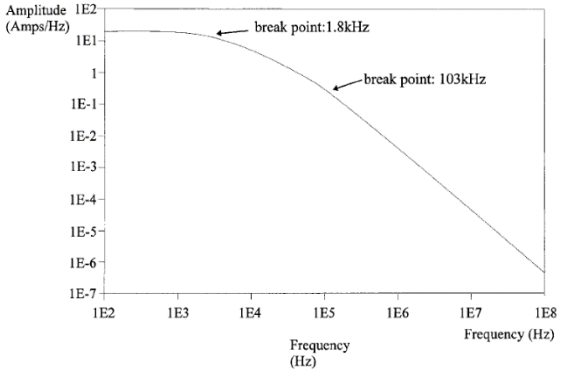


Figure 5: Frequency content [6].

3. RESULTS OF ANALYSIS

3.1. CURRENT DISTRIBUTION ANALYSIS

The computed current distributions without the composite skin are illustrated in Figure 6. Figure 7 shows current waveforms in individual diverters when the lightning strikes on Div 2.

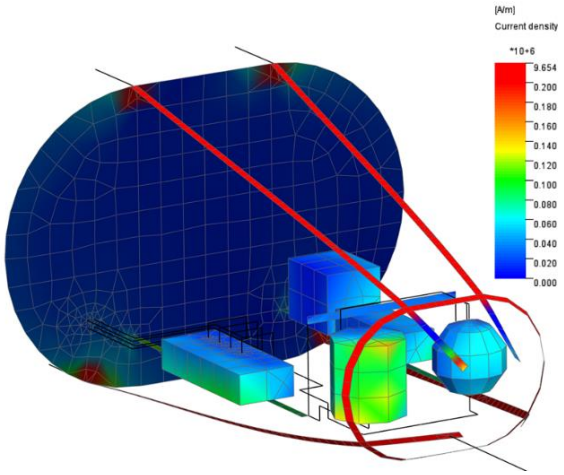


Figure 6: Current distribution without skin in time 6.4 μs.

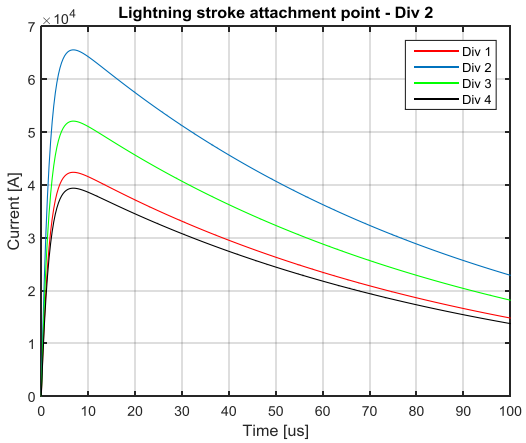


Figure 7: Current distribution – striking point Div 2.

3.2. COMPOSITE SKIN EFFECT

Adding the composite parts, we get more realistic model. As for the detection of carbon fiber composite (CFC) conductivity effects, the conductivity was set ten times higher, which can be achieved with appropriate orientation of carbon fibers. Figure 8 shows the distribution for the CFC skin value σ = 1 000 S/m (typical value) and Figure 9 for the value σ = 10 000 S/m (better conductivity for a compare).

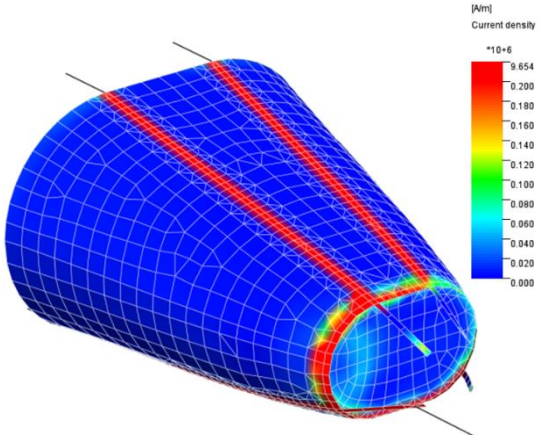


Figure 8: Current distribution with conductivity CFC σ = 1 000 S/m in time 6.4 μs.

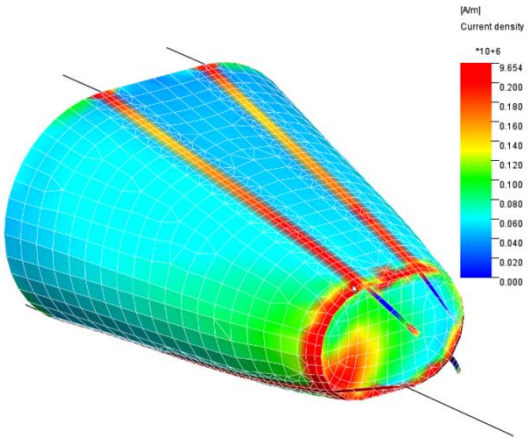


Figure 9: Current distribution with conductivity CFC σ = 10 000 S/m in time 6.4 μs.

The surface current evidently shows that current flows on the composite much more due to a higher conductivity. Focusing on the shape of current distribution itself, a change is evident. Figure 10 illustrates the comparison of current distribution on the metal strip with several conductivity adjustments – with and without composite. Composite materials significantly influence the current distribution on the surface of aircraft. That is the reason why the LPS must be included into design computations.

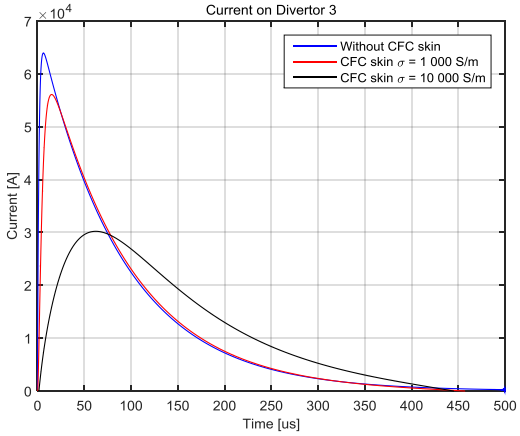


Figure 10: Comparison of time domain current distribution through divertor.

4. CONCLUSION

Nowadays, we can relatively quickly and accurately calculate the phenomena associated with airplane lightning strikes on powerful desktop computers. Thanks to consultations with designers, we can verify several possible lightning protections and integrate the best ones to aircraft prototype. The main task was focused on a suitable protection of built-in avionics and a hydraulic system. However, the lightest weight of protection as well as its mechanical strength is essential.

ACKNOWLEDGEMENT

The presented research was supported by Internal Grant Agency of Brno University of Technology project no. FEKT-S-14-2483. For research, infrastructure of the SIX Center was used. Special thanks belong to EVEKTOR Company for providing the aircraft model data.

REFERENCES

[1] Certification of Aircraft Electrical/Electronic Systems for the Indirect Effects of Lightning, ref. EUROCAE ED-81 &A1, August 1999.

[2] EUROCAE ED-14D/RTCA DO-160G (2010) – Section 22.

[3] Aircraft Accident Report, Boeing 707-12, N709PA Pan American World Airways, Inc., Near Elkton, Maryland, Dec. 8, 1963. Civil Aeronautic Board File n° 1-0015, http://en.wikipedia.org/wiki/Pan _Am_Flight_214, 1965.

[4] Wing Failure of Boeing 747-131 Near Madrid Spain May 9, 1976. Report n° NTSB-AAR-78-12.

[5] Airplane EV-55 [Online]. 2015[cit. 2015-4-22] Available on WWW: <http://www.evektor.cz/en/ ev-55-outback>.

[6] ED-84. Aircraft Lightning Environment and Related Test Waveforms. The European Organization for Civil Aviation Equipment, 1997. Including Amendment N. 1 - October 1999, Amendment N. 2 – May 2001 & Amendment N. 3 - June 2006.

SINGLE-ENDED FREQUENCY FILTER IN PRACTICAL REALIZATION WITH DIGITALLY ADJUSTABLE CURRENT AMPLIFIER

Josef Polak, Lukas Langhammer

Doctoral Degree Programme (3,4), FEEC BUT

E-mail: xpolak24@phd.feec.vutbr.cz, xlangh01@phd.feec.vutbr.cz

Supervised by: Jan Jerabek

E-mail: jerabekj@feec.vutbr.cz

Abstract: This paper deals with proposed 2nd order universal frequency filter. In design of this frequency filter, multiple-output current follower (MO-CF) and digitally adjustable current amplifier (DACA) have been used. The frequency filter is proposed with ability to control the pole frequency without disturbing the selected quality factor by parameter of DACA active elements in form of integrated circuit. Design is supported by simulation and also measurement results.

Keywords: DACA, MO-CF, frequency filter, controllable pole frequency

1. INTRODUCTION

Several papers dealing with design of the frequency filter with digitally adjustable current amplifier DACA have been published [1], [2], [3]. Very similar current active element to DACA is an adjustable current amplifier ACA [4]. Current gain of DACA can be controlled by 3-bit word, while current gain of ACA is controlled by DC current. Some of mentioned works deal only with design of frequency filters and this proposal is verified using suitable simulation program (in most cases PSpice). The other proposals of frequency filters are verified using not only simulation program and but also practical realization. For the practical realization of these filters, alternative solution of the DACA elements was frequently used in the past. Alternative circuit implementation of DACA element is really complex because it is realized using universal current conveyors UCC [5], universal voltage conveyor UVC [6] and EL2082 [7] circuits.

Verifying of correctness of output responses of the proposed frequency filter in this paper was carried out using PSpice, and also by practical realization with newly designed and manufactured active element DACA in form of the integrated circuit. This circuit was produced in 0.35 µm CMOS technology with digital control of current gain and it is designated as DACA_N [8].

2. DEFINITION OF USED ACTIVE ELEMENTS

In this section, used active elements are described. Each active element is presented by its schematic symbol and real implementation of the active element using suitable active element to practical realization (if necessary).

The first used active element is a digitally adjustable current amplifier (DACA). This element is controllable when using 3-bit digital bus, which serves for control current gain *A* of this active element. Current gain *A* can be controlled in range from 1 to 8. Fig. 1 shows a schematic symbol of the DACA active element. Relationships between low impedance current inputs and high impedance current outputs are described by the following equations:

$$I_{DIF\_IN} = I_{IN+} - I_{IN-}, \quad (1) \qquad I_{OUT+} = A(I_{IN+} - I_{IN-}), \quad (3)$$

$$I_{DIF\_OUT} = I_{OUT+} - I_{OUT-}, \quad (2) \qquad I_{OUT-} = -A(I_{IN+} - I_{IN-}), \quad (4)$$

$$I_{DIF\_OUT} = 2AI_{DIF\_IN}, \quad (5)$$

where I_{DIF_IN}, I_{DIF_OUT} are differential input and output currents, respectively. More detailed information about the DACA active element is in [9].

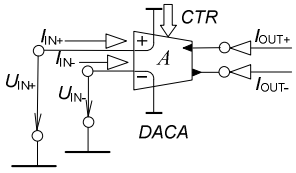


Figure 1: Digitally adjustable current amplifier schematic symbol.

The second used active element is a multiple-output current follower (MO-CF). This active element has one current input and four current output terminals and none of its parameters can be controlled electronically. The function of this element can be described by the following equation:

$$I_{IN} = I_{OUT1} = -I_{OUT2} = I_{OUT3} = -I_{OUT4}. \quad (6)$$

Schematic symbol and possible realization of MO-CF using universal current conveyor (UCC) [6] are illustrated in Fig. 2.

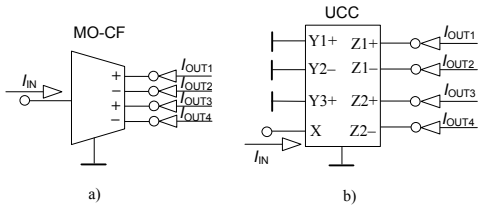


Figure 2: Multiple-output current follower a) schematic symbol, b) possible realization using UCC.

3. DEFINITION OF PROPOSED UNIVERSAL FREQUENCY FILTER

The filter presented in this paper is proposed as universal second order frequency filter with possibility to control the pole frequency f_0 without disturbing the selected quality factor Q. All filtering functions, i.e. inverting high pass (iHP), low pass (LP), inverting low pass (iLP), band pass (BP), inverting band stop (iBS) and inverting all pass (iAP) are taken directly from outputs of active elements. Scheme of the proposed frequency filter is shown in Fig. 3.

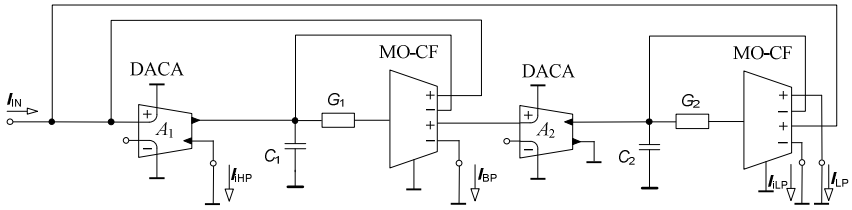


Figure 3: Scheme of the 2nd-order universal frequency filter.

The denominator of each of transfer functions of the proposed filter is as follows:

$$D(\mathrm{s}) = s^2 C_1 C_2 + sC_2 G_1 A_1 + G_1 G_2 A_1 A_2 , \tag{7}$$

The equation (8) for calculating quality factor (which is not electronically controllable in this particular case) is calculated from equation (7). The pole frequency is for this filter electronically controllable and the equation (9) shows calculation of pole frequency.

$$Q = \sqrt{\frac{C_1 G_2}{C_2 G_1}} , \tag{8}$$

$$f_0 = \frac{A}{2\pi}\sqrt{\frac{G_1 G_2}{C_1 C_2}}, \tag{9}$$

where $A = A_1 = A_2$.

The pole frequency of this filter is controllable by current gain $A = A_1 = A_2$ of two DACAs used.
The following equations represent the individual filtering functions:

$$\frac{I_{LP}}{I_{IN}} = \frac{G_1 G_2 A_1 A_2}{D(\mathrm{s})}, \tag{10}$$

$$\frac{I_{iLP}}{I_{IN}} = \frac{-G_1 G_2 A_1 A_2}{D(\mathrm{s})}, \tag{11}$$

$$\frac{I_{iHP}}{I_{IN}} = \frac{-s^2 C_1 C_2 A_1}{D(\mathrm{s})}, \tag{12}$$

$$\frac{I_{BP}}{I_{IN}} = \frac{sC_2 G_1 A_1}{D(\mathrm{s})}, \tag{13}$$

$$\frac{I_{iBS}}{I_{IN}} = \frac{-s^2 C_1 C_2 A_1 - G_1 G_2 A_1 A_2}{D(\mathrm{s})}, \tag{14}$$

$$\frac{I_{iAP}}{I_{IN}} = \frac{-s^2 C_1 C_2 A_1 + sC_2 G_1 A_1 - G_1 G_2 A_1 A_2}{D(\mathrm{s})}. \tag{15}$$

Inverting BS is obtained by sum of I_{iHP} and I_{iLP} outputs and then we receive output response corresponding to (14). Output response of iAP is represented by equation (15) and it is obtained by sum of I_{iHP}, I_{iLP} and I_{BP} outputs. Connections of other output responses are in Fig. 3.

4. COMPARISON OF SIMULATION AND MEASUREMENT RESULTS

The proposed frequency filtering functions were verified in simulation program OrCAD. PSpice simulation models of required active elements are mentioned in [9]. The results of practical measurement of this frequency filter are mentioned in comparison with simulation results for following values of passive elements.
The values of passive elements have been chosen for the ideal pole frequency f_0 = 1MHz and quality factor Q = 0.707 (Butterworth approximation). The values of capacitors were chosen as $C_1 = C_2$ = 100 pF and values of resistors R_1 = 1100 Ω, R_2 = 2200 Ω were calculated according the equations (16), (17).

$$R_1 = \frac{AQ}{f_0 2\pi C_1}, \tag{16}$$

$$R_2 = \frac{C_1 R_1}{Q^2 C_2}. \tag{17}$$

The functionality of the pole frequency tuning was verified by theoretical calculations (7), simulations and real measurements as demonstrated in Tab.1 for LP output response. Theoretical calculations are provided in two versions. One of them is truly ideal for ideal parameters of DACA element and the second is the case when the non-ideal current gain is taken into account (the worst case of gains used in PSpice model, derived from practically tested DACA_N integrated circuits). On this fact are based results in Tab. 1, where pole frequencies of measured values are nearer to theoretical (ideal) pole frequencies than simulated values of pole frequency. The results of simulated pole frequencies are nearer to theoretical (PSpice) pole frequencies.

Theoretical (calculated)				Simulated		Measured	
A (ideal)	f_0 [kHz]	A (PSpice)	f_0 [kHz]	A	f_0 [kHz]	A	f_0 [kHz]
1	1023	0.779	797	1	664	1	819
2	2046	1.594	1630	2	1405	2	1915
3	3069	2.489	2546	3	2287	3	2971
4	4092	3.299	3374	4	3069	4	4084

Table 1: Comparison of the pole frequency values of LP response.

Practical measurement results in comparison with simulation results are shown in the following graphs. Graphs in Fig. 4 and Fig. 5 compare the results of practical measurement (coloured solid curves) with simulation results (dashed black curves) for all filtering functions (iLP, iHP, BP, iBS and iAP) for A = 1. The pole frequency tunability of LP response is shown in Fig. 6 for values of gain corresponding to values mentioned in Tab. 1.

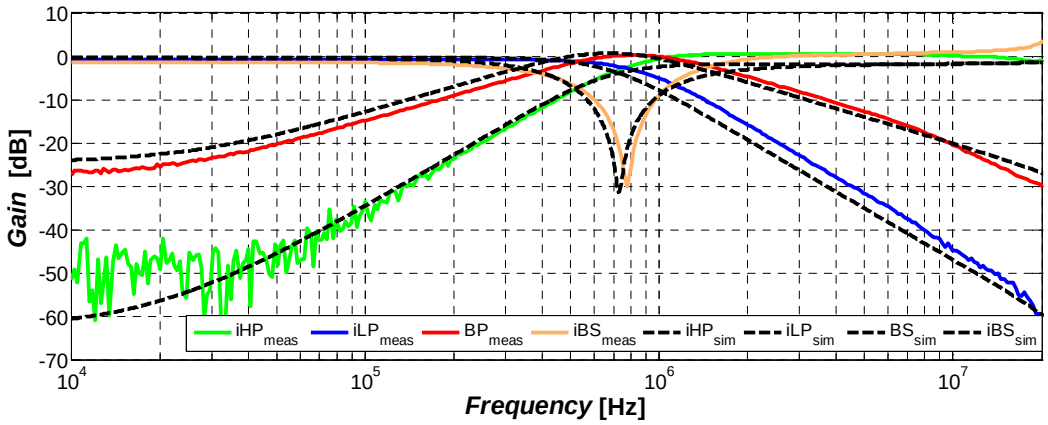


Figure 4: Comparison of magnitude responses (measured with simulated values) of iHP, iLP, BP, iBS (A = 1).

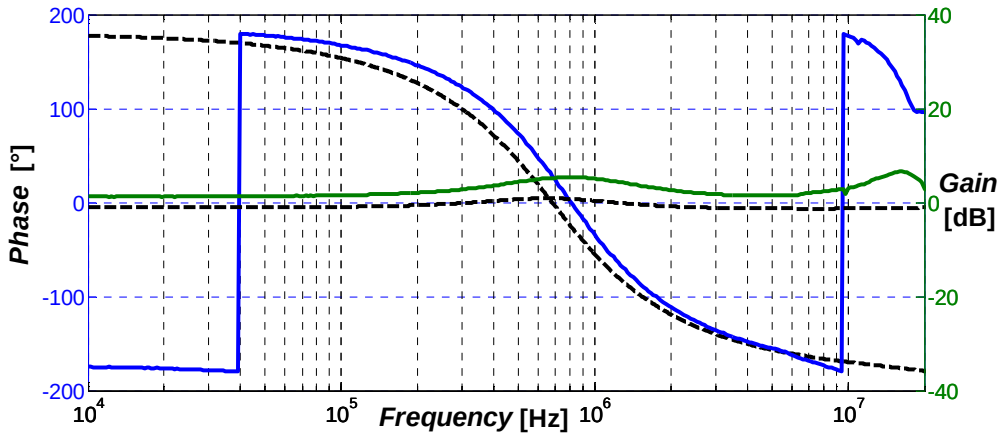


Figure 5: Comparison of magnitude and phase responses (measured with simulated values) of iAP (A = 1).

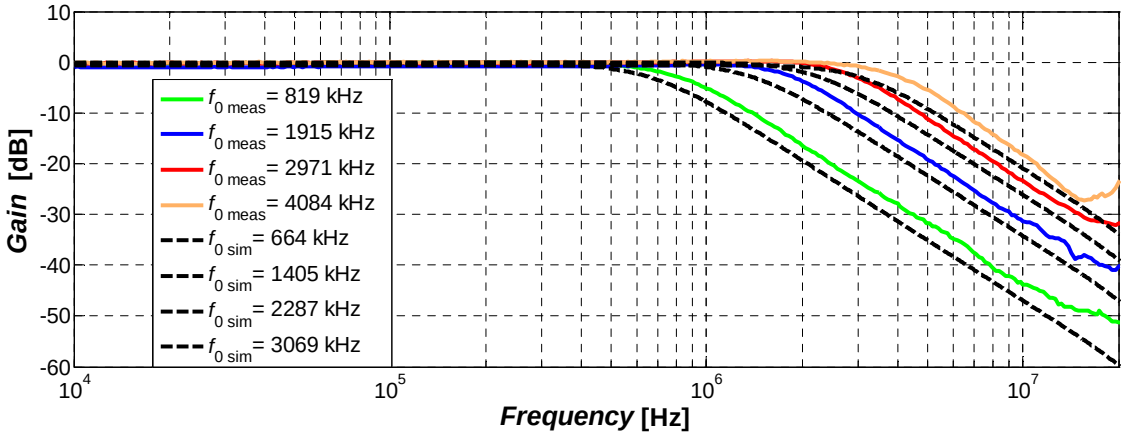


Figure 6: Demonstration of controlling the pole frequency f_0 of the LP filter (*A* from 1 to 4).

5. CONCLUSION

The new universal frequency filter of 2nd order is presented in this paper. The filter is designed to operate in current mode. The DACA current amplifier is used to control the pole frequency of this filter. Real experimental measurements of this filter were performed with DACA current amplifier in form of the integrated circuit and two additional MO-CFs. Simulation and experimental results are used to verify function of proposed frequency filter and some of them are included in this paper.

REFERENCES

[1] Jerabek, J., Sotner, R., Vrba, K.: Fully- Differential Universal Filter with Current Active Elements, In Proc of the 4th Int. Conf CSS'10, WSEAS, 2010, pp. 83-86

[2] Koton, J., Herencsar, N., Vrba, K., Jerabek, J.: Digitally Adjustable Current Amplifier and its Application in Fully Differential Current-Mode Band- Pass Filter Design, Elektrorevue Internet journal, 2010, No. 90, pp. 1-6, ISSN: 1213- 1539

[3] Langhammer, L., Jerabek, J.: Fully Differential Universal Current-Mode Frequency Filters Based on Signal- Flow Graphs Method, International Journal of Advances in Telecommunications, Electrotechnics, Signals and Systems, No. 1, pp. 1-12, ISSN: 1805- 5443

[4] Herencsar, N., Jerabek, J., Koton, J., Vrba, K., Mineai, S., Goknar, I. C.: Pole Frequency and Pass-Band Gain Tunable Novel Fully-Differential Current-Mode All- Pass Filter, In Proceedings of ISCAS, Portugal, 2015, pp. 2668-2671, ISBN: 978-1-4799-8391- 9

[5] Herencsar, N., Vrba, K., Koudar, I.: Universal Voltage Conveyor (UVC-N1C 0520), Data Sheet, 2010, online: http://www.utko.feec.vutbr.cz/~herencsar/UVC_N1C_0520.pdf

[6] Koton, J., Vrba, K., Koudar, I.: Universal Current Conveyor (UCC), Data Sheet, 2010, online: http://www.utko.feec.vutbr.cz/~koton/soubory/UCC_N1B_Rev0.pdf

[7] Intersil (Elantec). EL 2082CN Current-Mode Multiplier, 1996, http://www.intersil.com/data/fn/fn7152.pdf

[8] Jerabek, J., Vrba, K.—Koudar, I.: Dual Wide-Bandwidth Digitally Adjustable Current Amplifier (DACA), Data Sheet, 2010, online: http://www.utko.feec.vutbr.cz/~jerabekj/daca.pdf

[9] Polak, J., Langhammer, L., Jerabek, J.: Behavioral modeling of Digitally Adjustable Current Amplifier, International Journal Of Advances In Telecommunications, Electrotechnics, Signals And Systems, 2015, pp. 1-7, doi:10.11601/ijates.v4i1.104

TEXTILE SIW RECTANGULAR RING-SLOT ANTENNAS FOR OFF-BODY COMMUNICATION

Petr Vašina

Doctoral Degree Programme (3), FEEC BUT

E-mail: xvasin05@stud.feec.vutbr.cz

Supervised by: Jaroslav Láčík

E-mail: lacik@feec.vutbr.cz

Abstract: In this paper, a textile SIW rectangular ring-slot antenna for the ISM frequency band 5.8 GHz is proposed. The antenna radiates a linearly polarized wave with the maximum radiation in the perpendicular direction to the antenna plane. The antenna is fabricated on the soft-shell fabric in the two versions. In the first version, the conductive layers were fabricated by a self-adhesive copper foil. In the second version, the conductive layers were fabricated by the screen printing on a foil ironed on the soft-shell fabric. Experimental results prove that the proposed antenna fabricated by copper foil achieves the impedance bandwidth of 3.1 % for the reflection coefficient less than -10 dB and the gain of 6.91 dBi. While screen printing antenna achieves the impedance bandwidth of 2.93 % for the reflection coefficient less than -10 dB and the gain of 7 dBi, respectively.

Keywords: wearable antenna, slot antenna, screen printing, substrate integrated waveguide

1. INTRODUCITION

In recent years, technologies for wireless body centric communication at millimeter waves have been rapidly developed. Body-centric communication is considered as a part of fourth generation mobile communication system and might encounter concepts of a body area network (BAN) and personal area network (PAN) [1].The first introduction of an intelligent textile system embroidered into firefighters clothing was presented in [2]. The following generations of wearable antennas were textile antennas based on the substrate integrated waveguide (SIW) technology. In [3], a cavity-backed slot antenna for the communication between rescue workers was introduced. The antenna consists of a square SIW cavity and a slot etched in the top metal layer. The possibility of a SIW fabrication in the textile was published in [4, 5]. In [4], conducting fabric and silver-coated yarns were combined to realize a textile SIW. In [5], a flat flange cooper tube eyelets were applied to form a SIW cavity. Due to production of flexible, light-weight and robust body devices, a special attention to the antenna textile materials is given. Hence, their selection should be done with greatest care. In [6, 7], authors investigated the electrical behavior of some textile materials. Studies have shown that in terms of the low losses, it is the best to use textile material containing polyester.

In this contribution, we present a textile type of a linearly polarized SIW rectangular ring-slot antenna [8].

2. ANTENNA CONFIGURATION

The structure of the antennas is depicted in Fig. 1. The antenna consists of one layer SIW operating in the fundamental mode $TE_{10.}$ One end of the SIW is shorted. The rectangular ring-slot is located in the top conductive layer. The antenna is equipped by a microstrip line-to-SIW transition [9].

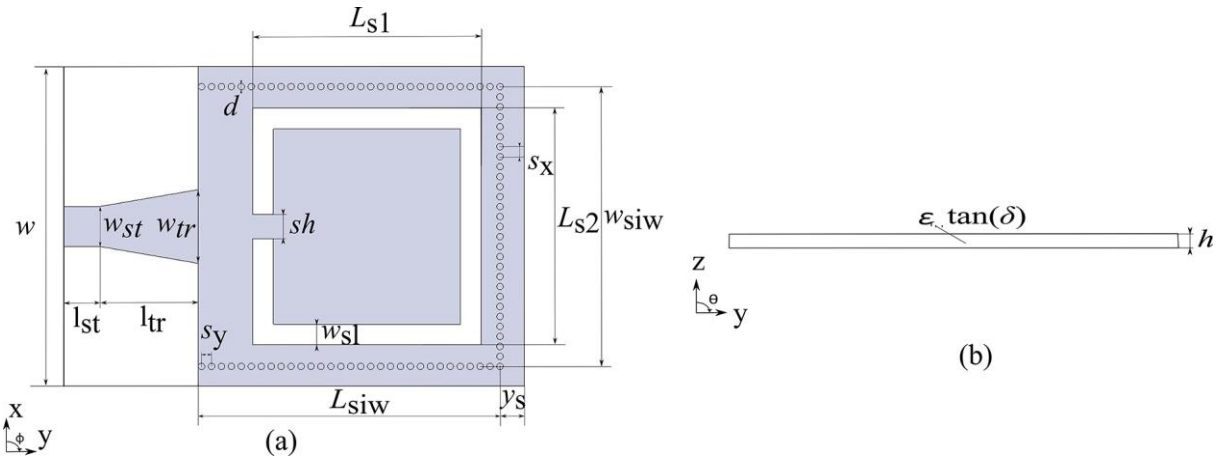


Figure 1: Geometrical configuration of proposed rectangular ring slot antenna: (a) Top side view, (b) Bottom side view.

The dielectric substrate of length $L_{siw}+y_s$, the width w, and the height h is on both sides covered by conductive sheets. The SIW is constructed by two rows of vias of the width w_{siw}, the diameter and the spacing s. In the proximity of the shorted SIW wall, the rectangular ring-slot radiator of the lengths *Ls1* x *Ls2* and the width w_{sl} is etched. The ring-slot is shorted by a narrow strip of the length *sh*.The feeding part of the antenna (Fig. 1(a)) consists of the microstrip line-to-SIW transition. The transition consists of the microstrip line (the width w_{st} and the length l_{st}) and the tapered microstrip line (the width w_{tr} and the length l_{tr} about λ/4 at the center frequency 5.8 GHz).

The proposed antenna (Fig. 1) was designed for the ISM frequency band 5.8 GHz with the help of CST Microwave studio on the soft-shell textile material with the relative permittivity ε_r = 1.4 and the thickness h = 1.27-1.37 mm. Relative permittivity was determined by reverse method. The width of the SIW (*siw_w* = 34.185 mm) is designed for the operating frequency 5.8 GHz in the fundamental mode TE_{10}. The cutoff frequency of the fundamental mode TE_{10} is 3.8 GHz. The resultant dimensions of the proposed antenna are summarized in Table I.

The proposed antenna radiates a linearly polarized wave. The operating frequency depends mainly on the length of the slot *Ls1* and slightly on the dimension of the slot *Ls2*. The impedance matching is influenced by the length of the slot *Ls2*, the width of the slot *Wsl* and the width of the short *sh*.

Table I: Parameters of the proposed antennas with microstrip line to SIW transition

Parameters	Dimensions [mm]	Parameters	Dimensions [mm]	Parameters	Dimensions [mm]
d	0.8	l_{tr}	11.82	w_{tr}	9.0
h	1.32	sh	2.63	w_{siw}	34.19
L_{s1}	26.54	s_x	1.20	w_{sl}	2.07
L_{s2}	28.13	s_y	1.22	y_s	3.02
L_{siw}	37.0	w	39.2		
l_{st}	4.75	w_{st}	4.9		

3. SIMULATED AND EXPERIMENTAL RESULT

The antenna was fabricated in two versions (Fig. 2). In the first version, the conductive layers were fabricated by a self-adhesive copper foil. In the second version, the conductive layers were fabricated by the screen printing on a Digiflex-Master foil ironed on the soft-shell fabric. For the printing, the Aurel C880 semiautomatic screen printer was used. The ESL 1901-S polymer silver con-

ductor paste with good electrical conductivity was used and cured to prevent temperature damage of the ironed foil and the fabric. Afterwards, the same fabrication procedure was applied for the bottom side of the antenna which was printed on the other side of the flexible textile fabric material. Walls of the SIW were created by the conductive thread. Antennas were measured merely in free space without the presence of phantom. Large differences can not be assumed because of the ground plane, which sufficiently suppresses interaction with the material placed under antennas.

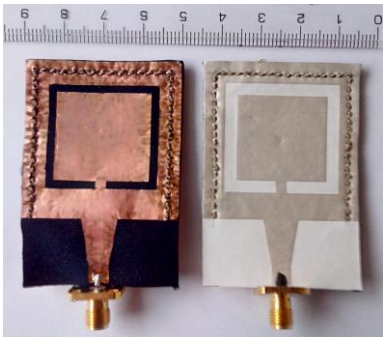


Figure 2: Fabricated samples of proposed textile antennas.

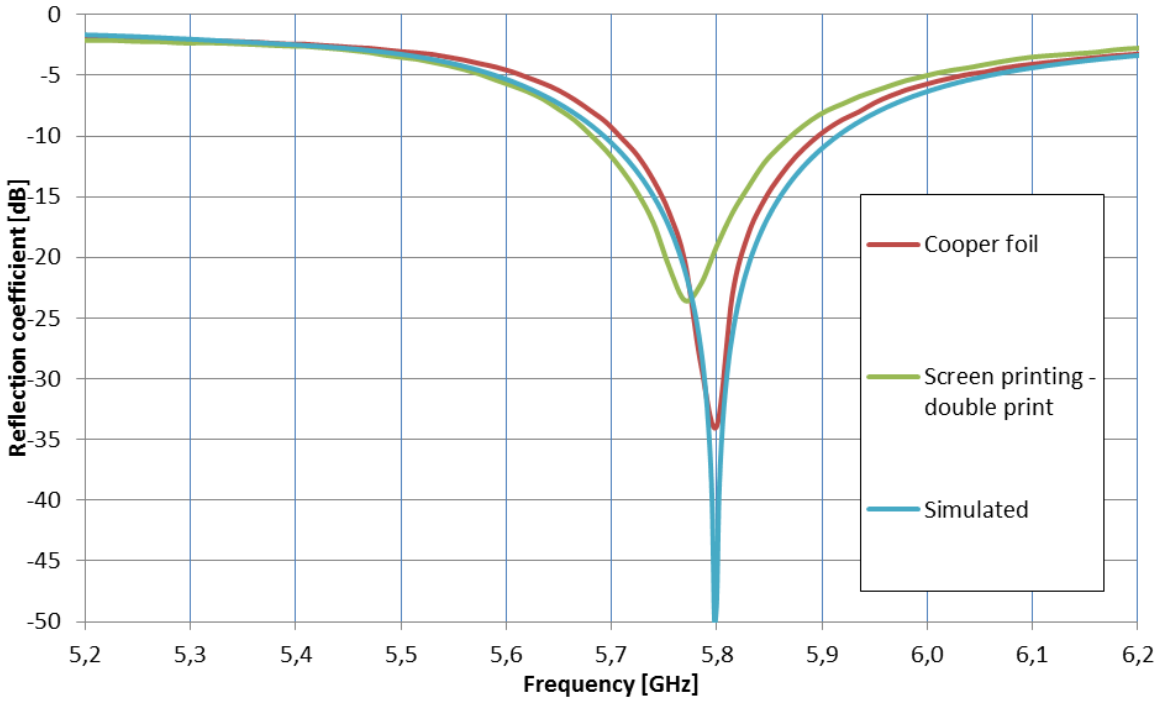


Figure 3: Reflection coefficient of proposed textile antenna realized in different versions.

The simulated reflection coefficient of the antenna is compared with fabricated antennas in Fig. 3. The SIW walls were made by hand, and therefore, the antennas are slightly detuned. However, despite slightly detuning the antennas covered the entire ISM frequency band 5.8 GHz (5.725–5.875 GHz) for the reflection coefficient less than -10 dB.

Simulated results of the radiation patterns were verified by measurement at the frequency 5.8 GHz (Fig. 4). The maximum radiation of the antennas is in the perpendicular direction to the antenna plane. In both planes we can see good agreement between measured and simulated data. In the XZ plane is shown small misalignment radiation characteristics, caused by the character wave propagation in the substrate and the radiation within the slot. At the bottom of the slot (with the short) flows through the lower magnetic current compared to the upper part and this creates non-symmetric radiation pattern.

The measured gain at the center frequency is 6.91 dBi (cooper antenna) and 7 dBi (screen printing antenna), 1 dB higher than simulated value due to a little bit lower dielectric losses of the soft-shell. From the known values of measured gain and simulated directivity was determined a radiation efficiency of proposed antennas. The radiation efficiency of cooper antenna is 74.1 % and 75.7 % for the screen printing antenna, respectively.

In comparison with the antenna [8], presented antenna achieves a smaller gain about 1.5 dBi which is due to the material used as dielectric substrate where soft-shell is more lossy than ARLON Cu Clad 217. Radiation pattern is the same for both antennas (perpendicular to the plane of the antenna) but antenna from [8] has a smaller angular displacement in the E-plane.

5.8 GHz	E-plane (xz-plane)	H-plane (yz-plane)

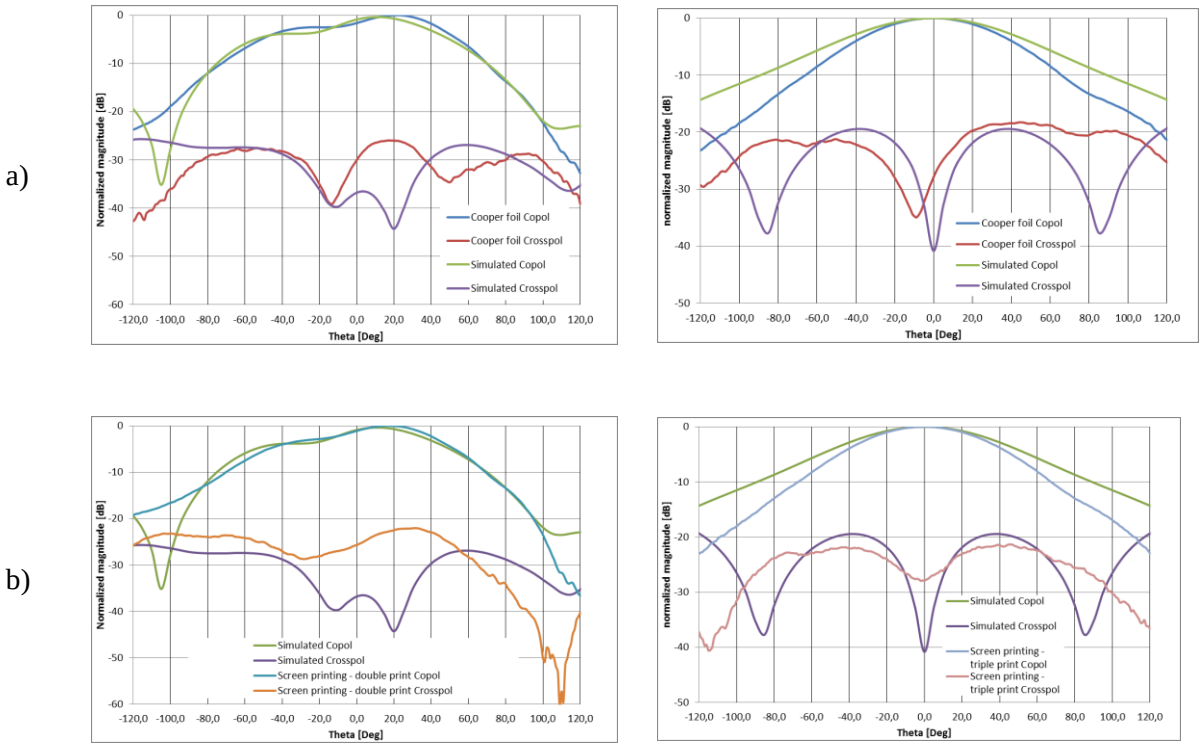


Figure 4: Normalized radiation pattern of proposed textile antenna at 5.8 GHz: (a) Cooper foil antenna, (b) Screen printing antenna – double print.

CONCLUSION

A textile rectangular ring slot antennas operating in the ISM frequency band 5.8 GHz has been proposed. The antennas were designed on a soft-shell material where conductive layers were fabricated by self-adhesive cooper foil and by screen-printing using the ESL 1901-S polymer silver conductor paste. The measured results showed that the antenna made of cooper foil achieved the impedance bandwidth of 3.1 % for the reflection coefficient less than -10 dB. The maximum of radiation was perpendicular to the antenna plane with gain 6.91 dBi. The screen printing antenna achieves the impedance bandwidth of 2.93 % for the reflection coefficient less than -10 dB and the gain of 7 dBi in the maximum of perpendicular radiation.

ACKNOWLEDGEMENT

The presented research was supported by the Czech Grant Agency project no. P102/12/1274, by the Czech Ministry of Education in frame of National Sustainability Program under grant LO1401, and by the Internal Grant Agency of Brno University of Technology project no. FEKT-S-14-2483. The research is the part of the COST Action IC1301 which is financially supported by the grant of the Czech Ministry of Education no. LD14057. For research, infrastructure of the SIX Center was used.

REFERENCES

[1] Hall, P.S.; Yang, Hao.: Antennas and Propagation for Body-centric Wireless Communications, 2nd ed. Boston: Artech House, 2012, XIII, 387 s. antenna and propagation series. ISBN 978-1-60807-376-4.

[2] Herleer C., Rogier H., Vallozzi L., L.V. Langenhove L.V.: A textile antenna for off-body communication integrated into protective clothing for firefighters, IEEE Trans. Antennas Propag., vol. 57, no. 4, pp. 919–925, Apr. 2009.

[3] Moro R., Agneessens S., Rogier H., Bozzi M.: Wearable textile antenna in substrate integrated waveguide technology, Electron. Lett., vol. 48, no. 16, pp. 985–986, 2012.

[4] Kaufmann T., Xu Z., Fumeaux C.: Wearable substrate-integrated waveguide with embroidered vias, in Proc. 8th Eur. Conf. Antennas Propag., The Hague, The Netherlands, 2014.

[5] Lemey S., Rogier H.: SIW textile antennas as a novel technology for UWB RFID tags, in Proceedings of RFID Technology and Applications Conference (RFID-TA). Tampere, Finland: IEEE, September 2014, pp. 256–260.

[6] Ouyang Y., Chappell W.J.: High frequency properties of electro-textiles for wearable antenna applications, IEEE Trans. Antennas Propag., vol. 56, no. 2, pp. 381–389, Feb. 2008.

[7] Asanovic K.A., Mihajlidi T.A., Milosavlejevic S.V., Cerovic D. D. Chappell J. R.: Investigation of the electrical behavior of some textile materials, Journal of Electrostatics., 2006.

[8] Lacik J., Mikulasek T.: Substrate integrated waveguide rectangular ring slot antenna, International Conference on Electromagnetics in Advanced Applications (ICEAA), pp.1164-1167, 12-16 Sept. 2011.

[9] Deslandes, D.: Design equations for tapered microstrip-toSubstrate Integrated Waveguide transitions, Microwave Symposium Digest (MTT), 2010 IEEE MTT-S International , vol., no., pp.704,707, 23-28 May 2010.

THE DESCRIPTION OF THE PROGRAM FOR FACTORIZATION OF LARGE NUMBERS

Vladimír Levek

Doctoral Degree Programme (5), FEEC BUT

E-mail: levek@feec.vutbr.cz

Supervised by: Pavel Šteffan

E-mail: steffan@feec.vutbr.cz

Abstract: The article describes the algorithm for factorization of large numbers. If there is the result of the product of two prime numbers, then the program can find the factors. The first part of the article generally introduces the problem of factoring large integers and its impact in the field of the cryptography. The next part describes the algorithm and the program for calculation. At the end of the article there is a summary of the possibilities of the program.

Keywords: Cryptography, Factorization, RSA, Prime number, Large Numbers, Order of magnitude

1. INTRODUCTION

The resistance of cryptography, especially asymmetric cryptosystems IF is based on the inability to decompose a large number into the product of smaller integers – factors for the reasonable time. Large number is a number in order of magnitude 10^{300} and higher and the "reasonable time" could be within several months. This task is currently considered insoluble, therefore asymmetric cryptography belongs to the category of highly resistant cryptosystems.

2. ASYMMETRIC CRYPTOSYSTEM RSA

Like any other asymmetric cryptosystems, RSA is primarily designated for encryption of short messages. It is given by computational demands of this system resulting into practical impossibility of the decryption in a short time. The message intended for the encryption must be shorter than numbers used for its security and decryption. The basic mathematical operations for encryption and decryption of messages are

$$C = Z^{VK} \bmod n, \tag{1}$$

$$Z = C^{SK} \bmod n, \tag{2}$$

where C is a message Z encrypted with a public key (VK, n). The message is decrypted with a private key SK. The condition: message Z must be shorter than n, which equals the product of p and q:

$$n = p \cdot q, \tag{3}$$

p a q are prime numbers larger than 10^{150}. These numbers are chosen in such a way that the conditions given below are valid. These numbers also serve for the calculation of modulo r (4) – which is used for the calculation of the private key SK (5). [2][3].

$$r = (p-1)(q-1) \tag{4}$$

$$(VK \cdot SK) \bmod r = 1. \tag{5}$$

Based on the relations above, there is a possibility to get the private key with factorization of the formula (3). As stated above p a q are large prime numbers, therefore it is obvious that the result n is final and therefore factorization will not have more solutions. If cybernetic attacker manages to express p and q from n, he could get r (4) and subsequently he could calculate the private key with the use of condition in (5).

3. FACTORING ALGORITHMS

Algorithm for the calculation of the factors results from eventual compilation of all the possibilities from the lowest order of magnitudes. In each order of magnitude beginning with the lowest, there is a calculation of possible option of factors of the proper order of magnitude. Previous results could be confirmed or disproved. General product A x B can be decomposed to elementary order of magnitudes with the use of modular mathematics.

The product is:

$$A \cdot B = \left(a_0.10^0 + a_1.10^1 + a_2.10^2 ... + a_n.10^n\right) \cdot \left(b_0.10^0 + b_1.10^1 + b_2.10^2 ... + b_n.10^n\right), \quad (6)$$

where A and B are factors and a_x and b_x are integer from 0 - 9. The decomposition of the formula (6) makes a mathematical equation:

$$A \cdot B = a_0 b_0.10^0 + \left(a_0 b_1 + a_1 b_0\right)10^1 + \left(a_0 b_2 + a_1 b_1 + a_2 b_0\right)10^2 ... + \left(a_0 b_n + a_1 b_{n-1} ... + a_{n-1} b_1 + a_n b_0\right)10^n . \quad (7)$$

The formula 7 shows the way the result in a particular order of magnitude is made. For example digit in the result in the second order of magnitude is created in this way:

$$c_2.10^2 = \left(a_0 b_2 + a_1 b_1 + a_2 b_0\right).10^2 . \quad (8)$$

Regarding the fact that the final digit must be between 0-9, the only result would be in the unit order of magnitudes. The whole result is made by modulo of tens. During this process the transmissions to the higher orders of magnitude should be applied. Generally, digits in a particular order of magnitude are made by these formulas:

$$\begin{aligned}
c_0 &= \left(a_0 b_0\right) \bmod 10 \\
c_1 &= \left(a_0 b_1 + a_1 b_0 + P_0\right) \bmod 10 \\
c_2 &= \left(a_0 b_2 + a_1 b_1 + a_2 b_0 + P_1\right) \bmod 10 \\
c_3 &= \left(a_0 b_3 + a_1 b_2 + a_2 b_1 + a_3 b_0 + P_2\right) \bmod 10 \\
c_4 &= \left(a_1 b_3 + a_2 b_2 + a_3 b_1 + P_3\right) \bmod 10 \\
c_5 &= \left(a_2 b_3 + a_3 b_2 + P_4\right) \bmod 10 \\
c_6 &= \left(a_3 b_3 + P_5\right) \bmod 10 \\
c_7 &= P_6 \\
c_8 &= 0
\end{aligned} \quad (9)$$

$$\begin{aligned}
P_0 &= \frac{a_0 b_{0_1} - \left(a_0 b_{0_1}\right) \bmod 10}{10} \\
P_1 &= \frac{\left(a_0 b_1 + a_1 b_0 + P_0\right) - \left(a_0 b_1 + a_1 b_0 + P_0\right) \bmod 10}{10} \\
P_2 &= \frac{\left(a_0 b_2 + a_1 b_1 + a_2 b_0 + P_1\right)}{10} \\
&\quad - \frac{\left(a_0 b_2 + a_1 b_1 + a_2 b_0 + P_1\right) \bmod 10}{10} \\
P_3 &= \frac{\left(a_0 b_3 + a_1 b_2 ... + a_3 b_0 + P_2\right)}{10} \\
&\quad - \frac{\left(a_0 b_3 + a_1 b_2 ... + a_3 b_0 + P_2\right) \bmod 10}{10} \\
P_4 &= \frac{\left(a_1 b_3 + a_2 b_2 + a_3 b_1 + P_3\right)}{10} \\
&\quad - \frac{\left(a_1 b_3 + a_2 b_2 + a_3 b_1 + P_3\right) \bmod 10}{10} \\
P_5 &= \frac{\left(a_2 b_3 + a_3 b_2 + P_4\right) - \left(a_2 b_3 + a_3 b_2 + P_4\right) \bmod 10}{10} \\
P_6 &= \frac{\left(a_3 b_3 + P_5\right) - \left(a_3 b_3 + P_5\right) \bmod 10}{10}
\end{aligned} \quad (10)$$

where *c* are final digits as results of modulo and transmissions *P* are defined by the overall result, which is reduced by modulo ten.

4. THE DECRYPTION OF THE PROGRAM

The program for the decomposition of the product was made for calculation of factorization. For the simplification of the whole problematics, it is important to realize that the product of prime numbers will end with digits: 1, 3, 7 or 9. Therefore, the result cannot be even and it is highly probable that it will not end with digit 5.

Number 5 is the only prime number which ends with digit 5. It is supposed that the factors are larger than single-digit. Program assumes that the order of magnitude of the factors is at least two times lower than the order of magnitude of the result.

The programs involves the following steps:

1. Defining the length of factors.

2. Estimation of zero order of magnitude of factors.

3. Calculation of the "corpus" of the order of magnitudes (factors of previous order of magnitudes).

4. Estimation of factors (factors of current order of magnitude).

5. Testing if assumption corresponds with the final calculation

6. If the condition in 5 is fulfilled, the order of magnitude is increased and program continues from the point 2

7. When the calculation in the highest order of magnitude is finished, program finishes, too.

In the first part of the program, the length of factors is calculated. The next part of the program consists of algorithm which determines digits that could be used in zero order of magnitude. "Factor" covers digits of factors in particular order of magnitude – a_x, b_x. The program subsequently tests in dynamic loop all possible variants in particular orders of magnitude until the result is attained. If the result is not attained, the examined number is prime number.

General calculation of each digit in particular order of magnitude is:

$$c_n = \left(a_0 b_n + a_1 b_{n-1} + a_2 b_{n-2} + \ldots + a_n b_0 + P_{n-1}\right) \bmod 10 \,. \tag{11}$$

Because the process of calculation proceeds from the lower orders of magnitude to higher, it is obvious that in 11 there are two unknown variables - a_n a b_n. Other numbers are calculated previously in lower orders of magnitude. Finally, the calculation of the orders of magnitude is divided into a part of "estimation of factors" and a part dealing with calculation of "corpus of the order of magnitude".

$$C_n = \begin{pmatrix} a_0 b_n + a_n b_0 & + \\ a_1 b_{n-1} + a_2 b_{n-2} + a_3 b_{n-3} \ldots + a_n b_0 & + \\ P_{n-1} & \end{pmatrix} \bmod 10 \,. \tag{12}$$

In formula (12) there are three important points:

- Estimation of factors
- Calculation of "corpus of the order of magnitude"
- Addition of the transmission from the lower order of magnitude

While the calculation of "the corpus" is determined by the sum of permutated factor of lower orders of magnitude, estimation of factors is determined by generating factors of current order of magnitude and their multiplying with zero factors.

$$a_0 b_n + a_n b_0 . \tag{13}$$

The relation (13) shows a calculation of newly generated digits, which are generated from 0. In each step it is tested whether the condition in (13) and (14) is valid.

$$a_1 b_{n-1} + a_2 b_{n-2} + a_3 b_{n-3} ... + a_n b_0 , \tag{14}$$

In the figure 1, there is a flowchart of the whole program.

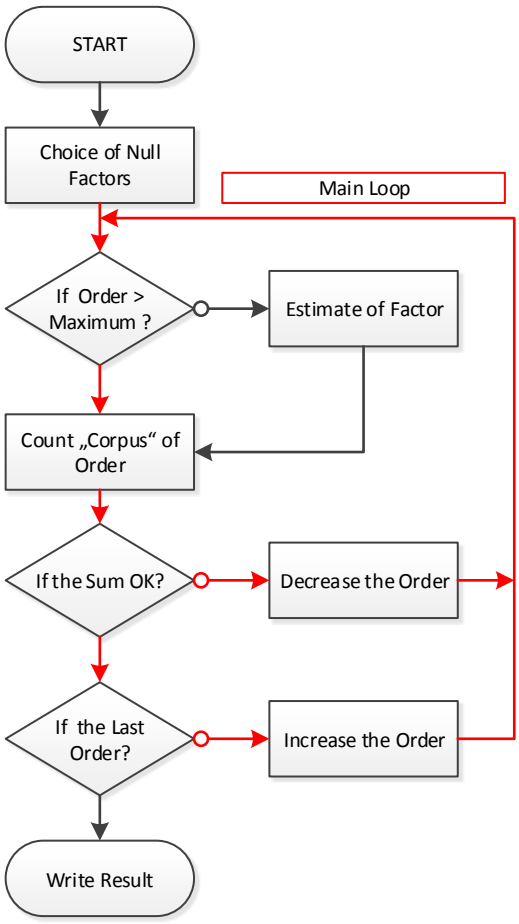


Figure 1: A Flowchart of program for factoring

Red arrows mark the working cycle of the program. Basically, in each step the program compares the final result with the assumption. If they correspond to one another, the program starts to calculate in higher order of magnitude. If the result does not correspond with the assumption, the program starts to calculate in lower order of magnitude where the combination of other factors is generated. Afterwards, the calculation proceeds to the higher order of magnitude again. If the attained order of magnitude is higher than the the reach of the order of magnitude higher than the maximum length of factors there would be no more factors estimated, only transmissions from lower orders of magnitude would be completed. If the program tests all the possibilities and no result is attained, the processed number is prime number.

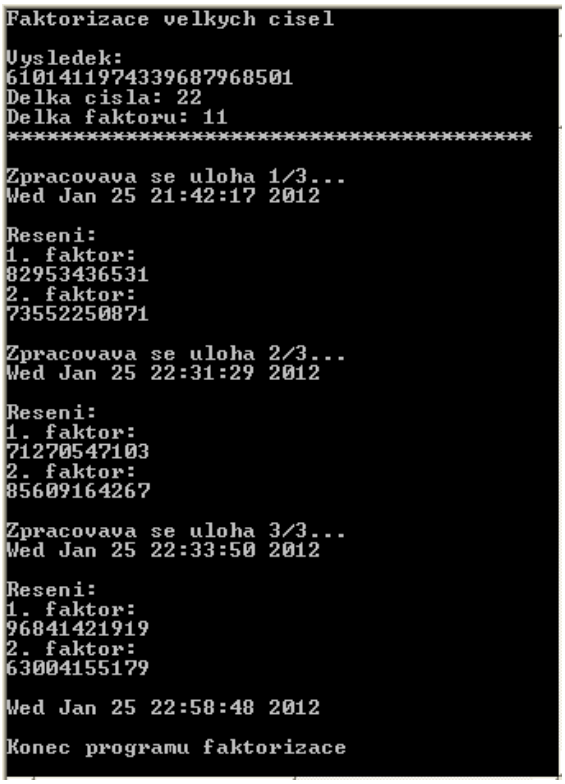


Figure 2: Demonstration of resulting console factorization 22-digit number

5. CONCLUSION

This article describes the formation of algorithm for factorization of large numbers. Figure 2 shows time needed for calculation of factorization 22-digit number (the time is 3000 sec). Obviously, this is not reasonable time for the practical use of this program. Program was made in language C++ in the environment of Microsoft Visual Studio 2010. When the program was on the run, the processor of the computer was fully occupied too, therefore it was impossible to perform any other tasks. Program was made for demonstrative and educational purposes and it definitely does not aim to be the significant tool for the break of cryptosystem RSA.

ACKNOWLEDGEMENT

The work was supported by the Brno University of Technology under project no. FEKT-S-14-2300: “New types of electronic circuits and sensors for specific applications”.

REFERENCES

[1] BURDA, K. Aplikovaná kryptografie (přednáška 7) Asymetrické kryptosystémy. VUT v Brně, rok neuveden

[2] Wikipedia: RSA. [online]. [cit. 2011-12-18]. Dostupné z: http://en.wikipedia.org/wiki/RSA

[3] RSA: RSA History. *RSA* [online]. [cit. 2011-12-18]. Dostupné z: http://www.rsa.com/node.aspx?id=2760

Doktorské projekty

Zpracování signálů, obrazu a dat

SPARSE SIGNAL RECOVERY VIA CONVEX OPTIMIZATION

Michaela Novosadová
Doctoral Degree Programme (2.), FEEC BUT
E-mail: xnovos11@stud.feec.vutbr.cz

Supervised by: Pavel Rajmic
E-mail: rajmic@feec.vutbr.cz

Abstract: We propose recovering 1D piecewice linear signal using a sparsity-based method consisting of two steps. The first step is signal segmentation via optimization algorithms solving sparsity based model. Second step consists of applying an ordinary mean square method on each detected segment of the signal. We show results of our experiments on two types of the signal.

Keywords: convex optimization, Douglas-Rachford algorithm, Forward-Backward algorithm, signal recovery, signal segmentation

1 INTRODUCTION

We explore signal recovering by use combination of convex optimization methods for segmenting of noisy piecewise linear signal and signal denoising by use of mean square method at each detected segment of the signal. The number of segments is considerably lower than the number of signal samples, which suggests using sparse signal processing techniques [1].

2 PROBLEM FORMULATION

The overparameterization model should be introduced first. We suppose 1D piecewise linear signal $\mathbf{f}$. Each element of the signal $\mathbf{f}$ can be described by two parameterization coefficients: constant offset (parameter a) and constant slope (parameter b). So the i-th element of the signal is defined as $f_i = a_i + b_i i$, alternatively

$$\mathbf{f} = \mathbf{a} + \mathbf{D}\mathbf{b} \quad \text{i.e.} \quad \mathbf{f} = [\mathbf{I}\mathbf{D}] \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \mathbf{A}\mathbf{x}, \tag{1}$$

where $\mathbf{a}$, $\mathbf{b} \in \mathbb{R}^N$, $\mathbf{I} = \mathbf{I}_N$ is the identity matrix, $\mathbf{D} = \mathrm{diag}(1, 2, \ldots, N)$ is the diagonal matrix with the values $1, 2, ..., N$ on its main diagonal. Note that vectors $\mathbf{a}$ and $\mathbf{b}$ are piecewise constant in each segment so they are sparse under the difference operation. Due to this assumption we can formulate the recovery problem using total variation (TV):

$$\hat{\mathbf{a}}, \hat{\mathbf{b}} = \underset{\mathbf{a},\mathbf{b}}{\arg\min} \; \frac{1}{2} \left\| \mathbf{y} - [\mathbf{I}\mathbf{D}] \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \right\|_2^2 + \tau_{\mathbf{a}} \mathrm{TV}(\mathbf{a}) + \tau_{\mathbf{b}} \mathrm{TV}(\mathbf{b}), \tag{2}$$

where $\mathbf{y}$ is observed signal which is corrupted by uncorrelated Gaussian noise $\mathbf{e}$ with zero mean and positive variance $\mathbf{y} = \mathbf{f} + \mathbf{e}$, and $\mathrm{TV}(\cdot)$ is the total variation functional defined as

$$\mathrm{TV}(\mathbf{z}) = \|\nabla \mathbf{z}\|_1 = \sum_{i=1}^{N-1} |z_{i+1} - z_i|, \tag{3}$$

and finally $\hat{\mathbf{a}}, \hat{\mathbf{b}}$ are achieved optimizers. The nonzero values in $\nabla\hat{\mathbf{a}}$ and $\nabla\hat{\mathbf{b}}$ indicate possible segment borders, the nonzero values should be on shared positions. Positive constants $\tau_{\mathbf{a}}, \tau_{\mathbf{b}}$ are regularization weights, with their values depending on the properties of the signal and noise level. The regularization weights should be carefully tuned.

3 METHODOLOGY

The solution of signal denoising is divided into two parts. First, the breakpoints are detected describing each segment of the signal. This is done by proximal splitting methods. Second is the denoising of each segment of the signal by least squared method.

3.1 SIGNAL SEGMENTATION

Signal segmentation is achieved by finding the breakpoints. To identify breakpoints, it is necessary to solve the optimalization problem (2). For solving this problem proximal splitting algorithms can be used. Proximal splitting algorithm is an iterative way to minimize a sum of convex functions by repetitive evaluation of their gradients or proximal operators. The basic optimalization problem is

$$\arg\min_{\mathbf{x}} \; f_1(\mathbf{x}) + f_2(\mathbf{x}). \tag{4}$$

For solving this optimalization problem, two appropriate algorithms are available, namely the Forward-Backward and Douglas-Rachford algorithm [2]. Is proven that proximal algorithms converge to the optimal value which is minimum in case that (4) is convex. The convergence is in practice influenced by the character of functions and the choice of parameters of algorithm.

3.1.1 Used proximal splitting algorithms

We assign: $f_1(\mathbf{x}) = \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2$, $f_2(\mathbf{x}) = f_2(\mathbf{a}, \mathbf{b}) = \tau_{\mathbf{a}}\,\mathrm{TV}(\mathbf{a}) + \tau_{\mathbf{b}}\,\mathrm{TV}(\mathbf{b})$.

Forward-Backward (FB) splitting algorithm solves problem (4) where $f_1(\mathbf{x})$ is convex and differentiable with a β-Lipschitz continuous gradient $\nabla f_1(\mathbf{x}) = \mathbf{A}^\top(\mathbf{A}\mathbf{x} - \mathbf{y})$ [3]. Proximal operator of $f_2(\mathbf{x})$ is $\mathrm{prox}_{f_2}(\mathbf{x}) = \mathrm{prox}_{f_2}(\mathbf{a}, \mathbf{b}) = \begin{bmatrix} \mathrm{prox}_{\tau_{\mathbf{a}}\mathrm{TV}(\cdot)}(\mathbf{a}) \\ \mathrm{prox}_{\tau_{\mathbf{b}}\mathrm{TV}(\cdot)}(\mathbf{b}) \end{bmatrix}$. The proximal operator of $\mathrm{TV}(\cdot)$ for 1D signals can be computed fast using the Condat's algorithm [4]. The algorithm consists of a forward (gradient) step using function f_1 and a backward (proximal) step using function f_2.

Douglas-Rachford (DR) algorithm does not require function f_1 having β-Lipschitz continuous gradient. This algorithm consists of two proximal steps using functions f_1 and f_2. Instead of the gradient step, proximal operator of f_1 is used, which is $\mathrm{prox}_{f_1}(\mathbf{x}) = (\mathbf{I} + \tau\mathbf{A}^\top\mathbf{A})^{-1}(\mathbf{x} + \tau\mathbf{A}^\top\mathbf{y})$ [3].

3.1.2 Detection of breakpoints

From the solution $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$, the breakpoints can be established. Nonzero values in $\nabla\hat{\mathbf{a}}$ and $\nabla\hat{\mathbf{b}}$ indicate possible segment borders. In theory, the nonzero values should share the same position, but in practice they are not and there are found more segments than there should be. This happens because it is complicated to set regularization parameters to achieve piecewise constant solution in $\mathbf{a}$ and $\mathbf{b}$ so that in $\nabla\hat{\mathbf{a}}$ and $\nabla\hat{\mathbf{b}}$ are found more nonzero values. Because of this we used thresholding with threshold λ_a for $\nabla\hat{\mathbf{a}}$ and threshold λ_b for $\nabla\hat{\mathbf{b}}$ to find appropriated segment borders. After thresholding is created vector of breakpoints as:

$$\mathbf{bp} = [0, sort(\mathbf{bp}_a \cup \mathbf{bp}_b), N], \tag{5}$$

where $\mathbf{bp}_a$ and $\mathbf{bp}_b$ are positions of nonzero values in $\nabla\hat{\mathbf{a}}$ satisfying condition $|\hat{a}_i| > \lambda_a$ and in $\nabla\hat{\mathbf{b}}$ satisfying condition $|\hat{b}_i| > \lambda_b$, respectively. N is length of signal $\mathbf{y}$.

3.2 SIGNAL DENOISING

In each detected segment we perform simple denoising of the signal by least squared method finding optimal parameters:

$$\beta_k = (\mathbf{X}_k^T \mathbf{X}_k)^{-1} \mathbf{X}_k^T \mathbf{y}(bp_k : bp_{k+1}), \tag{6}$$

where $\beta_k = \begin{bmatrix} \dot{a}_k & \dot{b}_k \end{bmatrix}^T$

is vector with parameter $\dot{a}$ and $\dot{b}$ (offset and slope) belonging to the k-th segment of signal $\mathbf{y}$, and $\mathbf{X}_k = \begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ bp_k & bp_k+1 & \dots & bp_{k+1}-1 & bp_{k+1} \end{bmatrix}^T$.

For each segment k we get parameter $\dot{a}_k$ and parameter $\dot{b}_k$. With obtained parameters we can reconstructed the signal $\mathbf{y}$, and observed the denoised signal $\hat{\mathbf{y}}$ according to:

$$\hat{\mathbf{y}} = \dot{\mathbf{a}} + \mathbf{D}\dot{\mathbf{b}}, \tag{7}$$

where $\dot{\mathbf{a}}$ and $\dot{\mathbf{b}}$ contain parameter $\dot{a}_k$ resp. parameter $\dot{b}_k$ for each element of k-th segment of $\mathbf{y}$.

4 EXPERIMENTAL RESULTS

Two experiments of signal denoising were performed. First experiment was performed with sawtooth signal and the second with the randomly generated signal.

4.1 RESULTS FOR SAWTOOTH SIGNAL

A periodic sawtooth signal of length $N = 150$ with line slope equal to 1 was generated. Vector $\mathbf{b} \in \mathbb{R}^N$ is a vector of ones. Vector $\mathbf{a} \in \mathbb{R}^N$ is, according to the assumption, piecewise constant, and for each subsequent segment, it takes value lowered by $L = 25$, which is the selected period length. Computing $[\mathbf{ID}][\mathbf{a}\,\mathbf{b}]^\top$ and adding Gaussian IID noise then synthesizes the noisy sawtooth.

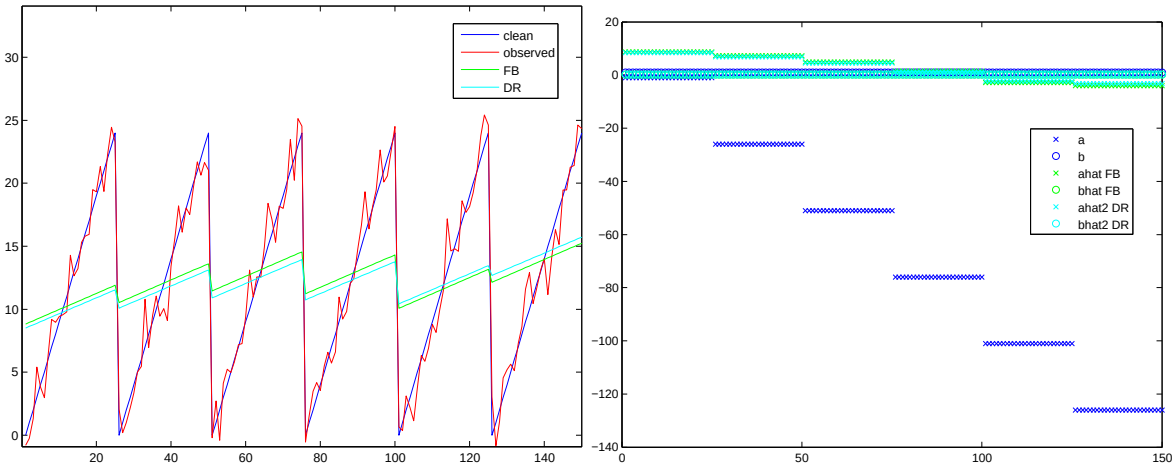


Figure 1: Approximating sawtooth signal after first step with $\tau_\mathbf{a} = 51.8$ and $\tau_\mathbf{b} = 12540$. We have $\mathrm{TV}(\mathbf{b}) = 0$, $\mathrm{TV}_{FB}(\hat{\mathbf{b}}) = 0$, $\mathrm{TV}_{CR}(\hat{\mathbf{b}}) = 0$, $\mathrm{TV}(\mathbf{a}) = 55.9$, $\mathrm{TV}_{FB}(\hat{\mathbf{a}}) = 6.29$, $\mathrm{TV}_{DR}(\hat{\mathbf{a}}) = 5,65$. Signal to noise ratio (SNR) of the observed signal is $SNR = 19.68dB$, recovered signal after first step has $SNR_{FB} = 6.83dB$, $SNR_{DR} = 6.77dB$.

Regularization parameters $\tau_\mathbf{a}, \tau_\mathbf{b}$ and threshold parameters λ_a and λ_b were carefully tuned to obtain stepwise parameter vector $\mathbf{a}$ producing signal that has breakpoints at the same positions as the clean one does. The resulting signal of first step and the parameters found are depicted in Fig. 1. Recovered sawtooth signal $\hat{\mathbf{y}}$ and its parameters are depicted in Fig. 2.

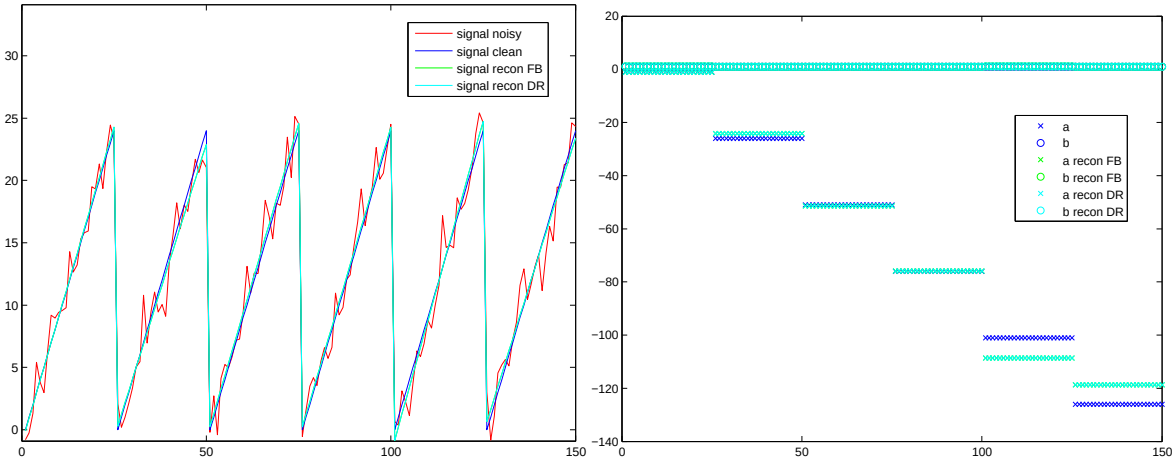


Figure 2: Approximating sawtooth signal after second step with $\lambda_{\mathbf{a}} = 0.075$ and $\lambda_{\mathbf{b}} = 0.075$. SNR is $SNR = 19.68dB$, of the recovered signal after second step $SNR_{FB} = 30.77dB$, $SNR_{DR} = 30.77dB$.

4.2 Results for Randomly Generated Signal

Second, we performed a similar experiment on a randomly generated signal of length $N = 150$ with five linear segments. Vectors $\mathbf{a} \in \mathbb{R}^N$ and $\mathbf{b} \in \mathbb{R}^N$ are, according to the assumption, piecewise constant. Computing $[\mathbf{ID}][\mathbf{a}\,\mathbf{b}]^\top$ and adding Gaussian IID noise then synthesizes the noisy signal. The resulting signal of breakpoints detection and the parameters found are depicted in Fig. 3. Recovered random signal $\hat{\mathbf{y}}$ and its parameters are depicted in Fig. 4.

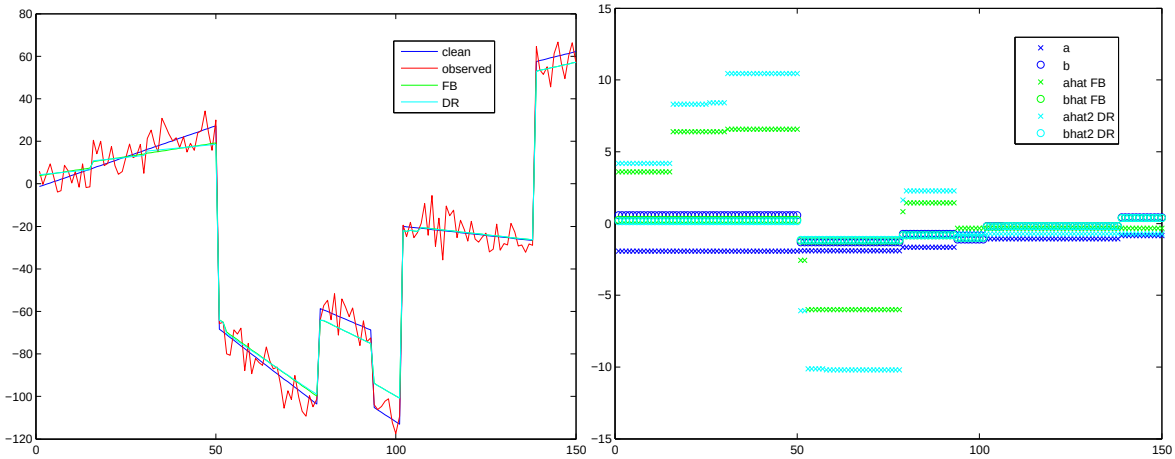


Figure 3: Approximating randomly generated signal after first step $\tau_{\mathbf{a}} = 43.27$ and $\tau_{\mathbf{b}} = 3850$. We have $\mathrm{TV}(\mathbf{b}) = 2.3$, $\mathrm{TV}_{FB}(\hat{\mathbf{b}}) = 1.8$, $\mathrm{TV}_{CR}(\hat{\mathbf{b}}) = 1.66$, $\mathrm{TV}(\mathbf{a}) = 1.1$, $\mathrm{TV}_{FB}(\hat{\mathbf{a}}) = 12.4$, $\mathrm{TV}_{DR}(\hat{\mathbf{a}}) = 21.45$. SNR of the observed signal is $SNR = 18.91dB$, recovered signal after first step has $SNR_{FB} = 21.86dB$, $SNR_{DR} = 21.8dB$.

5 CONCLUSION

The experiments present the fact that the suggested approach gives relatively good results. Detected positions of breakpoints in $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ are not the same although it is supposed that they should be on the same positions. Such a problem, however, can be solve by enforcing joint breakpoints using a group-sparse model, which can be provided by usage of a ℓ_{21} mixed norm in (2) instead of a used ℓ_1 norm. This will be the topic of further research.

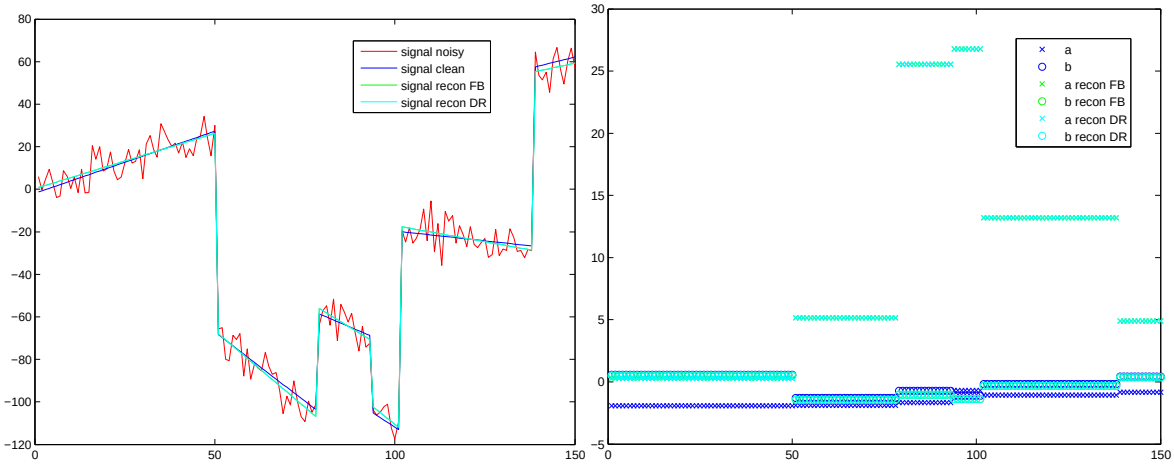


Figure 4: Approximating randomly generated signal after second step $\lambda_{\mathbf{a}} = 8$ and $\lambda_{\mathbf{b}} = 0.15$. SNR of the observed signal is $SNR = 18.91dB$, of the recovered signal after second step $SNR_{FB} = 31.03dB$, $SNR_{DR} = 31.03dB$.

REFERENCES

[1] Giryes, R.; Elad, M.; Bruckstein, A.M.: Sparsity Based Methods for Overparameterized Variational Problems, SIAM journal on imaging sciences, vol. 8, no. 3, pp. 2133 - 2159, 2015.

[2] Combettes, P.L.; Pesquet, J.C.: Proximal splitting methods in signal processing, Fixed-Point Algorithms for Inverse Problems in Science and Engineering, Springer, 2011, DOI: 10.1007/978-1-4419-9569-8_10

[3] Condat, L.: A Generic Proximal Algorithm for Convex Optimization—Application to Total Variation Minimization, Signal Processing Letters, IEEE, vol. 21, no. 8, pp. 985-989, 2014, DOI: 10.1109/LSP.2014.2322123

[4] Condat, L.: A Direct Algorithm for 1-D Total Variation Denoising, Signal Processing Letters, IEEE, vol. 20, no.11, pp. 1054-1057, 2013, DOI: 10.1109/LSP.2013.2278339

ANALYSIS OF SPEECH SIGNALS IN VOIP TRANSFERRED OVER A NETWORK CHANNEL

Josef Báňa

Doctoral Degree Programme (1), FEEC BUT

E-mail: banaj@feec.vutbr.cz

Supervised by: Hicham Atassi

E-mail: atassi@feec.vutbr.cz

Abstract: The article examines the impact of VoIP transmission channel characteristics on speech parameters. It seeks ways of emulating properties of a selected VoIP transmission channel and creating a network emulation environment. Several scripts have been created in the Matlab development environment and used to modify and divide a continuous recording into parts identical to the original speeches, before passing through the transmission channel. Subsequently a database of speech recordings is created, as affected by selected characteristics - jitter, bandwidth, loss. Within these databases, symptoms are sought as the most evident characteristics of the transmission channel. Using correlation, symptoms are selected that are best suited for automated determination of the properties of transmission channel characteristics such as jitter, loss and bandwidth.

Keywords: Speech parameter, VoIP, transmission channel, WAN network emulation, jitter, bandwidth, loss.

1. INTRODUCTION

The work described in this paper focuses on automatic detection of transfer channels and explores the effect of transfer channels on speech parameters (symptoms). Finding specific characteristics of transfer channels like GSM mobile network, cable phone network and VoIP (Voice over Internet Protocols). Knowledge of the specifics and characteristics of the transmission channels influencing the properties of speech brings options exercisable in large automated speech analysis.

The voice transfer in internet network – VoIP is an interesting part of transfer channels. In this article we focus on a deeper analysis of VoIP and explore specific, typical properties of the Ethernet network and network parasites affecting the quality of the resulting transmitted voice signal.

2. EXPERIMENTAL SETUP

2.1. AUDIO CODECS IN VOIP

VoIP – this technology uses the standard IP network protocol for transfer of digitized voice over the Ethernet network and data are encapsulated into UDP datagrams. In general terms it is a connectionless and unreliable service.

Audio codecs – these are used for digitalization and compression of an audio signal. The general purpose of codecs is to capture the input signal as faithfully as possible, and use as few bytes of digitized information as possible. After receiving and decompression of the signal we should ideally not recognize the original signal from the digitized when hearing it.

Features of some of the most commonly used audio codecs:

G.722 – A standardized broad-band codec, originally used in telephone lines. Previously, it was a proprietary. Since 2008 it has been free to use, its license expired. Coded audio bandwidth is

50 – 7000 Hz. Sampling frequency is 16 kHz (using 14 bits). Transmission speeds are 48, 56 a 64 kbps. [1]

G.722.2 (WB-AMR) – It is not a better version of G.722. WB-AMR (Adaptive Multi-Rate Wideband) is an independent codec, developed by Nokia and VoiceAge. The ITU standardized labelling is G.722.2. It is mostly used in mobile networks of the third generation. Sampling frequency is 16 kHz (using 14 bits). Transmission speeds are 6.6, 8.85, 12.65, 14.25, 15.85, 18.25, 19.85, 23.05, 23.85 kbps. Sound quality is same like G.722, but it's 2.5 times more efficient. [2]

Silk – A newer codec used since 2009 in Skype programme. Sampling frequency can be 8, 12, 16 or 24 kHz and changes dynamically (it depends on the connection speed, packet loss, etc.). Transmission speed is between 6 - 40 kbps. [3]

Speex – an open-source audio codec developed by the community. It uses 3 sampling rates for voice digitalization: 8 kHz for narrowband transmission channels, 16 kHz for broadband transmission channels and 32 kHz for ultra-wideband transmission channels in one data stream. Transmission speed is between 2 do 44 kbps. [4]

2.2. EMULATION OF NETWORK PARASITES

Used laboratory involvement of emulated network is shown on Figure 1**.** PC-play plays the audio source with speech playlist from the TIMIT database. There is a beep signal between every speech. Sound output from PC-play is connected to input Host 1. A connection is established from PC Host 1 to Host 2 over Ethernet network. Audio sound is encoded by G.722 codec. Connection between Host 1 and Host 2 is routed through PC-WANem, where emulated transmission parasites. WANem [5] is a program emulating network parameters, which is similar in WAN networks. They include characteristics like bandwidth, packet loss, delay, jitter or change of packet order. WANem is an Open Source program developed under GPLv2 licenses.

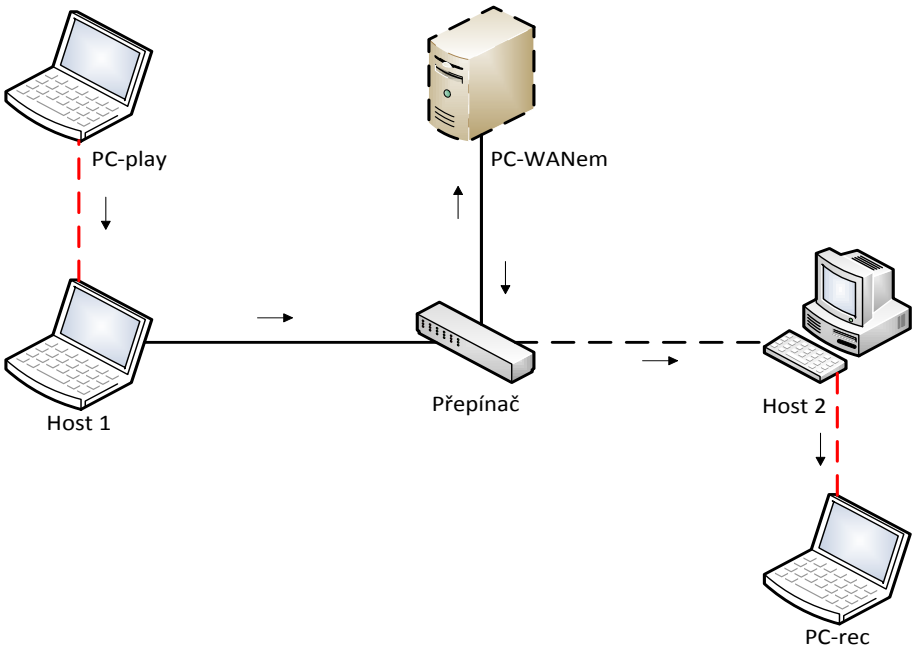


Figure 1: Emulated laboratory system

2.3. THE BASIC NETWORK PARAMETERS

The basic parameters affecting the audio signal transmission in the IP network using the SIP protocol are bandwidth, loss, delay of UDP datagrams, jitter, packet damage, reordering of sent and delivered datagrams.

Bandwidth – is the bit-rate of available or consumed information capacity expressed typically in metric multiples of bits per second.
Jitter – fluctuation of delay.
Losses – losses of packets.

In connection with the audio codecs, we mainly focus on codec resistance against greater losses in the network. Figure 2 shows a cut-out of the word "close" from speech number 58.

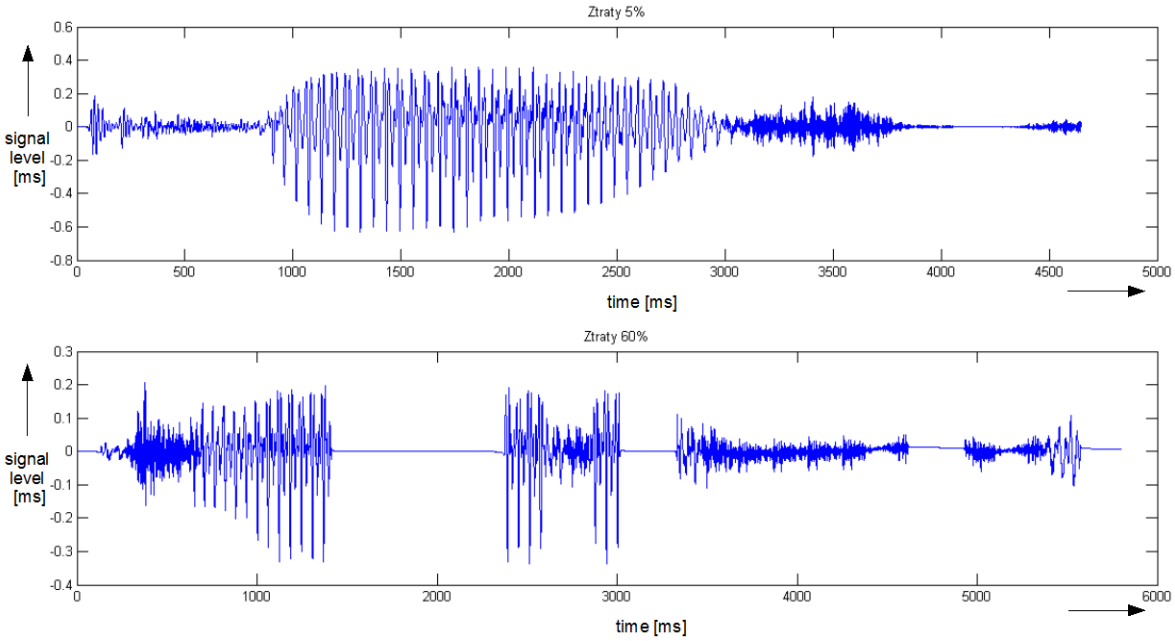


Figure 2: Cut out of the word "close" at a 5% loss (above) and 60% loss (below).

3. METHODOLOGY OF RECOGNITION PARAMETERS IN TRANSMISSION CHANNEL

3.1. REGRESSION TREES

Ranking classifies data into predefined categories, the algorithm learns the characteristics of the classes from the training set and on the data, for which we know their affiliation to groups. [6]

Decisions trees are widely used, in this work they were used as one of the possible classification models. The decision tree can be defined as a tree, where each non-leaf node represents a test for the attribute value and branches leading from this node are the possible results of this test. Classification runs through the record from tree root to leaf. In every step there will be a test and the journey continues along the test result branch. Once a recording reaches the leaf node, it is classified in categories identified by the value of tree leaves. [7]

3.2. SYMPTOMS

Out of different groups of symptoms we selected some representing symptoms and these were tested further. From perceptual spectral features we selected LMCC and MFCC symptoms. From the Energies and ZCR group we used Temporal Energy (TpEn) and Teager energy operator (TEO) symptoms for further work. The Spectral features group is kept as a separate feature.

3.3. SELECTION BASED ON CORRELATION

The equation for calculating the correlation coefficient R [8]:

$$R = \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \cdot \sum (y_i - \bar{y})^2}} \tag{1}$$

where *x* and *y* are tested variables.
Coefficient of Determination R^2 can be calculated according to the equation [8]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - y(x_i))^2}{\sum_{i=1}^{n} (y_i - \langle y_i \rangle)^2} \tag{2}$$

Where *x* is x x-coordinate of the data point,
y is y x-coordinate of the data point,
n is the number of data points,
$\langle y_i \rangle$ it can be calculated as follows:

$$\langle y_i \rangle = \frac{\sum_{i=1}^{n} y_i}{n} \tag{3}$$

According to the index value (coefficient), the correlation determines the degree of dependency. When Index (coefficient) correlation values mean:

R^2 = 0,0 – 0,3 weak dependence,
R^2 = 0,3 – 0,7 average dependence,
R^2 = 0,7 – 0,9 strong dependence,
R^2 = 0,9 –1,0 very strong dependence.

Mean absolute deviation is given by the equation [9]:

$$\text{Mean absolute deviation} = \frac{1}{n} \sum \left| x_i - \bar{x} \right|. \tag{4}$$

Mean-square deviation is given by the equation [9]:

$$\text{Mean - square deviation} = \frac{1}{n} \sum \left(x_i - \bar{x} \right)^2. \tag{5}$$

The table always features a symptom with the highest correlation (or a combination of symptoms having a positive impact on achieving correlation) shown in red text. For Jitter parameter we calculated correlation values with every symptom, as seen in Tab. 1.

JT [ms]	MFCC	LFCC	SpFe	TEO	TpEn
R^2	0,61	0,58	0,65	0,00	0,59
R	0,81	0,79	0,83	0,23	0,80

Table 1: Calculated correlation for symptoms [10].

Subsequently we calculated cross-correlation of the five symptoms. The symptoms were then separated based on those that do not affect the result of the correlation (or which negatively affect the result of the correlation). Only those combinations of symptoms that positively affect the correlation remained. Figure 3 for example shows a correlation with Spectral coefficients (SpeFe) symptoms for the Jitter (JT) parameter. [10] The best correlation prevails while using ca. 140 Spectral coefficients.

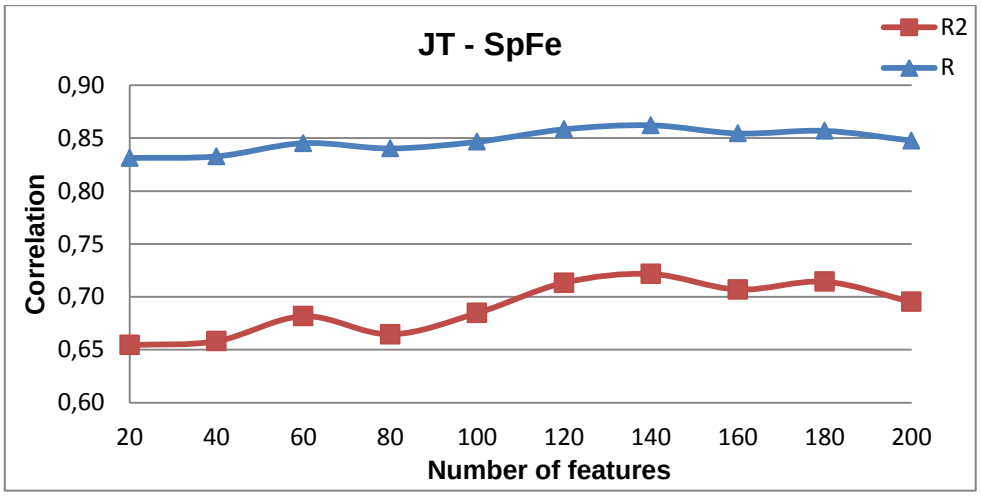


Figure 3: Correlation for symptom SpeFe (Spectral coefficients). [10]

4. CONCLUSION

We have explored the speech characteristics, transmission characteristics of the network channel and especially sound characteristics of a VoIP channel. A method for simulation of network parasites was proposed. A network simulation environment was realized by using a WANem simulator. The network simulation environment was created with many continuous recordings that were used after treatment in the training database. Three groups of databases were created: Losses - LS [%], Jitter JT - [ms] and bandwidth BW [kbps]. Every group contains 7 databases. Every database is created by many groups of loss levels. Every database contains 120 speeches.

In order to quickly determine the value of jitter or loss from the recordings, simply use the Spectral Features symptom. To determine the bandwidth value it is necessary to create a database with highest bandwidth of 256 kbps. However, it is difficult to divide such a database into separate speeches.

Results obtained in this work can be applied in police investigations for example. Police obtain an unknown recording, determine the transmission channel through speech analysis and need to find more parameters about this channel in short time.

ACKNOWLEDGEMENT

The presented research was supported by the Internal Grant Agency of the Brno University of Technology, project no. FEKT-S-14-2483 and INWITE project;) registration number LO1401 – Interdisciplinary Research of Wireless Technologies. The research was performed in laboratories supported by the SIX project; registration number CZ.1.05/2.1.00/03.0072, Operational Programme Research and Development for Innovation.

REFERENCES

[1] G.722 : 7 kHz audio-coding within 64 kbit/s: Recommendation G.722 (09/12). *International Communication Union* [online]. 2012-09 [cit. 2016-03-30]. Dostupné z: http://www.itu.int/rec/T-REC-G.722-201209-I/en.

[2] G.722.2 : Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB): Recommendation G.722.2 (07/03). ITU. *International Communication Union* [online]. 2003, 2004-09-07 [cit. 2016-03-30]. Dostupné z: https://www.itu.int/rec/T-REC-G.722.2-200307-I/en.

[3] SILK Speech Codec. KUNEŠ, Jan, ASTROM, Henrik a VOS, Koen. SKYPE LIMITED. Skype: *Skype developers* [online]. 2010 [cit. 2016-03-30]. Dostupné z: http://dev.skype.com/silk.

[4] Speex: A Free Codec For Free Speech. XIPH.ORG. *Speex* [online]. 2007, 2015 [cit. 2016-03-30]. Dostupné z: http://www.speex.org.

[5] WANem: The Wide Area Network emulator. TATA CONSULTANCY SERVICES LTD. *WANem* [online]. 2007, 2014 [cit. 2016-03-30]. Dostupné z: http://wanem.sourceforge.net

[6] DATA MINING KLASIFIKACE. *DMINA* [online]. 2009 [cit. 2016-03-30]. Dostupné z: http://eamos.pf.jcu.cz/amos/kat_inf/externi/kat_inf_21586/files/studijni_texty/klasifikace.pdf

[7] MAREK, RYCHLÝ. *Klasifikace a predikce* [online]. 2005 [cit. 2016-03-30]. Dostupné z: http://www.fit.vutbr.cz/~rychly/public/docs/classification-and-prediction/xhtml/classification-and-prediction.xhtml.

[8] SYNEK, VÁCLAV. Statistika. *UNIVERZITA JANA EVANGELISTY PURKYNĚ V ÚSTÍ NAD LABEM: FAKULTA ŽIVOTNÍHO PROSTŘEDÍ: FAKULTA ŽIVOTNÍHO PROSTŘEDÍ* [online]. 2013 [cit. 2016-03-30]. Dostupné z: http://fzp.ujep.cz/~synek/#_Skripta.

[9] RŮŽIČKOVÁ, Irena a Břetislav FAJMON. *Matematika 3* [online]. Brno, 2005 [cit. 2016-03-30]. Dostupné z: http://www.umat.feec.vutbr.cz/~novakm/matematika3.pdf. Elektronický text. VUT v Brně FEKT.

[10] BÁŇA, Josef. *Klasifikace přenosových kanálů na základě analýzy řečového signálu*. Brno, 2013. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Vedoucí práce Ing. Hicham Atassi.

ECG SIGNAL CLASSIFICATION BASED ON SVM

Radovan Smíšek

Doctoral Degree Programme (1), FEEC BUT

E-mail: xsmise00@stud.feec.vutbr.cz

Supervised by: Jana Kolářová

E-mail: kolarova@feec.vutbr.cz

Abstract: Cardiovascular diseases nowadays represent the most common cause of death in Western countries. Long-term ECG recording is modern method, because it allows to detect sporadically occurring pathology. We designed an automatic classifier to detect five pathologies (AAMI standard) by SVM method. The classifier was tested on the entire MIT-BIH Arrhythmia Database with an accuracy of 99.17 %. We also compared the quality of parameters entering the classifier.

Keywords: ECG classification, support vector machines, SVM, MIT-BIH database

1. INTRODUCTION

Classification of ECG is still the current task in the field of signal processing. Reliable classifier is needed especially for modern long-term ECG (Holter, telemetry). Cardiologist cannot see the entire signal that is often sensed several days. Classifier helps to find pathologies in the signal. The main aim of this study is to propose a reliable classification algorithm and test it on a standard publicly available database.

Several classification methods were developed to classify the ECG in other studies. Frequently used methods are support vector machines (SVM) [1, 2], artificial neural network (ANN) [3, 4], nearest neighbor (NN) [5, 6] and decision tree (DT) [7, 8]. In this study, we used SVM method.

Cycles of ECG are not usually entering the classifier, but parameters (called features) derived from the signals are entering. Features derived from time domain are usually used. This may be the width and height of the waves, skewness, kurtosis, area under the curves and so on. [1, 3, 5] RR intervals are also frequently used features. Other features may also be used in more advanced algorithms. It can be features derived from the frequency or time-frequency domain, features calculated using wavelet transform and higher order statistics features (HOS). [1, 2, 3, 4, 5] Selection of entering features is very important. The classifier is unable to distinguish categories, if there is only small number of features or features are chosen inappropriately. A large number of features leads to large computational complexity and the resulting success rate may be lower. Universal recommendation of number of features is not easy to create. One of the aims of this study is to compare the various types of features and determine their suitability.

MIT-BIH Arrhythmia Database [9, 10] was selected for testing the quality of the classifier. The database contains 48 records, which are about 30 minutes long. Each record has a two leads. Sampling frequency is 360 Hz, the resolution is 11-bit within a 10 mV range. Each QRS complex was classified by cardiologists into one of the fifteen heartbeat types. These groups can be united into 5 groups according to the AAMI standard [11]: supraventricular ectopic beat (S), ventricular ectopic beat (V), fusion beat (F), unknown beat (Q) and any heartbeat not in the S, V, F or Q classes (N). The entire database contains 90,510 N beats, 2774 S beats, 7703 V beats, 802 F beats and 8031 Q beats. In this work, we classified into these five groups. [9, 10]

The following section describes the test data, pre-processing, features extraction and description of the selected classifier. In chapter 3 are presented and discussed the results. Finally, the conclusion of the paper is in chapter 4.

2. METHODS

Basic steps of the proposed algorithm are shown in the block diagram in Figure 1. Each block will be described in detail in this section.



Figure 1: Block diagram of the proposed algorithm

2.1. PRE-PROCESSING

During the pre-processing was removed baseline wandering. It was removed by zeroing the spectral components before calculation of features. The advantage of this method compared to conventional filters is that it does not introduce delays. We removed components corresponding to the frequency band 0-0.5 Hz. Removing this band has no significant influence on the useful signal.

2.2. FEATURES EXTRACTION

Entry into the classifier is a set of features which characterize the QRS complex. Values of each feature should be the most different between groups that we want to separate and on the other hand, the values of feature should be approximately the same for complexes of the same type. Features used in this study were derived from the time domain, frequency domain, time-frequency domain and time-scale domain (calculated using the continuous wavelet transform). Meaningfulness of using these features is described in the following paragraph.

Features derived from the time domain to describe the morphology of the QRS complex (such as R wave amplitude, skewness and kurtosis of QRS complex, area under the QRS complex and QRS width) vary depending on place of origin of action potential and method of action potential conduction. QRS complex has sharp edges, and it is narrow at the physiological state. When the activation pulse is formed in the chamber or it is blocked at some point in its transmission, QRS complex becomes wider and has less sharp edges. The features in the frequency domain are thus also changing. If there are less sharp edges of the QRS complex, the energy of high frequency components decreases. Decomposition using wavelet transform (WT) allows detailed analysis the useful low frequency components, while high frequency components are not decomposed. A special group of features are RR intervals before, after and in the wider area of the analyzed QRS complex. Using the RR intervals is necessary to distinguish supraventricular ectopic beats, because these beats do not differ in shape from physiological QRS complex. The difference is only in the previous and subsequent RR interval. The importance of different groups of parameters is discussed in the section 3.

All parameters were calculated from both available leads. Features were normalized before entering the classifier. This means that the software centers and scales each feature by the mean and standard deviation [12].

2.3. CLASSIFICATION

Support vector machine (SVM) method was used for classification in this work. SVM is a supervised machine learning method. Groups which should be separated are not usually linearly separable in the original n-dimensional space of features. SVM transforms original n-dimensional space into m-dimensional space (m > n), in which groups can be separated. Kernel transformation is used

for this transfer. We used polynomial kernel transformation of second order in this work. The aim is to find group separator, when groups distance from the separator is largest possible. SVM use to search for separator support vectors, which are the points near the separator.

The basic version of the SVM is applicable for the separation of only two groups. There are several modified methods for applying SVM to separate multiple groups. Method one vs. one is used in this work. The classifier is gradually learned to distinguish each of two groups separately. We need therefore in total (K×(K-1)/2) binary learning, where K is the number of separated groups (in this case 5 and therefore total 10). [13]

3. RESULTS AND DISCUSSION

This chapter describes the achieved accuracy of five-category classification (supraventricular ectopic beat (S), ventricular ectopic beat (V), fusion beat (F), unknown beat (Q) and any heartbeat not in the S, V, F or Q classes (N)). In this chapter, it is also compared the importance of using different types of features.

Training and testing of the classifier is realized by ten-fold cross-validation. All QRS complexes of all signals are randomly divided into ten equally sized subgroups. Nine subgroups are used for training and one for testing the classifier. Gradually, each of the subgroups is selected as the test (there will be ten testing in total).

Classification accuracy is described by variables: accuracy (Acc), sensitivity (Se), specificity (Sp), positive predictive values (PV+) and negative predictive values (PV−). These variables are calculated as follows:

$$Acc = (TP_N + TP_S + TP_V + TP_F + TP_Q) / Y, \quad (1)$$

$$Se_i = TP_i / (TP_i + FN_i), \quad (2)$$

$$Sp_i = TN_i / (TN_i + FP_i), \quad (3)$$

$$PV+_i = TP_i / (TP_i + FP_i), \quad (4)$$

$$PV-_i = TN_i / (TN_i + FN_i), \quad (5)$$

where TP (true positive) is the number of correctly classified complexes into the analyzed group, TN (true negative) is the number of correctly classified complexes into other groups than the analyzed group, FP (false positive) is the number of incorrectly classified complexes into the analyzed group, FN is the number of incorrectly classified complexes into other groups than the analyzed group, Y is the number of all classified QRS. Subscript indicates the analyzed group (N, S, V, F, Q or i), where i stands for any group. Se, Sp, PV+ and PV− are calculated for each group separately.

Achieved accuracy is 0.9917 for the proposed classifier using all the features. Sensitivity, specificity, positive and negative predictive value of the proposed classifier for each group are shown in the Table 1. Hardest classificable groups are S and F, where the sensitivity is low. This means that large number of these complexes was classified into the wrong group. This may be due to small number of the complexes of S and F type compared with other types, so classifier was not learned sufficiently for these types. Confusion matrix is in Table 2. Above conclusions are confirmed. Confusion matrix shows that it is significantly more N complexes compared to other types of complexes. On the other hand, F complexes are very little. Therefore, classification of N complexes has a greater influence on the resulting classification accuracy than other complexes. On the other hand, classification of F complexes has little effect on the calculated accuracy of the classifier.

	N	S	V	F	Q
Se	0.9981	0.8641	0.9784	0.7968	0.9951
Sp	0.9695	0.9990	0.9985	0.9996	0.9998
PV+	0.9935	0.9592	0.9801	0.9342	0.9969
PV−	0.9911	0.9965	0.9984	0.9985	0.9996

Table 1: Sensitivity (Se), specificity (Sp), positive predictive values (PV+) and negative predictive values (PV−) of classifier for each type of QRS complex

		Output groups				
		N	S	V	F	Q
Correct groups	N	90342	83	60	9	16
	S	352	2397	24	1	0
	V	106	16	7537	35	9
	F	101	3	59	639	0
	Q	29	0	10	0	7992

Table 2: Confusion matrix

Using just certain groups, classification accuracy is shown in Table 3. Group All means all the features; KW is a group in which are used features selected by Kruskal-Wallis test from all features; TM are morphological features derived from the time domain; RR are features including RR intervals around QRS; Fr are features derived from the frequency domain; TF are features derived from the time-frequency domain and TS are features derived from the time-scale domain. Using ALL or KW is the most successful. Morphological features derived from the time domain are the most successful from separate groups. Conversely, classification is inappropriate by features derived only from the frequency domain or by RR intervals.

	All	KW	TM	RR	Fr	TF	TS
Acc	0.9917	0.9881	0.9770	0.7411	0.7243	0.9599	0.9179
Se	0.9265	0.8989	0.8496	0.3338	0.4239	0.7538	0.6618
Sp	0.9933	0.9894	0.9803	0.8330	0.8400	0.9613	0.9377
PV+	0.9728	0.9719	0.9266	0.3184	0.4461	0.9206	0.7344
PV−	0.9968	0.9960	0.9899	0.8324	0.8335	0.9861	0.9516

Table 3: Success of classifier using only selected group of features

The results of this study can be compared with previous studies. This comparison is shown in Table 4. Accuracy of the classifier described in this paper is approximately in the middle between studies mentioned here. The benefit of this study is comparing the above-mentioned types of features.

Classification method	Features	Number of target groups	Number of records	Accuracy [%]	Authors
SVM	RR, TM	5	33	98.8	Zidelmal et al. [1]
SVM	Samples	4	48	99.64	Thanapatay et al. [2]
ANN	Samples	6	32	94.26	Das et al. [3]
NN	Samples	4	44	99.61	Chang et al. [6]
NN	TM, Fr, HOS, TS	5	48	93.59	Kutlu et al. [5]
SVM	TM, RR, Fr, TF, TS	5	48	99.17	This article

Table 4: Comparison of the results of previous studies with the results obtained in this work

4. CONCLUSION

This article describes the classifier of QRS complexes in the ECG signal. We described all steps of the classification process. We used SVM method for classification and we used features derived from the time domain, frequency domain, time-frequency domain and time-scaling domain. Classification is realized into five types according to the AAMI standard. We tested the classifier on the entire MIT-BIH Arrhythmia database. The benefit of this study is comparing the types of features. Types of features were compared according to their suitability (it means classification accuracy using these features). Groups TM (Acc = 97.70 %) and TS (Acc = 95.99 %) were successful. Overall, the most successful was a group of all features (Acc = 99.17 %).

REFERENCES

[1] ZIDELMAL, Z, AMIROU, A, OULD-ABDESLAM, D and MERCKLE, J. ECG beat classification using a cost sensitive classifier. *Computer Methods and Programs in Biomedicine*. 2013. Vol. 111, no. 3, p. 570–577.

[2] THANAPATAY, Dusit, SUWANSAROJ, Chaiwat and THANAWATTANO, Chusak. ECG beat classification method for ECG printout with Principle Components Analysis and Support Vector Machines. *ICEIE 2010 - 2010 International Conference on Electronics and Information Engineering, Proceedings*. 2010. Vol. 1, no. Iceie, p. 72–75.

[3] DAS, Manab Kumar, MEMBER, Student and ARI, Samit. Electrocardiogram (ECG) Signal Classification using S-transform , Genetic Algorithm and Neural Network. *1st International Conference on Condition Assessment Techniques in Electrical Systems*. 2013.

[4] JIANG, Wei and KONG, Seong G. Block-based neural networks for personalized ECG signal classification. *IEEE Transactions on Neural Networks*. 2007. Vol. 18, no. 6, p. 1750–1761.

[5] KUTLU, Yakup and KUNTALP, Damla. A multi-stage automatic arrhythmia recognition and classification system. *Computers in Biology and Medicine*. 2011. Vol. 41, no. 1, p. 37–45.

[6] CHANG, Kuang-Chiung, WEN, Cheng, YEH, Ming-Feng and LEE, Ren-Guey. Comparison of Similarity Measures for Clustering of Qrs Complexes. *Biomedical Engineering: Applications, Basis and Communications*. 2005. Vol. 17, no. 06, p. 324–331.

[7] PARK, Juyoung and KANG, Kyungtae. PcHD: Personalized classification of heartbeat types using a decision tree. *Computers in Biology and Medicine*. 2014. Vol. 54, p. 79–88.

[8] SAMAD, Saleha, KHAN, Shoab a., HAQ, Anam and RIAZ, Amna. Classification of Arrhythmia. *International Journal of Electrical Energy* [online]. 2014. Vol. 2, no. 1, p. 57–61.

[9] MOODY, G. B. and MARK, R. G. The impact of the MIT-BIH arrhythmia database. *IEEE Engineering in Medicine and Biology Magazine*. 2001. Vol. 20, no. 3, p. 45–50.

[10] GOLDBERGER, Ary L, AMARAL, Luis A N, GLASS, Leon, HAUSDORFF, Jeffrey M, IVANOV, Plamen Ch, MARK, Roger G, MIETUS, Joseph E, MOODY, George B, PENG, Chung-kang and STANLEY, H Eugene. Current Perspective. 2000.

[11] CHAZAL, Philip De Chazal Philip De, O'DWYER, M. and REILLY, R.B. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE Transactions on Biomedical Engineering*. 2004. Vol. 51, no. 7, p. 1196–1206.

[12] TemplateSVM. *Mathworks* [online]. c1994-2016 [cit. 2016-03-06]. Dostupné z: http://www.mathworks.com/help/stats/templatesvm.html

[13] Fitcecoc. *MathWorks* [online]. c1994-2016 [cit. 2016-03-06]. Dostupné z: http://www.mathworks.com/help/stats/fitcecoc.html?s_tid=srchtitle

EVALUATION OF DISTORTION OF ECG SIGNALS AFTER COMPRESSION

Andrea Němcová

Doctoral Degree Programme (2), FEEC BUT

E-mail: xnemco01@stud.feec.vutbr.cz

Supervised by: Martin Vítek

E-mail: vitek@feec.vutbr.cz

Abstract: Evaluation of compression quality is essential part of data compression. No unified approach exists in this area and comparison of different algorithms is very difficult. The paper deals with 13 common efficiency and quality evaluation methods. Two compression algorithms are tested – improved single cycle fractal based algorithm and wavelet based SPIHT algorithm. All the quality evaluation methods have the same trend for both compression algorithms. For proper evaluation of compressed ECG signal quality, it is recommended to combine various evaluation algorithms.

Keywords: ECG, electrocardiogram, compression, distortion, quality evaluation, signal

1. INTRODUCTION

Quality evaluation is very important in compression of biosignals, especially electrocardiogram (ECG). The aim of ECG signal compression is reducing data size while preserving the diagnostic information. Some compression algorithms use quality evaluation as a control of compression. There are many different algorithms for compression evaluation. They can be divided in two big groups – compression effectiveness and compression quality algorithms. The compression quality algorithms are further divided in subjective (evaluation by cardiologist) and objective groups. Objective methods consist of those with diagnostic information and without diagnostic information. Objective methods can be global (one number for whole signal) and local (value for every beat). Local methods need QRS detector for segmentation of every beat.

In articles and conference papers, the use of evaluation methods is not unified. This fact makes the comparison between different algorithms difficult. Percentage root mean square difference (PRD) is the most common method, which belongs to the group of objective algorithms without diagnostic information. Moreover, some authors use PRD (without normalization) and some PRDN (with normalization). Normalization means subtraction of DC component before compression and quality evaluation or subtraction of DC component in PRDN equation. PRD is usually smaller than PRDN, which means, that the compressed signal (algorithm) evaluated by PRD seems better than the signal (algorithm) evaluated by PRDN. It can be misleading.

2. METHODS

2.1. COMPRESSION ALGORITHMS

In this work, two different algorithms for ECG signal compression are used – fractal based single cycle algorithm and wavelet based SPIHT (Set Partitioning in Hierarchical Trees) algorithm [1]. The fractal based single cycle algorithm is based on that, which was published on EEICT 2015 [2]. It was further improved by block division, Burrows-Wheeler transform and smoothing (for smooth blocks connection). Both algorithms are tested on 5 randomly selected signals from CSE database. We used first lead of orthogonal XYZ leads – W001, W036, W063, W099 and W122. The length of the signals is 10 s, sampling frequency 500 Hz and bit resolution 16 bits per sample.

2.2. EVALUATION ALGORITHMS

In this work, we used four different approaches for compression evaluation.

Evaluation of compression in terms of data size: This group involves especially compression ratio (CR), compression factor (CF) and average value length (avL) [3]. Each of these parameters can be calculated from another. Many authors use CR in their articles, but they often mean CF (according to the definition and equations in [3]). For this reason, using of avL is clear and it is used in this work (Eq. 1).

$$avL = \frac{size\ of\ the\ output\ stream}{original\ signal\ length}, [bps] \tag{1}$$

Global objective evaluation without diagnostic information: This group consists of normalized percentage root mean square difference (PRDN) [3], cross correlation coefficient (CC) [4], mean square error (MSE) [4], normalized mean square error (NMSE) [4], root mean square error (RMS) [3], [4], signal to noise ratio (SNR) [4] and standard error (Std_Err) [4].

Global wavelet based methods with diagnostic information: In this group belong wavelet-based weighted PRD (WWPRD) [5], wavelet-energy based weighted PRD (WEWPRD) [6], wavelet energy based diagnostic distortion (WEDD) [4] and multiscale entropy-based weighted PRD (MSEW-PRD) [7], which has two variants: relative mean wavelet subband energy (RMWSE) and relative wavelet subband energy (RWSE). These advanced methods are based on wavelet transform and weights. Distortion and weights are calculated for every scale. Final result (one number) is calculated as a weighted average of distortions in every scale. The algorithms differ in weights estimation.

Method based on delineation: This method needs delineation algorithm. In this work, we used delineation algorithm ECG SEEKER [8]. ECG signal is delineated; the output of the algorithm are positions of points of interest (QRS, QRSonset, QRSoffset, T, Toffset, P, Ponset and Poffset). Both original and compressed signals are delineated. Then the similarity is computed as shown in Eq. 2:

$$similarity = 100 - 100 \cdot \sum_{n=1}^{N} \frac{|PoI0(n) - PoI(n)|}{PoI0(n)}, [\%] \tag{2}$$

where *PoI* means points of interest, *N* is the length of every type of PoI vector (e.g. Ponset). When the compressed signal is distorted, the delineation algorithm can find less or more PoI of some type than in original signal (PoI0). Then the similarity should be adjusted; redundant PoI are deleted and penalization is used for this purpose as shown in Eq. 3.

$$similarity = similarity - (\frac{K}{N0} * 100), [\%] \tag{3}$$

where *K* is the difference between the length of PoI0 and PoI of one type (number of deleted PoI), *N0* is the length of PoI of one type in original signal.

Five signals were compressed by two above-mentioned algorithms and then evaluated by all mentioned evaluation algorithms. Then the mean value of all indexes were calculated and pictured.

3. RESULTS

Mean values of global objective indexes without diagnostic information are shown in Figure 1 and Figure 2. Figure 1 shows the trend of PRDN, CC, MSE, NMSE, RMS, SNR and STD_ERR in dependence of avL for improved single cycle fractal based algorithm. Figure 2 shows the same indexes for wavelet based SPIHT algorithm. From these figures is obvious, that the trend of all the indexes is the same for both methods. RMS is very similar to STD_ERR, that is why only one of these indexes can be used. CC and NMSE have very low sensitivity – their values change very little with avL (their trends are almost constant). The highest sensitivity has MSE, which decreases exponentially with

increasing avL (decreasing compression efficiency). From the figures follows, that both compression algorithms are comparable (the wavelet based SPIHT algorithm causes distortion comparable with fractal based method according to CC, NMSE and PRDN, a little bit lower distortion according to RMS, SNR and STD_ERR and lower distortion according to MSE).

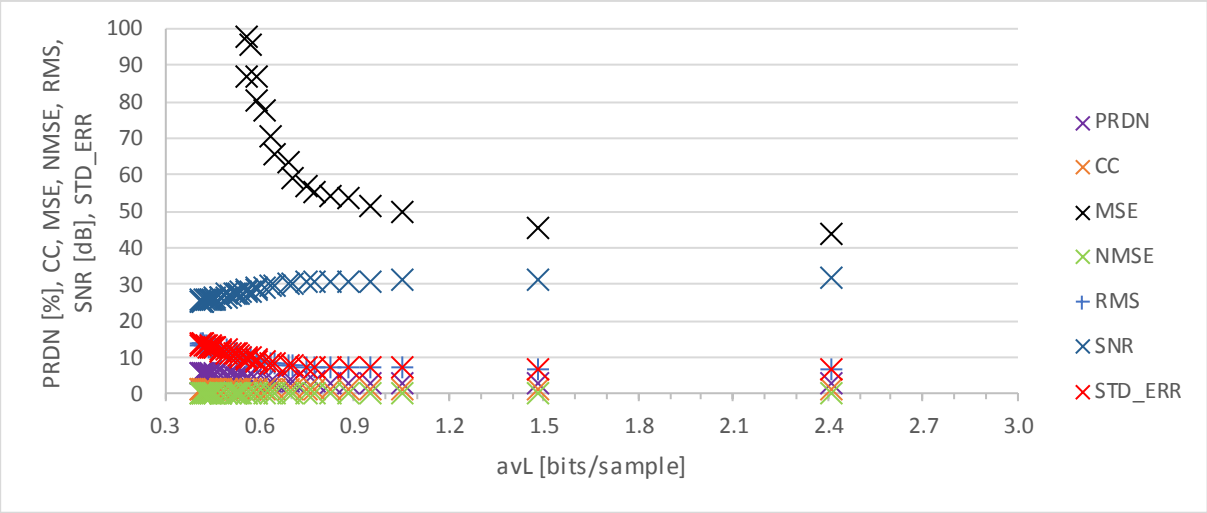


Figure 1: Trend of PRDN, CC, MSE, NMSE, RMS, SNR and STD_ERR in dependence of avL for improved single cycle fractal based algorithm.

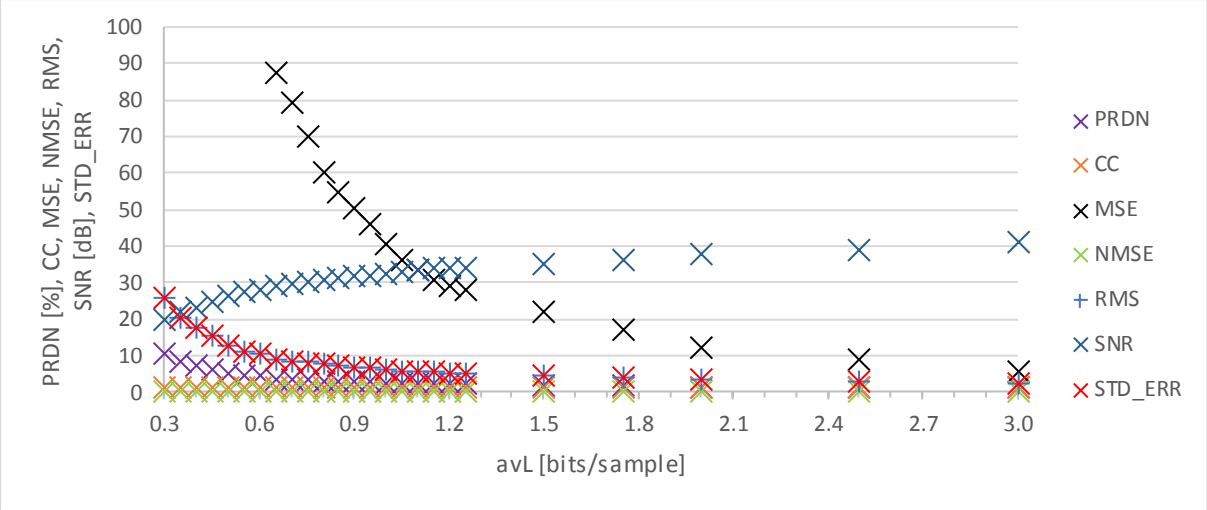


Figure 2: Trend of PRDN, CC, MSE, NMSE, RMS, SNR and STD_ERR in dependence of avL for wavelet based SPIHT algorithm.

Figures 3 and 4 show the trend of mean values of global wavelet based indexes with diagnostic information in dependence of avL for both compression algorithms. The trends of all the indexes are almost the same for both compression methods. All the indexes decrease exponentially with increasing avL. WEWPRD and WEDD have the same value for every avL, it means that one of these methods is redundant in this work. MSEWPRD_RMWSE and MSEWPRD_RWSE have lower sensitivity than other methods. From these figures (especially for avL from 0.3 to 0.6) is obvious, that the wavelet based SPIHT algorithm introduces lower distortion than fractal based algorithm.

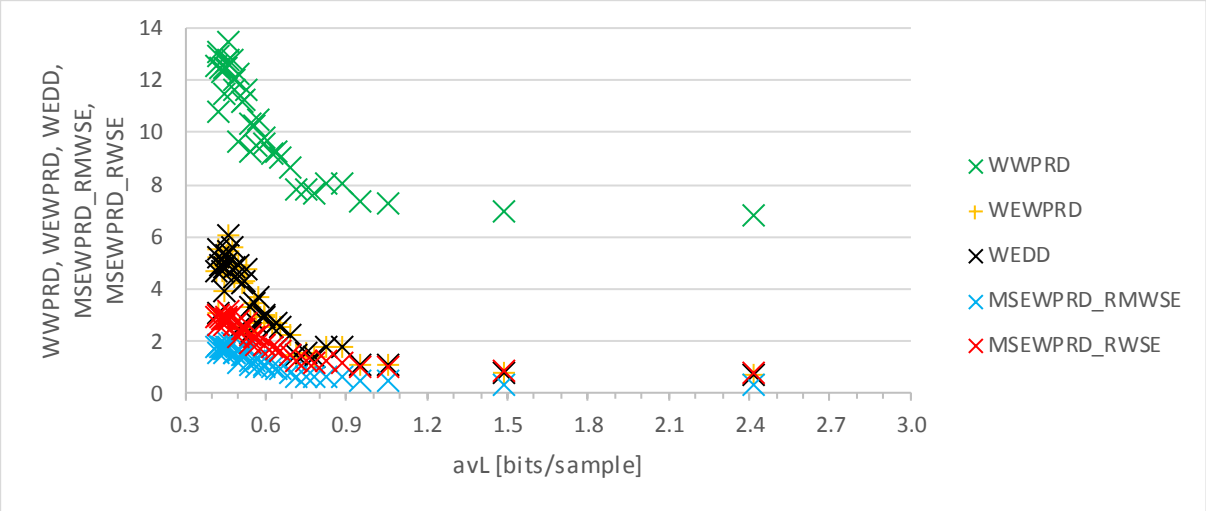


Figure 3: Trend of WWPRD, WEWPRD, WEDD, MSEWPRD_RMWSE and MSEWPRD_RWSE in dependence of avL for improved single cycle fractal based algorithm.

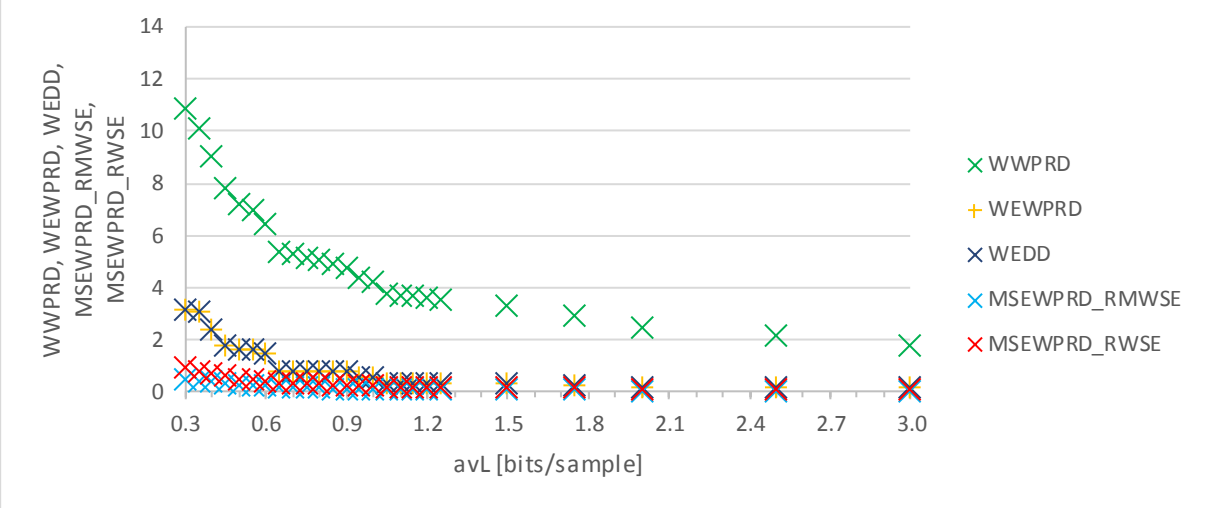


Figure 4: Trend of WWPRD, WEWPRD, WEDD, MSEWPRD_RMWSE and MSEWPRD_RWSE in dependence of avL for wavelet based SPIHT algorithm.

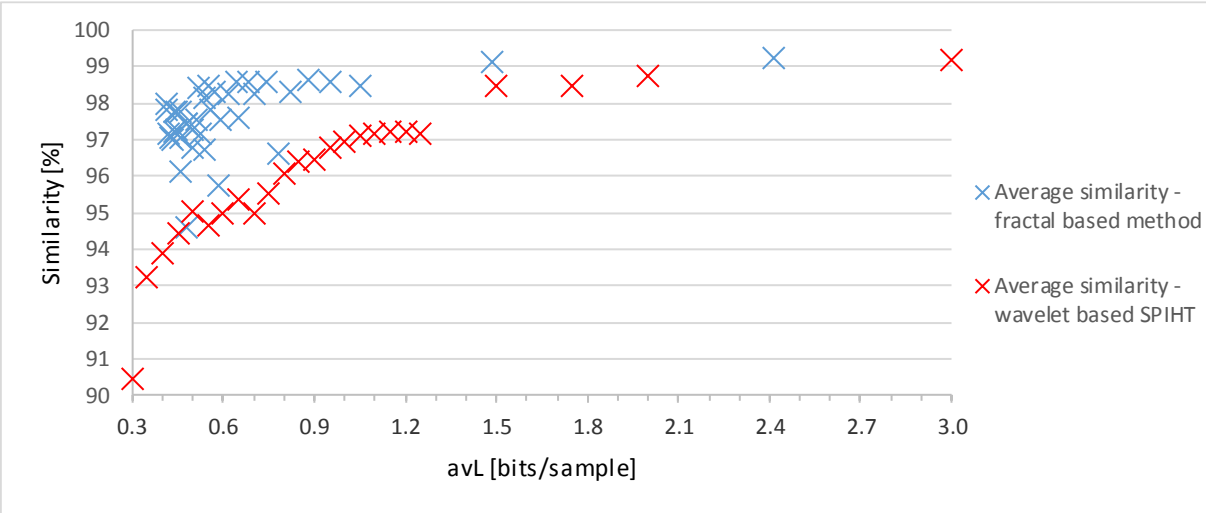


Figure 5: Trend of mean similarity of original and compressed signals in dependence of avL.

	Similarity of original and compressed signals [%]								
	QRS	QRSonset	QRSoffset	T	Toffset	P	Ponset	Poffset	Mean
Fractal	99.7535	98.8316	98.9608	99.5025	98.1467	94.5386	95.6116	95.9954	97.6676
SPIHT	99.9613	98.3999	98.4337	97.4318	94.0795	93.2525	94.5378	94.6389	96.3419

Table 1: Mean similarity of original and compressed signals.

The results of delineation of original and compressed signals are shown in Figure 5 and Table 1. In Figure 5 there are shown average similarity values (from 5 signals) for both compression methods. The trend is again almost the same for both compression methods. The similarity between original and compressed signal increases with avL. From Figure 5 follows, that signals compressed by fractal based method are more accurate and more similar to the original signals.

Conclusion followed from Figure 5 is confirmed by values in Table 1, where the average value for every type of PoI and method are shown. From the Table 1 is evident, that the highest similarity have QRS, QRSonset and QRSoffset. On the other hand, P, Ponset and Poffset have the lowest average similarity. It is probably caused by high magnitude of QRS complex and lower magnitude of P wave.

4. CONCLUSION

There are many types of algorithms for evaluation of compression effectiveness and quality. No unified approach for evaluation of compressed ECG quality exists. PRD or PRDN are the most commonly used indexes. As follows from results, only PRD (PRDN) is not enough for ECG signal quality evaluation. According to objective evaluation algorithms without diagnostic information, the wavelet based SPIHT algorithm is comparable with fractal based algorithm, according to wavelet based indexes with diagnostic information, the SPIHT algorithm is better than the fractal based one and according to similarity algorithm, the fractal based method is better than SPIHT. Using only one objective method, two compression algorithms cannot be compared correctly. It is good to combine a few evaluation methods of different type; visual inspection is useful as well. Another contribution of this paper is the fact, that all used indexes have similar trend for different compression methods.

REFERENCES

[1] VÍTEK, M., J. HRUBEŠ a J. KOZUMPLÍK: EKG KVANTUM; *Software pro filtraci, kompresi a rozměřování signálů EKG* [software]. 2009.

[2] NĚMCOVÁ, A. a M. VÍTEK. ECG signal compression based on fractals and RLE. In: *Proceedings of the 21st Conference STUDENT EEICT 2015* [online]. Brno. 2015.

[3] SALOMON, D. *Data Compression: The Complete Reference*. 4th ed. London: Springer, 2007, xxv, 1092 p. ISBN 978-1-84628-602-5.

[4] MANIKANDAN, M. S. a S. DANDAPAT. Wavelet energy based diagnostic distortion measure for ECG. *Biomedical Signal Processing and Control* [online]. 2007, **2**(2): 80-96. DOI: 10.1016/j.bspc.2007.05.001. ISSN 17468094.

[5] AL-FAHOUM, A.S. Quality Assessment of ECG Compression Techniques Using a Wavelet-Based Diagnostic Measure. *IEEE Transactions on Information Technology in Biomedicine* [online]. 2006, **10**(1): 182-191. DOI: 10.1109/TITB.2005.855554. ISSN 1089-7771.

[6] MANIKANDAN, M. S. a S. DANDAPAT. Effective quality-controlled SPIHT-based ECG coding strategy under noise environments. *Electronic Letters* [online]. 2008, **44**(20).

[7] MANIKANDAN, M. S. a S. DANDAPAT. Multiscale Entropy-Based Weighted Distortion Measure for ECG Coding. *IEEE Signal Processing Letters* [online]. 2008, **15**: 829-832. DOI: 10.1109/LSP.2008.2007620. ISSN 1070-9908.

[8] VÍTEK, M. a J. KOZUMPLÍK.: ECG SEEKER; *Software pro rozměření signálů EKG* [software]. 2010.

EVALUATION OF METHODS FOR AR COEFFICIENTS ESTIMATION USING MONTE CARLO ANALYSIS

Eva Klejmova
Doctoral Degree Programme (2.), FEEC BUT
E-mail: xklejm00@stud.feec.vutbr.cz

Supervised by: Jitka Pomenkova
E-mail: pomenkaj@feec.vutbr.cz

Abstract: Aim of this paper is to give recommendation for work with methods used for estimation of coefficients of autoregressive process. We applied Monte Carlo simulations to investigate performance of Burg, Yule-Walker and covariance methods. Evaluation of precision of spectral estimation is done with focus on signal length and lag order. The results are presented in graphical form and briefly discussed. Taking these results into account, Yule-Walker method shows better performance in case of long length signals and in case of overvalued lag order. Burg and covariance methods provide better results in case of short length signal and undervalued lag order.

Keywords: Autoregressive process, AIC, Burg method, Yule-Walker method, covariance method

1. INTRODUCTION

Autoregressive (AR) process as a method of time-frequency analysis is applied in many fields of science. It is a useful instrument in natural sciences, engineering, biology or medicine, social and economic [1], [2], [3] sciences. To obtain the best possible spectral representation it is crucial to select optimal method for estimation of AR coefficients with regard to the nature of input time signal. Therefore on the basis of simulation, we analyze the behavior of Yule-Walker, Burg and covariance method via Monte Carlo simulations. In case of AR process another key factor is the choice of optimal lag order on which the accuracy strongly depends. Therefore, it is good to investigate various optimization criteria for its optimal selection [4]. Advantages and disadvantages of these estimation methods and selection of appropriate lag order is examined and recommendation for its usage is formulated.

2. METHODICAL BACKGROUND

2.1. AUTOREGRESSIVE (AR) PROCESS

Autoregressive process as a method for time-frequency analysis uses parametric approach and creates a model generating an input signal. Application of this approach for estimation of spectral coefficients is possible based on observation that any linear process can be approximated by AR process of adequate order [4]. To compute spectrum, analyzed signal $s(n)$ is regarded as the output of a linear filter influenced by white noise w with variance σ_w^2. The output spectrum is then described as [5]

$$S_s(f) = \left|H\left(e^{j2\pi fT}\right)\right|^2 \sigma_w^2, \tag{1}$$

where $H\left(e^{j2\pi fT}\right)$ is a linear time variant filter.

The estimation of time-frequency model is then done according to formula [4]:

$$\widehat{S}_s(f) = \frac{\widehat{\sigma}_w^2}{\left|1 + \sum_{k=1}^{p} \widehat{a_p}(k) e^{-j2\pi fk}\right|^2}, \tag{2}$$

where $\widehat{a_p}(k)$ are estimates of the AR parameters, p is the lag order and σ_w^2 is white noise variance.

In order to construct spectral representation of the signal it is necessary to calculate AR model parameters. There are many ways to estimate the parameters. One of the method providing good results is ordinary least squares estimation (OLS). Unfortunately this method leads to multiple transpositions and multiplying of large matrices making OLS very complex for computation, especially for large value of lag order. This can be prevented by using Burg method, Yule-Walker method and covariance method [4]. Yule-Walker method is similar to OLS but estimates the autocorrelation from the data, which is then used to solve AR model parameters [4]. Burg method is based on least-squares lattice method and uses minimization of the forward and backward errors in linear predictors [4]. Covariance method is similar to Burg method and minimizes the forward prediction error in the least squares sense [6].

For accurate AR modeling the choice of adequate lag order p is essential. Selecting non optimal lag order leads to false spectral representation. In case of lag order distinctly smaller than optimal, the error manifests itself as excessive smoothing of the spectrum. Furthermore high level of lag order can significantly highlight non-significant spectral coefficients. Several information criteria can be used to ensure optimal selection. One of them, most commonly used and not computationally demanding is Akaike information criterion (AIC) [4]:

$$AIC(p) = ln\widehat{\sigma}_{wp}^2 + \frac{2p}{N}, \tag{3}$$

where σ_{wp}^2 is estimated variance of linear prediction error, p is the order and N is length of the signal. The order is selected as optimal when AIC reaches minimum.

2.2. Monte Carlo Analysis

The term Monte Carlo includes group of computer based statistical methods that can be applied to different scientific areas requiring statistical testing (eg. statistical physic, molecular simulation,...). The term itself comes from name of the city Monte Carlo, known for number of casinos that represent symbol of random number generator.

This suggest that the Monte Carlo experiment tries to replicate original data by generating large number of random realization of a stochastic process. Resulting simulated process is approximated by averaging these sub realizations in some way.

Monte Carlo methods usually contain following steps [7]:

1. Define required model
2. Generate random inputs
3. Perform a deterministic computation on the inputs
4. Analyze the result (using summary statistic)

Methods used to obtain AR model parameters contain white noise and therefor show different statistical properties when large variety of random processes is available. In this case Monte Carlo analysis is suitable tool to assess performance and quality of applied methods.

3. APPLICATION

3.1. SIMULATION

To evaluate performances of selected methods autoregressive process of order 20 was created. Selection of the order was done with regard to desired complexity of signal. The requirements were several spectral peaks with diverse amplitude and diverse frequency spacing. Therefore we chose level of 20 which provides this sufficient complexity with reasonable computational requirements. AR coefficients of our choice are shown in Table 1.

Table 1: AR coefficients

a_1	1	a_2	-1.04	a_3	0.78	a_4	-0.88	a_5	0.45
a_6	0.35	a_7	0.34	a_8	-0.04	a_9	-0.40	a_{10}	-0.10
a_{11}	0.04	a_{12}	0.36	a_{13}	0.38	a_{14}	-0.32	a_{15}	-0.35
a_{16}	-0.11	a_{17}	-0.08	a_{18}	0.26	a_{19}	0.15	a_{20}	-0.09
a_{21}	-0.31								

In next step w created signal in time domain using these coefficients as input for equation 1. Using the same AR coefficients, signals of two length were produced to assess the impact of signal lengths. Based on AR order of 20, we chose length of short signal to be 50 samples and length of long signal to be 500 samples. Then we took these two time signals and use them separately as an input for Monte Carlo analysis. Number of simulations was set to $N = 10, 100, 1000$. In each simulation covariance, Burg and Yule-Walker method was applied to obtain estimated coefficients. Using equation 2 we obtained corresponding spectral estimation. To asses performance of both methods in case of selecting non optimal lag order, the computation was done on selected range of order. The range was 1 to 30 in case of short signal and 1 to 300 in case of long signal. Mean square error (MSE) was calculated using theoretical and estimated spectrum representation for each process. This was done according following formula [4].

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(\widehat{Y}_i - Y_i\right)^2, \tag{4}$$

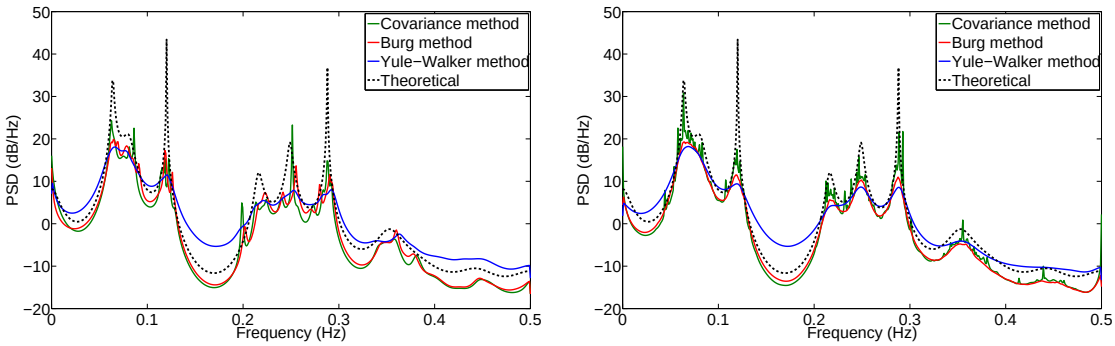
where $\widehat{Y}_i$ is value of theoretical spectral coefficient Y_i is its estimated counterpart. The lower the resulting number is the more accurate is the estimation.

3.2. RESULTS

Table 2 shows values of MSE in dB for selected methods. Lag order was set to optimum (i.e. 20, order of original signal). The results suggest that with increasing number of simulations precision of spectral estimation also increases. In general, results of all methods are significantly better for the signal of longer length. The MSE is smaller for Yule-Walker method in all considered cases. Covariance and Burg method show similar values in case of long signal. For short signal Burg method provides better results.

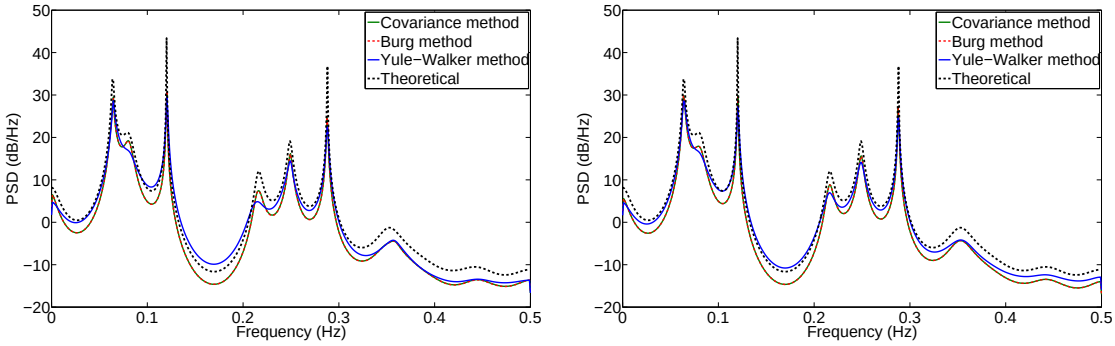
Table 2: Values of MSE in dB

	10 simulations		100 simulations		1000 simulations	
signal length	short	long	short	long	short	long
Yule-Walker method	27.35	7.60	23.79	5.91	22.95	5.77
Burg method	45.89	13.76	42.41	12.03	39.32	11.73
Covariance method	50.20	13.54	48.68	11.91	45.15	11.65



(a) 10 simulations **(b)** 1000 simulations

Figure 1: Theoretical and estimated spectrum of short signal



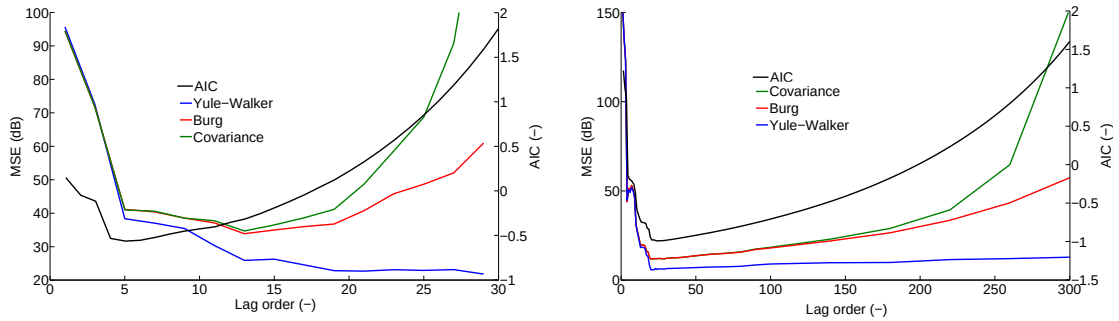
(a) 10 simulations **(b)** 1000 simulations

Figure 2: Theoretical and estimated spectrum of long signal

Figure 1 (a) and (b) shows theoretical and estimated spectrum of short signal for 10 and 1000 simulations. Lag order was set to optimal value of 20. The frequency of x-axis is relative to sampling frequency (value of 0.5 in figures corresponds to $0.5 f_s$). We can see that larger number of simulation leads to more precise estimation and smoothness of the graph. Neither of examined methods was able to adequately describe rapid changes of spectral peaks for measuring on 10 simulations. In case of 1000 simulations covariance method tends to be more accurate in case of describing amplitude of spectral peaks. However other two methods provide spectral line with greater smoothness. We can also see that the Burg method tends to be more accurate in capturing the shape of the graph than the Yule-Walker method.

In Figure 2 (a) and (b) theoretical and estimated spectrum of long signal for 10 and 1000 simulations and optimal lag order is shown. Larger number of simulations also leads to more precise estimation and smoothness of the graph but in is not as significant as in case above. Covariance and Burg method show almost the same results and tend to be more precise in capturing the shape of the graph than Yule-Walker method. However all methods were able to capture all significant spectral peaks.

In Figure 3 (a) and (b) we can see influence of lag order on precision of spectrum estimation. For signal of length 1000, the minimum of AIC and of all methods (lowest error) corresponds to the order of original AR process. We can see that the trend decrease rapidly to optimal lag order, after that the error tends to increase. In case of the short signal, the minimum lays on lower value that the optimal order. The trend also decreases rapidly to the lag order of 5. In case of Yule-Walker method the decrease continues, however error of Burg and covariance method increase exponentially.



(a) short signal **(b)** long signal

Figure 3: MSE and AIC of estimated spectrum for 1000 simulations

4. CONCLUSSION

In this work, we offer a comparison of different methods for power spectrum estimation. The performance of used methods is evaluated based on precision of the spectral estimation. For purposes of the study, we used a simulated AR signal with the order of 20. The performance of used methods vary, depending on the properties of the signal, namely on its length. Yule-Walker method proved to be more efficient when applied on signal of longer duration. On the other hand, Burg and covariance method offers better peak detection for short signals. Furthermore, the results showed that the performance of used methods also strongly depends on used lag order. Yule-Walker method performs best in scenarios, when higher lag orders are allowed. Burg and covariance method is more suited in cases, when we aim for lower lag orders.

ACKNOWLEDGEMENT

The research published in this submission was financially supported by the Brno University of Technology Internal Grant Agency under project no. FEKT-S-14-2177 (PEKOS)

REFERENCES

[1] Ding, H., Chao, B.F.: Detecting harmonic signals in a noisy time-series: the z-domain Autoregressive (AR-z) spectrum. *Geophysical Journal International*, 2015, vol. 201, no.3, p. 1287-1296. DOI:10.1093/gji/ggv077

[2] Blumenstein, J., Pomenkova, J., Marsalek, R.: Comparative Study Of Time-Frequency Analysis Approaches With Application To Economic Indicators. In 26th European Conference on Modelling and Simulation ECMS 2012. Germany: 2012. pp. 291-297

[3] Sebesta, V., Marsalek, R., Pomenkova, J.: The Modified Empirical Mode Decomposition Method For Analysing The Cyclical Behavior Of Time Series. In 27th European Conference on Modelling and Simulation ECMS 2013. Norway: 2013. pp. 288-292

[4] Proakis, J. G., Rader, Ch. M, Ling, F. L., Nikias, Ch. L., Moonen, M. and Proudler, J. K.: Algorithms for Statistical Signal Processing. Prentice Hall, 2002.

[5] Ján, J.:Číslicová filtrace, analýza a restaurace signálu. VUTIUM Brno, 2002.

[6] Stoica, P., and Moses R.L.: Introduction to Spectral Analysis. Prentice-Hall, 1997.

[7] Robert, C. P.; Casella, G.: Monte Carlo Statistical Methods, 2nd ed. New York: Springer, 2004.

IMPLEMENTING GUNSHOT DETECTION ALGORITHM ON DIGITAL SIGNAL PROCESSOR

Martin Hrabina

Doctoral Degree Programme (1), FEEC BUT

E-mail: hrabina@phd.feec.vutbr.cz

Supervised by: Milan Sigmund

E-mail: sigmund@feec.vutbr.cz

Abstract: Aim of this paper is implementation of developed gunshot detection algorithm on TMS320C6713 digital signal processor. Used algorithm is based on 3 LPC coefficients, energy in 3 spectral bands and correlation against template. Analyzed acoustical signal is segmented into 23 ms frames. If a segment contains gunshot, it is signalized by flashing LED diode. This method results in 82% correct gunshot detection and approximately 3% false alarms averaged in various non-gunshot sound categories.

Keywords: signal processing, gunshot detection, digital signal processor

1. INTRODUCTION

The aim of our effort is development of reliable gunshot detection algorithm suitable for digital signal processor which would be used as an addition to already existing GPS collars for elephants. Current collar periodically informs national park guard about elephant herd position. The idea is, to extend its possibilities to gunshot detection, which could alarm guard when poaching happens.

This application requires reliable detection algorithm, with emphasis on lower false alarm rate rather than higher detection rate. Additionally, low power consumption is required because of impossibility of frequent battery changes on collars. In contrast with similar existing algorithms, this application does not need to be able to classify different sounds, only to distinguish gunshot sounds from other possible sounds, such as breaking branches, splashing or sounds produced by animals. It is not crucial to calculate exact location of gunshot origin, as in some military applications, since GPS collars provide sufficient accuracy.

The main objective of this paper is implementation of previously developed gunshot detection algorithm [1] on digital signal processor. The algorithm was slightly modified with enlarging sound database. Whole algorithm is implemented using C programming.

There are many works dedicated to gunshot detection and several systems designed to detect gunshots together with exact shooter location, these are used mainly for military purposes. Among important publications on gunshot detection are [2] comparing different algorithms and [3] dedicated to feature selection for gunshot detection.

2. DETECTION ALGORITHM

Presented detection algorithm, in contrast with [1], is working in real time with normalized audio samples. Audio is processed in overlapping frames of length 1024 samples (~ 23ms at 44.1 kHz sampling frequency), with overlapping of 1/4 which corresponds to 256 samples. Frames are progressively tested for 3 sets of features, linear predictor coding coefficients (LPC), energy in 3 spectral bands and correlation against template. Each method was tested with several audio files in order to determine its computational demands and number of false alarms and correct recognitions. Obtained results are summarized in Table 1. In this table, correct event recognition is number of

events correctly flagged (gunshot or non-gunshot event) and number of non-gunshots flagged as gunshots in False alarms column. Clock cycles are measured using Profile Clock in Code Composer Studio. Less computationally demanding parameters are calculated first and if a frame does not pass the test, it is not processed further. Whole processing is depicted in Fig. 1.

Method	**Correct event recognition**	**False Alarms**	**Clock Cycles**
LPC	27	90	271k
Energy	61	54	2.7M
Correlation	58	59	4.0M

Table 1: Computational demands and false alarms

As can be seen in Table 1, each method separately gives many false alarms, with LPC coefficients in worst position. However LPC coefficients, due to their low computational demands, can still be used as event detector with unfit frames discarded with minimal effort. Also, as is shown in [1], they are needed in order to further reduce false alarms generated by correlation and energy criteria.

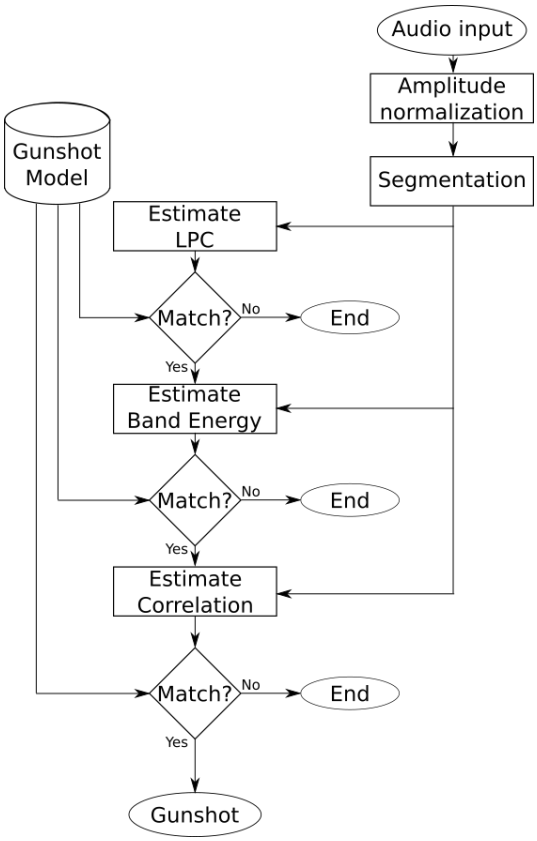


Figure 1: Block diagram of signal processing

3. PROCESSOR TMS320C6713

Digital signal processor was chosen in order to satisfy autonomous, real-time sound analysis and gunshot detection, in contrast with the method used in [1] where whole recordings were processed at once using Matlab. For testing purposes a DSK with floating point processor TMS320C6713 was chosen, with input data flow 705.6 kbits/s (corresponding to 44.1 kHz sampling frequency, 16 bits quantization and single channel). This processor uses DSP/BIOS real-time kernel with preemptive multitasking, which is used in switching between different tasks, such as filling buffer and computing features. Semaphores take care of atomic operations, with key functions SEM_pend (SEM_handle, Uns timeout) and SEM_post (SEM_handle). More information on DSK can be found in datasheet [4] and documentation.

Input audio sampling is granted by stereophonic TVL320AIC23 codec in combination with Multichannel Buffered Port (McBSP) and Enhanced Direct Memory Access (EDMA). McBSP is a communication channel between codec and processor, it is used during configuration and sample transport, with each of those roles having dedicated channel. Since codec is stereophonic it samples both left and right channel, resulting in 2 samples per sampling period, however algorithm works with only monophonic signal. Codec could also be configured to sample just one channel, but as will be seen later, input buffer with double length will be useful, in this case, we obtain length of 2048 samples, with signal samples interleaved by zeros. After filling first buffer, semaphore indicates start of features calculation while second buffer is being filled. Overlapping is implemented by preserving the last 1/4 of the frame and moving it to the start of new frame.

Two LED diodes are used for activity signalization. The first is signaling presence of any sound, with LED turning on if input sample exceeds 0.1 threshold, the second detected gunshot. If gunshot is present in analyzed frame, LED will turn on for approximately 1 second. This indicates alarm that will be sent to park guard.

4. FEATURES CALCULATION

This section deals with how the three feature types are calculated. The necessity of low power consumption and real-time signal processing led to preference of existing optimized functions for features calculation. The three features has each its own function which evaluates whether criteria has been met. If all criteria are met, frame is flagged as containing gunshot. If at least one criterion is not met, frame is dropped as not containing gunshot without proceeding to calculate remaining criteria. Operations are computed in floating-point with input data being converted to floating point using division by 32768.0.

4.1. LPC COEFFICIENTS

The algorithm calculates 8th order of LPC coefficients, but for decision uses only 3rd, 4th and 6th. Coefficients are calculated using Levinson-Durbin algorithm. This algorithm uses the Toeplitz autocorrelation matrix with a maximum shift equal to the order of coefficients calculated. Autocorrelation is calculated using digital signal processor library (Dsplib) for TMS320C67x. Levinson-Durbin algorithm uses iterative process with number of iterations equal to coefficients order, in this case eight. This algorithm works with real domain numbers, so it needs input matrix without interleaving zeros, i.e. length 1024. Output of whole function is one or zero, depending on whether criteria are met or not. Table 2 shows code fragment for computation of LPC coefficients. Input variable *out[]* is autocorrelation of input signal, its name indicates it is output from previous function. Variable *a[][]* is a matrix for recursive calculation of LPC coefficients, variables *E[]* and *k[]* are error and correction auxiliary variables. Resulting LPC coefficients are stored in *a[n][8]*.

```
1. for( p = 1; p<9; p++){
2.    a[0][p-1] = 1
3.    temp = 0;
4.    for( s = 0; s < p; s++)
5.        temp += a[s][p-1]*out[p-s];
6.     k[p] = -temp/E[p-1];
7.     a[p][p] = k[p];
8.     for( s = 0; s<p; s++)
9.         a[s][p] = a[s][p-1] + k[p]*a[p-s][p-1];
10.   E[p] = (1-k[p]*k[p])*E[p-1]; }
```

Table 2: Code fragment for LPC calculation

4.2. ENERGY IN SPECTRAL BANDS

Spectral energy is calculated mostly by using functions from Dsplib. The Fast Fourier Transform (FFT) function supplied by library needs twiddle factors, which can be generated by another function from library and subsequently bit-reversed so they are in correct order. Twiddle factors are generated first time when whole algorithm starts and later are only looked up. FFT function works with complex input vector, complex numbers are represented as consecutive real and imaginary numbers, so input is of length 2*N*, i.e. 2048, with each even (imaginary) element being zero. After FFT, bit reversal and division by *N* is needed, both functions are supplied by Dsplib. Energy is calculated as the sum of squared complex modules in 5 spectral bands of 3 kHz each in range 0 Hz to 15 kHz. In our algorithm, only three bands in range 3 kHz to 12 kHz are used. This function again outputs one when all three bands comply with requirements, or zero if at least one does not.

4.3. CORRELATION

Correlation can be calculated both in the time domain and in the frequency domain. If calculation in time domain was chosen, there is a possibility of using slightly modified Dsplib function for convolution, which would require *Nm* multiplications, where *N* is signal vector length and *m* required shift, in this case equal to *N*. Computation in frequency domain requires calculation of complex conjugate, but it is enough to calculate complex conjugate of template signal and load it each time it is required, as opposed to calculating complex conjugate of input signal each frame. Since FFT of input signal is already calculated in previous step, this algorithm uses calculation in frequency domain, due to only *N* needed multiplications.

5. EXPERIMENTAL RESULTS

Achieved results are presented in Table 3 with information about output of each stage displayed in Fig. 1. True detections represent percentage of successfully detected gunshots out of all gunshots and false detections percentage of recordings flagged as gunshots out of all non-gunshot sounds. Table 4 shows distribution of false alarms in percentages in different non-gunshot categories for final stage of algorithm, i.e. output of whole device.

Sound database used consists of four different audio categories, gunshots, barking dogs, sounds with wood (e.g. cracking, chopping …) and ultimately other, such as splashing, elephant sounds or cattle bells. Each category contains more than 50 events. Sound database was recorded in open space shooting range and forest using commercial sound recording device.

	True Detection	False Detection
Stage 1 LPC	100%	92%
Stage 2 LPC + Energy	88%	21%
Stage 3 LPC + Energy + Correlation	82%	3%

Table 3: True and false detections at different stages

	Wood	Dogs	Other
False detections	5.5%	0%	4.3%

Table 4: False alarms rates

6. CONCLUSION

Algorithm presented in this paper achieved 82% successful gunshot detection and 3% overall false alarm rate. This specific application prefers minimal false alarm rate to more true detections because of multiple gunshots during hunting and possible adverse effects of frequent false alarms such as money wasted on checking false alarms and distrust of the system in the eyes of park guards. In future work, resistance to non-stationary noise used in speech recognition [5] will be tested. Further development should also focus on reduction of energy consumption due to difficult battery replacement in collars.

ACKNOWLEDGEMENT

Research described in this paper was financed by Czech Ministry of Education in frame of National Sustainability Program under grant LO1401. For research, infrastructure of the SIX Center was used. The research was also financially supported by the Brno University of Technology Internal Grant Agency under project no. FEKT-S-14-2177 (PEKOS).

REFERENCES

[1] Hrabina, M., Sigmund, M.: Acoustical Detection of Gunshots. In: Proceedings of 25th International Conference Radioelektronika, Pardubice, Czech Republic, 2015, pp. 150-153.

[2] Chacon-Rodriguez, A., Julian, P., Castro, L., Alvarado, P., Hernandez, N.: Evaluation of Gunshot Detection Algorithms. IEEE Transactions on Circuits and Systems I: Regular Papers, 2011, vol. 58, no. 2, pp. 363-373. DOI: 10.1109/TCSI.2010.2072052

[3] Kiktova-Vozarikova, E., Juhar, J., Cizmar, A.: Feature selection for acoustic events detection. Multimedia Tools and Applications, vol. 74, no. 12, pp. 4213-4233. DOI: 10.1007/s11042-013-1529-2

[4] Datasheet for TMS320C6713 [online].
Available: http://www.ti.com/lit/ds/symlink/tms320c6713.pdf

[5] Zelinka, P., Sigmund, M.: Hierarchical classification tree modeling of nonstationary noise for robust speech recognition. Information Technology and Control, 2010, vol. 39, pp. 202-210.

INSTRUCTION MAPPING PROCESS ON THE VLIW ARCHITECTURES

Roman Mego

Doctoral Degree Programme (4), FEEC BUT

E-mail: roman.mego@phd.feec.vutbr.cz

Supervised by: Tomas Fryza

E-mail: fryza@feec.vutbr.cz

Abstract: This paper deals with the process of instruction mapping on the digital signal processors. This process is used by the newly developed tool, which is designed for generating low-level assembly code for very long instruction word processors. The tool is suitable for creating cores of the signal processing algorithms.

Keywords: instruction mapping, low-level, digital signal processing, very long instruction word

1. INTRODUCTION

The advanced real time signal processing applications are realized mostly using the digital signal processors (DSP) instead of the analogue circuits. The complexity of the digital signal processing algorithms is still increasing, which puts higher requirements on the DSP resources. The main aspects influencing the performance of the final application are:

- the architecture suitability,
- the software optimization.

The DSPs from the architecture point of view are highly optimized for the frequent access to the memory, executing instruction in the loops and performing special operations typical for the DSP algorithms, such as multiply and accumulate (MAC) [1].

The software optimization is the second aspect influencing the final performance of the application. Several years ago, critical parts of the software was optimized by the hand in the low-level assembly language. This approach was chosen because the high-level compilers were not generating efficient code. This leads to the longer development time and therefore to higher cost.

Nowadays the situation has changed. The modern compilers are able to optimize the final code which is practically not worth to spend more time on hand optimization. This especially applies for the scalar processors. There are also special architectures, where the results are less effective and there is still worth to do low-level parts of software. They are mainly using the instruction level parallelism. This also includes the very long instruction word (VLIW) architecture [2], where the creation of low-level code is more complex than on the traditional scalar architectures. There are also retargetable compiler frameworks, such as [3], which are more efficient than commercially available compilers, but creating desired model of the processor is not an easy task. For this reason, the tool for efficient instruction mapping is developed.

This paper will introduce the approach used by the mapping tool. This tool allows to implement DSP algorithms on VLIW processors in the efficient way. It uses the signal-flow graph description of the algorithm which is transformed low-level assembly code.

2. REFERENCE PLATFORM

The VLIW architecture, as it was mentioned above, takes the advantage from the instruction parallelism. This means, that single processor core contains multiple functional units able to execute operations in the same time [2]. The difference between VLIW and superscalar processor is the point of decision, which operations will be performed at the same time (figure 1). The superscalar processors contain the dispatch buffer, which dynamically allocates the functional units during the runtime [4]. The instructions on VLIW architectures are allocated statically during the software compilation. The compiler creates the instruction packets, which are directly sent into the functional units. This creates the space for the core simplification, optimization or extension for the new functionality.

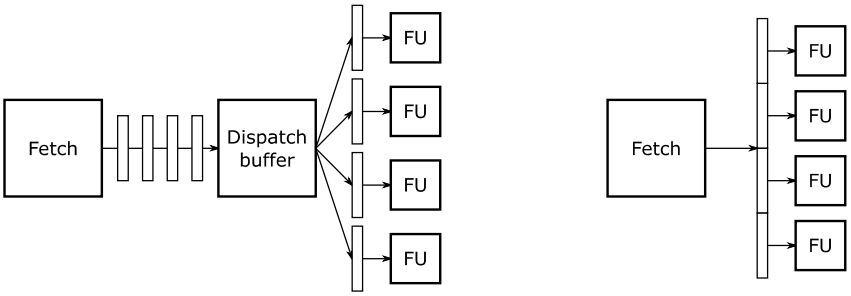


Figure 1: Superscalar (left) and VLIW (right) difference.

The reference processor for the tool design is the TMS320C6678 from the Texas Instruments [5]. This is the 8-core fixed/floating-point DSP. Each core consists of the two identical data paths, where each data path has four functional units and the set of thirty-two 32-bit registers (figure 2). The functional units are not identical in the same data path. Each of them was designed for different purpose. The processor was not chosen, because it is multi-core, but for the diversity in the core structure.

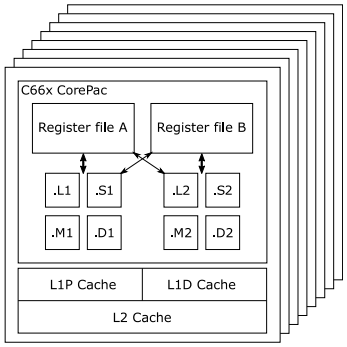


Figure 2: Simplified structure of the C66x core.

3. TOOL IDEA

The tool will be used for optimizing the signal processing operations, not the branch or function call control. The idea is based on the main weakness of the compiled code for this kind of processors. The high-level notation of algorithms is sequential and compiler tries to keep the same order of the instructions. The result is that the final code uses the limiting number of the functional units in the same time.

The tool uses the signal-float graph approach for the algorithm description. This description, as the analogue circuit, can contain multiple parallel signals and operations. The order of the processing will not be as it is in the source code text file, but with the relations of the signals and operations in the algorithm. This should be decided automatically by the tool without the user interaction.

4. INSTRUCTION MAPPING PROCESS

The whole process of the instruction mapping from the reading of the input files to the generating low-level assembly code can be divided into the following steps:

- reading and parsing input files
- finding the relations between the operations and signals
- sorting the operations and signals
- mapping the operations and signals to the functional units and registers
- generating output file with low-level assembly code

4.1. READING AND PARSING INPUT FILES

When the tool is executed, it needs to know the algorithm which will be mapped and also the architecture. The information is stored separately in two text files. The first keeps the information about the target processor, which consists of:

- the list of data paths,
- the list of functional units,
- the list of registers,
- the list of the instructions.

The second file contains the algorithm description, which is independent on the architecture. The main elements are:

- signals,
- operations.

Signals are represented with the data type and its functionality. The functionality divides the signals on three groups, input, output and internal signal. Operations are the mathematical or logical actions performed on signals. It can be recognized as the node in the signal graph with signals on its inputs and output.

4.2. FINDING RELATIONS BETWEEN THE SIGNALS AND OPERATIONS

When the algorithm notation is parsed, there is time to finding the relations. This step can be used to validate the connections. Signals practically represent the variables, which are used in the mathematical operations. The difference in comparison with the ordinary programming languages is that the value to the signal can be assigned only once. This is similar to the hardware description languages (HDL). All internal signals must have assigned its creator (operation) and it must be used as the input at least on one operation. The input and output signals do not need to respect all of these conditions. The example is shown on figure 3, which represents single butterfly of the Fast Fourier Transform (FFT).

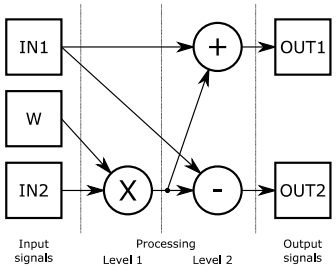


Figure 3: Single FFT butterfly represented by signal-flow diagram.

In this step we can also connect the operations in the algorithm with the suitable instructions supported by the processor. The instruction is chosen on the operation itself and the data type of the input and output signals. The tool now creates internally new structure, which references the architecture independent parsed data and extends it with other architecture dependent information.

4.3. SORTING THE OPERATIONS AND SIGNALS

Sorting is the first step of the mapping itself. It is preparation of the possible execution order, which will be assigned to the functional units in the next steps. This order is based on the relations between the items in the algorithm. The main condition is that the operation must be on higher level than the highest level from the creators of the all input signals. For example, if all of the operation input signals are input signals of the algorithm, this operation is placed on first level. If one signal is the input signal of the algorithm and second signal is created by the operation on second level, the operation is placed on third level. Multiple operations can be placed on the same level without any restriction given by the architecture (figure 3). There can be additional sorting on the individual levels for better results.

4.4. MAPPING OPERATIONS AND SIGNAL

The mapping of the operations and signals is realized with the two 2-dimensional maps (figure 4). First dimension represents time in instruction cycles in both maps. Second dimension represents hardware resources, functional units or registers. During the mapping, the order of the operations is reflected in the process. The operations are mapped as the first. The placing is similar to first-fit method from the first instruction cycle, so in final different levels can be executed simultaneously if there is a free space in the map. Operations are mapped only on one data path to avoid the usage of the cross paths.

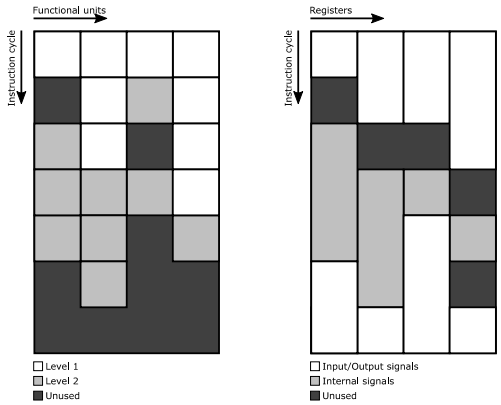


Figure 4: Example of functional unit and register map.

After the operations are mapped to the functional units, the signals must be allocated to the registers as well. In first step, the lifetime of the signal is determined from the position of operations in time. The lifetime of the signal starts at the end of the operation which creates it and ends when in it passed into the last operation. This means, that the registers are allocated for the signal only when necessary, not for the whole time of the algorithm. The input and output signals has its lifetime given by the start and end instruction cycle of the final algorithm. After that, the signals are mapped in the similar way like the operations.

4.5. GENERATING OUTPUT FILE

The last step it to generate output file with the low-level assembly file. All information is provided in the referenced structures form the maps. Each instruction in the architecture definition contains the template for human readable notation, where the functional unit and registers will be added. The tool goes through operation map from the first instruction cycle and joins available operations together into the instruction packets.

5. EXAMPLE OF THE GENERATED MAP

The example of the mapped algorithm is showed on figure 5. The first part is the source code of the 4-point Fast Fourier Transform without the signal definitions. The next two part are maps of the functional units and registers. More information about the mapped algorithms can be found in [6].

```
B1_RE = ADD(A1_RE, A2_RE)
B1_IM = ADD(A1_IM, A2_IM)
B2_RE = SUB(A1_RE, A2_RE)
B2_IM = SUB(A1_IM, A2_IM)
B3_RE = ADD(A3_RE, A4_RE)
B3_IM = ADD(A3_IM, A4_IM)
B4_RE = SUB(A3_RE, A4_RE)
B4_IM = SUB(A3_IM, A4_IM)

C1_RE = ADD(B1_RE, B3_RE)
C1_IM = ADD(B1_IM, B3_IM)
C2_RE = ADD(B2_RE, B4_IM)
C2_IM = SUB(B2_IM, B4_RE)
C3_RE = SUB(B1_RE, B3_RE)
C3_IM = SUB(B1_IM, B3_IM)
C4_RE = SUB(B2_RE, B4_IM)
C4_IM = ADD(B2_IM, B4_RE)
```

Cycle	.L	.S	.M	.D
1	ADD	SUB		ADD
2	ADD	ADD		SUB
3	SUB			SUB
4	ADD	SUB		ADD
5	SUB	SUB		ADD
6	ADD			SUB

Cycle	A0	A1	A2	A3	A4	A5	A6	A7	A8	A9
1	A1_RE	A1_IM	A2_RE	A2_IM	A3_RE	A3_IM	A4_RE	A4_IM		
2	B1_RE								B2_RE	B2_IM
3		B3_RE	B1_IM	B3_IM						
4							B4_RE	B4_IM		
5	C1_RE	C1_IM			C3_RE	C3_IM				
6			C2_RE	C2_IM			C4_RE	C4_IM		
7										

Figure 5: Source code of 4-point FFT and generated map for resource usage.

6. CONCLUSION

This paper presented the way of the instruction mapping using the signal-flow graph approach instead of sequential notation. It means that the desired algorithm is described only with the relations between signals and the tool determines the order of the execution. This approach was adjusted for the VLIW architectures. The tool can be also used for scalar processors, but probably without any benefits. The efficiency of the generated code for TMS320C6678 model is close to the hand-optimized code. The disadvantage of the tool is, that it can be used only for cores of the algorithms, because the limiting factor is the number of the resisters.

ACKNOWLEDGEMENT

The research published in this submission was financially supported by the Brno University of Technology Internal Grant Agency under project no. FEKT-S-14-2177 (PEKOS).

REFERENCES

[1] Smith, Steven W. The scientist and engineer's guide to digital signal processing. 1st ed. San Diego, Calif.: California Technical Pub., 1997. ISBN 0966017633.

[2] Philips Semiconductors. An Introduction to Very Long Instruction Word Computer Architecture. Publication no. 9397-750-01759.

[3] Rajagopalan, Subramanian et al. A retargetable VLIW compiler framework for DSPs with instruction-level parallelism. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. 2002. ISSN 02780070.

[4] Shen, John Paul. Modern processor design: fundamentals of superscalar processors. Long Grove: Waveland Press, 2013. ISBN 9781478607830.

[5] Texas Instruments Incorporated. TMS320C6678 Multicore Fixed and Floating-Point Digital Signal Processor. [Online] 2014. [Cited: 14 March 2016]. <http://www.ti.com/lit/gpn/tms320c6678>.

[6] Mego, Roman and Fryza, Tomas. Efficiency of the signal processing algorithms using signal-flow based mapping tool. International Conference Radioelektronika. 2015. ISBN 978-1-4799-8117-5.

LET THE CROWD BE HEARD - A CROWDSOURCING-BASED STUDY ON VIDEO CODING

Ondrej Zach
Doctoral Degree Programme (3), FEEC BUT
E-mail: ondrej.zach@phd.feec.vutbr.cz

Supervised by: Martin Slanina
E-mail: slaninam@feec.vutbr.cz

Abstract: In this paper, we present a crowdsourcing-based study on video codec comparison. We compared the perceived quality of sequences encoded using H.264 and H.265 algorithms. The main advantage of the study is, that the users could watch the videos in their natural environment using their own equipment. The results show, that H.265 may not bring a significantly better quality in SD scenarios, however, in other scenarios, the use of H.265 might be preferred.

Keywords: QoE, video codec, AVC, HEVC, H.264, H.265

1 INTRODUCTION

Video streaming and other video-based services represent majority of the overall internet traffic. According to a Cisco study [1], video services take the share of more than 67% of the overall traffic and this share is expected to grow up to 80% in 2019. The behavior of the users also tends to higher quality of the video. This has also a significant impact on the amount of the data transferred. To offer higher quality of the video while preserving the same data rate, a change in video compression scheme is necessary.

Currently, the majority of video content on the Internet is encoded using H.264/AVC (Advanced Video Coding). AVC performs best in HD (720p) and Full HD (1080p) scenarios with required bit rates in the range from 2 Mbps to 5 Mbps. However, AVC might be insufficient for streaming higher quality video content (e.g. UHD). Therefore, a successor to AVC was developed in order to fulfill this demand. High Efficiency Video Coding (HEVC), which is again a result of join collaboration of ISO Motion Picture Experts Group (MPEG) and ITU-T Video Coding Experts Group (VCEG), standardized jointly as ITU-T H.265 [2] and ISO/IEC 23008-2 [3] is often stated to need just 50% of the bitrate of AVC while preserving the same visual quality [4].

The motivation of this paper is to offer a comparison of AVC and HEVC video coding standard using crowdsourcing. Crowdsourcing platform enables us to gather the ratings from the users watching the videos in their home environment using their own equipment. The rest of the paper is divided as follows: Sec. II offers a brief description of related work, Section III. describes the test setup. In Section IV., a discussion of gathered results is offered and the last section concludes.

2 RELATED WORK

AVC is used in the majority of video applications nowadays and serves as a benchmark for video coding of this time. However, AVC may not be able to fulfill the upcoming challenges (e.g. UHD streaming, higher quality with limited bandwidth when mobile, ...). It is predicted, that 21% of the global Video on Demand (VoD) traffic will be in UHD in 2019, [1], which can cause H.264 to have insufficient performance. One of the solution, how to satisfy customers' needs might be HEVC.

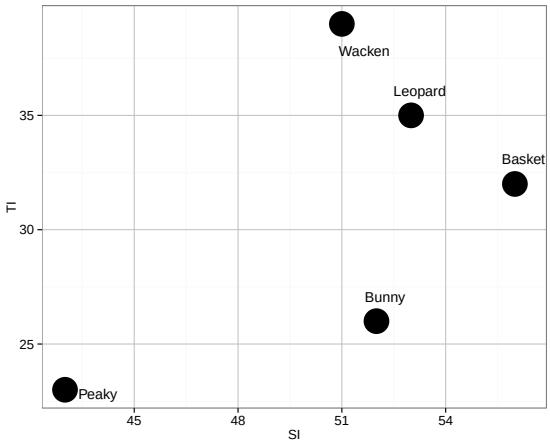


Figure 1: Spatio-temporal activity of sequences used

H.265/MPEG-H Part 2, HEVC is a product of joint collaboration between ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG) [2]. HEVC is built on the basics of its predecessor AVC, however, new features and improvements in individual coding tools enable for approximately two times higher bitrate reduction while preserving the same quality compared to AVC. Further information on HEVC can be found in [2, 4].

Although HEVC is a direct successor to AVC, there are not many studies comparing perceived subjective quality of both codecs. The authors of [5] present a comparative study of AVC and HEVC, however, the quality is compared based on objective metrics PSNR and SSIM only. Another study, [6], presents the comparison results of HEVC with AVC and the VP9 algorithm developed by Google. However, only objective metrics and bitrate measurements are available.

We present a study with a realistic scenario when the users are watching the videos using their own equipment (PC, laptop). In order to do so, we use crowdsourcing as a tool to acquire a sufficient number of participants. In this scenario we use the most typical bitrates and standard resolution only as streaming higher resolution (1080p, 4K) might not be feasible for many users due to the limitations of either displaying devices or their Internet connections.

3 TEST SETUP

For the purposes of a subjective quality evaluation a dataset of encoded sequences was created. We used 5 video sequences - a short animated movie Big Buck Bunny (further denoted as *Bunny*), a *Basket* sequence showing a basketball match and a *leopard* sequences available from the Consumer Digital Video Library[1]; and 2 real-life sequences *Wacken*, which is a short clip from a music concert and *Peaky* - a short clip from BBC TV Series Peaky Blinders. These sequences were acquired from a high-quality blu-ray disc. One frame of each content is shown in Fig. 2. These sequences cover wide variety of characteristics, as can be seen from the Spatio-Temporal Indeces in Fig. 1.

All source video sequences were originally available in FullHD (1080p) resolution at 25 frames per second, the length of the sequences was adjusted to 10 seconds. As the purpose of this study is to compare the perceived video quality of both H.264/AVC and H.265/HEVC using crowdsourcing, we had to deal with the problem, how to play the HEVC coded sequences on the computers of the participants. As there is no simple browser-based solution for playing HEVC video yet, we compressed HEVC sequences again with close to lossless H.264. Because of this fact, we had to use lower res-

[1] http://www.cdvl.org



a) Basket b) Bunny c) Leopard d) Peaky e) Wacken

Figure 2: Used sequences

olution of the sequences used in the study. Our tests revealed that in order to achieve acceptable file sizes after the H.265→H.264 re-encoding, we could use standard resolution (576p) only. Therefore, the source video sequences were downscaled using *ffmpeg*[2] tool to meet this limitation.

3.1 ENCODING

All sequences were encoded compliant to the H.264 and H.265 video coding standards. For encoding to H.264 and H.265 we used the *x264*[3] and *x265*[4] encoder implementations, respectively. Both encoders were set to medium presets to use their most common features. All source video sequences were encoded to 5 quality levels. These levels were determined by bitrate and the selected values were {500, 800, 1000, 2000 and 3000} kbps. This range of bitrates is commonly used on the Internet for video streaming of Standard Definition video. Together, our database consisted of 50 processed video sequences.

3.2 WEB-BASED APPLICATION

A web-based test application similar to [7] was implemented, in order to gather users' opinions. The frameworks utilizes various types of monitoring the users' activities to avoid or to record cheating or not focusing on the test properly. The test consisted of a set of user-oriented questions (e.g. nationality, occupation, use of video streaming services etc.) and of the video quality evaluation part. Each user was shown 5 video sequences, 1 for each content. Afterwards, the user was asked to rate the quality on a 5 grade ACR scale. If the behavior of the users was evaluated as in-consistent by the framework (e.g. switching to another tab while watching the clip), the ratings of this user were not further processed.

The study was available as a micro job on the Microworkers[5] crowdsourcing platform. Every user could participate and was rewarded with 0.30$ upon completion of the test. More detailed information on the testing procedure can be found in [8].

4 RESULTS

For comparing the video quality of sequences encoded using both codec, we will use both the objective and subjective metrics.

In video quality assessment, typicaly full-reference (FR) objective metrics like PSNR or others are widely used. The PSNR (peak signal-to-noise ratio) is plotted for all sequences in Fig. 3 grouped by content and ordered by bitrate. It can be seen that the different contents result in different PSNR ranges for the compressed sequences, which again indicates the diverse characteristics of the videos. For each content, as expected, the PSNR increases with increasing bitrate. However, when directly

[2]http://www.ffmpeg.org
[3]http://www.videolan.org/developers/x264.html
[4]http://www.x265.org
[5]https://microworkers.com

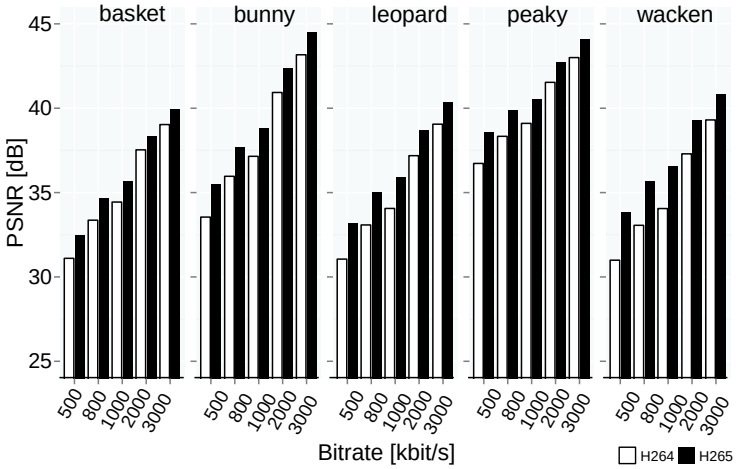


Figure 3: Objectively measured quality - PSNR.

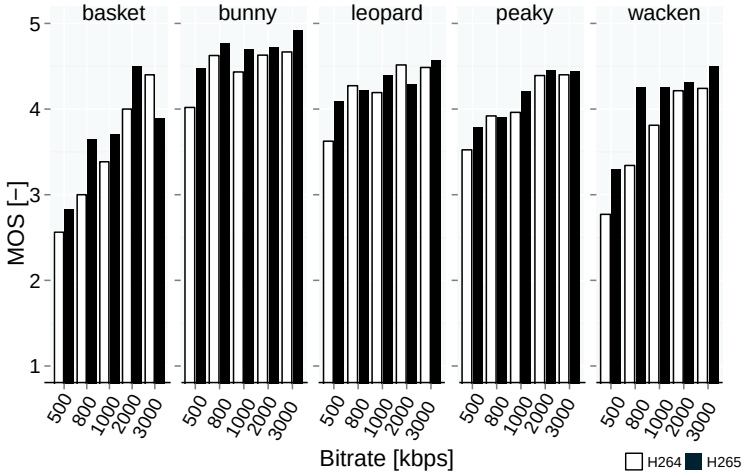


Figure 4: MOS values for used sequences, CI values are not depicted as the lengths of CI varied only slightly and were in the range of 0.5-1 .

comparing H.264 clips (white bars) to H.265 sequences (black bars), H.265 is able to always achieve a higher PSNR with the same bitrate. This effect is strongest for Leopard and Wacken sequences, which have generally lower PSNR level for the H.264 sequences as they required highly lossy compression to reach the desired bitrates. Further details on results of objective quality metrics is described in [8].

Figure 4 shows the results of the subjective crowdsourcing study. The x-axis depicts the sequences grouped by content and ordered by bitrate. White bars represent H.264 clips and black ones indicate H.265 clips. The MOS is plotted on the y-axis.

When taking a look at the MOS ratings in Figure 4, we see an increase of the MOS for each content and codec, when the bitrate increases. The influence of bitrate is significant with p value of 0.0101 for a t-test on all ratings. However, the Bunny clip, which is an animated cartoon clip, reaches high ratings almost regardless of the bitrate. This is due to the fact that the cartoon can be well compressed without many visible artifacts. A similar observation holds for the Peaky Blinders clip and its rather dark images.

When we take a look at the difference between H.264 and H.265 codecs, also a less clear result can be seen. Although most of H.265 clips reach a higher MOS than their H.264 counterparts, the CIs

nearly always overlaped. This means that we cannot easily see if the results for the two codecs are significantly different. The t-test resulted in p value of 0.0782, therefore, the influence of codec used cannot be considered as significant for the typical probability level of 5%. A reason for this could be that the study setting (i.e., bitrate range and resolution) is already well covered by H.264, such that H.265 brings only small quality improvements. In future work, the study has to be repeated for different scenarios.

4.1 CONCLUSION

In this work, we compared the perceived visual quality of videos encoded with AVC and HEVC using crowdsourcing. The most beneficial outcome of the study is, that we gathered the opinions of the real users watching videos in their home environment. However, in the proposed scenario, the influence of the codec used was not statistically significant, although HEVC encoded sequenced reached usually higher scores. Therefore, a future study should focus on other scenarios with higher resolutions. However, an issue with smooth HEVC coded videos playback must be solved in order to be able to use crowdsourcing platform.

ACKNOWLEDGEMENT

The research published in this submission was financially supported by the Brno University of Technology Internal Grant Agency under project no. FEKT-S-14-2177 (PEKOS).

REFERENCES

[1] Cisco, "The Zettabyte Era: Trends and Analysis," Cisco, Tech. Rep., 2015.

[2] International Telecommunication Union, "ITU-T Recommendation H.265: High efficiency video coding," 2015.

[3] International Standards Organization/International Electrotechnical Commission (ISO/IEC), "23008-2:2015 Information Technology – High Efficiency Coding and Media Delivery in Heterogeneous Environments – Part 2: High Efficiency Video Coding," 2015.

[4] G. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 12, pp. 1649–1668, Dec 2012.

[5] H. Koumaras, M.-A. Kourtis, and C. Skianis, "Compression performance and video quality comparison of hevc and avc," in *Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), 2014 IEEE 19th International Workshop on*, Dec 2014, pp. 75–79.

[6] D. Grois, D. Marpe, A. Mulayoff, B. Itzhaky, and O. Hadar, "Performance Comparison of H. 265/MPEG-HEVC, VP9, and H. 264/MPEG-AVC Encoders," in *Proceedings of the 30th Picture Coding Symposium (PCS)*, San Jose, CA, USA, 2013.

[7] T. Hoßfeld, R. Schatz, M. Seufert, M. Hirth, T. Zinner, and P. Tran-Gia, "Quantification of YouTube QoE via Crowdsourcing," in *Proceedings of the IEEE International Workshop on Multimedia Quality of Experience - Modeling, Evaluation, and Directions (MQoE)*, Dana Point, CA, USA, 2011.

[8] O. Zach, M. Seufert, M. Hirth, M. Slanina, and P. Tran-Gia, "Can You See the Difference? Comparison of H.264/AVC and H.265/HEVC Video Streaming," in *Quality of Multimedia Experience (QoMEX), 2016 Eighth International Workshop on*, submitted, 2016, pp. 1–6.

PREDEFINED QUALITY PROFILES IN DIFFERENT VERSIONS OF HEVC ENCODER– INFLUENCE OF PC ARCHITECTURE OF PROCESSOR

Jan Kufa

Doctoral Degree Programme (2), FEEC BUT

E-mail: xkufaj00@stud.feec.vutbr.cz

Supervised by: Tomas Kratochvil

E-mail: kratot@feec.vutbr.cz

Abstract: This paper describes a coding efficiency of the High Efficiency Video Coding (HEVC) encoder with different predefined quality profiles. To increase the encoding speed of the video at the same bitrate we can use either other implementation of the encoder or we can set different predefined quality profiles. The x265 is the implementation of the coding standard HEVC. This article complements of our previous exploration of predefined x265 profiles. There is explored the influence of processor architecture and version of x265 implementation on the speed of encoding and on the performance of video quality.

Keywords: H.265, HEVC, x265, processor architecture, PSNR, SSIM, VQM

1. INTRODUCTION

Nowadays, customers' interest for video services in a high quality is increasing. It is necessary to use encoders with higher coding efficiency because the transmission bandwidth is limited. Advanced video codecs are more flexible but also more complexes and need more time to encode video. The H.265 encoder, known as High Efficiency Video Coding (HEVC), is used for IPTV Netflix and will be used in the upcoming 2nd generation Digital Video Broadcasting – Terrestrial (DVB-T2) standard. Real-time encoding of videos is necessary in various systems but it is not necessary in home conditions. Why examine predefined setting of HEVC encoder? If a compression of video needs five hours instead of one hour, it should be proved that such time consumption has significant effect on the final video quality. The video quality highly depends on encoders profiles [1]. Hence, unanswered questions are: What is the performance of the new version of encoder? Can the new processor architecture improve the speed of video encoding?

This paper is organized as follows. Brief description of the HEVC encoder, video sequences and objective video quality metrics are outlined in Section II. Results from objective metrics are evaluated, compared and discussed in Section III. Finally, Section IV concludes the paper.

2. ENCODED VIDEOS AND CODEC

2.1. CODEC HEVC AND ITS IMPLEMENTATIONS

The x265 is the most widespread implementation of HEVC encoder. It is often used to rip the movies from DVD or Blu-ray to the personal computer (PC). The HM is the reference implementation of HEVC for video encoding however, it is not often used (slow encoding). This is the main reason why x265 is prioritized implementation because speed of encoding is higher and its impact on overall video quality is negligible. This implementation is still under the research and its new releases are available in [3]. Parameters of the used HEVC codec implementations are in Table 1.

Implementation	x265	x265
Built date	24.1.2015	24.1.2014
Compiler	GCC 4.6.3	GCC 4.6.3
Encoder version	1.4	0.6

Table 1: Parameters of used HEVC codec

Parameters of predefined quality profiles of HEVC encoder are in Table 2. Our implementation is able to measure the time needed to encode video by itself. There were defined only bitrates and profiles without any other system parameter modification and without any tune of the encoder. Dispersion of the set bitrate was about 3 %. Such dispersion is sufficiently small.

Profile	Ultrafast	Superfast	Veryfast	Faster	Fast	Medium	Slow	Slower	Veryslow	Placebo
CTU	32	32	32	64	64	64	64	64	64	64
Bframes	4	4	4	4	4	4	4	8	8	8
RC	10	10	15	15	15	20	25	30	40	60
Refs	1	1	1	1	3	3	3	3	5	5
Me	dia	hex	hex	hex	hex	hex	star	star	star	star
Merange	25	44	57	57	57	57	57	57	57	92

Table 2: Parameters of predefined profiles for encoder x265 [1].

Name	CPU	Architecture	Core/Threat	Frequency	RAM	SP Flops
C2D	C2D-E6700	Conroe	2/2	2.66 GHz	4GB 800MHz	42,4 GFlops
i5	i5-3550	Ivy Bridge	4/4	3.3 GHz	8 GB 1600 MHz	220,7 GFlops
i7	i7-2600	Sandy Bridge	4/8	3.4 GHz	16GB 1333MHz	219,7 GFlops

Table 3: Parameters of used computers

Hardware parameters of used computers are in Table 3. Computer C2D is an old PC appropriate only for office work. Computers i5 and i7 are up to date average powerful PCs. Performance of all PC configurations was calculated by parameter FLOPs. FLOPS is an acronym for floating-point operations per second and it is a benchmark of computer performance. Single precision flops were measured. The performance of C2D computer is five times weaker than the performance of i5 and i7. Computer i5 is newer than i7 but i5 is for a common user while i7 has high-end performances.

2.2. USED VIDEOS

For a set of tested videos, it was necessary to choose various kinds of videos (see Figure 1). The experiment was performed on two uncompressed video clips in RAW format. Such short video sequences have various spatial and temporal indexes (SI and TI) and different frames per second (see Table 4) [2].

Name	Frames per second [Fps]	Frames [-]	Time [s]	SI [-]	TI [-]
Tree	50	500	10	36.0	11.3
Life	30	825	27.5	30.4	23.2

Table 4: Parameters of videos used in comparison



Figure 1: Used videos “Tree” and “Life”.

2.3. OBJECTIVE QUALITY METRICS

In our experiments, only full reference objective metrics were used to evaluate the video quality. All objective video quality metrics were calculated by VQM [4] and VQMT [5] programs. The used objective video metrics were Peak Signal-to-Noise Ratio (PSNR), Structural Similarity (SSIM) and Video Quality Metric (VQM) [6].

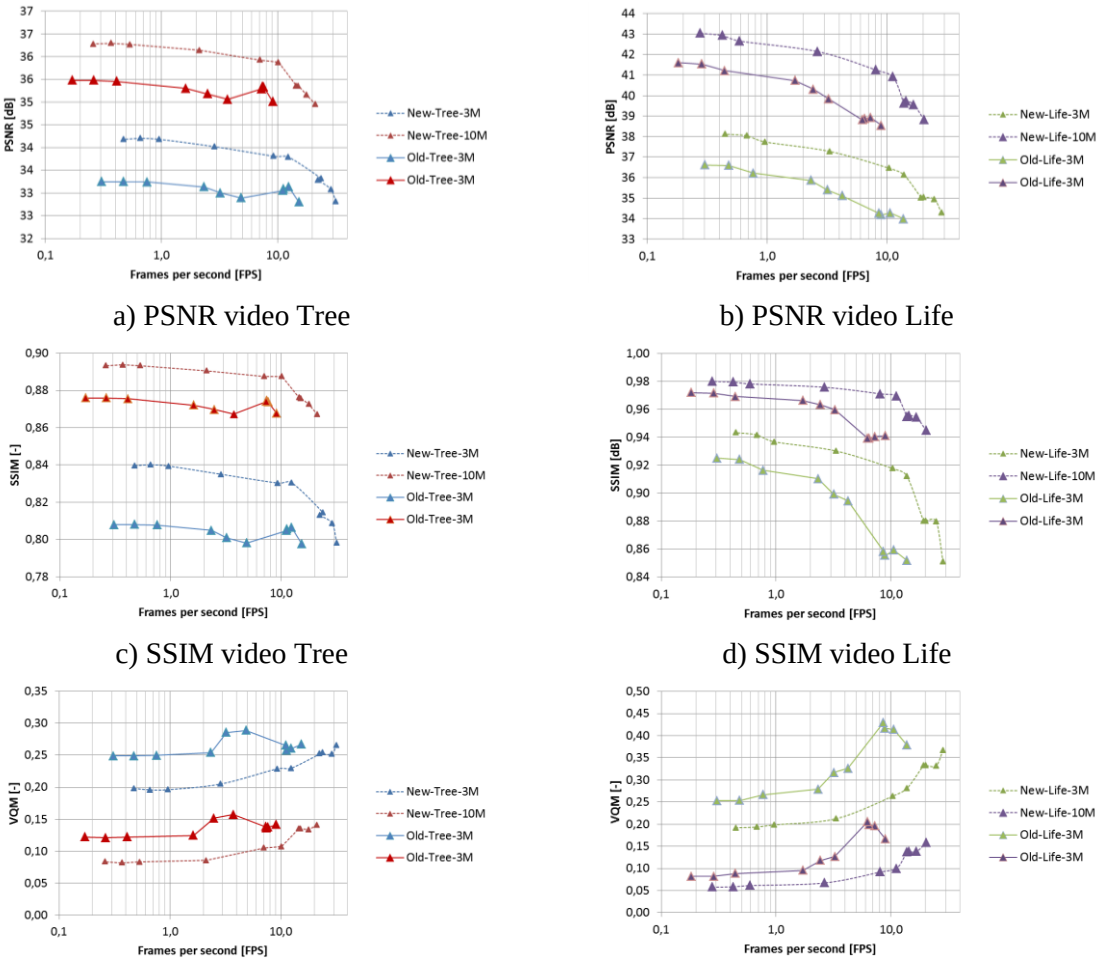
3. RESULTS AND EVALUATION

3.1. PREDEFINED QUALITY PROFILES

In this article, we also study the effect of predefined profiles on the video quality and on the sped of encoding. Similar analysis of this problem was described in [1] in detail. In our previous study, we explored that the higher quality profiles have not significant impact on the video quality, related to very slow encoding speed. When the speed of video encoding has the highest priority then it is recommended to set "Superfast" profile. The "Very Fast" and "Faster" profile provide very similar video quality like a "Superfast" but they are slower.

3.2. VERSIONS OF HEVC ENCODER

Two differently old versions of the same HEVC encoder implementation were chosen. The time difference between both versions is exactly one year (see in Table 1). Dependency of coding efficiency and the speed of encoding on the version of the encoder are shown in Figure 2.



a) PSNR video Tree

b) PSNR video Life

c) SSIM video Tree

d) SSIM video Life

e) VQM video Tree

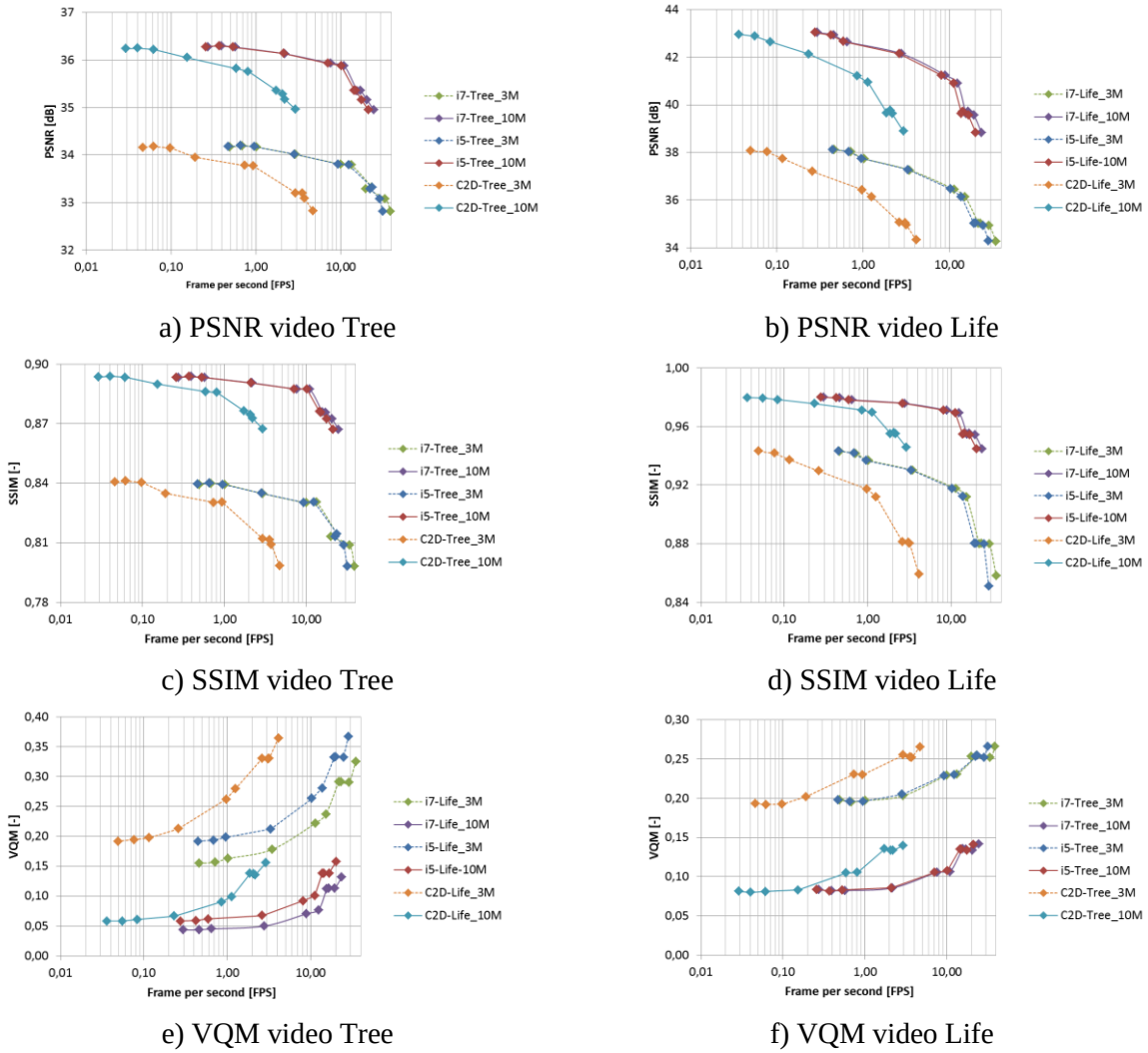
f) VQM video Life

Figure 2: Impact of the version of the encoder on the speed of encoding and video quality.

The best quality profiles (e.g. "Placebo") start on the left side of graphs and gradually, results for faster profiles (e.g. "Ultra Fast") are indicated on the right side. Name of dependencies (Legend) in figures is presented as follows: „version of the video codec_name of the video_bitrate". Results show that different versions of the same encoder have a different influence on the speed of encoding and on the final video quality. For example, the speed of the encoding at older encoder with profile "Medium" for video "Tree" at bitrate 10 Mbps was 2.5 fps whereas, with the new version of the encoder the speed of encoding was 7 fps (see Figure 2 a)). The PSNR is increased from 35.0 dB (old version of the encoder) to 35.8 dB (new version of the encoder). The performance of the old version of the encoder is not uniform. Especially, in the case of "Ultrafast" profile, the speed of encoding is increased from 9 fps to 22 fps for video "Tree" while the changes in video quality are negligible. This is very close to real-time encoding.

3.3. PROCESSOR ARCHITECTURE OF PC

In this subsection, we focus on the study of the influence of processor architecture from the point of encoding speed. Furthermore, we explore the impact of newer instruction set of the processor on the video quality. Advanced Vector Extensions (AVX) and Streaming SIMD Extensions 4 (SSE4) are new extensions of the instruction set which are supported only by the PC with processor i5 and i7. These extensions can be helpful from the point of video encoding speed. The speed of encoding as dependence of fps vs. objective metric on the processor architecture is shown in Figure 3. In this study, only the new version of HEVC encoder was used.



a) PSNR video Tree

b) PSNR video Life

c) SSIM video Tree

d) SSIM video Life

e) VQM video Tree

f) VQM video Life

Figure 3: Impact of the architecture of processor on the speed of encoding and video quality.

We know that, theoretically, the processor with C2D has approximately five times weaker performance than other two processors (see Table 3). But what effect will this have on the encoding performance? The impact of processor cores for some implementations is insignificant, for example for the default HM implementation. This implementation can handle only with one core.

Name of dependencies (Legend) in figures is presented as follows: "the name of the PC_name of the video_bitrate". For example, the speed of the encoding on C2D with quality profile "Medium" for video "Tree" at bitrate 10 Mbps was just 0.6 fps whereas, with the processor i5 and i7 it is 7 fps and 7.5 fps, respectively (see Figure 3 a)). When we compare the speed of encoding on C2D and on i5, where i5 has clock frequency higher about 24 %, the encoding speed of i5 is faster by 1200 %. It is an extreme speed up of the encoding. In the profile "Fast" the difference between i5 and i7 is only few percent. In faster profiles the difference between new processors is about 15 %. Processor i7 can fully take advantage of the Hyper-Threading. In these profiles C2D is eight times slower than i5 and i7 processors.

4. CONCLUSION

In this paper, the performance of the HEVC encoder and its predefined quality profiles on the video quality was analyzed. Our study was especially focused on the changes in coding efficiency due to the version of the encoder and on the influence of the architecture of the processor in PC. We have used three different computers. As it was mentioned, many papers deal only with different video codecs and their implementation, but study of predefined quality profiles of one implementation of the same codec is still missing. The main goal of this paper was to present a complete view about predefined quality profiles of x265 codec and explore the influence of its most probable settings on the video quality. Obtained results show that a new version of HEVC encoder has higher influence on the video quality and speed of the encoding is also higher. It was proved that PC with advanced architecture (but with the same clock frequency and with the same number of cores) could grow the speed of encoding by five times in comparison with old PC. Differences between the results obtained by various objective quality metrics are negligible (seen in Figure 2 and Figure 3).

ACKNOWLEDGEMENT

This work was supported by the BUT project no. FEKT-S-14-2177 (PEKOS). Research described in this paper was financed by Czech Ministry of Education in frame of National Sustainability Program under grant LO1401. For research, infrastructure of the SIX Center was used.

REFERENCES

[1] Kufa, J., Kratochvil, T.: Impact of Predefined Quality Profiles in H. 265 to Video Quality and Speed of Encoding. In Sborník příspěvků studentské konference Kohútka 2015. Brno, Czech Republic: VUT v Brně, 2015. pp. 32-34. ISBN: 978-80-214-5239-8.

[2] Xiph.org: Derf\'s Test Media Collection. Xiph.org Video Test Media. [online]. [cit. 2016-02-02]. Available from: https://media.xiph.org/ video/derf

[3] x265 HEVC High Efficiency Video Coding H.265. [online]. [cit. 2016-02-02]. Available from: http://x265.org/index.html

[4] Software – ITS. Video Quality Metric (VQM) Software. [online]. [cit. 2016-02-02]. Available from: http://www.its.bldrdoc.gov/resources/video-quality-research/software

[5] Software - VQMT: Video Quality Measurement Tool (École polytechnique fédérale de Lausanne). [online]. [cit. 2016-02-02]. Available from: http://mmspg.epfl.ch/vqmt

[6] Kufa, J., Kratochvil, T.: Comparison of H.265 and VP9 coding efficiency for full HDTV and ultra HDTV applications, Radioelektronika, 2015 25th International Conference, Pardubice, pp. 168-171, 2015. DOI: 10.1109/RADIOELEK.2015.7128999

RULES FOR DETERMINATION OF EXPECTED P WAVE TYPE IN ECG SIGNALS

Lucie Maršánová

Doctoral Degree Programme (1), FEEC BUT

E-mail: xmarsa08@stud.feec.vutbr.cz

Supervised by: Martin Vítek

E-mail: vitek@feec.vutbr.cz

Abstract: The aim of this work is to improve existing possibilities of automatic detection of pathological P wave in electrocardiogram. For detection of different types of P wave is crucial to choose suitable algorithm. In this paper, establishing of rules for determining which P wave type is in given situation probably present is done. This rules are derived from positions of QRS complexes. Therefore, accurate QRS detector based on a signal filtering by bandpass filter and finding the maximum in a moving window is designed.

Keywords: P wave, detection of QRS, P wave type, P wave detector

1. INTRODUCTION

Due to the ease of use, non-invasiveness and cheapness, ECG (electrocardiogram) is nowadays still the most available and widely used method for the cardiovascular system examination. The ECG reflects two heart activities: the atrial activity (represented by P wave) and ventricular activity (represented by QRS complex and T wave). The automatic and reliable detection of atrial activity is still an unsolved problem. This applies especially in the case of pathological situation [1]. In specific pathology, P wave does not respect sequence normal time ordering P, QRS and T. P wave can be dissociated from the QRS complex, hidden in the QRS complex or T wave or can be even missing. It means, that it is necessary to take into account all of these situations during P wave detection.

Correct detection of pathological P waves is necessary for automatic identification of pathologies such as 2nd and 3rd degree atrioventricular (AV) block, atrial fibrillation or flutter, supraventricular tachycardia, AV (nodal) re-entrant tachycardia, junctional ectopic tachycardia, ventricular tachycardia, idioventricular rhythm etc. Many types of software solutions are nowadays used for automatic analysis of long-term ECG e.g. [2] in practise. However, none of them can automatically detect pathologies, whose identification is based on correct detection of pathological P waves. It means, that these types of pathologies must cardiologist find in large volume of data manually.

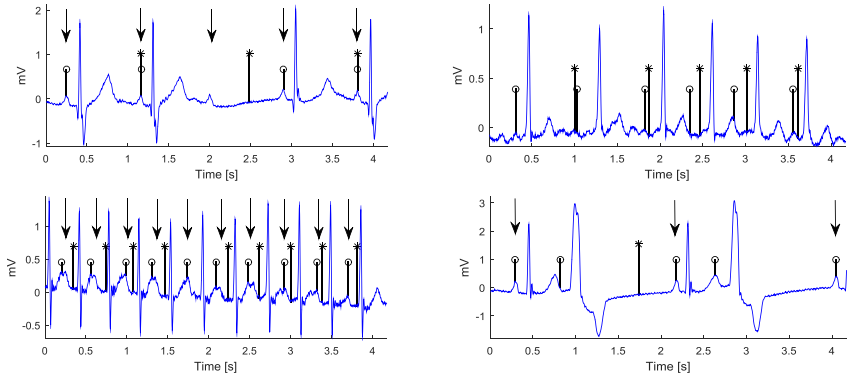


Figure 1: Illustration of wrong detection of P waves by two different P wave detectors (1. – marked by *o*, 2. – marked by *). Arrows show correct positions of P wave.

In Figure 1, examples of results of two P wave detectors are shown. First searches maximum in the area (200-60 ms) before QRS complex and marks it as P wave, second is taken from [3]. It can be seen, that both detectors fail in all cases of pathology. The reason is the nonexistence of universal algorithm which could correctly detect all types of pathological P waves (according to [1], the accuracy of P wave detection e.g. in case of AV block is about 66 % and for nodal rhythm 33 %). It follows, that for accurate detection of different pathological P waves is necessary to use different detection approaches and rules. In a given situation, it is crucial for the correct choosing of suitable P wave detection approaches to distinguish which type of P wave may just occur. In this paper, the rules for determination of which types of P wave are in a given situation probably present will be established.

2. METHODS

In Figure 2 is a scheme, which expresses a complete process of P wave detection in electrocardiogram. In this work, first three blocks are solved. The block of signal filtering is performed within the block of QRS complex detection. The principle of QRS complex detection and derivation of rules for possible P wave (pathology) type determination are described below. The rules are derived from the previous correctly detected QRS positions. The performance of correct P wave detection heavily relies on the accuracy of the used QRS detector.

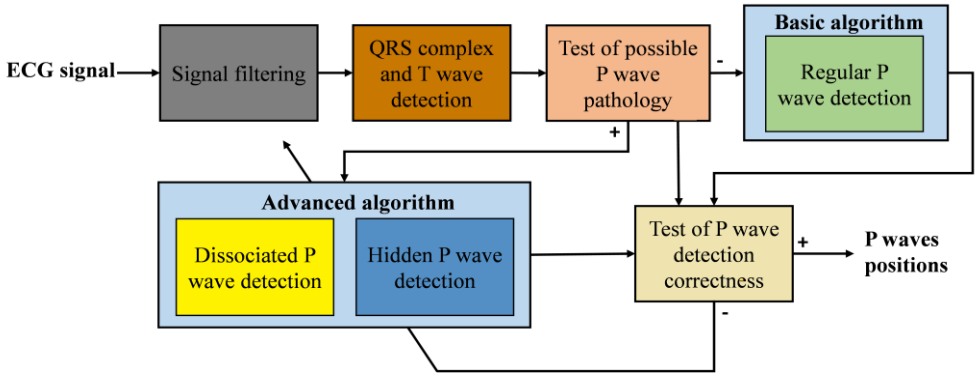


Figure 2: Complete process of P wave detection in electrocardiogram

2.1. QRS DETECTION

The designed algorithm is based on filtering ECG by bandpass filter transmitting only a limited range of frequencies and finding the maximum in a sliding window. Main frequency components of a physiological QRS complex range from about 10 Hz to 25 Hz. Therefore, almost all QRS detection algorithms use a bandpass filter that transmits a range of frequencies from 10 Hz to about 25 Hz [4]. In this work, the FIR (Finite Impulse Response) filter transmitting only a range of frequencies from 30 Hz to 35 Hz is used (see Figure 3a). This filter completely attenuate other signal components (P wave and T wave) and also most of the artefacts (baseline drift, muscle noise). The vast majority of signal from this specific range of frequencies (30-35 Hz) originates from QRS complex. Moreover, this range of frequencies is also present in pathological QRS complexes (like ventricular extrasystole or bundle of branch blocks) which detection is usually problematic.

After signal filtering, the absolute value of signal is calculated (shown in Figure 3b). It allows finding QRS complexes having a higher amplitude of S wave than R wave. Then, finding maxima in a sliding window is performed. The length of this window is key for the correct detection and it is established to maximize the number of true positives (TP), while keeping the number of false negatives (FN) and false positives (FP) detections on minimum. If the value of the considered sample is the highest from all values of samples in a given window, it is stored in output feature vector. In other case is to feature vector stored zero (see Figure 3c). After finding maxima, thresholding of feature signal is performed. If the value of a feature signal is higher than the threshold, the position of the QRS complex is determined. Threshold is determined according to a standard deviation,

which is computed in the sliding window. The detected QRS complexes marked by *o* are shown in Figure 3d.

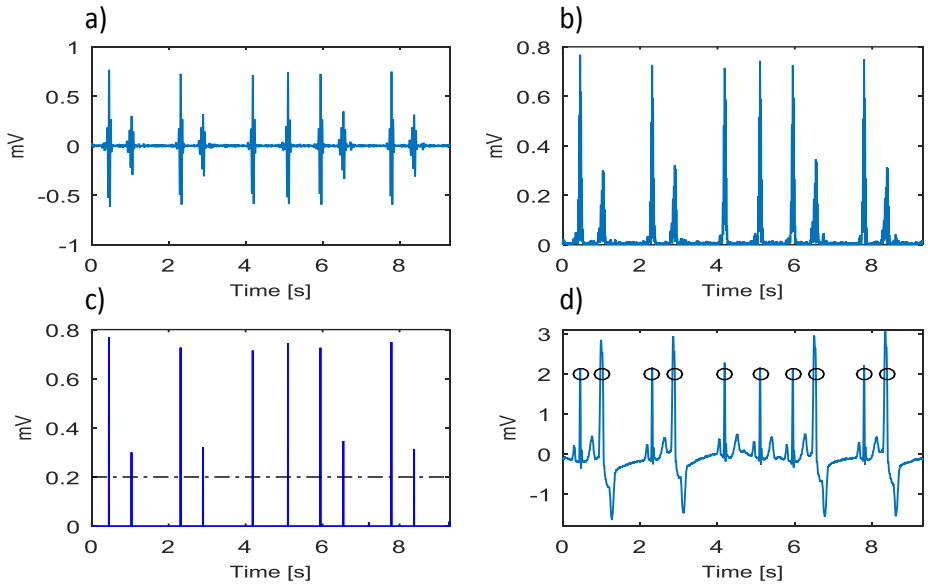


Figure 3: Complete process of QRS detection a) signal filtered by bandpass filter (30-35 Hz), b) absolute value of signal a), c) signal after finding maxima in sliding window, d) results of QRS detection.

2.2. RULES FOR P WAVE TYPE DETERMINATION

As was mentioned above, for correct detection of all types (pathological) of P waves in ECG signals is necessary to predict which type o P wave is in a given situation probably present. In this chapter, rules for determination of possible P wave type are derived [5]. According to the results of these rules, appropriate algorithm for the specific P wave type detection will be chosen.

There are 4 possible P wave types:

- **Case 0 - No P wave:** P wave is not present (atrial fibrillation, atrial flutter, ventricular extrasystole)
- **Case 1 - Physiological P wave**: P wave is present before QRS complex (physiological situation (sinus rhythm), supraventricular extrasystole)
- **Case 2 - Dissociated P wave:** P wave is dissociated from QRS complex (2^{nd} and 3^{rd} degree AV block)
- **Case 3 - Hidden P wave:** P wave is hidden in the T wave or QRS complex (AVRT, AVNRT, junctional rhythm, junctional extrasystole)

The rules for the prediction of which type of P wave is in a given situation probably present are based on knowledge of heart physiology and manifest of pathology. They are derived from intervals between two QRS complexes (RR intervals). These rules are:

I. RR_n and RR_{n-1} intervals are shorter than threshold $T_1 = 400$ msec → **Case 3**

II. RR_n is shorter than $T_2 = (0.3*RR_{n-1})$ and $(RR_n + RR_{n+1}) > (1.9*RR_{n-1})$ → **Case 0**

III. RR_n interval is longer than threshold $T_3 = (1.5*RR_{n-1})$ → **Case 2**

IV. In Case 1, not found P wave and heart rate (HR) is higher than $T_4 = 60$ → **Case 0**

V. In Case 2, not found P wave → **Case 3**

VI. In Case 1, not found P wave and (RR_n, $RR_{n-1,}$ RR_{n-2}) > 1000 msec → **Case 3**

VII. Other situation → **Case 1**

3. RESULTS AND DISCUSSION

3.1. TESTING OF QRS DETECTION

The proposed real-time ECG detector was evaluated on the MIT-BIH Arrhythmia Database, which is the most commonly used database for QRS detector evaluation. The MIT-BIH Arrhythmia Database contains 48 half-hour two-channel records. The sampling frequency is 360 Hz. Two benchmark parameters are calculated to evaluate designed QRS detection algorithm and to compare it with previous works. The sensitivity (SE) and positive predictivity (PP) of the QRS detection algorithm are computed as: SE=TP/(TP+FN) and PP=TP/(TP+FP). The sensitivity reports the percentage of true QRS that were correctly detected by the algorithm. The positive predictivity reports the percentage of QRS detections which were in reality true beats. Results are listed in Table 1.

Table 1: Result of QRS detector testing on the MIT-BIH Arrhythmia Database (SE-sensitivity, PP-positive predictivity)

Record	SE [%]	PP [%]	Record	SE [%]	PP [%]	Record	SE [%]	PP [%]	Record	SE [%]	PP [%]
100	99.17	99.15	113	100	100	201	97.4	100	217	97.92	97.96
101	99.79	99.73	114	99.84	99.79	202	99.53	100	219	99.95	100
102	97.94	97.94	115	100	100	203	95.81	99.58	220	100	100
103	99.95	100	116	99.3	99.83	205	99.28	100	221	99.13	100
104	98,65	98.48	118	100	100	207	99.52	90.56	222	99.92	99.92
105	99.46	99.19	119	93.1	93.46	208	93.84	97.81	223	99.62	99.96
106	99.56	99.95	121	100	97.39	209	100	100	228	99.76	99.42
107	99.67	99.63	122	100	100	210	97.85	99.96	230	100	100
108	99.04	98.26	123	100	100	212	100	100	231	100	100
109	99.84	99.96	124	99.57	99.57	213	98.74	99.72	232	100	99.89
111	99.91	99.95	117	100	100	214	99.69	99.87	233	99.71	100
112	100	100	200	99.85	100	215	99.67	100	234	99.89	100
									TOTAL	**99.22**	**99.31**

The mean accuracy (SE=99.22 %, PP=99.31 %) of the proposed QRS detector is comparable to other algorithms reported in the literature. In the study [6], which used a similar approach was achieved mean SE=99.59 % and PP=99.56 %. However, not in the most problematic record. In signal 108 achieved only SE=96.28 % and PP=92.17 %, while our detector achieved SE=99.04 % and PP=99.26 %. Advantage of designed detector is also that is not using a "search-back" strategy to search the ECG again in the case that interval between two detected QRS is too long. It means, that proposed algorithm can operate in real-time unlike [6].

Moreover, further improvements of the algorithm can be achieved by using adaptive window length respecting the heart rate for maxima searching. The next possible improvements can be achieved by using more bandpass filters transmitting different frequency ranges and combining their results.

3.2. TEST OF RULES FOR P WAVE TYPE DETERMINATION

The rules (defined in chapter 2.2) for determination of expected P wave type were tested on the ECG signal shown in the Figure 4 which is composed from parts of pathological ECG signals from MIT-BIH arrhythmia database. Signal with determined areas of expected P wave types according

to the rules I-VII is shown in Figure 4. This signal contains supraventricular tachycardia, sinus rhythm, AV block of 2st degree, sinus rhythm, atrial fibrillation and junctional rhythm, exactly in this order. Expected types of P waves are distinguished by different colour. The area of expected physiological P wave is marked by green, dissociated P wave by yellow, hidden P wave is marked by blue and the part where is probably no P wave is marked by grey.

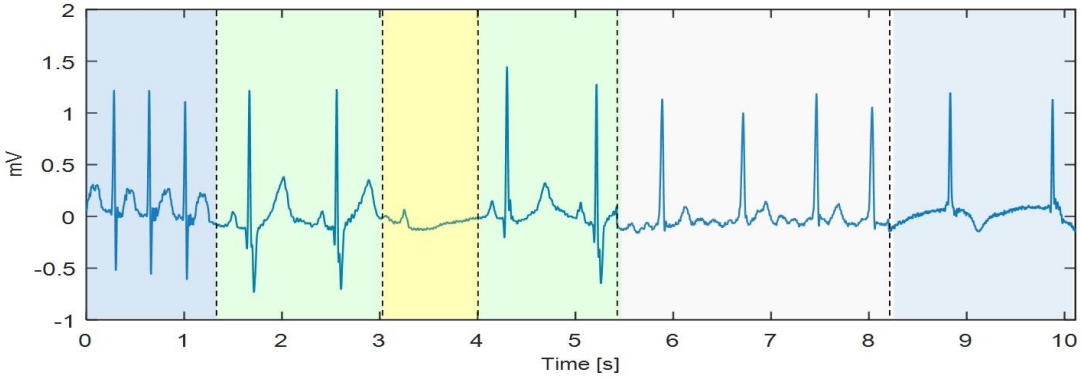


Figure 4: The result of rules for expected P wave type demarcation: blue = case 3 (hidden P wave), green = case 1 (physiological P wave), yellow = case 2 (dissociated P wave), grey = case 0 (no P wave)

4. CONCLUSION

In this work, the rules for the determination of which P wave type is in given situation probably present were established. This rules are improving existing possibilities of pathological P wave automatic detection in electrocardiogram. According to the rules, it is possible to distinguish four P wave types and then chose suitable algorithm for its detection. This rules are derived from QRS complexes positions, knowledge about heart function and manifest of its pathology. Therefore, accurate QRS detector tested on MIT-BIH arrhythmia database with mean SE=99.22 % and PP=99.31 % based on a signal filtering by bandpass filter and finding the maxima in a sliding window was designed.

REFERENCES

[1] PORTET, F. P wave detector with PP rhythm tracking: evaluation in different arrhythmia contexts. *Physiological Measurement* [online]. 2008, 29(1), 141-155 [cit. 2016-03-04]. DOI: 10.1088/0967-3334/29/1/010. ISSN 0967-3334.

[2] Biomedical Systems Century C3000 Holter System Specifications. *METEC: Marketing of speciality products for cardiology laboratories and hospital wards in Denmark and Sweden* [online]. ©2015.

[3] VÍTEK, M.; KOZUMPLÍK, J.: ECG SEEKER. Software for delineation of ECG signals. Department of biomedical engineering, Brno university of technology, Kolejní 2906/ 4, 612 00 Brno, Czech republic

[4] KOHLER, B.-U.; C. HENNIG; R. ORGLMEISTER. The principles of software QRS detection. *IEEE Engineering in Medicine and Biology Magazine* [online]. 21(1), 42-57 [cit. 2016-02-26]. DOI: 10.1109/51.993193. ISSN 07395175.

[5] KUSUMOTO, F. *ECG Interpretation : To Clinical Application ECG Interpretation :* 2009. ISBN 9780387888798.

[6] AFONSO, V.X.; W.J. TOMPKINS; T.Q. NGUYEN a S. LUO. ECG beat detection using filter banks. *IEEE Transactions on Biomedical Engineering* [online]. 46(2), 192-202 [cit. 2016-03-04]. DOI: 10.1109/10.740882. ISSN 00189294.

Doktorské projekty

Kybernetika a automatizace

COMPARISON OF THE METHODS USED TO LOW LEVEL DC VOLTAGE MEASUREMENT

Lucie Obšilová

Doctoral Degree Programme (1), FEEC BUT

E-mail: xobsil01@stud.feec.vutbr.cz

Supervised by: Zdeněk Bradáč

E-mail: bradac@feec.vutbr.cz

Abstract: This paper deals with comparison of three methods used to measure low level DC voltage. It is about Reference Step Method, potentiometric method and direct method. Measurement was held in cooperation with the Czech metrology institute. Comparison of the methods consist of a graphic comparison and a key comparison. Results showed good agreement of all methods except for potentiometric method at one range of a device under test.

Keywords: DC voltage, Reference Step Method, potentiometric method, uncertainty

1. INTRODUCTION

Reference step method, potentiometric method and direct method are three methods which can be used to measure low level DC voltage. Reference step method evaluates gains of ranges of a calibrator and uses only commonly available instruments. Potentiometric method is based on a potentiometer Measurement International 8000A which divides the voltage of a Zener reference. Direct method used a voltmeter which is calibrated by Josephson Voltage Standard.

All methods were used to calibrate a device under test Fluke 5720A. Measurement by each method was repeated several times. Measured data were corrected and the uncertainties were evaluated. The graphs for comparison were created and the key comparison were evaluated.

2. METHODS USED TO LOW LEVEL DC VOLTAGE MEASUREMENT

Direct method is based on the use of a voltmeter and a calibrator as a device under test (DUT). Voltmeter measures the absolute value of the DUT. During comparison the voltmeter was Keithley 2182A and the tested calibrator was Fluke 5720A. Before measurement the voltmeter was calibrated by Josephson Voltage Standard. The Josephson Voltage Standard is described in detail in [1].

After calibration the gain of voltmeter Keithley 2182A was evaluated. Then the errors of Keithley 2182A were calculated and the data measured by Keithley 2182A were corrected. The standard expanded uncertainty was evaluated according GUM for all measured values. Ten measured data were used for type A evaluation of uncertainty.

Reference Step Method (RSM) was developed by L. A. Christian at New Zealand [2]. The method evaluates gains of ranges of device under test. The main advantage of RSM is that it uses commonly available instruments. Basic prerequisite is very good linearity of the voltmeter.

Principle of the method consists in measurement the ratio of gains of different ranges of device under test. RSM uses two calibrators and a voltmeter. The first calibrator is a source of reference voltage and the second calibrator represents DUT. The voltmeter is used for measurement the difference between two calibrators. The connection arrangement is shown in **Figure 1**. The device under test was calibrator Fluke 5720A, the reference calibrator was a Datron 4808 and voltmeter was a HP 3458A.

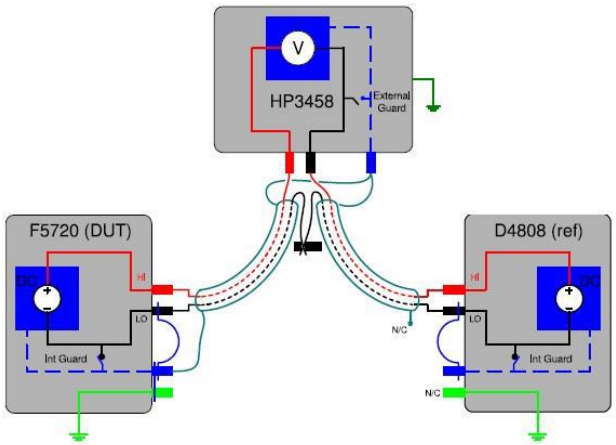


Figure 1: Connection arrangement for Reference Step Method [2]

At the end of whole measurement other method was used to calibrate one range of DUT. During comparison it was the precision Zener reference which measures absolute values of one range of the DUT. Then the absolute values of voltage for all ranges of DUT were calculated. The uncertainties were evaluated by Monte Carlo method which was used because of complicated evaluation uncertainties according GUM. Matlab was used for evaluation uncertainties by Monte Carlo method. RSM is fully automated, but it is time consuming. The RSM is described in detail in [2].

The last method is potentiometric method. This method is based on using the potentiometer Measurement International 8000A [3]. The potentiometer divides the voltage of a Zener reference, which is calibrated by Josephson Voltage Standard. The voltmeter measures difference between divided value of the Zener reference and the DUT. The circuit arrangement is displayed in **Figure 2**. The Zener reference used during comparison was Fluke 732A, the voltmeter was Keithley 2182A and the DUT was Fluke 5720A.

The measurement is fully automated. The uncertainties were evaluated according GUM by potentiometer software because of the large amount of variables. Ten measured data were used for type A evaluation of uncertainty. The potentiometric method is described in detail in [3].

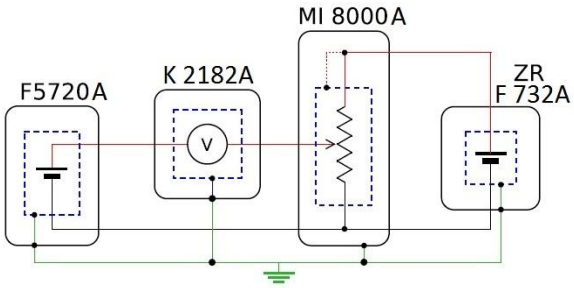


Figure 2: Circuit arrangement for the potentiometric method [2]

3. COMPARISON

Comparison was made for two ranges of the calibrator Fluke 5720A. Calibration points were selected from 1mV to 10mV for both polarities at range 220mV. Calibration points were from 3V to 10V for both polarities at range 11V. The device under test was measured three times by direct method, four times by RSM and twice or three times by potentiometric method. The reason for the difference in the number of measurement by each method was organization of work at CMI. All measurements lasted about 6 days (140 hours). During this 6 days were alternated methods due to the avability of devices at CMI. It was not necessary to deal with drift of the calibrator Fluke 5720A because of its very good stability. [5]

The first part of comparison was a graphic comparison. Graphs show differences between absolute values of measured voltage and nominal values of voltage including expanded uncertainty. The x-axis labels give time in hours when the measurement was made. Uncertainties were evaluated for a level of confidence of 95.45 % for every calibration point. The second part of comparison was the key comparison which is described in detail in [6].

4. ACHIEVED RESULTS

All methods are in good agreement at range 11 V of the calibrator Fluke 5720A for each calibration point and both polarities. It is displayed in Figure 3. Data measured by direct method are very different, which is shown in Figure 4. These different values are in good agreement with other methods because of the larger uncertainty of direct method. The larger uncertainty is consisting especially of the uncertainty calculated from specification of the voltmeter Keithley 2182A. The uncertainty of Keithley 2182A depends on the measured value. Voltmeter Keithley 2182A is better for measurement of voltage lower than 1 V because of the stable gain. It will be better to use for instance Fluke 8508A instead of Keithley 2182A for voltages higher than 1 V.

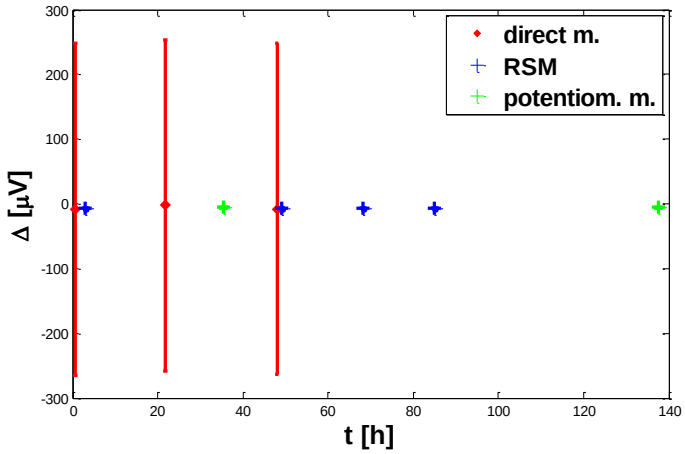


Figure 3: Difference *Δ* between absolute value of measured voltage and nominal voltage 4 V, uncertainties are at level of confidence of 95.45 %

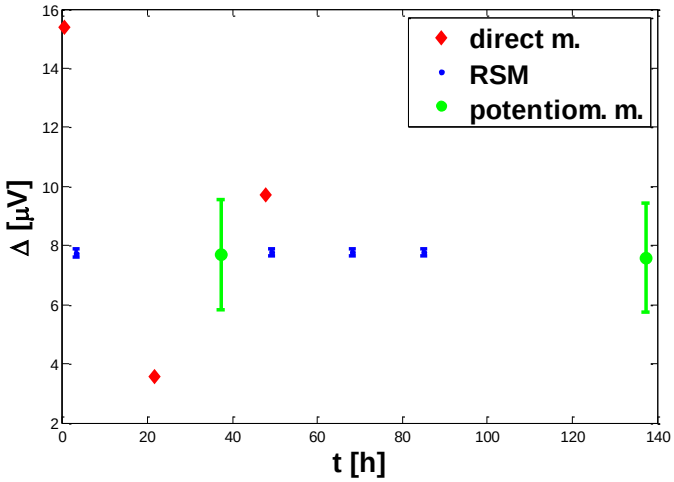


Figure 4: Difference *Δ* between absolute value of measured voltage and nominal voltage -9 V, uncertainties are at level of confidence of 95.45 %, the differences of direct method are displayed without uncertainties

RSM and direct method are in good agreement at range 220 mV of the calibrator Fluke 5720A. Data measured by potentiometric method are different from other methods for nominal voltage lower than 10 mV. Potentiometric method is in good agreement only with direct method (Figure 5) except for three measured points (1 mV, 2 mV, 3 mV). Figure 6 shows that potentiometric method is not in good agreement with other methods for nominal value 2 mV. The difference between data measured by potentiometric method and RSM is approximately 35 ppm for nominal value 10 mV and approximately 400 ppm for nominal value 1 mV. It can be deduced that potentiometric method is more different from other methods with decreasing voltage. Therefore, it was measured voltage for nominal value 50 µV. Difference between direct method and potentiometric method was approximately 800 ppm. Consequently, difference between potentiometric method and other methods increase with decreasing voltage.

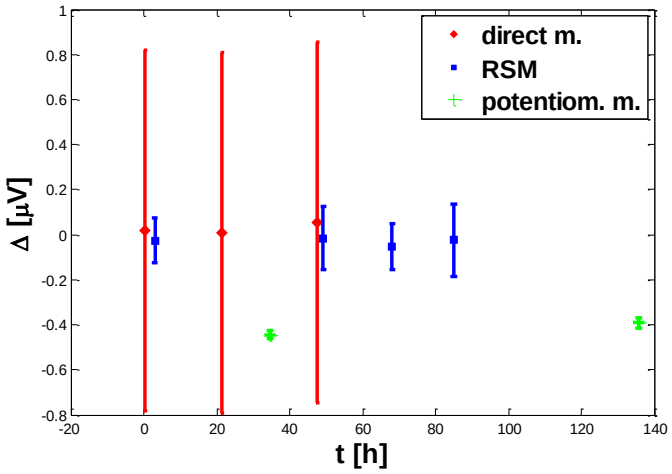


Figure 5: Difference *Δ* between absolute value of measured voltage and nominal voltage 9 mV, uncertainties are at level of confidence of 95.45 %

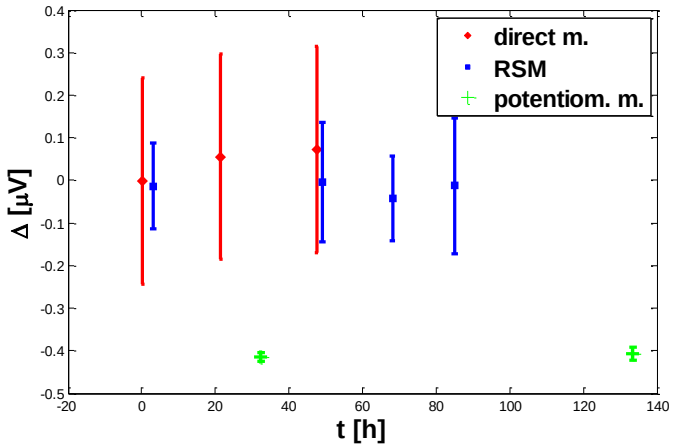


Figure 6: Difference *Δ* between absolute value of measured voltage and nominal voltage 2 mV, uncertainties are at level of confidence of 95.45 %

The uncertainties of potentiometric method are very small for measured points smaller than 10 mV. It is the second reason why potentiometric method is not in good agreement with other methods. The uncertainties of potentiometric method were evaluated by potentiometer software. When the potentiometer measures small value it uses larger dividing ratio so its uncertainty should be larger. Therefore, it was found that potentiometer software do not evaluate the uncertainty correctly.

The second part of comparison was Procedure A of the key comparison for both ranges of the calibrator Fluke 5720A. The most accurate method at range 11 V of Fluke 5720A was the Referen-

ce Step method. At range 220 mV of Fluke 5720A it was evaluated that potentiometric method is the most accurate method but this result was not correct. The reason was the very low uncertainty of the potentiometric method. Therefore, the potentiometric method was eliminated from the key comparison. Then the most accurate method was the direct method at range 220 mV of the calibrator Fluke 5720A.

5. CONCLUSION

The results of the comparison showed that all methods are in good agreement at range 11 V of the calibrator Fluke 5720A. Data measured by voltmeter Keithley 2182A are very different but it corresponds to greater uncertainty of direct method.

Direct method and RSM are in good agreement at range 220 mV of the calibrator Fluke 5720A. Potentiometric method is in good agreement only with direct method. For calibration point 1 mV, 2 mV and 3 mV potentiometric method is not in good agreement with both other methods. It was deduced that difference between potentiometric method and other methods increase with decreasing voltage. It was found that potentiometer software do not evaluate the uncertainty correctly. Therefore, it should be corrected the errors in software.

According to the key comparison the most accurate method at range 11 V of the calibrator Fluke 5720A was the Reference Step method. The direct method was the most accurate method at range 220 mV of calibrator Fluke 5720A.

ACKNOWLEDGEMENT

The completion of this paper was made possible by the grant No. FEKT-S-14-2429 - „The research of new control methods, measurement procedures and intelligent instruments in automation" financially supported by the Internal science fund of Brno University of Technology.

REFERENCES

[1] Jeanneret, B. & S. P. Benz. Application of the Josephson effect in electrical metrology. *The European Physical Journal Special Topics* [online]. 2009, vol. 172, issue 1, p. 181- 206, [cit. 2014-10-14]. DOI: 10.1140/epjst/e2009-01050-6, http://link.springer.com

[2] Sira, M., Early, M., Streit, J. & L. Christian. Comparison of the Reference Step Method with a MI 8000/8001 potentiometer. *2012 Conference on Precision electromagnetic Measurements* [online]. 2012, p. 540-541 [cit. 2014-11-05] DOI: 10.1109/cpem.2012.6251042.

[3] Early, M., Sira, M., Anderson, B., Christian, L., Gunnarsson, O., Rydler, K. & J. Streit. A Simple Build-Up Method for the DC Voltage Scale of a Source. *IEEE Transactions on Instrumentation and Measurement* [online]. 2013, vol. 62, issue 6, p. 1600-1607 [cit. 2014-11-08]. DOI: 10.1109/tim.2012.2230734, http://ieeexplore.ieee.org/xpls/abs all.jsp?arnumber=6251021&tag=1

[4] Honig, R., Sira, M. & J. Streit. Use of the BVD for Traceability of Bipolar DC Voltage Scale from 1 mV Up to 1200 V. *NCSL International Workshop and Symposium*. San Antonio, Texas, USA, 2009.

[5] Obšilová, L. *Low level DC voltage measurement*. Brno University of Technology, Faculty of Electrical Engineering and Communication, 2015. 89p. Master´s thesis supervisor was Ing. Marie Havlíková, Ph.D.

[6] Cox, M. G. The evaluation of key comparison data. *Metrologia* [online]. 2002, vol. 39, issue 6, p. 589-595 [cit. 2014-12-06]. DOI: 10.1088/0026-1394/39/6/10.

CONTINUOUS REALIZATION OF MONOCULAR STRUCTURE FROM MOTION

Jan Klečka

Doctoral Degree Programme (2), FEEC BUT

E-mail: klecka@feec.vutbr.cz

Supervised by: Karel Horák

E-mail: horak@feec.vutbr.cz

Abstract: This paper describes a concept and practical aspects of a realization of an algorithm used for 3D scene reconstruction from multiple views captured by single moving camera. Position and orientation of the camera are a priory unknown and are calculated on-the-fly from tracked point correspondences.

Keywords: Structure from motion, multiple view reconstruction, visual SLAM, monocular SLAM

1. INTRODUCTION

Reconstruction of an environment by processing its pictures taken from different views has become a popular topic in recent years. Two major reasons for this are: their use in mobile robotic, where a single camera can be used e.g. for improving a navigation of an unmanned vehicle, and augmented and virtual reality applications.

Algorithm described in this paper can be referred as SLAM (Simultaneous Localization and Mapping) algorithm, because camera position and orientation are calculated simultaneously with 3D reconstruction, although in this particular realization trajectory is not being considered as part of the result. For single realizations of SLAM algorithms a varied group of sensors is being used e.g. matrix TOF, RGB-D cameras, camera stereo pair. The main difference between usage a single camera and the other sensors is an ambiguity of the final reconstruction. While data from mentioned sensors can be usually processed into reconstruction ambiguous only in a link to the global coordinates system (i.e. zero point and direction has to be chosen) in data captured by single camera can be find only monocular decencies and due to this the reconstruction is, in addition, ambiguous in overall scale (i.e. the distance unit has to be chosen). This property of monocular reconstruction can be useful – same device and algorithm can be used in dimensionally diverse environments.

2. ALGORITHM OUTLINE

Scene reconstruction from camera picture can be done in several ways. In this realization, I use a technique, based on reobservation a set of projections of 3D points, called passive triangulation. This technique can compute points 3D coordinates from at least two observation of its projection (correspondences). Such a reconstruction have a singular state when reconstructed point and camera centers lie all on the same line. This state occurs for example if observations are made from the same center of projection or reconstructed point is in camera movement direction.

So, the reconstruction problem leads to the correspondence search problem which is in dense variant computationally demanding and its computational complexity rapidly increase with the image count; that's why I decided to use only two pictures for reconstruction. However, two consecutive images captured by moving camera will be, most probably, close to the singular state of reconstruction, because their centers of projection will be close together. To avoid this, I develop a fast meth-

od, based on sparse correspondence points tracking, which decide whether the images are plausible for reconstruction.

Several point correspondences (specifically 8 or more) are also necessarily for another triangulation assumption - known camera projection matrix for every used image. I utilize the epipolar geometry constraints to get this information.

Epipolar geometry can be, also, used for facilitation of dense correspondence search by the process called rectification, therefore, deform processed images in such way that all correspondence will be in the same row (or column) of both deformed pictures i.e. correspondence can be searched only in one dimension instead of two.

Algorithm can be sum up into four points:

1. Point tracking, until camera movement, is significant to scene dimensions
2. Computation of camera position and orientation
3. Images rectification
4. Disparity map computation and scene reconstruction

2.1. POINT TRACKING

Realization of point tracking has been done by pyramidal modification Lucas-Kanade algorithm described in [4]. It's a standard feature tracking algorithm, reliable enough for this application. As a feature detector, I use the Harris detector.

Detection whether camera movement is significant to scene dimensions I treated by checking how tracked correspondences fit the 2D homography model. Let's have an image I_1 with feature vector $\mathbf{x}$ and image I_2 with feature vector $\mathbf{x}'$ which corresponds to $\mathbf{x}$ both feature vectors are in homogenous coordinates. Decision whether I_1 and I_2 should be processed by reconstruction algorithm is based these two equations:

$$err = \min_{\mathbf{H}} \left(\sum \left\| \mathbf{x}' \times \mathbf{H}\mathbf{x} \right\| \right) \tag{1}$$

$$J = \mathrm{Med}\left(\mathrm{d}\left(\mathbf{x}', \mathbf{H}\mathbf{x} \right) \right) \tag{2}$$

The equation (1) represents homography estimation by minimization of algebraic error it can be computed fast because it's in homogenous coordinates linear. The equation (2) express the decision criterion so the median of geometric error. The value of this criterion is checked for every picture and when it exceeds predefined threshold pictures are assumed as appropriate for scene reconstruction. Plausible threshold values can differ for different cameras but in my experiments optimal value lie between 20 and 80 most often, I use 50.

2.2. POSITION AND ORIENTATION COMPUTATION

Computation of camera position and orientation thru constraints of epipolar geometry can be done in several ways. I decided to utilize essential matrix decomposition as described in [1]. This approach works with so-called normalized image coordinates $\mathbf{x}_n$ which can be computed from image coordinates using the internal camera parameters.

$$\mathbf{x}_n = D^{-1}(r)\mathbf{K}^{-1}\mathbf{x} \tag{3}$$

Where: $\mathbf{K}$ is the matrix of linear internal camera parameters and $D(r)$ is non-linear distortion function dependent on the distance from camera principal point.

The essential matrix $\mathbf{E}$ determined by the following equation:

$$\mathbf{x}'_n \mathbf{E} \mathbf{x}_n = 0 \tag{4}$$

This matrix can be scaled and decomposed using singular value decomposition into:

$$\mathbf{U}\,\mathrm{diag}(1,\ 1,\ 0)\mathbf{V}^T = \mathbf{E} \tag{5}$$

Where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrixes.

The decomposition then can be used for evaluation the external parameters of suitable projection matrixes $\mathbf{P}_n = [\mathbf{I} \mid \mathbf{0}]$ and $\mathbf{P}'_n = [\mathbf{R}_{rk} \mid \mathbf{t}_{rk}]$ in following way: Let's have constant matrix $\mathbf{W} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, then:

$$\mathbf{R}_{rk} = \{\mathbf{U}\mathbf{W}\mathbf{V}^T; \mathbf{U}\mathbf{W}^T\mathbf{V}^T\} \qquad \mathbf{t}_{rk} = \pm s\mathbf{u}_3 \tag{6}$$

Where: $[\mathbf{R}_{rk} \mid \mathbf{t}_{rk}]$ is the relative coordinate transformation between camera coordinates system in step $k-1$ and k, $\mathbf{u}_3$ is the third column of $\mathbf{U}$ and s is the scale factor.

As we can see this procedure leads to four possible solutions (two possible $\mathbf{R}_{rk}$ and two possible $\mathbf{t}_{rk}$ cause four combinations of $\mathbf{P}'_n$) the rights solution is the one which causes 3D reconstruction with positive depth in coordinates systems of both cameras. I check this using randomly chosen point correspondence.

Actual camera position is determined by global coordinates system transformation which is given by combination of all previous relative transformations:

$$[\mathbf{R}_{gk} \mid \mathbf{t}_{gk}] = \prod_{i=1}^{k} [\mathbf{R}_{ri} \mid \mathbf{t}_{ri}] \tag{7}$$

Finally, the scale factor should be determined. Because in my solution the true "metric" scale factor is ambiguous I fix the factor value for the initial image pair to 1. In the other pairs I hold the reconstruction scale by minimizing following criterion:

$$S_{crit} = \sum d\left([\mathbf{R}_{gk} \mid \mathbf{t}_{gk}]\hat{\mathbf{X}}, \mathbf{x}'_n\right)^2 \tag{8}$$

Where $\hat{\mathbf{X}}$ is the 3D reconstruction of $\mathbf{x}'_n$ made from image $k-2$ and $k-1$.

2.3. RECTIFICATION

The goal of rectification is to transform both images geometrically in such way that all epipolar lines will be parallel. Todays are widely used two rectification algorithms. The first is the linear rectification which is based on deforming both pictures by homography transformations which map epipoles into the infinity. But this concept will fail when one of the epipoles lie inside processed image. The second algorithm is the polar rectification [2]. This approach has no limitation and is based on remaping pictures into polar coordinates with the center in epipole; every corresponding epipolar lines are mapped on the same row. I chose to use the polar rectification because it can handle the general movement of the camera.

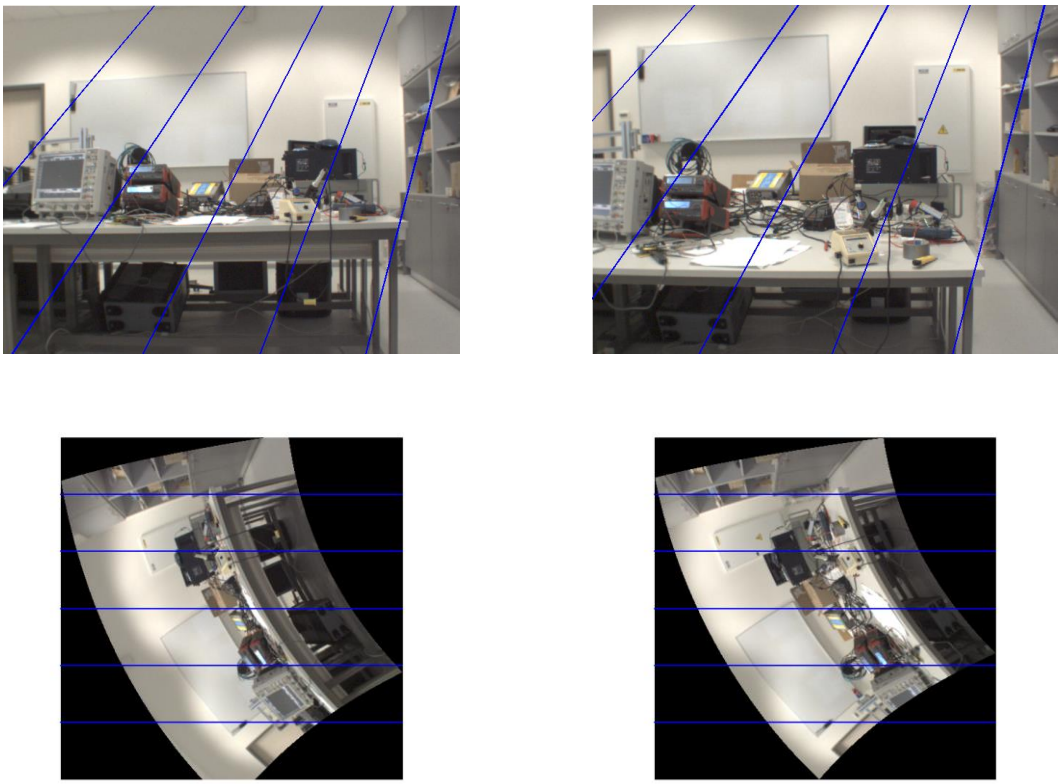


Figure 1: Example of polar rectification with few epipolar lines

2.4. RECONSTRUCTION

After rectification, the dense correspondence search can be performed. The result of such pixel-wise search in rectified coordinates is usually called disparity map. It exists several methods for disparity map generation I decided to use the semi-global method described in [3]. Because in rectified coordinates the correspondences can be found only in the same row (eventually column) disparity map is usually presented as a grayscale image of the differences in the column (eventually row) coordinate.

With known correspondences reconstruction can be done by homogenous triangulation method described in [1]. This requires a recalculation correspondences from polar coordinates and search for null space vector of 4x4 rank three matrix.

$$\begin{bmatrix} x\mathbf{p}^{3\mathrm{T}} - \mathbf{p}^{1\mathrm{T}} \\ y\mathbf{p}^{3\mathrm{T}} - \mathbf{p}^{2\mathrm{T}} \\ x'\mathbf{p}'^{3\mathrm{T}} - \mathbf{p}'^{1\mathrm{T}} \\ y'\mathbf{p}'^{3\mathrm{T}} - \mathbf{p}'^{1\mathrm{T}} \end{bmatrix} \mathbf{X} = \mathbf{0} \tag{9}$$

Where: $\mathbf{p}^{i}$ is the i-th row of the matrix $\left[\mathbf{R}_{g(k-1)} \mid \mathbf{t}_{g(k-1)}\right]$, $\mathbf{p}'^{i}$ is the i-th row of the matrix $\left[\mathbf{R}_{gk} \mid \mathbf{t}_{gk}\right]$ and x, y, x', y' correspondence coordinates transformed from polar rectification.

3. RESULTS

I implemented whole described algorithm in C++ using the OpenCV libraries and perform the real data experiment. As image source, I use a calibrated CCD camera with wide-angle lens, resolution 640x480. I move the camera manually by my hand.



Figure 2: Example of result of the real data experiment

4. CONCLUSION

I developed an algorithm which can conceptually provide multiple view reconstruction from monocularly captured image sequence. The main weakness of this algorithm is a drift of the scale estimation which is caused by poor $\hat{\mathbf{X}}$ estimate in equation (8). Further work will focus mainly on improving the quality of these point reconstructions.

ACKNOWLEDGEMENT

The completion of this paper was made possible by the grant No. FEKT-S-14-2429 - „The research of new control methods, measurement procedures and intelligent instruments in automation" financially supported by the Internal science fund of Brno University of Technology.

REFERENCES

[1] HARTLEY, Richard a Andrew ZISSERMAN. *Multiple view geometry in computer vision*. 2nd ed. Cambridge: Cambridge University Press, 2003. ISBN 05-215-4051-8.

[2] POLLEFEYS, M., R. KOCH a L. VAN GOOL. A simple and efficient rectification method for general motion. In: *Proceedings of the Seventh IEEE International Conference on Computer Vision* [online]. IEEE, 1999, s. 496-501 vol.1 [cit. 2016-03-04]. DOI: 10.1109/ICCV.1999.791262. ISBN 0-7695-0164-8. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=791262

[3] HIRSCHMÜLLER, Heiko. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence* [online]. 2008, **30**(2), 328-341 [cit. 2015-02-26]. DOI: 10.1109/TPAMI.2007.1166. ISSN 01628828.URL: http://ieeexplore.ieee.org.ezproxy.lib.vutbr.cz/xpl/articleDetails.jsp?arnumber=4359315

[4] BOUGUET, Jean-Yves. Pyramidal Implementation of the Lucas Kanade Feature Tracker. *Intel Corporation Microprocessor Research Labs* [online]. 2000 [cit. 2016-03-07]. URL: http://robots.stanford.edu/cs223b04/algo_tracking.pdf

CONTROL PLATFORM FOR MULTIPHASE MOTORS

Matúš Kozovský

Doctoral Degree Programme (1), FEEC BUT

E-mail: xkozov00@stud.feec.vutbr.cz

Supervised by: Petr Blaha

E-mail: blahap@feec.vutbr.cz

Abstract: This article focuses on the design and subsequent implementation platform for controlling multi-phase motors. Article describes the functionality of the assembled platform and the possibility of its usage to control either the individual motors or a single multi-phase motor. Platform is able to control Maximum nine-phases motor. The platform uses the three core AURIX TriCore microcontroller TC275. This paper also describes the distribution control program to initialization of device and separated control application generated by the Simulink automatic code generation. Finally interconnection these partial programs into a single control application.

Keywords: TC275, AURIX, TriCore, nine-phases power stage,

1. INTRODUCTION

At the present, for conversion electrical energy into mechanical energy are frequently used AC motors with different structures, the squirrel cage, winding in the rotor, with permanent magnets etc. These motors with advances in technology, improve efficiencies [1] For this reason, there are requirements for continuous improvement of control methods. This is made possible by the increase in computing power of processors. These nowadays integrates all the necessary tools for motor control, measurement and computing control algorithms.

Conventional three-phase motors are suitable for common use. However, if higher reliability is desired, we need to draw attention to the multi-phase motors. This allows to control the magnetic field in the stator in the event of failure. It may be the failure of one or more phases (depending on the total number of phases) and the motor can continues to operate, but power output is lower. Three-phase AC motor in case of failure cannot be correctly controlled. In case of failure of one of the phases magnetic field vector stops to rotate. In this case, we can only control magnitude of this vector and direction from negative to positive values.

2. ASSEMBLING PLATFORM

The platform is compiled from the Evaluation Board KIT_AURIX_TC275_TFT and three 3-Phase BLDC/PMSM Low Voltage Power Stages from Freescale. Internal structure of used microcontroller can be seen on Figure 1. To interconnect all boards together, specialized PCB was created. This board contain the separating interface for Incremental rotary coder (IRC) which provides information of the rotor current position. Pulse width or their count over time can also be recalculated to speed of rotation. Board also includes voltage reference for processor A/D (Analog to Digital) convertors. These A/D convertors are used to measure actual current values. For this reason, voltage reference must be connected to measuring circuits of power stage boards. Current is measured by shunt resistors. Gain and offset is configured to current range -10 to 10 A. Measured value of voltage is in rage 0 to 3.3 V. Power Stage allows to control power transistors individually or using fixed deadtime. In this case value of deadtime can be only changed during platform initialization.

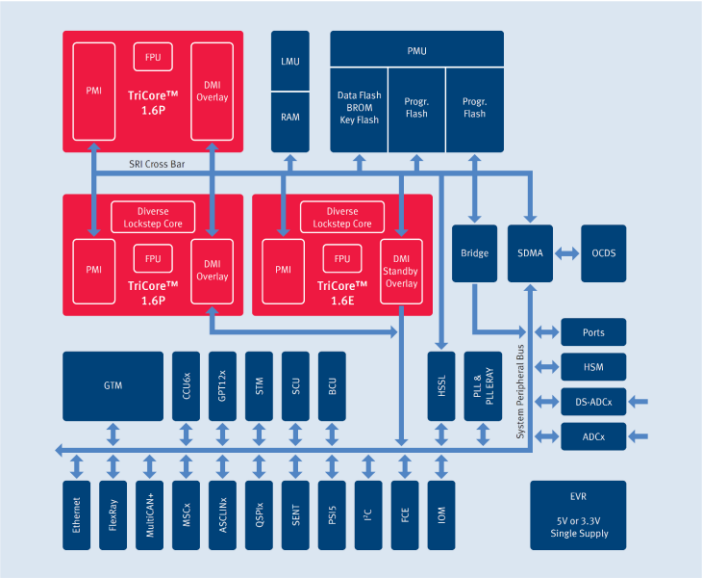


Figure 1: The internal structure of the processor [4]

These Low Voltage Power Stages is also able to diagnose some failures as overcurrent or under-voltage of power stage. There is used serial peripheral interface (ISP) for initialization of these powers stages. This transfer bus is also used to read actual status of power stage. Output voltage from power stage is controlled by pulse width modulation (PWM). Requested duty cycle is calculated by control algorithm and transferred to power stage.

Fullbridge is typically used for three-phase motor [2]. Fullbridge is controlled by six PWM signals. Three signals control top transistors and another three control bottom transistors. Switching time of PWM signals for complementary transistor must be shifted because of transistor's closing and opening time. In our case it is a grouping of three individual fullbridge into one functional unit, this created a nine-phase bridge. All three boards have same voltage. Their DC buses are also interconnected. For controlling this platform, 18 independent PWM outputs of the control board are required. Block diagram can be seen in the Figure **2**2.

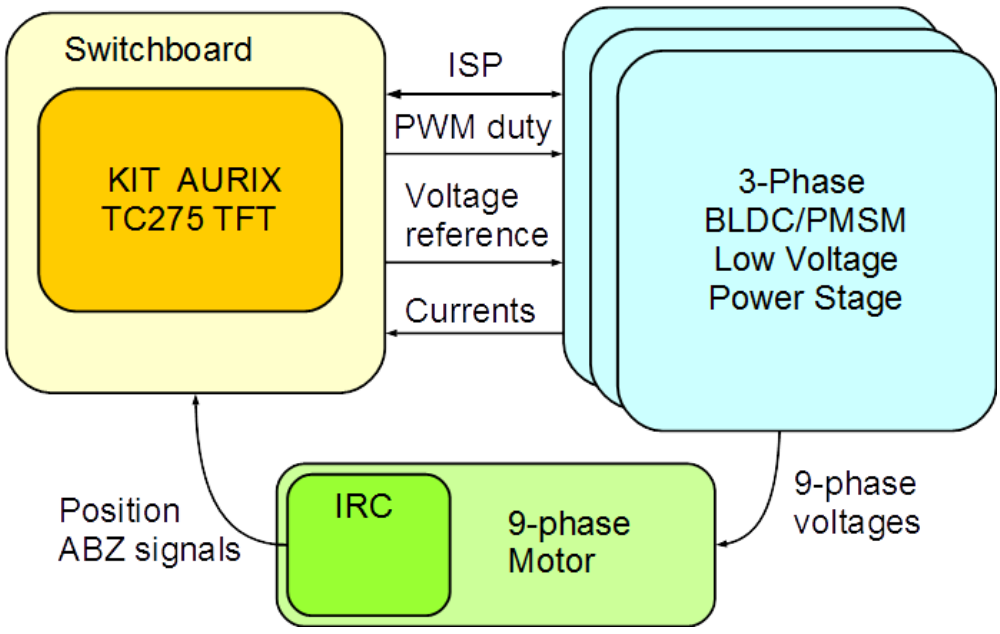


Figure 2: Block diagram

3. SOFTWARE

The main program is divided into two parts, the first is used to initialize the component and the processor peripheral interface. The second part is generated by Simulink through automatic code generation. This block includes the motor control algorithm and also it is possible to implement communication with a superior system

3.1. INITIALIZATION OF PERIPHERALS

After starting the processor it is necessary to configure its peripherals. There is necessary to send PWM signals from the control unit. TC275 includes a module for generating PWM signals. There is Timer Output Module (TOM) which allows to create necessary 18 synchronized PWM signals. Switching off and switching on time can be configured for every signal independently of the other. During initialization there is configured deadtime and PWM frequency. TOM signal also synchronize processor versatile Analog-to-Digital Converter (VADC).

Analog values representing current should be measured at start of PWM period. At this time all the lower transistors are opened and it is guaranteed that each coil current passes through the shunt resistor. VADC is configured for 12bit-conversion. There is also required nine phase currents and DC bus voltage. Value is converted after measurement to float numbers and normalized to measured range. Current can be measured on several shunt resistors or by single current sensor. Using only one sensor requires a complex measurement process [3]. In our case, for each measured current is used individual shunt resistor.

During initialization is also configured General Purpose Timer Unit (GPT12). This unit is configured as counter. Value is changed on both edges of encoder AB signals. Speed can be measured by width of AB pulses or by counting pulses over time.

There are two different interrupts generated in software. First interrupt is generated by GTP and is used to recalculate actual rotor rotation speed. Second interrupt is generated by VADC. After current and DC bus voltage measuring. In this interrupt should be recalculated PWM duty. Maximum execution time depends on PWM period time.

3.2. AUTOMATIC CODE GENERATION

Nowadays there is many possibilities to create control algorithm. Preferred properties are particularly transparency, the possibility of additional adjustments and fast implementation. All of these requirements can be fulfilled using automatic code generation. Simulink provides the ability to create C functions from Simulink's block diagram. During automatic code generation can be specific function for example C function "*sin()*" replaced by special variations of these functions "*fast_sin()*" optimized for target HW.

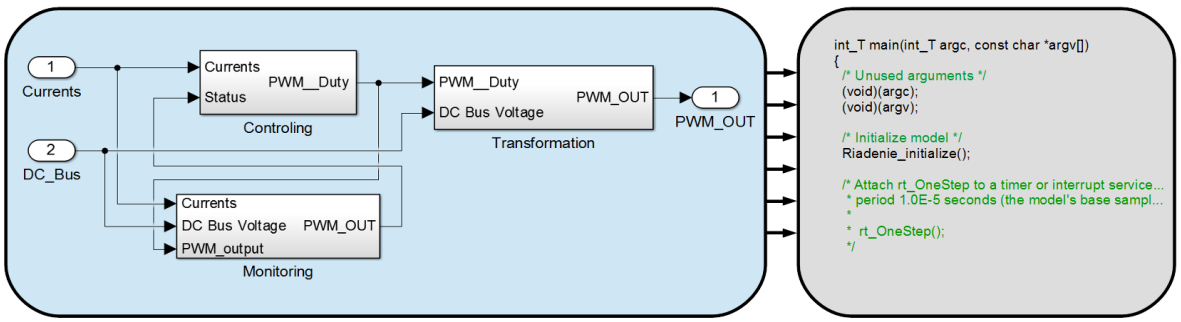


Figure 3: Automatic code generation

The control block diagrams in Simulink is clearly arranged and can be easily adjusted. The result code can be quickly embedded into first part of software. There is initialization function which configure control algorithm initial state. Second function should be called periodically and calculate new PWM state in every period of PWM cycle. The Figure 3 shows the principle.

4. FUNCTIONALITY VERIFICATION

Individual components were interconnected and were sequentially verified. Each functionality was tested. The basic software for time-synchronized measuring of analog values and rotor position was created. The program includes all the features mentioned below. For testing it was first created a simple program which periodically changed the PWM duty. The individual signals were observed by oscilloscope. It was important to be sure that individual signals correspond to requirements even in extreme conditions. It must also be fulfilled that the deadtime is always inserted correctly to avoid possible damage to transistors. After basic testing, motor was connected to platform. Connection to the motor allows to check the current measurements and also measure the position and speed. It was finally implemented a simple program for motor control. Program contain simple Field oriented control (FOC) containing PI controllers for current control in the DQ coordinates. Transformation uses the current rotor position measured by encoder. This program allows control the current value. Superior controller allows set speed. Final assembled platform can be seen in the following Figure 4.

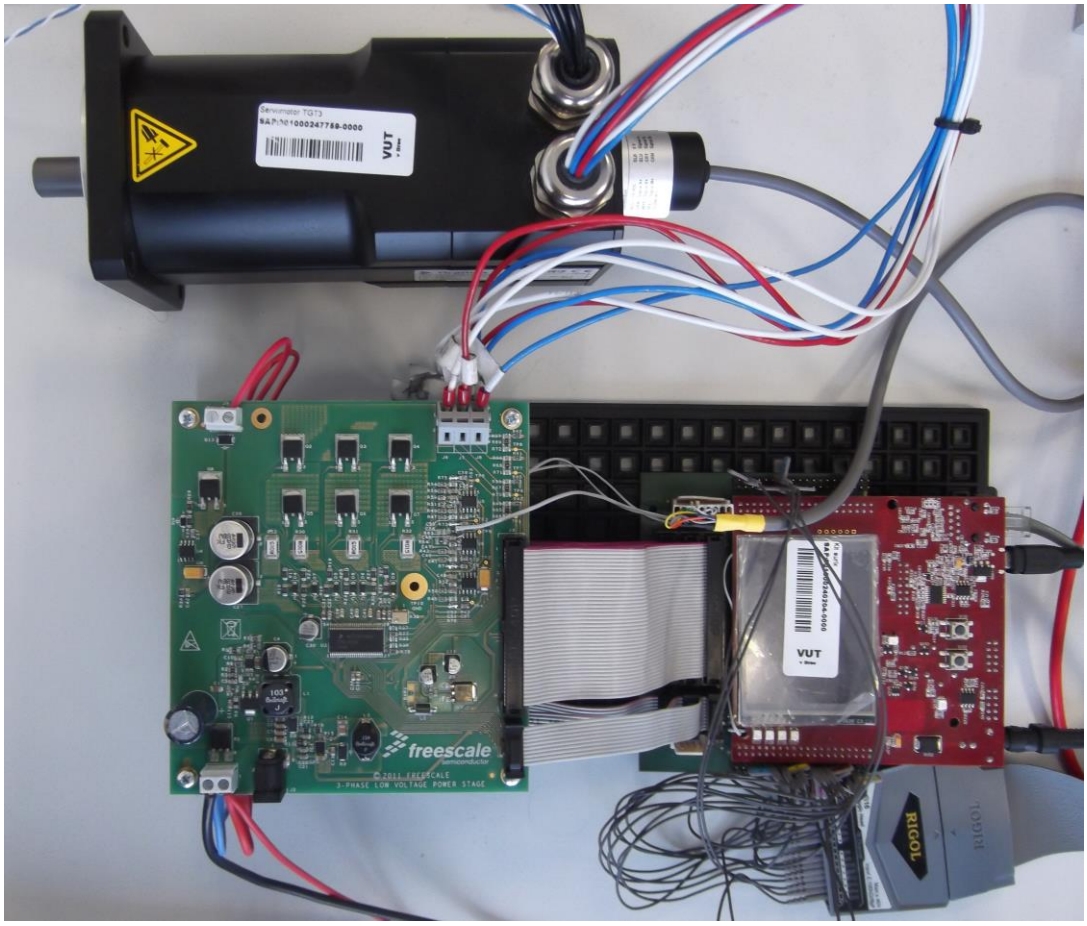


Figure 4: Final implementation

5. CONCLUSION

This work designed platform allows the controlling of a multi-phase motor for a total of nine phases. Basic software for device initialization was tested, and following checked the output signals as required. For basic functionality it has also been checked current measurement and measurement of the position and speed. Basic functions for set PWM duty also correctly insert deadtimes. Possibility of automatically code generation and its insertion into the whole software was tested on the platform. Simple control algorithm that allows to adjust the speed or torque of the motor was implemented. This method of implementation shows that it is easy to change the final control structure as well as adding new features.

ACKNOWLEDGEMENT

The completion of this paper was made possible by the grant No. FEKT-S-14-2429 -„The research of new control methods, measurement procedures and intelligent instruments in automation" financially supported by the internal science fund of Brno University of Technology.

REFERENCES

[1] J. Malinowski, J. McCormick, and K. Dunn, "Advances in construction techniques of AC induction motors preparation for super-premium efficiency levels," in *IEEE Industry Applications Society 50th Annual Petroleum and Chemical Industry Conference, 2003. Record of Conference Papers.*, pp. 205–211.

[2] F. A. Samman, T. Waris, T. D. Anugerah, and M. N. Z. Mide, "Three-phase inverter using microcontroller for speed control application on induction motor," in *2014 Makassar International Conference on Electrical Engineering and Informatics (MICEEI)*, 2014, pp. 28–32.

[3] Y. Cho, T. LaBella, and J.-S. Lai, "A Three-Phase Current Reconstruction Strategy With Online Current Offset Compensation Using a Single Current Sensor," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 7, pp. 2924–2933, 2012.

[4] Infineon Technologies, 'Highly Integrated and Performance Optimized', 2016, p. 54 <http://www.infineon.com/dgdl/TriCore_Family_BR-2014.pdf?fileId=db3a30431f848401011fc664882a7648>

DISTRIBUTED FRAMEWORK FOR TESTING MACHINE LEARNING METHODS

Jan Klusáček
Doctoral Degree Programme (4) , FEEC BUT
E-mail: xklusa00@stud.feec.vutbr.cz

Supervised by: Václav Jirsík
E-mail: jirsik@feec.vutbr.cz

Abstract: When designing new Machine learning (ML) methods, solid testing is very important part of process. This article describes a framework that was created for automated testing and comparison of different ML methods. This framework allows to automate most of tedious and recurrent tasks related to comparison and testing of new methods.

It consists of two parts. First part is intended for work with results of ML methods. It allows to compare results of different methods with different settings. It allows to create tables and graph from these results and it also perform statistical tests on these results.

It is often necessary to perform considerable number of test runs on different datasets and with different settings for purpose of comparison and statistical test. For these reasons it is advisable to automate these tasks as much as possible. Automation of these tasks is purpose of second part of this framework. It allows to divide, plan and execute tasks on remote machines.

Whole framework is written using Python and Django framework allowing to easily extend customize it for particular task.

Keywords: Machine learning, Python, Parallelization, Web interface

1 INTRODUCTION

Regular approach when developing Machine learning (ML) methods[1] is to compare them to the existing method. It is also possible to search for best parameters of single method by comparing its results to results gained with different settings. In both cases, approach is pretty similar. It consists from creation of model using particular settings and training datasets. This model is than used to predict values of the testing dataset. Resulting predictions are then compared with correct labels and performance of this model is evaluated.

1.1 COMPARING RESULTS OF MACHINE LEARNING METHODS

Generally it is not sufficient to perform just one test on one dataset to compare performance of different models[1]. Improvement in one test can be easily coincidence. For this reasons, methods are usually compared on multiple different datasets. One of problems when evaluating performance of model is division of the dataset between training and testing subsets[2]. Using bigger training dataset improve quality of model, but resulting smaller testing set means lower reliability of performance estimation and vice versa. Cross validation can be used to improve reliability of these tests further.

[1]Methods considered in this work are belongs to group of supervised learning. Described framework can work with unsupervised methods but evaluation of quality of these methods is different from supervised ones

[2]It is desirable to use two disjoint datasets for training and testing model. If one dataset is used for both training and testing, resulting performance estimates are optimistic and underestimate error. This is a result of ignoring overfitting error (lack of a generalization).

The main principle behind cross-validation is division of original dataset to multiple disjoint subsets[2]. Different combinations of these subsets are subsequently used to create models and evaluate their performance. Performances acquired this way are subsequently used to estimate the performance of model created using the whole dataset. This approach allows to use all data for creation of a model and for testing (not in same time, but in many different models models). The result is a better a model with a good estimation of its performance. Cross validation also allows to estimate a stability of a method (its sensitivity to changes in input data) by comparing changes of performance on slightly different datasets.

1.2 TASK PARALLELIZATION

From the previous description, it's obvious that a simple comparison of two methods require considerable number computations. Depending on character of used models and datasets, these computations are often very time consuming and can take many hours or days when executed on a single processor of consumer computer, for this reason parallelization of testing is desired. One of possible solutions of this problem is to use high performance machine. This approach clearly mean additional costs, but today it is not necessary to buy own hardware. It is possible to rent desired hardware for hourly fee[3]. Another solution is usage of ordinary computers when they are not used. This approach doesn't require additional investments to hardware, but it is limited by available time (computers can be used only when they are not required for their primary function) and it is not suitable for some tasks[3].

1.3 PARALLELIZATION TYPES

As mentioned previously, parallel tasks can be divided according to their requirements for exchanging intermediate results (inter-process communication). Most demanding tasks require frequent communication and are very sensitive on the communication latency. These tasks are called fine grained.

Fine grained tasks can be run on multicore microprocessors or on multiprocessor machines, that allows very fast (high bandwidth and short latency) communication between individual cores. Depending on type of task it's possible to run these tasks on computers interconnected with high performance network such as Infiniband with various performance penalties.

Coarse grained tasks also needs to exchange intermediate results, but not so often as fine grained tasks therefore requirement on the communication between machines are much lower and Ethernet is usually sufficient.

Some of the tasks require almost no communication between individual tasks. These tasks are called embarrassing parallel tasks.

Tasks used for testing of ML methods are mostly in class of embarrassingly parallel tasks because they are independent of each other. Each task creates one model with specific settings and a dataset and evaluate its performance. Only necessary communication is concentration of all results required for comparison and statistical tests.

2 TESTING FRAMEWORK

To address problems introduced above simple framework was created. This framework is written in Python. Its goal is not to create a complete application that can be easily deployed, requiring complex configuration depending on a particular usage. It is rather framework that takes care of tedious tasks and let user focus on a real problem. Instead of using complex configurations that try to address all possible scenarios, this framework implements a functionality required by all tasks and a customization is achieved by inheriting from prepared classes. This way it is easy to add a custom behavior, without need to prepare support for high configurability. A drawback of this approach is

[3]Some tasks require fast communication between threads, that is not achievable using a network. Another problem is if task require big amounts of memory

that user needs at least basic programing skills, but it should be no problem for people designing ML methods.

Whole testing framework is written in Python 2.7. using Django framework for access to database and visualization. Python was chosen because it allows a fast development and it is very widespread in machine learning community. It also allows to write a multiplatform code[4]. Basic diagram depicting frameworks function is in Fig. 1

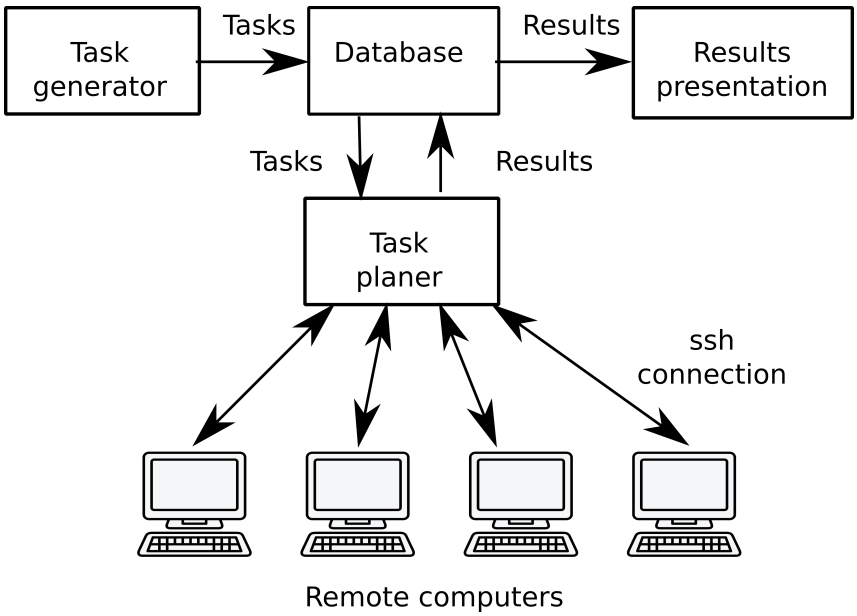


Figure 1: Framework diagram

All test configurations and their results are stored in database[5]. It allows to easily access configuration using web interface and at the same time it simplify presentation of results.

2.1 TASK CREATION

First part of process is a creation of tasks. All necessary functionality for one task are implemented in class `Task`. This class is derived from Django `Model` class, this mean that it is possible to easily load and save this class to database. Most important property of this class is `task_data`. This property contains a pickled[6] object that will be transferred and executed on a remote machine. User of the framework have to prepare an object with method `work` and save it to this property. It will be automatically executed on an available remote machine and a results will be automatically transferred back. `Task` class also allows to define some additional properties. One task can for example require another task finished before it is executed (this is useful if one task require some results from a previous task). Every task can also add requirements on remote machine (minimal available RAM).

2.2 TASK PLANER

Another part of the framework is a task planner. This python module can run as a thread in a Django server, or it can be run as a standalone daemon. It has two main tasks. It maintain list of available re-

[4]Described testing framework is currently able to run on both Windows and Linux without any modifications

[5]SQLite is currently used as a database engine but thanks to Django framework, it is possible to easily switch to a more powerful database engine if higher performance is required

[6]cPickle is a python library that allows to serialize a generic python object to a binary string and later recreate it using this string. It has some limitations (for example it is not possible to serialize file handles, sockets and similar resources), but it is very comfortable way of transferring an object to a different process, possibly on a different machines

mote machines and their state (e.g. ready, waiting, offline) and it assigns `Task` loaded form database to the available remote machines.

Access to remote machines is provided by ssh with usage of private keys. A list of all unfinished tasks is loaded every time one of previous tasks is finished and suitable task is selected from list for the available machine. Pickled data of task are then transferred to the remote machine, they are unpickled and executed. When the computation is finished its results are saved to the database and a next task is selected and executed.

Several possibilities for communication with remote machine were considered before selecting ssh. It would be possible to send task data to remote machine using http or https protocols, but it would require installation http server on each remote machine. Another possibility is usage of telnet or custom tcp/ip protocol. These possibilities were ruled out because they would require unnecessary amount of work and would take long time to debug, especially if they should be secure to use over internet. For these reasons ssh was selected. It not only allows to securely send data over untrusted networks but also allows to control remote machines. It is also heavily used and tested, so there should be minimal number of bugs. Its drawback is that its encryption can slow down transfer of large amounts of data. This can be overcome by disabling encryption (which would require recompiling ssh client and server), or selecting faster cipher (Blowfish, ARC4).

2.3 RESULTS VISUALIZATION

Last part of the framework is visualization of the results. Similarly to previous parts the framework does not provide any complete solution for visualization as is, but it is very easy to use existing libraries to create tables and graphs as necessary. Used Django framework provides easy way to create web interface.

All results are saved in the database, so it is quite easy to access them and make simple statistics directly using SQL functions. Whenever more complex processing is necessary, it is possible to use existing python modules such as ScyPy. Tables can be easily created using library tables2 (provided as part of Django framework). MatPlotLib can be used for creating various graphs.

3 EXAMPLE OF USAGE

The described testing framework is currently used for development of the new method of transformation of feature space transformation. This task is little different from testing of classification methods, but it is still possible to use this framework. In this case, the task form chain of requirements longer than usual. The first task uses a whole dataset to find transformation of feature space. Subsequent task divides original dataset to ten parts for purpose of a cross-validation and apply the transformation found in previous task. Following tasks prepare different models with different settings on datasets from previous task and evaluate their performances. All these tasks are repeated for five different original datasets. Example output is in Fig. 2.

4 CONCLUSION

The new framework introduced in this article enables easy parallelization when testing and comparing ML methods. It has also prepared methods that allows easily process results, apply statistical tests and present results in form of tables and graphs using a web interface. Access through web interface allows to access from remote computer so it is possible to have one continuously running computer and check progress and results of testing from different machines in different places. Web interface is also most cross-platform way of creating GUI.

Designed framework is currently used for developing new ML method as described in previous chapter. Remote machines used for computations are currently heterogeneous set of different available

Sumary table for best result			
-		Method	
Model	Dataset Name	IWFD	Gram. Evol. IWFD
K1	numerals	0.984(18)	0.973(20)
	sensvehicle	0.709(16)	0.689(15)
	spectro	0.882(16)	0.859(20)
	isolet	0.738(20)	0.847(20)
	optdig	0.965(20)	0.937(20)
	scene	0.522(8)	0.639(20)
	satelite	0.893(20)	0.853(17)
AMLP	numerals	0.971(15)	0.966(19)
	sensvehicle	0.711(5)	0.590(3)
	spectro	0.853(10)	0.878(18)
	isolet	0.759(20)	0.827(14)
	optdig	0.895(20)	0.867(20)
	scene	0.582(8)	0.673(20)
	satelite	0.797(3)	0.830(2)
	numerals	0.971(15)	0.966(19)
	sensvehicle	0.711(5)	0.590(3)
	spectro	0.853(10)	0.878(18)

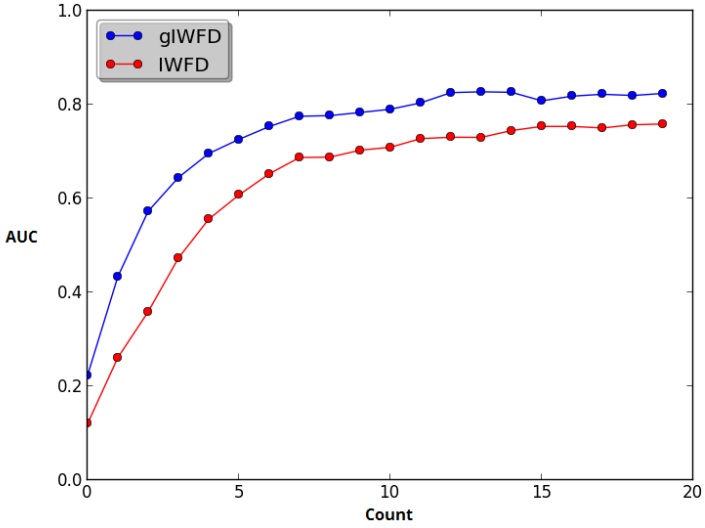


Figure 2: Example of web interface output comparing accuracy of two methods using table and graph.

computers. This state is far from ideal for further testing it will be necessary to replace these computers with higher number of (ideally) same computers. Therefore it is desirable to automate preparation of system for these computers. Ideal solution will be booting these computers from a network because it allows to easily prepare one central image with a preconfigured operating system for all these computers. Second advantage of this approach is that it does not require any changes to settings of current OS, meaning that these computers can be easily used for their original purpose just by rebooting from their hard drive.

Acknowledgement: The completion of this paper was made possible by the grant No. FEKT-S-14-2429 - „The research of new control methods, measurement procedures and intelligent instruments in automation" financially supported by the Internal science fund of Brno University of Technology.

REFERENCES

[1] Thomas G Dietterich. Approximate statistical tests for comparing supervised classification learning algorithms. *Neural computation*, 10(7):1895–1923, 1998.

[2] Ron Kohavi et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, volume 14, pages 1137–1145, 1995.

[3] EC2 Instances Pricing. `https://aws.amazon.com/ec2/pricing/`. Accessed: 2016-03-15.

FMUTOOLBOX FOR MATLAB/SIMULINK

Jan Glos
Doctoral Degree Programme (1), FEEC BUT
E-mail: xglosj00@stud.feec.vutbr.cz

Supervised by: Pavel Václavek
E-mail: vaclavek@feec.vutbr.cz

Abstract: Acausal (physical) modeling can be advantageously used to construct dynamic models, which are legible and easily modifiable for user. Even highly complex models can be easily assembled and the model remains clear. The most widespread language for acausal modeling is Modelica. Unfortunately Matlab/Simulink does not support Modelica language and that is why FMUtoolbox for Matlab/Simulink was established. It adds possibility to simulate Functional Mock-Up Units (exported Modelica models) to Matlab/Simulink environment. FMUtoolbox was practically proved during verification of control algorithms for stepper motor and in the design of heat pump control.

Keywords: Modelica, Functional Mock-up Interface, FMI, Functional Mock-up Unit, FMU, Matlab, Simulink, Model Exchange, Co-Simulation

1 INTRODUCTION

Modelica is an object-oriented language designed for modeling heterogeneous physical systems [1]. It is intended for multi-domain modeling and aspires to be general-purpose modeling tool [2]. It provides large amount of free libraries for modeling and simulation of electrical, magnetic, mechanic, thermal and other physical systems. Moreover there are lots of commercial libraries for special purposes (like refrigeration, engines, batteries, pneumatics etc.). The user can also create own library using existing models or create completely new one with own developed models, which can be build using equations or by connecting existing objects.

Modelica language uses acausal modeling, which brings significant advantages for user. Compared to causal description (like in Simulink using integrators, gains etc.) the acausal approach provides better legibility of model, resulting in less chance of mistakes [3]. Also the model can be easily modified or extended.

Functional Mock-up Interface (FMI) describes standardized way to exchange models between modeling and simulation tools [4]. It is supported by most of Modelica tools (Dymola, OpenModelica etc.). Matlab/Simulink supports neither Modelica language nor this standard and that's why FMUtoolbox for Matlab/Simulink was established. It adds functionality for FMU simulation into Matlab/Simulink environment, what brings the possibility of using Modelica models in Matlab/Simulink.

Very first version of FMUtoolbox was presented in my previous article [5]. This paper describes the second version of the toolbox, whose most important improvements include support of FMI 2.0, support of Co-Simulation, notable simulation speed-up, native Matlab installer and others.

2 MODELICA LANGUAGE

This chapter briefly describes the Modelica language, in depth as needed in the following chapters. More details about Modelica language can be found in [1], [6], [7], [8] and others.

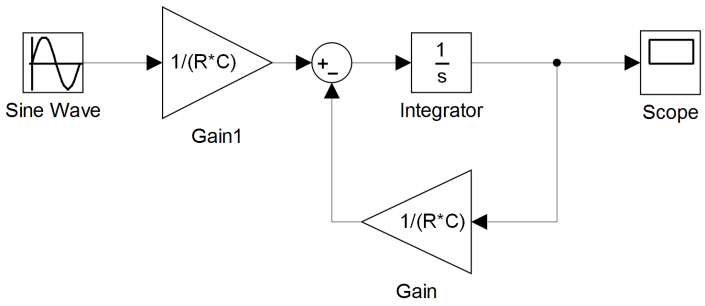
Every Modelica model consists of three basic sections. The first of them is placed behind keyword `model` (with the model name) and contains model variables, inheritance and others. The second section of the model contains initial equations, which are used to compute initial conditions for state variables. The third part of the model includes model equations and is prefaced by keyword `equation`. There can be both the real equations (like $a+b-c=0$) and connections of blocks (like `connect(objA.output, objB.input)`). The model is terminated with keyword `end` with the name of the model.

A basic example of simple Modelica model is presented below.

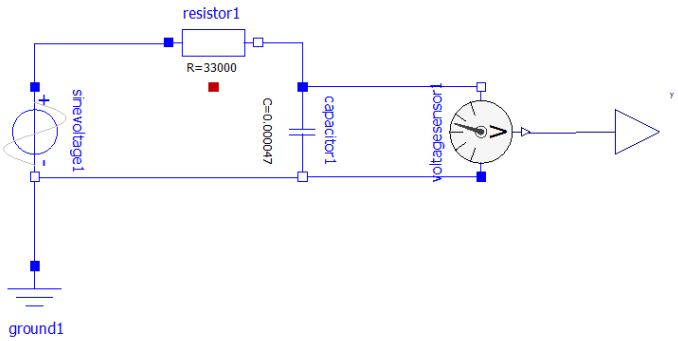
```
model simpleExample
    Real x "State variable";
    parameter Real x0 = 5 "initial value of x";
initial equation
    x = x0;
equation
    0 = der(x) + 3*x;
end simpleExample;
```

The name of the model is `simpleExample` and using parameter `x0` and initial equation the start value of `x` can be set.

Figures 1a and 1b illustrate the main difference between causal and acausal modeling, especially Simulink and Modelica. Simulink model in Fig. 1a is strictly based on differential equations and it is not very legible for user. Conversely the model in Modelica (Fig. 1b) is based on objects and their connections, so the model is legible and it can reflect the reality.



(a) in Simulink



(b) in Modelica

Figure 1: RC circuit model

3 FMI 2.0

FMI 2.0 is the second major version of the standard defined to normalize Model Exchange and Co-Simulation between Modelica (and other) modeling and simulation tools [4]. In this version standards for Model Exchange and Co-Simulation were merged, which simplifies the solution needed for handling model import, export and Co-Simulation.

Almost every Modelica tool is capable to export a model into Functional Mock-up Unit (FMU), a special file type consisting of model description in XML file and model equations as source code or binary form (dynamic-link library or shared object). There might be additional files packed into FMU (like tables, maps and images). In fact, FMU is a folder with defined structure, which is packed into zip and then renamed to have an extension `.fmu`.

In comparison with the previous version multiple changes were performed. As mentioned above, Model Exchange and Co-Simulation are now based on a common core and only necessary parts are different. Regarding model description clarification of some terms was accomplished (especially model variable properties `causality` and `variability`). Further an item `ModelStructure` was added, which describes model outputs, derivatives and initial unknown. It is an advantageous improvement, for example in version 1.0 it was needed to find outputs between all model variables (sometimes thousands and more). Moreover items describing FMU capabilities were added (`ModelExchange` and `CoSimulation`). Regarding Application Programming Interface (API) several improvements were made. For example possibility to reset FMU can be mentioned, hereafter logging categories and other improvements.

4 REALIZATION OF FMUTOOLBOX

FMUtoolbox is a tool allowing simulation of Modelica models (exported into FMU format) in Matlab/Simulink environment. This tool was created by me and is continuously improved and perfected.

New version of FMUtoolbox had to be created due to new version of FMI standard, performance issues and because of the need for Co-Simulation support. FMUtoolbox 1.0 was created previously and described in my last article [5]. The first version provided basic FMU simulation capabilities for Matlab/Simulink.

The core of FMUtoolbox 2.0 is derived from the previous version. It is based on Matlab class, which provides methods and data storage for single FMU. Class methods are responsible for loading FMU into Matlab, loading XML file with model description, partly for user interface and others.

FMUtoolbox 2.0 provides user interface in two ways - graphical user interface (GUI) for FMU usage in Simulink and command line interface for simulation using Matlab command line.

The most important changes were performed on Simulink library, blocks and S-function. Originally Level 2 Matlab S-Function was used to provide the interface between Matlab and FMU (additionally there was a dynamic-link library or shared object between Matlab and FMU to provide needed functionality). Actual solution employs Level 2 C S-Function as the interface between Matlab and FMU. No other DLL or SO is necessary, all needed functionality is formed inside C S-function (or is included into it). This solution brought significant simulation speed-up and simplification of communication between Matlab/Simulink and FMU. Actually there are four C S-Functions created (Fig. 2), two of them are for FMI 1.0 (the first one for Model Exchange, the second for Co-Simulation) and two for FMI 2.0 (the first one for Model Exchange, the second for Co-Simulation). This division ensures the possibility to simulate any combination of FMUs concurrently - for example several the same FMUs with different parameters, one Model Exchange FMU 1.0 with Co-Simulation FMU 2.0 and so on.

One of the most important improvement is added support for Co-Simulation. In some cases Co-

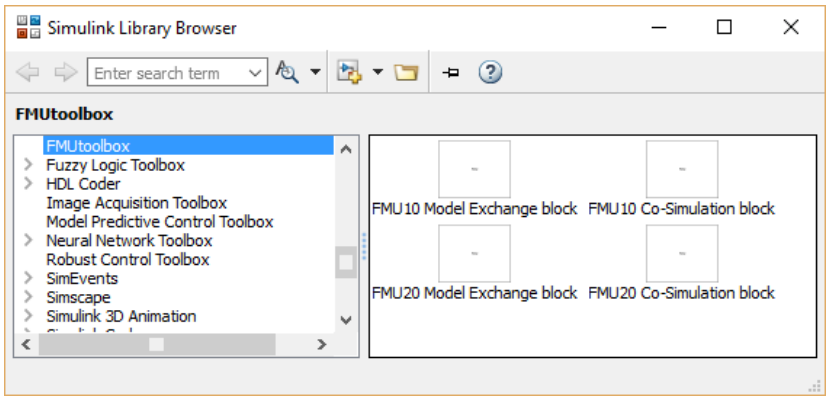


Figure 2: FMUtoolbox library for Simulink

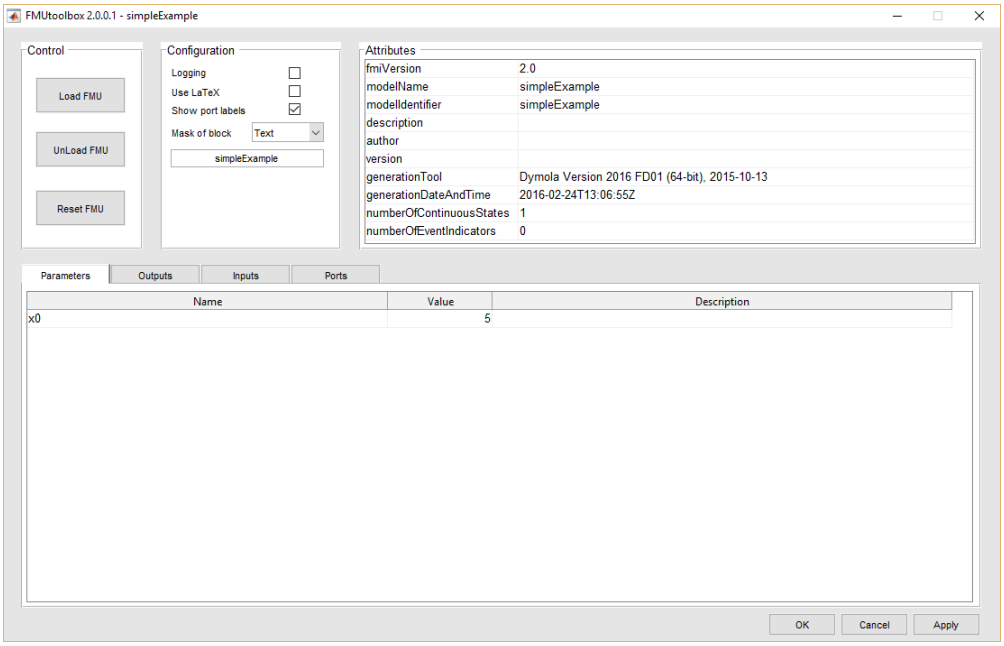


Figure 3: FMUtoolbox GUI

Simulation is more suitable as it can bring simulation speed-up or increased accuracy. There are two possible ways to perform Co-Simulation - communication between different tools and packing solver into FMU. The second option was chosen because of the compatibility reasons and because the communication between tools is not addressed in FMI 2.0 standard. It was found that for some models Dymola solvers are more appropriate than Matlab solvers - in those cases Co-Simulation is very useful.

Another important change covers installation of FMUtoolbox. In previous version the installers were prepared for each platform (Windows, Linux, OS X) separately according to different possibilities of these platforms. For FMUtoolbox 2.0, standard shared installation procedure for all platforms was introduced. It uses Matlab Add-Ons to install FMUtoolbox.

Regarding graphical user interface (GUI) some performance improvements were realized and the GUI was adjusted for FMI 2.0 usage. In Fig. 3 GUI with loaded FMU is shown - it is an example Modelica model from section 2 exported into FMU and then loaded into Matlab. The value of parameter `x0` can be adjusted before simulating the model.

5 CONCLUSION

In this paper FMUtoolbox for Matlab/Simulink was described, especially the second version of this tool. This version brought support of FMI 2.0, Co-Simulation and some other features. Main differences between version FMI 2.0 and FMI 1.0 were described as well as differences between FMUtoolbox 1.0 and 2.0.

Most importantly, the toolbox has been verified in practical use and it was found satisfactory for the validation of control algorithms. The first successful use was in the design of control loop for hybrid stepper motor (HSM). The accurate model of HSM was constructed in Modelica language and then it was exported into FMU. Then the model was loaded into Matlab/Simulink using FMUtoolbox and the testing of control algorithms could be done. The model contained both the HSM and the power circuit, so the simulations results were similar to behaviour of the real system.

FMUtoolbox was also used during designing heat pump control. As usually the model of heat pump was formed using Modelica language and then it was loaded into Matlab/Simulink for further use. In this case co-simulation was advantageously employed, because Dymola solvers were found to be much more faster and accurate when used with heat pump Modelica model. Moreover it is almost impossible to construct a heat pump model in Matlab/Simulink environment and conversely Dymola is not so suitable for the design of control algorithms as Matlab.

According to the above examples it is evident that FMUtoolbox is a suitable tool for Modelica models simulation in Matlab/Simulink environment.

ACKNOWLEDGEMENT

The completion of this paper was made possible by the grant No. FEKT-S-14-2429 - „The research of new control methods, measurement procedures and intelligent instruments in automation" financially supported by the Internal science fund of Brno University of Technology.

REFERENCES

[1] Modelica Association, "Modelica - a unified object-oriented language for systems modeling," 2014.

[2] S. E. Mattsson, H. Elmqvist, and M. Otter, "Physical system modeling with Modelica," *Control Engineering Practice*, vol. 6, no. 4, pp. 501 – 510, 1998.

[3] J. Kofránek, M. Mateják, P. Privitzer, and M. Tribula, "Causal or acausal modelling: Labour for humans or labour for machines," *Technical Computing Prague*, 2008.

[4] MODELISAR consortium and Modelica Association Project "FMI", "Functional Mock-up interface for Model Exchange and Co-Simulation," 2014.

[5] J. Glos, "Využití modelů v jazyce Modelica v prostředí Matlab-Simulink," in *Proceedings of the 21st Conference STUDENT EEICT 2015*, (Brno), pp. 310–312, Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2015.

[6] P. A. Fritzson, *Principles of object-oriented modeling and simulation with Modelica 2.1*. New York: Wiley-Interscience, c2004.

[7] M. Tiller, *Modelica by Example*. 2014.

[8] M. Tiller, *Introduction to physical modeling with Modelica*. Boston: Kluwer Academic Publishers, c2001.

IMAGE PROCESSING ALGORITHMS VISUALISER

Daniel Davídek
Doctoral Degree Programme (1), FEEC BUT
E-mail: davidek@feec.vutbr.cz

Supervised by: Karel Horák
E-mail: horak@feec.vutbr.cz

Abstract: A highly visual way to understand and develop computer vision algorithms is proposed. Image processing realization commonly sustains of successive function calls over the output image from the previous function. In the later phases (i.e. contour analysis, operations over segments) when the output argument of a function is not always an image, there is often a way to visualize the arguments by an image. The developed program described in this article is capable of executing basic computer vision algorithms over the input image sequentially and visualize individual steps outputs as image queue. The initialization of functions chain is possible with simple text input of individual function names with an easy option to program new algorithms.

Keywords: computer vision, visualisation

1 PROLOGUE

From the time when the computer could load a digitalized image, there have been sustained development in the field of computer vision algorithms and image processing software. One of nowadays examples of continuously developed proprietary CV library is the Image Processing Toolbox (IPT) from MathWorks®. It is one of many other toolboxes from MATLAB® environment. The MATLAB environment has its high-level programming language which is consistent across the toolboxes. In spite of its higher price, it is commonly used in many different subjects in schools and universities all over the world.

The Image Processing Toolbox capabilities are also suitable for education of CV, as the time-demanding functions are precompiled. Nevertheless, the MATLAB environment is not commonly used for out-of-school projects.

When you want to run the program, you either have the MATLAB environment to interpret the program code on the fly or to compile the script into a binary file. The second option is burdened with the need of buying MATLAB Compiler™. The compiled program also starts longer than a program in some other high-level programming language even if it contains the same simple algorithms. The overhead is caused by starting the MCR[1]. Once the MCR interpreter is running, the code should execute as fast as in MATLAB environment although its execution can be slower than other lower-level language programs.

Another framework for CV algorithm development is HALCON from MVTec. It has its IDE with the possibility to program in the textual and graphical interface. It also has interfaces with other languages e.g. C, C++ and C#. Thus, HALCON programming environment is for free only for a limited period; then the user must buy a license that comes again with a higher price.

If we want to teach students something they can use in their projects, we should consider another CV library. From the year 2000, there is a continuous development of open-source CV library written in

[1] MATLAB Runtime

C++ named OpenCV. Its algorithms can be called from within a C or C++ program directly, and there are wrappers to use it within high-level programming language Python.

There is an abundance of projects using OpenCV library. Not only in technical, but also in art and other disciplines. This huge community can also contribute to OpenCV code not only through bug fixing but also by developing brand new CV algorithm[2]. These conditions create a positive feedback loop which guarantees OpenCV robustness and further development.

Currently, there is a stable OpenCV 3.1 (as of 21-12-2015) for which the Python 3.4 wrappers exists. The OpenCV library is also multiplatform and can be used on Windows, Linux, Mac and also on Android and iOS which further scales its applicability.

2 INTRODUCTION

This article discusses IPAV[3] a program with two main goals. Firstly it serves as a CV algorithms visualiser processing an image or video stream[4]. This function can be used as a teaching aid in CV learning classes as a visual representation of individual steps of CV algorithm i.e. resizing, grayscaling, filtering, etc. It also can help during the development of such chain of steps which creates the desired CV algorithm during the design phase.

Secondly, it can help the students to learn how to use the OpenCV library from Python. They can build upon the IPAV to develop own CV algorithm chains, and it smoothens the learning curve of OpenCV as it is already bundled with GUI library Kivy. Similarly as with the MATLAB IPT, this solution also enables rapid application development with similar execution times. The higher-level Python language helps not to sink inevitably into the lower-level memory management and lets developers contemplate about the CV algorithms. Besides, the program developed with the OpenCV library can be released with BSD License[2].

3 PROGRAM FUNCTION

3.1 PROGRAM STRUCTURE

The program IPAV is written in Python 3.4 and OpenCV 3.0. The GUI is created through Kivy multiplatform open source UI framework [4].

Separate threads are used for GUI, CV calculations, and real-time video stream capture to ensure stable and high-speed performance.

3.2 ALGORITHM CHAINS AND STEP METHODS

Every high-level CV algorithm is implemented as a chain of individual CV step methods. This way IPAV deploys conventional CV algorithms as they are traditionally made of parts e.g. preprocessing, segmentation, feature extraction, and classification.

Currently, the algorithm chain can only form a topological daisy chain; multiple children nodes are not possible. The chain is defined by a string which contains individual names of its step methods. Explicitly it is a step name strings list. The chain selection then ensures successive calling of defined step methods over the image. All applicable step methods are defined and initialized at the program start to be ready whenever needed.

[2] e.g. ORB (Oriented FAST and Rotated BRIEF) which was created by OpenCV Labs in 2011 [3]

[3] IPAV– Image processing algorithms visualiser

[4] video stream from a file or a connected camera

3.3 FRONTEND

Main program window has a control panel on the left to enable video stream switching and toggling of capturing and computing. There is also a slider to control input image resolution division and multiplication and a camera preview.

In the main section on the right, there is step widgets daisy chain graphical representation. Each step method is visualized by its name, output image and an average time of execution. There is a possibility to add an interactive text overlay to each step separately to show more information on demand. This will be connected later with inline python snippet creation to let the user show any relevant data i.e. histogram analysis, morphological and other statistics over an image and step data.

Currently, only OpenCV openable video stream input is supported, e.g. USB camera, ethernet camera, etc. The output is showed only in the widgets.

4 EXAMPLE

4.1 EDGE DETECTORS

There are many algorithms for edge detection implemented in OpenCV library, and there can be a step implementation created for every one of them. The algorithm chains of Laplacian and Sobel (vertical) edge filters with 5x5 kernels are shown in following figures. Laplacian on figure 2 on the following page, Sobel vertical on figure 1. More on the topic of edge filtering can be found for example in Medical Image Processing book[1].

The whole chain consist of preprocessing steps which perform – original, resizing, grayscaling, thresholding and separate filtering step method. The widgets are displayed in two columns sorted in rows with a retrospective order (later-first is the default view).

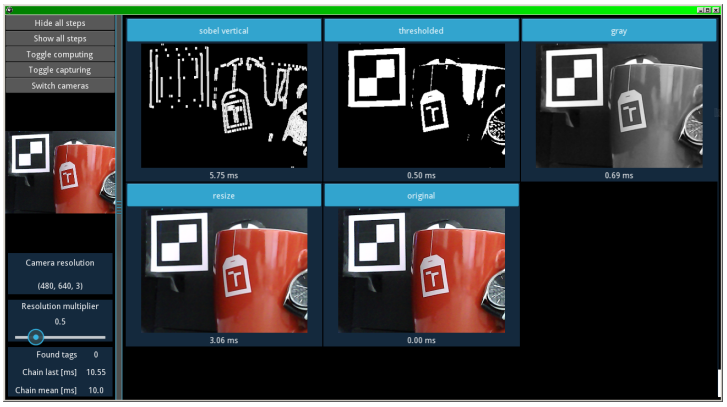


Figure 1: Sobel filtering chain example (5x5 kernel S_x)

4.2 MARKUP-BASED VISION

In the following example, the algorithm chain is designed to find the markup square in the image and calculate its homography transformation. Individual step methods are shown retrospectively. They chronologically consist of – original, resized, grayscaled, inverted threshold, removing a frame, flooded with black, white flooded, inverted and findTags algorithm.

More output from the findTags algorithm is shown in the right column. The green background represents successful match, and the text overlay shows valid and false tag detection counts. Flawlessly located and back-projected square tags are displayed in the bottom-right corner.

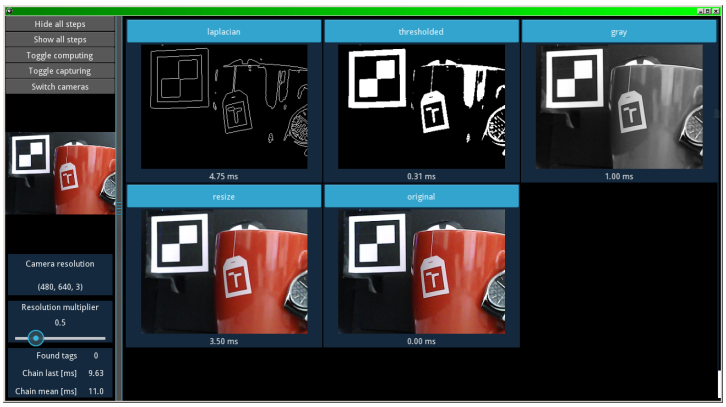


Figure 2: Laplacian filtering chain example (5x5 kernel)

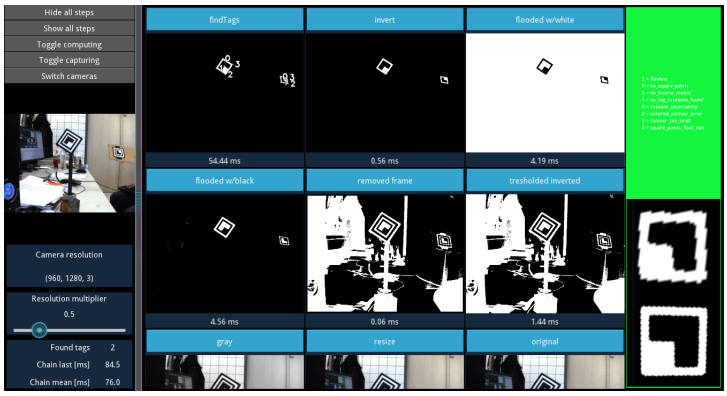


Figure 3: Markup-tag homography calculation algorithm chain example

4.3 Design flow

To create a new chain of steps user can insert string of comma separated method names after clicking Load chain button as seen in figure 4 on the following page. The list of all available step method names is also accessible through the popup window.

5 APPLICATION

The IPAV was used to develop the algorithm for markup-based vision depicted in the second example section 4.2 on the previous page. This example also demonstrates the intrinsical simplicity of understanding CV algorithm flow when the user deals with visual data.

6 CONCLUSION

It was discussed that a highly visual way of representing CV algorithms in the IPAV can positively affect the simple CV algorithm development time. Furthermore, IPAV delivers simple visual data of higher-level CV algorithm which can lead to the less steep learning curve of CV methods.

Lastly, the program can be a stepping stone for any other project as it enables programming with OpenCV library in higher-level Python language with robust Kivy UI framework. The project source

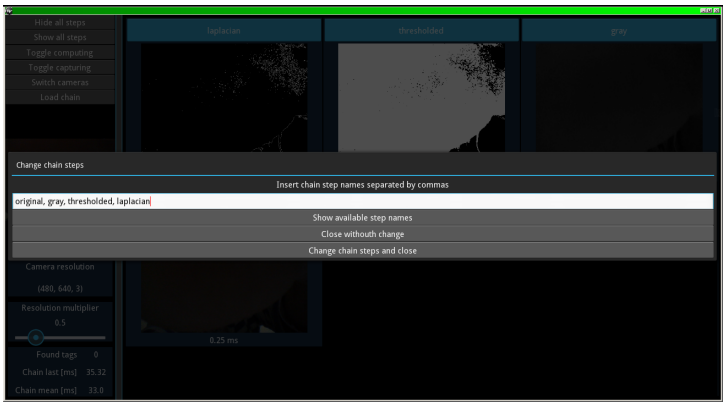


Figure 4: Insertion of new step chain string

code is currently hosted on git repository at bitbucket[5]

7 FUTURE WORK

In the next development cycle; there is planned to add a textual interface with sentence comprehension. Continuous algorithm chain creation with the step name suggestion by tabulator key. The following step would be to save and load the text string during execution.

As it was said, currently the algorithm chain can only form a topological daisy chain. Multiple output nodes for any step method would account for further simplicity in understanding different CV step methods. How to effectively compute and synchronize forked chains by multi-thread step method calling is one of the implementation questions which arises from non-daisy chain structure.

The innovation of GUI elements and widgets is also planned. The introduction of ROI zoom to individual step widgets and fullscreen preview for selected step method as well as saving of individual step outputs as an image or video stream are proposed. Also, the inline creation of Python snippets over individual step text and image data will be supported.

ACKNOWLEDGEMENT

The completion of this paper was made possible by the grant No. FEKT-S-14-2429 - „The research of new control methods, measurement procedures and intelligent instruments in automation" financially supported by the Internal science fund of Brno University of Technology.

REFERENCES

[1] Jan, J., Medical Image Processing, Reconstruction and Restoration – Concepts and Methods, 2006, s. 601-604, ISBN 0-8247-5849-8

[2] Itseez, OpenCV 3-clause BSD License, url: `http://opencv.org/license.html`

[3] Itseez, OpenCV Labs, ORB, 2011, `http://docs.opencv.org/3.0-beta/doc/py_tutorials/py_feature2d/py_orb/py_orb.html#theory`

[4] kivy.org, Kivy UI framework, `https://kivy.org/`

[5] `https://bitbucket.org/gr4viton/kivycv`

MODERN METHODS FOR IMAGE QUALITY ASSESSMENT

Petr Nováček

Doctoral Degree Programme (1), FEEC BUT

E-mail: novacekp@feec.vutbr.cz

Supervised by: Karel Horák

E-mail: horak@feec.vutbr.cz

Abstract: This paper deals with the comparison of Image Quality Assessment (IQA) algorithms. The test gallery of colour calibre images were captured for comparison of IQA algorithms. Test gallery was captured by two cameras with a different type of sensors – CMOS with Bayer mask and a Foveon sensor.

Keywords: Image Quality Assessment, Mean Squared Error, Peak Signal to Noise Ratio, Universal Quality Index, Structural Similarity Index, Edge Strength Similarity

1. INTRODUCTION

Image Quality Assessment (IQA) algorithms are very useful instrument to compare image quality. This paper is devoted to the comparison of IQA algorithms which are focused on structural and brightness differencies between images. Algorithms are tested on the real problem which is a comparison of two camera sensors with optical colour calibre. The first camera is Nikon D3100 with CMOS sensor and the second camera is Sigma DP2 Merill, which has Foveon sensor. Test images of these two cameras were compared to a synthetic image of an optical calibre which has defined colours in RGB space.

2. IMAGE QUALITY ASSESSMENT ALGORITHMS

Image Quality Assessment (IQA) algorithms [1] can be divided into three parts according to the use of reference image. The first parts of these algorithms are algorithms which don't use reference image (No-Reference IQA - NR-IQA). The second part is a group of algorithms that using part of the reference image or extracted features of the reference image and third part are algorithms that use the whole reference image. This article is focused on that part of IQA algorithm which uses the whole reference image.

2.1. MEAN SQUARED ERROR (MSE)

This algorithm [1] computes mean squared error of the tested image with the reference image. Very easy computation is an advantage of this algorithm but there is a slightly similarity with human visual system assessment of image quality. The MSE is calculated as (1)

$$MSE(f,g) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\left[f(x,y) - g(x,y)\right]^2 \qquad (1)$$

where *f(x,y)* is tested image, *g(x,y)* is the reference image, *x* is an index of pixel row, *y* is an index of pixel column and *MxN* is the size of both images. Lower result value means that the images are more similar.

2.2. PEAK SIGNAL TO NOISE RATIO (PSNR)

The PSNR algorithm [1] is principally similar to MSE algorithm with some extension. PSNR is calculated as (2) or as (3)

$$PSNR(f,g) = 10\log_{10}\left(\frac{\left(2^B-1\right)^2}{\frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\left[f(x,y)-g(x,y)\right]^2}\right) \quad (2)$$

$$PSNR(f,g) = 10\log_{10}\left(\frac{\left(2^B-1\right)^2}{MSE(f,g)}\right) \quad (3)$$

where B is the bit depth of images. The result value can acquire values ≥ 0 where 0 means that images are the same. Because of the PSNR algorithm is based on MSE this algorithm is slightly similar to human visual system assessment of image quality too.

Higher result value means that the images are more similar.

2.3. UNIVERSAL QUALITY INDEX (UQI)

The Universal Quality Index [2] measures change (decrease) of correlation between reference and tested images. UQI can be calculated as (4)

$$UQI(f,g) = \frac{4\sigma_{fg}\mu_f\mu_g}{\left(\sigma_f^2+\sigma_g^2\right)\left(\mu_f^2+\mu_g^2\right)} \quad (4)$$

where μ_f and μ_g are mean luminance of tested and reference images, σ_f and σ_g are standard deviations of these images and σ_{fg} is cross covariance between both images calculated as (5).

$$\sigma_{fg} = \frac{1}{N-1}\sum_{i=1}^{N}\left(f_i-\mu_f\right)\left(g_i-\mu_g\right) \quad (5)$$

UQI can acquire values between -1 and 1, where 1 is value which means that tested and reference images are the same.

This algorithm is in comparison with MSE or PSNR more similar to human visual system assessment of image quality.

2.4. STRUCTURAL SIMILARITY INDEX (SSIM)

The human visual system is highly adapted to obtain information from structural differences in image. The SSIM algorithm [1] striving to simulate this fact and assesses dependence between values of pixels in image, where can be found dependence on values which lies in their neighbourhood. The SSIM algorithm compares luminance, contrast and structural components of tested and reference images. This algorithm can be calculated as (6)

$$SSIM(f,g) = \left[l(f,g)\right]^{\alpha}\cdot\left[c(f,g)\right]^{\beta}\cdot\left[s(f,g)\right]^{\gamma} \quad (6)$$

where $l(f,g)$ is luminance component, calculated as (7), $c(f,g)$ is contrast component, calculated as (8), $s(f,g)$ is structural component, calculated as (9), f is tested image, g is reference image and $\alpha>0$, $\beta>0$, $\gamma>0$ are constants, which means relative importance of these components.

$$l(f,g) = \frac{2\mu_f\mu_g+C_1}{\mu_f^2+\mu_g^2+C_1} \quad (7)$$

$$c(f,g) = \frac{2\sigma_f\sigma_g+C_2}{\sigma_f^2+\sigma_g^2+C_2} \quad (8)$$

$$s(f,g) = \frac{\sigma_{fg} + C_3}{\sigma_f \sigma_g + C_3} \tag{9}$$

The parameters μ_f and μ_g are mean luminance of tested and reference images, σ_f and σ_g are standard deviations of these images and σ_{fg} is cross covariance between both images calculated as (5). The constants C_1, C_2 and C_3 can be used to adjustment of equations in case those are expressions $\mu_f^2 + \mu_g^2$, $\sigma_f^2 + \sigma_g^2$ or $\sigma_f \sigma_g$ too close to 0 and are calculated as $C_1 = (K_1 L)^2$, $C_2 = (K_2 L)^2$ and $C_3 = \frac{C_2}{2}$, where $K_1 << 1$, $K_2 << 1$ and L is dynamic range of one pixel (e.g. for 8bit grayscale image is $L = 255$).The result value of this algorithm is between 0 and 1, where value 1 means that images are the same.

2.5. EDGE STRENGTH SIMILARITY (ESSIM)

The ESSIM index [1] is based on fact that human visual system is more sensitive to direction which is in image represented by stronger edges. We can use e.g. high-pass filter to boost edges in the image and then extract edges with use of gradient operator (Sobel, Prewitt, etc.). Edge detection is used as horizontal and vertical and as diagonal detection. Higher value of these two is used as edge strength for every pixel. The ESSIM index is then calculated as (10)

$$ESSIM(f,g) = \frac{1}{N} \sum_{i=1}^{N} \frac{2E(f,i)E(g,i) + C}{(E(f,i))^2 + (E(g,i))^2 + C} \tag{10}$$

where f is tested image, g is the reference image, N is number of pixels in every image, $E(f,i)$ is edge strength of i-th pixel in tested image and $E(g,i)$ edge strength of i-th pixel in reference image. The parameter C is scale and is calculated as $C = (BL)^2$, where B is constant and L is dynamic range of edge strength.

The IQA based on ESSIM is very similar to human visual system assessment of image quality as well as SSIM. The result values can acquire same values as SSIM too.

3. COMPARISON OF IQA ALGORITHMS

Tested IQA algorithms have been compared on a real example problem. Images of colour optical calibre captured by Nikon D3100 (CMOS sensor with Baer mask) has been compared with images captured by Sigma DP2 Merill (Foveon sensor). The synthetic image of defined RGB values was used as the reference image for comparison of these test galleries.

The monolithic colour area (defined RGB code 255;171;14) was used in example results in this paper as reference image which has been compared with original images taken by two tested cameras.

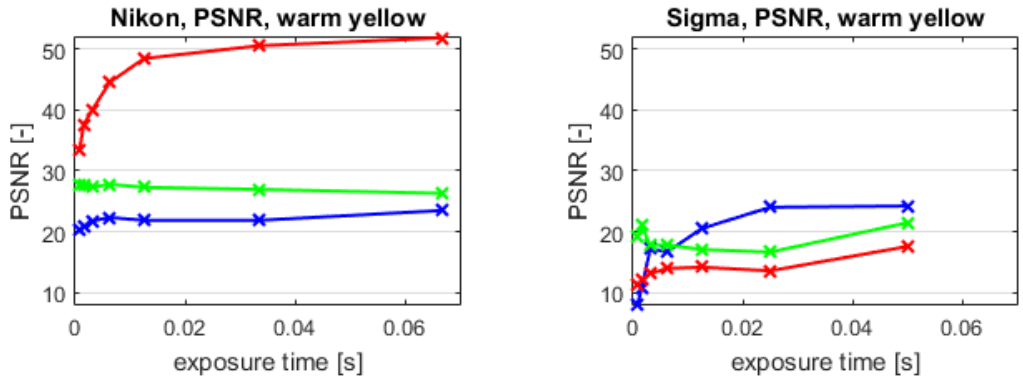
3.1. RESULTS

The output of MSE algorithm can be in some cases much higher than at other test images. The PSNR algorithm has the MSE output in a denominator. This is a reason why the extremely high MSE output values causes very small output values (Graph 1) of the PSNR algorithm. Thanks to this extremely changes, the outputs of these two algorithms have a very small dispersion of values.

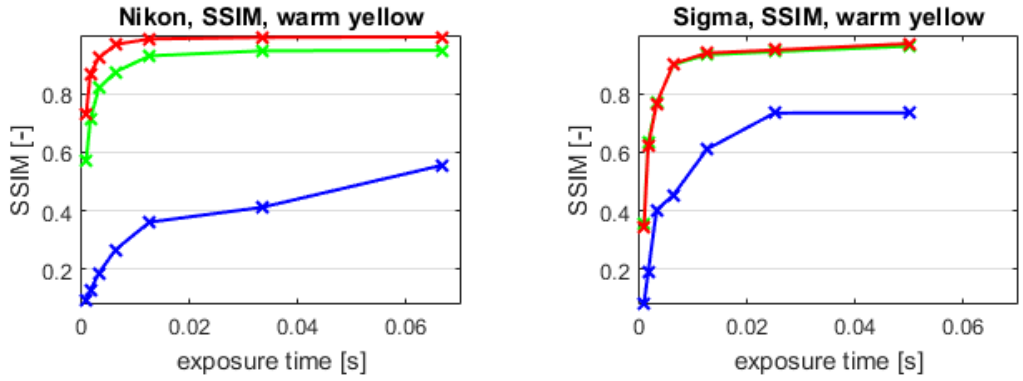
The UQI algorithm is very similar to the SSIM algorithm but has no correction constants. Because of that, this algorithm fails on images with mean luminance or standard deviations which are very close to zero or are zero. This is the case of one the tested synthetic image. The UQI algorithm returns only zero for calculations with this images.

The SSIM algorithm has against to UQI correction constants C_x and can fix problems with mean luminance or standard deviations which are very close to zero. Because of that output values of SSIM (Graph 2) are more stable and allow comparison of test images where one of which is the synthetic image.

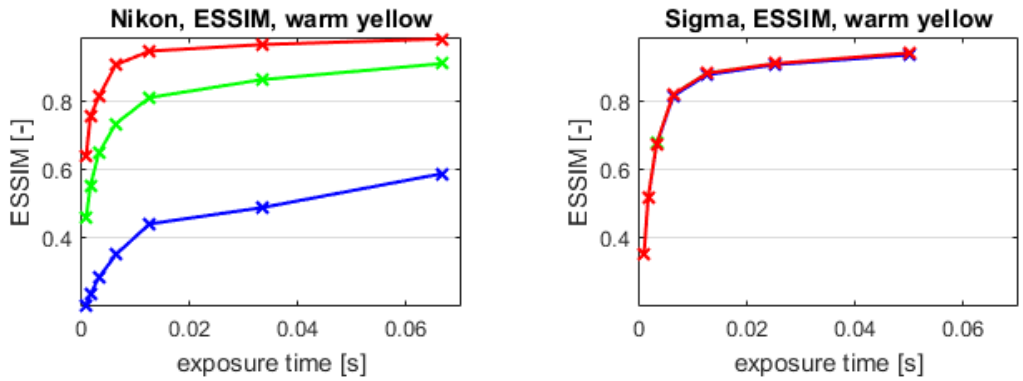
Algorithm of ESSIM type uses edge detectors to calculate the output. This fact can cause problems of distinguishing between areas in the test images that have monolithic colours or of synthetic images. The output values of ESSIM (Graph 3) are very similar to SSIM output in other causes.



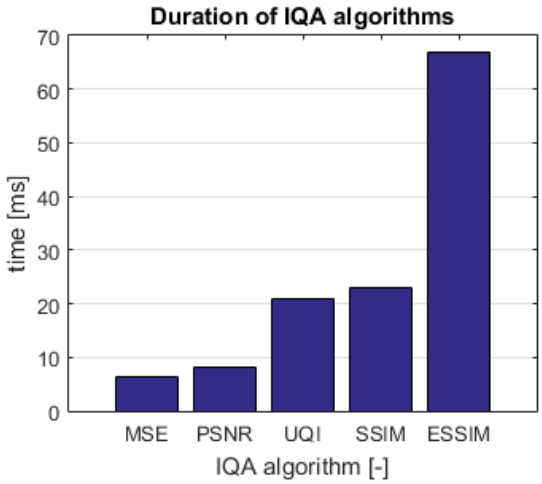
Graph 1: A graph of image quality according to exposure time, alg. PSNR, RGB channels



Graph 2: A graph of image quality according to exposure time, alg. SSIM, RGB channels



Graph 3: A graph of image quality according to exposure time, alg. ESSIM, RGB channels



Graph 4: Duration of IQA algorithms, average of 1008 repeats, image size 400x400 px, measured on CPU Intel Core i3-330M

Graph (Graph 4) shows average duration of tested IQA algorithms. Considering outputs of tested algorithms and time duration, the SSIM algorithm seems to be the best choice for comparison of synthetic and tested image combinations.

4. CONCLUSION

The MSE and PSNR algorithms are the fastest for computation. The UQI is almost as quickly as SSIM and ESSIM algorithm is significantly slowest. The UQI and ESSIM algorithms are not suitable for assesment of synthetic images or images with monolitic colours. The MSE and PSNR algorithms may have very small dipersion of output values so they are advantageous for simplier applications that are focused on speed. The SSIM algorithm has been chosen as the best solution for quality assessment of test image gallery. Duration of calculation and outputs of the SSIM algorithm had best results for my example problem.

Future work will be focused on No-Reference IQA algorithms and both of these types will be used as fitness functions of evolutionary algorithms in my dissertation thesis.

ACKNOWLEDGEMENT

The completion of this paper was made possible by the grant No. FEKT-S-14-2429 - „The research of new control methods, measurement procedures and intelligent instruments in automation" financially supported by the Internal science fund of Brno University of Technology.

REFERENCES

[1] JOY, K. R. GOPALAKRISHNA SARMA, E. Recent Developments in Image Quality Assessment Algorithms. Journal of Theoretical and Applied Information Technology, July 2014. Vol. 65 No.1. ISSN 1992-8645., cit. 7. 3. 2016. Accessible from: http://www.jatit.org/volumes/Vol65No1/20Vol65No1.pdf

[2] WANG, Z. BOVIK, C. A. A Universal Image Quality Index. IEEE Signal Processing Letters, March 2002., cit. 7. 3. 2016. Accessible from: https://ece.uwaterloo.ca/~z70wang/publications/quality_2c.pdf

[3] WANG, Z. BOVIK, C. A. SHEIKH, H. R. SIMONCELLI, E. P. Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE Transactions of Image Processing, Vol. 13, No. 4, April 2004., cit. 30. 3. 2016. Accessible from: http://www.cns.nyu.edu/pub/lcv/wang03-preprint.pdf

POWER OPTIMIZED CONTROL OF SYNCHRONOUS RELUCTANCE MACHINE

Zbyněk Mynář
Doctoral Degree Programme (2), FEEC BUT
E-mail: xmynar03@stud.feec.vutbr.cz

Supervised by: Pavel Václavek
E-mail: vaclavek@feec.vutbr.cz

Abstract: Synchronous reluctance motor (SynRM) is becoming more and more popular alternative of AC induction motor (ACIM) for its low cost, rugged structure and possibility of higher power efficiency. Power optimized control of SynRM is, therefore, important topics of research. This paper proposes power optimizing control algorithm incorporating active search of optimal current angle according to the MTPA criterion using the DC current injection. Results of implementation on real machine are presented.

Keywords: SynRM, MTPA, power optimization, injection

1 INTRODUCTION

The synchronous reluctance motor is an AC machine with typical harmonically distributed stator windings and rotor with significant salience of inductance in dq axis, but no other windings or permanent magnets. Main advantages of SynRM are rugged structure, "cold" rotor, low price (due to absence of permanent magnets) and synchronous operation. Thanks to the absence of rotor losses and recent progress in rotor design, the SynRM can reach higher efficiency than AC induction machine and therefore it is becoming popular in applications like fans or pumps. Because the power efficiency is important advantage of SynRM, a lot of attention has been paid to power optimizing algorithms. The Max Torque Per Ampere (MTPA), which is briefly described in section two of this article, is a very popular power optimization criterion. Various approaches were proposed for implementation of MTPA strategy to Vector Oriented Control (VOC) of both SynRM and interior permanent magnet synchronous motor (IPMSM), which are in many ways similar. Most of these algorithms rely on knowledge of machine parameters, however the inductances are difficult to obtain in case of SynRM. The active search algorithms on the other hand are parameter independent at the price of torque ripple. Algorithm proposed in this article belongs to active MTPA search class of power optimizing algorithms. The algorithm utilizes DC current injection in stator reference frame and amplitude demodulation principle to obtain MPTA operating point. Results of implementation on real SynRM are shown in section four. [1, 5]

2 POWER OPTIMAL VECTOR CONTROL OF SYNCHRONOUS RELUCTANCE MACHINE

To obtain sufficiently simple, yet precise mathematical model of SynRM, the dq synchronous reference frame rotated by angle θ_e and aligned with rotor axis with highest inductance will be considered. The core losses are going to be neglected. The stator voltage $u_{dq} = [u_d, u_q]$ can than be expressed as

$$u_d = \frac{\mathrm{d}}{\mathrm{d}t}\left[L_d(i_d, i_q) i_d\right] + R_s i_d - \omega_e L_q(i_d, i_q) i_q, \tag{1}$$

$$u_q = \frac{\mathrm{d}}{\mathrm{d}t}\left[L_q(i_d, i_q) i_q\right] + R_s i_q + \omega_e L_d(i_d, i_q) i_d, \tag{2}$$

where $i_{dq} = [i_d, i_q]$ is stator current, $L_d(i_d, i_q)$ and $L_q(i_d, i_q)$ are inductances in d and q axis, R_s is stator resistance and ω_e is synchronous electrical rotor speed. The inductances are nonlinear functions of both d and q axis current because of direct and cross-saturation effect, which is significant for SynRM. The rotor speed can be modeled as

$$\frac{\mathrm{d}\omega_e}{\mathrm{d}t} = \frac{P_p}{J}(T - T_L), \tag{3}$$

where P_p is number of pole pairs, J is moment of inertia, T_L is load torque and torque generated by machine T can be calculated as

$$T = \frac{3}{2}P_p[L_d(i_d, i_q) - L_q(i_d, i_q)]i_d i_q. \tag{4}$$

In well known VOC algorithm for SynRM, the speed and torque of machine are usually controlled by PI or PID controller. Its output is the stator current amplitude reference $I_m^{\mathrm{req}} = \sqrt{i_d^2 + i_q^2}$. The stator current references in d and q axis then can be calculated as

$$i_d^{\mathrm{req}} = I_m^{\mathrm{req}} \cos\theta_I, \tag{5}$$

$$i_q^{\mathrm{req}} = I_m^{\mathrm{req}} \sin\theta_I, \tag{6}$$

where θ_I is the stator current angle. [4]

The MTPA criterion optimizes stator current angle θ_I so the stator current amplitude I_m is minimal for given torque as shown in Figure 1. This leads to minimization of stator copper losses. If the core losses are neglected, the MTPA operating point is equal to point of maximal power efficiency. If we for simplicity consider the inductances to be constant, the optimal point of criterion can be expressed from equation (4) as

$$\frac{\mathrm{d}T}{\mathrm{d}\theta_I} = \frac{3}{4}P_p I_m^2 (L_d - L_q)\cos 2\theta_I = 0 \tag{7}$$

In this ideal case the stator current angle θ_I is always equal to $\pi/4$. In reality however, the inductances saturates heavily and the angle θ_I is always higher than $\pi/4$.[1, 2]

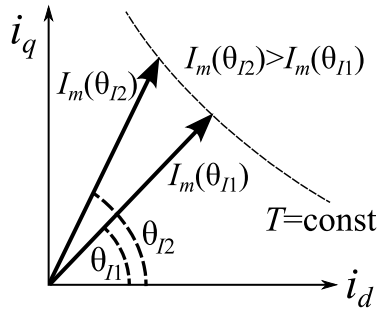


Figure 1: Max torque per ampere power optimization principle

3 PROPOSED CURRENT ANGLE CONTROL ALGORITHM

The proposed algorithm conducts active search for optimal angle θ_I using the injection of DC current vector of amplitude I_{DC} into stator coordinates. This will naturally add current vector rotating with electrical rotor speed ω_e to stator current vector in rotating dq reference frame. The situation is shown in Figure 2. The DC current can easily be created by the current PI controllers in rotating reference frame because the current control frequency bandwidth is generally larger than maximal values of ω_e. This is done by modification of stator current references (5) and (6) as

$$i_d^{\mathrm{req}} = I_m^{\mathrm{req}} \cos\theta_I + I_{DC} \sin\theta_e \cos\theta_I - I_{DC} \cos\theta_e \sin\theta_I = I_m^{\mathrm{req}} \cos\theta_I + i_d^{\mathrm{inj}}, \tag{8}$$

$$i_q^{\mathrm{req}} = I_m^{\mathrm{req}} \sin\theta_I + I_{DC} \sin\theta_e \sin\theta_I + I_{DC} \cos\theta_e \cos\theta_I = I_m^{\mathrm{req}} \sin\theta_I + i_q^{\mathrm{inj}}, \tag{9}$$

where $i_{dq}^{\text{inj}} = [i_d^{\text{inj}}, i_q^{\text{inj}}]$ is injected current vector reference. Rotation of injected current vector by angle θ_I is necessary for extraction of information on location of MTPA operating point. It should be noted, that no extra sine or cosine must be calculated to generate DC current according to (8) and (9).

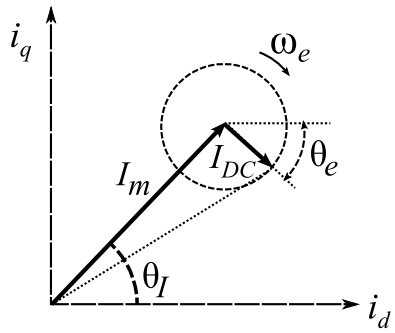


Figure 2: Rotating current vector in *dq* reference frame caused by DC current injection

To obtain information on position of optimal current angle θ_I, let us consider the voltage equation (1) and (2) without voltage drop on stator resistance R_s. For sake of simplicity the inductances will be considered to be constant for current deviation caused by injected current. This leads to

$$\begin{aligned} u_d - R_s i_d = L_d \frac{\mathrm{d}i_d}{\mathrm{d}t} - L_q \omega_e i_q = & L_d I_{DC} \omega_e \cos\theta_e \cos\theta_I + L_d I_{DC} \omega_e \sin\theta_e \sin\theta_I - L_q \omega_e I_m \sin\theta_I - \\ & L_q \omega_e I_{DC} \sin\theta_e \sin\theta_I - L_q \omega_e I_{DC} \cos\theta_e \cos\theta_I, \end{aligned} \tag{10}$$

$$\begin{aligned} u_q - R_s i_q = L_q \frac{\mathrm{d}i_q}{\mathrm{d}t} + L_d \omega_e i_d = & L_q I_{DC} \omega_e \cos\theta_e \sin\theta_I - L_q I_{DC} \omega_e \cos\theta_e \cos\theta_I + L_d \omega_e I_m \cos\theta_I + \\ & L_d \omega_e I_{DC} \sin\theta_e \cos\theta_I - L_d \omega_e I_{DC} \cos\theta_e \cos\theta_I. \end{aligned} \tag{11}$$

Use of amplitude demodulation principle on frequency ω_e on both equation (10) and (11) leads to

$$u_{d,S}^{\text{dmd}} = \text{LPF}\{(u_d - R_s i_d)\sin\theta_e\} = \frac{1}{2}(L_d - L_q)\,\omega_e I_{DC} \sin\theta_I, \tag{12}$$

$$u_{d,C}^{\text{dmd}} = \text{LPF}\{(u_d - R_s i_d)\cos\theta_e\} = \frac{1}{2}(L_d - L_q)\,\omega_e I_{DC} \cos\theta_I, \tag{13}$$

$$u_{q,S}^{\text{dmd}} = \text{LPF}\{(u_q - R_s i_q)\sin\theta_e\} = \frac{1}{2}(L_d - L_q)\,\omega_e I_{DC} \cos\theta_I, \tag{14}$$

$$u_{q,C}^{\text{dmd}} = \text{LPF}\{(u_q - R_s i_q)\cos\theta_e\} = -\frac{1}{2}(L_d - L_q)\,\omega_e I_{DC} \sin\theta_I, \tag{15}$$

where LPF stands for Low-Pass Filter. Because the frequency on injected current is not constant, the bandwidth of low-pass filters must be set below minimal expected rotor speed ω_e. To get approximate current angle θ_I information according to MTPA criterion (7), the following formula can be expressed

$$e_{\theta I} = u_{d,C}^{\text{dmd}} u_{q,S}^{\text{dmd}} + u_{d,S}^{\text{dmd}} u_{q,C}^{\text{dmd}} = \frac{1}{4}(L_d - L_q)^2\,\omega_e^2 I_{DC}^2 \cos 2\theta_I \simeq K \frac{3}{4} P_p I_m^2 (L_d - L_q)\cos 2\theta_I, \tag{16}$$

where K is general scalar constant. The equation (16) is then used as error input of PI controller, whose output is the angle θ_I. Block diagram of final proposed current angle control algorithm is in Figure 3. Found solution is unfortunately sub-optimal as the inductance difference term is in second power in equation (16). Due to general non-linear nature of inductances, the analytical expression of error between estimated and true optimal angle θ_I would be complicated. Both simulation and implementation in real conditions however showed that, thanks to relative flatness of constant torque curve around MTPA optimal point (see Figure 1), the error is acceptable.

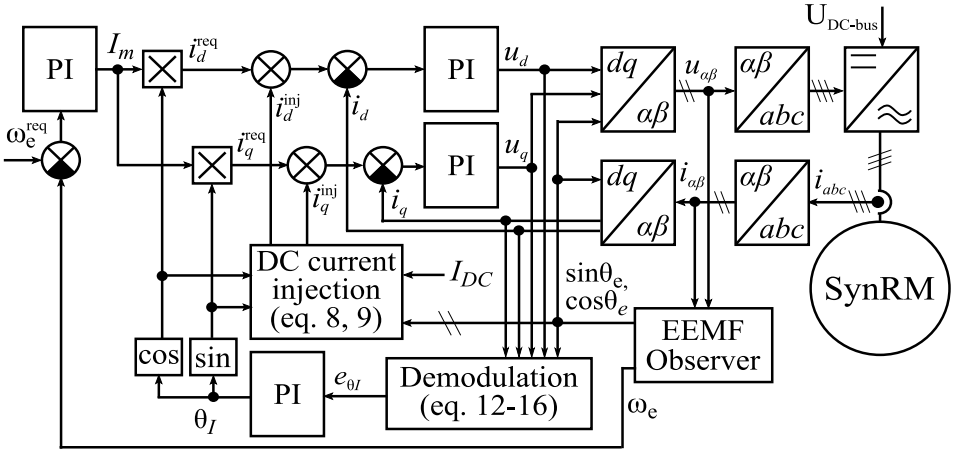


Figure 3: Block diagram of the proposed current angle control algorithm

4 EXPERIMENTAL RESULTS

The algorithm was implemented on NXP MKV46F256 Cortex-M4 microcontroller. The parameters of 550 W SynRM are stated in Table 1 and Figure 4. Amplitude of injected current was $I_{DC} = 50$ mA. Extended-EMF (EEMF) observer was used for position and speed estimation. [3]

Table 1: Parameters of SynRM

P_{nom}	550	W
N_{nom}	1500	rpm
U_{nom}	350	V
I_{nom}	1.6	A
P_p	2	-
R_s	9.68	Ω
J	$1.64 \cdot 10^{-3}$	kg.m^2

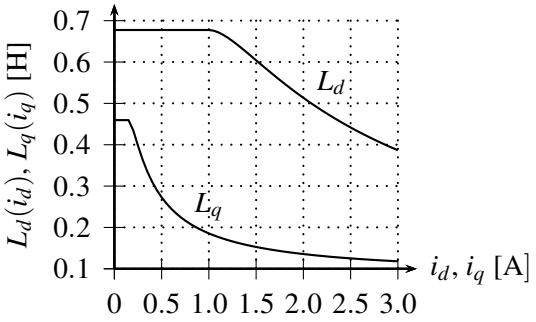


Figure 4: Current dependency of inductances

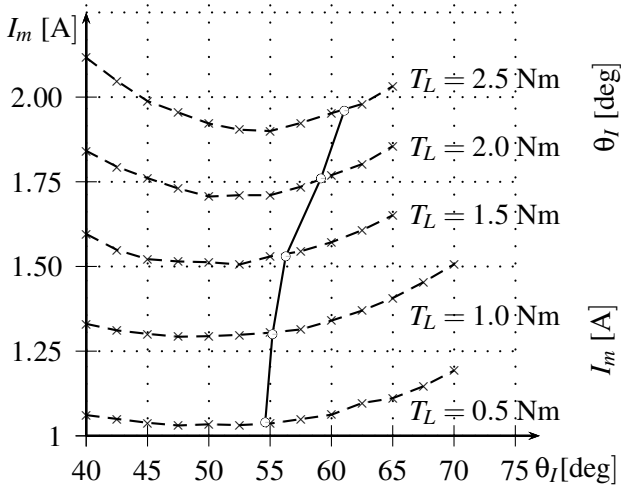


Figure 5: MTPA tracking with proposed algorithm

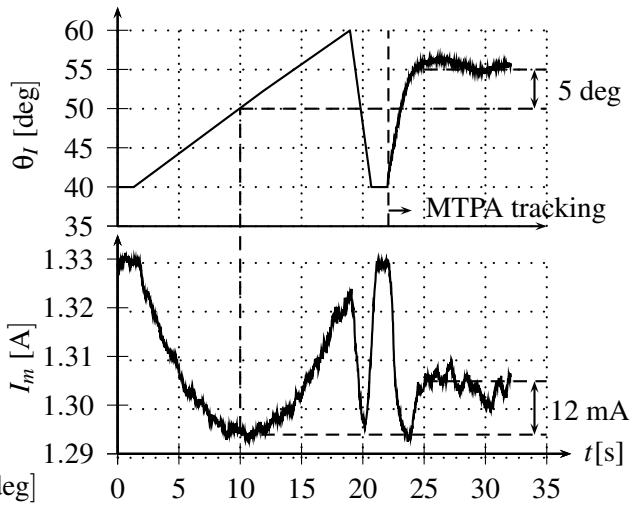


Figure 6: Transient response at $T_L = 1$ Nm

Figure 5 shows stator current amplitude I_m dependency on current angle θ_I for loading torque from 0.5 Nm to 2.5 Nm. As expected, the optimal current angle was always found above $\pi/4$. The current angle trajectory found by proposed algorithm was marked as well. It can be seen, that MTPA operating point is not tracked exactly and found angle is always higher than optimal one. Main reason is

probably the fact, that equations (16) and (7) are not exact match. Another causes might be neglected core losses and insufficient separation of frequency spectrum of injected signal and vector control. Overall, the current amplitude I_m was however lower than with classic angle setup $\theta_I = \pi/4$.

Figure 6 shows transient performance of proposed algorithm at $T_L = 1.0$ Nm. The current angle θ_I sweep is conducted up to 22 second in order to show minimum of I_m, which is located around $\theta_I = 50$ degrees. Then the proposed tracking algorithm is activated and after several seconds the current angle settles at value $\theta_I = 55$ degrees, which correspond to 12 mA higher value of I_m compared to optimal point. Just like all MTPA active search algorithms, the proposed algorithm is suited for steady states only, otherwise additional error is introduced. Data for both Figure 5 and Figure 6 were obtained at speed $\omega_e = 104.72$ rad/s.

5 CONCLUSION

MTPA active search power optimizing algorithm utilizing small DC current injection for SynRM was proposed in this paper. As stated in section three, the algorithm employs amplitude demodulation principle to obtain useful information on approximate location of MTPA operating point. Most notable benefits of algorithm are relative calculation simplicity and independency on motor parameters (except stator resistance). It is expected, that DC current would also allow for stator resistance measurement, however further research is needed. The disadvantages are unaviodable torque ripple caused by injected current, problematic low-speed operation and fact, that found solution is sub-optimal. Section four presents experiment results. Figure 5 shows MTPA operating point tracking performance of proposed algorithm. It can be seen that found solution does not match the optimal one, however it is generally better than standard setup $\theta_I = \pi/4$ for all loads.

ACKNOWLEDGEMENT

The completion of this paper was made possible by the grant No. FEKT-S-14-2429 - „The research of new control methods, measurement procedures and intelligent instruments in automation" financially supported by the Internal science fund of Brno University of Technology.

REFERENCES

[1] BOLOGNANI, Silverio, Roberto PETRELLA, Antonio PREARO and Luca SGARBOSSA. Online tracking of the MTPA trajectory in IPM motors via active power measurement. *The XIX International Conference on Electrical Machines - ICEM 2010*. IEEE, 2010, 1-7. DOI: 10.1109/ICELMACH.2010.5607843. ISBN 978-1-4244-4174-7.

[2] HUAI-CONG, L., K. IN-GUN and J. TAE-CHUL. Magnetic saturation and iron loss influence on max torque per ampere current vector variation of synchronous reluctance machine. *EVS28 International Electric Vehicle Symposium and Exhibition*. EVS28, 2015

[3] ICHIKAWA, S., M. TOMITA, S. DOKI a S. OKUMA. Sensorless Control of Synchronous Reluctance Motors Based on Extended EMF Models Considering Magnetic Saturation With Online Parameter Identification. *IEEE Transactions on Industry Applications*. 2006, 42(5), 1264-1274. DOI: 10.1109/TIA.2006.880848. ISSN 0093-9994.

[4] XU, L., X. XU, T.A. LIPO and D.W. NOVOTNY. Vector control of a synchronous reluctance motor including saturation and iron loss. *IEEE Transactions on Industry Applications*. IEEE, 1991, 977-985. DOI: 10.1109/28.90356. ISSN 00939994.

[5] ZENGCAI QU and M. HINKKANEN. Loss-minimizing control of synchronous reluctance motors — A review.*IEEE International Conference on Industrial Technology (ICIT)*. IEEE, 2013, 350-355. DOI: 10.1109/ICIT.2013.6505697. ISBN 978-1-4673-4569-9.

SUPPRESSION OF RESONANCES FROM VIBRATION SIGNAL USING VOLD-KALMAN FILTER

Martin Čala
Doctoral Degree Programme (2), FEEC, Brno University of Technology
E-mail: xcalam00@stud.feec.vutbr.cz

Supervised by: Petr Beneš
E-mail: benesp@feec.vutbr.cz

Abstract: Presented paper describes usage of Vold-Kalman filter of the first generation compared with time-frequency method as the short-time Fourier transform or order tracking technique based on resampling. Advantages of Vold-Kalman filter will be showed on extraction of order signal for variable rotor speed. Additionally, ability to extract multiple orders will be demonstrated as well as utilization of removing resonances from a real vibration signal during the rotor startup and coastdown.

Keywords: Vold-Kalman filter, resonance, crossing orders, order analysis

1 INTRODUCTION

There are multiple methods for analysis of vibration or acoustic signals from rotary machines. General time-frequency methods like the short-time Fourier transform (STFT) offer easy calculation of spectrogram but have limited resolution in time and frequency. In this context, rotor speed must be kept constant or changing only a little [1], otherwise the examined signal components are smeared. Methods using averaging and synchronized sampling are also used but the only application is for stable speed with small fluctuations. To overcome this limitation, order analysis methods based on resampling the vibration signal according to the rotational frequency are used. Instead of frequency, so called orders are used, which mean multiples of rotational frequency. The orders are constant thus spectrogram smearing is greatly reduced even when the speed is changed rapidly.

During the rotor startup and coastdown, orders are crossing the resonances originating from mechanics of the stator part. To determine condition of the machine, the energy of the orders is evaluated. The resonances distort the orders and highly reduce interpretation of such signals. Vold-Kalman filter is beneficial in such situations when the crossing orders need to be split [2].

2 VOLD-KALMAN FILTER

There are two generations of Vold-Kalman filter. Only first generation (VKF1) published by Vold and Leuridan will be described and used in this paper [3]. This filter is capable of extracting a sinusoidal waveform of time-varying frequency, also called the order. Moreover, this extraction is possible for close or crossing orders [4] as will be showed in experimental section of the paper.

Main idea of the method is covered by structural and data equations. Structural equation describes how the order x is modulated in frequency and amplitude as sinusoidal signal

$$x_n + c_n x_{n-1} + x_{n-2} = \varepsilon_n \,, \tag{1}$$

where n indexes data samples with total length N. Error term ε_n enables slight changes in amplitude and frequency. The order phase c_n is usually calculated from instantaneous frequency f and sampling period T_s

$$c_n = -2\cos(2\pi f_n T_s) \,. \tag{2}$$

Usually, more than one order needs to be extracted, as showed in data equation

$$y_n = \sum_{p=1}^{P} x_{p,n} + \eta_n \,, \tag{3}$$

where P means number of all orders to be extracted. Error term η includes remaining signal as noise or other non-harmonic components. The goal of the filter is to minimize error terms and thus find the best fit according to the equation

$$J = r^2 \varepsilon^T \varepsilon + \eta^T \eta \,. \tag{4}$$

Following matrix equation leads to a solution minimizing the criterion function from Eq. (4)

$$\begin{bmatrix} \mathbf{B}_1 & \mathbf{E} & \cdots & \mathbf{E} \\ \mathbf{E} & \mathbf{B}_2 & \cdots & \mathbf{E} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{E} & \mathbf{E} & \cdots & \mathbf{B}_P \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_P \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \\ \vdots \\ \mathbf{y} \end{bmatrix}, \quad \mathbf{A}_p = \begin{bmatrix} 1 & c_{p,1} & 1 & 0 & \cdots & 0 \\ 0 & 1 & c_{p,2} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & c_{p,N-2} & 1 \end{bmatrix}, \tag{5}$$

$$\mathbf{B}_p = r_p^2 \mathbf{A}_p^T \mathbf{A}_p + \mathbf{E} \,. \tag{6}$$

where matrices $\mathbf{B}_p$ are described as multiplication of matrix $\mathbf{A}_p$, $\mathbf{E}$ is identity matrix. Factor r adjusts filter bandwidth, higher value makes it narrower [5]. Equation (5) contains multiplication of the main matrix with dimensions NP by NP, solution vectors $\mathbf{x}$ has dimensions NP by 1, vector in the right side has also dimensions NP by 1. Mentioned matrix system contains nonzero values only on particular diagonals, this makes it sparse. The best option for solving such system is to use method of conjugate gradients with preconditioning [2], [6]. Beneficial property of VKF1 is zero phase delay.

3 EXPERIMENT

Signals used in this section are measured on a small balancing jig driven by an electric motor. Piezoelectric accelerometer is used to measure vibrations perpendicular to the rotation axis and tacho probe to measure rotor speed using reflective mark attached to the rotor. Sampling frequency for both channels is 25.6 kHz. There were measured two sets of data, one for unbalanced and one for balanced rotor. The main goal of this experiment is to show extraction of multiple orders on real data because this property is usually showed on synthetic signals [2], [4].

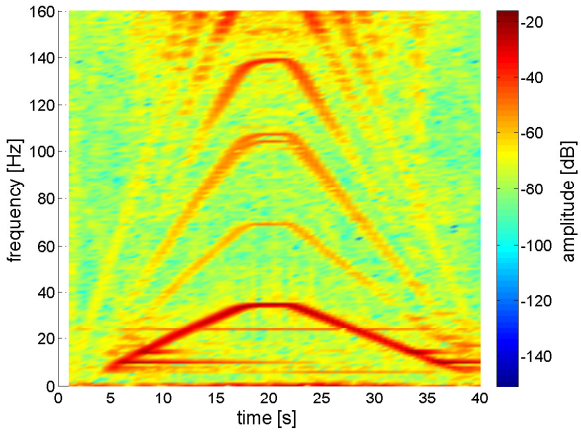


Figure 1: STFT spectrogram of an unbalanced rotor during startup and coastdown.

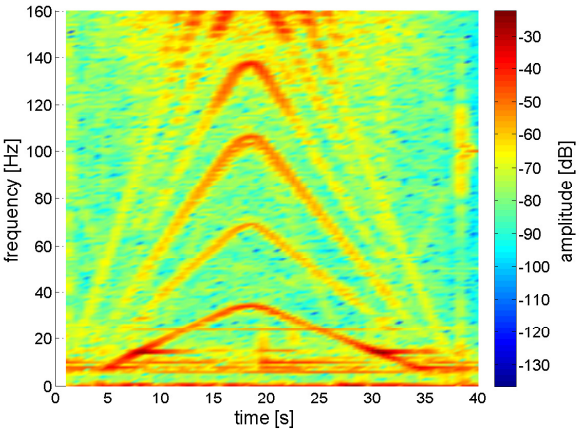


Figure 2: STFT spectrogram of a balanced rotor during startup and coastdown.

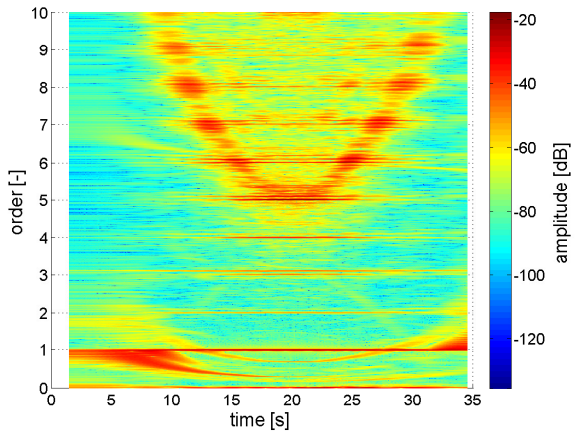


Figure 3: Order spectrogram of an unbalanced rotor.

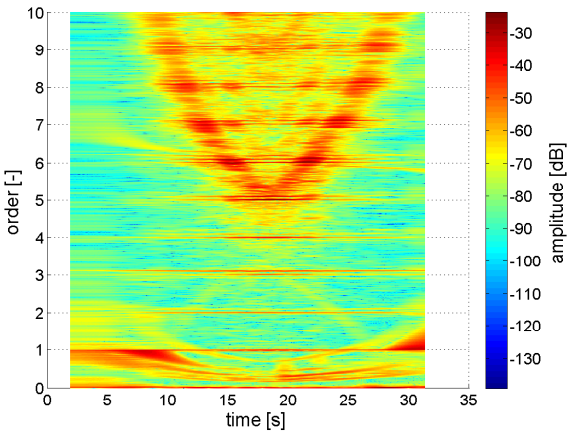


Figure 4: Order spectrogram of a balanced rotor.

To make the first view on the measured vibration signal, standard spectrogram using STFT was created. Since the trade-off between resolution in the time and in the frequency must be made, window of 51200 samples (2 seconds) was used and this window was applied on signals from both balanced and unbalanced rotors with shifts of 12800 samples (i.e. 1/4 of the window length) to make spectrogram smoother. Resulting spectrograms are showed in Figs. 1 and 2. It is clearly visible that the rotational speed is approximately linearly increasing to about 35 Hz and then decreasing toward zero as visible in the first, the strongest harmonic. Higher harmonics are also visible as integer multiples of base (first) harmonic. Apart from harmonics, there are resonances as disturbances at constant frequency. Effect of correct balancing is clearly visible in Fig. 2 since the first harmonic is significantly weaker than the one in Fig. 1.

Even that the spectrograms in Figs. 1 and 2 are smoothed, there are visible square artefacts, especially in rapid frequency changes. These correspond to resolution limits in the time and the frequency. For analysis purposes, it is better to use orders instead of frequency, because the order represents specific multiple of the rotational speed displayed as a straight line in the order spectrogram. To achieve this, vibration signal is resampled with respect to the rotational speed in the way that there is constant number of samples in one revolution, also called even angle signal. Order spectrogram is then created applying the STFT on even angle signal. Since the spectral smearing is avoided by the resampling, rectangular window is utilized. Resulting order spectrograms are showed in Figs. 3 and 4. The first harmonic is now straight line corresponding to the first order, which is not affected by the resolution as much as STFT spectrogram. Originally constant resonances are now hyperbolically shaped.

STFT and order spectrograms are widely used in vibration analysis because they allow easy overview of measured vibrations. Since the levels in spectrograms are distinguished by colors and there are effects of windows, it is impossible to see subtle details in the spectrograms. To look deeper and find even more, it is beneficial to see the time waveform of the particular order. This can be achieved by the Vold-Kalman filter.

For instance, higher amplitude of the first order of the vibration signal shows how precisely the rotor is balanced. To see exact amplitude as a number, it is needed to extract the waveform with the same frequency as the speed profile called order waveform. Higher amplitude will show that the rotor is worse balanced. VKF1 with bandwidth of 1 Hz is used and there were extracted waveforms from vibration signals measured on both balanced and unbalanced rotors. Such bandwidth will cover only the order signal without any noise or other components. Additionally, an envelope of both signals was calculated using Hilbert transform for simpler analysis. Figs. 5 and 6 clearly show that

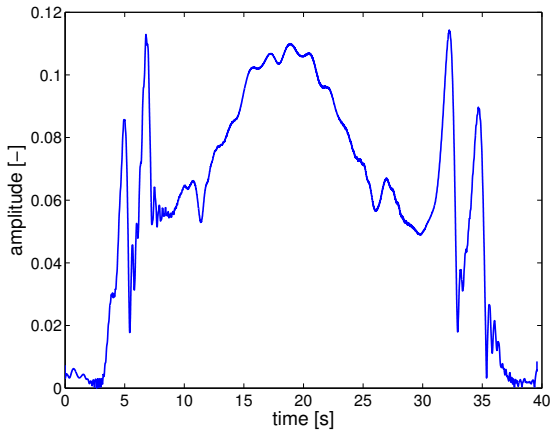


Figure 5: Result of VKF1 extraction only of 1st order for unbalanced rotor.

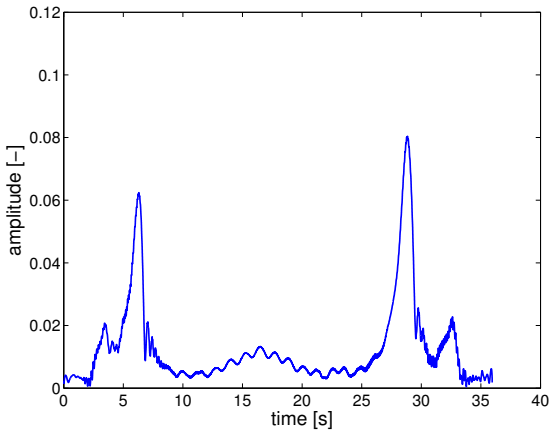


Figure 6: Result of VKF1 extraction only of 1st order for balanced rotor.

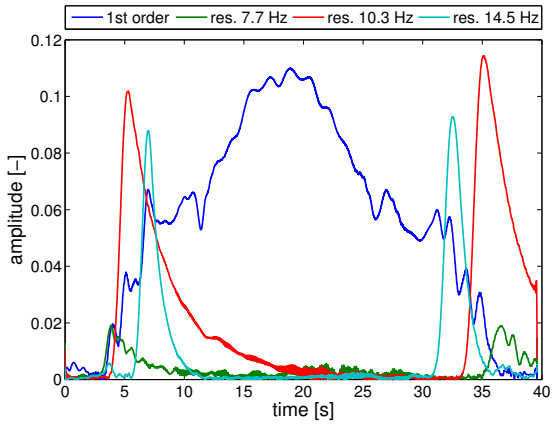


Figure 7: Result of VKF1 extraction of 1st order and major resonances for unbalanced rotor.

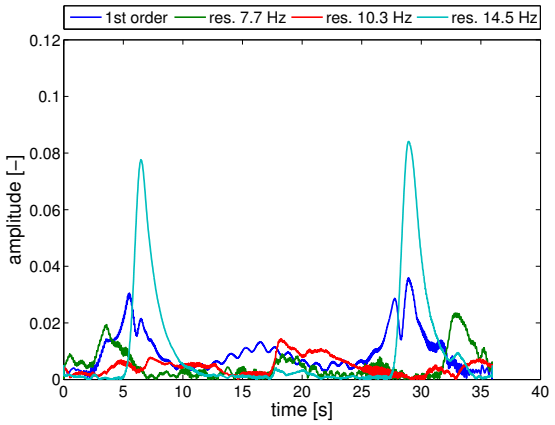


Figure 8: Result of VKF1 extraction of 1st order and major resonances for balanced rotor.

the instantaneous amplitude of the first order of unbalanced rotor is about ten times higher than for balanced one. Moreover, substantial resonances crossing the order during the startup and coastdown put unwanted spikes to the order envelope. Resonances amplitude is about the same level as the amplitude of the order. Removal of such resonances would make it possible to obtain more descriptive order signal.

The simplest way how to suppress the resonances is to create their model and extract multiple orders from the signal using VKF1. Apart from speed profile of previously used first order, additional speed profiles with constant frequency of the disturbing resonances will be used. STFT spectrogram using longer window of 8 seconds to increase the frequency resolution is used to find the strongest resonances. These resonances are located at 7.7 Hz, 10.3 Hz and 14.5 Hz. Filter bandwidth associated to all four extracted signals is 1 Hz, thereby other components are significantly suppressed and more importantly, filter convergence is dramatically increased. Suppression of the resonances is clearly visible in Figs. 7 and 8. The first order envelope for unbalanced rotor is significantly enhanced.

4 CONCLUSION

The paper shows possible procedure for vibration analysis of rotating machinery where combination of multiple methods enables to inspect vibration signal in detail. The main accent is focused on re-

moving the resonances from the signal based on modelling and extraction using Vold-Kalman filter. An assumption of knowledge of the resonances frequency is used as well as theirs unchanging frequency. VKF1 effectively split the crossing orders. This was showed on the experimental data rather than on some synthetic signals usually used in the relevant literature. Filter bandwidth is critical since it significantly speeds up filter convergence for method of conjugate gradients with preconditioning used in the experiment.

Presented graphs show that suppressing of the resonances is possible when these are properly modelled. The main advantage of this method becomes apparent for rotating devices during the startup and coastdown, even with dramatic speed transitions. Such scenarios are suitable for final tests in production areas.

ACKNOWLEDGEMENT

The completion of this paper was made possible by the grant No. FEKT-S-14-2429 – "The research of new control methods, measurement procedures and intelligent instruments in automation" financially supported by the Internal science fund of Brno University of Technology.

REFERENCES

[1] BRANDT, Anders. *Noise and Vibration Analysis: Signal Analysis and Experimental Procedures*. Chichester: John Wiley & Sons, 2011. ISBN 978-0-470-74644-8.

[2] TŮMA, Jiří. *Vehicle gearbox noise and vibration: measurement, signal analysis, signal processing and noise reduction measures*. Chichester, England: Wiley, 2014, 261 p. Automotive series (Wiley). ISBN 978-1-118-79761-7.

[3] VOLD, Håvard a Jan LEURIDAN. High resolution order tracking at extreme slew rates using Kalman tracking filters. *Shock and vibration*. New York, NY: John Wiley, 1995, Vol. 2 (6), pp. 507-515.

[4] PAN, M.-Ch. a Y.-F. LIN. Further exploration of Vold–Kalman-filtering order tracking with shaft-speed information—I: Theoretical part, numerical implementation and parameter investigations. *Mechanical Systems and Signal Processing*. 2006, Vol. 20 (5), pp. 1134-1154.

[5] TŮMA, Jiří. Setting the passband width in the Vold-Kalman order tracking filter. In: *International Conference on Sound and Vibration (ICSV12)*. Lisbon, Portugal, 2005.

[6] FELDBAUER, Christian a Robert HÖLDRICH. Realization of a Vold-Kalman tracking filter - a least squares problem. In: *Proceedings of the COST G-6 Conference on Digital Audio Effects*. Verona, Italy, 2000.

TRANSFORMATION OF UAV ATTITUDE AND POSITION FOR USE IN DIRECT GEOREFERENCING

Petr Gabrlik
Doctoral Degree Programme (4), FEEC BUT
E-mail: xgabrl00@stud.fit.vutbr.cz

Supervised by: Ludek Zalud
E-mail: zalud@feec.vutbr.cz

Abstract: The direct georeferencing in aerial photogrammetry requires precise measurement of position and attitude of an image sensor, because the object point position is computed directly using this data. But for several reasons, data from onboard inertial navigation system can not be used for this purpose directly. Especially GNSS antenna displacement and image sensor misalignment have to be compensated and different rotation and coordinate system conventions have to be taken into account. This paper describes the transformations of onboard GNSS/INS data for the needs of the direct georeferencing and demonstrates it on a real UAV photogrammetric mission.

Keywords: direct georeferencing, photogrammetry, UAV, 3D transformation

1 INTRODUCTION

There are many possible applications for aerial photogrammetry performed by micro and light[1] Unmanned Aerial Vehicles (UAV). They can be used for archaeological sites mapping [2], in agriculture for the monitoring of crop growth [3] [4], or in building industry for precise 3D surface/building model creation [3]. The main advantages of these vehicles are their low price, low operating costs, high level of autonomy and safe and fast operation. Due to their size and payload limits, there is typically no place for precise photogrammetric equipment such as calibrated frame camera, high precise Global Navigation Satellite System (GNSS) and Inertial Navigation System (INS). For this reason, indirect georeferencing approach is widely used in practice using this category of UAVs. This approach eliminates the need of precise onboard GNSS/INS equipment and performs well even with low-cost cameras. Only disadvantage of this technique may be the need of Ground Control Points (GCP) for terrain model or orthophoto georeferencing.

Sometimes, e.g. during disasters or in inaccessible or dangerous areas, it is not possible to use GCPs for georeferencing. This paper deals with some issues, which relate to the direct georeferencing approach. At first, section 2 describes the basic principle of the direct georeferencing and compares rotation conventions used in airborne navigation and photogrammetry. Section 3 is aimed to a real photogrammetric mission and some results of the direct georeferencing are presented there.

2 DIRECT GEOREFERENCING

In photogrammetry, the direct georeferencing means the technique, which uses known exterior orientation of image sensor (3 position coordinates $[X_0, Y_0, Z_0]$ of projection center O_i and 3 attitude coordinates $[\omega, \phi, \kappa]$) for the computation of absolute object point position $[X, Y, Z]$. Every object point P have to be visible at least at two overlapped photographs, where its position is given by 2 coordinates $[\xi, \eta]$. Except that, camera parameters called interior orientation have to be known. The

[1] UAV weight classification according [1]. The weight limit for *micro* UAV is 5 kilograms and the limit of 50 kilograms is for *light* UAV category.

most important is focal length c, but there are many other parameters, which mainly describes lens geometrical distortion. The principle of the direct georeferencing is illustrated in Figure 1.

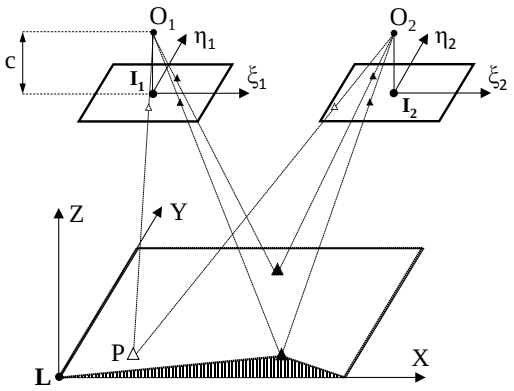


Figure 1: Direct georeferencing.

2.1 ROTATION AND COORDINATE CONVENTIONS

The aforementioned 3 attitude coordinates $[\omega, \phi, \kappa]$ represent photogrammetry Euler angles, which describe the attitude of the image sensor [5]. Since Euler angles are ambiguous in attitude description (in contrast to quaternions and rotation matrix), the sequence of rotations have to be known. Equations (1), (2) and (3) describe rotations about three principle axis: rotation of α radians about x-axis, rotation of β radians about y-axis and rotation of γ radians about z-axis.[2] [3]

$$\mathbf{R_x}(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c\alpha & -s\alpha \\ 0 & s\alpha & c\alpha \end{bmatrix} \quad (1) \qquad \mathbf{R_y}(\beta) = \begin{bmatrix} c\beta & 0 & s\beta \\ 0 & 1 & 0 \\ -s\beta & 0 & c\beta \end{bmatrix} \quad (2) \qquad \mathbf{R_z}(\gamma) = \begin{bmatrix} c\gamma & -s\gamma & 0 \\ s\gamma & c\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

In photogrammetry, there is no international convention for attitude representation, but this paper will follow conventions described in [5]. Three intrinsic rotations against a local system are performed in following order: ω radians about x_i image axis, ϕ radians about y_i image axis and finally κ radians about z_i image axis[4]. The resultant rotation matrix (4) is the composition of rotation matrices (1), (2) and (3). Another photogrammetry rotation conventions are mentioned in [6].

$$\mathbf{R_I^L}(\omega, \phi, \kappa) = \mathbf{R_{x_I}}(\omega)\mathbf{R_{y_I}}(\phi)\mathbf{R_{z_I}}(\kappa) \quad (4)$$

In airborne navigation, there is typically used another sequence of rotations. This convention is described in many publications (e.g. [7]) and by several international standards (e.g. ISO 1151–2:1985). An airplane attitude is performed by the sequence of the following three intrinsic rotations against navigation system: Ψ radians around z_b body axis (which is identical to z_n navigation axis), Θ radians around y_b body axis and Φ radians around x_b body axis[5]. The resultant rotation matrix has form

[2] All rotations are performed according to the right hand rule; positive rotation is counterclockwise, when looking along the positive part of the axis towards the origin.

[3] The functions *sine* and *cosine* are abbreviated by s and c in this equations.

[4] The angles of rotation about three distinct axes are also called Tait-Bryan angles or Cardan angles (or e.g. x-y-z for extrinsic and x-y'-z" for intrinsic rotations). But in general, this angles are called Euler angles.

[5] Rotations about z_b, y_b and x_b axes, which describe the attitude of an aircraft, are often called Yaw, Pitch and Roll or Heading, Elevation and Bank.

showed in equation (5).

$$\mathbf{R}_{\mathbf{B}}^{\mathbf{N}}(\Phi,\Theta,\Psi) = \mathbf{R}_{\mathbf{z_B}}(\Psi)\mathbf{R}_{\mathbf{y_B}}(\Theta)\mathbf{R}_{\mathbf{x_B}}(\Phi) \tag{5}$$

The difference between described rotation sequences can be seen in gimbals shown in Figure 2 and 3. Since the navigation angles do not correspond to the photogrammetric angles, appropriate transformation have to be applied (in more details in [8] and [6]).

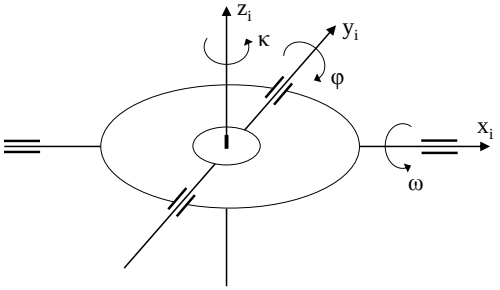


Figure 2: Photogrammetry rotations in gimbals.

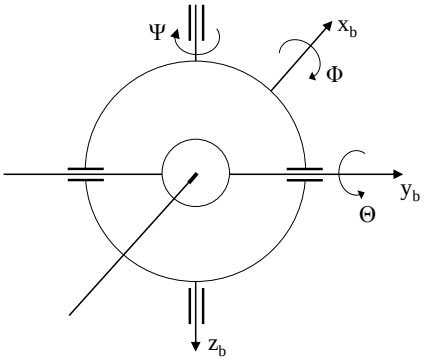


Figure 3: Airborne rotations in gimbals.

Except aforementioned rotations, different coordinate systems can be used. In airborne navigation there is typically used North-East-Down (NED) coordinate system and earth fixed local system in normally defined as East-North-Up (ENU). The appropriate rotation matrix have to be used for the system transformation [6].

2.2 COLLINEARITY EQUATIONS

Once the photogrammetric Euler angles are known and all position coordinates are in proper coordinate system, transformation between image sensor system I and local system L can be applied. This 7-parameters transformation (3 translations $[X_0, Y_0, Z_0]$, 3 rotations $[\omega, \phi, \kappa]$ and scale λ) based on the central projection is called spatial similarity (Helmert) transformation. Equation (6) shows the Helmert transformation complemented with sensor boresight alignment matrix $\mathbf{R}_{\mathbf{I}}^{\mathbf{B}}$ and image position offset vector $[X_I, Y_I, Z_I]^T$ in the body system B. An image point position in the system I is given by the coordinates ξ and η and the focal length c.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_L = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix}_L + \mathbf{R}_{\mathbf{B}}^{\mathbf{L}}(\omega,\phi,\kappa)\left(\begin{bmatrix} X_I \\ Y_I \\ Z_I \end{bmatrix}_B + \lambda \mathbf{R}_{\mathbf{I}}^{\mathbf{B}}(\Delta\omega,\Delta\phi,\Delta\kappa) \begin{bmatrix} \xi \\ \eta \\ -c \end{bmatrix}_I \right) \tag{6}$$

The transformation can be also expressed by collinearity equations (7) and (8) (derived in [5])[6]. This form clearly shows the link between image and object coordinates and it is suitable form for least square algorithm implementation.

$$X = X_0 + (Z - Z_0)\frac{r_{11}\xi + r_{12}\eta + r_{13}c}{r_{31}\xi + r_{32}\eta + r_{33}c} \tag{7}$$

$$Y = Y_0 + (Z - Z_0)\frac{r_{21}\xi + r_{22}\eta + r_{23}c}{r_{31}\xi + r_{32}\eta + r_{33}c} \tag{8}$$

[6]Elements r_{ij} mean elements of rotation matrix $\mathbf{R}_{\mathbf{I}}^{\mathbf{L}}$ where i represents row index and j column index.

3 MEASUREMENT

Described photogrammetric mission was performed in November 2015 in Aalborg, Denmark (N 57.060938, E 10.036612 - WGS84). The aim of the mission was to capture high amount of aerial photographs together with onboard navigation data for the direct georeferencing performance assessment. As the UAV platform was used fixed-wing aircraft Cumulus One (Figure 4) manufactured by Little Smart Things company, which was specially designed for fully autonomous airborne mapping missions. Since this UAV belongs to the *micro* category, it has very limited payload capability about 600 grams (max. take-off weight 2.2 kg, wing span 165 cm), therefore the UAV can fly only equipped with a compact camera (Sony RX-100 III in this case). The UAV is controlled by Pixhawk control unit, which contains navigation-grade GPS/INS sensors. For this reasons Cumulus One is suitable for indirect georeferencing, where the accuracy of 2.5 cm can be achieved.



Figure 4: UAV Cumulus One.



Figure 5: Ground Test Point.



Figure 6: An image from 60 meters altitude.

During one flight hundreds of near-vertical aerial photographs from different altitudes (60 to 100 meters) and directions were captured. For the purpose of accuracy assessment, 10 ground Test Points (TP), which accurate position was known, were used. Figure 5 shows one TP in the measured area and in Figure 6 there is sample photo captured from 60 meters altitude.

3.1 PROCESSING AND RESULTS

The processing was done in a semi-automatic fashion using Matlab. At first, the position of all visible TPs in every image were manually measured and position/orientation data from the flight log were assigned to the each photo. For every strip of photographs with around 80 % overlap the overdetermined system of equations (based on equations (7) and (8)) was compiled. This systems of equations were solved using Gauss-Newton least squares algorithm and multiple position results for every TP were estimated. Another approach, direct georeferencing using complex SW tools, is described in [4].

Figure 7 presents the results of the position estimation of 10 TPs in X-Y plane, when the position and attitude data from onboard INS were directly used in the collinearity equations (7) and (8). Even if all photographs were near-vertical (Roll and Pitch angles were close to zero), different rotation and coordinate system conventions cause high variance in the direct georeferencing results. Figure 8 shows position results of the same TPs, when the conversion from navigation to photogrammetry convention was applied (section 2.1).

4 CONCLUSION

This paper summarized the problem of different rotation and coordinate conventions used in airborne navigation and photogrammetry. Data from onboard INS/GNSS can not be used directly for the direct georeferencing, because the collinearity equations are not typically based on the same conventions. Since the Euler angles are ambiguous in attitude description, more information about the sequence of

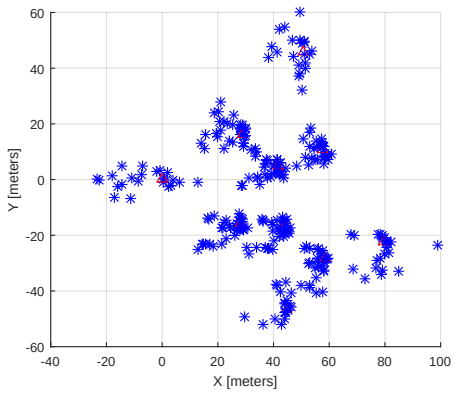


Figure 7: Direct georeferencing of 10 TPs using navigation data.

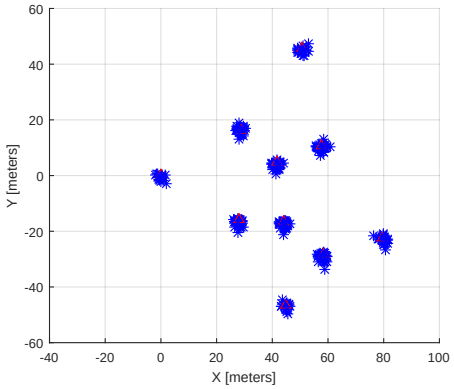


Figure 8: Direct georeferencing of 10 TPs after rotation and coordinate transformations.

rotations have to be specified. Described problem was presented on the real photogrammetry mission, where the impact of different conventions on the direct georeferencing is evident.

ACKNOWLEDGEMENT

The completion of this paper was made possible by the grant No. FEKT-S-14-2429 - „The research of new control methods, measurement procedures and intelligent instruments in automation" financially supported by the Internal science fund of Brno University of Technology. The author also would like to thank Anders la Cour-Harbo for his methodical help and assistance in collecting field data.

REFERENCES

[1] M. Arjomandi, S. Agostino, M. Mammone, M. Nelson, and T. Zhou, *Classification of Unmanned Aerial Vehicles*. Report for Mechanical Engineering class, University of Adelaide, Adelaide, Australia, 2006.

[2] G. Verhoeven, M. Doneus, C. Briese, and F. Vermeulen, "Mapping by matching: a computer vision-based approach to fast and accurate georeferencing of archaeological aerial photographs," *Journal of Archaeological Science*, vol. 39, pp. 2060–2070, July 2012.

[3] R. Haarbrink and H. Eisenbeiss, "Accurate DSM Production from Unmanned Helicopter Systems," in *International Archives of Photogrammetry and Remote Sensing*, vol. 21, pp. 1259–1264, 2008.

[4] D. Turner, A. Lucieer, and L. Wallace, "Direct Georeferencing of Ultrahigh-Resolution UAV Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, pp. 2738–2745, May 2014.

[5] K. Kraus, *Photogrammetry: Geometry from Images and Laser Scans*. Walter de Gruyter, 2007.

[6] M. Baumker and F. J. Heimes, "New calibration and computing method for direct georeferencing of image and scanner data using the position and angular data of an hybrid inertial navigation system," in *Integrated sensor orientation: test report and workshop proceedings*, (Hannover), pp. 197–212, 2002.

[7] W. F. Phillips, *Mechanics of Flight*. John Wiley & Sons, Jan. 2004.

[8] M. Cramer, *Performance of GPS/Inertial Solutions in Photogrammetry*. 2001.

WEB CONTROL OF THE ROBOTIC ARM

Jakub Arm
Doctoral Degree Programme (2), FEEC BUT
E-mail: xarmja00@stud.feec.vutbr.cz

Supervised by: Zdenek Bradac
E-mail: bradac@feec.vutbr.cz

Abstract: In this paper, control of the 5-DOF stationary robotic arm is presented. This robot uses RC servo-motors which are controlled using NXP LPC1756 at low level. This system is connected to server made up from Raspberry Pi running Java application server Glassfish. The robotic arm is controlled using only web browser using HTML5 and Javascript technology through internet connection. First purpose of this project is to popularize this type of robotic arm platform as a functional model of a true machine for hobbyists and researchers. Second purpose is to propose another approach of economic web based HMI for end user machines.

Keywords: Robotic arm, LPC1756, Raspberry Pi, RC servo, HTML5, Javascript, websocket

1 INTRODUCTION

One of the main scope of the industrial automation is to control robotic manipulators such as crane, arm, etc. These machines are used for many purposes, e.g. manipulation, cutting, or welding. Many hobbyists and researches focus on Cartesian machines of the portal type because of simple basic control [2]. The amount of hobby machine is growing among hobbyists, researchers and also end users. Most of this machines need to install specialized software which does not have to be supported by user OS. Therefore, web based HMI is good and easy solution for them.

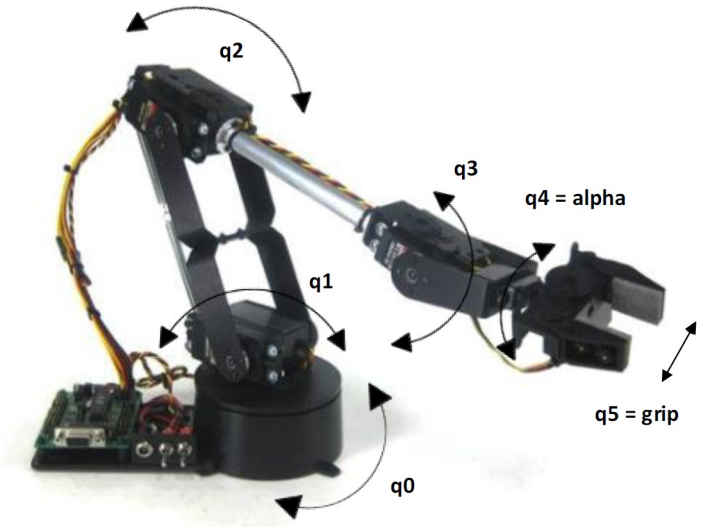


Figure 1: The robotic arm model [7]

2 RC SERVO MOTORS

RC (Remote Control) servo motor is a component for modellers which angle is controlled using DC motor by incoming PWM (Pulse Width Modulation) signal. This motor is controlled in closed loop which is an advantage because of low price of the whole component. A disadvantage is that the control loop tries to regulate the error using maximum motor power. Next disadvantage comes from the model construction. Some parts of the model are heavy enough to maintain the precise servo motor angle. There are some errors in RC servo motor precision caused by its construction, i.e. error in zero angle.

3 KINEMATIC EQUATIONS

The kinematic equations are needed for transformation of the coordinates. While motion is performing, the transformation from the Cartesian system to joint values is performed. This is called inverse kinematics. On the other hand, the operator wants to know the current Cartesian coordinates of the tool. So the current joint values are transformed to the Cartesian coordinates. This is called direct kinematics.

Regarding inverse kinematics, there are three issues. Firstly, the elbow joint value can be computed using two methods - using cosine rule or computing intersection points of two circles. The law of cosine is used in this project because it is computationally lesser demanding than the other method [3]. Secondly, points on Z axis and some other points can be reached by many movements according to base joint rotation. So when this kind of position is needed, the rotation of the whole robotic arm is not performed and the current rotation is used. Thirdly, all equations are computed using floating point data type, and trigonometric functions are computed by CPU using lookup tables and fast numeric approximation methods. All equations in velocity loop (indirect and direct kinematics) are solved within $(381 \pm 54)\ \mu s$ which is enough to meet deadline coming from velocity loop period 10 ms.

The direct kinematic equations $\overline{P} = f(\overline{q})$ are [7] :

$$\begin{aligned} x &= cos(q_0)[l_1 + l_3 cos(q_1+q_2) + l_2 cos(q_1) + l_4 cos(q_1+q_2+q_3)][mm] \\ y &= sin(q_0)[l_1 + l_3 cos(q_1+q_2) + l_2 cos(q_1) + l_4 cos(q_1+q_2+q_3)][mm] \\ z &= l_0 - l_3 sin(q_1+q_2) - l_2 sin(q_1) - l_4 sin(q_1+q_2+q_3)[mm] \\ \alpha &= q_4[^\circ] \\ \beta &= q_1+q_2+q_3[^\circ] \\ g &= q_5[mm] \end{aligned} \tag{1}$$

4 VELOCITY CONTROL

In this moment, the velocity is controlled according to ramp function. So instead of maximal velocity coming from low level RC servo loop, the velocity is controlled according to basic motion function in software. To overcome using only maximal velocity of the RC servo motor, the RC-motor control signal goes through points in high rate (velocity loop period) to make desired motion at desired speed. The result motion can be smooth because of motor load inertia and motor regulator time constant. When ramp motion function is used, there is a issue caused by jerk in consequence of the acceleration function. To eliminate high peaks in the jerk function, the S-curve motion profile can be used which is computationally more demanding.

The velocity controller is part of motion controller. Motion controller takes care about motion in a two axis plane defined by user command (G-code or manual command). In this moment, the motion control can only do the linear interpolation using one of the simplest algorithm. This algorithm

uses also not much of computational resources during motion because it computes a lot of before. The algorithm flow :

- compute velocities in each axis
- find the axis with the longest motion
- recompute motion profile for the other axis (lowering velocity)

5 MCU SUBSYSTEM

The MCU subsystem is based on LPC1756 from NXP. LPC1756 has ARM-Cortex M3 architecture without FPU (Floating Point Unit). Thus every equation with float types has to be done using main processor. The MCU subsystem is communicating with the server via UART using special G-code like protocol and commands for manual control. The MCU takes care about math (motion planning, direct and inverse kinematic) and generating PWM signals for RC servo-motors. The MCU subsystem is located on PCB created on VUT before.

In the MCU subsystem, FreeRTOS is used as a real-time operating system to meet hard deadlines and to make the software easy to extend, e.g. for a command mailbox with a independent command manager. This condition, used MCU and careful written application code ensures deterministic behaviour of this part of the system. This is essential for low level signal generation and when the velocity control loop is created in software.

The MCU subsystem cares also about inappropriate movements. So when a planned moved is out of scope of robot possibilities, the movement is canceled.

6 SERVER SUBSYSTEM

The server subsystem consists of the Raspberry Pi development board. There is Raspbian OS installed. The prebuilt image includes also some components like SSH, USB WiFi adapter driver, Java, etc. The Java application server called Glassfish4 has to be installed. Then the server runs the server side of the web application. This server side provides dynamic (HTML5 + Javascript) web pages and communicates actively with the web client. The received commands are then transmitted over serial line to the MCU subsystem. The whole system is described on Fig. 2.

The server can communicate with clients either using raw TCP or using websockets. Raw TCP needs implementation in some programming language and the handshake is faster than using websockets [9]. On the other hand, websockets can be directly implemented in dynamic web pages which provides comfort for programmers. Websockets have also less overhead than HTTP protocol [8].

Other advantage is that websocket supports message encoding and decoding which can be used for making a trusted connection using some fast cipher method, e.g. AES. Also some other web based proven authentication method can be used before the actual control connection is established. The control connection is then full-duplex client-server type.

Using public internet, low signal connection or connection through many active network components can bring some other issue regarding the latency. If the communication latency with the server is higher than 2 s, the current model information will not be accurate, new user commands will not be executed and the client will be disconnected. The MCU subsystem performs last command and will stay waiting for next command. The server will be waiting for new connection which can be only one at the same time.

Using Java programming language for real-time application is not recommended. On the other hand, this system is considered as firm real-time system and the computational requirements for server

subsystem are lesser thanks to MCU subsystem and human perception of real-time. Server source code is multi-platform across OS and Java provides programming comfort which saves costs.

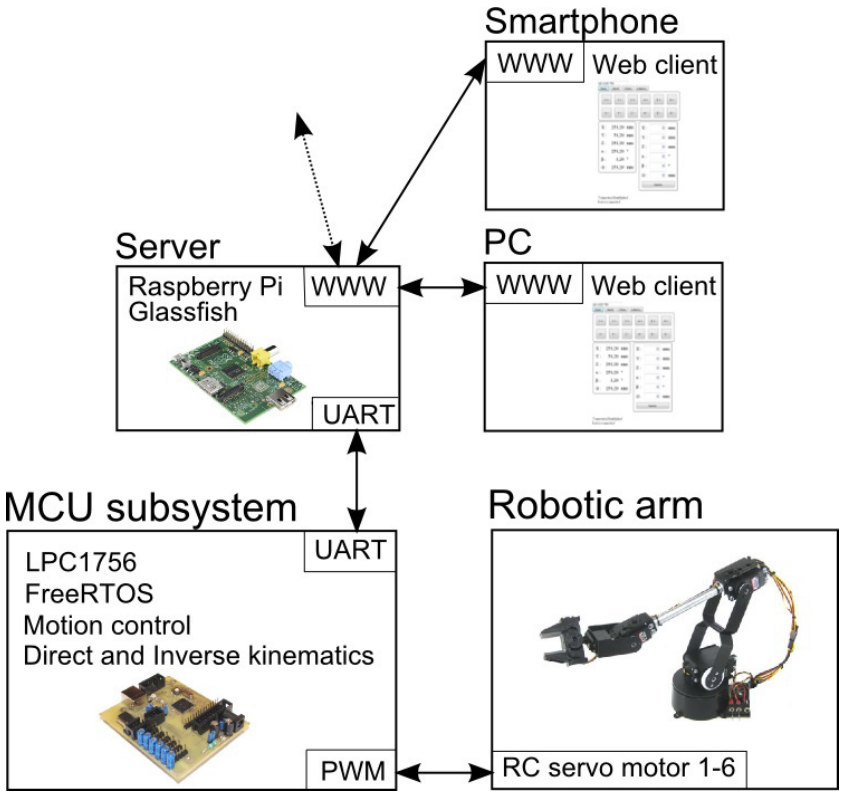


Figure 2: Block diagram of the system

7 WEB CLIENT

The web client is from category of thin clients. It runs completely in the web browser. It uses HTML5, CSS and Javascript technology to make pages dynamic and to communicate with the server nearly real-time. Real-time requirements come from human perception. In this project, a value 250 ms is considered as a soft deadline which is the average time of human reaction [1]. The reaction of the system has been measured using poll method (request followed by answer) with 60 character message as $(2,445 \pm 12,046)$ *ms* on the same network and $(4,550 \pm 29,402)$ *ms* on smartphone when five routers were traced. The highest latencies were $203,156$ *ms* on the same network and $937,687$ *ms* on smartphone. The message poll was done 10000 times per measurement.

This approach is intended to use in hobby low cost end user machines on local network. Some similar project [4] suggests using websockets in industry but only for data acquisition.

In this moment, it supports linear G-code commands and manual control of the robotic arm in the Cartesian coordinates. The current Cartesian axis values and the current joint values are monitored. The web client can run on any device with the web browser supporting Java and HTML5, e.g. PC (Firefox 15 and higher) or smartphone (Android 4.4 and higher).

8 CONCLUSION

This project aims to the firm real-time web control of the robotic arm built of the RC servo-motors using HTML5 and Javascript technologies. Some theory like kinematic equations, RC servo-steering,

motion control and real-time control is presented. The proposed approach of web based HMI can be used in any other end user machines. It is economic, multi-platform and client does not have to install any specialized software. In present, user can monitor the robot state, and control it manually or use G-code commands. So it behaves like CNC hobby machine with basic functions. In future, motion control is going to be improved, and 3D model of the robotic arm for the web client is going to be implemented or the system will be extended by camera feedback to make telepresence system. The communication latencies between the web client and the server on the same network have been measured and found sufficient.

ACKNOWLEDGEMENT

The completion of this paper was made possible by the grant No. FEKT-S-14-2429 - "The research of new control methods, measurement procedures and intelligent instruments in automation" financially supported by the Internal science fund of Brno University of Technology.

REFERENCES

[1] Thomas Burger. How fast is realtime? human perception and technology. *www.pubnub.com*, 2015.

[2] Josef Cernohorsky. Zaklady robotiky. *Liberec : Technicka Univerzita v Liberci. Fakulta mechatroniky, informatiky a mezioborovych studii.*, 2015.

[3] Jan Crha, Miroslav Kis, and Jan Mares. Inverse kinematics - tool for robotic manipulator control. 2013.

[4] C. M. Grigorescu, S. A. Moraru, and C. Grama. Industrial software monitoring system extension for mobile devices based on glassfish and phonegap. *Electrical and Power Engineering (EPE), 2012 International Conference and Exposition on*, pages 756 – 761, 2012.

[5] V. Pimentel and B. G. Nickerson. Communicating and displaying real-time data with websocket. *IEEE Internet Computing*, pages 45 – 53, 2012.

[6] Timotej Pochlopen. Robotic arm control. *Brno University of Technology, Faculty of Electrical Engineering and Communication.*, 2012.

[7] Michal Semrad. 6 axis manipulator control. *Brno University of Technology, Faculty of Electrical Engineering and Communication.*, 2014.

[8] K. Shuang, X. Shan, Z. Sheng, and C. Zhu. An efficient zigbee-websocket based m2m environmental monitoring system. *Dependable, Autonomic and Secure Computing (DASC), 2014 IEEE 12th International Conference on*, pages 322 – 326, 2014.

[9] D. Skvorc, M. Horvat, and S. Srbljic. Performance evaluation of websocket protocol for implementation of full-duplex web streams. *Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2014 37th International Convention on*, pages 1003 – 1008, 2014.

[10] Ondrej Sterba. Draughts playing robot. *Brno University of Technology, Faculty of Electrical Engineering and Communication.*

[11] Miroslav Vaclavek. 6 axis robot control in rtos. *Brno University of Technology, Faculty of Electrical Engineering and Communication.*, 2009.

Doktorské projekty

Silnoproudá elektrotechnika a elektroenergetika

ANALYSIS OF ECCENTRICITY AND BROKEN ROTOR BAR FAULTS EFFECT ON THE INDUCTION MOTOR SPECTRUMS OF CURRENT AND STRAY FLUX

Ielyzaveta Ishkova

Doctoral Degree Programme (2), FEEC BUT

E-mail: xishko00@stud.feec.vutbr.cz

Supervised by: Ondřej Vítek

E-mail: viteko@feec.vutbr.cz

Abstract: In this paper the analysis of effect of eccentricity and broken rotor bar faults on the spectrums of current and magnetic flux density outside the motor are presented. The behavior and origination of characteristic frequencies is presented and analysed. The results are presented in the form of figures illustrating the advantages of analysed diagnosis methods.

Keywords: induction motor, MCSA, stray flux, eccentricity, broken rotor bars

1. INTRODUCTION

There are a number of faults that can occur during the motor operation and they can be classified as: stator faults that result in the opening or shortening of a stator phase windings, abnormal connection of the stator windings, broken rotor bars or cracked rotor end-rings, static, dynamic, or mixed air-gap irregularities, bent shaft, shorted rotor field winding, bearing and gearbox failures.

Of the above types of faults the most prevalent ones are the eccentricity and the broken rotor bar faults. The result of presence of these faults in the induction motor can be the changes and unbalances in the voltages and line currents, increased torque pulsations, decreased average torque, overheating, etc.

In the present research the presence of the broken rotor bars in the motor as well as the presence of all of the eccentricity types (static, dynamic, mixed) were investigated. The condition monitoring of investigated motors was carried out using the motor current signature analysis (MCSA) as well as the monitoring of the magnetic flux density outside the motor [1, 2].

For clear understanding of the occurrence and presence of characteristic frequencies and other peaks in the spectrums, the behavior of the magnetic flux density depending on the time, together with spatial variation in the air-gap were considered.

2. FAULT EFFECT ON FREQUENCY CONTENT

It is well known that stator current spectrum of squirrel cage induction motors contains a wide range of harmonic components of different magnitudes and frequencies caused by the winding distribution, stator and rotor slot permeance, pole pair numbers, saturation, air-gap eccentricity, etc. Among these components the characteristic frequencies as well as rotor slot harmonics or principal slot harmonics are used as the indicators of the fault in the motor.

2.1. ECCENTRICITY FAULT EFFECT ON THE CURRENT AND MAGNETIC FLUX DENSITY SPECTRUMS OF MOTOR

The harmonic content in the spectrum of magnetic flux density may be described using analysis of the magnetic flux density in the airgap of the motor as a function of time and angle in the same ti-

me [3]. Such consideration allows to see the direct dependence of the principal slot harmonics (PSH) as well as fundamental frequency and its multiples on the fundamental number of pole pairs as well as the number of stator or rotor slots.

Considering the length of the motor airgap, its change due to the presence of the fault and its inverse function the value of the magnetic flux density in the motor with dynamic eccentricity is expressed as

$$B(\theta_m, \mathrm{t}) = \frac{\mu_0}{\delta} \cdot \left(F_{\mathrm{m}\nu} \cdot \cos(\nu\, \mathrm{p} \cdot \theta_m - \omega_1\, \mathrm{t}) \right) - \\ - \frac{k_{\varepsilon d} \cdot F_{m\nu}}{2} \left[\cos\left((\nu p - 1)\theta_m - (\omega_1 - \omega_r)\, \mathrm{t} \right) + \right. \\ \left. \cos\left((\nu p + 1)\theta_m - (\omega_1 + \omega_r)\, \mathrm{t} \right) \right] \quad (1)$$

where $k_{\varepsilon d}$ – is the level of dynamic eccentricity, p - is the number of pole pairs.

Magnetic flux density for the case of static eccentricity is:

$$B(\theta_m, \mathrm{t}) = \frac{\mu_0}{\delta} \cdot \left(F_{\mathrm{m}\nu} \cdot \cos(\nu\, \mathrm{p} \cdot \theta_m - \omega_1\, \mathrm{t}) \right) - \\ - \frac{k_{\varepsilon s} \cdot F_{m\nu}}{2} \left[\cos\left((\nu p - 1)\theta_m - \omega_1\, \mathrm{t} \right) + \right. \\ \left. \cos\left((\nu p + 1)\theta_m - \omega_1\, \mathrm{t} \right) \right] \quad (2)$$

where $k_{\varepsilon s}$ – is the level of static eccentricity.

For the consideration of current harmonic content there exist two methods of detection of air-gap eccentricity. In most cases it may be detected with the help of certain high frequency components in the stator line current. The first method allows to monitor behavior of the current at the sidebands of the slot frequencies. The sideband frequencies associated with an eccentricity are

$$f_{slot+ecc} = f_s \left[(kQ_r \pm n_d) \left(\frac{1-s}{p} \right) \pm n_\omega \right] \quad (3)$$

where $n_d = 0$ in case of static eccentricity, and $n_d = 1,2,3\ldots$ in case of dynamic eccentricity (n_d is known as eccentricity order), f_s is the fundamental supply frequency, Q_r is the number of rotor slots, s is the slip, k is any integer, and n_ω is the order of stator time harmonics that are present in the power supply driving the motor ($n_\omega = 1,3,5$, etc).

The second method is based on the monitoring of the behavior of the current at the sidebands of the supply frequency and its odd multiples. These frequencies of interest are given by

$$f_{ecc} = k_1 \cdot f_s \pm m \cdot f_s \cdot \left(\frac{1-s}{p} \right) \quad (4)$$

where m – is any integer.

2.2. BROKEN ROTOR BAR FAULT EFFECT ON THE SPECTRUMS OF THE MOTOR

Generation and presence of frequencies related to the presence of broken rotor bar faults can be described as follows. Due to the asymmetry in the rotor circuit, the rotor currents will produce positive- and negative-sequence rotor MMFs. Since the angular frequency of the rotor currents is $s\omega_1$, where ω_1 is the angular stator supply frequency and s is the slip, the angular speed of the positive-sequence rotor MMF with respect to the stator is

$$\omega_r + s\omega_1 = \omega_1 (1-s) + s\omega_1 = \omega_1 \quad (5)$$

and the negative-sequence rotor MMF with respect to the stator is

$$\omega_r - s\omega_1 = \omega_1(1-s) - s\omega_1 = (1-2s)\omega_1 \tag{6}$$

If the symmetrical three-phase stator winding of a symmetrical three-phase A.C. machine with smooth airgap is supplied by a symmetrical three-phase voltage system, due to the effects of space-harmonics, the airgap flux density distribution around the periphery is not sinusoidal, but can be described by

$$\begin{aligned} B(\theta_1,t) &= B_1\cos(\omega_1 t - p\theta_1) + B_5\cos(\omega_1 t - 5p\theta_1) \\ &-B_7\cos(\omega_1 t - 7p\theta_1) + B_{11}\cos(\omega_1 t - 11p\theta_1) - \ldots \end{aligned} \tag{7}$$

Where θ_1 is an angle around the periphery (θ_1 =0 is in the direct-axis of the stator), t is the time.

The frequencies of the induced rotor currents can be obtained by expressing (4) in the reference frame fixed to the rotor and thus by considering that

$$\theta_1 = \omega_r t + \theta_2 \tag{8}$$

where ω_r is the angular rotor speed and θ_2 is the initial position of the rotor at *t*=0,

$$\begin{aligned} B'(\theta_2,t) &= B_1\cos(s\omega_1 t - p\theta_2) + \\ &B_5\cos\left[(6-5s)\omega_1 t + 5p\theta_2\right] - \\ &B_7\cos\left[(7s-6)\omega_1 t - 7p\theta_2\right] + \\ &B_{11}\cos\left[(12-11s)\omega_1 t + 11p\theta_2\right] \end{aligned} \tag{9}$$

is obtained, where *s* is the slip. It follows that even in a symmetrical three-phase machine, due to space harmonics, the induced rotor currents will have slip-dependent frequencies.

The fundamental airgap flux density wave can be resolved into two counter-rotating waves. The positive-sequence wave has the normal amplitude of magnetic flux density, and the amplitude of negative-sequence one is changed due to the level of asymmetry.

In the case of asymmetries on rotor of an induction machine, when there is a broken rotor bar and only the fundamental space harmonic is considered, rotor currents will produce positive- and negative-sequence rotor MMFs, and their frequencies are $s\omega_1$.

Magnetic flux density in the case of asymmetry with relation to rotor is

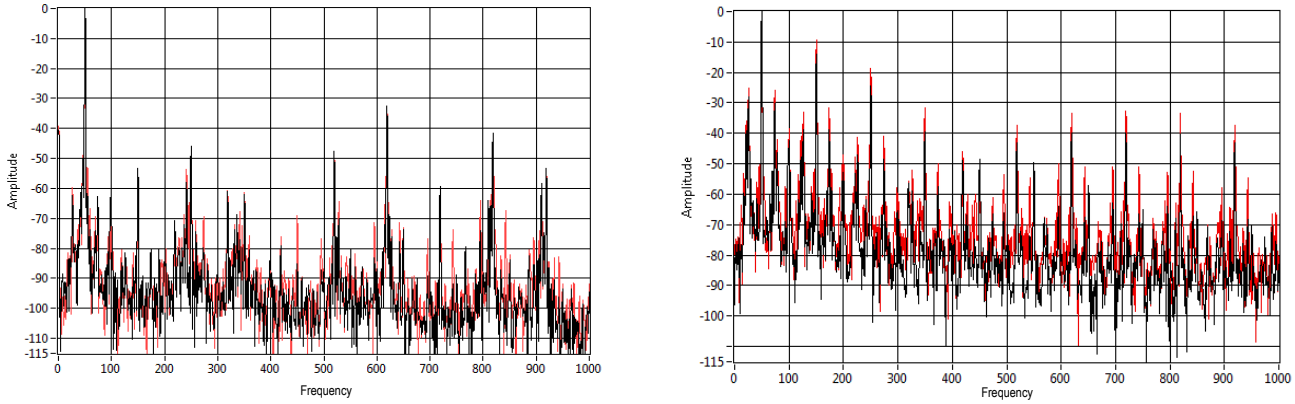
$$\begin{aligned} B'(\theta_2,t) &= B_1\cos(s\omega_1 t - p\theta_2) + \\ &B_{1F}\cos(-s\omega_1 t - p\theta_2) \end{aligned} \tag{10}$$

So the magnetic flux density formula considering the presence of rotor asymmetry in the stator reference frame is

$$\begin{aligned} B(\theta_1,t) &= B_1\cos(\omega_1 t - p\theta_1) + \\ &B_{1F}\cos((1-2s)\omega_1 t - p\theta_1) + \\ &B_\nu\cos(\nu\omega_1(1-s) - s\omega_1 - p\theta_1) + \\ &B_{\nu F}\cos(\nu\omega_1(1-s) + s\omega_1 - p\theta_1) \end{aligned} \tag{11}$$

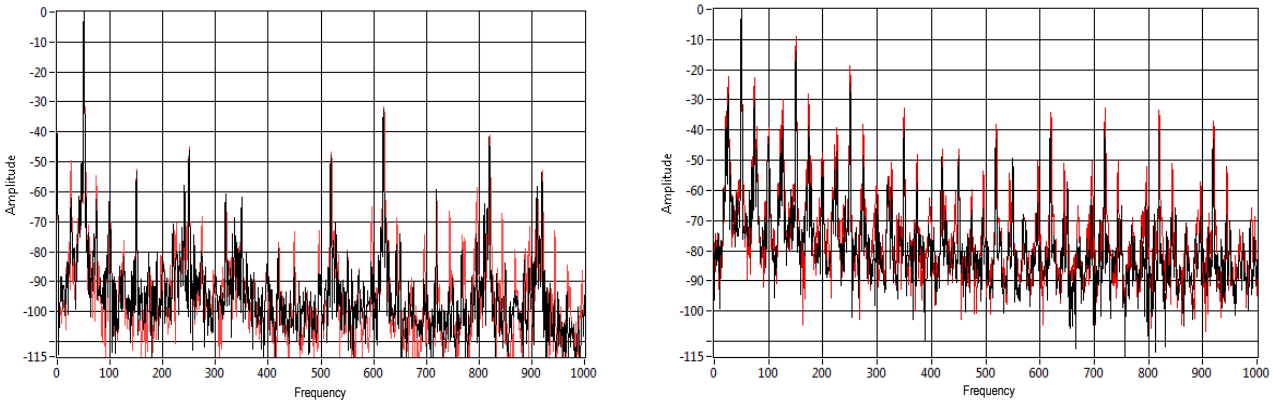
From stated above expression of the magnetic flux density the presence of the fundamental frequency can be clearly seen, as well as the presence of the sideband at frequency $2s\omega_1$. This is twice the slip frequency modulation of the supply current. Such a cyclic variation in the current reacts back onto the rotor to produce a torque variation at twice the slip frequency that, if the rotor does not have an infinitely high inertia, gives rise to $2s\omega_1$ variation on mechanical vibration that can be used for fault detection. This speed effect reduces the lower sideband current swing and produces an upper sideband at $(1+2s)\omega_1$, enhanced by modulation of the third harmonic flux in the stator and it can be seen that other sidebands at $(1\pm2s)\omega_1$ are also found [4].

3. EXPERIMENTAL RESULTS



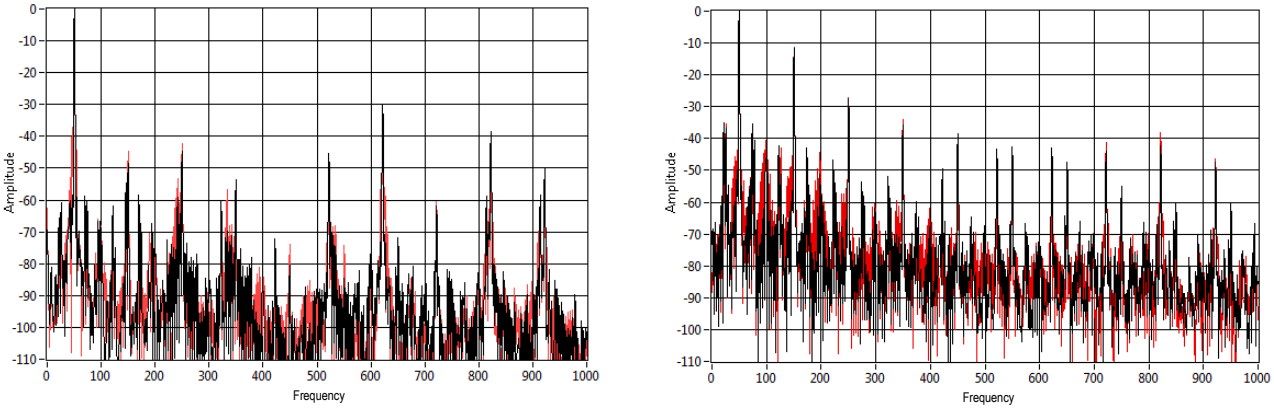
a) b)

Figure 1: Measured spectrums of current (a) and magnetic flux density (b) of the healthy motor (black) compared to the motor with dynamic eccentricity (red)



a) b)

Figure 2: Measured spectrums of current (a) and magnetic flux density (b) of the healthy motor (black) compared to the motor with mixed eccentricity (red)



a) b)

Figure 3: Measured spectrums of current (a) and magnetic flux density (b) of the healthy motor (black) compared to the motor with 2 adjacent broken rotor bars(red)

The current of motor as well as its stray flux were monitored and analyzed using the experimental installation with F.W. Bell 7030 Gauss/Teslameter and the magnetic flux sensor.

There were measured the 4-pole squirrel cage induction motors with static, dynamic and mixed eccentricities as well as the motors with different ammount and position of broken rotor bars. Measured results were supported by the data obtained from the somulation models. All obtained data were processed using the LabView Software, Ansys Maxwell and MatLab software.

From the obtained results it can be seen that the monitoring of the current spectrum of the motor is useful for the detection of the presence of eccentricity and broken rotor bar faults. Analysis of magnetic flux density in the case of these faults show that the monitoring of spectrum of stray flux of motor is useful in the case of eccentricity fault, that can be seen from the Figure 1(b) and Figure 2(b). But in the case of the broken rotor bar fault magnetic flux density is not useful, that is proved by the Figure 3 (b) as well as by the presented formulas.

4. CONCLUSIONS

In this paper the effects of eccentricity and broken rotor bar faults in induction motor were analysed. Data obtained from the measurement of real motors as well as from simulation in Ansys Maxwell was used for the analysis. The paper is focused on the effect of fault on the spectrums of current and magnetic flux density outside the motor. Methods used for the analysis of these effects are the MCSA with the monitoring of the stray flux of motor. The paper consist of the theoretical part where the formulas of magnetic flux density and there changes due to the presence of the fault are presented, as well as the experimental part. Obtained results are presented in the form of figures.

From the obtained results it is clearly seen that for the detection of presence of eccentricity fault monitoring of current spectrum is useful as well as the analysis of the magnetic flux density spectrums. Characteristic frequencies for this type of fault can be clearly destinguished in the spectrums, that allows to recognize the fault in the motor. For the diagnosis of broken rotor bar fault it is better to use the monitoring of the current spectrum, spectrum of magnetic flux density is not useful in this case.

ACKNOWLEDGEMENT

This research work has been carried out in the Centre for Research and Utilization of Renewable Energy (CVVOZE). Authors gratefully acknowledge financial support from the Ministry of Education, Youth and Sports of the Czech Republic under NPU I programme (project No. LO1210).

REFERENCES

[1] Romary, R., Pusca, R., Lecointe, J.P., Electrical Machines Fault Diagnosis by Stray Flux Analysis. In: IEEE Workshop on Electrical Machines Design, Control and Diagnosis, Paris, France, 2013, pp.247-256

[2] Ishkova, I., Vitek, O., Diagnosis of Eccentricity and Broken Rotor Bar Related Faults of Induction Motor by Means of Motor Current Signature Analysis. In: 16th International Scientific Conference on Electric Power Engineering, Kouty nad Desnou, Czeck Republic, 2015, pp.682-686

[3] Boldea, I., Nasar, S., The Induction Machine Handbook. Boca Raton: CRC Press, 2002

[4] Tavner, P., Ran, L., Penman, J., Sedding, H., Condition Monitoring of Rotating Electrical Machines. Stevenage: Institution of Engineering and Technology, 2008, ISBN 978-0-86341-739-9

CALCULATION AND MEASUREMENT OF THERMOSTATIC BIMETAL DEFLECTION IN MOLDED CASE CIRCUIT BREAKER

Michal Zelenka

Doctoral Degree Programme (1), FEEC BUT

E-mail: xzelen04@stud.feec.vutbr.cz

Supervised by: Bohuslav Bušov

E-mail: busov@feec.vutbr.cz

Abstract: This paper deals with a thermal overload trip unit in automatically operated electrical switching devices. The first part describes an application of thermostatic bimetal in thermal overload trip unit. The main part is devoted to analytical calculation and measurement of specific thermostatic bimetal type contained in thermal trip unit of molded case circuit breaker. It is measurement of deflection depending on temperature change using photoelectric laser sensor. The analytical results are compared with the measured results at the end of the paper.

Keywords: thermostatic bimetal, molded case circuit breaker, measurement

1. INTRODUCTION

Overcurrents generate excessive heat, leading to the risk of damage to electrical equipment. There are many components in automatically operated electrical switching devices used for scanning and evaluating overcurrents. Currently, the most common such component is a thermostatic bimetal, electromagnet and electronic evaluation of overcurrent using the current transformer and Rogowski coil. Thermostatic bimetal is a classic element mainly used for detection of small overcurrents caused by overloads. It is the main structural component of the thermal overload trip unit and thermal overload relays. In general, overcurrent trip unit is the main functional element of automaticly operated switching devices. It scans and evaluates the current flowing through the device (circuit breakers, motor starters) and instructs the device to trip acting directly on its switching mechanism.

2. THERMOSTATIC BIMETAL IN ELECTRICAL SWITCHING DEVICES

Thermostatic bimetal is made of two different metals bonded together. The two metals have different thermal expansion coefficients, so the bimetal bends when heated. Therefore, the basic thermostatic bimetal characteristics is an ability to change its shape due to a temperature change. In electrical switching devices this ability is used in thermal trip unit for a protection against overloads. If an obstacle (a tripping latch in circuit breaker) prevents the deformation of bimetallic strip as an opposite force to the internal forces causing the deformation, these internal forces can do a mechanical work on the track of deflection, release the latch of tripping mechanism and trip the circuit breaker. [2]

Thermal trip unit works on the principle of thermal effect of the current. The greater current, the hotter the bimetallic strip becomes, and the more it bends. Heat generated by the strip is converted to the force releasing the latch of the tripping mechanism. A tripping characteristic of thermal trip unit is time dependent. The greater the overcurrent through the breaker, the shorter the time it takes to interrupt the current. If the overcurrent of certain value passes for a certain time (depending on the tripping characteristic of the breaker), the bimetallic strip heats up and bends. After bending a predetermined distance, the strip makes contact with the tripper bar activating the tripping mecha-

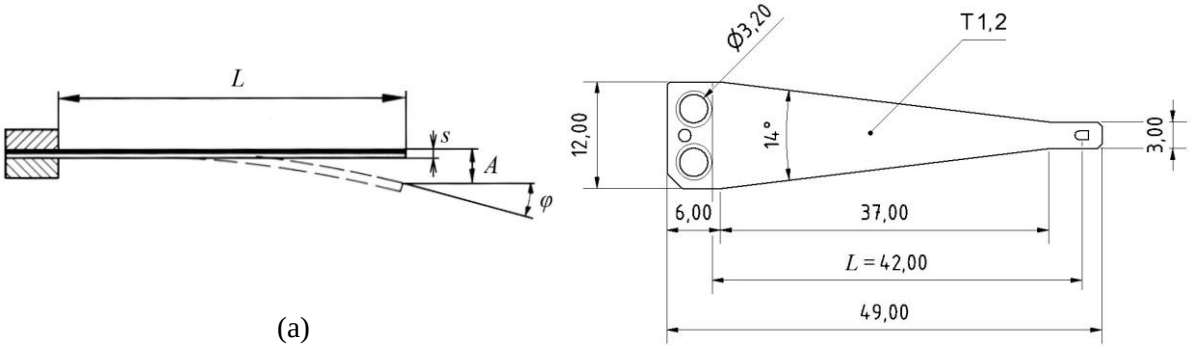
nism. Then, a power contacts of the breaker are automaticly disconnected using an energy stored in the switching mechanism springs. [5]

In the switching devices the bimetal element is most often shaped into straight, rectangular or trapezoidal cantilever. It is heated directly by passing current through itself or indirectly by a heating strip or winding with a defined resistance. Another possibility is a combined heating of the bimetal using the two previous methods. [5]

3. CALCULATION OF THERMOSTATIC BIMETAL DEFLECTION

Knowledge of the temperature dependence of deflection, a heating characteristic and a temperature dependence of force developed by the bimetallic strip (for partially or fully restrained deflection caused by the tripping mechanism) is important to determine the time needed to release the tripping latch by the strip (for the certain overcurrent). Deflection is not dependent on a width (shape) of the strip.

The calculation is performed for the bimetal used in overcurrent trip unit of a Sentron 3VA12 molded case circuit breaker with a nominal current of 250 A, intended for a line protection against overload and short circuit. The breaker is equipped with a thermal-magnetic overcurrent trip unit, which is a separate removable block. The trip unit consists of a thermal trip unit for protecting against overload, and a magnetic trip unit for protecting against short circuits. The both trip unit components are series-connected [3]. In the each pole of the thermal trip unit a straight trapezoidal bimetallic strip with an interlayer of copper is used. Its dimensions are shown in **Figure 1** (b). It is heated indirectly by a parallel current path to which is fixed.



(a) (b)

Figure 1: Deflection of straight thermostatic bimetal cantilever strip (a) and dimensions of the strip in trip unit (b)

Deflection of straight thermostatic bimetal cantilever strip for the temperature change is calculated according to [1] by following equation:

$$A_s = \frac{a\Delta T L^2}{s} = \frac{0{,}53k\Delta T L^2}{s} \qquad (1)$$

Where A_s is deflection [mm], a is specific thermal deflection [K^{-1}], k is specific thermal curvature [K^{-1}], ΔT is temperature change [K], L is active length [mm] and s is thickness [mm].

The deflection of the strip was calculated using equation (1). The specific thermal curvature according to the manufacturer [1] is $k \pm 5$ %. The calculation considers an active strip length of L = 42 mm, because of a measurement method. The calculated deflection A_s for temperature change $\Delta T = (0 \div 160)$ K considering $k \pm 5$ % is shown in **Table 1:** and a graphical dependence is shown in **Figure 3**.

Temperature change [K]	Deflection [mm]		
	k	*k* - 5%	*k* + 5%
0	0,00	0,00	0,00
20	0,58	0,55	0,61
40	1,15	1,10	1,21
60	1,73	1,64	1,82
80	2,31	2,19	2,42
100	2,88	2,74	3,03
120	3,46	3,29	3,63
140	4,04	3,83	4,24
160	4,61	4,38	4,84

Table 1: Calculated values of bimetallic strip deflection depending on temperature change

4. MEASUREMENT OF THERMOSTATIC BIMETAL DEFLECTION

4.1. MEASUREMENT METHOD

The aim of the measurement was to verify the calculated bimetallic strip deflection depending on temperature change. The measurement was performed in a Professor List Technology Park FEEC BUT.

A hole for sensing the position of the bimetallic strip free end was drilled into the 3VA12 circuit breaker plastic cover. The strip deflection was measured using a special extension piece with a free-hanging plastic tube and a measuring disc. It was mounted directly on a plastic cap of the strip free end. To the strip passive layer, two thermocouples for the temperature measurement were fastened. The first was fastened to the free end, another to the fixed end. The both thermocouples for the bimetal temperature measurement and also a thermocouple for an ambient temperature measurement were connected to an universal multi-channel logger.

For the deflection measurement, the tripping mechanism of the breaker was removed. The breaker was mounted to a vertical stand in a classic working position and a pole, in which the strip temperature and deflection was measured, was connected to a current source. For the disc position measurement a photoelectric laser sensor was fastened below the breaker and the sensor beam was aimed precisely to the center of the disc. In this measurement method, it can be considered that the measured disc deflection corresponds to the true strip deflection in the trip unit. The laser sensor was powered by a laboratory source. The analog laser output representing the detected distance in its operational measurement range was measured by a digital multimeter. The current flowing through the breaker was measured by a current transformer and digital multimeter. Scheme of the strip deflection measurement is shown in **Figure 2**.

The temperature dependence of the strip deflection was measured for heating by current of 262.5 A and 325 A and subsequent cooling. These are the values of conventional non-tripping and tripping current of molded case circuit breaker tripping characteristic. According to [5] conventional non-tripping current (I_{nt} = 1.05 x I_n = 262,5 A) is a specified value of current which the thermal unit can carry for a specified time (2 hours) without operating. Conventional tripping current (I_t = 1.3 x I_n = 325 A) is a specified value of current which causes the thermal unit to operate within a specified time (2 hours).

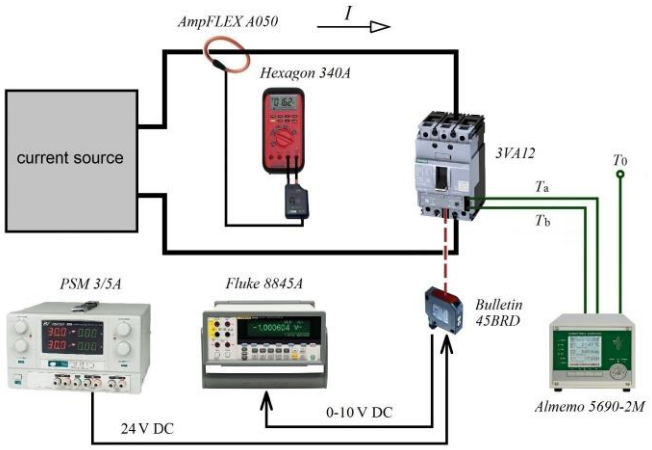


Figure 2: Measurement of thermostatic bimetal deflection

The measured values of the laser sensor DC output were converted to the corresponding distance of the disc in the sensor measuring range. Instantaneous strip temperature T is the mean temperature of the free end T_a and fixed end T_b of the strip. Instantaneous temperature change of the strip is:

$$\Delta T = T - T_i \qquad [\mathrm{K}] \qquad (2)$$

where T_i is measured initial mean temperature of the strip. Deflection of the strip (disc) is:

$$A_m = l_i - l \qquad [\mathrm{mm}] \qquad (3)$$

where l_i is measured initial distance of the disc in the sensor measuring range. Absolute error of the deflection meaurement is:

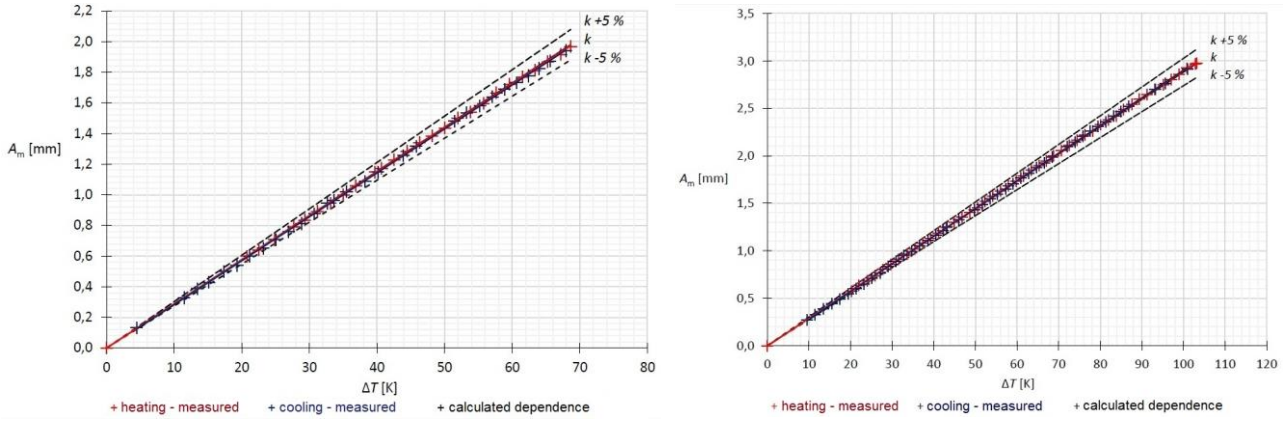
$$\Delta_A = \left| A_m - A_s \right| \qquad [\mathrm{mm}] \qquad (4)$$

and relative error:

$$\delta_A = \frac{\Delta_A}{A_s} \cdot 100 \qquad [\%] \qquad (5)$$

where A_s is the calculated deflection.

4.2. MEASURED DATA



(a) (b)

Figure 3: Measured bimetallic strip deflection depending on temperature change for heating by current of 262,5 A (a) and 325 A (b)

5. CONCLUSION

As can be seen in **Figure 3** for heating by the both currents 262.5 A and 325 A and subsequent cooling the bimetallic strip deflection is directly proportional to the temperature change. A whole range of the measured values is within a linearity range (-20 ÷ 200) °C given by the strip manufacturer in [1]. According to [1], the linearity range is the temperature range within which the strip deflection does not vary more than ± 5% from the deflection as calculated from the specific thermal curvature and the thickness.

The measured deflection for heating by current 262.5 A and subsequent cooling does not vary more than 2.9 % from the calculated deflection. The steady state temperature change of 68.7 K corresponds to the deflection of 1.97 mm. In the case of heating by currnet 325 A and subsequent cooling the measured deflection does not vary more than 2.7 % from the calculated deflection. The steady state temperature change of 103.2 K corresponds to the deflection of 2.98 mm. The measured dependence $A_m = f(\Delta T)$ therefore corresponds to the calculated dependence $A_s = f(\Delta T)$ when considering the active strip length of 42 mm. However, the measurement was affected by the following measurement errors:

- The measurement method. In the whole measurement range it is considered that the measured disc deflection corresponds to the true strip deflection.
- Uneven distribution of temperature when heating by the current. It is contemplated that the strip temperature is the mean value of the free and fixed end of the strip.
- According to the manufacturer the strip is straight at the temperature of 20°C. There is no deflection at this temperature. In this measurement it is contemplated that the strip is straight at the initial temperature T_i before heating.
- Instruments error.

ACKNOWLEDGEMENT

This research work has been carried out in the Centre for Research and Utilization of Renewable Energy (CVVOZE). Authors gratefully acknowledge financial support from the Ministry of Education, Youth and Sports of the Czech Republic under NPU I programme (project No. LO1210)

REFERENCES

[1] UHLIG, Wolfgang, 2007, "Thermostatic Metal: Manufacture and Application," Auerhammer Metallwerk GmbH, Aue, Germany, pp. 200

[2] HAVELKA, Otto, 1984, *Stavba elektrických přístrojů I*. Brno, Czechoslovakia

[3] *3VA Molded Case Circuit Breakers* [online]. 2014. Regensburg (Germany): Siemens AG, Infrastructure & Cities Sector Low and Medium Voltage Division Low Voltage & Products. ISBN E86060-K8220-E480-A3-7600

[4] ZELENKA, Michal, 2015, *Analysis of a bimetal release*. Master's thesis. Brno: Brno university of technology. Faculty of electrical engineering and communication

[5] OEZ, *Příručka elektrotechnika I*. Letohrad, Czech Republic, 2011

ELECTRICITY LOAD FORECASTING USING AUTO-REGRESSIVE AND ARTIFICIAL NEURAL NETWORK MODEL

Vaclav Vycital

Doctoral Degree Programme (1), FEEC BUT

E-mail: xvycit01@stud.feec.vutbr.cz

Supervised by: Petr Toman

E-mail: toman@feec.vutbr.cz

Abstract: In this paper a short review of two forecasting models Autoregressive and Artificial neural network is presented. Both models were used to demonstrate its superior performance in load forecasting issues. In the third section the results of load forecasting experiment are given. For obtained forecasted results mean absolute percentage error for autoregressive model was 0.644 % and for artificial neural network model 2.31 %. In this paper error distribution for both models is also shown.

Keywords:Load, load forecast, load prediction, autoregressive model, artificial neural network

1. INTRODUCTION

A knowledge of electricity load and its changes in time is one of crucial information necessary for operation and planning decision for distribution and transmission operators companies. It is important to know the expected changes in electricity load or "load forecasting (LF)" to make appropriate arrangements like to specify the suitable time for maintenance, load switching or infrastructure development. In deregulated electricity market this information is also necessary for all market participants (energy suppliers, financial institutions, transmission and distribution operators) to help them make important decision on purchasing electricity, generation capacity and spinning reserve etc. Thus it is essential to develop accurate methods and techniques for electricity load forecasting. Accurate load forecasts can improve the overall economic of electricity power system and also increase the reliability of the network. Load forecasts can be divided into three categories [1]: short term load forecast (STLF) which is usually from one hour to one week, medium term load forecast which is usually from one week to one year and long term forecast which is usually from one year to longer forecasts. STLF can help to estimate optimal load flow or avoid overloading. Also in recent years with the deregulation of market and increasing number of renewable energy sources (RES) the importance of accurate LF methods is growing. However with the rollout of smart grid implementation plan and smart meters the importance of LF methods grew even further.

Several methods were developed for STLF [1] that are usually based on statistical or artificial intelligence approach like Similar day approach, Polynomial regression, Fuzzy logic, Support vector machine or Expert systems. In this paper STLFs were based on the other ones:

Time series – Time series are statistically considered as time sampled values (usually discrete or continuous). Those methods are based on assumption that the data have an internal structure, like an autocorrelation, trend or seasonal variation. Time series forecasting methods learn such structure. For single input and single output models, the forecasted value is obtained as a linear combination of any amount (order) of previously observed values – i.e. simple autoregressive (AR) model. Other time series models are (ARMA, ARIME, ARMAX, ARIMAX, FARMAX etc.)

Neural networks – Artificial neural network (ANN or NN) is a method that allows to overcome the need of functional notation. With the development of computer technology and its computation ca-

pacity this method can be used for many different applications. The idea and behaviour of ANN is derived from biological neural network. This fitting/prediction model can be used to forecast or to fit any linear or nonlinear dependency. For load forecasting use, multiple different inputs can be considered as weather, day of week, hour etc.

2. AR AND ANN FORECASTING METHODS DESCRIPTION

Notation of single input and single output autoregressive model AR(p) of order p used for prediction can be written as [2]

$$Y_t = \sum_{i=1}^{p} (a_i \cdot Y_{t-i}) + e_t \tag{1}$$

Where Y_t is a load value at time t (i.e. forecasted value), $Y_{t\text{-}i}$ are load values at time t-i, p is an order of AR model and that also means from how many previous values is the next value calculated, a_i are the unknown coefficients and e_t is a white noise. The unknown coefficients of AR model can be determined by using ordinary well known least mean square (LMS) algorithm.

Another used model for STLF in this paper was ANN because of its ability to compute complex problems with multiple inputs and outputs without the need of finding exact functional notation [3]. ANN can be seen as a network of connected neurons (nodes). ANN consist of multiple layers with the first one input layer and the last one output layer. Between input and output layer, there can be any amount of hidden layers, simple network architecture is depicted in Figure 1 a).

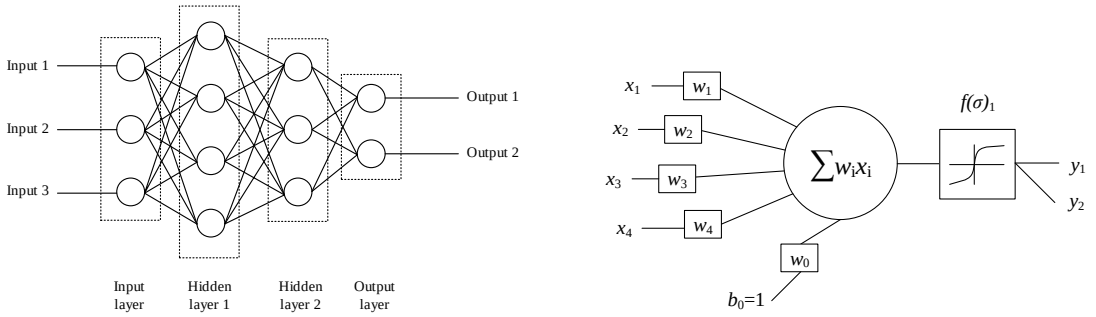


Figure 1: Multilayer NN architecture a) and detail of one node in hidden layer b)

For most of the problems ANN with only one hidden layer is capable to learn any nonlinear dependency between inputs and outputs. Each node of the network has its respective weight vector $\boldsymbol{w}_i$, summation function $\sum$ and activation or threshold function $f(\sigma)_1$. Before the inputs can be applied to the ANN they have to be pre-processed. In this process the input data are transformed between interval 0 to 1 just that the network will be able to easily learn the pattern and generate better results.

$$f(\sigma) = \frac{1}{1+e^{-\sigma}} \tag{2}$$

$$\frac{\partial f}{\partial \sigma} = f(\sigma)[1 - f(\sigma)] \tag{3}$$

Vector of weights is usually at the beginning of learning process initialized with random values from interval between 0 to 1. As a threshold function the most commonly used function is a sigmoid function (2) or hyperbolic sigmoid function. This function makes the network possible to learn nonlinearity between inputs and outputs and this function is also used because it is mathematically convenient (is differentiable (3)). According to the path of the information going through the network there are two basic network types. Feedforward network is the simplest ANN. All inputs to the layer are first multiplied by respective weights, summed together and applied to threshold function and makes output of the layer. Other type may be the ANN with feedback loop.

The main task of ANN is the learning process of respective weights. In this case supervised type of learning is briefly described. In supervised learning technique, the learning algorithm tries to min-

imize the mean square error between the inputs and already known outputs. For this purpose an error function is introduced (4) in accordance with [3]

$$E = \frac{1}{N}\sum_{i=1}^{N}(T_i - A_i)^2 \tag{4}$$

Where E is the sum of mean square errors for all inputs, T_i is the target output that should be achieved and A_i is the actual output that can differ from T_i and N is the number of samples. It can be derived that the error function depends on the weight vectors. According to the gradient descent algorithm the local/global minimum of such error function can be than calculated in iterative loop matter, when in each iteration step weight vector is updated with the negative proportion of gradient of the error function (5), where index k represents the k-th loop value and k-1 the previous loop value

$$\boldsymbol{w}_{i,k} = \boldsymbol{w}_{i,k-1} - \nabla \boldsymbol{E}_k = \boldsymbol{w}_{i,k-1} - \frac{\partial E_i}{\partial w_i} \tag{5}$$

To determine the accuracy of forecasting models the Absolute Percentage Error (APE) and its mean value over all samples (Mean Absolute Percentage Error MAPE) (6) was used. Root mean square error or others methods would also work. *MAPE* is given as

$$MAPE = \frac{1}{N}\cdot\sum_{i=1}^{N} APE_i = \frac{1}{N}\sum_{i=1}^{N}\frac{100\cdot|T_i - A_i|}{T_i} \tag{6}$$

The learning process of ANN is finished when error function reaches desired minimum value.

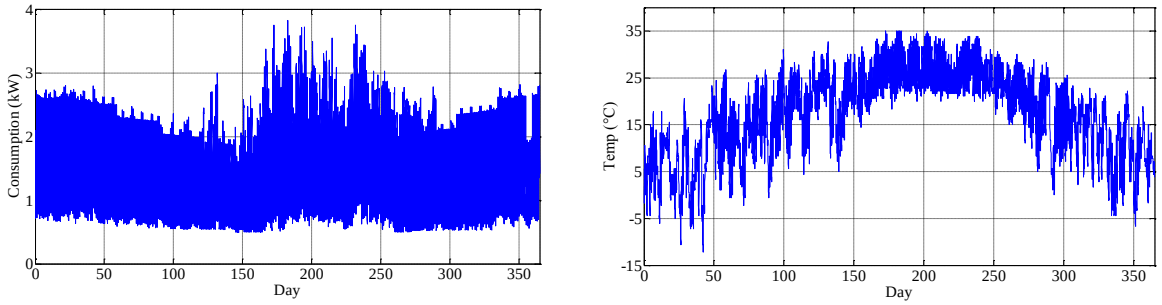


Figure 2: Annual consumption and temperature curve

3. EXPERIMENT DESCRIPTION AND RESULTS

In this section the results of week ahead load forecast by two previously described forecaster models are presented - Autoregressive and ANN. As the input data for the load forecast were used residential hourly load data simulated by EERE [5]. The simulation was performed for slightly more than one thousand locations in US and is based on statistical data for typical meteorological year (TMY3). The usage of energy, water, electricity etc. were considered according to the weather and Building America House Simulation Protocol [6]. The consumption curve with its temperature curve data are depicted for entire year in Figure 2.

To set the AR and ANN model forecasters the data was divided into three groups. The first group contained load data for the first 49 days of the year. This data starts with Sunday 1st of January and ends after 7 weeks on Saturday 18th of February. This data was considered for the learning purpose and consists of more than 11 hundred time samples. For the forecasting purpose another one week data was used starting from Sunday 19th of February and ending on 25th of February and represented 168 time samples to be forecasted. The real load data form 1st of January to 25th of February (i.e. learning data and one week data to forecast) are depicted in Figure 3. The rest of the data wasn't used. The coefficients a_i of autoregressive model AR (1) was estimated with the use of Matlab tool AR. The order of AR model was set to 25 according to empiric estimation and also with respect to autocorrelation function. This Matlab tool also allows to use different algorithms to calculate the least squares (Burg´s method, Yule-Walker method (YW), Forward-Backward method etc.). Ex-

cept YW method other gave almost the same results but bit better than YW, thus Forward-Backward method was chosen.

Model of ANN was obtained by using Matlab and its NN tool nftool. This tool enables to create feedforward neural network with input, output and one hidden layer. The number of neurons in hidden layer can be arbitrarily changed by the user. To avoid overfitting and to improve the performance of training, the data were divided into three categories-training, validation and testing. These categories were set as default (70 %, 15 %, 15 %) and its influence can be examined in further research. To minimize the square error (error function) Levenberq-Marquardt algorithm was used, which is similar but more robust algorithm than simple gradient descent. According to empirical experiments and also trial and error method the number of neurons in hidden layer was set to 15. For higher number of neurons the performance didn't improve or was even worse. Probably because of overfitting problem. As the inputs for the ANN were chosen day of week, logical information if its weekend or not, hour, outside temperature and dew point temperature (DP). As paper [3] proposes DP value should theoretically has impact on the load curve.

In Figure 4 and 5 the Absolute Percentage Error for fitted and forecasted data is depicted. From these figures can be also seen the quality of regression (fitting) because of both models being used also on learning dataset.

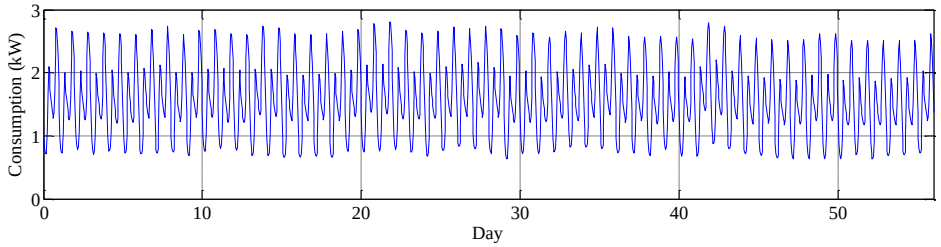


Figure 3: Real load data used for learning and forecasting

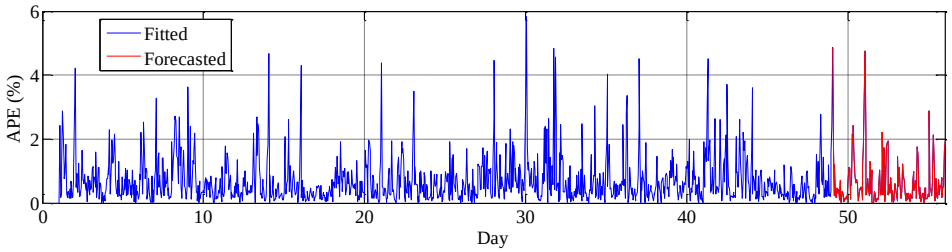


Figure 4: Distribution of Absolute Percentage Error (APE) among fitted and forecasted consumption for AR model

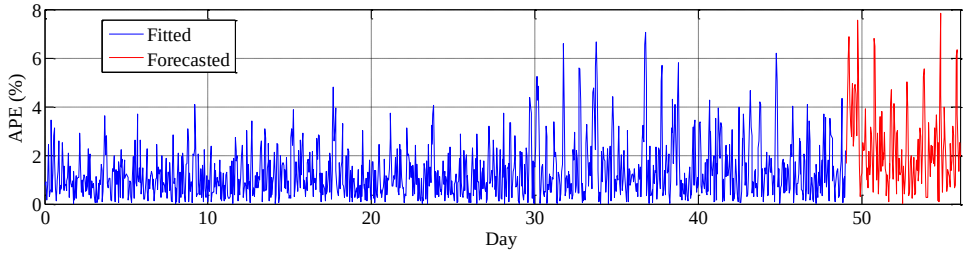
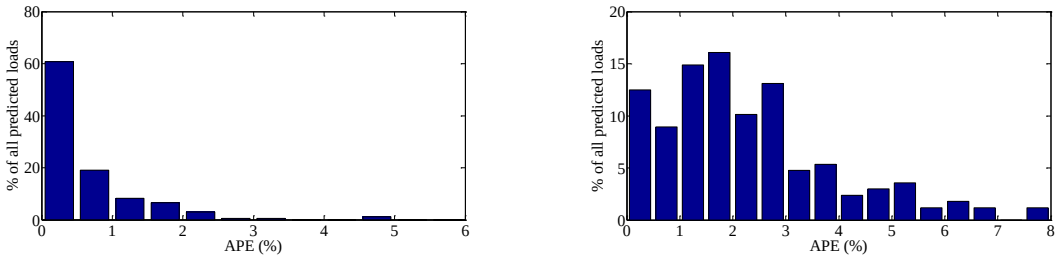


Figure 5: Distribution of Absolute Percentage Error (APE) among fitted and forecasted consumption for ANN model

For AR model the MAPE for forecasted values was 0.644 % whereas for ANN MAPE was 2.31 %. The maximum absolute percentage error for AR was 4.87 % and for ANN 7.84 %. The distribution of error among forecasted values for both methods is depicted in Figure 6. The MAPE was calculated only from data of 7 forecasted days, not including the fitted data (day 1-49 that are also ploted in Figures 4 and 5).



a) b)

Figure 6: Distribution of error among forecasted data, AR a), ANN b)

4. CONCLUSION

In this paper the suitability of AR and ANN models for electricity consumption was introduced. From obtained values of week ahead hourly load forecast, it can been seen that AR model achieves superior accuracy to ANN model. From the experiment it was also found that the AR model is quite simple to learn. ANN model is quite different but allows to consider other factors that can influence the forecasted load. Different input setups were tested, and as [3] proposes including DP temperature slightly improved the results reducing the maximal APE by 1% but not influencing overall MAPE. However to achieve better accuracy ANN model had to be trained more times to find best fit. For future development it is supposed to try other Autoregressive models and also for ANN extend the inputs by information of past load (i.e. the same day last year, same day last week).

ACKNOWLEDGEMENT

This research work has been carried out in the Centre for Research and Utilization of Renewable Energy (CVVOZE). Author gratefully acknowledges financial support from the Ministry of Education, Youth and Sports of the Czech Republic under NPU I programme (project No. LO1210).

REFERENCES

[1] Chow, J., Felix F. WU., Momoh, J. A. *Applied mathematics for restructured electric power systems: optimization, control, and computational intelligence*. New York: Springer, 2005. Power electronics and power systems (Springer). ISBN 03-872-3470-5.

[2] Singh, A. K., Khatoon, I. S., Muazzam, Md. *An Overview of Electricity Demand Forecasting Techniques. Journal on Network and Complex Systems* National Conference on Emerging Trends in Electrical, Instrumentation & Communication Engineering. 2013, 3(3), 38-48. ISSN 2225-0603.

[3] Raza, M. Q., Khosravi, A. *A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings*. *Renewable and Sustainable Energy Reviews*. 2015, 50, 1352-1372. DOI: 10.1016/j.rser.2015.04.065. ISSN 13640321.

[4] Kamel, N., Baharudin, Z. *Short term load forecast using Burg autoregressive technique*. *2007 International Conference on Intelligent and Advanced Systems*. IEEE, 2007, 912-916. DOI: 10.1109/ICIAS.2007.4658519.

[5] OFFICE OF ENERGY EFFICIENCY & RENEWABLE ENERGY (EERE). *Commercial and Residential Hourly Load Profiles for all TMY3 Locations in the United States*. 2012. Available: http://en.openei.org/datasets/files/961/pub/

[6] Hendron, B., Engebrecht, Ch., NATIONAL RENEWABLE ENERGY LABORATORY. *Building America House Simulation Protocols*. 2010. Available: http://www.nrel.gov/docs/fy11osti/49246.pdf

EXPERIMENTAL ANALYSIS OF LOW VOLTAGE ARC BETWEEN CARBON ELECTRODES WITH UNKNOWN CORES

David Šimek

Doctoral Degree Programme (1), FEEC BUT

E-mail: xsimek18@stud.feec.vutbr.cz

Supervised by: Vladimír Aubrecht

E-mail: aubrecht@feec.vutbr.cz

Abstract: The paper is focused on an electric arc spectroscopy diagnostics related to electric low voltage apparatuses. Measurement of radiation spectra of the electric arc burning between carbon electrodes with unknown cores is presented. Measured radiation spectrum has a lot of suitable atomic spectra lines for comparison purposes. The problems connected with the measurements are discussed. The main part of the paper deals with a prediction of composition of the measured arc plasma. Atomic lines database of National Institute of Standards and Technology was used as a spectral data source for the calculations.

Keywords: switching arc; arc plasma composition; atomic emission spectroscopy.

1. INTRODUCTION

Atomic emission spectroscopy is a powerful tool for a prediction of an unknown plasma composition. A hot plasma between electrodes radiates a lot of energy. If the plasma contains free atoms/ions, its radiation spectra consist of narrow energy peaks. These peaks are called spectral lines. If the atom/ion accepts amount of energy which is given by the difference between two neighbouring energy levels, then the electron can jump to the higher one. This phenomenon is called excitation. Excited state is unstable. If there is lower energy income than necessary to stay excited, electron moves spontaneously back to the lower energy level while irradiating energy specific to the radiation transfer between two energy levels. We can measure this radiation energy via spectrometer. There is two main methods of the spectra measurement. The first method which is called emission spectroscopy is based on self-radiation of plasma. On the other hand in the absorption spectroscopy is used supplementary source of radiation. The radiation of this source is measured after passing through measured object. Output data from spectrometer are emission or absorption spectra, see Fig 1.

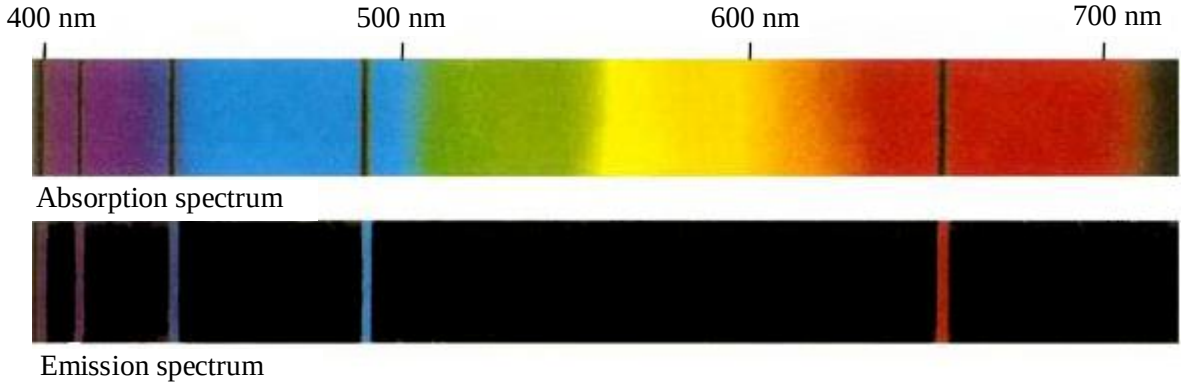


Figure 1: Absorption and emission spectra of hydrogen [1].

Atomic spectra can be described by Bohr model of atom and wave mechanics. Bohr used for his theory classic laws of motion and quantum conditions. Basic idea was placed by equation (1).

$$\Delta E = \frac{hc}{\lambda} \tag{1}$$

where h is Planck constant,
c speed of light
and λ wavelength [1, 2].

Atomic emission spectroscopy is more suitable for an electric arc spectra measurement. It allows calculation of a plasma composition, temperature and other physical properties, respectively. An advantage of this method is that the measured object is not affected by the measuring devices. Block diagram of measuring circuit is shown in Fig. 2.

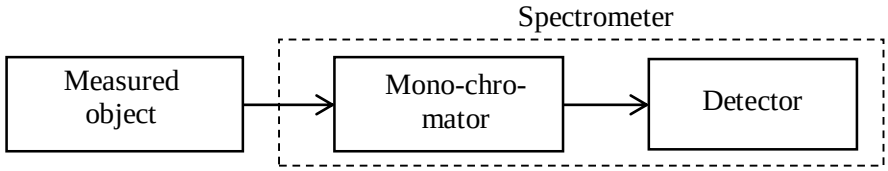


Figure 2: Block diagram of emission spectra measurement.

Spectrometer was used as measuring instrument. CCD chip is commonly used as a detector. Output from this measurement is a line spectrum which contains peaks of specific wavelengths. Peaks look like lines in global view. Shape of a spectral line is shown in Figure 3. In the Figure 3 there are labelled important properties of spectral line: maximum relative intensity of line (I_{max}), full width at half maximum ($\Delta\lambda$) and characteristic wavelength of line (λ_0).

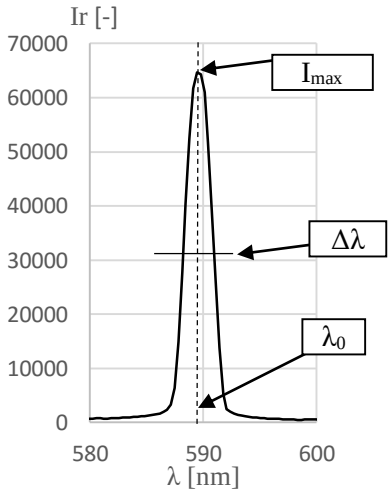


Figure 3: Spectral line shape [2].

2. EXPERIMENTAL SETUP

Measuring chain was arranged according to Fig. 4. Dynamo was used as DC power supply. Series stabilization resistor was used for the current control. Both current and voltage between electrodes were measured. Electric arc was burning between carbon electrodes with unknown stabilization

cores. The electrodes were specially treated by dipping in a salt solution because of detection of spectral shift. Load current was of 5.26 A, voltage on the arc of 56 V and power supply voltage was of 100 V. An arc was projected through set of lens to a shade. Optical fibre was used for necessary separation of measurement channel from the power circuit and transference of light from specific position to the input slit of the spectrometer. For the spectroscopy measurements we used Avantes Avaspec ULS-3648TEC spectrometer. The signal from the spectrometer was transferred via interface USB 2.0. Data were stored and processed using AvaSoft 7.7 software.

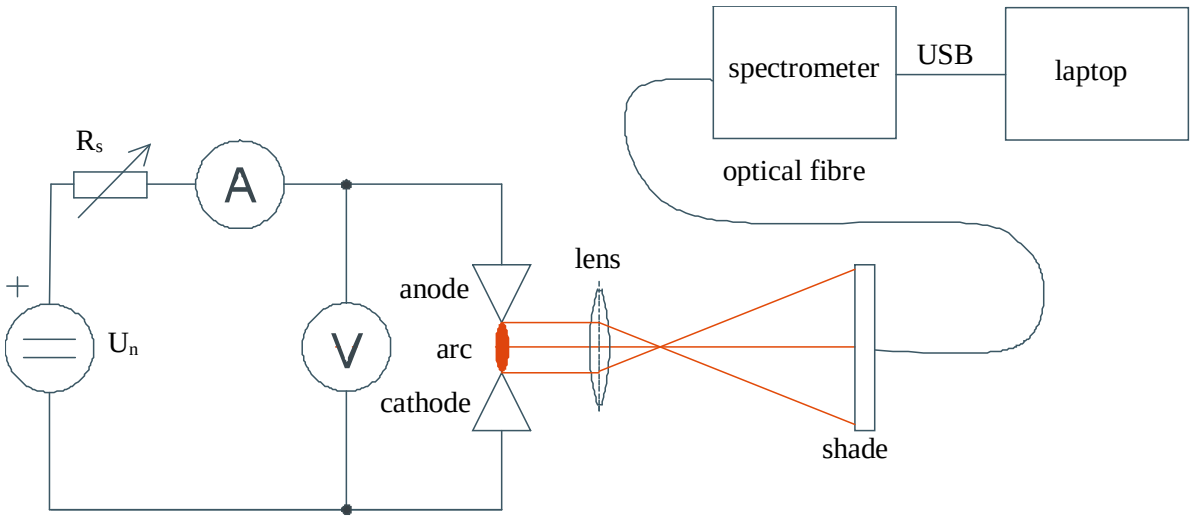


Figure 4: Simplified scheme of measurement chain [4].

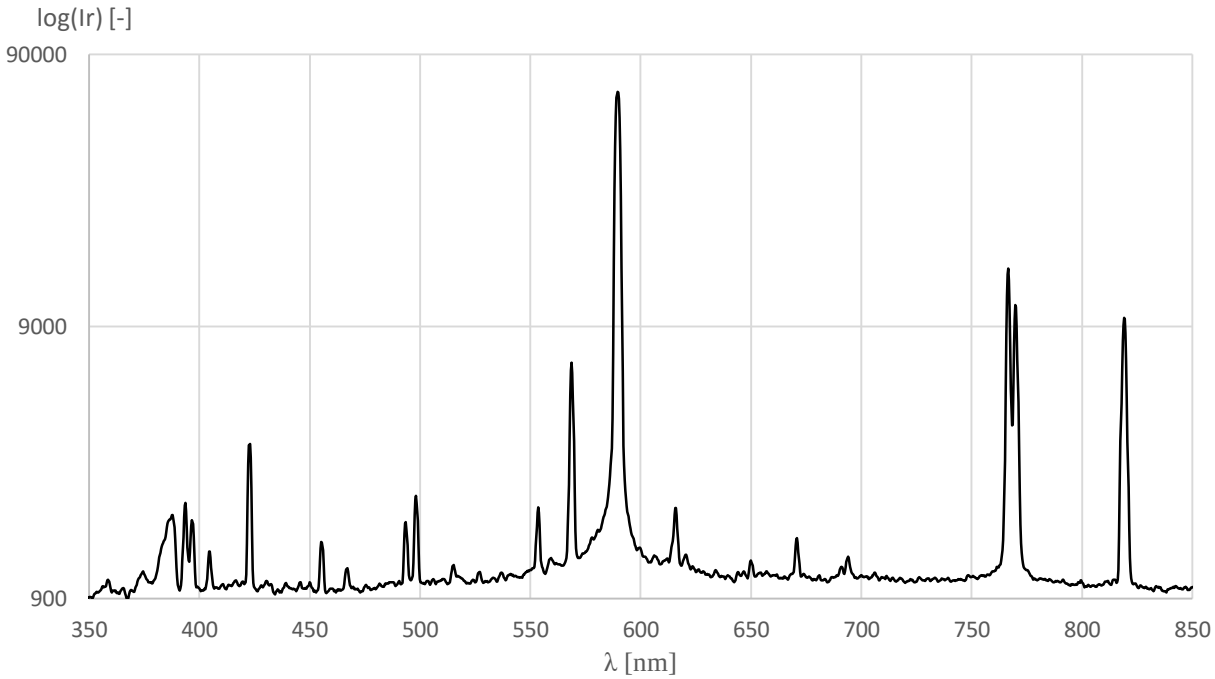


Figure 5: Radiation spectrum of the arc plasma between carbon electrodes.

The measured atomic emission spectrum of plasma is presented in Fig. 5. The plasma composition can be found by comparison of the measured spectral lines with spectral database NIST [3]. Elements assignment and corresponding wavelengths are shown in Table 1.

λ – measured	λ – database	assigned element
[nm]	[nm]	[-]
378.84 – 390.69	-	molecule of nitrogen
393.27	393.26	iron
396.56	396.55	iron
404.38	404.39	iron
422.47	422.45	iron
455.22	455.25	iron
493.28	493.21	carbon
498.02	498.17	titanium
553.40	553.43	carbon
568.39	568.65	iron
589.45	588.99	sodium
589.45	589.59	sodium
615.72	615.77	iron
766.42	766.43	iron
769.83	770.02	iodine
819.24	819.12	fluorine

Table 1: Comparison of measured wavelengths with spectral database NIST [3] and corresponding elements contained in the arc plasma.

3. RESULTS AND DISCUSSION

Results of elements assignment is presented in Table 1. If the arc burns in the air, then it contains molecular nitrogen N_2. Fluorine, iodine and sodium come from the salt, which was evaporated from the electrodes. Carbon corresponds to the material of electrodes.

It is important to mention that the identification of plasma composition may be affected by a few complications. The first one is a spectral lines shift. It depends on spectrometer resolution. We have found the shift of 0.8 nm via spectra measurement of known plasma. This shift is already presented in the Fig. 5.

The second problem corresponds to database using. There is a lot of nearby lines in the spectrum which is sometimes very difficult to distinguish.

Another complication is connected with database using again. In the database, there are presented spectral lines for neutral elements, but also for once, twice and triple ionised elements. Probability of occurrence of neutral species is very high. For once ionised species there is a probability of their occurrence much lower. Probability that the free burning arc with temperature below 10 000 K and without any foreign ionization source contains more than once ionised species is insignificant because of high density of air. Average free path of the electron is short to reach necessary energy due to the high density of air. Also the probability of collision between electron and ionised atom is insignificant.

The problem would be also in that the NIST database cannot cover all experiment conditions in which the spectral line are measured. That means that there are lines measured under different pressures, temperature etc.

4. CONCLUSION

Aim of the measurement was to find out composition of the electrode core. From the measurements follows that the core is probably made of pressed iron powder, titanium powder and carbon powder.

ACKNOWLEDGEMENT

Author gratefully acknowledges financial support from the Ministry of Education, Youth and Sports under project No FEKT S-14-2342.

REFERENCES

[1] FANTZ, U: Plasma Sources Science and Technology: Basics of plasma spectroscopy, [online], 2006. Vol. 15, no. 4p. S137-S147. [Accessed 08 March 2016]. DOI 10.1088/0963-0252/15/4/S01. Retrieved from: http://stacks.iop.org/0963-0252/15/i=4/a=S01?key=crossref.55c2687ff5694740c0db81c4b07cacd0

[2] REICHL, J.: Spektrum atomu vodíku, [online]. 2006. [Accessed 08 March 2016]. Retrieved from: http://fyzika.jreichl.com/main.article/view/752-spektrum-atomu-vodiku

[3] NIST: National Institute of Standards and Technology: Atomic Spectra Database Lines Form, [online], 2012. [Accessed 08 March 2016]. Retrieved from: http://physics.nist.gov/PhysRefData/ASD/lines_form.html

[4] ŠIMEK, D: Master Thesis: Působení spínacího oblouku na konstrukci výkonového přístroje nn, BUT, 2015.

FINITE ELEMENT ANALYSIS OF HIGH SPEED INDUCTION MOTOR WITH AXIALLY SLITTED ROTOR

Jiří Klíma

Doctoral Degree Programme (2), FEEC BUT

E-mail: xklima14@stud.feec.vutbr.cz

Supervised by: Ondřej Vítek

E-mail: *viteko@feec.vutbr.cz*

Abstract: This paper deals with an analysis of three phase, two pole high speed induction machine with axially slitted solid rotor. The finite element model of induction motor with axial slits has been developed and motor has been analyzed for several slips in order to get performance characteristics. Setting the simulation is crucial and will be discussed. All-important fields and performance characteristics are presented.

Keywords: high-speed, induction motor, solid rotor, finite element method

1. INTRODUCTION

Nowadays, high-speed solid rotor motors have found its place in the market. In order to boost power density or reduce gear boxes, high speed motors supply from frequency inverter are used. Increasing shaft rotating speed by higher input supply frequency leads to better shaft power with the same active motor volume [4]. Unfortunately, increasing shaft rotating speed brings motor to a certain limitation. Mechanical limitation is caused by centrifugal force and therefore solid rotor, made of one piece of steel is used. There are several feature modification in rotor construction, see Fig. 1. Very popular modification is copper coated rotor (see. Fig 1a), where the rotor yoke is coated with thin layer of copper or other high conductivity material. Our subject of interest is axially slitted solid rotor (see Fig. 1c). Solid rotors, especially smooth (see Fig. 1b), are characterized with low power factor and great eddy currents losses due higher harmonics. There are several ways how to reduce these harmonics. One of them is increasing the air gap or use a ferromagnetic wedges with small relative permeability [1].

Unlike classical induction motor, smooth solid rotor cannot be analyzed classical numerical method due to separated electrical and magnetical path. Calculation method describe in [1] is called multilayer transfer matrix (MLTM) method, where the idea is that the transfer matrix calculates the strength of the field in each thin metal layer from the center to the surface of the rotor. Smooth solid rotor (Fig. 1b) can be easily analyzed with finite element approach, but it takes a lot of time to determine the working characteristics. It was shown in many studies [1], [4] and [5] that magnetic flux, due to magnetic skin effect, do not penetrate very deep towards the center of the rotor. Current density is mostly concentrated on the surface as well. Therefore, this concept has a large rotor resistance and need a great slip to produce required torque.



a) b) c)

Figure 1: Solid rotor topologies: (a) Coated solid rotor, (b) smooth solid rotor, (c) slitted solid rotor

One way, how to improve efficiency is by slitting rotor surface in axial direction. This method will cause better penetration of magnetic flux and electrical current density. The depth of the slits should be choose wisely with respect to mechanical toughness at the bottom of the rotor teeth. As a matter of fact, the depth of the slit is trade-off between electromagnetic and mechanical design [2], [4]. The width of the slits is important as well. The narrower slit is, the less saturated teeth will be and better performance can be obtained.

2. SIMULATION OF HIGH SPEED INDUCTION MOTOR

Presented motor have 24 stator slots and 28 rotor teeth. Material with label M235-35A was used for stator sheets and its BH curve for frequency f =1 kHz and f =500 Hz was defined from datasheet up to B =1.3 T. Rotor is made of one piece of steel, labeled S355J2G3. BH curve of rotor steel was measured at the low frequency of 15 *Hz*. The steel conductivity σ =2.72 MS/m at temperature 150 °*C*, was defined in model. Main design and rated parameters are presented in table 1.

Machine rated and design parameters	Value	Unit
Line to line voltage	400	[V]
Rated power	12 000	[W]
Air gap length	1	[mm]
Rated frequency	750	[Hz]
Number of stator slots/rotor slits	24/28	[-]
Stator outer/bore diametr	120/65	[mm]
Width/depth of rotor slit	1/12	[mm]
Rotor stack & one steel end ring length	80 & 35	[mm]

Table 1: Main Motor parameters

Mesh and sampling time are very important aspects. Mesh and sampling time have to be in balance between computing time and accuracy. Motor is supply by frequency of f = 750 Hz. Our sampling time, to provide sufficient results, was set to be T_s = 2.66 μs. The sampling time was set intentionally to make N = 500 samples during one period and include higher harmonics. The mesh was dense in area of air gap and motor was analyzed with more than 40000 mesh elements.

2.1. END-EFFECT CORRECTION FACTOR

Rotor modification by axial slits leads to better performance. The assumption is that all electrical current conduct in the teeth and therefore resistance and leakage inductance of the rotor end rings is extra calculated and added in the model. The resistance R_{kn} between two teeth is calculate by follow equation.

$$R_{kn} = \rho_{Fe150} \cdot \frac{\pi \cdot D_{kn}}{Q_2 \cdot S_{kn}} = \frac{1}{2.88 \cdot 10^6} \cdot \frac{\pi \cdot 51 \cdot 10^{-3}}{28 \cdot 4.2 \cdot 10^{-4}} = 4.73 \cdot 10^{-6} \ \ \Omega \tag{1}$$

Where D_{kn} is average diameter of steel ring, Q_2 is number of rotor teeth, S_{kn} is cross area of steel end-ring and ρ_{Fe150} is specific resistivity of solid steel.

The leakage reactance was calculate by well-known equation from [3] and it include magnetic conductivity coefficient of rotor end-rings which are directly attached to rotor stack.

$$\lambda_{ei} = \frac{2{,}3 \cdot D_{kn}}{4 \cdot Q_2 \cdot l_{fe} \cdot \sin^2 \cdot \left(\frac{\pi \cdot p}{Q_2} \right)} \cdot \log \frac{4{,}7 \cdot D_{kn}}{2 \cdot a_{kn} + b_{kn}} = 0.636 \tag{2}$$

Where l_{Fe} is stator stack, a_{kn} is length of rotor end ring in radial direction and b_{kn} in axial direction. Magnetic conductivity coefficient is then divided by two for one end-ring and leakage inductance of one end-ring between two teeth is calculated, L_r=31.98 nH. The conductivity of rotor core was

reduce by Russel correction factor, described in []. The final value of rotor conductivity material is σ =1.49 MS/m

3. FIELD RESULTS

Calculated field distribution for axially slitted rotor at 2% slip are shown in Fig.2 and Fig.3. It can be seen from Fig. 3 that current density and eddy current loss are mainly concentrated on the edge of rotor teeth, under the surface. Fig.4a and 4b shows flux density distribution. The fundamental harmonic has value of 0.6 T.

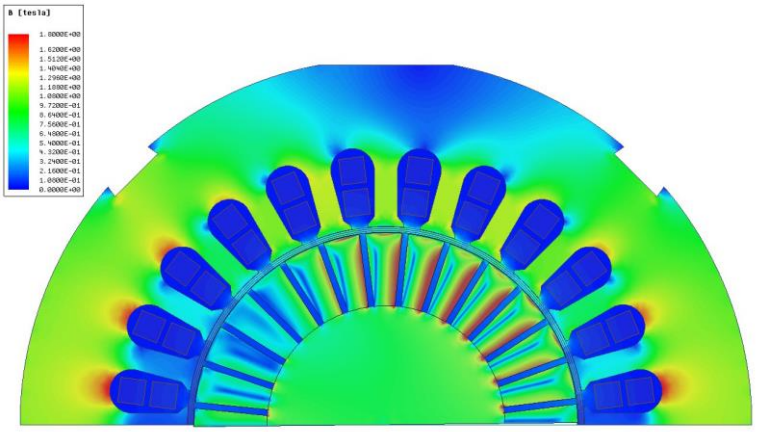
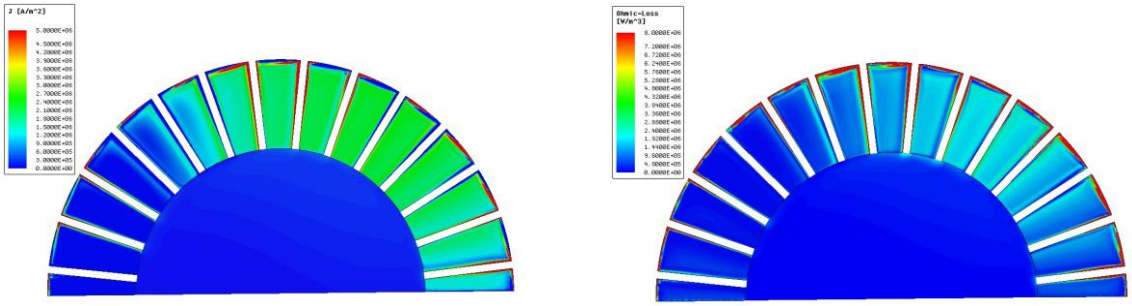
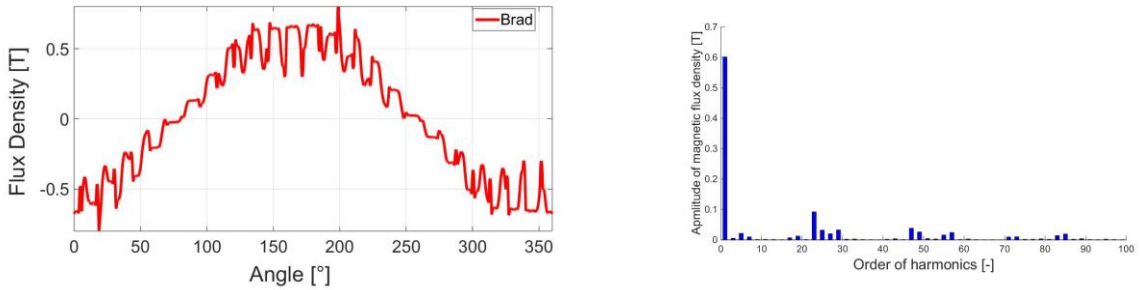


Figure 2: Flux density distribution in cross-section area of axially slitted rotor at 2% slip



a) b)

Figure 3: Density distribution in cross-section area at 2% slip a) Current density b) Eddy current loss

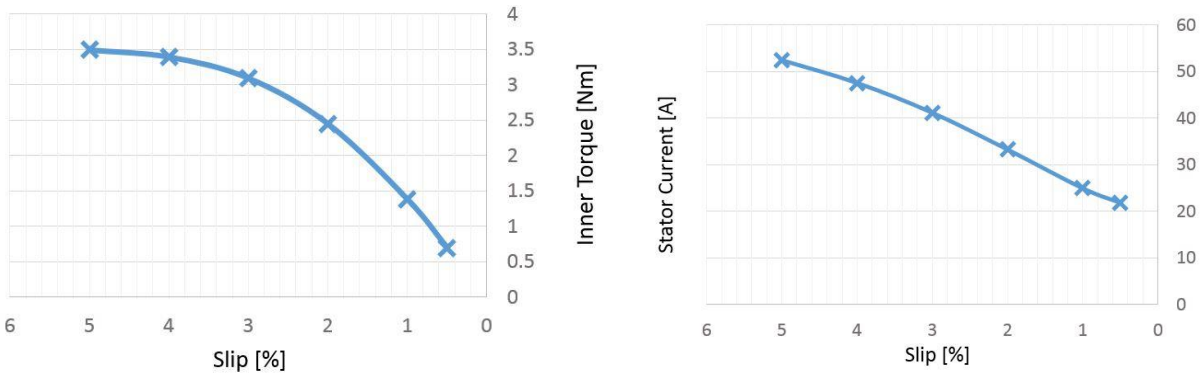


a) b)

Figure 4: Air-Gap flux density distribution a) with respect to angular position b) its harmonic spectrum, at 2% slip

4. PERFORMANCE CHARACTERISTICS

Performance characteristics, can be seen in Fig.5 and Fig 6, represent motor behavior at different slip. Fig.5a and 5b shows the inner torque and stator input current of motor in dependence on rotor slip. Fig. 6a represent values of stator and rotor losses with respect of rotor slip.



a) b)

Figure 5: Motor performance characteristics a) torque vs. slip characteristics b) current vs. slip characteristics

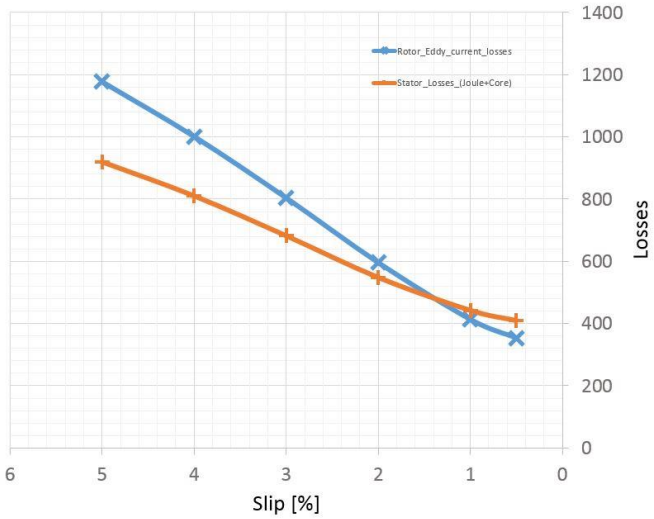


Figure 6: Stator and rotor loss vs. slip characteristics

5. CONCLUSION

Performance motor characteristic and field distribution has been demonstrated. From field distribution, it can be seen that, eddy current loss are mainly concentrated at rotor surface. Due to, feature modification, the magnetic flux density penetrates deeper than in solid smooth rotor. The analysis by finite element method in ANSYS Maxwell software has shown that axially slitted rotor can be easily examine and therefore it can be start point for other design aspects, for example mechanical and thermal analysis. Our motor is designed to have output power 12kW. Operating point will be located slightly above 2% rotor slip. Mechanical losses caused by friction and ventilation were not included in motor characteristics.

ACKNOWLEDGMENT

This research work has been carried out in the Centre for Research and Utilization of Renewable Energy (CVVOZE). Authors gratefully acknowledge financial support from the Ministry of Education, Youth and Sports of the Czech Republic under NPU I programme (project No. LO1210).

REFERENCE

[1] HUPPUNEN, J.: *High-speed solid-rotor induction machine - electromagnetic calculation and design*, PhD thesis, Lappeenranta University of Technology, Finland, 2004.

[2] T. AHO, J. NERG, and J. PYRHÖNEN, *Experimental and finite element analysis of solid rotor end effects*, in IEEE Int. Symp. Industrial Electronics, Vigo, Spain, 2007, pp. 1242–1247

[3] KOPYLOV, I.P. *Stavba elektrických strojů*, celostátní vysokoškolská učebnice pro elektrotechnické fakulty vysokých škol technických. 1. vyd. Praha 1: Státní nakladatelství technické literatury, 1988, 685 s. ISBN 04-532-88

[4] UZHEGOV, N. *High-speed solid-rotor induction machine – calculation program*. Lappeenranta University of technology, Helsinki, 2012. Master's Thesis. Supervised by Professor Juha Pyrhönen and Dr. Janne Nerg.

[5] PYRHÖNEN J. *The high speed induction motor: Calculating the effects of solid rotor material on machine characteristics*. Acta Polytechnica Scandinavica Electrical Engineering Series no. 68. Diss. Lappeenranta University of Technology, Helsinki, 1991, 84 p., ISBN 951-666-332-X.

[6] Maxwell 2D Online Help; ANSYS Inc, 628p

FREQUENCY CHARACTERISTICS OF THE INDUCTION MACHINE

Martin Kroupa

Doctoral Degree Programme (1), FEEC BUT

E-mail: xkroup07@stud.feec.vutbr.cz

Supervised by: Čestmír Ondrůšek

E-mail:ondrusek@feec.vutbr.cz

Abstract: The paper focuses on the induction machine as the transfer member and on its behavior in the frequency domain. The linearized mathematical model of induction machine is introduced and the issue of its frequency transmissions is discussed. Finally, the appropriateness of the motor use as diagnostic tool for analysis of time-varying load torque is mentioned.

Keywords: Induction machine, transmission function, frequency characteristic, control theory

1. INTRODUCTION

Since the 80s of last century, the issue of dealing with the diagnosis of electromechanical objects is very often discussed topic. Each electromechanical system usually consists of a drive (electric motor) and the driven load. Driven load can be for example: fun, hydrodynamic pump, screw compressor and other. The task of diagnosis is evaluation the behavior of one subsystem from analysis of second subsystem. In most cases, this means evaluating driven load by the analysis of the electric motor quantities.

At present, the major phenomenon of diagnostic method is analysis named Motor Current Signature Analysis (MCSA). This method commonly allows to perform the diagnosis of electric motor or diagnosis of driven load today. There are already a lot of literature in which this diagnostic methodology have been published and papers which this method use for functional diagnosis. However, almost any literature does not view on the motor as a part of a larger system, as on "the coupling element" which has its inputs (power grid voltage, load torque), "outputs" (stator and rotor currents, electromagnetic torque) and its characteristic behavior in the frequency domain. The characteristic behavior affects produced variables and input quantities of motor. Therefore, the aim of this article is to introduce the behavior of the motor as the transmission element.

From the point of view of simple construction and high efficiency, the majority of all installed motors are induction machines (IM) today, this is why attention is paid to the induction motor in this paper.

2. INDUCTION MACHINE FROM THE PERSPECTIVE OF CONTROL THEORY

2.1. MATHEMATICAL MODEL OF INDUCTION MACHINE

The induction machine is multi parametric nonlinear dynamic system which can be described by a differential equation of order *n*. For the numerical solution, the differential equation is often converted on *n* differential equations of the first order – into state equations. These equations can be obtained from the two-phase equivalent mathematical model of IM based on the theory of general machine. All derivatives are converted to one side of the equation and dependent variables

are eliminated using mathematical adjustments. The state equations for mathematical description of squirrel cage IM in coordinate system rotating angular velocity – ω are [1]:

$$\frac{d\psi_{Rd}}{dt} = -\frac{R_r}{L_r}\cdot\psi_{Rd} + (\omega-\omega_e)\cdot\psi_{Rq} + \frac{L_m\cdot R_r}{L_r}\cdot i_{sd} \tag{1}$$

$$\frac{d\psi_{Rq}}{dt} = -(\omega-\omega_e)\cdot\psi_{Rd} - \frac{R_r}{L_r}\cdot\psi_{Rq} + \frac{L_m\cdot R_r}{L_r}\cdot i_{sq} \tag{2}$$

$$\begin{aligned}\frac{di_{sd}}{dt} = {} & \frac{L_m\cdot R_r}{L_r\cdot\left(L_s\cdot L_r - {L_m}^2\right)}\cdot\psi_{Rd} + \frac{L_m}{L_s\cdot L_r - {L_m}^2}\cdot\omega_e\cdot\psi_{Rq} - \\ & -\frac{{L_r}^2\cdot R_s + {L_m}^2\cdot R_r}{L_r\cdot\left(L_s\cdot L_r - {L_m}^2\right)}\cdot i_{sd} + \omega\cdot i_{sq} + \frac{L_r}{L_s\cdot L_r - {L_m}^2}\cdot u_{sd}\end{aligned} \tag{3}$$

$$\begin{aligned}\frac{di_{sq}}{dt} = {} & -\frac{L_m}{L_s\cdot L_r - {L_m}^2}\cdot\omega_e\cdot\psi_{Rd} + \frac{L_m\cdot R_r}{L_r\cdot\left(L_s\cdot L_r - {L_m}^2\right)}\cdot\psi_{Rq} - \\ & -\omega\cdot i_{sd} - \frac{{L_r}^2\cdot R_s + {L_m}^2\cdot R_r}{L_r\cdot\left(L_s\cdot L_r - {L_m}^2\right)}\cdot i_{sq} + \frac{L_r}{L_s\cdot L_r - {L_m}^2}\cdot u_{sd}\end{aligned} \tag{4}$$

$$\frac{d\omega_e}{dt} = \frac{3}{2}\cdot\frac{L_m}{L_r}\cdot\frac{{p_p}^2}{J}\cdot\left(\psi_{Rd}\cdot i_{sq} - \psi_{Rd}\cdot i_{sq}\right) - \frac{p_p}{J}\cdot M_z \tag{5}$$

Where R_s and R_r are resistances of the stator and rotor circuit, L_s a L_r are total inductances of the stator and rotor circuit. L_m represents the magnetizing inductance. p_p – number of pole pairs and J is moment of inertia. The parameters are listed in the documentation for the motor. ω_e is mechanical angular velocity in electrical degrees and M_z represents load torque. Expressions on the left sides of equations are derivatives of state variables – rotor and stator flux linkage and rotor and stator current. Finally, equations still contain supply voltage – u_{sd}, u_{sq}. As shown for a complete description of IM in the state space, (five equations after distribution to the „*d*“ and „*q*“). In matrix form, the mentioned system of five differential equations can be written as Eq. (6).

$$\dot{\bar{x}} = \boldsymbol{A}\bar{x} + \boldsymbol{B}\bar{u} \tag{6}$$

$$\overline{y} = \boldsymbol{C}\bar{x} + \boldsymbol{D}\bar{u} \tag{7}$$

Where vector $\bar{x}$ is state vector, $\boldsymbol{A}$ is characteristic matrix of systems which determines behavior of the entire system. For stable systems, all eigenvalues of matrix are negative. In the case of complex numbers, eigenvalues have to be with negative real part. Matrix $\boldsymbol{B}$ is input matrix of system, $\boldsymbol{C}$ is output matrix of system and matrix $\boldsymbol{D}$ expresses direct effect of inputs to output variables, $\bar{u}$ is vector of variables which enter to system and $\overline{y}$ is vector of output variables. This vector is often identical with the state vector. Note: Matrix $\boldsymbol{D}$ is zero for evaluation of motor frequency characteristics.[2]

2.2. LINEARIZED MATHEMATICAL MODEL OF INDUCTION MACHINE

Like most real systems, IM is nonlinear system and its dynamic behavior changes with increasing loads or slip. In order to the motor could be correctly described, range of speed will be divided into several intervals – from zero load to rated load. The linear behavior of IM is supposed at these intervals and with small changes of system variables. This is a local approximative linearization.

If the system properties are studied from linearized model - in certain steady state, it will have to take in mind the local validity and the possibility of major changes by the larger deviations of the system variables from the "operating point". [3] The local linearization is carried out by so called

„perturbation variable". Then the system of linearized differential state equations in matrix form can be written as [2]:

$$\dot{\bar{x}} + \Delta\dot{\bar{x}} = \boldsymbol{A}(\bar{x} + \Delta\bar{x}) + \boldsymbol{B}(\bar{u} + \Delta\bar{u}) \tag{8}$$

$$\overline{y} = \boldsymbol{C}\bar{x} + \boldsymbol{D}\bar{u} \tag{9}$$

It means that all state and input variables that are not constant and are able to be changed are replaced by constant characterizing the linearized point and perturbation variable (labeled with „Δ").

3. FREQUENCY CHARACTERISTICS OF INDUCTION MACHINE

It is evident that the frequency characteristics of IM are dependent on the parameters of the equivalent mathematical model. Therefore, each motor used as a diagnostic „tool" may have different frequency characteristics. The characteristics may be different in: cutoff frequency, resonance amplification, number of resonance amplification and other. In this paper, the frequency characteristics of 3 kW IM will be investigated. The parameters and some values of IM are listed in the Tab. 1.

$P = 3$ kW	
R_s	1.993 Ω
R_r'	1.735 Ω
L_s	0.01134 H
L_r'	0.01134 H
L_m	0.21 H
p_p	2
J	0.0062 kg/m^2
M_n	19.967 Nm
n_n	1437 min^{-1}
U_n	400 V

Table 1: The parameters of IM

It is possible to compile 18 combinations of motor transfer functions from equations of state. However in the article, only stated transmissions of motor are investigated, concretely:

$$F_1 = \frac{i_{sd}}{m_z} \quad ; \quad F_2 = \frac{i_{sq}}{m_z} \quad ; \quad F_3 = \frac{\omega_{mec\,h}}{m_z} \quad ; \quad F_1 = \frac{m_i}{m_z} \tag{10}$$

The transfer functions are chosen in correlation with the use of the motor for detecting the load torque oscillation. Another interesting transfer functions are dependent on the voltage change – u_{sd}, u_{sq}.

3.1. RESULTS

The frequency characteristics of the selected IM are drawn for zero load motor and for its rated load. Of course in practice, it is necessary to know all the characteristics at any load. (It should be noted that the frequency characteristics in this article provide information on the dynamic changes of the examined variables. Medium values of variables respecting the operating point of the motor appear to be zero. This fact is based on Eq. (8) and Eq (9), where derivatives of constants by time are zero.

On Fig. 1 there are frequency characteristics and time responses to a step change for four variables of motor – i_{sd}, i_{sq}, ω_{mech} a m_i. The frequency characteristics of all the studied variables are related to

a step change in load torque, therefore three of the named variables begin with negative amplification. It is clearly seen from the diagrams that the motor behaves as a low pass filter with a cutoff frequency of about 28 Hz. The amplification of all values fall sharply after this frequency. The cutoff frequency is exactly 28.3 Hz for zero load motor, but after a load up to the rated torque, the cutoff frequency decreases slightly to about 27.2 Hz. In detail, this fact is shown in Fig. 2. The frequency characteristics of motor variables include the resonant increase at the cutoff frequency. This fact is really unpleasant and entails a certain risk. Not only that the stator currents will be greatly distorted at frequencies around 28 Hz, but the modulations of stator currents will achieve uncomfortably high levels. This can significantly affect machine life. Therefore, this has to be taken into account.

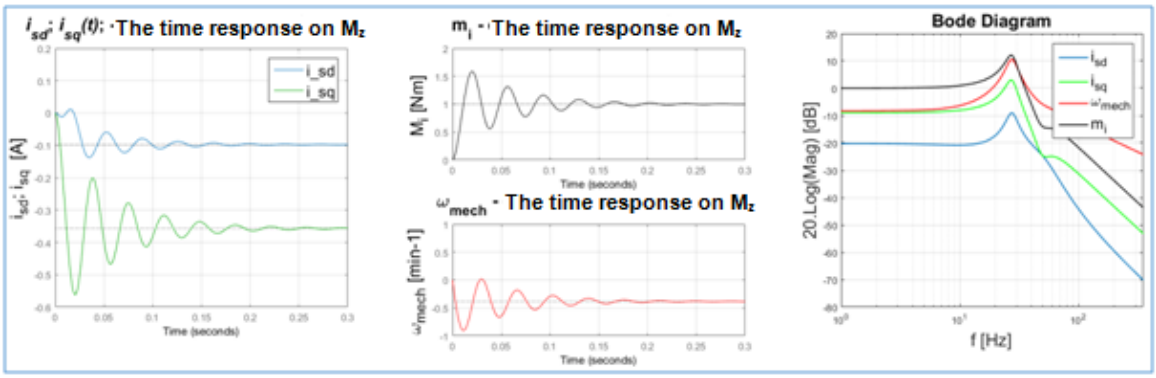


Figure 1: Motor response to a step change M_z [M_z = (0.0, 1.0) M_n]

To complement, the dependence of resonant increase of motor variables is showed on the following figure. The cutoff frequency of all motor variables decreases due to the average load.

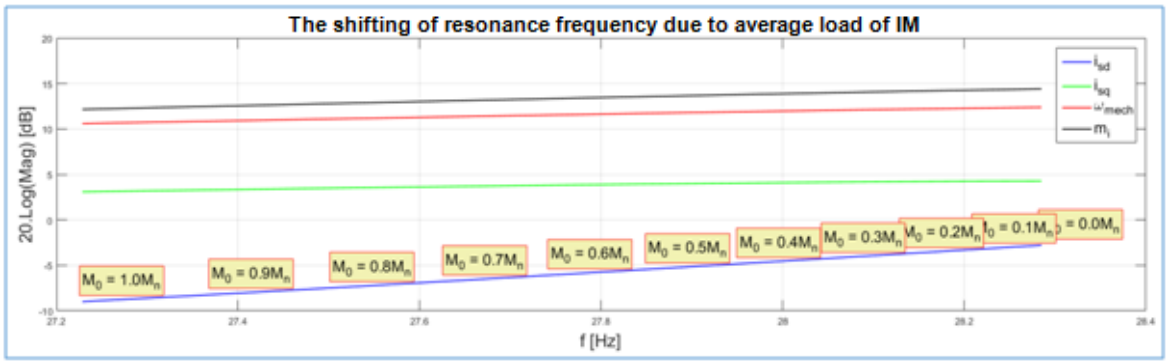


Figure 2: The shifting of resonance frequency due to average load

Frequency transfers of IM are dependent on all the parameters of the equivalent mathematical model, the number of poles and the moment of inertia. The parameters of equivalent motor model are constant after neglecting the temperature dependences, notably winding resistance, neglecting of the magnetizing inductance dependence on the average load and assuming steady state motor operation. All frequency characteristics of the motor were created for these simplifications. The only thing that may change is the moment of inertia. From the electrical point of view, the moment of inertia forms with resistance and inductance of IM analogy resonant RLC circuit. In fact, the moment of inertia is naturally bigger after connecting the load. It must be calculated with this for the purpose of diagnosis and for increased accuracy of frequency characteristics.

This fact is visible on Fig. 3. The diagrams are created for rated load but the moment of inertia of the rotating masses is changed. This time, the amplifications of motor quantities do not relate to 1 Nm. On the left, the moment of inertia is the same as moment of inertia specified in the catalog for the selected motor. Toward the right, the moment of inertia is raised up on two times greater value than rated one. The cutoff frequency is slightly shifted downward from the initial 27.2 Hz to about

23 Hz and pass band of IM was restricted even more. More importantly, the resonance peaks don't reached already such amplification and the effects of variable load torque in all investigated motor variables are mitigated at cutoff frequency.

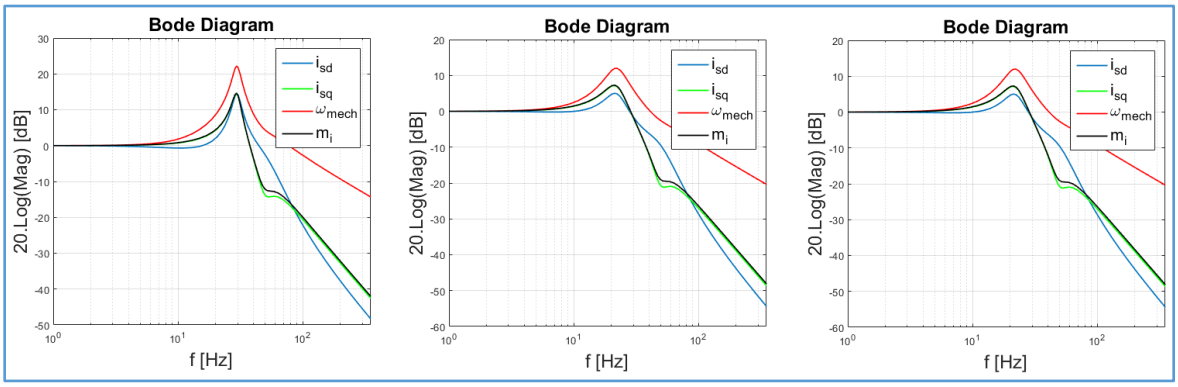


Figure 3: The frequency characteristics of 3 kW IM [J = (1.0, 1.5, 2.0) J_n]

4. CONCLUSION

The article introduced induction machine from a different perspective. Its characteristics in the frequency domain were described in detail, namely the behavior of the motor at a time variable load. The study foresees a further extension, where it will be examined the behavior of IM depending on supply voltage variation and the properties of several motors similar type of performance will be compared.

REFERENCES

[1] ALEXANDROVITZ, A. a S. LECHTMAN. *Dynamic Behaviour of Induction Motor Based on Transfer-Function Approach*. In: 17th Convention of Electrical and Electronics Engineers in Israel [online]. 1991 [cit. 2015-04-07]. DOI: 10.1109/eeis.1991.217702.

[2] SKALICKÝ, Jiří. *Teorie řízení*. vyd. Brno: FAKULTA ELEKTRONIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ, 2002, 113 s.

[3] MELICHAR, Jiří. *Lineární systémy 1*. vyd. ZČU Plzeň: FAKULTA APLIKOVANÝCH VĚD, 2010,124 s.

LABORATORY WORKBENCH FOR EDUCATIONAL PURPOSES

Vladimír Ondrejček

Doctoral Degree Programme (1),FEEC BUT

E-mail: xondre03@stud.feec.vutbr.cz

Supervised by: Pavel Vorel

E-mail: vorel@feec.vutbr.cz

Abstract: This work deals with the development and implementation of new laboratory device for laboratory exercises in electric drives. Reasons and requirements for this workbench are evaluated. In the next section a prototype design of laboratory stand is introduced. Finally the possibilities of extending and use in other subjects and applications are given.

Keywords: EEICT, power stage, control unit, driver, electric drives, laboratory stand

1. INTRODUCTION

The problem nowadays is reduction of the amount of practical training. This situation leads to the decrease of general practical knowledge and skills. Low level of knowledge impairs future jobs at companies and industries. For this reason it is necessary to increase volume of practical and laboratory teaching. This is in contradiction with the requirement to reduce the cost of teaching. It follows that to ensure quality teaching in sufficient numbers, demands on teachers and on laboratory equipment must be increased. One possible solution to this problem is the introduction of a parallel teaching in laboratory exercises. In parallel learning is it possible to conduct laboratory work with one teacher and students work individually at identical workplaces.

These facts lead to the development and implementation of a unified workplace for laboratory exercises in large numbers, which must be as universal and cheap as possible. The optimal solution is the following laboratory workbench, which will be presented in this article.

2. REQUIREMENTS OF LABORATORY STAND

Laboratory workplace for educational purposes will be operated mainly by students. This fact leads to the definition of basic workplace requirements. The most important requirement is safety of the device. Due to this a safe voltage level was chosen to supply the laboratory stand. Another requirement is to prevent unauthorized and illegal operations. Unauthorized manipulations could result in damage or complete destruction of a part of the workplace. In laboratory practice, this is often a result of improper handling by students. For this reason, it is necessary that the entire laboratory workbench must be very durable. In the event of damage or failure of any component, it is necessary to replace the faulty part and repair it. These requirements greatly facilitate unification of the laboratory stand. Unification of parts allows simplifying the design of the entire workplace in a compact element and extends variability of device. An example is the use of processor board, which is already used in more applications and projects.

Because it is a workplace to support teaching in laboratories of electric drives, entire device must be made clearly. Students can gain more knowledge and experience in the design of electric drive - electric engine with the inverter and control unit as a whole device. Laboratory stand for support of education must be designed to meet all requirements of laboratory exercises and should work in two basic modes.

- In bachelor's degree program, students discover the basic parameters of electrical machines. Students also perform measurements on a DC motor and design current controller and speed controller. A control panel is used for operation of the laboratory stand.
- In master's degree program, students in addition to previous tasks implement vector control of induction machine. Computer software Freemaster is used for operation of laboratory stand.

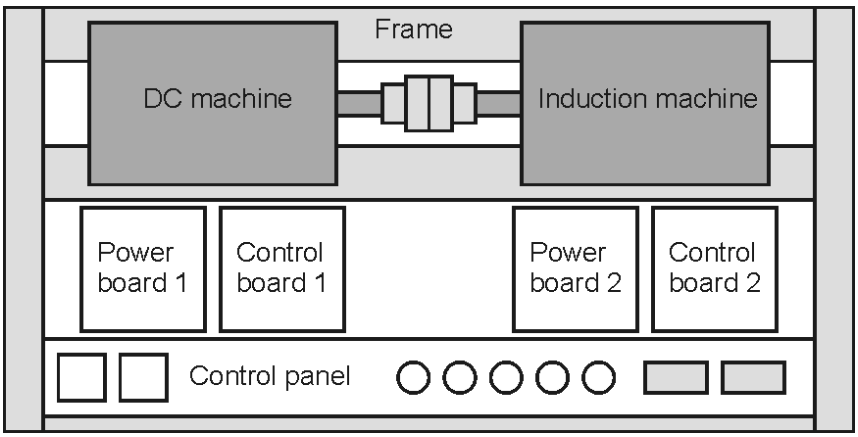


Fig. 1 Block diagram

3. PRACTICAL REALIZATION

This universal workbench is formed by an aluminum frame on which all of the components necessary for the function of the laboratory stand are mounted. There are mainly two machines and boards with power and control electronics. A machine set, consisting of a DC machine and an induction machine, is fixed on the frame. DC machine is designed for 24V with output power of 400W. Induction machine is designed for 3x24V with output power of 180W. The machines are connected by a flexible shaft coupling. A resolver for measuring speed and rotor position is mounted on the shaft of the induction machine. Both machines are supplied from a transistor inverter controlled by Freescale microcontrollers. The DC motor can be replaced by another type of machine which has the same mounting, such as a synchronous motor.

3.1. POWER STAGE

The power part of the laboratory stand is realized by two identical boards. Each board contains DC link capacitors and three-phase inverter from MOS-FET transistors produced by Semikron. For current measurements, the board has two sensors manufactured by LEM. These sensors are mounted in phase A and C of the three-phase inverter. Only one current sensor is used in the power stage which supplies the DC machine. If necessary, a second current sensor can be added. The power board can also be used for other applications. The board is designed for 100V and 80A.

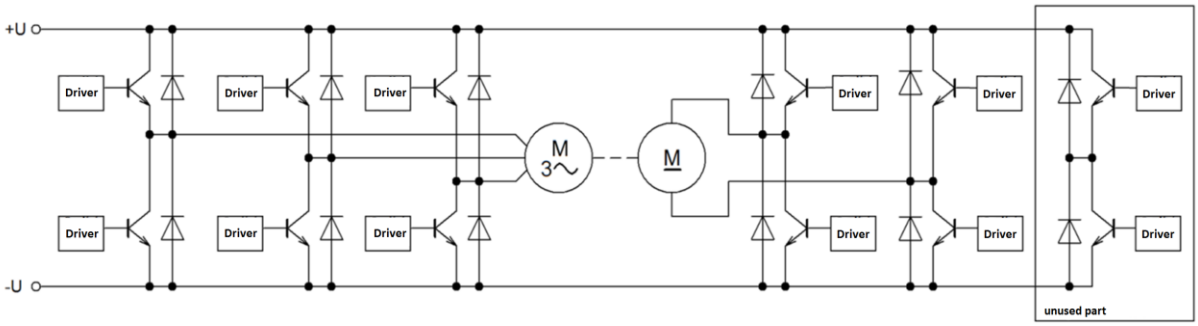


Fig. 2 Power stage

3.2. CONTROL UNIT

The control part is designed to control all the functions of the laboratory stand. It contains all the electronic components needed for proper operation of the entire workplace. Two power boards are used in the laboratory stand – for this reason, it is necessary to use two control units.

The control unit is composed of the following units:

- Processor unit
- Motherboard
- Drivers

Processor board is the main control part of the laboratory stand. The unit provides all the control and regulation commands to achieve the desired functions. The entire laboratory stand comprises a total of two units. Each unit is controlling one power stage, which is connected to the respective machine. The most important component of this unit is a processor produced by Freescale Semiconductor. This is a microcontroller, which includes all the circuits necessary for its operation, such as digital inputs / outputs for processing the buttons and switches installed on the control panel, analog converters and interface for serial communication. A circuit for communication over USB is on this board. A position (speed) signal of engines from resolver, which is installed on the shaft of the induction machine, is also coupled there. The most important part is the PWM outputs for controlling drives. LEDs are connected on PWM output signals for monitoring the activities of these signals, and the operation of the inverter control. For controlling basic functions, three switches are placed on the board. This unit is very versatile and fully programmable due to the design of the processor board.

The motherboard serves as the main interface unit for all board contained in a laboratory stand. It ensures the continuity of all devices that work together. The board produces the required logical dependencies. It includes power circuits for all control circuits and drivers. A circuit for position (speed) sensing from resolver is placed on the motherboard. This unit also provides measurement of analog signals - currents and voltages. To ensure high durability of the entire laboratory stand, there are circuits which can block PWM signals and handle fault signals. Additional circuits are used to evaluate other signals and connect these signals to the input of processor boards. Input and output connectors for communication between two control units, control panel and for encoder output are also located on the motherboard. The processor unit and MOS-FET drivers are mounted on this board.

To ensure transmission of control signals to the transistors, MOS-FET drivers are used. Driver provides galvanic isolation between control and power section. Driver performs switching of the transistor based on the control signals and the protection of the transistor using saturation protection. Driver must be very small, simple and reliable.

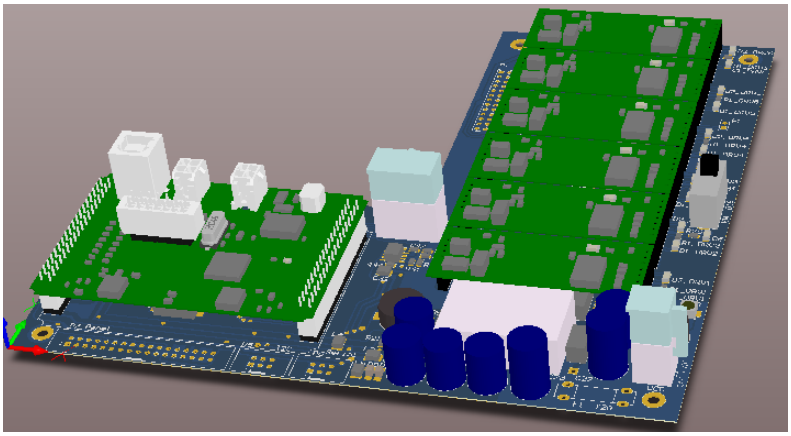


Fig. 3 Control unit

4. OTHER OPTIONS

Workbench for educational purposes can be extended or customized for another application. If necessary, the DC machine can be replaced by a synchronous machine. This modification leads to further laboratory experiments and use in the classroom. Prototype versions of the laboratory workbench will be completed by a control panel. The panel will be used to control the whole laboratory stand using buttons and switches. Analog signals can be entered by potentiometers. A four line alphanumeric display which displays all of the settings and the overall status of the device will placed on the control panel. In the future, the workbench will be completed by a computer using control program with a graphical user interface for easy operation, and which will replace oscilloscope.

Universal design of laboratory workbench allows using some components in other applications. Power stage with the control unit can be used as a compact device in various industrial applications and other projects, which are solved at department of power electrical and electronic engineering.

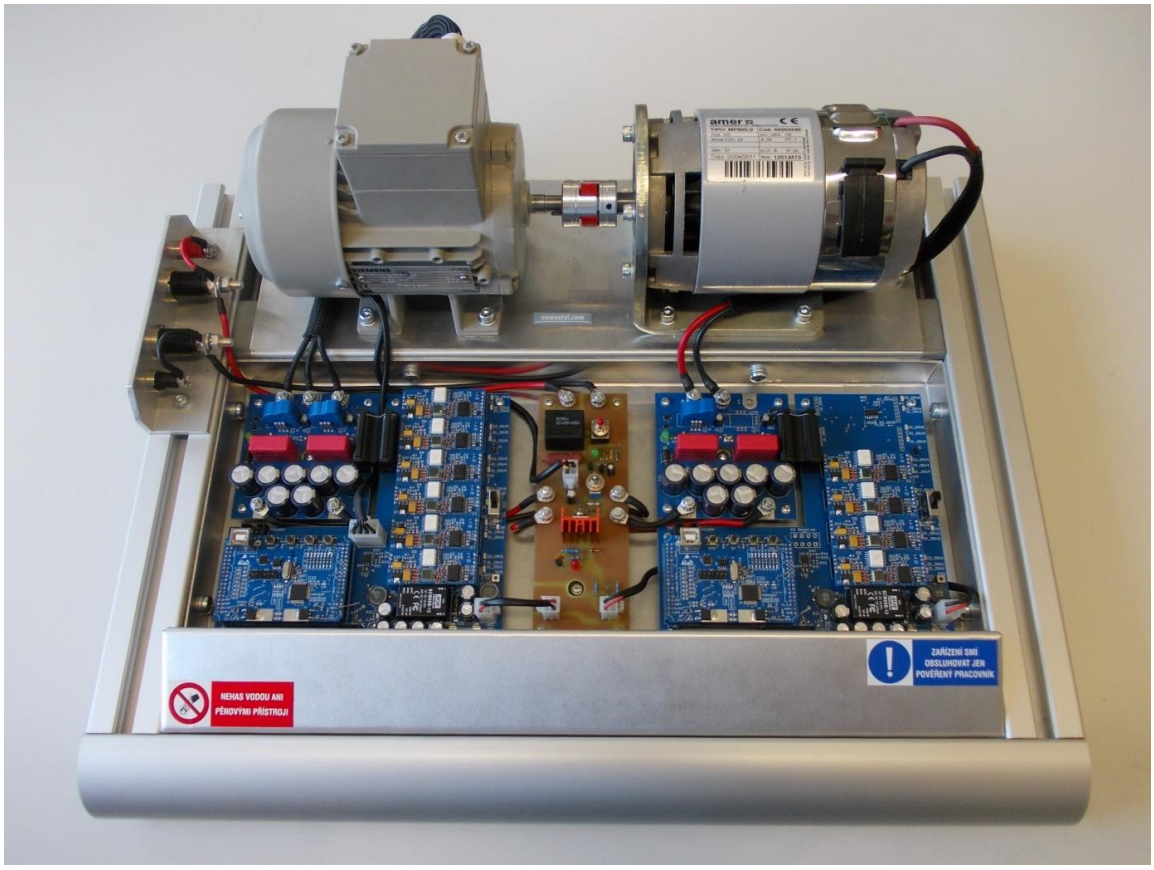


Fig. 4 Laboratory workbench

5. CONCLUSION

This laboratory workbench presents innovation in teaching of electrical drives. The introduction of this device into the laboratory exercises will reduce the cost of teaching. The workbench can be used in other subjects, which is taught at department of power electrical and electronic engineering, for example electrical machines or microprocessor technology. The advantage is the use of one universal device for all applications in the laboratory exercises in electric drives.

REFERENCES

[1] Hudák, O. *Laboratorního soustrojí s asynchronním a stejnosměrným motorem.* Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2012. 59s.

[2] ONDREJČEK, V. Laboratorní výukový stand s měniči a motory. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2013. 40 s.

[3] ONDREJČEK, V. *Řídicí systém laboratorního standu pro výukové účely.* Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2014. 63 s.

THERMAL MODEL OF INDUCTION MACHINE AND ITS PRACTICAL VERIFICATION

Marek Toman
Doctoral Degree Programme (1), FEEC BUT
E-mail: xtoman33@stud.feec.vutbr.cz

Supervised by: Pavel Vorel
E-mail: vorel@feec.vutbr.cz

Abstract: This article deals with making of thermal models of electric machines using so-called thermal network. The first part of this article describes reasons why thermal models of electric machine should be used. In the second part the heat transfer basics and thermal network basics are described. The last part is focused on simulation of a thermal network of a specific induction machine and its verification of correct functionality by measurements.

Keywords: Temperature, thermal model, heating curve, induction machine

1 INTRODUCTION

It is well-known that one of the main adverse effect in electrical machines is their heating-up during the operation. The calculation of temperature rise is necessary already in the design of machines because the temperatures have a major impact on the life of the machine. The most delicate part of an electric machine for exceeding of permissible temperature is definitely a winding. Overheating of winding causes its damage and therefore irreversible damage to the machine. Maximum permissible temperature of the winding is dependent on used insulation class. The maximum working temperatures of different insulation classes are shown below.

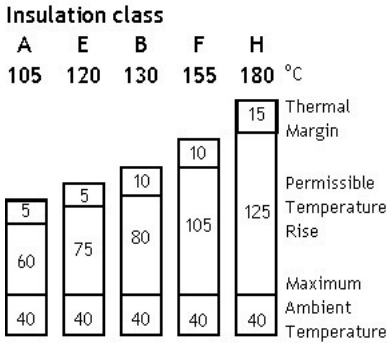


Figure 1: Insulation classes [1]

The increased temperature affects not only the lifetime of the machine but also its electrical parameters especially resistances of stator and rotor. These resistances can vary quite a lot with temperature, which is dependent on the loading of the machine. A typical example is the measurement of torque characteristics.

2 HEAT TRANSFER BASICS

For the calculation of thermal processes in electrical machines a so-called thermal network method is very often used. This method is based on the analogy between electrical and thermal circuits. The heat flux $\dot{Q}$ flow through a thermal resistance R_ϑ is directly proportional to the temperature difference $\Delta\vartheta$ and inversely proportional to the thermal resistance, see eq. (1).

$$\dot{Q} = \frac{\Delta\vartheta}{R_\vartheta} \tag{1}$$

We can see that this is a certain analogy to Ohm's law. Comparing thermal and electrical circuits we can see that the heat flow represents the equivalent of an electric current, temperature difference represents the equivalent of an electric voltage difference and the thermal resistance represents the equivalent of an electric resistance. This means that for the calculation of thermal processes, we can use the basic rules for the calculation of electrical circuits. It will be demonstrated that it is even possible to use the thermal capacitance which represents the equivalent of an electrical capacitance and it allows the progress of temperature versus time.

A key step in the calculation is the thermal resistances determination. There are three fundamental modes of heat transfer, namely conduction, convection and radiation. In the electric machines only conduction and convection heat transfer is usually present. Relationship for the calculation of thermal resistance is dependent on the mode of heat transfer.

Thermal resistance for the heat transfer caused by conduction is expressed as

$$R_\vartheta = \frac{l}{kA} \tag{2}$$

where l is the length of the material, k is the thermal conductivity of the material and A is the cross-sectional area [2].

Thermal resistance for the heat transfer caused by convection is expressed as

$$R_\vartheta = \frac{1}{hA} \tag{3}$$

where h is the convection heat transfer coefficient and A is the heat transfer surface area [2].

2.1 THERMAL NETWORK BASICS

Thermal network method is used to calculate of unknown temperatures in the individual parts of the machine. It is based on a creation of electric scheme in which the analogy between electrical and thermal circuit is used. This means that the individual resistors are the thermal resistances between different parts of the machine. Individual parts of the machine are represented by nodes. Each node represents a temperature in individual part of the machine. Each node may be a source of heat flow (in the case that in this node the heat loss are generated). Each node may also have its thermal capacitance (in the case we want to know the progress of temperature versus time).

Network configuration is usually based on the geometric shape of the machine. It is very important that the thermal network represents a system with lumped parameters. Therefore calculated temperatures in the nodes represent the mean temperatures of individual parts of the machine. Because the temperature curves in the machine are generally nonlinear (it is caused mainly by the internal generated heat flux) it may cause that in some part may be higher temperature then the mean calculated temperature. This can be compensated by a higher number of nodes.

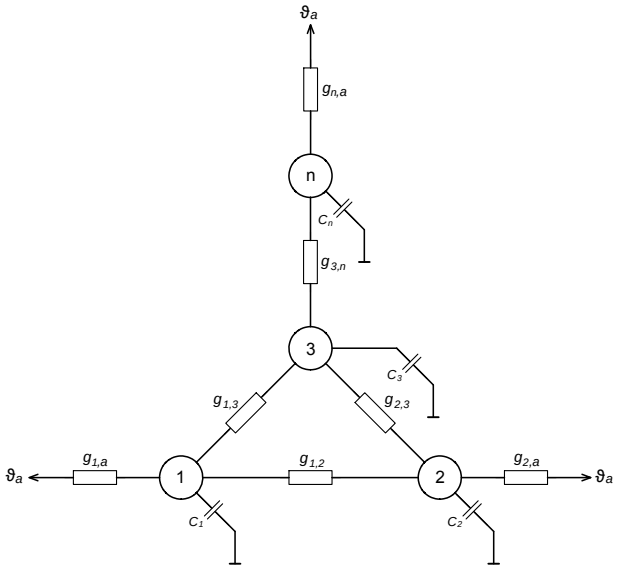


Figure 2: General thermal network

Example of general thermal network can be seen in Figure 2. Using this network will be described the basic rules to obtain equations which are need to calculation of unknown temperatures. The total number of equations is given by the number of nodes of the thermal network.

$$\begin{pmatrix} G_1 & -g_{12} & -g_{13} & \cdots & -g_{1n} \\ -g_{21} & G_2 & -g_{23} & \cdots & -g_{2n} \\ -g_{31} & -g_{32} & G_3 & \cdots & -g_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -g_{n1} & -g_{n2} & -g_{n3} & \cdots & G_n \end{pmatrix} \cdot \begin{pmatrix} \vartheta_1 \\ \vartheta_2 \\ \vartheta_3 \\ \vdots \\ \vartheta_n \end{pmatrix} = \begin{pmatrix} \dot{Q}_1 + g_{1a}\vartheta_a \\ \dot{Q}_2 + g_{2a}\vartheta_a \\ \dot{Q}_3 + g_{3a}\vartheta_a \\ \vdots \\ \dot{Q}_n + g_{na}\vartheta_a \end{pmatrix} - \begin{pmatrix} C_1 & 0 & 0 & \cdots & 0 \\ 0 & C_2 & 0 & \cdots & 0 \\ 0 & 0 & C_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & C_n \end{pmatrix} \cdot \begin{pmatrix} \dot{\vartheta}_1 \\ \dot{\vartheta}_2 \\ \dot{\vartheta}_3 \\ \vdots \\ \dot{\vartheta}_n \end{pmatrix} \tag{4}$$

Arbitrary thermal network can be described by system of equations, see in eq. (4). It can be written in matrix form

$$\boldsymbol{G} \cdot \boldsymbol{\vartheta} = \dot{\boldsymbol{Q}} - \boldsymbol{C} \cdot \dot{\boldsymbol{\vartheta}} \tag{5}$$

where $\boldsymbol{G}$ is a square matrix of thermal conductivities, $\boldsymbol{\vartheta}$ is a matrix (column vector) of unknown temperatures, $\dot{\boldsymbol{Q}}$ is a matrix (column vector) of heat fluxes, $\boldsymbol{C}$ is a square matrix of thermal capacitancies and $\dot{\boldsymbol{\vartheta}}$ is a matrix (column vector) of a time derivatives of temperatures. Element $\dot{Q}_i$ in $\dot{\boldsymbol{Q}}$ matrix represents heat loss of $i-$th node and product $g_{ia}\vartheta_a$ represents heat dissipation to an ambient from this node.

Diagonal matrix element G_i is given by the sum of all thermal conductivities entering to $i-$th node. Thermal conductivity g_{ij} is expressed using thermal resistivity as

$$g_{ij} = \frac{1}{R_{\vartheta ij}}. \tag{6}$$

Thermal conductivity g_{ij} is the same as g_{ji}. Where i is a number of row and j is a number of column.

To obtain the solution the highest derivative must be expressed, see eq. (7).

$$\dot{\boldsymbol{\vartheta}} = \boldsymbol{C}^{-1} \cdot (\dot{\boldsymbol{Q}} - \boldsymbol{G} \cdot \boldsymbol{\vartheta}) \tag{7}$$

3 SIMULATION AND MEASUREMENTS

By using the procedures described in previous section the simulation of heating curve of induction motor was done and the results were compared with measurements in the laboratory. Parameters of used motor can be seen in table below

ATAS T22VT512	
Voltage:	400 V/Y
Frequency:	50 Hz
Nominal power:	600 W
Rated speed:	2830 min^{-1}
Rated Current:	1.6 A
Efficiency:	82 %
No. of poles:	2

Table 1: Parameters of motor

3.1 SIMULATION

For this machine, a specific thermal network with 18 nodes was created for the simulation. This network was created based on the information from [3]. After the specific network was made, all the thermal resistances and thermal capacitances of this network were calculated. Furthermore, the heat losses in individual parts of the machine (stator winding Joule losses, rotor Joule losses, iron losses and mechanical losses) had to be calculated. Then these losses were divided into the corresponding nodes of the network. After that, the simulation was started. The simulation was performed in MATLAB/Simulink software and the results can be seen in Figure 3.

3.2 MEASUREMENTS

To verify correct function of simulation, measurements in the laboratory were taken. Measurements of heating curve were taken. These measurements are performed so that the motor is loaded at nominal torque, and the temperatures versus time are measured. During the measurements, temperatures on different parts of the machine were taken. Specifically, five PT100 sensors were used. Two sensors were placed on stator end-windigs and one sensor was placed on stator slot. These sensors were used to measure the stator winding temperatures. Fourth sensor was used for measurement of inner air temperature and last sensor was used for measurement of frame temperature. The measured temperature curves are in Figure 3.

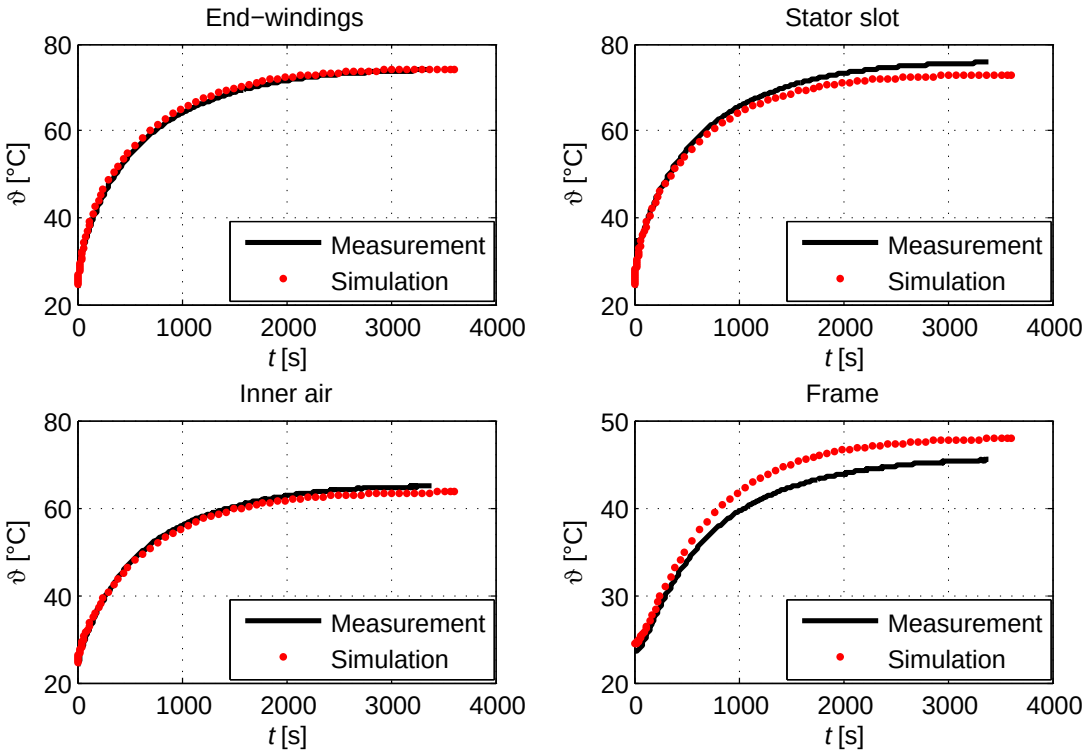


Figure 3: Temperature curves

Figure shows the resulting curves of simulated and measured temperatures versus time. Each of the four charts displays comparison of simulated and measured temperature of individual part of the machine. The temperature curves of both end-windings were averaged. It can be summarized from comparing the measured and simulated graphs that used thermal network gives very good results.

4 CONCLUSION

This article deals with the analytical calculation of the temperatures in electrical machines. The calculation method based on using so-called thermal network was described. Using the presented procedures a specific thermal network for induction machine was created and simulation of heating curve in MATLAB/Simulink was done. Correct functionality of the model was validated by comparison with measured temperature curves using a particular induction motor. Differences of simulated and measured temperature curves were very low. It can be summarized that used thermal model works very good and in the future it will be used to compensation of error caused by increased electrical resistances resulting from heating during torque curve measurement.

REFERENCES

[1] MyElectrical Engineering. *Motor Insulation* [online]. [cit. 2016-03-28]. Retrieved from: http://myelectrical.com/notes/entryid/154/motor-insulation

[2] ÇENGEL, Y. A., A. J. GHAJAR. *Heat and mass transfer: Fundamentals and applications*. Fifth Edition. New York: McGraw-Hill Education, 2015. ISBN 978-007-3398-181.

[3] HYNST, R., D. NECASOVA, L. POKORA, J. CHARVAT a J. LONDIN. *Vypocet chlazeni asynchronniho motoru nakratko, provedeni IP 44, s zebrovanou kostrou, bez vnitrniho obehu vzduchu*. Brno: VUES, 1969.

PROBLEMS WITH MAGNETIC MEASUREMENT OF 41CRMO4 STEEL

Tomáš Bulín

Doctoral Degree Programme (2), FEEC BUT

E-mail: xbulin01@stud.feec.vutbr.cz

Supervised by: Čestmír Ondrůšek

E-mail: ondrusek@feec.vutbr.cz

Abstract: This article deals with a problem of measuring a soft magnetic material with different geometrical dimensions. Steel 41CrMo4 was chosen as tested solid material. Two geometrical dimensions of this material were measured from the view of magnetic properties and compared each other. It is very difficult to determine proper commutation and hysteresis curve because eddy currents occur in every solid material.

Keywords: Magnetic measurement, losses, solid material, eddy currents.

1. INTRODUCTION

Magnetic measurement is not an easy task as it seems. Every difference of geometry dimensions gives various results. Similar results can be obtained on the samples composed of sheets with the same measuring method. But when it is desirable to measure solid material there are differences because of eddy currents. These currents are generally undesirable. Electric machines have worse efficiency due to the greater value of losses. Eddy currents can be restrained using sheets instead of solid material but sometimes it is necessary to use the solid material for example in high-speed induction motors with a solid rotor. It is desired to determine the correct magnetic parameters for the proper analytical design and subsequent simulation.

2. MAGNETIC MEASUREMENTS

2.1. TYPE OF MAGNETIC MEASUREMENTS

It is very important to know the magnetic properties of selected material because this knowledge is a prerequisite for the correct design of the electrical machine. The most suitable method for measuring of solid material is the inductive method, which is based on the principle of Faraday law of electromagnetic induction. Other measuring methods [1] are not so much suitable. For example some methods can be based on a change of material properties in presence of magnetic field. There are very small specimen and not enough volume to closure eddy-currents of the material. Induction law represents the following equation:

$$v_i = -N\frac{d\phi}{dt}, \qquad (1)$$

where v_i is induced voltage, Φ is magnetic flux and N is number of turns. Measurement of magnetic flux through the coil is based on formula $\phi = B \cdot S$, so it is necessary to know the size of the cross sectional area S. Then induced voltage can be easily measured. The growth of the vector magnetic flux density $\boldsymbol{B}$ causes the growth of the induced voltage v_i.

The calculation of magnetic field strength H_z from magnetizing current can be calculated according to equation:

$$H_z = \frac{N_1 \cdot I_1}{l_s}, \qquad (2)$$

where N_1 is a number of turns of magnetizing windings, I_1 is magnetizing current and l_s is mean magnetic path length. Manufacturer of measuring instruments is using equation (1) and (2) as introduction to problem of magnetic measurement. Generally the secondary winding is wound closely on the test specimen so the air flux is insignificant over the range of magnetic field strength 0 to 4 kA/m. At higher values of magnetic field strength, an air flux correction should be applied. The amount of total core losses P_S in the specimen depends on hysteresis losses, eddy current core losses and the excess losses. Hysteresis and eddy current core losses are applied at low magnetic flux densities B approximately until 1.0 T. It is necessary to calculate with excess losses for higher saturation which is in the teeth of stator and rotor in electric machine. The equation for calculation of magnetic total core loss is:

$$P_S = P_h + P_c + P_e = k_h f B_m^2 + k_c f^2 B_m^2 + k_e (f B_m)^{1,5}, \qquad (3)$$

where B_m is amplitude of AC flux component, f is frequency, k_h is hysteresis core loss coefficient, k_c is eddy-current core loss coefficient and k_e is excess core loss coefficient. The loss coefficients k_h, k_c, k_e are constants for the conventional model and they are dependent on the material property. Amount of losses is changing with frequency f and/or flux density B_m. This equation is valid for materials under sinusoidal flux conditions.

2.2. MEASURED DATA

All measurements were carried out at the measuring station made by Magnet-physik Dr.Steingroever GmbH. The station consists of two devices called RemaCOMP and RemaGRAPH C-710. These devices are able to measure static hysteresis curve (for DC excitation) and dynamic hysteresis curves (for AC excitations). There is a signal generator connected to a power amplifier, which supplies current to the primary winding of the specimen. The field strength H can be calculated from the current and the magnetic path length l_m. The secondary induced voltage u_{i2} is proportional to the time change of the magnetic flux $d\phi/dt$. It is numerically integrated and then calculated flux density B. Difference between AC and DC measurement is only in use of another device but the specimen and winding is the same for both types of measurements. DC measurement is using a very small change of supply current so there must be special integrator like fluxmeter, which is a precision integrator for slowly variable signals because it calculates the flux density B [2].

41CrMo4 steel was chosen for measurement because this material is used for construction of solid rotors of high-speed motors. It is a very strong material with good mechanical properties. There is a lot of stresses due to extreme centrifugal forces in high-speed machines, which material has to withstand and it must have a good magnetic conductivity in the same time. Unfortunately, these parameters are difficult to find and then it does not have to be accurate. Two specimens in the shape of toroid were used for measurement. Geometrical dimensions are in Table 1:

	Larger	Smaller	
Outer diameter	120	44.3	mm
Inner diameter	100	38.1	mm
Height	10	3	mm
Primary winding	220	101	
Secondary winding	200	20	
Mass	267	9.07	g

Table 1: Geometrical dimensions of specimens

Magnetic measurement can be affected in different directions. For example the dimensions of the specimen has to be square. The magnetic measurement can be affected by temperature change and mechanical stress of specimen.

On the Figure 1 there can be seen comparison of hysteresis curves for a smaller sample for different frequencies. The area of hysteresis curves is increasing with higher frequencies too, due to eddy currents in the sample.

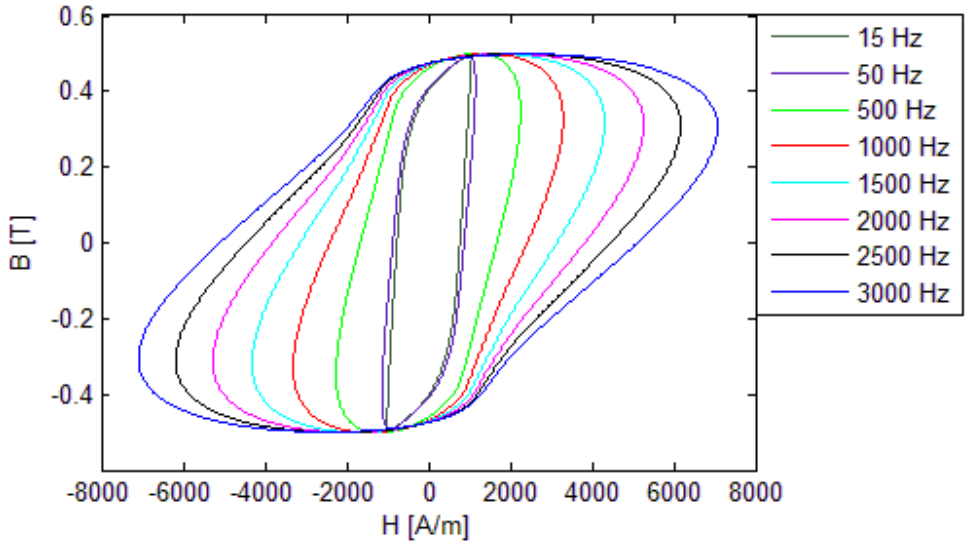


Figure 1: Comparison of hysteresis curves (smaller sample) for different frequencies

In the following graphs there can be seen how the value of magnetic losses *Ps* (Figure 2 and Figure 3) or permeability (Figure 4 and Figure 5) is changing in dependence on magnetic flux density.

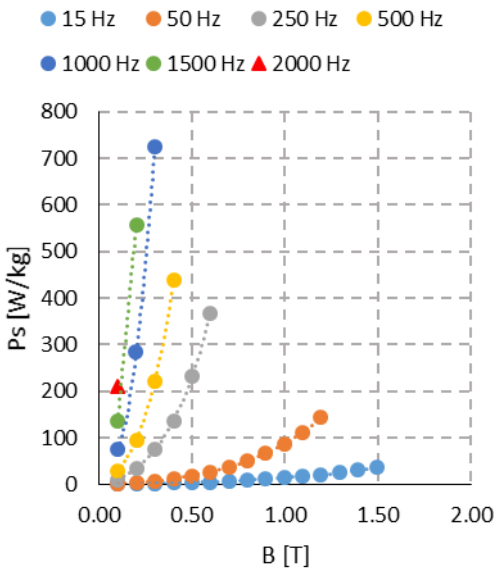


Figure 2: Dependency amount of magnetic losses on magnetic flux density Larger sample

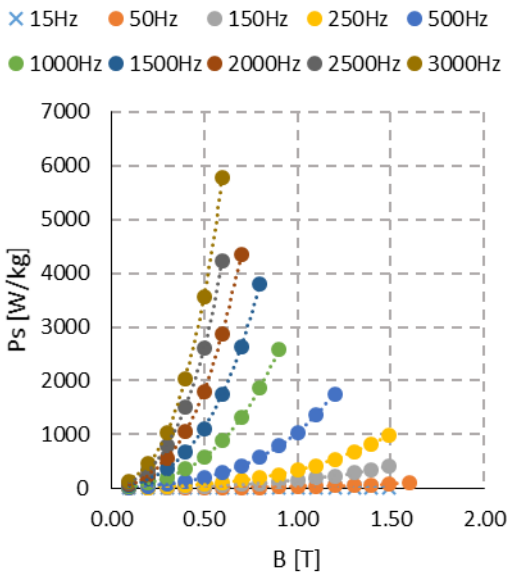


Figure 3: Dependency amount of magnetic losses on magnetic flux density Smaller sample

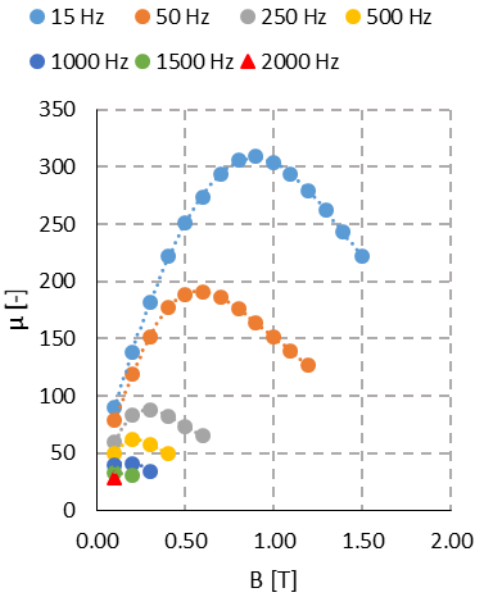


Figure 4: Dependency size of permeability on magnetic flux density Larger sample

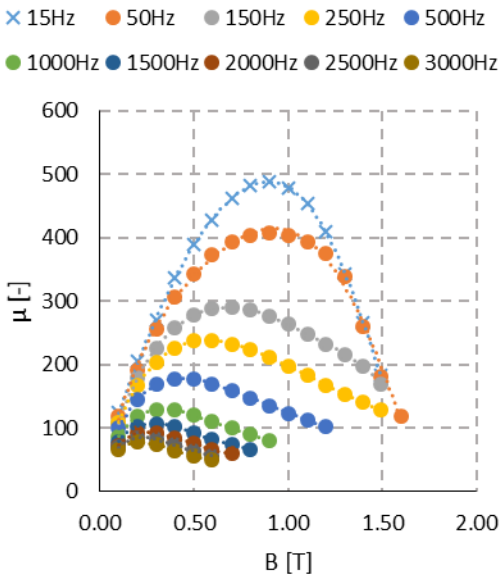


Figure 5: Dependency size of permeability on magnetic flux density Smaller sample

Comparison of larger and smaller sample for magnetic losses (Figure 6 and Figure 7) shows that in the larger sample there is a greater amount of eddy current and due to this fact there are higher total magnetic losses.

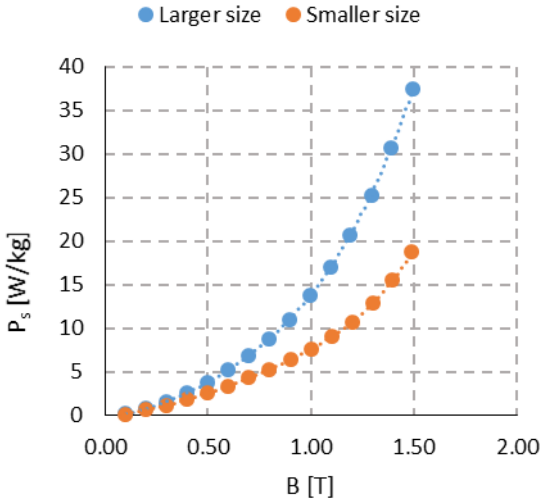


Figure 6: Comparison of magnetic losses at frequency 15 Hz

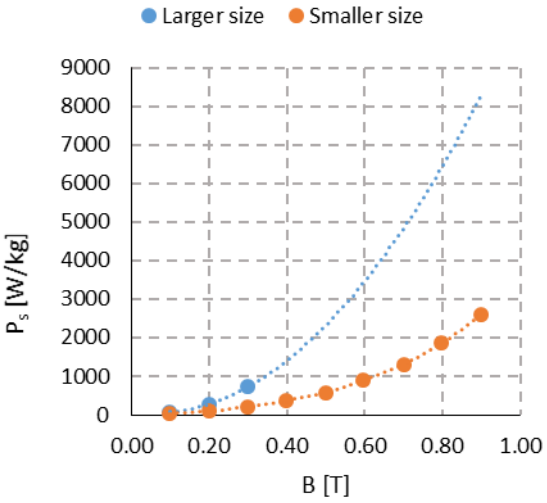


Figure 7: Comparison of magnetic losses at frequency 1000 Hz

Total magnetic core losses are rapidly exponentially increasing with magnetic flux density and frequency. The Larger sample could not be measured for the same excitation as a smaller sample because of measuring device limitation. But even for measured values excitation, there are significant differences between both samples. The curve of the total magnetic loss (Figure 7) depending on magnetic flux density was interpose at larger sample by the polynomial and estimated for higher excitation because of the better display.

After measuring the magnetic properties at frequency samples were measured with the quasi-static device too. This device has the advantage that measuring frequency is near to 0 Hz so there are no

additional losses, which come from eddy currents. Comparison of quasi-static and AC measurement can be seen in Table 2. higher value of coercivity H_c represents larger area of hysteresis curve and consequently higher total magnetic core losses. Parameter H_c is dramatically increasing with the growth of supply frequency. Device which serves for magnetic measurement at 0 Hz is not able to measure total magnetic losses, so coercivity H_c is a better parameter to compare.

	Larger		Smaller	
	f = 0 Hz	f = 50 Hz	f = 0 Hz	f = 50 Hz
	H_c [A/m]	H_c [A/m]	H_c [A/m]	H_c [A/m]
B = 0.5 T	865	1668	664	875
B = 1.0 T	1097	3792	783	1340
B = 1.5 T	1275	-	881	2196

Table 2: Comparison quasi-static and AC measurement

In Table 3, there can be seen magnetic total loss *Ps* for different excitation and frequencies. The deviation between Smaller and Larger sample is significant and is increasing with both changing parameters. Even if the higher frequency (250 Hz) is only five times greater than fundamental frequency (50 Hz) deviation is larger by 138 % for investigated samples at 0.6 T.

	f = 50 Hz			f = 250 Hz	
	B = 0.2 T	B = 0.6 T	B = 1.2 T	B = 0.2 T	B = 0.6 T
	P_s [W/kg]	P_s [W/kg]	P_s [W/kg]	P_s [W/kg]	P_s [W/kg]
Larger	3.83	27	143	35	367.4
Smaller	2.27	13.4	47.5	15,7	108.3
Deviation	68.7 %	101.5 %	201.1 %	122.9 %	239.2 %

Table 3: Comparison of values of magnetic measurements

3. CONCLUSION

The magnetic measurement was performed on toroidal shape samples. The properties of the material were measured for different frequencies from 0 Hz to 2000 Hz. Unfortunately, it was not possible to measure high magnetic flux density at higher frequencies because of measuring device limitation. The differences are caused by different geometrical dimensions of samples and consequently generation of larger amount of eddy currents. The same results were expected at zero frequency, but coercivity *Hc* differs too and area of hysteresis curve is also greater for the large sample. Appropriate geometrical dimensions with respect to more precise results, which can be used as right input parameters of simulation, should be the subject of further investigation

ACKNOWLEDGEMENT

This research work has been carried out in the Centre for Research and Utilization of Renewable Energy (CVVOZE). Authors gratefully acknowledge financial support from the Ministry of Education, Youth and Sports of the Czech Republic under NPU I programme (project No. LO1210).

REFERENCES

[1] FIORILLO, Fausto. Measurements of magnetic materials. *Metrologia* [online]. 2010, roč. 47, č. 2, s. S114–S142. ISSN 0026-1394. Available from: doi:10.1088/0026-1394/47/2/S11

[2] STEINGROEVER, Dr. Erich a Dr. Gunnar ROSS. MAGNET-PHYSIK DR. STEINGROEVER GMBH. *Magnetic measuring techniques*. Köln, Germany, 2009.

THE EDDY CURRENT LOSSES OF HIGH-SPEED SOLID ROTOR INDUCTION MACHINE AS FUNCTION OF THE RADIAL AIR-GAP LENGHT

Jan Bárta

Doctoral Degree Programme (2), FEEC BUT

E-mail: jan.barta@phd.feec.vutbr.cz

Supervised by: Čestmír Ondrůšek

E-mail: ondrusek@feec.vutbr.cz

Abstract: This paper introduces that the appropriately selected air-gap width have major impact in terms of efficiency for high-speed induction machines. Machine of interest is high-speed solid rotor cage induction machine intended for helium turbo-circulator. Proposed study is done from electromagnetic and thermal point of view. Presented results can be used as a guideline for air-gap length selection for others high-speed solid rotor induction machines.

Keywords: AC machine; High-speed drive; Induction motors; Finite element analysis

1. INTRODUCTION

Rapid grow can be noticed in development, research and applications of high-speed electrical machines in recent years. High efficiency level of new high-speed motors has to be reached to minimize their energy losses.

Electrical machine under investigation is shown in Fig. 1. This machine is three-phase, 2 poles high-speed solid rotor cage induction machine (HSSRCIM). Short-pitched double-layer winding for suppressing 5^{th} and 7^{th} air-gap flux density harmonics is distributed within 12 stator slots. Solid rotor structure instead of laminated one is used due to high centrifugal forces. To achieve necessary centrifugal force resistance the containment sleeve is fitted over the short-circuit rings. Parameters of the machine under investigation are listed in table I.

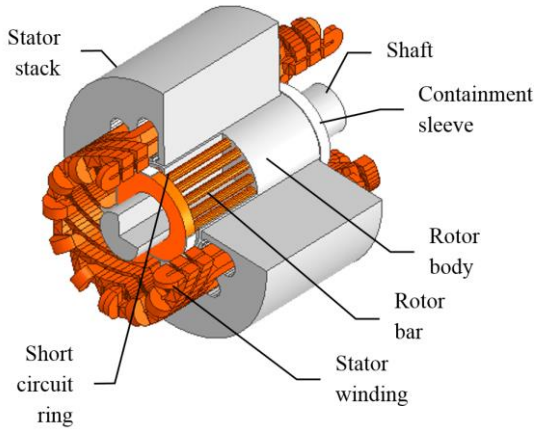


Figure 1: Solid rotor cage induction machine 6 kW, 120 000 rpm

Specification	Unit	Rating values
Number of pole pairs		1
Number of phases		3
Number of stator slots		12
Number of rotor slots		18
Stator external diameter	mm	90
Stator inner diameter	mm	32.5
Active length	mm	50
Rotor diameter	mm	30.7
Rated line-to-line voltage	V	350
Stator winding connection		Wye
Stator stack material		M250-35A
Rotor stack material		41CrMo4
Squirrel cage material		Glidcop – Al 15

Table 1: Main parameters of machine under investigation

2. THEORETICAL BACKGROUND

High-speed electrical machines are powered from high-frequency power sources. If we considered constant magnetic flux in the air-gap, then we can write following equation

$$P_{loss} = P_{fund} + P_{hyst} + P_{eddy} + P_{fric} \approx k_{fund} + k_{hyst}\omega + k_{eddy}\omega^2 + k_{fric}\omega^3 \quad (1)$$

where P_{loss} are electrical machine total losses, P_{fund} are losses nondependent on frequency (e.g. additional losses), P_{hyst} are losses dependent on frequency (e.g. hysteresis losses), P_{eddy} are losses dependent on square of frequency (e.g. eddy current losses) and P_{fric} are losses dependent on cube of frequency (e.g. friction losses). Therefor special attention must be given to reduction of friction and eddy current losses in high-speed electrical machines (HSEM).

Higher harmonic content of air-gap flux density is the main source of eddy current losses in solid rotor cage induction machines. These higher harmonics are mainly caused by slotting. Generally, there are two sources of harmonics related to slotting in electrical machines. First are harmonics caused by windings concentrated in slots and another one are harmonics related to permeance variation due to slotting. The first one is dependent on the loading of machine and the second one is typically considered as load-independent. Both sources of harmonics have the same wave numbers and frequencies. This mean that these harmonics are superposing on each other. Heller and Hamata proposed in [1] following equation for the slot harmonics:

$$B_S^S(\phi^S, t) = B_+ \cos\left[(Q_s + p)\phi^S - \omega_s t\right] + B_- \cos\left[(Q_s - p)\phi^S + \omega_s t\right] \quad (2)$$

where Q_s is number of stator slots, p is number of pole-pairs, ϕ is polar angle, ω_s is synchronous angular speed, t is time and B+, B- are amplitudes of slot harmonics. First one rotates in the same direction as a main flux, second one rotates in negative direction. These harmonics have different length wave which varies at the same frequency as fundamental harmonic. Solid rotor surface is exposed to field in rotor frame of reference. After transformation from stator to rotor frame of reference the equation is as follow

$$B_r^r(\phi^r, t) = B_+ \cos\left[(Q_s + p)\phi^r + Q_s(1-s)\frac{\omega_s}{p}t - s\omega_s t\right] + B_- \cos\left[(Q_s - p)\phi^r + Q_s(1-s)\frac{\omega_s}{p}t + s\omega_s t\right] \quad (3)$$

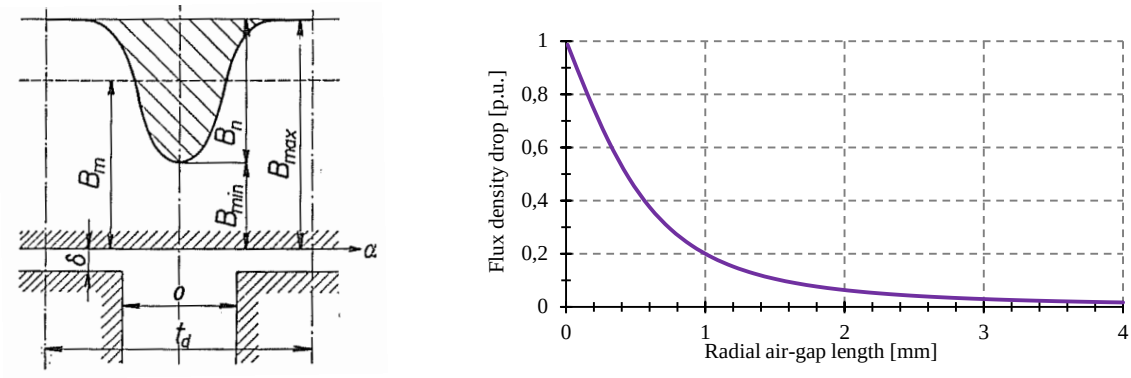
where s is rotor slip. It can be noticed that harmonics vary at much higher frequency in rotor frame of reference than in stator frame. Therefore, slot harmonics are very efficient in producing eddy-current losses. The slot harmonics can be seen as a drop in flux density distribution in air-gap. Higher drop is resulting into the higher slot harmonics components B_+ and B_- and therefore into higher eddy current losses. Distribution of flux density above slot opening on a surface of the rotor is shown in Fig. 2a. The amplitude of flux density drop B_n can be, according to Heller and Hamata [1], derived by following equation:

$$B_n = 2\beta \; B_{max} \quad (4)$$

where B_{max} is the value of flux density without slotting and β is function of the ratio between slot opening and the radial air-gap length:

$$\beta = \frac{1 + x^2 - 2x}{2(1+x)} \quad \text{where,} \quad x = \frac{o}{2\delta} + \sqrt{1 + \left(\frac{o}{2\delta}\right)^2} \quad (5)$$

Where o is slot opening and δ is radial air-gap length. If we assume that slot opening is o=1.5 mm then the per unit flux density drop can be expressed by function plotted in Fig. 2b. Since eddy currents are generated by changes of magnetic flux the total value of rotor eddy current losses should follow this function. Amount of eddy current losses might be significantly reduced by varying air-gap length. Therefore air-gap length has significant influence on machine efficiency with respect to equation (1).



(a) (b)

Figure 2: a) Distribution of flux density above a slot [2] b) Per unit flux density drop as a function of radial air-gap length

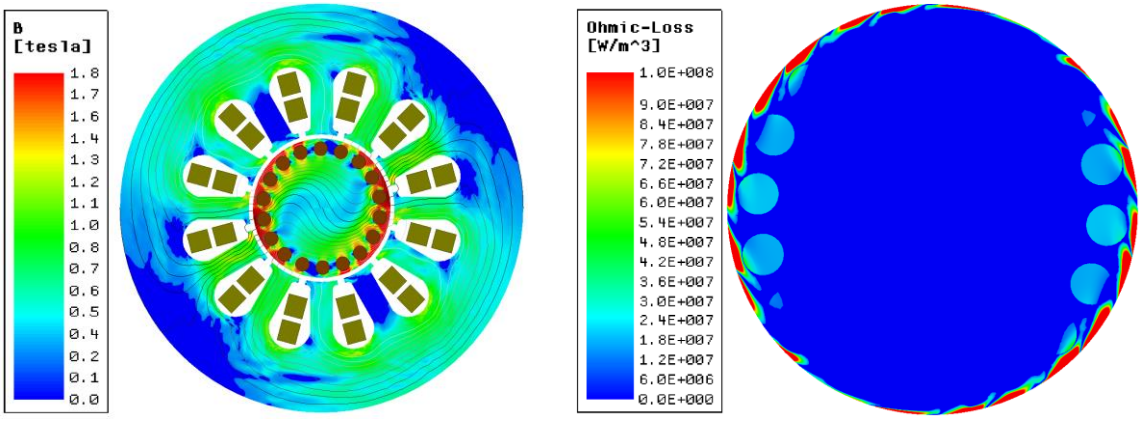
3. RESULTS OF ELECTROMAGNETIC CALCULATIONS

Model for electromagnetic calculations has been made according to [2] in Ansys Maxwell. Performed calculations are based on time-step, 2D finite element analysis. For reduction of computation time the sinusoidal power source is considered. Impedances of end windings and short-circuit rings are modeled through external circuits. The elements dimensions close to rotor surface are smaller than the skin depth associated with slot harmonics and material characteristics. This set-up is necessary for accurate eddy current losses calculation. Skin depth can by expressed as follow:

$$\delta_{\text{skin}} = \sqrt{\frac{2}{\omega\sigma\mu}} \tag{6}$$

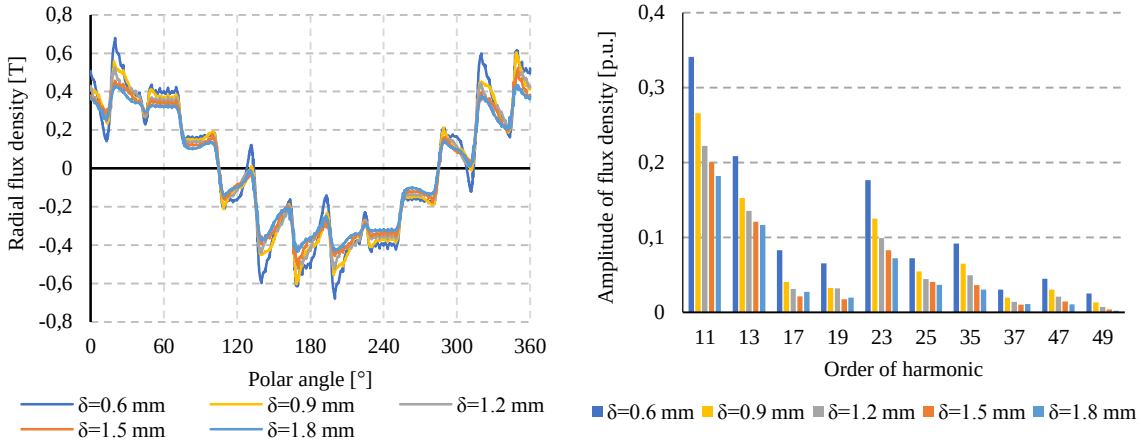
where ω is harmonic angular frequency, σ is electric conductivity and μ is permeability. Calculated field and rotor ohmic loss distributions are both shown in Fig 3.

Flux density drops caused by slotting can be noticed in normal component of radial flux density shown in Fig. 4a. Per unit flux density amplitudes of first- and second-order slot harmonics of the stator
(11^{th}, 13^{th} and 23^{rd}, 25^{th}) and rotor (17^{th}, 19^{th} and 35^{th}, 37^{th}) are shown in Fig. 4b. In the stator frame of reference, the stator slot harmonics vary at power supply frequency. The frequency of the rotor slot harmonics is in the kilohertz range.



(a) (b)

Figure 3: a) Flux density distribution at the rated load condition and radial air-gap length 0.9 mm b) Ohmic-loss distribution in the rotor at the rated load condition and radial air-gap length 0.9 mm



(a) (b)

Figure 4: a) Flux density normal component distribution in the middle of the air gap b) Per unit amplitudes of the air-gap flux density slot harmonics

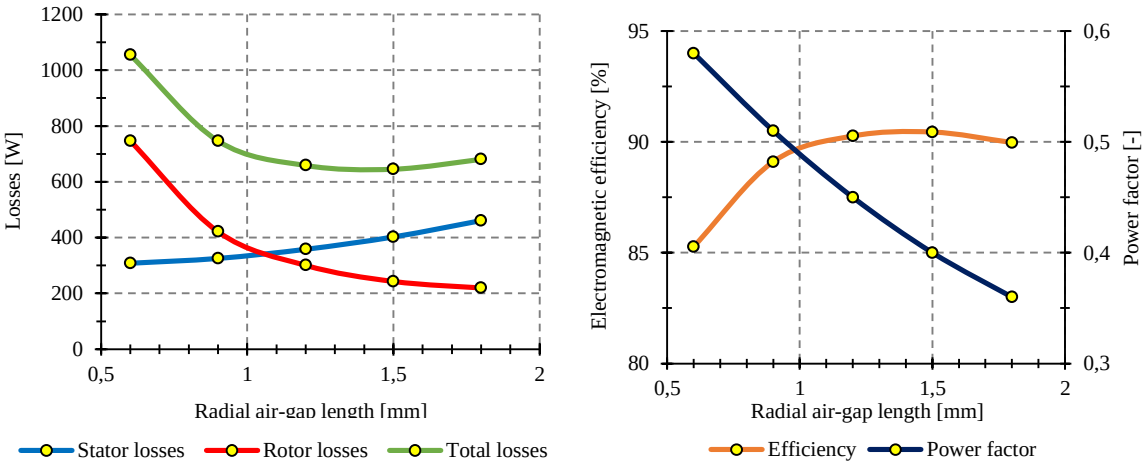
Calculated losses for rated load condition as a function of the radial air-gap length are shown in Fig. 5a. Results were obtain by number of iteration with thermal calculations. Iterations with thermal calculation are necessary since material resistances are thermal dependent. It can be noticed that rotor losses have similar radial air-gap length dependency as a function of flux density drop shown in Fig. 2b. It is because the eddy-current losses are decreased with smaller flux density drop. In simplified form, rotor losses dependency can be expressed as follows:

$$P_{rot} = \frac{1}{k_{rot}\delta^{x}} \tag{7}$$

where k_{rot} and x are constants dependent on motor design and δ is radial air-gap length. The stator losses increase with increase of air-gap. That is due to increase of magnetization current of the motor. Stator losses in simplified form can be expressed as follows:

$$P_{stat} = k_{stat}\delta^{x} \tag{8}$$

where k_{stat} and x are constants dependent on motor design and δ is radial air-gap length. Radial air-gap length has impact on HSSRCIM electromagnetic efficiency as is shown in Fig. 5b. However, large air-gap leads to poor power factor.



(a) (b)

Figure 5: a) Losses as a function of radial air-gap length b) Electromagnetic efficiency and power factor as a function of radial air-gap length

4. RESULTS OF THERMAL CALCULATIONS

For internal thermal distribution, a thermal equivalent circuit has been created to model heat transfer inside the designed HSEM. The machine cooling is done through indirect water flow in cylindrical channels located inside of stator housing in combination with Helium which flows through electric machine air-gap. It is assumed that ambient temperature is 40 °C.

The calculated temperature dependency on air-gap length is shown in Fig. 6. The stator winding insulation class must be selected with respect to stator winding temperature. The rotor eddy current losses generates high temperature rise in area of rotor surface. This generated heat is efficiently removed by helium flow. Without the helium, the stator would be affected much more by rotor temperature rise.

Stator temperature is increasing for wider air-gap. That is due higher stator losses caused by poor power factor. However, it can be noticed that this temperature rise has no significant impact on temperature distribution inside electrical machine.

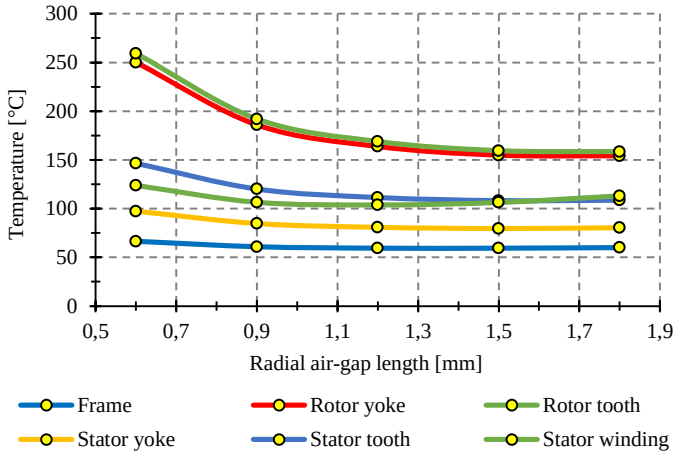


Figure 6: Temperatures as a function of radial air-gap length

5. CONCLUSION

It has been shown that significant additional losses are generated by slotting and slot harmonics in solid rotor. This paper presents reduction of their effects by using wider air-gap. Another possibility how to reduce their effects is by using closed slot, magnetic slot wedges and skewing of rotor or stator slots.

It can be concluded that air-gap design has major impact on machine performance in terms of efficiency. Presented results can therefore be used as a guideline for radial air-gap length design.

ACKNOWLEDGEMENT

This research work has been carried out in the Centre for Research and Utilization of Renewable Energy (CVVOZE). Authors gratefully acknowledge financial support from the Ministry of Education, Youth and Sports of the Czech Republic under NPU I programme (project No. LO1210)

REFERENCES

[1] Heller B., Hamata V.: Harmonics Field Effects in Induction Machines, Academia, Publishing House of the Czechoslovak Academy of Sciences, 1977

[2] Barta J., Ondrusek C., "Finite Element Analysis of High-Speed Solid Rotor Induction Machine with Copper Cage," Advanced Mechatronics Solutions, vol. 393, pp.195-200, Jan. 2016.

THE PROCESS OF CROSS SECTION GENERATION FOR REACTOR CORE SIMULATIONS

Filip Novotny

Doctoral Degree Programme (2), FEEC BUT

E-mail: xnovot66@stud.feec.vutbr.cz

Supervised by: Karel Katovsky

E-mail: katovsky@feec.vutbr.cz

Abstract: Reactor core safety analysis is a complex problem and needs processed input data and cannot be performed without it. For this purpose several nuclear programs exist. A suitable nuclear programs as well as the motivation of the generation are given in this study. This article also summarizes methods and steps needed for the homogenized cross-section generation and provides basic instruction and settings needed to provide these simulations and accomplished this research and gives basic instruction and information about problems and programs which can occur during the cross-section generation. Finally, there is proposed cross-sections structure suitable for safety analysis with reactor core simulator PARCS.

Keywords: SCALE, lattice physics models, cross-sections, GENPMAXS, PARCS

1. INTRODUCTION

Nuclear power plants (NPP) which has been operating since 80s were designed during late 60s and 70s of 20's century. The computer resources together with the mathematical methods weren't that involved like in the present. Based on uncertainties of the simplified models the projects of nuclear power plants were design very conservative. The nuclear power plant is the best of the art project of those days. Nevertheless, since that time there has been invented new technologies, methods and materials which can improve these projects. Therefore several parts of the nuclear power plant can be modernized and better utilization of the technology can be reach. Every improvement of the project, especially changes in the primary circuit need to be proven by simulation before implementation. When the reactor core is chosen to improvement, the modernization can be for example enhanced by modification of the nuclear fuel.

These modifications of nuclear fuel or reactor core have to be proven by several simulations of reactor core behavior. Several transient scenarios are established for this purpose and their common feature is the homogenized cross-sections need. Main aim of this article is to introduce steps and methods for homogenized cross-section generation which is necessary for 3D reactor core simulations, e.g. for safety analysis. Reactor core simulation is performed by reactor core simulator PARCS. Therefore the proper data which is requested by PARCS has to be calculated. At the beginning of the whole process the user has only geometric and material properties about the reactor itself and about the fuel assemblies. At this time, the user has information about fuel assemblies and reflector and about PARCS. According to Figure 1 the user has to fill several blank spots in the calculation chain. To fill these spots the nuclear program SCALE/TRITON and code GENPMAXS are very convenient to use.

With the definition of the used programs the user can begin with gathering the required information and preparing the fuel assemblies models (Figure 1 – green rectangles). Information about material properties of fuel assembly, fuel assembly dimensions and table of future problems under investigation are the most important information. This information is normally contained in several documents. After the study of documents the user usually has not all requested information and has to

adopt several assumptions. Every assumption should be handling very carefully and should be tested.

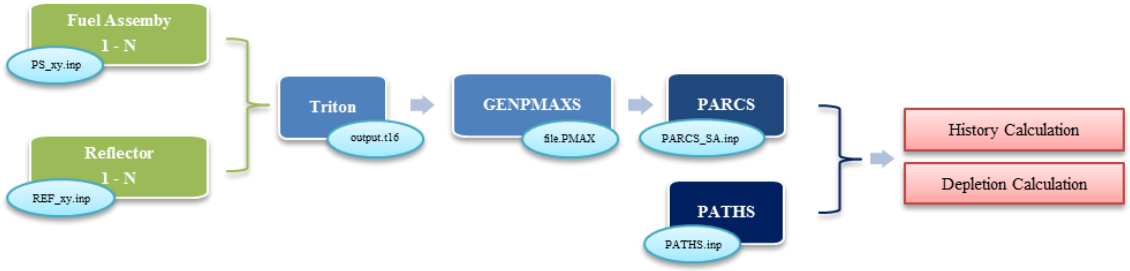
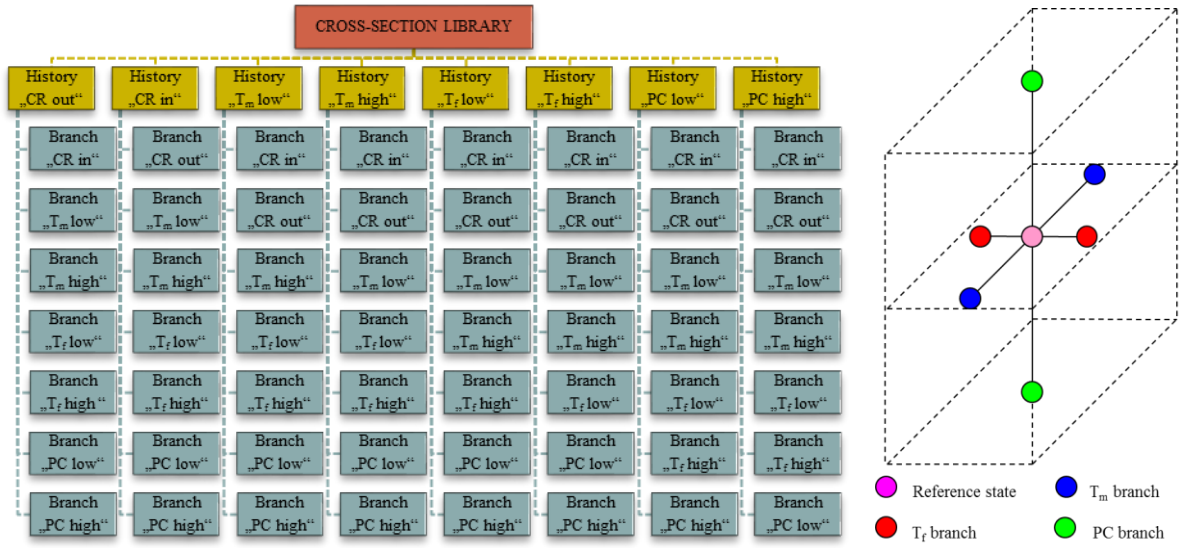


Figure 1: Calculation chain.

The reactor core of NPP Temelin with TVSA-T fuel consists of 6 different fuel assemblies and the reactor core reflector is represented by additional 4 reflector assemblies. The following step is the branch structure and history structure determination. Determination strongly depends on future use of homogenized cross-sections. The most general but not the most accurate structure consists of 14 branch calculation and 8 history calculation per fuel assembly. General history and branch structure is shown in the Figure 2.a where structure of branches is simplified i.e. branch "CR in" consists of another 7 branch calculations. Figure 2.b shows cuboid of possible cross-section which can be obtained from this cross-section library. Depending on the requested accuracy, more detailed structure can be determined but the time of the calculation is rising very quickly. All together there has to be calculated 1120 (10 assemblies multiple by 14 branches multiple by 8 histories) transport calculations to obtain homogenized cross-sections for single reactor core. The results of calculation performed with lattice physics model has to be compared to Monte Carlo simulation as a verification of the prepared models.



a) b)

Figure 2: History and branch structure (a), 3D interpretation of possible cross-section (b) [1].

At this time, the models and calculation scheme is determined and the brief description of the used codes (Figure 1 – blue rectangles) is given in the following sections.

2. HOMOGENIZED CROSS-SECTIONS LIBRARY GENERATION

Cross-section generation is a first step of the complex safety analysis. It consists of several phases such as microscopic cross-section preparation and processing, transport calculation, burnup calculation and branches calculation. Each phase is very important because this is the beginning of the whole calculation chain and the errors can accumulate through the whole chain.

2.1. MICROSCOPIC CROSS-SECTION PREPARATION

As a source for this sequence there is Evaluated Nuclear Data File (ENDF/B) library. At present, the best and the most accurate library is ENDF/B-VII.1. Pointwise data from this library has to be prepared for multi group calculation. Many different codes for this preparation exists e.g. NJOY. Dependent on the problem which has to be solved the weighted functions are chosen. Cross-sections data is very sensitive to this weighted functions and this makes the correct cross-section preparation very difficult. One of the advantages of SCALE nuclear program is that the sequence AMPX exists and the program developer provides calculated and tested multigroup nuclear data. The user of the SCALE nuclear program can choose between better accuracy with 252 groups energy structure or faster 56 groups structure. In the Figure 2 the illustration of the 238 groups structure is shown. [2] [3]

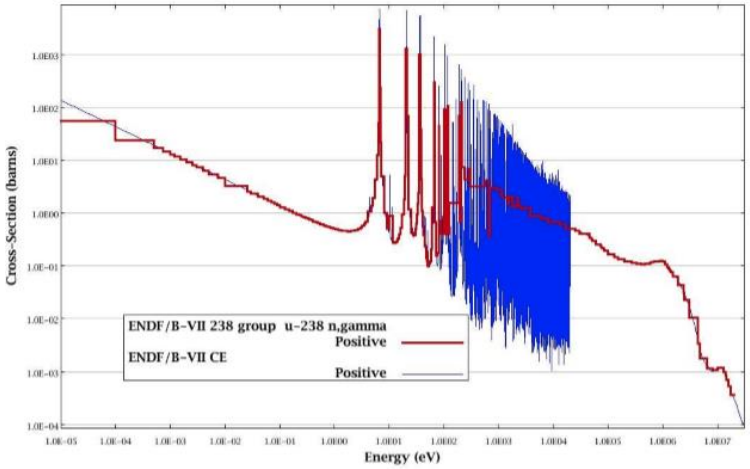


Figure 3: Comparison of the multi group structure with continuous energy structure. [3]

2.2. CROSS-SECTION PROCESSING

Cross-section processing utilizes several codes which is included in the TRITON sequence. As a first code the 1D cross-section processing code CENTRM is called together with self-shielding calculation code BONAMI. CEMTRM determinates problem-specific fluxes for processing resonance-shielded multigroup data. This data is stored in the main AMPX library and then this data is sent to higher dimension calculation. Within this nuclear package the transport code NEWT is utilized to higher dimension calculation. [3]

2.3. TRANSPORT CALCULATION

The transport calculation is provided by 2D discrete ordinated transport code NEWT. This transport code determinates spatial discretization by Extended Step Characteristic and this approach makes the NEWT very powerful tool to solve real-world non-orthogonal problem domains. Complex geometry in the transport calculation is built from several basic body shapes. The available body shapes are rectangle, circle, hexagon, rotated hexagon and triangle. Several boundary conditions such as reflective, periodic, white or vacuum can be used for modeling proper behavior of the model. [2] [4]

2.4. BURNUP CALCULATION

Material composition isn't constant during the operation of the nuclear reactor. Burnup is rising with the time of the operation and the uranium is depleted and fission products deposit in the fuel. Burnup calculation is held by code ORIGEN. Main aim of this code is to calculate time-dependent concentrations and activities for large amount of isotopes from ENDF library. ORIGEN tracks every nuclide which is generated or transformed during neutron fission, transmutation or radioactive decay. [2]

3. GENPMAXS TRANFORMATION

Prepared macroscopic cross-sections by lattice physics code are the information source for nodal 3D eigenvalue and transient calculations. Nevertheless every lattice physics code has unique output file format. Transformation of the lattice physics output file to readable file for PARCS code is obtained by macroscopic cross-section processing code GENPMAXS. The macroscopic cross-section in PARCS is constricted with the assumption of a linear superposition. During the transformation the exact value of cross-section has to be recalculated to the partial derivatives. During the partial derivatives is calculated three interpolation possibilities can occur. First and the easiest is interpolation of one branch. The interpolation is calculated by equation 1 where X is branch variable, i is state of the variable and r represents reference state.

$$\left.\frac{\partial \Sigma}{\partial x_k}\right|_{X^m} = \frac{\Sigma(X^i) - \Sigma(X^r)}{x_k^i - x_k^r} \tag{1}$$

An illustrative example for density of coolant branch (DC) is shown in the Figure 4. Six branches (DC^1, DC^2, DC^3, DC^4, Ref, Cr) are calculated by lattice physics code Reference state is stored in PMAXS file and partial derivatives (DC^{1m}, DC^{2m}, DC^{3m}, DC^{4m}) of coolant density are calculated using Equation 1 from original six branches. These partial derivatives are written into PMAXS file like four density branches. [1]

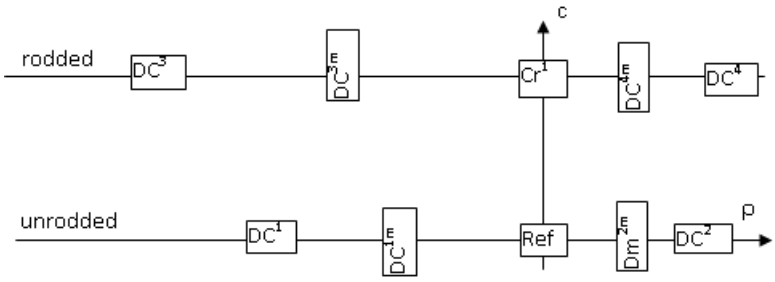


Figure 4: Illustrative example of partial derivatives calculation. [1]

4. CONCLUSION

Research in the area of reactor core safety analysis is based on sufficient input data. Many different materials, complicated geometry in the reactor core together with higher utilization of nuclear fuel in the last years lead to need of new analysis of reactor core behavior. Input data for reactor core analysis are one of the most important parts of simulation. The scheme of homogenized cross-section generation based on geometric and material data of TVSA-T nuclear fuel used in NPP Temelin is described in detail. Necessary nuclear programs and codes together with their sequences are described. The most general calculation, branch and history structure is proposed. The structure consists of 10 assemblies, 14 branches and 8 histories calculations. Last step of the cross-section processing is summarized in the last section. All calculations have to be combined and the partial derivatives have to be calculated. After this calculation, properly structured PARCS input has to be generated. This structured PARCS input file is the main aim of the whole process. Models based on

this scheme are developed at this time but the calculation is very time consuming and the future development will be focused on speedup and verification of the used models.

ACKNOWLEDGEMENT

This research work has been carried out in the Centre for Research and Utilization of Renewable Energy (CVVOZE). Authors gratefully acknowledge financial support from the Ministry of Education, Youth and Sports of the Czech Republic under NPU I programme (project No. LO1210). This paper contains the results of research works funded from project of specific research program of Brno University of Technology No. FEKT-S-11-19/1446.

REFERENCES

[1] Ward A., Xu Y., Downar T.: Code for Generating the PARCS Cross Section Interface File PMAXS, User`s manual, (2015), 105.

[2] Scale: A Comprehensive Modeling and Simulation Suite for Nuclear Safety Analysis and Design, ORNL/TM-2005/39, Version 6.1, June 2011. Available from Radiation Safety Information Computational Center at Oak Ridge National Laboratory as CCC-785.

[3] Scale: Lattice Physics and Depletion Training Course, Training manual, Paris, (2016), 342.

[4] B. J. Ade, "SCALE/TRITON Primer: A Primer for Light Water Reactor Lattice Physics Calculations," ORNL/TM-2011/21, NUREG/CR-7041, Oak Ridge National Laboratory, Oak Ridge, TN, January 2012.

Doktorské projekty

Biomedicínské inženýrství a bioinformatika

ACOUSTIC ANALYSIS OF SENTENCES COMPLICATED FOR ARTICULATION IN PATIENTS WITH PARKINSON'S DISEASE

Tomáš Kiska
Doctoral Degree Programme (1), FEEC BUT
E-mail: kiskatomas@phd.feec.vutbr.cz

Supervised by: Jiří Mekyska
E-mail: mekyska@feec.vutbr.cz

Abstract: This paper deals with acoustic analysis of hypokinetic dysarthria. Hypokinetic dysarthria is a speech motor dysfunction that is present in approximately 90% of patients with Parkinson's disease (PD). The work is mainly focused on parameterization techniques that can be used to diagnose or monitor this disease as well as estimate its progress. Acoustic analysis can be used to estimate a grade of hypokinetic dysarthria in fields of phonation, articulation, prosody and speech fluency. Regarding the parameterization, new features based on RASTA method were proposed. The analysis is based on parametrization of sentences complicated for articulation. Experimental dataset consists of 101 PD patients with different disease progress and 53 healthy controls. For the purpose of feature selection we employed mRMR (minimum Redundancy Maximum Relevance) method.

Keywords: Parkinson's disease, hypokinetic dysarthria, speech parameterization, speech signal processing, objective analysis, diagnosis, monitoring, progress estimation

1 INTRODUCTION

One of the most frequent neurodegenerative disease is Parkinson's disease (PD) [1]. This disease was firstly described by James Parkinson in his essay entitled "An Essay on the Shaking Palsy" in 1817. PD is a chronic idiopathic disorder with the unknown aetiology characterized by the progressive loss of dopaminergic neurons in part of basal ganglia called *substancia nigra pars compacta* [2]. This disorder is associated with motor deficits like resting tremor, muscular rigidity, postural instability and bradykinesia. In addition to these clinical signs, there is also a reduction of cognitive abilities, develop of behavioural alternations and many worldwide studies showed that one of the most frequent manifestation sign of patients with PD is speech disorder. In 60-90 % of these patients, we speak about hypokinetic dysarthria (HD) [3], which is characterized by monopitch, monoloudness, harsh or breathy voice and especially bad articulation.

Articulation disorders are the common manifestations of HD in patients with PD. This kind of disorder mainly occurs in consonants pronounciation. It is caused by an incorrect functionality of articulation organs such as the lips, tongue tip, middle of the tongue, the tongue, glottis or larynx [4].

Articulation problems of sentence units are explored in the work of Weismer et al. [5] In this case, patients affected by Amyotrophic lateral sclerosis (ALR) and mainly PD were analysed. The authors observed smaller vowel space area for PD then healthy speakers. Pell et al. [4] monitored pronunciation of sentence units and found out that the PD patients were characterized by more pronunciation errors than the healthy controls. Thanks to the articulation parameterization, they were able to differentiate PD patients and healthy speakers with 80 % accuracy.

This paper mainly deals with the analysis of sentences complicated for articulation. The aim of this study is to differentiate PD and healthy speech with the highest discrimination power in Czech language and introduce new speech features based on the Relative Spectral Transform (RASTA) method.

The article is organized as follows. Section 2 presents a dataset and methods used in this study. Feature extraction and analysis techniques are described here. Experimental results are given in section 3, and, finally, section 4 provides a discussion and some conclusions

2 PATIENTS AND METHODS

2.1 PATIENT DATABASE AND DATA ACQUISITION

A total of 154 Czech native speakers were tested. Altogether, 101 PD patients (60 men, 41 women; mean age 67.52 ± 8.29 years) and 53 gender matched healthy controls (26 men, 27 women; mean age 63.96 ± 9.21 years) were enrolled at the First Department of Neurology, St. Anne's University Hospital in Brno, Czech Republic. The healthy participants had no history or presence of speech disorders or brain diseases such as neurological and psychiatric illnesses. All participants signed an informed consent form that had been approved by the Ethics Committee of St. Anne's University Hospital in Brno. Each of the participants uttered 2 sets of five sentences complicated for the articulation. In the first set we monitored intelligibility of five sentences (9.2-1 to 5), while in the second set (another five sentences), we focused on fatigue of vocal organs (9.2-l-1 to 5). Both sets consist of the the same five sentences, however, the second one was uttered after approximately fifteen-minute speech training:

in Czech	English translation
9.2(-l)-1 *Do čtvrt hodiny tam byla smršť.*	*In a quarter of an hour there was a whirlwind.*
9.2(-l)-2 *Prohovořte to s ním dopodrobna.*	*Talk to him in detail.*
9.2(-l)-3 *Při ústupu pluku duní bubny.*	*Drums are pounding during the retreat of regiment.*
9.2(-l)-4 *Kuchařští učni nejsou jak zlatničtí.*	*Apprentices of cookery school are not as those from goldsmith one.*
9.2(-l)-5 *Celý večer se učí sčítat.*	*He is learning to add the whole evening*

All these sentences were read aloud. Speech samples were digitized by sampling frequency $f_s = 48\,\text{kHz}$ and consequently resampled to 16 kHz.

2.2 METHODOLOGY

The feature extraction stage included calculation of conventional parameters such as fundamental frequency F_0 [7], 6 kinds of shimmer and 5 kinds of jitter, short-time energy STE, Teager-Kaiser energy operator TKEO, low short-time energy ratio LSTER, median of power spectral density MPSD, formants F_1–F_3 and their bandwidths BW_1–BW_3, percentual pause ratio PPR, interpause speech duration ISD, total speech rate TSR, net speech rate NSR [8], articulation rate AR and harmonic-to-noise ratio HNR.

We also calculated features based on the method speech enhancement called RASTA. This method of uses digital bandpass filters in the range from 1 Hz to 16 Hz applied in speech spectrogram over time axis. Thanks to this, these fetaures quantify changes in position of the speech muscles during speech production. Therefore this method appears to be suitable to describe the articulation in patients with PD. However, in the case of our newly proposed method, we applied this filtering to the power spectral density of first and second formant. These features are labeled as E RASTA(F_x-y), where x represents order of formant and y number of filter bank (A total of 5 bands).

If the feature has been represented by vector, we have employed transformation to scalar value based on median, standard deviation (std), 1st percentile (1p), 99th percentile (99p) and interpercentile range (ir) defined as $99\text{p} - 1\text{p}$. For the purpose of speech feature extraction, we used Praat acoustic analysis software [9] and functions developed in MATLAB software.

To evaluate discrimination power of the sentences complicated for articulation we calculated mutual information (MI) and Spearman's rank sum correlation between resulting speaker's label and feature vectors. Then we employed Mann-Whitney U test and finally classification based on random forests.

Table 1: Preliminary individual statistical analysis of top features calculated for both genders separately.

Task code	Feature	ρ	MI	p	ACC [%]	SEN [%]	SPE [%]	TSS
Female participants								
9.2-1	BW_2 (min)	−0.0719	0.4801	0.0038	50.00	65.85	70.37	1.70
9.2-2	BW_2 (median)	−0.1521	0.2846	0.0171	47.06	65.85	77.78	1.75
9.2-3	BW_1 (max)	0.0698	0.2095	0.0035	57.35	**75.61**	66.67	1.75
9.2-4	TKEO (median)	0.1455	0.3750	0.0151	**60.29**	**75.61**	59.26	1.67
9.2-5	TKEO (median)	−0.1591	0.5983	0.0186	45.59	60.98	74.07	1.68
9.2-l-1	E RASTA(F_2-3)	0.0551	0.4801	0.0022	55.88	70.73	62.96	1.68
9.2-l-2	NST_{50}	−0.0220	0.2846	0.0004	52.94	70.73	70.37	1.74
9.2-l-3	E RASTA(F_1-5)	−0.0683	**0.8638**	0.0034	48.53	58.54	62.96	1.58
9.2-l-4	B_2 (ir)	0.0046	0.7271	**0.0000**	52.94	63.41	62.96	1.63
9.2-l-5	voicing nbreaks	**−0.3297**	0.0682	0.0888	41.18	65.85	**92.59**	**1.81**
Male participants								
9.2-1	relF_0SD	−0.0474	**1.0000**	0.0016	55.81	70.00	73.08	1.76
9.2-2	E RASTA(F_1-2)	−0.1026	**1.0000**	0.0078	53.49	68.33	76.92	1.77
9.2-3	F_0 (median)	0.0628	**1.0000**	0.0028	60.47	71.67	65.38	1.71
9.2-4	E RASTA(F_2-5)	−0.2083	0.3037	**0.0000**	60.47	**78.33**	76.92	**1.84**
9.2-5	TPT_{50}	0.1171	0.6147	0.0096	**63.95**	76.67	65.38	1.74
9.2-l-1	STE (min)	**−0.2693**	0.5176	0.0571	44.19	58.33	**84.62**	1.71
9.2-l-2	TKEO (var)	0.0027	0.6147	**0.0000**	59.30	73.33	73.08	1.78
9.2-l-3	F_0 (max)	0.0246	0.8693	0.0004	59.30	71.67	69.23	1.74
9.2-l-4	F_0 (max)	−0.1100	0.5176	0.0089	51.16	61.67	73.08	1.68
9.2-l-5	TKEO (1p)	−0.3112	0.2631	0.0763	40.70	53.33	**84.62**	1.65

ρ – Spearman's rank correlation coefficient; MI – mutual information; p – p-values of Mann–Whitney U test; ACC – classification accuracy; SEN – sensitivity; SPE – specificity; TSS – trade-off between sensitivity and specificity; 9.2-1 to 9.2-5 – speech tasks used for the intelligibility evaluation; 9.2-l-1 to 9.2-l-5 – speech tasks used for the analysis of the fatigue of vocal organs

The results were expressed by classification accuracy (ACC), sensitivity (SEN), specificity (SPE) and trade-off between sensitivity and specificity (TSS), which can be defined as:

$$\mathrm{TSS} = 2^{\sin(\frac{\pi \cdot \mathrm{SEN}}{2})\sin(\frac{\pi \cdot \mathrm{SPE}}{2})}. \tag{1}$$

A main problem of data analysis is presence of very large number of features, which inhibits detection of useful patterns underlying the data. Therefore, in the last step we performed classification along with minimum Redundancy Maximum Relevance (mRMR) algorithm [10], in order to reduce the dimensionality of input feature set and remove non-relevant features.

3 EXPERIMENTAL RESULTS

Altogether, we calculated more than 1100 speech features per subject. As can be seen in Tab. 1, the highest discrimination power for females (evaluated by accuracy) has TKEO (median) (ACC = 60.29 %, SEN = 75.61 %) extracted from task 9.2-4. voicing nbreaks with SPE = 92.59 % and TSS = 1.81 (9.2-l-5) is significant as well. On the other hand, in the case of males the most significant features is TPT_{50} with ACC = 63.95 % (9.2-5) and feature E RASTA(F_2-5) with the highest TSS = 1.84 (9.2-4). In 4 cases the most significant were new features based on RASTA.

The results of more complex evaluation based on feature selection can be seen in Tab. 2. In the first

Table 2: Classification results (using feature selection by mRMR) for both genders separately.

Task code	ACC [%]	SEN [%]	SPE [%]	TSS	No.
Female participants					
9.2-1	70.59	65.85	**77.78**	1.75	14
9.2-2	67.65	65.85	70.37	1.70	4
9.2-3	**82.35**	**90.24**	70.37	**1.84**	86
9.2-4	73.53	75.61	70.37	1.78	37
9.2-5	76.47	80.49	70.37	1.80	47
9.2-l-1	**73.53**	**78.05**	**66.67**	**1.76**	24
9.2-l-2	72.06	**78.05**	62.96	1.72	82
9.2-l-3	66.18	65.85	**66.67**	1.68	33
9.2-l-4	63.24	63.41	62.96	1.63	2
9.2-l-5	63.24	68.29	55.56	1.59	39
9.2-1 to 5	**82.35**	**90.24**	**70.37**	**1.84**	106
9.2-l-1 to 5	69.12	73.17	62.96	1.70	26
all	**82.35**	**90.24**	**70.37**	**1.84**	113
Male participants					
9.2-1	**79.07**	**83.33**	69.23	**1.81**	57
9.2-2	76.74	78.33	**73.08**	**1.81**	10
9.2-3	68.60	73.33	57.69	1.65	3
9.2-4	56.98	58.33	53.85	1.51	3
9.2-5	73.26	81.67	53.85	1.64	19
9.2-l-1	63.95	63.33	65.38	1.64	18
9.2-l-2	63.95	63.33	65.38	1.64	1
9.2-l-3	**69.77**	**73.33**	61.54	1.68	20
9.2-l-4	66.28	65.00	**69.23**	**1.69**	40
9.2-l-5	56.98	58.33	53.85	1.51	18
9.2-1 to 5	**74.42**	**76.67**	**69.23**	**1.77**	39
9.2-l-1 to 5	63.95	63.33	65.38	1.64	1
all	**76.74**	**78.33**	**73.08**	**1.81**	38

ACC – classification accuracy; SEN – sensitivity; SPE – specificity; TSS – trade-off between sensitivity and specificity; No. – number of selected features; 9.2-1 to 9.2-5 – speech tasks used for the intelligibility evaluation; 9.2-l-1 to 9.2-l-5 – speech tasks used for the analysis of the fatigue of vocal organs

scenario (individual task classification) the best results for males and females were obtained analysing speech tasks used for the intelligibility evaluation. Especially task 9.2-3 for female (ACC = 82.35 %, SEN = 90.24 %, SPE = 70.37 %) and 9.2-1 for male participants (ACC = 79.07 %, SEN = 83.33 %, SPE = 69.23 %). In the second scenario (classification within each task set) the best results were again obtained analysing the first set of sentences for both genders. Finally in the last scenario (classification using all task realizations) we obtained the best results using 113 features (ACC = 82.35 %, SEN = 90.24 %, SPE = 70.37 %) for females and 38 features for males (ACC = 76.74 %, SEN = 78.33 %, SPE = 73.08 %).

4 CONCLUSION

The aim of this study was analyse of sentences complicated for articulation in patients with PD. We studied 2 sets of five sentences tasks (the intelligibility evaluation, the analysis of the fatigue of vocal organs) of 101 PD patients and 53 gender matched healthy controls.

We have achieved all goals of this work. We have found out that the task set for the intelligibility evaluation outperformed the analysis of the fatigue of vocal organs. We have proved that the classification accuracy can be significantly improved by the new features based on RASTA. It was shown that articulation features can be used for diagnosis of PD with classification accuracy over 82 % for female and 76 % for male participants.

In the near future we are going to introduce and analyse new articulation features. To better select speech features that provide maximal clinical information we will use the sequential forward feature selection algorithm (SFFS). Articulation features will be also analysed in terms of PD progression monitoring.

ACKNOWLEDGEMENT

Research described in this paper was financed by the National Sustainability Program under grant LO1401 and by project COST IC1206. For the research, infrastructure of the SIX Center was used.

REFERENCES

[1] L. O. Ramig, C. Fox, S. Sapir, Speech treatment for Parkinson's disease, Expert Rev Neurother 2 (2008) 297–309

[2] A. M. Goberman, C. Coelho Acoustic analysis of parkinsonian speech I: speech characteristics and L-Dopa therapy, NeuroRehabilitation 17 (2002): 237–246.

[3] S. Cantiniaux, M. Vaugoyeau, D. Robert, C. Horrelou-Pitek, C. Mancini, Comparative analysis of gait and speech in Parkinson's disease: hypokinetic or dysrhythmic disorders?, J Neurol Neurosurg Psychiatry 81 (2010) 177-184.

[4] J. Mekyska, Z. Smekal, M. Kostalova, M. Mrackova, S. Skutilova, I. Rektorova, Motor aspects of speech imparment in parkinson's disease and their assessment, Cesk Slov Neurol N 74 (2011) 662–668.

[5] G. Weismer, J. Y. Jeng, J. S. Laures, R. D. Kent, J. F. Kent, Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders, Folia Phoniatrica et Logopaedica 1 (2001) 1-18.

[6] M. D. Pell, H. S. Cheang, C. L. Leonard, The impact of parkinsons disease on vocal-prosodic communication from the perspective of listeners, Brain and Language 2 (2006) 123–134.

[7] I. Rektorova, J. Barrett, M. Mikl, I. Rektor, T. Paus, Functional abnormalities in the primary orofacial sensorimotor cortex during speech in Parkinson's disease, Mov Disord 22 (2007) 2043–2051.

[8] S. Skodda, H. Rinsche, U. Schlegel, Progression of dysprosody in Parkinson's disease over time – a longitudinal study, Movement Disorders 5 (2009) 716-722.

[9] P. Boersma, D. Weenink, Praat: Doing Phonetics by Computer (Version 5.3.10), 2012.

[10] A. Tsanas, M. A. Little, P. E. McSharry, A methodology for the analysis of medical data, Chapter 7 in Handbook of Systems and Complexity in Health, Springler, 2013.

COMPARATIVE ANALYSIS OF COMMON AND UNIQUE TARGETS IN DRUG RESISTANT STRAIN OF *BORRELIA BURGDORFERI*

Pavlína Koščová

Doctoral Degree Programme (2), FEEC BUT

E-mail: koscova@phd.feec.vutbr.cz

Supervised by: Ivo Provazník

E-mail: provaznik@feec.vutbr.cz

Abstract: The number of drug-resistant strains of *Borrelia burgdorferi* necessitated the identification of potential drug targets specific to the strain of interest. The chromosomal and plasmid genes of *B. burgdorferi* strain B31 were compared with erythromycin-resistant *B. burgdorferi* strain N40 to find common (core) and unique (strain-specific) genes in present study. *In silico* analysis of genomic data showed total number of unique genes higher in strain N40. The presence of higher number of unique genes in N40 signifies their role in drug resistance mechanism. Furthermore, human proteome was compared with proteome of these strains to find target protein specific to the strain of interest and not present in host. In conclusion, identification of unique genes in these strains provided on differences in drug resistance potential.

Keywords: *Borrelia burgdorferi*, Lyme disease, erythromycin, resistance

1. INTRODUCTION

Lyme borreliosis (Lyme disease) is an infectious disease of multistage disorders. There are 12 (out of the 37) species of genus *Borrelia* (*B.*) known to cause Lyme disease or borreliosis. The major *Borrelia* species causing Lyme disease in Europe and Asia are *B. burgdorferi*, *B. afzelii*, and *B. garinii*, while *B. burgdorferi* is also known as the main causative agent of Lyme disease in North America [1]. This strain was described as the agent of Lyme disease in 1984. This spirochete bacterium is transmitted to a wide range of mammals by the *Ixodes* genus [2].

Lyme borreliosis is difficult to diagnose at any stage and difficult to treat during the later symptoms as well. Early symptoms typically occur after contact with the infected thick and include fever, chills, myalgia, fatigue, headache, and malaise. Late stage has many manifestations including intermittent pain and swelling of joints, occasionally, neurologic manifestations can occur as well [3].

Many different antibiotics are used during the early and late stages of the disease. Moreover, it has been indicated, that antibiotic treatment of borreliosis during the early stages of the disease such as azithromycin, ceftriaxone or doxycycline is very efficacious and has a long-term benefit for the patient [4]. For intolerant patients of first-line therapeutics, macrolide antibiotics as erythromycin can be recommended [5]. However, it has been shown that some *B. burgdorferi* strains are able to express resistance to erythromycin [6]. Some studies have been performed for characterization of this resistance [7] but further analysis are still needed.

Use of the information provided by genome sequences to better understand host-pathogen interactions has proved to be the key for new vaccinology approaches [8]. There is an extensive need to find strain-specific genes responsible for differences in antibiotic resistance in *B. burgdorferi*. The aim of present study was to use specific computational tools and identify strain-specific unique genes important for differences in antibiotic resistance *B. burgdorferi* strains. We performed *in silico* analysis to find the differences in erythromycin-resistant strain of *B. burgdorferi* responsible for

varied drug response to this antibiotic [6]. Our major interest was to know about genetic differences in erythromycin-resistant strain N40 with other erythromycin-sensitive strain B31. First, we identified common and unique genes present in both strains. Next, both strains were compared with human proteome to find potential drug targets in each strain.

This analysis is based on the research project described in the chapter "Comparative Analysis of Common and Unique Targets in Drug Resistant Strains of *Staphylococcus aureus*" published in the book "Frontier Discoveries and Innovations in Interdisciplinary Microbiology" [9].

2. MATERIALS AND METHODS

2.1. GENOMIC DATA ANALYSIS

Information about chromosomal and plasmid genes of *B. burgdorferi* strains N40 (accessions, NC_017418.1 and NC_017401.1), and B31 (accessions, NC_001318.1 and NC_001903.1) were collected from the NCBI genome database (http://www.ncbi.nlm.nih.gov/genome/) (Table 1). Differences between strains at the genome level were elucidated through *in silico* analysis - the strains were compared for common and unique genes using CGUG 3.5. The tool CGUG is available at http://binf.gmu.edu/geneorder.html [10], and provides iterative BLASTp analyses using a reference genome and up to four query genomes. Genomic sequences of studied strains were compared using CGUG 3.5 with BLASTp threshold set at 75.

organism	strain	chromosome	plasmid	size (Mb)	specification
B. burgdorferi	N40	NC_017418.1	NC_017401.1	1.33954	erythromycin-resistant
B. burgdorferi	B31	NC_001318.1	NC_001903.1	1.52121	erythromycin-sensitive

Table 1: Summary of studied strains

2.2. *IN SILICO* DRUG TARGET PREDICTION

Protein-coding genes of *B. burgdorferi* strains were compared against human non-pathogenic microbes using UniDrug-Target tool which is intended to identify bacterial pathogen-specific drug targets along with unique target's functionality information [11]. The provided tool was used to find out common and specific target in given sequence. This server is currently available at http://14.139.240.55/UDT/main.html and stand-alone application source codes are available at http://www.bioinformatics.org/ftp/pub/bioinfojuit/UDT.rar. Host proteome microbes were compared against *B. burgdorferi* strains N40 and B31 to identify strain-specific potential drug targets.

2.3. *IN SILICO* ANALYSIS OF ESSENTIAL GENES

BLASTp analysis against human proteome was executed to find essential genes in Database of essential genes (DEG) [12].

3. RESULTS

3.1. *IN SILICO* IDENTIFICATION OF COMMON AND STRAIN-SPECIFIC GENES IN ERYTHROMYCIN-RESISTANT STRAIN OF *BORRELIA BURGDORFERI*

In *B. burgdorferi* strain N40, the chromosomal genes showed several unique genes of following functions: lipoprotein releasing factor, phosphatidate cytidylyltransferase, acriflavin resistance protein, nucleotidyltransferase, endonuclease, haloacid dehalogenase, 23S rRNA (guanosine(2251)-2'-O)-methyltransferase RlmB, NAD+ synthetase, TIGR02757 family protein, KTSC domain-containing protein, exodeoxyribonuclease V subunit gamma, DNA primase, pseudouridine synthase and seven genes of hypothetical proteins, whereas plasmid genes showed only two genes of hypothetical proteins. In *B. burgdorferi* strain B31, the chromosomal genes showed one gene of hypothetical protein and plasmid genes showed no unique genes. Comparative analysis data is

shown in Table 2 and Fig. 1. The number of computed genes of hypothetical proteins was 187 in *B. burgdorferi* strain N40 and 218 in *B. burgdorferi* strain B31.

Name of the strain	**Strain-specific genes in chromosome and their function**	**Strain-specific genes in plasmid and their function**
Accession	NC_017418.1	NC_017401.1
B. burgdorferi strain N40	GI:504353063 hypothetical protein GI:488734875 lipoprotein releasing factor GI:530323927 phosphatidate cytidyl-yltransferase GI:504353489 acriflavin resistance protein GI:504353512 nucleotidyltransferase GI:504353520 hypothetical protein GI:488739146 hypothetical protein GI:504353563 hypothetical protein GI:488733631 endonuclease GI:488732658 hypothetical protein GI:488740701 haloacid dehalogenase GI:497940846 23S rRNA (guano sine(2251)-2'-O)-methyltransferase RlmB GI:488738110 NAD+ synthetase GI:504353602 TIGR02757 family protein GI:488732889 hypothetical protein GI:488620397 KTSC domain-containing protein GI:740589237 hypothetical protein GI:504353651 exodeoxyribonuclease V subunit gamma GI:504353678 DNA primase GI:754989102 pseudouridine synthase	GI:497941997 hypothetical protein GI:497942008 hypothetical protein
Accession	NC_001318.1	NC_001903.1
B. burgdorferi strain B31	GI:15594483 hypothetical protein	x

Table 2: List of strain-specific/unique genes in chromosomal and plasmid genes of *Borrelia burgdorferi* strains N40 and B31

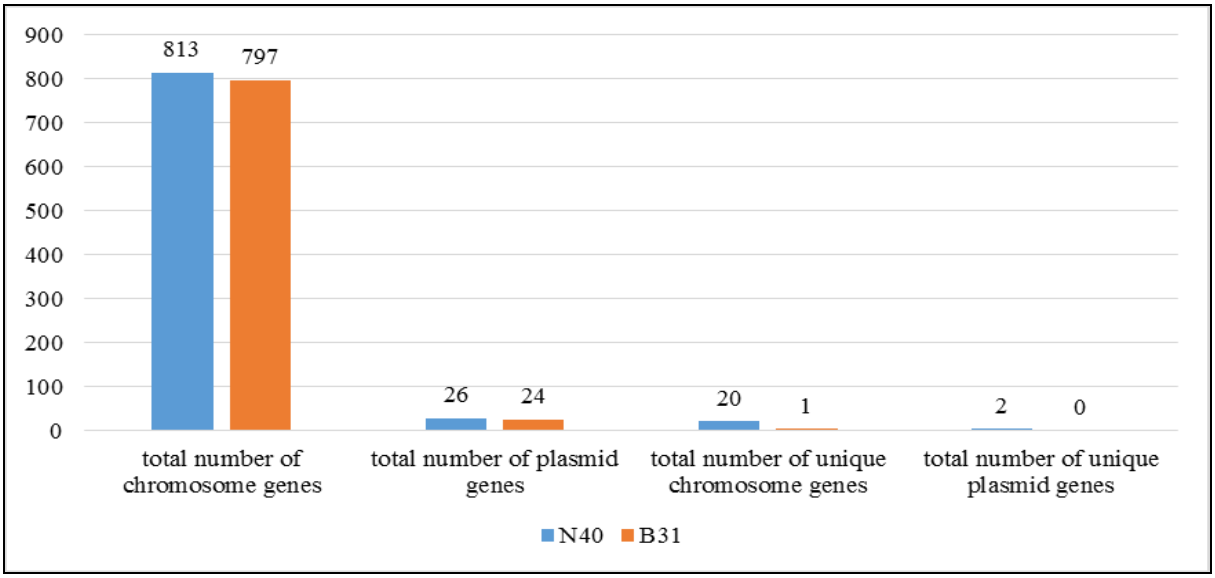


Figure 1: Chromosomal genes data of *B. burgdorferi* strain N40 compared through analysis with strain B31

3.2. COMPARATIVE GENOMICS ANALYSIS OF *BORRELIA BURGDORFERI* WITH HUMAN NON-PATHOGENIC MICROBES

DEG database searches showed high number of homologue genes. Results are shown in Table 3. UniDrug Target tool analysis of *B. burgdorferi* strains showed high number of unique proteins (with regard to human non-pathogenic proteomes). Results are shown in Table 4.

Name of the strain	Number of homologs (chromosome genes)	Number homologs (plasmid genes)
B. burgdorferi strain N40	1243	214
B. burgdorferi strain B31	1390	210

Table 3: *In silico* prediction of homologue proteins with respect to human proteome (DEG)

Name of the strain	Unique proteins with regard to human proteome	Matching similarity score in unique proteins with non-pathogenic proteome
B. burgdorferi strain N40	302	0-29.6 %
B. burgdorferi strain B31	294	0-28.8 %

Table 4: *In silico* prediction of drug targets in *Borrelia burgdorferi* with regard to human non-pathogenic proteome (UniDrug-Target tool)

3.3. DISCUSSION

Drug resistance development of *B. burgdorferi* strains to available antibiotics is an important point of interest among health practitioners. According to NCBI genome database information (http://www.ncbi.nlm.nih.gov/genome/genomes/738), the total number of genome sequences of *B. burgdorferi* strains assembled and annotated is currently 42. Genome database showed 252 plasmid annotation data among genome sequences of *B. burgdorferi* strains annotated till date.

B. burgdorferi strain N40 showed erythromycin-resistance in previous study and was taken as a reference strain to compare with strain B31 which showed erythromycin-sensitivity in previous study [6].

Understanding the differences between *B. burgdorferi* genomes and the controls that govern these changes could help to predict new drug targets of these strains. The present study identified the differences in terms of presence and absence of drug-resistant genes in the erythromycin-resistant *B. burgdorferi* strain N40. Several genes for specific proteins were found as strain-specific genes in *B. burgdorferi* strain N40. Further *in silico* analysis identified the potential drug targets.

4. CONCLUSION

The present study identified the gene specific differences in the studied *Borrelia* strains. The analysis identified the potential unique target proteins in these strains, which are non-homologues to human. We found more numbers of hypothetical proteins in N40 strains, suggesting their functional role in the organism responsible for erythromycin-resistance. Functional characterization of these hypothetical proteins could help in determining their importance in drug resistance differences in these *Borrelia* strains in the future and lead to the discovery and development of novel vaccines.

Bacterial infections remain a major cause of deaths and disabilities in the world. Patients treated with the proper antibiotic dose respond to the treatment the most effectively. This study demonstrates the usefulness of the predictive tools for better characterization of diseases' mechanisms. Anyway, there is still highlighted need to validate computed predictions experimentally.

REFERENCES

[1] JOHNSON, Russell C., George P. SCHMID, Fred W. HYDE, A. G. STEIGERWALT a Don J. BRENNER. *Borrelia burgdorferi* sp. nov.: etiologic agent of Lyme disease. *International Journal of Systematic and Evolutionary Microbiology*. 1984, roč. 34, č. 4, s. 496–497.

[2] BARBOUR, ALAN G. Isolation and cultivation of Lyme disease spirochetes. *The Yale journal of biology and medicine*. 1984, roč. 57, č. 4, s. 521.

[3] BRATTON, ROBERT L, MD; WHITESIDE, JOHN W, MD; HOVAN, MICHAEL J, MD; ENGLE, RICHARD L, MD; EDWARDS, FREDERICK D, MD. Diagnosis and Treatment of Lyme Disease. *Mayo Clinic Proceedings*. 2008, roč. 83, č. 5, s. 566–571.

[4] STRAUBINGER, Reinhard K., Alix F. STRAUBINGER, Brian A. SUMMERS a Richard H. JACOBSON. Status of *Borrelia burgdorferi* Infection after Antibiotic Treatment and the Effects of Corticosteroids: An Experimental Study. *The Journal of Infectious Diseases*. 2000, roč. 181, č. 3, s. 1069–1081. ISSN 0022-1899, 1537-6613.

[5] GARY P. WORMSER, ROBERT B. NADELMAN, RAYMOND J. DATTWYLER, DAVID T. DENNIS, EUGENE D. SHAPIRO, ALLEN C. STEERE, THOMAS J. RUSH, DANIEL W. RAHN, PATRICIA K. COYLE, DAVID H. PERSING, DURLAND FISH a BENJAMIN J. LUFT. *Practice Guidelines for the Treatment of Lyme Disease*. B.m.: Clinical Infectious Diseases Society of America. 2000

[6] TEREKHOVA, D., M. L. SARTAKOVA, G. P. WORMSER, I. SCHWARTZ a F. C. CABELLO. Erythromycin Resistance in *Borrelia burgdorferi*. *Antimicrobial Agents and Chemotherapy*. 2002, roč. 46, č. 11, s. 3637–3640. ISSN 0066-4804, 1098-6596.

[7] JACKSON, Charlene R., Julie A. BOYLAN, Jonathan G. FRYE a Frank C. GHERARDINI. Evidence of a conjugal erythromycin resistance element in the Lyme disease spirochete *Borrelia burgdorferi*. *International Journal of Antimicrobial Agents*. 2007, roč. 30, č. 6, s. 496–504. ISSN 09248579.

[8] MORIEL, Danilo Gomes, Maria SCARSELLI, Laura SERINO, Marirosa MORA, Rino RAPPUOLI a Vega MASIGNANI. Genome‐based vaccine development: A short cut for the future. *Human Vaccines*. 2008, roč. 4, č. 3, s. 184–188. ISSN 1554-8600, 1554-8619.

[9] KAHLON, Amandeep Kaur, Mahendra P. DAROKAR a Ashok SHARMA. Comparative Analysis of Common and Unique Targets in Drug Resistant Strains of Staphylococcus aureus. In: Pratyoosh SHUKLA, ed. *Frontier Discoveries and Innovations in Interdisciplinary Microbiology*. New Delhi: Springer India, 2016, s. 193–205. ISBN 978-81-322-2609-3.

[10] MAHADEVAN, Padmanabhan, John F KING a Donald SETO. CGUG: *in silico* proteome and genome parsing tool for the determination of „core" and unique genes in the analysis of genomes up to ca. 1.9 Mb. *BMC Research Notes*. 2009, roč. 2, č. 1, s. 168. ISSN 1756-0500.

[11] CHANUMOLU, Sree Krishna, Chittaranjan ROUT a Rajinder S. CHAUHAN. UniDrugTarget: A Computational Tool to Identify Unique Drug Targets in Pathogenic Bacteria. *PLoS ONE*. 2012, roč. 7, č. 3, s. e32833. ISSN 1932-6203.

[12] ZHANG, R. DEG: a database of essential genes. *Nucleic Acids Research*. 2004, roč. 32, č. 90001, s. 271D–272. ISSN 1362-4962.

EEG SIGNAL ANALYSIS BASED ON EMD AND DISCRETE ENERGY SEPARATION ALGORITHM

Tomáš Potočňák

Doctoral Degree Programme (3), FEEC BUT

E-mail: xpotoc00@stud.feec.vutbr.cz

Supervised by: Jiří Kozumplík

E-mail: kozumplik@feec.vutbr.cz

Abstract: This paper deals with spectral analysis of nocturnal EEG signal from apnoea/hypopnea patients. Main goal is to employ methods independent to Fourier Transform, because of nonstationary character of signal, to better description of frequency changes. For this purpose, analysis based on Empirical Mode Decomposition and Discrete Energy Separation Algorithm was tested. This method is similar to commonly used Hilbert Huang Transform, but can provide higher time and frequency resolution due to algorithms based on Teager-Keiser Energy Operator, which can work with very short time window.

Keywords: nocturnal EEG, Empirical Mode Decomposition, Teager-Keiser Energy Operator, Discrete Energy Separation Algorithm

1. INTRODUCTION

Nowadays, there are still some unsolved tasks in electroencephalogram (EEG) analysis, especially in frequency and time-frequency domain. Due to the considerable non-stationary and nonlinear character of EEG, commonly used methods based on Fourier transform (FT) provide the resulting frequency spectrum with only little physical sense [1]. Therefore, several methods to analyse spectrum of such signals have been developed in past decades.

Hilbert Huang Transform (HHT) is one of these methods. HHT is based on empirical mode decomposition (EMD), followed by Hilbert transform to compute Hilbert spectrum [1]. This method forms a significant inspiration for our work. The main goal was to replace rather difficult Hilbert transform with simpler frequency determination method. Discrete Energy Separation Algorithm (DESA) can be considered as such [2], [3].

Implemented methods are described in Chapter 2. Chapter 3 considers practical use on real EEG signal.

2. METHODS

2.1. EMPIRICAL MODE DECOMPOSITION

EMD is an empirical method used to decompose a multicomponent signal into a number of signal components in the time domain called intrinsic mode functions (IMF). Each IMF represents a bandwidth of frequencies of the signal, so the EMD method can be used for removing unwanted components of the signal being analysed. Linearity of this method can be considered as strong aspect, as well as decomposition based on signal character and not just on uniform frequency bandwidths [1], [2].

The EMD procedure for extracting IMF is called the *sifting process* and consists of the following steps [1], [2]:

- finding of local maxima and local minima of signal *x(t)*,
- constructing of upper (maxima) envelope *u(t)*, and lower (minima) envelope *l(t)*,
- estimation of mean envelope $m_1(t)$ as a mean value of *u(t)* and *l(t)*, then first possible component is given by the equation $h_1(t) = x(t) - m_1(t)$.

The component $h_1(t)$ is accepted as the first IMF only if two conditions are satisfied. First, the number of extrema and the number of zero crossings over the entire length of the IMF must be equal or differ at most by one, and second, at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero. Otherwise $h_1(t)$ is treated as the new data set and $m_{11}(t)$ is estimate, that is used to calculate $h_{11}(t)$. The *sifting process* is repeated until the component $h_{1k}(t)$ is accepted as an IMF of the signal *x(t)* and is denoted by $C_1(t)$

$$C_1(t) = h_{1k}(t) = h_{1(k-1)}(t) - m_{1k}(t). \tag{1}$$

- IMF is subtracted from original signal *x(t)* resulting residual signal $r_N(t)$. After that, *sifting process* repeats with residual signal as input. The process ends when the last residual signal, $r_N(t)$ is obtained and it is a constant or a monotonic function.

2.2. DISCRETE ENERGY SEPARATION ALGORITHMS

These algorithms are based on the Teager-Kaiser Energy Operator (TKEO). Assume that TKEO for discrete case is [3]:

$$\psi[x(n)] = x^2(n) - x(n-1)\cdot x(n+1) \approx A^2\omega^2, \tag{2}$$

where $x(n) = A\cdot\cos(\Omega(n))$ is input signal with amplitude *A* and $\Omega(n) = \omega\cdot T_s$ (T_s – sampling period) according to [3].

TKEO offers excellent time resolution, because only three samples are required for the energy computation at each time instant. From (4), it is obvious that TKEO reflects not only amplitude of analysed signal, but also its instantaneous frequency $\omega = 2\pi f$. Based on this knowledge, DESA algorithms can be estimated [2], [3].

First DESA-2 algorithm, where '2' implies the approximation of first-order derivatives by differences between samples whose time indices differ by 2, estimates time varying frequency and amplitude envelope according to [3], as:

$$\Omega(n) \approx \arcsin\left(\sqrt{\frac{\psi[x(n+1)-x(n-1)]}{4\cdot\psi[x(n)]}}\right);\ |A(n)| \approx \frac{2\cdot\psi[x(n)]}{\sqrt{\psi[x(n+1)-x(n-1)]}} \tag{3}$$

By implementation of *arcsin (u)*, usage of DESA-2 was limited to frequencies ≤ *1/4* of the sampling frequency (assumes that $|u| \le \pi/2$).

An alternative is DESA-1 algorithm, which according to [3] replaces derivatives with backward and forward differences:

$$y(n) = x(n) - x(n-1),\ y(n+1) = x(n+1) - x(n), \tag{4}$$

such replacement yields new estimations as

$$\Omega(n) \approx \arccos(G(n)),\ |A(n)| \approx \sqrt{\frac{\psi[x(n)]}{1-G^2(n)}}, \tag{5}$$

where

$$G(n) = 1 - \frac{\psi[y(n)] + \psi[y(n+1)]}{4\cdot\psi[x(n)]}, \tag{6}$$

which provides benefits within frequency range ≤ *1/2* of the sampling frequency (*arccos (u)*) and according to previous studies slightly lower error rate [3].

DESA-1 algorithm in theory can be considered as significant improvement compared to DESA-2. In this study, both algorithms were tested.

2.3. EMD – DESA SPECTROGRAM

Final algorithm combines EMD and DESA to compute amplitude 'spectrogram' of EEG signal. First, entering segment of EEG was decomposed to various number of IMFs based on conditions mentioned in chapter 2.1. Then DESA algorithm to estimate instantaneous frequency and amplitude characteristics for each IMF was applied. Finally, spectrogram matrix was constructed based on instantaneous frequency, amplitude and time position from each IMF for the same EEG segment.

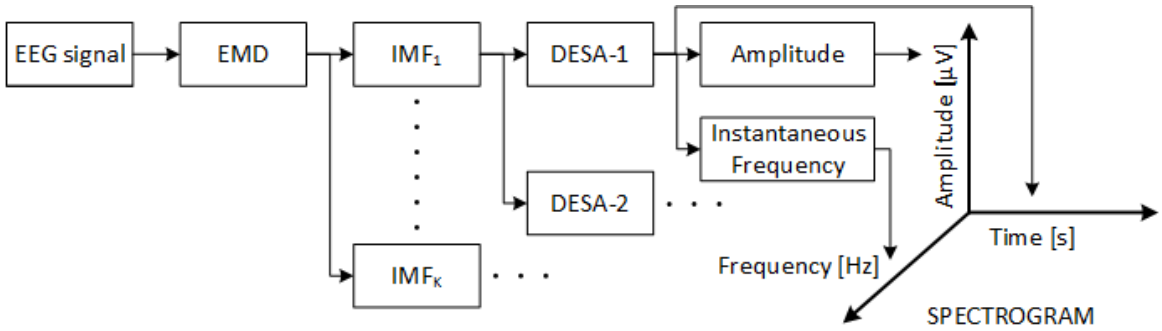


Figure 1: Calculation of spectrogram using EMD – DESA algorithm diagram

Mentioned algorithms were tested on real EEG signals from nocturnal polysomnographic examination of patients with sleep apnoea/hypopnea. All EEGs were sampled with frequency $f_s = 256\,Hz$ and filtered (lowpass at 70 Hz, highpass at 0.3 Hz and notch filter at 50 Hz). As additional pre-processing, another filtration with bandpass filter (bandwidth 0.5 – 30 Hz) was used. Segments (sleep epochs lasting 30 s) of EEG signal with high power noise (mostly due to patient movement) and identified as unclassified by PSG expert were excluded from analysis.

Obtained spectrograms were compared with that calculated using short-time Fourier transform (STFT) with Hamming window of a length 256 samples and 50% overlap (Figure 4).

3. RESULTS

EEG epochs were processed by basic EMD algorithm (chapter 2.1). Number of outcome IMFs (Figure 2 - right) varies trough the phase of sleep and patients from 9 to 16. EMD provides data with less noisy and more homogeny character than original EEG (Figure 2 - left) except 1st and sometimes 2nd IMF usually representing high frequency noise.

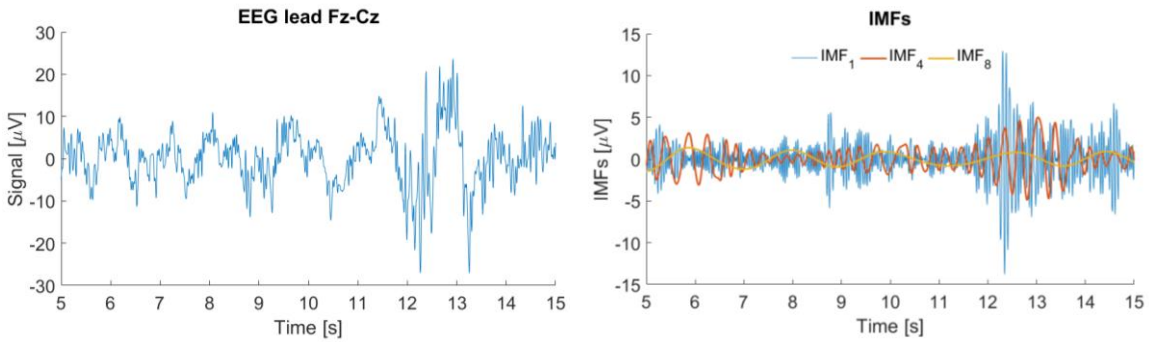


Figure 2: Detail of EEG segment and example of outcome IMFs

For each calculated IMF, instantaneous frequency and amplitude characteristics were estimated according to DESA algorithms (Chapter 2.2). DESA-1 and DESA-2 were tested. Because both of these algorithms works with 5-sample window, first and last two samples of epoch ware approximated by the closest value. Frequency precision was set on 0.1 Hz by simple rounding to one decimal place. Amplitude precision was set on 0.01 μV same way. DESA-1 and DESA-2 algorithms yield mostly same results with occasional local extremes.

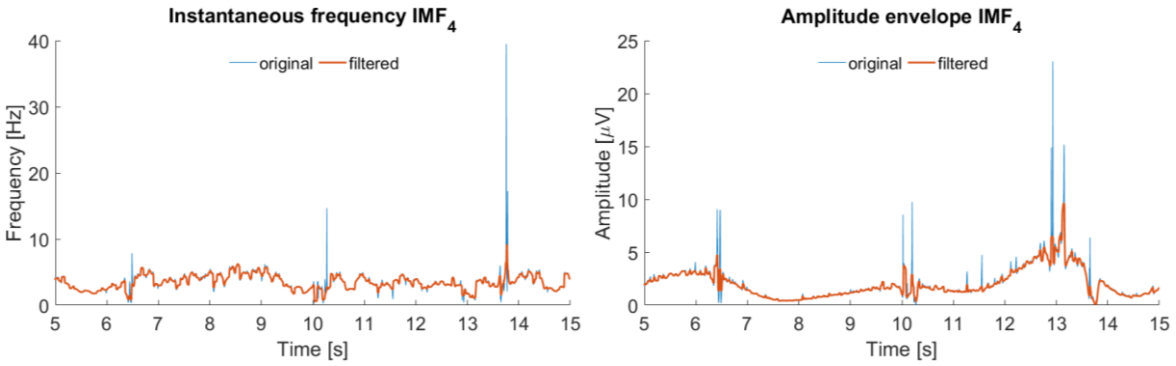


Figure 3: Detail of frequency and amplitude characteristics of IMF_4

After estimation of frequency and amplitude characteristics, final amplitude spectrogram (Figure 4) was created. To do that, frequency and time parts of IMF were used as the positions for x and y-axes, respectively. Amplitude values (z) were calculated as average values through all amplitudes associated with particular frequency and time instances. Median filtration (with 9 samples window) was applied as first choice to minimize mentioned local extremes (Figure 3, Figure 5).

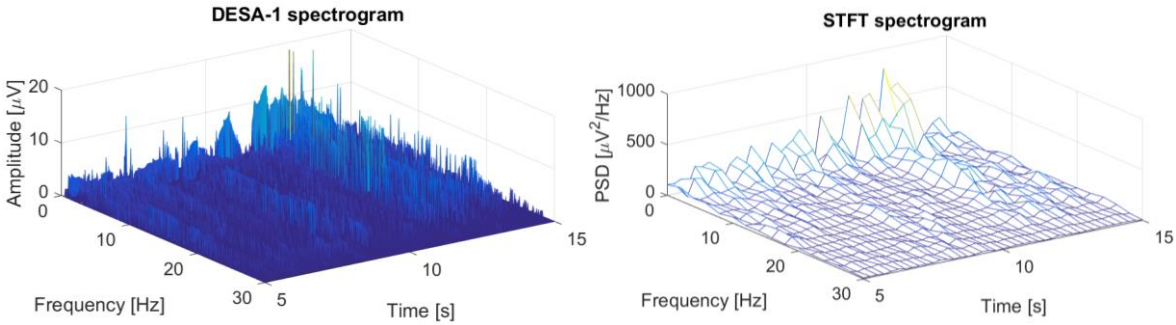


Figure 4: Detail of resulting EMD – DESA and STFT spectrograms

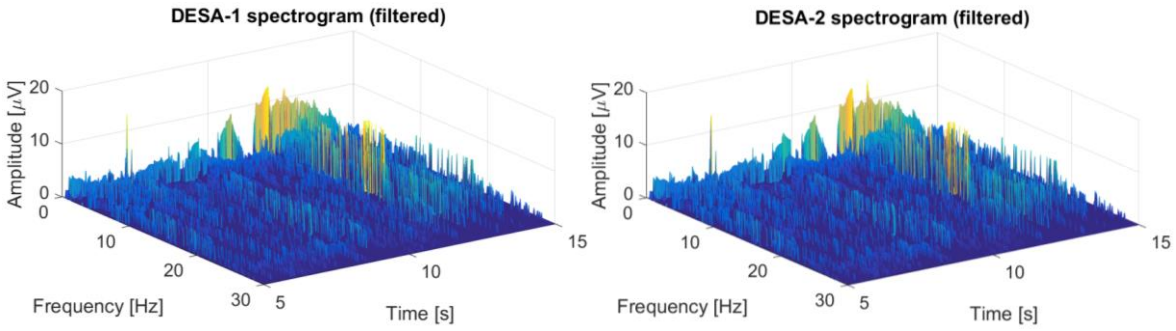


Figure 5: Detail of filtered EMD – DESA spectrograms

4. CONCLUSION

The problem of correct estimation of spectrogram for nonstationary signals was introduced. Main issue consists in use of Fourier Transform, which expects stationary harmonic signal as input. The EMD – DESA spectrogram can provide one of the possible solutions of this problem. In this paper,

theoretical background and practical implementation of mentioned method were presented. From comparison of results obtained by EMD-DESA and STFT(Figure 4), it is evident, that both types of spectrograms have very similar character, even with different representation of z-axes. Nevertheless, EMD – DESA spectrogram provides higher frequency and time resolution compared with STFT spectrogram.

Mentioned benefits of proposed method can provide us powerful tool for time frequency analysis of nonstationary signals such as EEG. It may have great potential in such areas as automatic sleep stages recognition, where various parameters are calculated from EEG spectrogram and used as inputs for classification algorithm. High resolution of EMD-DESA spectrogram provides better localization of some important patterns, which correlates with different sleep phases (e.g. K-complexes, spindles, etc.). For example, low frequency components of original EEG (Figure 2) are better distinguished even in first 5 seconds of EMD – DESA spectrogram (Figure 5) opposite to STFT spectrogram (Figure 3) which suffers frequency leakage. Such quality can be crucial for reaching successful performance of sleep stages distinguishing, which is main objective of following study.

In other studies, DESA algorithms were successfully applied in speech analysis and multi-tone signal detection [3], [5]. Combination of EMD and DESA was tested in industry sector [2] and compared with HHT, where higher time resolution and less computation complexity of EMD – DESA spectrogram was confirmed. Main restrictions of method are limited frequency range (Chapter 2.2), occasional noisy character (Figure 4), and "low frequency error" [4]. As was shown above, noisy character of resulting amplitude and frequency characteristics can be supressed by simple filtration. Ensemble EMD (EEMD) can be applied in future to improve noise separation during decomposition process.

REFERENCE

[1] HUANG, N. E., Z. SHEN, S. R. LONG, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*. IEEE, 1998, 454(1971), 903-995. DOI: 10.1098/rspa.1998.0193. ISBN 0-7803-2127-8. ISSN 1364-5021

[2] ANTONIADOU, I., G. MANSON, W.J. STASZEWSKI, T. BARSZCZ a K. WORDEN. A time–frequency analysis approach for condition monitoring of a wind turbine gearbox under varying load conditions. *Mechanical Systems and Signal Processing*. 2015, 64-65(10), 188-216. DOI: 10.1016/j.ymssp.2015.03.003. ISSN 08883270

[3] MARAGOS, P., J.F. KAISER, T.F. QUATIERI, et al. On separating amplitude from frequency modulations using energy operators. [Proceedings] ICASSP-92: 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing. IEEE, 1992, 454(1971), 1-4 vol.2. DOI: 10.1109/ICASSP.1992.226135. ISBN 0-7803-0532-9. ISSN 1364-5021

[4] LIPSEY, M.J., J.P. HAVLICEK, W.J. STASZEWSKI, T. BARSZCZ a K. WORDEN. On the Teager-Kaiser energy operator "low frequency error". *The 2002 45th Midwest Symposium on Circuits and Systems*, 2002. MWSCAS-2002. IEEE, 2002, 64-65(10), III-53-III-56. DOI: 10.1109/MWSCAS.2002.1186968. ISBN 0-7803-7523-8. ISSN 08883270

[5] VELEZ, E.F., J.P. HAVLICEK, W.J. STASZEWSKI, T. BARSZCZ a K. WORDEN. Detection of multi-tone signals based on energy operators. *Proceedings of IEEE-SP International Symposium on Time- Frequency and Time-Scale Analysis*. IEEE, 1994, 64-65(10), 229-232. DOI: 10.1109/TFSA.1994.467251. ISBN 0-7803-2127-8. ISSN 08883270

POTENTIAL OF PROSODIC FEATURES TO ESTIMATE DEGREE OF PARKINSON'S DISEASE SEVERITY

Zoltán Galáž
Doctoral Degree Programme (2), FEEC BUT
E-mail: z.galaz@phd.feec.vutbr.cz

Supervised by: Jiří Mekyska
E-mail: mekyska@feec.vutbr.cz

Abstract: This paper deals with non-invasive and objective Parkinson's disease (PD) severity estimation. For this purpose, prosodic speech features expressing monopitch, monoloudness, and speech rate abnormalities were extracted from recordings of stress-modified reading task acquired from 72 patients with idiopathic PD. Using a single feature regression (esimating values of subjective clinical rating scales) with classification and regression algorithm, following performance in terms of root mean squared error was achieved: 10.72 (UPDRS III), 2.16 (UPDRS IV), 4.76 (FOG-Q), 17.89 (NMSS), 2.13 (RBDSQ), 6.43 (ACE-R), 1.41 (MMSE), and 4.82 (BDI). These results show a promising potential of prosodic speech features in the field of objective assessment of PD severity.

Keywords: Parkinson's disease, hypokinetic dysarthria, dysprosody, objective assessment

1 INTRODUCTION

Parkinson's disease (PD) is a frequent neurodegenerative disorder [1] characterized by a substantial reduction of dopaminergic neurons especially in the part of the brain named *substancia nigra pars compacta* [2]. Dopaminergic loss subsequently leads to a malfunction of the central nervous system resulting into variety of associated motor and non-motor symptoms. According to the previous studies [3], approximately 60 – 90 % of PD patients suffer from speech disorder, defined by Darley, Aronson, and Brown back in 1969, referred to as hypokinetic dysarthria (HD) [4]. HD affects all aspects of human verbal communication, namely: phonation, articulation, prosody, speech fluency, and faciokinesis [5, 6, 7]. In summary, voice tremor, hypophonia, dysphonia, hypernasality, dysprosody, articulation problems with unorthodox speech silence and speech rate alterations have been observed [8]. Dysprosody is characterized by variable speech rate, monopitch and monoloudness [9] leading to flat speech melody, which reduce perceived naturalness of the speech and also the ability to express the acoustic contrast and emotions of the speaker.

At present, objective and reliable assessment of PD is still not available. To diagnose, rate, and monitor the disease, neurologists and clinical psychologists have developed numerous subjective rating scales such as Unified Parkinson's Disease Rating Scale (UPDRS III: motor examination; UPDRS IV: Modified Hoehn and Yahr Staging), Non-Motor Symptoms Scale (NMSS), Beck Depression Inventory (BDI), Freezing Of Gait questionnaire (FOG-Q), The REM sleep Behaviour Disorder Screening Questionnaire (RBDSQ), Mini-Mental State Examination (MMSE) or Addenbrooke's Cognitive Examination-Revised (ACE-R) [6]. Nevertheless, this particular approach has several disadvantages. The most critical one is the natural subjectivity of the examination. Therefore, researchers have investigated non-invasive paraclinical methods of PD assessment that would provide clinicians objective insight into motor and non-motor features of the examined patient [6, 10]. These methods could also be used to monitor the disease and efficiency of the treatment. One of the most advances methods that has been developed in this field of science is the acoustic analysis of dysarthric speech.

In this work, we perform quantitative prosodic analysis of HD in PD to build a regression model capable of PD severity estimation. Degree of PD severity is assessed according to 8 aforementioned subjective clinical rating scales. For this purpose, we investigate the speech recordings acquired from the specialized reading task in which the speakers read 3 sentences (indicative, interrogative, and imperative) using the prosodic speech features that express monopitch, monoloudness and speech rate deficits in HD. The rest of this paper is organized as follows. Section 2 presents the dataset and the methodology. Experimental results are presented in section 3, and section 4 provides discussion and some conclusions.

2 MATERIALS AND METHODS

2.1 DATA ACQUISITION

Altogether, speech recordings of 72 (47 men and 25 women) Czech native speakers with idiopathic PD (age: 67.50 ± 8.08, PD duration: 7.47 ± 4.17 years) were analysed. The patients were enrolled at the First Department of Neurology, St. Anne's University Hospital in Brno, Czech Republic. All patients were examined approximately 1 hour after their regular dopaminergic medication. None of the patients had a disease affecting the central nervous system other than PD or suffered from hallucinations, illusions, or psychosis. All patients signed an informed consent form that had been approved by the Ethics Committee of St. Anne's University Hospital in Brno. At first, patients were examined by clinical neurologist and psychologist according to 8 previously mentioned rating scales.

Afterwards, the speech recordings were acquired and subsequently checked by a trained acoustic engineer who discarded the recordings with an undesirable acoustic noise. The speech protocol consisted of a specialized reading task (reading of indicative, imperative, and interrogative sentences). It was composed of 22 words (87 characters). The task required the stress-control to emphasize the emotions according to the sentences [9]. The speech task: in Czech – *Teď musíš být chvíli trpělivý, než to dokončíme. Už mě to nebaví, dej mi už konečně pokoj! Tak co, jak to dopadlo?*, English translation – *Now, you have to be patient until we finish it. I'm tired of it already, leave me alone! So, how did it go?*

2.2 SPEECH FEATURES

Adopting the methodology from our previous research [9], we extracted a set of prosodic speech features quantifying: a) monopitch – Standard Deviation of fundamental frequency ($F0_S$), Relative Standard Deviation of F0 ($F0_{r.\,S}$), Variation Range of F0 ($F0_R$), Relative Variation Range of F0 ($F0_{r.\,R}$); b) monoloudness – Standard Deviation of Squared Energy Operator/Teager-Kaiser Energy Operator (SEO_S/TEO_S), Relative Standard Deviation of SEO/TEO ($SEO_{r.\,S}$/$TEO_{r.\,S}$), Variation Range of SEO/TEO (SEO_R/TEO_R), Relative Variation Range of SEO/TEO ($SEO_{r.\,R}$/$TEO_{r.\,R}$); c) speech rate abnormalities – Total Speech Time (TST), Net Speech Time (NST), Total Pause Time (TPT), Total Speech Rate (TSR), Net Speech Rate (NSR), Total Pause Time (pauses longer than 50 ms) (TPT 50), Articulation Rate (AR) and SPeech Index of Rhytmicity (SPIR). For further description of these features, see [11, 12, 8].

For the purpose of F0 contour extraction, we used Praat acoustic analysis software [13] [available at: `http://www.praat.org` (Phonetic Sciences, University of Amsterdam, The Netherlands)]. The prosodic features were computed using Neurological Disorder Analysis Tool (NDAT) [5, 8] written in MATLAB and developed at the Brno University of Technology.

2.3 STATISTICAL ANALYSIS

First, the speech features were normalized before the analysis on a per-feature basis to have 0 mean and a standard deviation of 1. To estimate a degree of linear dependence between the prosodic features and PD severity, we computed Pearson's correlation coefficient (r) between feature vectors and associated clinical diagnosis denoted by the values of selected clinical rating scales. To evaluate a prediction power of the prosodic features, univariate regression models were designed. For this purpose, Classification And Regression Trees algorithm (CART) in a supervised learning setup (10-fold validation approach with 5 repetitions) was used. The performance of the models were determined by Root Mean Squared Error (RMSE).

Classification And Regression Trees (CART) is a non-parametric supervised machine learning method based on landmark mathematical theory introduced in 1984 by Breiman, Freidman, Olshen, and Stone [15], which can be used either for classification or regression. One of the advantages of CART algorithm in comparison to other regression-based techniques is its robustness to outliers. The output of this algorithm is a collection of rules inferred from the training data. The rules are used to recursively grow the tree so that data in each subset are the ones with the maximum homogeneity. When the algorithm detects no further gain by partitioning the data, the pruning step is applied to reduce the size and complexity of the decision tree and hence to improve its final prediction power [16].

3 RESULTS

In this paper, we performed quantitative prosodic analysis to estimate degree of PD severity. For this purpose, we adopted the methodology from our previous research of parkinsonian dysprosody [9] and extracted a set of prosodic features to describe monopitch, monoloudness, and speech rate abnormalities in HD. To express linear dependence between the prosodic features and degree of PD severity denoted by 8 previously mentioned subjective clinical rating scales, we computed Pearson's correlation coefficient (r). In addition, we used CART algorithm in a supervised learning setup (10-fold approach validation with 5 repetitions) to built prediction models evaluating a potential of prosodic features to assess HD in PD. Table 1 summarizes the results of the analysis. The table shows top three prosodic features sorted according to the lowest RMSE achieved during evaluation of the associated regression models.

Table 1: Statistical analysis of top 3 prosodic features computed for the reading task.

scale	features	r	p	RMSE	scale	features	r	p	RMSE
UPDRS III	$F0_S$	-0.3738	0.0105	10.7203	UPDRS IV	$TEO_{r.\,R}$	0.2417	0.1056	2.1604
	SEO_S	0.3566	0.0150	10.7983		$TEO_{r.\,S}$	0.2298	0.1244	2.1668
	AR	0.3300	0.0251	10.9105		TEO_R	0.1856	0.2169	2.1878
FOG-Q	TSR	0.3081	0.0373	4.7629	NMSS	$F0_S$	-0.2775	0.0619	17.8900
	TST	-0.3053	0.0391	4.7673		AR	0.2582	0.0831	17.9897
	NST	-0.2752	0.0641	4.8130		TST	-0.2139	0.1535	18.1904
RBDSQ	TEO_S	0.4182	0.0038	2.1250	ACE-R	TST	-0.1772	0.2387	6.4280
	TEO_R	0.3461	0.0185	2.1949		TEO_S	0.1767	0.2402	6.4286
	$F0_{r.\,S}$	0.1715	0.2544	2.3047		TPT 50	-0.1729	0.2505	6.4330
MMSE	$F0_R$	0.3234	0.0284	1.4061	BDI	SEO_S	0.2260	0.1310	4.8236
	TPT 50	-0.1842	0.2204	1.4605		$TEO_{r.\,S}$	0.2240	0.1345	4.8259
	TPT	-0.1621	0.2818	1.4663		SPIR	-0.1952	0.1935	4.8565

Table notation: r – Pearson's product-moment correlation coefficient; p – significance level of correlation; RMSE – root mean squared error; R – variation range; S – standard deviation; r. R/r. S – relative values of range and standard deviation (normalised by the mean value); UPDRS III – Unified Parkinson's Disease Rating Scale (part III), UPDRS IV – Unified Parkinson's Disease Rating Scale (part IV), FOG-Q – Freezing Of Gait Questionnaire, NMSS – Non-Motor Symptoms Scale, RBDSQ – REM Sleep Behavior Disorder Screening Questionnaire, ACE-R – Addenbrooke's Cognitive Examination-Revised, MMSE – Mini-Mental State Examination, BDI – Beck Depression Inventory.

As can be seen, following performance in terms of RMSE has been observed: UPDRS III = 10.7203 for $F0_S$, UPDRS IV = 2.1604 for $TEO_{r.R}$, FOG-Q = 4.7629 for TSR, NMSS = 17.8900 for $F0_S$, RBDSQ = 2.1250 for TEO_S, ACE-R = 6.4280 for TST, MMSE = 1.4061 for $F0_R$, and BDI = 4.8236 for SEO_S. Maximum values of the rating scales: UPDRS III = 108, UPDRS IV = 23, FOG-Q = 20, NMSS = 360, RBDSQ = 13, ACE-R = 30, MMSE = 30, and BDI = 63. These features are conventional in the field of acoustic analysis of dysarthric speech [11, 12, 14]. Comparing the RMSE achieved during the analysis and the maximum values of given rating scales, it can be seen that searching a proper combination of the features could improve the precision of PD severity estimation. Nevertheless, the potential of prosodic features to assess PD is evident.

4 CONCLUSION

The results of this paper show a potential of prosodic features to assess HD in PD. Therefore, this work is a necessary step towards future investigation and development of methodology capable of reliable objective PD severity estimation. Further research in this field of study should be focused on the analysis of relationship between HD and PD (onset, progression, etc.). Quantification of speech prosody can later be used as a method to objectively rate and monitor motor (e.g. freezing of gait), and non-motor (depression, dementia) symptoms associated with PD.

ACKNOWLEDGEMENT

Research described in this paper was financed by the National Sustainability Program under grant LO1401 and by project COST IC1206. For the research, infrastructure of the SIX Center was used.

REFERENCES

[1] M. C. de Rijk, et al., Prevalence of parkinsonism and Parkinson's disease in europe: the europarkinson collaborative study. european community concerted action on the epidemiology of Parkinson's disease, J Neurol. Neurosurg. Psychiatry 1 (1997) 10–5.

[2] O. Hornykiewicz, Biochemical Aspects of Parkinson's Disease, Neurology 51 (1998) S2–S9.

[3] L. O. Ramig, C. Fox, S. Sapir, Speech treatment for Parkinson's disease, Expert Rev. Neurother. 8 (2008) 297–309.

[4] F. L. Darley, A. E. Aronson, J. R. Brown, Differential Diagnostic Patterns of Dysarthria, J. Speech Hear. Res. 12 (1969) 246–269.

[5] I. Eliasova, J. Mekyska, M. Kostalova, R. Marecek, Z. Smekal, I. Rektorova, Acoustic evaluation of short-term effects of repetitive transcranial magnetic stimulation on motor aspects of speech in Parkinson's disease, J. Neural Transm. 120 (2013) 597–605.

[6] Mekyska, J.; Galaz, Z.; Mzourek, Z.; Smekal, Z.; Rektorova, I.; Eliasova, I.; Kostalova, M.; Mrackova, M.; Berankova, D.; Faundez-Zanuy, M.; Lopez-de-Ipina, K.; Alonso-Hernandez, J.B., "Assessing progress of Parkinson's disease using acoustic analysis of phonation," in Bioinspired Intelligence (IWOBI), 2015 4th International Work Conference on, pp. 111-118, 10-12 June 2015.

[7] N. Elfmarkova, M. Gajdos, M. Mrackova, J. Mekyska, M. Mikl, I. Rektorova, Impact of Parkinson's disease and levodopa on resting state functional connectivity related to speech prosody control, Parkinsonism. Relat. Disord. 22 Suppl 1 (2016) 52–5.

[8] J. Mekyska, E. Janousova, P. Gomez-Vilda, Z. Smekal, I. Rektorova, I. Eliasova, M. Kostalova, M. Mrackova, J. B. Alonso-Hernandez, M. Faundez-Zanuy, K. L. Ipiña, Robust and complex approach of pathological speech signal analysis, Neurocomputing 167 (2015) 94–111.

[9] Z. Galaz, J. Mekyska, Z. Mzourek, Z. Smekal, I. Rektorova, I. Eliasova, M. Kostalova, M. Mrackova, D. Berankova, Prosodic analysis of neutral, stress-modified and rhymed speech in patients with parkinson's disease, Comput. Methods. Programs. Biomed. (2016) (in print).

[10] A. Tsanas, M. A. Little, P. E. McSharry, L. O. Ramig, Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity, J. R. Soc. Interface 8 (2010) 842–855.

[11] S. Skodda, U. Schlegel, Speech rate and rhythm in Parkinson's Disease, Mov Disord 23 (2008) 985–92.

[12] S. Skodda, H. Rinsche, U. Schlegel, Progression of dysprosody in Parkinson's disease over time–a longitudinal study, Mov. Disord. 24 (2009) 716–722.

[13] P. Boersma, D. Weenink, Praat: Doing Phonetics by Computer (Version 5.3.10), 2012.

[14] S. Skodda, W. Gronheit, U. Schlegel, Intonation and speech rate in Parkinson's Disease: General and dynamic aspects and responsiveness to levodopa admission, J Voice 25 (2011) 199–205.

[15] L. Breiman, J. Friedman, R. Olshen, C. Stone, Classification and Regression Trees. Boca Raton, FL: CRC Press, 1984.

[16] T. Hastie, R. Tibshirani, J. Friedman, T. Hastie, J. Friedman, and R. Tibshirani, The Elements of Statistical Learning. New York: Springer, 2009.

QUANTIFICATION OF PROSODIC IMPAIRMENT IN PATIENTS WITH IDIOPATHIC PARKINSON'S DISEASE

Zoltán Galáž
Doctoral Degree Programme (2), FEEC BUT
E-mail: z.galaz@phd.feec.vutbr.cz

Supervised by: Jiří Mekyska
E-mail: mekyska@feec.vutbr.cz

Abstract: This paper deals with quantitative analysis of prosodic impairment in idiopathic Parkinson's disease (PD). Experimental dataset consisted of 97 PD patients and 55 healthy speakers. The prosodic features expressing monopitch, monoloudness and speech rate deficits are extracted from stress-modified reading task. Classification accuracies of 70.71 % for females, 70.03 % for males, and 63.20 % for a mixture of both gender were achieved. According to permutation test (1000 permutations, $\alpha = 0.01$), the models were shown statistically significant. Promising potential of prosodic features to identify HD was confirmed.

Keywords: Parkinson's disease, hypokinetic dysarthria, dysprosody, speech processing

1 INTRODUCTION

Parkinson's disease (PD) is a chronic neurodegenerative disorder [1] affecting approximately 1.5 % of people aged over 65 years [1]. It is associated with a substantial reduction of dopaminergic neurons especially in *substancia nigra pars compacta* [2] resulting into variety of motor and non-motor symptoms. According to the previous studies [3], approximately 60–90 % of PD patients suffer from perceptually distinctive motor speech disorder referred to as hypokinetic dysarthria (HD) [4].

HD affects many aspects of speech, namely the area of phonation, articulation, prosody, speech fluency and faciokinesis [5, 6, 7, 8]. Detrimental impact of HD on human verbal communication and daily social life has been observed [9]. Prosodic impairment (dysprosody) is a common speech flaw in HD. It is mainly characterized by variable speech rate, reduced variations in pitch (monopitch) and intensity (monoloudness) [10] leading to unnatural and unintelligible speech. Dysprosody has been observed even in the early stages of PD [11].

At present, PD can not be definitely cured. Therefore, the current medicine is focused on treating its cardinal motor symptoms. However, effectiveness of the treatment depends upon the stage of the disease during which it is initiated [6]. Therefore, the accurate and early diagnosis is necessary for the clinicians to efficiently treat the patients. Nevertheless, the early diagnosis of PD requires a complex understanding of its manifestations on human body. Acoustic analysis of dysprosody is non-invasive, paraclinical method that has a great potential to provide clinicians an objective assessment of motor symptoms of PD.

The aim of this work is to evaluate methods of quantitative analysis of dysprosody that can identify HD and indirectly presence of PD. For this purpose, a variety of prosodic speech features are calculated. These features assess monopitch, monoloudness and speech rate deficits in HD. The features are extracted from the stress-modified reading task specially designed to capture dysprosody. Consequently, the features are used to train the classification model capable of HD identification. The rest of this paper is organized as follows. Section 2 presents the dataset and the methodology. Experimental results are presented in section 3, and section 4 provides discussion and some conclusions.

2 MATERIALS AND METHODS

2.1 DATA ACQUISITION

In this work, we investigated the speech recording acquired from 152 Czech native speakers. The speakers comprised 97 PD patients (53 men/44 women; mean age 67.52 $\pm$ 8.29 years; mean disease duration 7.80 $\pm$ 4.42 years) and 55 healthy speakers (29 men/22 women; mean age 63.96 $\pm$ 9.21 years). The participants were enrolled at the First Department of Neurology, St. Anne's University Hospital in Brno, Czech Republic. The healthy controls (HC) had no history or presence of speech disorders, brain diseases, including neurological and psychiatric illnesses. All patients were examined on their regular dopaminergic medication approximately 1 hour after the L-dopa dose. All patients signed an informed consent form that had been approved by the Ethics Committee of St. Anne's University Hospital in Brno.

The speech task was composed of 3 sentences (indicative, imperative, and interrogative). It comprised 22 words (87 characters). Regarding speech prosody, the task requires the stress-control in order to emphasize the emotions according to the sentences. The participants were asker to read the following sentences: in Czech – *Teď musíš být chvíli trpělivý, než to dokončíme. Už mě to nebaví, dej mi už konečně pokoj! Tak co, jak to dopadlo?*, English translation – *Now, you have to be patient until we finish it. I'm tired of it already, leave me alone! So, how did it go?* The speech signals were sampled with sampling frequency $f_s = 48\,\text{kHz}$ and subsequently downsampled to 16 kHz.

2.2 SPEECH FEATURES

Following our recent research of parkinsonian dysprosody [10], we extracted prosodic speech features that quantify: a) monopitch; b) monoloudness; c) speech rate abnormalities. Completely, we computed Standard Deviation of fundamental frequency ($F0_S$), Relative Standard Deviation of F0 ($F0_{r.\,S}$), Variation Range of F0 ($F0_R$), Relative Variation Range of F0 ($F0_{r.\,R}$), Standard Deviation of Squared Energy Operator/Teager-Kaiser Energy Operator (SEO_S/TEO_S), Relative Standard Deviation of SEO/TEO ($SEO_{r.\,S}$/$TEO_{r.\,S}$), Variation Range of SEO/TEO (SEO_R/TEO_R), Relative Variation Range of SEO/TEO ($SEO_{r.\,R}$/$TEO_{r.\,R}$), Total Speech Time (TST), Net Speech Time (NST), Total Pause Time (TPT), Total Speech Rate (TSR), Net Speech Rate (NSR), Total Pause Time (pauses longer than 50 ms) (TPT 50), Articulation Rate (AR) and SPeech Index of Rhytmicity (SPIR). These features are conventional in the field of dysarthric speech analysis [13].

To extract F0 contour, Praat acoustic analysis software [14] was used. To extract the rest of speech features, Neurological Disorder Analysis Tool (NDAT) [6, 12] written in MATLAB and developed at the Brno University of Technology was used.

2.3 STATISTICAL ANALYSIS

The features were normalized before the analysis on a per-feature basis to have 0 mean and a standard deviation of 1. Random Forests (RF) [15] classifier was used to investigate the power of prosodic features to discriminate healthy and dysarthric speech. Forward selection approach (modified version of Sequential Floating Forward Selection [16] algorithm) was employed to find a non-redundant combination of the features with the maximum clinical information about deterioration of speech prosody in HD. Matthew's correlation coefficient [17] (MCC $= TP \times TN + FP \times FN / \sqrt{N}$) was computed as a criterion of the feature selection. $N = (TP+FP)(TP+FN)(TN+FP)(TN+FN)$, TP (true positive) and FP (false positive) represents the number of correctly identified PD subjects and a number of subjects identified as PD, but being healthy. Similarly, TN (true negative) and FN (false negative) represent the total number of correctly identified HC, and PD patients identified as HC. Additionally, classification accuracy (ACC), sensitivity (SEN) and specificity (SPE) were computed.

Table 1: Statistical analysis of the prosodic features computed for the analysed speech task.

Feat.	G	MCC	ACC [%]	SEN [%]	SPE [%]	p	No.
F1	F	0.3064 ± 0.3507	64.3810 ± 15.6312	67.5000 ± 25.8775	60.3333 ± 30.8405	0.0990	2
	M	0.2129 ± 0.3776	61.7460 ± 17.7139	63.8000 ± 21.0281	57.6667 ± 31.2640	0.1280	2
	A	0.1458 ± 0.2642	58.3714 ± 13.1348	61.6000 ± 17.2150	52.7333 ± 19.6071	0.1560	2
F2	F	0.3995 ± 0.2622	69.6667 ± 12.0142	72.0000 ± 21.2132	65.6667 ± 26.6007	0.0240	3
	M	0.2412 ± 0.4092	63.3016 ± 18.2413	64.7333 ± 20.6810	59.6667 ± 33.5182	0.0360	4
	A	0.2004 ± 0.2135	61.9512 ± 9.1573	66.3556 ± 13.5546	53.8667 ± 21.5747	0.1150	2
F3	F	0.1492 ± 0.3861	58.9048 ± 15.3989	63.5000 ± 19.6980	51.6667 ± 35.9910	0.4570	2
	M	0.2430 ± 0.3193	64.4444 ± 14.1973	68.7333 ± 18.6493	55.6667 ± 30.4185	0.0650	1
	A	0.1300 ± 0.2503	58.1595 ± 12.3655	62.0222 ± 17.1806	51.1333 ± 21.8800	0.6840	2
F4	F	0.4229 ± 0.3514	70.7143 ± 16.2431	71.0000 ± 19.1397	70.3333 ± 30.9139	0.0001	1
	M	0.3708 ± 0.3446	70.0317 ± 16.0522	73.5333 ± 19.2891	63.0000 ± 29.4103	0.0120	5
	A	0.2497 ± 0.2586	63.2024 ± 12.4402	65.0667 ± 14.9209	60.0000 ± 21.0711	0.0130	3

Table notation: G – gender ([F] females, [M] males, [A] combination of both genders); F1 – monopitch features; F2 – monoloudness features; F3 – speech rate features; F4 – general prosodic features; MCC – Matthew's correlation coefficient; ACC – classification accuracy; SEN – classification sensitivity; SPE – classification specificity; No. – Number of selected features; p – p-values of classification calculated by permutation test (1000 permutations).

In this work, the classifier validation was conducted using 10-fold validation approach with 5 repetitions. Furthermore, to determine if the classification results were obtained by chance or by the actual relationship between the class labels and the data values, non-parametric statistical test referred to as permutation test was used [18]. To achieve sufficient statistical relevancy, 1000 permutations was performed. Significance level α was chosen equal to 0.01. Classification models with p values smaller than α was considered statistically significant.

3 RESULTS

In this work, we analysed dysprosody in patients with idiopathic PD. For this purpose, a set of prosodic speech features quantifying monopitch, monoloudness, and speech rate disturbances was extracted [10]. Non-relevant features were excluded by forward selection algorithm. For determining the power of the features to distinguish between healthy and speech with prosodic impairment, random forests classifier in a supervised learning setup (10-fold validation/5 repetitions) was used. Additionally, 1000 permutations were performed during the evaluation of the models. The results of classification process are summarized in Table 1.

As can be seen from the table, the best classification performance in terms of ACC was achieved by $\mathrm{F4-ACC} = 70.71\,\%$ using a single feature ($\mathrm{TEO_{r.\,R}}$) extracted from the reading of female participants. This model was also proven to be statistically relevant ($p < \alpha$). Regarding male speakers, almost identical classification performance was achieved $\mathrm{F4-ACC} = 70.00\,\%$, however the model was trained with 5 features ($\mathrm{F0_S}$, TPT, $\mathrm{F0_{r.\,R}}$, $\mathrm{SEO_S}$, $\mathrm{SEO_{r.\,R}}$) and also less statistically significant. Combining both genders in one group resulted into classification model of the following performance $\mathrm{F4-ACC} = 63.20\,\%$ ($\mathrm{TEO_R}$, TST, TSR). Of note is the fact that all prosodic submodels (F1 – F3) did not achieve sufficient statistical significance (see Table 1 $p > \alpha$). On the contrary, the model based on general prosodic features (monopitch, monoloudness and speech rate) was shown to be statistically significant.

The results of this paper show the promising potential of prosodic features for quantification of HD in PD. However, subsequent research is warranted to fully understand HD and its manifestation on human speech prosody. Moreover, despite relatively low classification accuracy of the prosodic submodels, further investigation of monopitch, monoloudness and speech rate deficits can lead to development of more robust speech features and therefore more precise HD identification.

4 CONCLUSION

In this paper, we performed a quantitative analysis of dysprosody in a set of 97 patients with idiopathic PD and 55 HC using the speech task based on the stress-modified reading of indicative, interrogative and imperative sentences. The presence of prosodic impairment in patients with PD was observed, which is in accordance with the previous research [19, 13, 20]. To emphasize the possible gender-related discrepancies in prosodic deterioration in HD, the analysis was conducted for both gender separately and also for a mixture of both. Based on the results of this study, we conclude that dysprosody in PD is likely to be gender-differentiated. This observation was considered reasonable taking into account the anatomic differences (e.g. length of the vocal tract, etc.) between the speech production system of men and women.

We showed that acoustic analysis of dysprosody can discriminate healthy and dysarthric speech with the classification accuracy over 70 %. Important thing to note is the simplicity and clinical interpretability of selected speech features. In our opinion, these facts compensate the lack of classification performance compared to more sophisticated methods of vocal pathology identification, such as the non-linear features used to quantify dysphonia [21]. The future investigation in this field of science should lead to a fusion of the two approaches to comprehensively describe the manifestation of PD on human speech production.

In our future study we will focus on deeper statistical analysis of dysprosody in patients with PD. We aim to estimate PD severity using the analysis of prosodic features. Furthermore, a longitudinal study of prosodic impairment in HD is planned. Deeper understanding of dysprosody in PD may lead to development of more robust features capable of assessing motor and non-motor aspects of PD and consequently result into precise, early diagnosis, disease tracking and efficiency of treatment monitoring.

ACKNOWLEDGEMENT

Research described in this paper was financed by the National Sustainability Program under grant LO1401 and by project COST IC1206. For the research, infrastructure of the SIX Center was used.

REFERENCES

[1] M. C. de Rijk, et al., Prevalence of parkinsonism and Parkinson's disease in europe: the europarkinson collaborative study. european community concerted action on the epidemiology of Parkinson's disease, J Neurol. Neurosurg. Psychiatry 1 (1997) 10–5.

[2] O. Hornykiewicz, Biochemical Aspects of Parkinson's Disease, Neurology 51 (1998) S2–S9.

[3] L. O. Ramig, C. Fox, S. Sapir, Speech treatment for Parkinson's disease, Expert Rev. Neurother. 8 (2008) 297–309.

[4] F. L. Darley, A. E. Aronson, J. R. Brown, Differential Diagnostic Patterns of Dysarthria, J. Speech Hear. Res. 12 (1969) 246–269.

[5] J. Mekyska, I. Rektorova, Z. Smekal, Selection of Optimal Parameters for Automatic Analysis of Speech Disorders in Parkinson's Disease, in: Telecommunications and Signal Processing (TSP), 2011 34th International Conference on, pp. 408–412.

[6] I. Eliasova, J. Mekyska, M. Kostalova, R. Marecek, Z. Smekal, I. Rektorova, Acoustic evaluation of short-term effects of repetitive transcranial magnetic stimulation on motor aspects of speech in Parkinson's disease, J. Neural Transm. 120 (2013) 597–605.

[7] Mekyska, J.; Galaz, Z.; Mzourek, Z.; Smekal, Z.; Rektorova, I.; Eliasova, I.; Kostalova, M.; Mrackova, M.; Berankova, D.; Faundez-Zanuy, M.; Lopez-de-Ipina, K.; Alonso-Hernandez, J.B., "Assessing progress of Parkinson's disease using acoustic analysis of phonation," in Bioinspired Intelligence (IWOBI), 2015 4th International Work Conference on , vol., no., pp. 111-118, 10-12 June 2015.

[8] N. Elfmarkova, M. Gajdos, M. Mrackova, J. Mekyska, M. Mikl, I. Rektorova, Impact of Parkinson's disease and levodopa on resting state functional connectivity related to speech prosody control, Parkinsonism. Relat. Disord. 22 Suppl 1 (2016) 52–5.

[9] M. D. Pell, H. S. Cheang, C. L. Leonard, The impact of parkinson's disease on vocal-prosodic communication from the perspective of listeners, Brain Lang. 97 (2006) 123–134.

[10] Z. Galaz, J. Mekyska, Z. Mzourek, Z. Smekal, I. Rektorova, I. Eliasova, M. Kostalova, M. Mrackova, D. Berankova, Prosodic analysis of neutral, stress-modified and rhymed speech in patients with parkinson's disease, Comput. Methods. Programs. Biomed. (2016) (in print).

[11] J. Rusz, R. Cmejla, H. Ruzickova, E. Ruzicka, Quantitative Acoustic Measurements for Characterization of Speech and Voice Disorders in Early Untreated Parkinson's Disease, The Journal of the Acoustical Society of America 129 (2011) 350–367.

[12] J. Mekyska, E. Janousova, P. Gomez-Vilda, Z. Smekal, I. Rektorova, I. Eliasova, M. Kostalova, M. Mrackova, J. B. Alonso-Hernandez, M. Faundez-Zanuy, K. L. Ipiña, Robust and complex approach of pathological speech signal analysis, Neurocomputing 167 (2015) 94–111.

[13] S. Skodda, H. Rinsche, U. Schlegel, Progression of dysprosody in Parkinson's disease over time–a longitudinal study, Mov. Disord. 24 (2009) 716–722.

[14] P. Boersma, D. Weenink, Praat: Doing Phonetics by Computer (Version 5.3.10), 2012.

[15] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32.

[16] J. Pohjalainen, O. Rasanen, S. Kadioglu, Feature selection methods and their combinations in high-dimensional classification of speaker likability, intelligibility and personality traits, Computer Speech and Language (2014).

[17] B. W. Matthews, Comparison of the predicted and observed secondary structure of t4 phage lysozyme, Biochim Biophys Acta (BBA) 405 (1975) 442–51.

[18] B. Phipson, G. K. Smyth, Permutation p-values should never be zero: Calculating exact p-values when permutations are randomly drawn, Statistical Applications in Genetics and Molecular Biology 9 (2010).

[19] S. Skodda, U. Schlegel, Speech rate and rhythm in Parkinson's Disease, Mov Disord 23 (2008) 985–92.

[20] S. Skodda, W. Gronheit, U. Schlegel, Intonation and speech rate in Parkinson's Disease: General and dynamic aspects and responsiveness to levodopa admission, J Voice 25 (2011) 199–205.

[21] A. Tsanas, M. A. Little, P. E. McSharry, L. O. Ramig, Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quanti?cation of average Parkinson's disease symptom severity, J. R. Soc. Interface 8 (2010) 842–855.

RECOMMENDATIONS FOR ECG ACQUISITION USING BITALINO

Andrea Němcová*, Lucie Maršánová and Radovan Smíšek

Doctoral Degree Programme (1), *Doctoral Degree Programme (2), FEEC BUT

E-mail: xnemco01@stud.feec.vutbr.cz, xmarsa08@stud.feec.vutbr.cz, xsmise00@stud.feec.vutbr.cz

Supervised by: Martin Vítek and Jana Kolářová

E-mail: vitek@feec.vutbr.cz, kolarova@feec.vutbr.cz

Abstract: Cardiovascular disorders are still the most common cause of death in Western countries. In addition to cardiologists´ care, patient self-examination is the growing area nowadays. For this purpose, mobile devices for ECG signal acquisition are suitable. From these devices, we selected low-cost system called BITalino. In this work, we tested the quality of ECG signals acquired with BITalino under various conditions. We tested 2 values of sampling frequency, 12 electrodes placements and 6 types of exercises. We recommend using sampling frequency of 1,000 Hz, two electrodes placements – on chest and on wrists for both resting and exercise ECG.

Keywords: BITalino, ECG, electrocardiogram, acquisition, sampling frequency

1. INTRODUCTION

The areas of telemedicine and especially patient self-examination using mobile applications are developing very fast nowadays. For this purpose, mobile and ideally wireless devices are necessary for acquiring physiological data. The data are wirelessly sent to access point (smartphone, tablet, PC), special application process them and the user is informed about their health condition. Another possibility is that the data can be sent from access point to database via the Internet and doctors evaluate them. In our work, we deal with the first part of this process, which is common for both approaches – data acquisition. We focus on electrocardiogram (ECG) signal only.

There are several commercial systems for ECG signal acquisition (e.g. [1], [2] and [3]). For our purpose (developing mobile application for wide spectrum of users), we need accessible, low-cost, open (we need the signal in a standard file format, such as *.txt*) and accurate system. The acquisition accuracy is important because in our future work we will evaluate pulse, heart rate variability and the most common heart pathologies (premature atrial or ventricular contractions, atrial fibrillations, sinoatrial blocks, atrioventricular blocks and left and right bundle branch blocks). BITalino acquisition system fulfills these conditions that is why we use it for ECG signal acquisition. Moreover, the system is small, user-friendly, and on the Internet there are available APIs (Application Programming Interfaces) for different platforms for free. [3]

In this work, we tested the quality of ECG signals acquired with BITalino under various conditions, namely sampling frequency, electrodes placements and quality of recorded signals. BITalino is described in chapter 2.1. The conditions are mentioned in chapters 2.2, 2.3 and 2.4. Chapter 3 deals with results. In the end, we recommend proper conditions for a good quality ECG acquisition.

2. METHODS

2.1. BITALINO ACQUISITION SYSTEM

BITalino is a multimodal system for biomedical data acquisition. It consists of microcontroller unit, Bluetooth (BT) module for wireless communication and power block with a Lithium Ion Polymer

battery (Figure 1). In one board, 6 various measurement sensors for bioelectrical and biomechanical data acquisition are integrated (Figure 1). BITalino enables acquisition of electrocardiogram, electromyogram (EMG), electrodermal activity (EDA) and accelerometry (ACC). Light sensor (LUX) and Light-Emitting Diode (LED) are included as well. Light sensor enables ambient light monitoring. Both Light sensor and LED enable synchronization with third party equipment. [4], [5], [6]

ECG signal acquisition with BITalino is noninvasive; ECG is sensed by three Ag-AgCl electrodes. BITalino enables one lead measurement with sampling frequency (f_s) of 10/100/1,000 Hz. The bandwidth is 0.5-40 Hz, in this range the accuracy is guaranteed. The system transfers frequencies from 40 Hz to the half of the sampling frequency as well, but with smaller magnitude. The resolution of ECG signal is 10 bits. The acquired voltage range is limited to (-1.5, +1.5) mV. Data can be visualized by the software OpenSignals, which enables real-time data acquisition and offline browsing. Recorded data can be stored in a standard ASCII file format (*.txt*) or in the HDF5 format. [4], [5], [6]

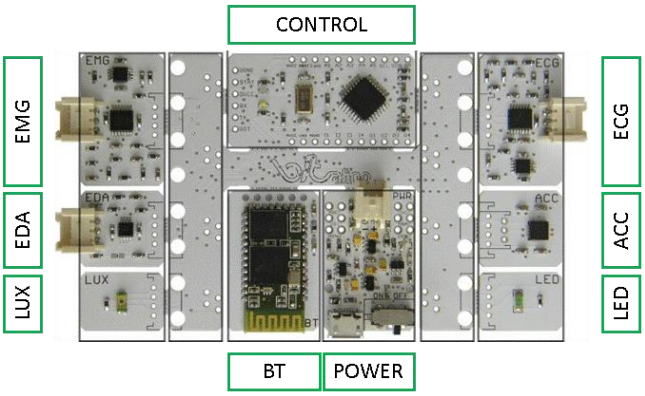


Figure 1: BITalino board consists of microcontroller unit (CONTROL), Bluetooth module (BT), power block (POWER) and 6 sensors for various data acquisition. [5]

2.2. SAMPLING FREQUENCY

BITalino generally enables to set 3 values of sampling frequency (10 Hz, 100 Hz and 1,000 Hz). For ECG acquisition, only 100 Hz and 1,000 Hz are considered. We tested both of these frequencies in terms of preserving the diagnostic information.

2.3. ELECTRODES PLACEMENTS

We tested 12 electrodes placements in terms of signal quality and patient comfort. The exact location of each electrode is given in Table 1. The best electrodes placement(s) will be determined as a compromise between signal quality (amount of diagnostic information) and patient comfort. Hands are more accessible and comfortable body parts for electrodes placement than chest. On the other hand, we expected that the signal from hands has lower quality and magnitude than the signal sensed from chest.

The electrodes placements were determined with respect to standard ECG leads, body part accessibility, assumed amplitudes and design of BITalino. The lead set is 30 cm long, which means that the distance between two electrodes (body parts) should be no more than 60 cm. From standard ECG leads, we take into consideration only limb leads (because they are bipolar). Moreover, the distance between hand and leg is greater than 60 cm, so the sensing is very uncomfortable. The only standard lead, which can be used is limb lead I (left hand, right hand).

The signals were acquired for 1 minute with f_s = 1,000 Hz. The patient was sitting on chair with hands laid on thighs.

All the tests in this paper were performed on three healthy subjects.

No.	Plus	Minus	Reference
1	center of right palm	center of left palm	upper part of right palm
2	under right clavicula	under left musculus pectoralis major	under left clavicula
3	right wrist	left wrist	right wrist closer to pinky
4	right wrist	left wrist	right wrist closer to thumb
5	right wrist	left wrist	left wrist closer to pinky
6	right wrist	left wrist	left wrist closer to thumb
7	sternum	under left pectoral muscles	under sternum
8	sternum	serratus anterior	under sternum
9	sternum	latissimus dorzi	under sternum
10	sternum	under left pectoral muscles	serratus anterior
11	sternum	latissimus dorzi	serratus anterior
12	right thumb	left thumb	right wrist closer to pinky

Table 1: 12 tested electrodes placements.

2.4. EXERCISE ECG

We tested the signal quality while the patient was moving (exercise ECG). We tried a few types of movements, because they can influence the signal in a different way. It depends especially on electrodes placement. Based on chapters 2.2 and 2.3 and results 3.1 and 3.2, we set the sampling frequency on 1,000 Hz and choose two proper electrodes placements (No.2 – chest and No.6 – hands).

The subject was doing each exercise for one minute, than they had 10 seconds to get ready for the next exercise. The acquisition was continual for one electrodes placement and all the exercises (total 6 minutes and 50 seconds). BITalino was hanged on the neck.

To quantify quality of the signal, we used segmentation algorithm [7] based on SNR estimation. The algorithm segmented signal into 3 quality groups. The 1st group – "low noise" (1) has very low level of noise, which means that the signals can be delineated. In the 2nd group – "medium noise" (2), level of noise is higher but still enables to use the signal for QRS detection. The 3rd group – "high noise" (3) includes signals with small SNR; they are not appropriate for further processing.

3. RESULTS

3.1. TESTING OF SAMPLING FREQUENCY

The acquired signals were compared visually by two experts in terms of details. In case of f_s = 1,000 Hz, more details are preserved, but the signal is 10× longer. In case of f_s = 100 Hz the signal is smoother, but some details are lost (rapid changes). These details are important for diagnosing some pathologies (e.g. bundle branch blocks). The sampling frequency should be set with respect to required diagnostic information. F_s = 100 Hz is enough for rhythm monitoring. For morphology assessment, f_s = 1,000 Hz is necessary.

3.2. TESTING OF ELECTRODES PLACEMENTS

After signals acquisition, we evaluated them manually (results are shown in Table 2) in terms of:

- presence of three important components (P, R, T) – 0 means absence, 1 means presence and the component is quite small, 2 means presence and the component is significant (every component in each cycle always belongs in one group (0, 1, 2) within one signal),
- trimming of R wave – Y (yes)/N (no),
- amplitudes of R and T,
- noise - 0 absence, 1 very low level, 2 low level, 3 high level (different scale than in the segmentation algorithm [7]),
- isoline – shift of isoline, drift (D).

Electrodes Placement	Presence of components									Trimming of R wave			Amplitude [mV]						Noise			Isoline [mV]		
No.	P			R			T						R			T								
	S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3
1	1	0	2	2	2	2	2	2	2	N	N	N	0.7	1.4	0.25	0.2	0.4	0.1	1	1	1	0	-0.05	-0.05
2	2	2	2	2	2	2	2	2	2	Y	N	N	> 1.5	1.5	1	0.4	1	0.3	0	0	1	-0.05	-0.1	-0.1
3	1	0	2	2	2	2	2	2	2	N	N	N	0.7	1.4	0.3	0.3	0.4	0.1	1	1	1	0	-0.05	-0.05
4	1	0	2	2	2	2	2	2	2	N	N	N	0.7	1.4	0.3	0.2	0.4	0.1	1	1	1	-0.05	-0.05	0
5	2	0	2	2	2	2	2	2	2	N	N	N	0.7	1.4	0.3	0.2	0.4	0.1	1	1	1	0	-0.05	0
6	1	1	2	2	2	2	2	2	2	N	N	N	0.6	1.4	0.3	0.2	0.4	0.1	2	1	1	-0.05	-0.05	0+D
7	2	1	2	2	2	2	1	1	2	Y	Y	Y	> 1.5	> 1.5	> 1.5	0.1	0.3	0.2	1	1	0	0	0	0+D
8	2	2	1	2	2	2	1	0	1	Y	Y	Y	> 1.5	> 1.5	> 1.5	0.1	0.1	0.1	1	1	1	0	0	0
9	2	1	1	2	2	2	1	2	1	Y	Y	Y	> 1.5	> 1.5	> 1.5	0.1	0	0.1	1	1	2	0	-0.05	0+D
10	2	2	1	2	2	2	2	2	1	Y	Y	Y	> 1.5	> 1.5	> 1.5	0.1	0.3	0.1	0	1	1	0	0	0+D
11	2	1	1	2	2	2	1	1	0	N	Y	Y	1.4	> 1.5	> 1.5	0	0	0	1	1	2	0	0	0
12	2	0	1	2	2	2	2	2	2	N	N	N	0.7	1.3	0.25	0.25	0.4	0.1	1	2	1	-0.05	0	0

Table 2: Quality of the signals sensed with different electrodes placements.

According to the results in Table 2, it was determined that the best signals were achieved with electrodes placement No.2 and No.6. In case of electrodes placement No.2, positive electrode is placed under the right clavicula, negative electrode under the left musculus pectoralis major and reference electrode under the left clavicula. In case of electrodes placement No.6, positive electrode is placed on the right wrist, negative electrode on the left wrist and reference electrode is placed on the left wrist closer to the thumb. The signals sensed with the two electrodes placements show significant presence of all components (P, Q, R, S, T). The R wave was trimmed in one subject in signal sensed with electrodes placement No.2. The trim was more significant while sensing from other places on chest (electrodes placement No.7-10). In case of electrodes placement No.2, components have higher amplitudes (mean value: 1.3 mV for R wave and 0.7 mV for T wave) than in case of electrodes placement No.6 (mean value: 0.75 mV for R wave, 0.23 mV for T wave). In both cases, the signals have very low or low level of noise. Shift of isoline is not significant.

The amplitude of signal components varies from subject to subject. It depends especially on subject´s constitution, skin impedance and physical condition. BITalino has limited voltage range (-1.5, +1.5) mV, that is why the signal with higher amplitude is trimmed. For further analysis is appropriate to use signal which is not trimmed. These are the reasons, why we determine two electrodes placements. For subjects with higher amplitudes of ECG signal, we recommend electrodes placement No.6. For subjects with smaller amplitudes, we recommend electrodes placement No.2.

3.3. TESTING THE QUALITY OF EXERCISE ECG

Table 3 shows the signal quality during exercise ECG acquisition. Presence of noise was evaluated using the method proposed in [7] and briefly described in chapter 2.4 of this paper. The algorithm classified each whole signal into one group (it means the signal has constant level of noise). The algorithm evaluated all signals sensed from chest as "low noise" (signal quality 1). Signal quality was determined as "low noise" or "medium noise" (1 or 2) in the case of acquiring from wrists. The signal sensed from wrists has worse quality (2) when the subject is squatting and dumbbells lifting. Worse quality is caused by greater movements of electrodes and lead set attached to hands than in case of chest electrodes. From this follows, that sensing from chest gives better results. None of the acquired signals (from any person and exercise) was evaluated as worse quality (3).

Performing a visual verification, we have found that the algorithm [7] fails when the evaluated signal is trimmed. Due to the limited voltage range of BITalino, signal is trimmed when the subject is walking, squatting and standing up and sitting down. According to visual verification of two ex-

perts, signal quality is of 2 when signals are acquired while the subject is walking and 3 while the subject is squatting and standing up and sitting down.

Electrodes placement	Exercise	Signal quality [%]											
		1 (low noise)				2 (medium noise)				3 (high noise)			
		S1	S2	S3	∑	S1	S2	S3	∑	S1	S2	S3	∑
Chest (2)	walking	100	100	100	300	0	0	0	0	0	0	0	0
Chest (2)	squating	100	100	100	300	0	0	0	0	0	0	0	0
Chest (2)	standing up and sitting down	100	100	100	300	0	0	0	0	0	0	0	0
Chest (2)	chest rotation	100	100	100	300	0	0	0	0	0	0	0	0
Chest (2)	dumbbells lifting	100	100	100	300	0	0	0	0	0	0	0	0
Chest (2)	deep breathing	100	100	100	300	0	0	0	0	0	0	0	0
Wrist (6)	walking	100	100	100	300	0	0	0	0	0	0	0	0
Wrist (6)	squating	0	0	0	0	100	100	100	300	0	0	0	0
Wrist (6)	standing up and sitting down	0	0	100	100	100	100	0	200	0	0	0	0
Wrist (6)	chest rotation	0	100	100	200	100	0	0	100	0	0	0	0
Wrist (6)	dumbbells lifting	0	0	0	0	100	100	100	300	0	0	0	0
Wrist (6)	deep breathing	100	0	100	200	0	100	0	100	0	0	0	0

Table 3: Quality of exercise ECG (S1-S3 refer to subjects 1-3).

4. CONCLUSION

In this paper, we recommend proper conditions for ECG signal acquisition using BITalino. The sampling frequency should be set with respect to required diagnostic information. For morphology assessment f_s = 1,000 Hz is necessary. For rhythm monitoring f_s = 100 Hz is sufficient. We recommend two possibilities of electrodes placements: 1) + under right clavicula, - under left musculus pectoralis major, ref. under left clavicula; 2) + right wrist, - left wrist, ref. left wrist closer to thumb. The choice of one of these electrodes placements is individual and depends on subject´s signal magnitude. If the subject has R wave < 1.5 mV, we recommend using chest electrodes placement. If the subject has R wave > 1.5 mV, we recommend using wrist electrodes placement. If these recommendations are not followed, the R waves may be trimmed or the amplitudes may be small (such signal makes further processing difficult or even impossible). Trimming of R wave is the biggest disadvantage identified in the tested acquisition system BITalino. From the exercise ECG follows, that signal is less affected by motion when the signal is sensed from chest. The highest noise was caused by squating and dumbbells lifting in case of sensing from wrists.

REFERENCES

[1] *AliveCor* [online]. c2016. Dostupné z: http://www.alivecor.com/home.

[2] SEEQ Mobile Cardiac Telemetry System. MEDTRONIC. *Cardiac Diagnostics & Monitoring* [online]. ©2016.

[3] PLUX. *BITalino: DiY biosignals* [online]. Lisbon. Dostupné z: http://www.bitalino.com/.

[4] Hugo Silva, Ana Fred and Raúl Martins. Biosignals for Everyone. *IEEE Pervasive Computing*, 13(4):64-71, 2014.

[5] Hugo Silva, André Lourenço, Ana Fred and Raúl Martins. BIT: Biosignal Igniter Toolkit. Computer Methods and Programs in Biomedicine, 115(1):20-32, 2014.

[6] GUERREIRO, José, Raúl MARTINS, Hugo SILVA, André LOURENÇO a Ana FRED. A multimodal Platform for Physiological Computing. *International Conference on Informatics in Control, Automation and Robotics* [online]. 2013.

[7] SMITAL, Lukáš a Martin VÍTEK. *ECG Signal Quality Segmentation (Assessment): Device of a fast classification of ECG signals for real-time and offline waveform analysis* [submitted to IEEE TBME]. 2016.

RETROSPECTIVE-GATING MYOCARDIAL T1 MAPPING IN RATS WITH DOXORUBICIN CARDIOMYOPATHY

Lenka Dvořáková
Doctoral Degree Programme (2), FEEC BUT
E-mail: xdvora0c@stud.feec.vutbr.cz

Supervised by: Radovan Jiřík, Eva Dražanová, Peter Scheer, Jana Hložková, Roman Panovský, Zenon Starčuk, jr.
E-mail: jirik@feec.vutbr.cz

Abstract: In vivo T_1 mapping of rat myocardium can be challenging given to the high heart and respiratory rate and small size of the heart. We propose a new method for in vivo T_1 estimation based on retrospectively gated inversion-recovery fast low-angle shot (IR FLASH) pulse sequence. The method was tested on an animal model of dilated cardiomyopathy. The T1 estimates were mostly in line with the effects of doxorubicin and estradiol described in literature.

Keywords: T_1 mapping, CMR, doxorubicin, dilated cardiomyopathy, postprocessing gating

1. INTRODUCTION

The longitudinal relaxation time T_1 of a tissue indicates how fast protons recover after an excitation radio-frequency pulse. The T_1 value within the myocardium is altered in various disease states due to increased water content or other changes of the local molecular environment. Changes in both native T_1 and T_1 following administration of gadolinium (Gd) based contrast agents are considered important biomarkers [1]. Nevertheless, in vivo imaging of animal models can be challenging due to a rapid heart beat and respiration and due to a small size of the heart. Fairly few methods have been introduced for in vivo T_1 mapping of myocardium in small animal models. They are based on the Look-Locker inversion recovery acquisition [3, 2] and the variable-flip-angle method [4]. The disadvantage of the Look-Locker inversion recovery methods is their dependence on a good quality of the ECG related to their prospective-gating character. The variable-flip-angle methods can use retrospective gating which avoids deterioration due to noisy ECG, however they suffer from long acquisition time and they are sensitive to the imperfections of the excitation radio-frequency pulse[4, 5]. In this contribution, a new method for T_1 quantification is proposed which provides advantages of both retrospective gating and fast acquisition time. It is based on inversion recovery fast low-angle shot (IR FLASH) pulse sequence.

The method was applied and validated in a doxorubicin study. Doxorubicin cardiomyopathy is a serious adverse effect occurring when doxorubicin is used as a chemotherapeutic agent in oncology. It is manifested as diffuse myocardial fibrosis which is known to affect native and post contrast T_1 [1].

2. METHODS

2.1. T_1 MAPPING

The acquisition method is based on the IR FLASH pulse sequence. Nonselective adiabatic inversion pulse is applied immediately after a respiration trigger. It is followed by 1000 FLASH excitation/acquisition cycles (TR=4 ms). Every second TR cycle has a zero-size phase encoding pulse and is used as a navigator. Phase encoding of the remaining TR cycles is fixed to the same value corresponding to the given IR pulse. Hence, the number of IR pulses corresponds to the number of phase

encoding steps (here 128). The repetition time of IR pulses is 5000 ms. The acquisition time of the sequence is approximately 11 minutes. [6]

Following the acquisition, images for various TI values are reconstructed (implemeted in Matlab, Mathworks, Natick, MA, USA). From the mean values of each navigator echo, a heart-motion signal is created for each phase-encoding step. It is then band-pass filtered to discard frequency bands out of the presumed heart rate. This quasi-harmonic heart-motion signal is then used for determination of the phase of the cardiac cycle. Only echoes acquired at local maxima and minima of the heart-motion signal are selected, representing two heart-cycle phases – systole and diastole. The echoes of maxima and minima are then linked together to form k-space data for both local maxima and minima and for each TI interval. An image is then reconstructed with TI being the mean TI of all phase-encoding-step echoes. By visual inspection, diastolic phase images (corresponding to local optima or minima of the heart-motion signal depending on TI and its location on the global T_1 recovery curve) are selected and images degraded by respiratory-motion are discarded. For our setup, this resulted in 10-20 images with different TI values (200 – 4000 ms). These image sets are then used for pixel-by-pixel T_1 estimation according to [6]:

$$S_\theta(TI) = k\rho(1 - 2 \cdot e^{-\frac{TI}{T_1}}), \tag{1}$$

where S_θ(TI) is the acquired signal and $\theta = \{k\rho, T_1\}$. To obtain θ, least-square approximator using Equation 1 is applied (as implemented in the Matlab Optimization Toolbox).

2.2. Study design and MR protocol

The doxorubicin study included eighteen female Wistar healthy rats (250-300 g) divided into two groups. The first group of animals (N=9) received a weekly dose of doxorubicin (2.5 mg/kg/week) intraperitoneally for six weeks. The second group (N=9) received the same dose of doxorubicin and moreover a dose of estradiol (agofolin, 5μg/week per toto) subcutaneously. In both groups the age of the animals was 10 weeks when the administration was started. Animals were scheduled to undergo MRI examinations at baseline and then again 18 weeks after weekly receipt of doxorubicin (and estradiol). Rats were anesthetized with isoflurane and placed on a surface coil with heart centered on the coil. A respiratory sensor and a rectal temperature probe (SA Instruments, Stony Brook, NY, USA) were applied for monitoring of the animal's state. The body temperature was kept at 36.5$\pm$1.5°C throughout the experiment. All experiments were in accordance with national legislature.

All imaging was performed on a 9.4T NMR system (Bruker-Biospec 94/30 USR). FLASH IntraGate was used for scout imaging. Three sets of anatomical images in long axis, short axis and four-chamber view were acquired. Images for T_1 mapping were acquired using the retrospectively gated IR FLASH sequence described above. The slice was positioned in the middle of the left ventricle in the short axis view. The field of view was 4$\times$4 cm – 6$\times$6 cm, depending on the size and position of the heart, the slice thickness was 1.0 mm. A bolus of Gd-based contrast agent Magnevist (Bayer HealthCare Pharmaceuticals Inc., Berlin, Germany) was administrated intravenously via the tail vein (dose a 0.02 mmol/kg of body weight) followed by 0.2 mL of saline per toto. Approximately thirty minutes after the contrast injection images for post contrast T_1 mapping were obtained using the same sequence.

2.3. Statistical analysis

For statistical analysis a region of interest (ROI) was drawn manually for each T_1 map in the left ventricle wall as shown in Figure 1. The results are expressed as mean value $\pm$ standard deviation in Table 1. During the experiment eight rats had died due to the doxorubicin treatment. One rat was excluded from the experiment due to artifacts from an incorrectly placed identification microchip.

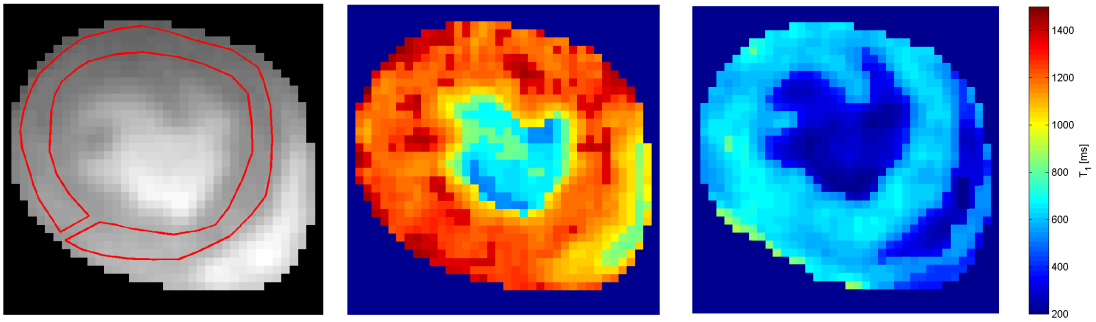


Figure 1: Left: A mean image from all the TI images of rat number 04 examined on the 18th week after the doxorubicin treatment. Center: A native (pre contrast) T_1 map. Right: A post contrast T_1 map.

	baseline		18th week		
Rat No.	native T_1	post contrast T_1	native T_1	post contrast T_1	estradiol
01	1242 ± 60	1125 ± 60	1380 ± 45	749 ±35	1
02	1263 ± 30	986 ± 67	1438 ±141	-	1
03	1157 ± 84	1010 ± 79	1372 ± 56	795 ±130	1
04	1180 ± 39	988 ± 45	1235 ± 72	612 ± 60	1
05	1158 ± 51	1081 ± 53	1216 ±111	1108 ±314	0
06	1200 ± 13	986 ± 52	1460 ± 78	1088 ± 45	0
07	1212 ±113	887 ± 82	1207 ± 82	911 ± 70	0
08	1317 ±158	1036 ± 214	1307 ± 54	1062 ± 58	0
09	1180 ± 39	988 ± 45	1244 ± 72	1125 ±165	0

Table 1: Outcomes of the T_1 estimation for the nine rats at the baseline and at the 18th week after receipt of weekly dose of DOX. In the column "estradiol" the receipt of estradiol is determined, number 1 marks rat with estradiol dosages. All T_1 values are in ms.

For rat number 02 at examination after 18 weeks of treatment, the contrast agent administration was not successful, so only a native T_1 is stated for this case.

The mean values were imported to the Satistica software (Dell software, Tulsa, Oklahoma, USA). The data were not normally distributed thus the non-parametric tests were applied. All analyzes that examined changes before and after the contrast agent bolus and before and after the doxorubicin treatment were evaluated using Wilcoxon matched pairs test. To asses the effects of estradiol administration on native and post contrast T_1, the Mann-Whitney U test was performed.

3. RESULTS

Figure 1 shows the example of T_1 maps estimated by our method. In Table 1 the ROI mean value for each animal is stated. At baseline (before the doxorubicin treatment), the T_1 estimates of all animals were similar and both the native and post contrast T_1 were consistent. The mean value of native and post contrast T_1 at baseline overall both groups was 1212±53 ms and 1009±67 ms respectively. We

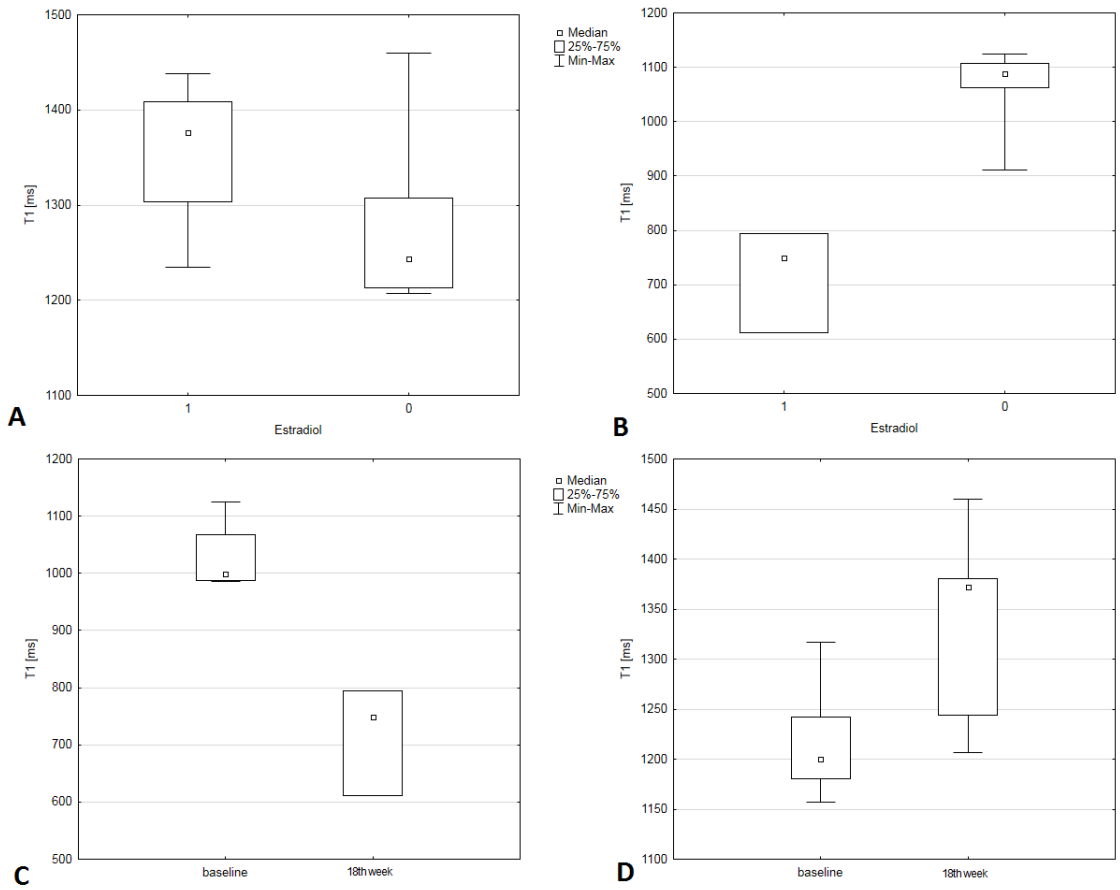


Figure 2: A: Box and whiskers plot by group for the native T_1 at the 18th week, p=0.2. B: Box and whiskers plot by group for the post contrast T_1 at the 18th week, p=0.02. C: Box and whiskers plot for the group with estradiol administration, post contrast T_1 at the baseline and at the 18th week, p=0.054. D: Box and whiskers plot for both groups, native T_1 at the baseline and at the 18th week, p=0.01.

observed a statistically significant decrease in the T_1 value after the contrast agent was administrated (p=0.004). There was no significant change between the group with estradiol administration and without in the native T_1 at the 18th week as seen in Figure 2 A. Given that fact we evaluated the changes in native T_1 for the whole population. Relative to the baseline, there was an increase of the native T_1 at the 18th week (p=0.01) (Figure 2 D). As shown in Figure 2 B, we observed a significant difference between groups with and without estradiol in the post contrast T_1 at the 18th week (p=0.02). In the group with estradiol treatment the post contrast T_1 decreased at the 18th week when collated with the baseline (p=0.054) as seen in Figure 2 C. We observed a significant increase in the post contrast T_1 between the baseline and the measurement at the 18th week (p=0.01).

4. DISCUSSION AND CONCLUSIONS

In this work, a fast IR FLASH method with retrospective gating for in vivo T_1 mapping of rat myocardium was introduced and tested on a rat doxorubicin study. The observed increase in native T_1 due to doxorubicin treatment is in line with literature [7]. This increase can be explained by increased water content and increased concentration of collagen macromolecules in diffuse myocardial fibrosis [1]. Also the observed decrease of post contrast T_1 in the group with estradiol dosage (although not

significant) is in accordance to literature and can be explained by elevated extracellular space in a fibrotic tissue, hence higher content of extravased contrast agent leading to a more pronounced shortening of T_1. The significant increase of the post contrast T_1 in the group without estradiol dosage was not expected. This might be due to unknown biological effects which could be revealed by histology (which was missing in this study). Another reason could be insufficient accuracy and precision of our T_1 estimation method in combination with a small number of the animals in the study. One source of errors was the inconsistency in the time between the contrast agent administration and acquisition for T1-quantification image data (the deviation of about 5 minutes).

Our hypothesis is that the administration of estradiol is affecting the effect of doxorubicin treatment and is increasing the cardiotoxicity. The significant difference at post contrast T_1 between group with estradiol administration and without supports this hypothesis. However, no histopathology analysis was conducted to confirm this theory.

To draw stronger conclusions, more animals will be included in the study. In addition, the pulse sequence will be further modified to acquire a navigator signal following each excitation. This will enable us to double the TI sampling and hence increase the robustness of the estimation algorithm.

ACKNOWLEDGEMENT

The study was supported by Ministry of Education, Youth, and Sports of the Czech Republic (project No. LO1212) and The International Clinical Research Center of St. Anne's University Hospital Brno (project No. 17-2015).

REFERENCES

[1] Moon, J.C., T. A. Treibel, E. B. Schelbert: T_1 Mapping for Diffuse Myocardial Fibrosis *European Radiology* [online] 2013, 23(7), 1871-1881. DOI: 10.1007/s00330-013-2801-6. ISSN 0938-7994.

[2] Vandsburger, M. H., R. L. Janiczek, Y. Xu, et al.: Improved arterial spin labeling after myocardial infarction in mice using cardiac and respiratory gated look-locker imaging with fuzzy C-means clustering. *Mag Res Med*. 2010, 63(3), 648-657. DOI: 10.1002/mrm.22280. ISSN 07403194.

[3] Messroghli, D.R., S. Nordmeyer, M. Buehrer, et al.: Small Animal Look-Locker Inversion Recovery (SALLI) for Simultaneous Generation of Cardiac T1 Maps and Cine and Inversion Recovery–prepared Images at High Heart Rates: Initial Experience. *Radiology* [online]. 2011, 261(1), 258-265. DOI: 10.1148/radiol.11101943. ISSN 0033-8419.

[4] Weingätner, S., M. Akcakaya, S. Roujol, et al.: Free-Breathing Post-Contrast Three-Dimensional T_1 Mapping: Volumetric Assessment of Myocardial T_1 Values. *Mag Res Med* [online]. 2015, 73(1), 214-222. DOI: 10.1002/mrm.25124. ISSN 07403194.

[5] Smit, H., R.P. Guridi, J. Guenoun, et al.: T1 mapping in the rat myocardium at 7 tesla usin a modified CINE inversion recovery sequence. *JMRI* [online]. 2014, 39(4),901-910. DOI: 10.1002/jmri.24251. ISSN 10531807.

[6] Bernstein, Matt A. and Kevin F. King. *Handbook of MRI pulse sequences*. 1st ed. Amsterdam: Academic Press, c2004. ISBN 0-12-092861-2.

[7] Moon, J.C., T. A. Treibel, E. B. Schelbert: T_1 Mapping for Diffuse Myocardial Fibrosis *Journal of the American College of Cardiology* [online]. 2013, 62(14), 1288-1289. DOI: 10.1016/j.jacc.2013.05.077. ISSN 0.735-1097.

SEGMENTATION OF THE COMMON CAROTID ARTERY INTIMA MEDIA USING ACTIVE CONTOUR MODELS

Branislav Hesko
Biomedical electronics and biocybernetics 1, FEEC BUT
E-mail: xhesko00@stud.feec.vutbr.cz

Supervised by: Vratislav Harabiš
E-mail: harabis@feec.vutbr.cz

Abstract: Ultrasound measurements of the human carotid artery intima media thickness are conventionally obtained using manual tracing between tissue layers. This article consists of different approach, when intima media thickness is determined automatically. Tissue layers are detected by use of active contour segmentation. This method does not use any type of gradient, only intensity based image is needed. Therefore, robustness to speckle noise is present. A total of 15 carotid artery ultrasound images is analyzed to determine wall thickness. For evaluation, each of the images contains manually added marks by an expert.

Keywords: Ultrasound imaging, Active contour methods, Segmentation, Carotid artery, Intima media thickness

1 THEORETICAL BACKGROUND

1.1 INTIMA MEDIA THICKNESS

Intima media thickness (IMT) of the common carotid artery can serve as a parameter of early indication of the development of some types of cardiovascular desease. According to the literature, IMT is directly associated with an increased risk of myocardial infarction and strokes, observed in elderly adults, without any previous history of cardiovascular diseases [3]. Moreover, IMT is correlated with atherosclerosis present in other body arteries. As the disease progresses, IMT initially increases along the artery and in later stages develops forms of discrete lesions or plaques. These continue to grow and obstruct blood flow. Physiological IMT value is about 0.75 mm. IMT greater than 1 mm is likely indicative of increased risk of cardiovascular disease [1].

Ultrasound imaging is a non-harmful method, which is capable of real time analysis. It is well established in the assessment of IMT. With image quality getting better in later acquisition systems, the problem of imaging artefacts or speckle noise remains present, causing IMT measurements difficult. Furthermore, in elderly patients, acoustic holes (echo-dropouts) in the adventitia layer are present as another source of difficulties [1], [3].

1.2 ACTIVE CONTOUR METHODS FOR SEGMENTATION

Active contour based segmentation is a set of methods using an initially defined curve, which is evolving in time. The curve may be defined as a set of points, spline interpolating function or parametrically defined curve. The last case may be, in some literature denoted as geometrical contour based segmentation.

The physical principle of these sets of methods is curve shape changing due to the forces acting on this curve, in time. Most commonly, there are two types of acting forces. Internal forces are derived from

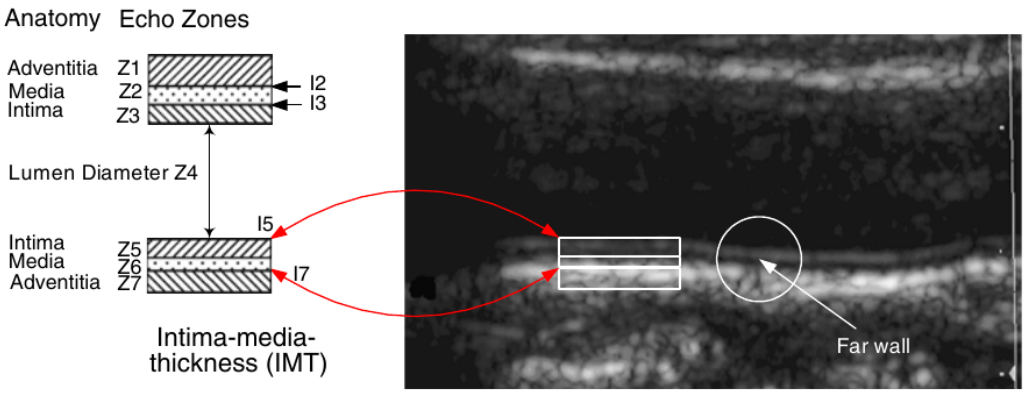


Figure 1: Common carotid artery anatomy echo zones. IMT is defined as distance between blood intima interface and the adventitia interface line. Right part of the image shows labelling real artery echo zones [1].

contour properties, such as tension or strain. On the other side, external forces are image dependent. Most often, gradient is used to calculate external force.

In this paper, active contour method without using gradient is used. Contour is defined in parametric space, in the form of a matrix of the same dimension as the original image with a scalar value assigned to each point. The zero level set then defines the curve. Points with a value greater than zero are representing area inside the segmented object, points with a value lower than zero, the area outside the object. Internal energy for each point of the parametric space is calculated using four-neighborhood of the point. An external force is then calculated using original intensity image. Finally, the image is divided into two distinct areas. Initial shape is always automatically placed circle in the middle of the image with a radius dependent on image size [2].

2 METHOD

2.1 PREPROCESSING

The only manual step is image cropping, because image borders contain useless information, as active contour methods uses all pixels to calculate external force. More concretely, mean intensity of an area outside object defined by the zero level set and inside an object is calculated at each step. Image cropping is the first step, implemented as the initial GUI window.

2.2 SEGMENTATION

At the input of the segmentation, the original image without any preprocessing a part of cropping is used. After this step, every part of the method is automatic with a possibility of modifying method parameters, which are set to their default value based on empirical testing.

The initial curve is defined as a parametric space $\phi_{i,j}$ with equation:

$$\phi^0_{i,j} = (-\sqrt{(i-y_s)^2+(j-x_s)^2+r})/k, \tag{1}$$

where i, j are coordinates in parametric space, which is defined as a matrix of image size. $[x_s, y_s]$ are coordinates of the middle of the circle, r is radius and k is used to lower values in parametric space resulting in better curve growing. The whole method is computed with following an equation:

$$\frac{\phi_{i,j}^{n+1}-\phi_{i,j}^{n}}{\Delta t} = H(\phi_{i,j}^{n})[\frac{\mu}{h^2}\Delta_{-}^{x}\cdot\left(\frac{\Delta_{+}^{x}\phi_{i,j}^{n+1}}{\sqrt{\Delta_{+}^{x}\phi_{i,j}^{n}/(h^2)+(\phi_{i,j+1}^{n}-\phi_{i,j-1}^{n})^2/(2h)^2}}\right)+$$
$$\frac{\mu}{h^2}\Delta_{-}^{y}\cdot\left(\frac{\Delta_{+}^{y}\phi_{i,j}^{n+1}}{\sqrt{\Delta_{+}^{y}\phi_{i,j}^{n}/(h^2)+(\phi_{i+1,j}^{n}-\phi_{i-1,j}^{n})^2/(2h)^2}}\right)- \tag{2}$$
$$\nu-\lambda_1(u_0(i,j)-c_1(\phi^n))^2+\lambda_2(u_0(i,j)-c_1(\phi^n))^2],$$

where μ, ν, λ_1, λ_2, h and Δt are method constants. $u_0(i,j)$ are pixels from the original intensity image, located at the coordinates i,j. In this equation, multiple variables need to be explained. $\Delta_{+}^{x},\Delta_{+}^{y},\Delta_{-}^{y}$ and Δ_{-}^{x} are difference equations calculated from neighbourhood points in parametric space with equations:

$$\Delta_{-}^{x}\phi_{i,j} = \phi_{i,j}-\phi_{i-1,j} \tag{3a}$$
$$\Delta_{+}^{x}\phi_{i,j} = \phi_{i+1,j}-\phi_{i,j} \tag{3b}$$
$$\Delta_{-}^{y}\phi_{i,j} = \phi_{i,j}-\phi_{i,j-1} \tag{3c}$$
$$\Delta_{+}^{y}\phi_{i,j} = \phi_{i,j+1}-\phi_{i,j} \tag{3d}$$

c_1, respectively c_2 represents the mean value inside and outside segmented object, where $\phi_{i,j}>0$ inside the object and $\phi_{i,j}<0$ outside the object. Their mathematical representation is:

$$c_1(\phi) = \frac{\sum_{\phi\geq 0}u_0(i,j)}{\sum_{\phi\geq 0}1} \tag{4}$$

$$c_2(\phi) = \frac{\sum_{\phi<0}u_0(i,j)}{\sum_{\phi<0}1} \tag{5}$$

Last variable is an alternatively expressed Heaviside function with the same behaviour in limits, but with continuous character in interval $[0,1]$. Its equation is:

$$H_k(x) = \frac{1}{2}(1+\frac{2}{\pi}\arctan(\frac{x}{k})) \tag{6}$$

Looking at the equation (2), it is obvious, that parametric value at the next step $\phi_{i,j}^{n+1}$ is located on both sides of the equation. Thus, this fact does not allow direct calculation of the new parametric value. Therefore, iterative calculation is used and $\phi_{i,j}^{n+1}$ is at the first step on the right side of the equation substituted by its best estimate, which is previous value of the n step, $\phi_{i,j}^{n}$. After k iterations, the result of the method is parametric space ϕ, where values greater than zero represent segmented objects and values lower than zero, background.

2.3 POSTPROCESSING

After finishing the iteration process, multiple steps are needed. First of all, a binary image is created from parametric space according to point belonging to inside of the object or to the background. This is distinguished by the sign of the concerned point value in parametric space.

Many small objects are present in the resultant binary image. To eliminate them, a morphological operation of closing is applied to get rid of them. This consists of dilatation followed by erosion. To achieve small objects suppressing, the erosion structural element is bigger concerning size than the dilatation one. Finally, the last step of postprocessing is contour detection. It consist of passing

through every point in binary object marked as inside of the object and evaluating, whether the point is on the boundary or not. This evaluation is based on number of neighbours that are also marked as inner, e.g. point with 8 neighbours marked as inner is not placed on the boundary and point with 3 inner neighbours is evaluated as boundary point. Finally the example of an image with highlighted contour looks like the image below.

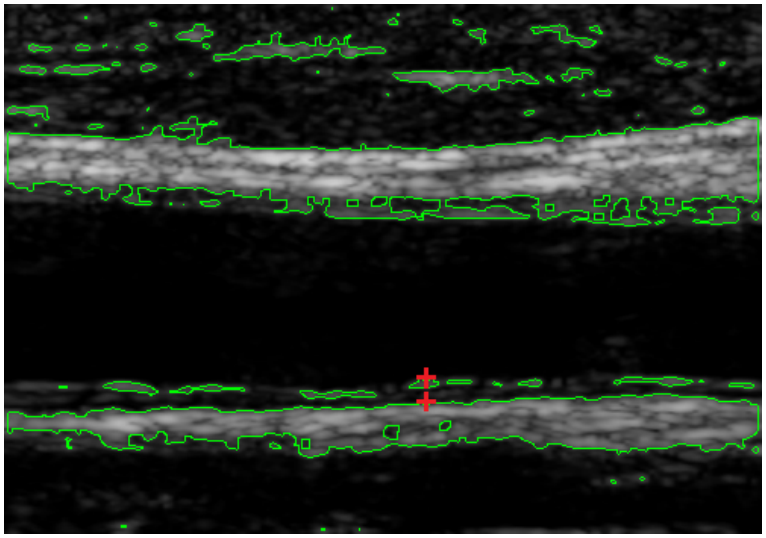


Figure 2: Example of an testing image of the carotid artery with highlighted contour.

2.4 IMT CALCULATION

The last step is to evaluate IMT, which is evaluated at the bottom part of the carotid artery. As it could be seen from the figure 2, the segmentation process ends with segmenting adventitia as the dominant object on the image. Some pixels above, small objects representing intima are segmented. Then, IMT is the distance between top of these objects and the first pixel of adventitia contour.

First of all, the original image is cropped the way, that only bottom part of the artery is available for the next step. Supposing that the lumen of the artery has minimal average intensity values, using average intensity projection along x axe, following by an averaging filter, the lumen can be identified easily. This is the origin, where image is cropped. The number of colons remains constant.

It is important to point out, that object boundary is one pixel thin. This property is used to calculate IMT as this distance in pixels is the difference between the first and third point of the contour along y axe. The algorithm systematically searches for the first, second and a third contour pixel column by column. The highest distance between first and third point is finally labelled as IMT. This computation method could fail when we realise that intima is not contoured along all the axe x. Looking for the distance difference, then would lead to meaningless results. But if we suppose that lumen does not contain any segmented objects, we could assume that the contour pixel with the highest position belongs to the intima. Then we could add a condition that while looking for the first pixel, his y coordinates cannot differ more than a few pixels from the mentioned contour pixel with the highest position.

3 EVALUATION OF ACHIEVED RESULTS

Implemented method has been tested on a set of 15 images with expert marks [1]. Resolution or in other words distance between two pixels was $d = 0.06$ mm. Distance between marks is computed,

followed by multiplication by the distance between two pixels resulting in IMT value. IMT calculated by the algorithm is then compared to the expert values in a table below.

Table 1: Difference between expert marked values and algorithm computed values of IMT.

	IMT_{expert}		$IMT_{computed}$		ΔIMT
i	[pixel]	[mm]	[pixel]	[mm]	[mm]
1	18	1.08	20	1.20	-0.12
2	12	0.72	12	0.72	0
3	13	0.78	11	0.66	0.12
4	11	0.66	11	0.66	0
5	12	0.72	7	0.42	0.3
6	10	0.60	10	0.60	0
7	9	0.54	9	0.54	0
8	10	0.60	11	0.66	-0.06
9	10	0.6	14	0.84	-0.24
10	18	1.08	19	1.14	-0.06
11	13	0.78	7	0.42	0.36
12	11	0.66	9	0.54	0.12
13	16	0.96	14	0.84	0.12
14	10	0.6	9	0.54	0.06
15	9	0.54	9	0.54	0

We found that the average difference between computed and labelled values is $\Delta IMT = 0.04 \pm 0.15$ mm. This value shows, that method is sufficiently precise, because labelling by many experts could show similar differences and the uncertainty of the measured value lies in the range of physiological values. Moreover, it is sometimes difficult to distinguish intima or adventitia interface and active contour methods fails, because of the low contrast.

4 CONCLUSION

This paper deals with the active contour method for common carotid artery segmentation and following intima media thickness evaluation. An automatic method has been implemented, the only manual step remains image cropping. Achieved results show good precision comparing to expert generated results. Potential method failure is caused by low contrast images, where active contour segmentation does not recognise individual tissues. Another source of persisting problems is the eventual presence of objects in the lumen.

In the future work, it would be useful to fully automatize the method and to enhance image contrast by other preprocessing. To achieve better results, it is possible to detect adventitia interface, which has shape of a line, by the Hough transform and then to look at the distance from this interface to objects above.

REFERENCES

[1] LOIZOU, C.P. et al.: *Snakes based segmentation of the common carotid artery intima media*. Medical Biological Engineering & Computing. 2007. 14p.

[2] CHAN, T.F. and VESE, L.A.: *Active contours without edges* . 2001. Los Angeles, USA. 12p. ISSN: 1057-7149

[3] LAMONT, D. et al.: *Risk of cardiovascular disease by carotid intima-media thickness at age 49-51: life course study*. 2000. BMJ 320:273-278.

SIGNAL BASED FEATURE SELECTION FOR FAST CLASSIFICATION OF SEQUENCES IN METAGENOMICS

Karel Sedlář

Doctoral Degree Programme (3), FEEC BUT

E-mail: sedlar@feec.vutbr.cz

Supervised by: Ivo Provazník

E-mail: provaznik@feec.vutbr.cz

Abstract: The rapid development in DNA sequencing techniques brings completely new possibilities into metagenomics research. No longer is the whole metagenome sequencing an issue. On the other hand, this progress lies new demands on bioinformatics tools indented to process this kind of data. Unlike the amplicon based sequencing where every sequence represents a particular gene, the whole metagenome sequencing produce sequences that are random pieces of genomes in the metagenome. Therefore, the reference database for identification of these sequences cannot be used. Here, we present fast feature selection based on genomic signal processing for alignment-free classification of sequences in the metagenome.

Keywords: metagenomics, WGS, genomic signal

1. INTRODUCTION

According to estimates, the number of cells belonging to microorganisms located in a human body exceeds the number of human cells in a ratio of 10 to 1. Although most of these organisms do not cause any harm to the body, not inconsiderable part of them is associated with the occurrence of dangerous diseases and antibiotic resistance. Accurate and rapid identification of these organisms is therefore fundamental in the research of diseases for fast deployment of appropriate treatment and therefore in saving of human lives [1]. In the research of farm animals, it can help prevent significant economic losses through preventing diseases of entire herds.

The rapid development of DNA sequencing techniques is dramatically changing the way in which metagenomic research is conducted, i.e. research of the genomes of different organisms found in the common environment. Such studies are nowadays closely associated with the use of advanced bioinformatics techniques for processing biological sequences [2]. Compared to two metagenomic studies published in 2006, thousands of studies are produced in these days. Thus, continuous development of new and faster bioinformatics tools that will be able to handle this amount of data is necessary [3].Older but still used metagenomic approach stands on targeted amplicon sequencing of one selected representative gene [4]. Sequence identification is then performed by comparing the sequences against the reference database and data can be visualized using the dimensionality reduction techniques, usually analysis of principal co-ordinates (PCoA) using UniFrac metric [5,6]. The current trend, however, lies in shotgun sequencing of the entire metagenome. Unfortunately this approach brings several issues. Firstly, using this approach, one is not able to identify the sequences directly, because they represent different parts of the genome, for which there is still no complete reference database. Secondly, the number of sequences prevents sufficient comparison against a reference database even with the heuristic algorithm BLAST [7,8].

Current bioinformatics methods for metagenomic data processing are based on fast classification of sequences into clusters that represent the different types of organisms and whose subsequent identification is no longer an issue. However, there are several different approaches. The simplest of them use locally sensitive hash function and compare the two sequences locally by particular

words, so called *k*-mers [9,10]. Advanced techniques are based on a selection of different feature vectors. In such cases the classification of metagenomic data is normally carried out using clustering techniques [11]. It is assumed that individual clusters will contain the feature vectors derived from the same organism and the overall architecture of the clusters will give us information about taxonomic diversity of studied species. Unfortunately, current techniques for creating such a specific feature vector are based only on the character processing techniques and are therefore very limited. On the other hand, a large selection of techniques to generate a sufficiently specific vector can be provided by a progressively evolving discipline of bioinformatics called genomic signal processing [12]. As it has been already deduced, phase signals are species specific and can be used for example for a sequence alignment [13] or comparison of species even after a massive downsampling [14]. Here, we present reconstruction of species specific feature vector based on combination of slopes of several different phase signal representations of genomic data.

2. MATERIALS AND METHODS

2.1. TEST DATASET

In order to test the proposed method, we created 3 test dataset containing 2,500 sequences each, derived from 5 different organisms representing 4 different bacterial species. The whole genome sequences were obtained from Genbank database at NCBI (http://www.ncbi.nlm.nih.gov/genbank/), see summary in Table 1. Sequencing reads were simulated as random fragments derived from complete genome sequences and 500 reads were generated from every genome. A half of the sequences were further modified as reverse complementary to better represent real sequencing data. Moreover, 3 dataset according to length of reads (500 nt, 1,000 nt and 5,000 nt) were prepared to examine the influence of read length on the succession of classification and to represent data from next-generation sequencing (NGS) as well as third-generation (TGS) sequencing platforms.

Table 1: Summary of organisms used for test dataset

species	organism	accession no.
E. coli	*Escherichia coli* UTI89	NC_007946.1
E. coli	*Escherichia coli* str. 'clone D i14'	NC_017652.1
C. C. ruddii	*Candidatus Carsonella ruddii* PV	NC_008512.1
G. obscurus	*Geodermatophilus obscurus* DSM 43160	NC_013757.1
R. prowazekii	*Rickettsia prowazekii* str. Dachau	NC_017051.1

2.2. SIGNAL REPRESENTATION

Although we have already proved the cumulated phase signal representing purine-pyrimidines (R-Y) and strong-week (S-W) ratio to be species specific, this feature was examined only on a large (whole genome) scale [15]. To examine the features on a smaller (reads) scale, we decided to use all 3 possibilities in which nucleotides {A, C, G, T} are assigned phases {$\pi/4$, $-3\pi/4$, $3\pi/4$, $-\pi/4$} to present R-Y and S-W, {$3\pi/4$, $-3\pi/4$, $\pi/4$, $-\pi/4$} to present R-Y and amino-keto (M-K), {$3\pi/4$, $-3\pi/4$, $-\pi/4$, $\pi/4$} to present S-W and M-K. Moreover, we examined both, cumulated phase signals representing cumulative sum of nucleotides along sequences as well as unwrapped phase representing transition and transversions along sequences [16].

For every read, all 6 signals (3 cumulated phase and 3 unwrapped phase signals) were reconstructed and their slopes were computed. A feature vectors representing reads were constructed using those 6 values representing the slopes.

2.3. CLUSTERING AND STATISTICS

Feature vectors were clustered using Ward's hierarchical clustering technique with utilization of Euclidean metric. The resulting clusters were compared to real taxonomy and statistics for each organism in form of sensitivity (*sensitivity=TP/(TP+FN)*), specificity (*specificity=TN/(TN+FP)*),

precision (*precision = TP/(TP+FP)*) and accuracy (*accuracy = (TP+TN)/(TP+FN+FP+FN)*) were computed.

3. RESULTS AND DISCUSSION

Firstly, we reduced the feature vectors by omitting the signals of unwrapped phases representing S-W and M-K. This kind of signals is identical to unwrapped phase representing R-Y and M-K from its definition and therefore it does not bring any additional information.

Secondly, by computing the above mentioned statistics for particular signal representations, the slope of cumulated phase signals representing R-Y and S-W were found to be unable to distinguish different species. Therefore, we omitted also this feature from the vector. The example of cumulated phase signals with (S-W and M-K) and without (R-Y and S-W) discriminative information is shown in Figure 1.

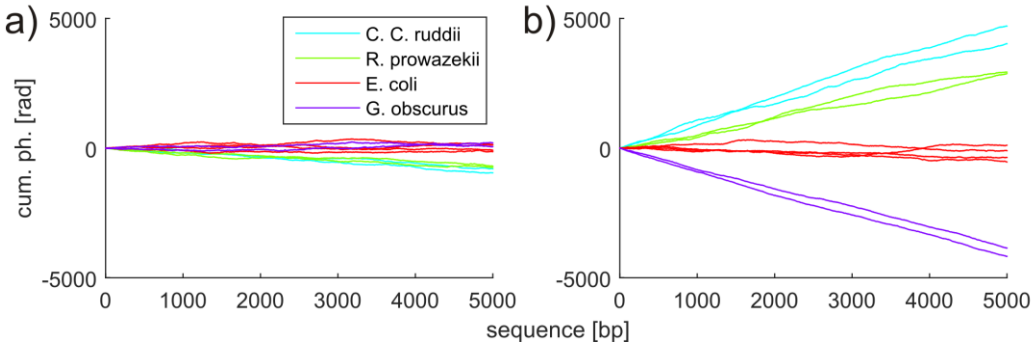


Figure 1: Cumulated phase signals for randomly selected sequences from the test dataset presenting a) S-W and M-K information b) R-Y and S-W information

The resulting vectors containing 4 slopes were further analyzed by hierarchical clustering and the resulting clusters were used for statistics evaluation. The tree reconstructed from feature vectors derived from 5,000 nt long sequences is presented in Figure 2. There are 4 evident clusters representing 4 species.

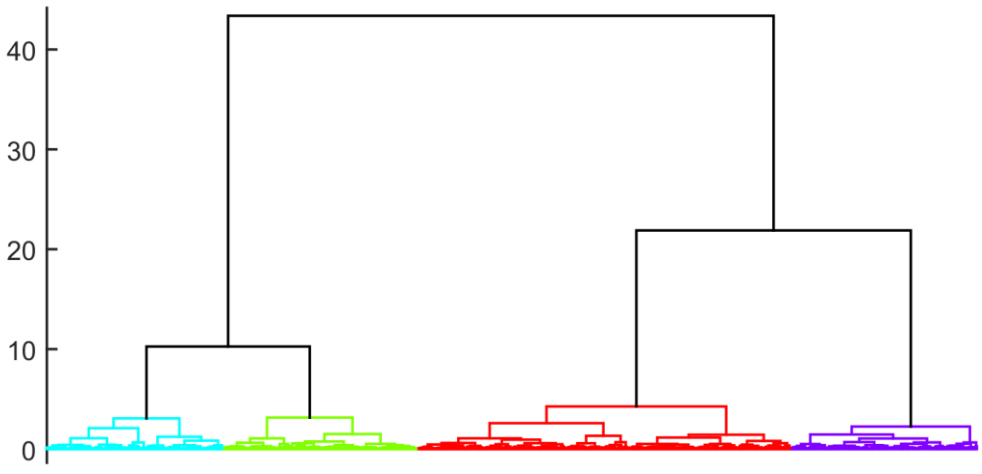


Figure 2: The tree reconstructed from the feature vectors using Ward method and Euclidean metric

The cluster containing *E. coli* vectors is larger because two organisms for this species were included in the test dataset. Only a small number of sequences were misclassified as shown below. However, the longer the sequences are, the higher amount of information they contain. Although 5,000

nt long reads are typical for TGS, we tried to test if the method works for NGS data too. Therefore, we repeated the analysis with shorter reads, typical for NGS. The results of classification are summarized in Table 2 and Table 3.

Table 2: The results of classification (sensitivity and specificity)

Organism	500 nt		1000 nt		5000 nt	
	Sensitivity	**Specificity**	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**
E. coli	98,28	94,64	99,78	93,63	99,90	99,60
C. C.ruddii	75,00	98,14	97,45	97,98	100,00	99,21
G. obscurus	99,80	99,60	99,40	99,95	100,00	100,00
R. prowazekii	74,29	92,08	77,67	99,31	95,78	99,95
Average	**86,84**	**96,12**	**93,58**	**97,72**	**98,92**	**99,69**

Table 3: The results of classification (precision and accuracy)

Organism	500 nt		1000 nt		5000 nt	
	Precision	**Accuracy**	**Precision**	**Accuracy**	**Precision**	**Accuracy**
E. coli	91,60	96,00	89,80	95,84	99,40	99,72
C. C.ruddii	93,00	92,40	91,80	97,88	96,80	99,36
G. obscurus	98,40	99,64	99,80	99,84	100,00	100,00
R. prowazekii	67,60	88,84	97,40	93,88	99,80	99,08
Average	**87,65**	**94,22**	**94,70**	**96,86**	**99,00**	**99,54**

The results show that the method is very successful for TGS data classification, however, it is also applicable on NGS data with satisfactory results. The main advantage of the method over the current techniques is its low computational complexity. While the most of current techniques suffer from quadratic complexity and the fastest of them are no better than $O(n \log n)$ [17], our method is computationally very efficient, based only on basic mathematical operations. In combination with recently published algorithm *TwisterTries* for hierarchical clustering [18], the whole pipeline would be only $O(n)$.

4. CONCLUSION

A new method for classification of metagenomic data without prior alignment to a reference database is introduced in this paper. Unlike the current techniques, our approach relies fully on genomic signal processing making it computationally very efficient. The sequences are first transformed into 4 different phase signal representations and then the feature vector using the slopes of the signals is reconstructed. The only computationally demanding part remains in final vector clustering. However, this part can be updated with recently published algorithm for hierarchical clustering with only linear complexity. Such a method fully meets the current needs when millions of sequences are produced during every sequencing run. We demonstrated our method to work not only on TGS data, but also on currently most widely used NGS.

ACKNOWLEDGEMENT

Supported by BUT intern grant agency, interfaculty junior project FEKT/FIT-J-16-3335 Advanced methods for fast identification and visualization of metagenomic data.

REFERENCES

[1] NIH Human Microbiome Project [online]. 2016 [cit. 2016-03-10]. Dostupné z: http://www.hmpdacc.org/

[2] SIMON, C. a R. DANIEL. Metagenomic Analyses: Past and Future Trends. Applied and Environmental Microbiology. 2011, 77(4), 1153-1161.

[3] REDDY, T. B. K., et al. The Genomes OnLine Database (GOLD) v.5: a metadata management system based on a four level (meta)genome project classification. Nucleic Acids Research. 2015, 43(D1), D1099-D1106.

[4] KLINDWORTH, A., et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Research. 2012, 41(1), e1-e1.

[5] LOZUPONE, C. a R. KNIGHT. UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. Applied and Environmental Microbiology. 2005, 71(12), 8228-8235.

[6] HAMADY, M., C. LOZUPONE a R. KNIGHT. Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. The ISME Journal. 2009, 4(1), 17-27

[7] SCHOLZ, M. B., C. LO a P CHAIN. Next generation sequencing and bioinformatic bottlenecks: the current state of metagenomic data analysis. Current Opinion in Biotechnology. 2012, 23(1), 9-15.

[8] ALTSCHUL, S. F., W. GISH, W. MILLER, E. W. MYERS a D. J. LIPMAN. Basic local alignment search tool. Journal of Molecular Biology. 1990, 215(3), 403-410.

[9] EDGAR, R. C. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010, 26(19), 2460-2461.

[10] RASHEED, Z. a H. RANGWALA. A Map-Reduce Framework for Clustering Metagenomes. In: 2013 IEEE International Symposium on Parallel. IEEE, 2013, s. 549-558.

[11] MANDE, S. S., et al. Classification of metagenomic sequences: methods and challenges. Briefings in Bioinformatics. 2012, 13(6), 669-681

[12] ANASTASSIOU, D. Genomic signal processing. IEEE Signal Processing Magazine. 2001, 18(4), 8-20.

[13] SKUTKOVA, H., M. VITEK, K. SEDLAR a I. PROVAZNIK. Progressive alignment of genomic signals by multiple dynamic time warping. *Journal of Theoretical Biology*. 2015, 385, 20-30.

[14] SEDLAR, K., H. SKUTKOVA, M. VITEK a I. PROVAZNIK. Set of rules for genomic signal downsampling. *Computers in Biology and Medicine*. 2016, 69, 308-314.

[15] SEDLAR, K., H. SKUTKOVA, M. VITEK a I. PROVAZNIK. Prokaryotic DNA Signal Downsampling for Fast Whole Genome Comparison. *Advances in Intelligent and Soft Computing*, 2014, 373, 273-291.

[16] CRISTEA, Paul Dan. Large scale features in DNA genomic signals. *Signal Processing*. 2003, 83(4), 871-888.

[17] LACZNY, C., et al. Alignment-free Visualization of Metagenomic Data by Nonlinear Dimension Reduction. *Scientific Reports*. 2014, **4**(1), e4516.

[18] COCHEZ, M. a H. MOU. Twister Tries. In: Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data - SIGMOD '15. New York, USA: ACM Press, 2015, 505-517.

Doktorské projekty

Komunikační technologie

DEPENDENCY OF THE CONVERGENCE RATE MEAN EXTENT OF VARIATION ON THE REPETITIONS NUMBER IN WEAKLY CONNECTED TOPOLOGIES

Martin Kenyeres

Doctoral Degree Programme (2), FEEC BUT

E-mail: kenyeres@phd.feec.vutbr.cz

Bohumil Novotný

Doctoral Degree Programme (3), FEEC BUT

E-mail: xnovot59@stud.feec.vutbr.cz

Supervised by: Vladislav Škorpil

E-mail: skorpil@feec.vutbr.cz

Abstract: This paper deals with the stochastic distributed algorithm – the push-sum protocol. We examine the effect of experiments repetitions on the mean of the convergence rates quantities. The main goal of the executed experiments is to show how many repetitions of the push-sum protocol are necessary to achieve a statistically credible representative of the obtained set of data. Within this paper, we have focused on weakly connected structures.

Keywords: Distributed computing, The push-sum protocol, The convergence rate mean extent of variation

1. INTRODUCTION

The centralized way of computation, which was frequently used in the past, is substituted by the novel solution – distributed computing [1]. The systems whose functionality is executed in this way are labelled as the distributed systems [1]. The agents forming these systems are characterizable by a limited awareness of the other elements within the same system as well as the whole system.

This paper continues the paper entitled as the dependency of the convergence rate mean extent of variation on the repetitions number in strongly connected topologies, where the convergence rate mean extent of variation was examined in a fully-connected mesh topology executing the push-sum protocol. Within this paper, our attention is focused on a line topology, which is a representative of weakly connected structures. We examine the effect of experiment repetitions number on the mean of the convergence rates quantities. Although the experiment executed more times ensures more statistically credible data, too extensive execution presents a demanding computing process. Therefore, the main motivation of this paper is to show how the number of the repetitions of the same experiment affects the credibility of the obtained results.

The experiment within this paper is executed in a line topology with the size of eight agents. The number of repetitions is 10, 100, 1000, 10 000 and 100 000. The experiment is repeated five times for each number of repetitions. At the end, the obtained results are discussed and compared with the ones obtained in the strongly connected topology.

2. THE PUSH-SUM PROTOCOL

The push-sum protocol is a multi-purpose gossip-based aggregation algorithm, which fulfills its functionality by an iterative pair-wise distribution of the aggregated values among particular agents. It is a stochastic distributed protocol, which means that the outputs are supposed to differ

when the experiment is repeated despite the same inputs. As mentioned, the push-sum protocol is supposed to solve more than one problem. In this paper, we have focused on the calculation of the average determined by all the initial values. Below is described the execution of the protocol [1]:

- The initial inner states of all the nodes represent the values from which the average is calculated. The initial weights are set to 1 for each node.
- Each node sends a half of its current state and a half of its weight to one of its neighbors chosen uniformly at random. The same value is stored in the inner memory and participates in further computation. This is repeated at each iteration.
- The sum of the inner state and the inner states sent by the adjacent nodes is the inner state for the next iteration. The sum of the weight stored in the inner memory and the weights sent by the adjacent nodes is the weight for the next iteration.
- The ratio of these two sums (the sum of the inner states is the numerator) poses the estimation of the average.
- The previously described procedure is repeated until the consensus is reached

3. THE EXAMINED TOPOLOGY

Within this paper, we focus our attention on a line topology. This topology was chosen as a representative of weakly connected structures, in which the protocol is supposed to reach slower convergence rates. The mentioned topology consists of eight agents and is shown in Fig. 2.



Figure 1: The examined line topology

4. EXPERIMENTS

The results obtained from the numerical experiments are shown in this section.The experiments were executed in order to show the behavior of the convergence rate mean extent of variation [2] of a data set varies when the number of repetitions changes. The number of the repetitions were changed and the experiment was repeated five times for each number.

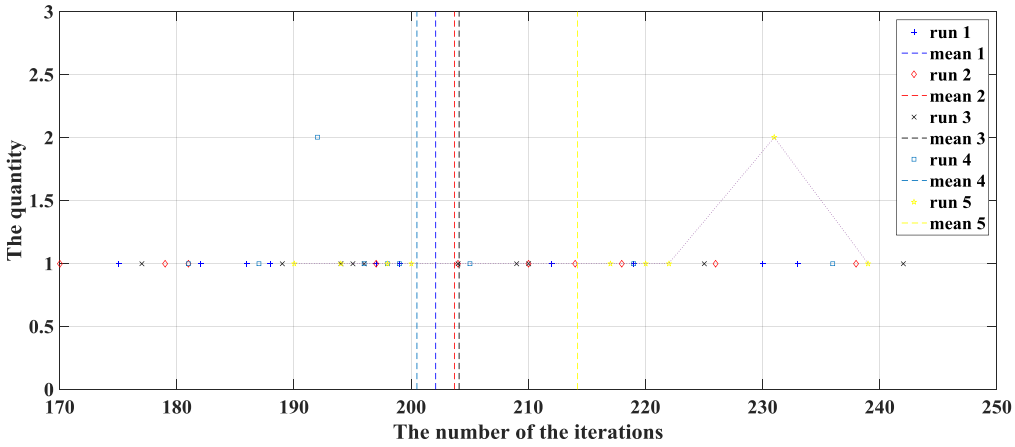


Figure 2: The results for 10 repetitions

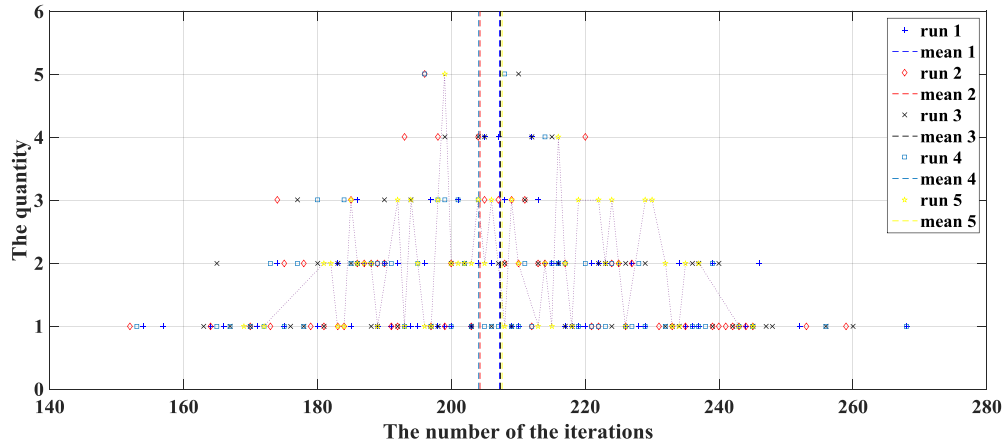


Figure 3: The results for 100 repetitions

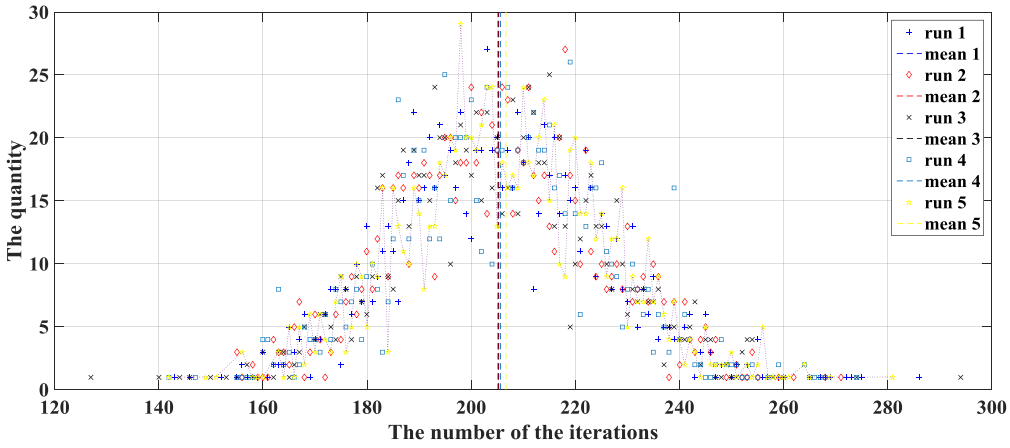


Figure 4: The results for 1 000 repetitions

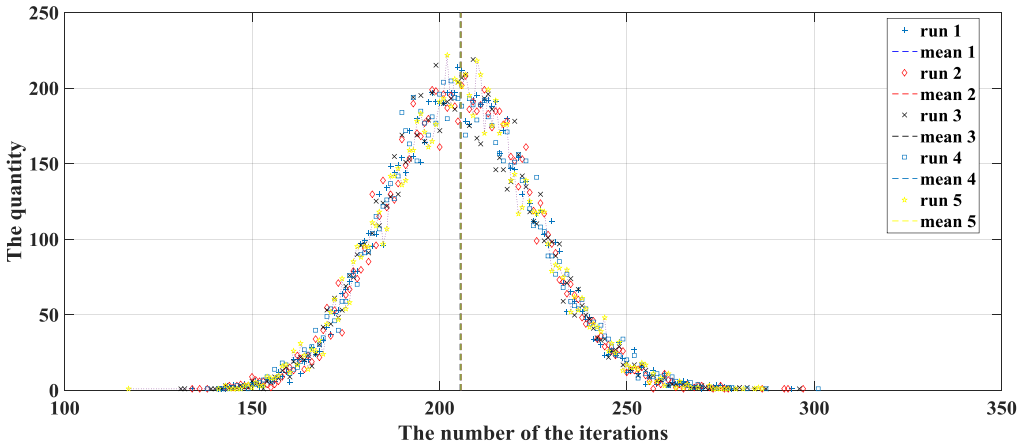


Figure 5: The results for 10 000 repetitions

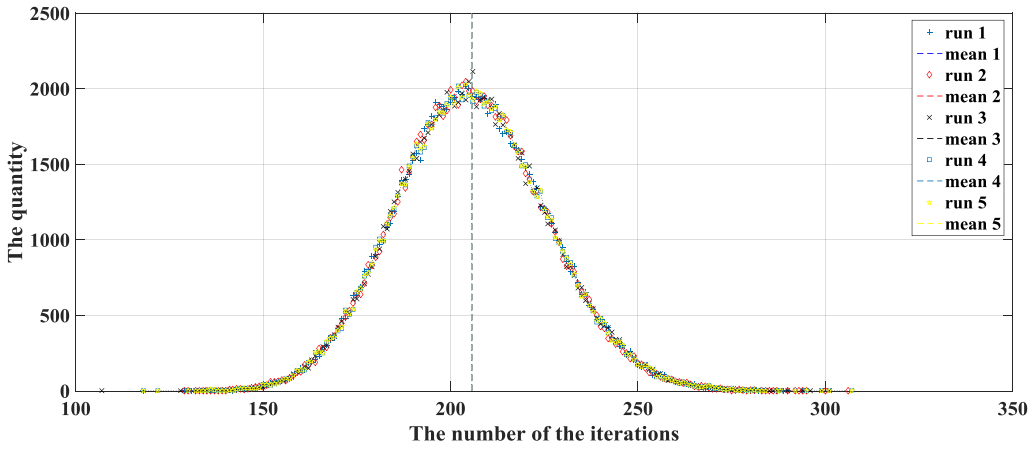


Figure 6: The results for 100 000 repetitions

	10	100	1 000	10 000	100 000
Run 1	202.1	207.2	205.298	205.7752	205.6805
Run 2	203.7	204.29	205.263	205.7373	205.6190
Run 3	204.1	207.34	205.165	205.5086	205.6778
Run 4	200.5	204.1	205.646	205.7110	205.6724
Run 5	214.2	207.59	206.677	205.7225	205.6736
Extent of variation	13.7	3.49	1.512	0.2666	0.0615

Table 1: The convergence rate mean extent of variations for the particular numbers of the repetitions

As we can see from the results, the behavior is same as in the strongly connected topology. The more time the experiment is repeated, the smaller the convergence rate mean extent of variation is for a particular number of the repetitions. Therefore, like in the previous experiment, a higher number of the repetitions results in more credible results.

5. COMPARISON WITH STRONGLY CONNECTED TOPOLOGY

In the following part, we have focused on a comparison of the convergence rate mean extent of variation in strongly (fully-connected mesh) and weakly connected (line) topology. Both structures are of the same size and had the same initial conditions. In Fig. 8, the convergence rate mean extent of variation for both topologies and for 10, 100, 1000, 10 000 and 100 000 repetitions has been shown. As we can see from the graph, the values are decreasing as the number of the repetitions is increasing. Furthermore, the weakly connected structure achieves worse results than the better-connected one.

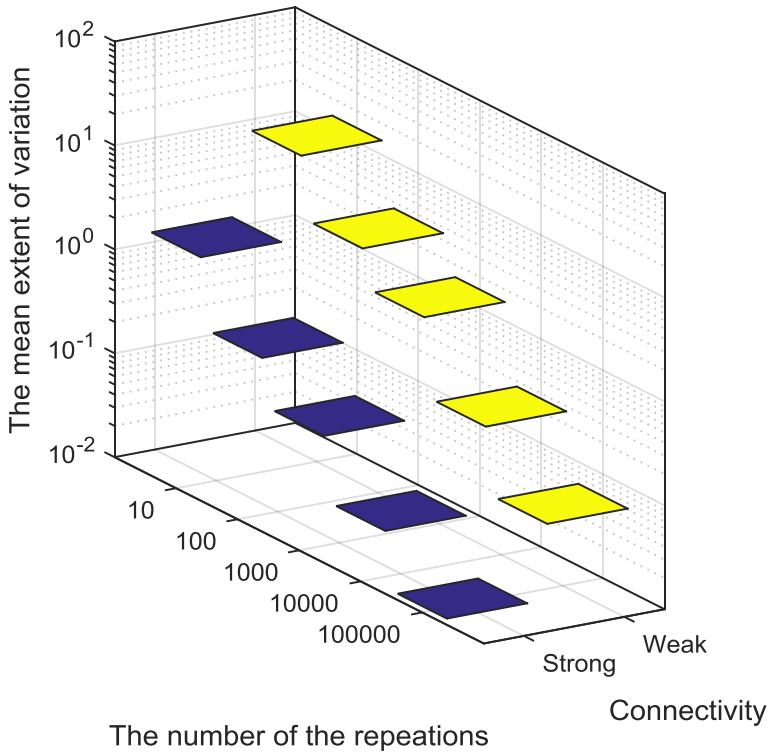


Figure 7: The results for 100 000 repetitions

6. CONCLUSION

In this paper, the impact of the numbers of the push-sum protocol repetitions on the convergence rate mean extent of variation is examined. The experiments were executed on a line topologies with the size of eight agents. The obtained results prove that increasing number of repetitions enhances statistical credibility of the gained results. The behavior of the examined phenomena is same as in the strongly connected topologies with the difference that the mentioned parameter achieves higher value in the weakly connected structure.

ACKNOWLEDGEMENT

Research described in this paper was financed by the National Sustainability Program under grant LO1401. For the research, infrastructure of the SIX Center was used.

REFERENCES

[1] Kempe, David, Alin Dobra, and Johannes Gehrke. "Gossip-based computation of aggregate information." *Foundations of Computer Science, 2003. Proceedings. 44th Annual IEEE Symposium on*. IEEE, 2003.

[2] Steiner, Stefan H., and R. Jock MacKay. *Statistical Engineering: An Algorithm for Reducing Variation in Manufacturing Processes*. Vol. 1. ASQ Quality Press, 2005.

DEPENDENCY OF THE CONVERGENCE RATE MEAN EXTENT OF VARIATION ON THE REPETITIONS NUMBER IN STRONGLY CONNECTED TOPOLOGIES

Martin Kenyeres

Doctoral Degree Programme (2), FEEC BUT

E-mail: kenyeres@phd.feec.vutbr.cz

Bohumil Novotný

Doctoral Degree Programme (3), FEEC BUT

E-mail: xnovot59@stud.feec.vutbr.cz

Supervised by: Vladislav Škorpil

E-mail: skorpil@feec.vutbr.cz

Abstract: This paper deals with the stochastic distributed algorithm – the push-sum protocol. We examine the effect of experiments repetitions on the mean of the convergence rates quantities. The main goal of the executed experiments is to show how many repetitions of the push-sum protocol are necessary to achieve a statistically credible representative of the obtained set of data. Within this paper, we have focused on strongly connected structures.

Keywords: Distributed computing, The push-sum protocol, The convergence rate mean extent of variation

1. INTRODUCTION

In the last decades, the evolution within the computer science has been strongly affected by a novel manner of communication among computers – so-called the distributed way [1]. The centralized systems, dominating in the past, are being substituted by the systems exploiting this novel computation manner. The systems executing distributed computing are labelled as the distributed systems and are usually formed by a set of the agents whose awareness of the other elements (as well as the system as the whole) is significantly limited.

In this paper, the authors‘ attention is focused on the stochastic distributed algorithm – the push-sum protocol. We examine the effect of experiments repetations on the mean of the convergence rates quantities. Even thought the experiment with more repetitions results in more statistically credible data, too extensive execution poses a demanding computing process. Thus, the motivation of this paper is to show how the number of the repetitions of the same experiment affects the credibility of the obtained results. The experiments is executed on a fully-connected topology, which is a representative of a strongly connected network.

2. THE PUSH-SUM PROTOCOL

The push-sum protocol is classified as a multi-purpose gossip-based aggregation algorithm, whose functionality is based on an iterative pair-wise distribution of the aggregated values among particular agents. Its execution is not deterministic; therefore, repetitions of an experiment may result in various outputs in spite of the constant inputs. As mentioned, the push-sum protocol can solve more problems after a smaller modification. Below is described the execution of the protocol [1]:

- The initial inner states of all the nodes represent the values from which the average is calculated. The initial weights are set to 1 for each node.

- Each node sends a half of its current inner state and a half of its weight to one of its neighbors chosen uniformly at random. The same value is stored in the inner memory and participates in further computation. This is repeated at each iteration.
- The sum of the inner state and the inner states sent by the adjacent nodes is the inner state for the next iteration. The sum of the weight stored in the inner memory and the weights sent by the adjacent nodes is the weight for the next iteration.
- The ratio of these two sums (the sum of the inner states is the numerator) poses the estimation of the average.
- The previously described procedure is repeated until the consensus is reached

3. THE EXAMINED TOPOLOGY

Within this paper, we focus our attention on the fully-connected mesh. As mentioned, the choice of this particular topology was related to its character. It is maximally connected structure, where the protocol is assumed to achieve fast convergence rates. Within this paper, a mesh topology with the size of eight agents is assumed.

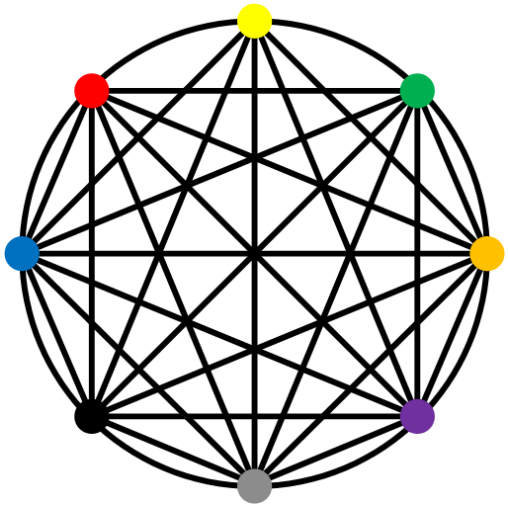


Figure 1: The examined fully-mesh topology

4. EXPERIMENTS

In this chapter, the results obtained within the numerical experiments are presented. The goal of these experiments is to show how the convergence rate mean extent of variation [2] of a data set varies when the number of repetitions changes. We change the number of the repetitions and repeat the experiment five times for each instance. In each figure, one of the repetitions is highlighted by a solid line in order to show the character of the obtained data.

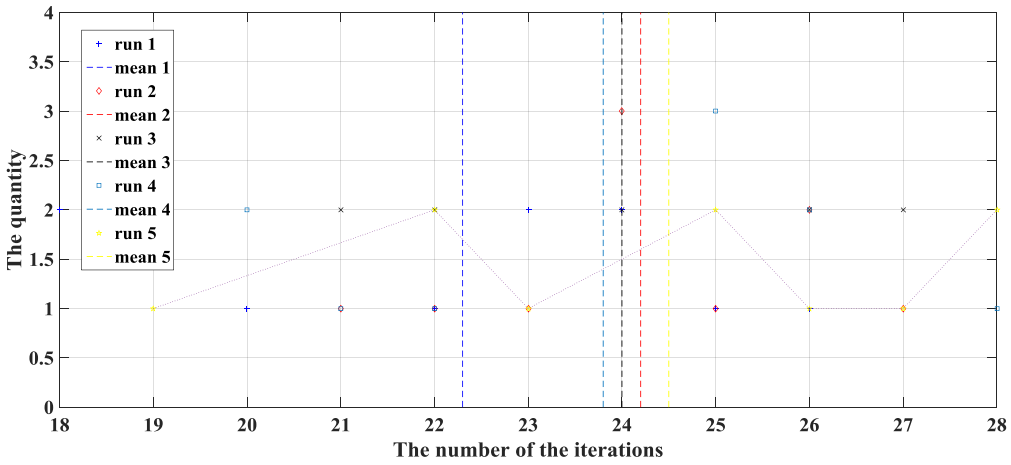


Figure 2: The results for 10 repetitions

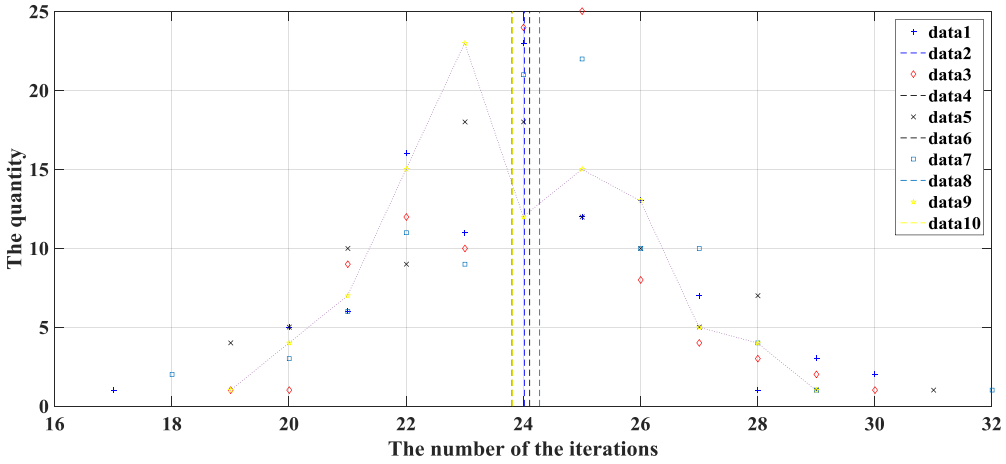


Figure 3: The results for 100 repetitions

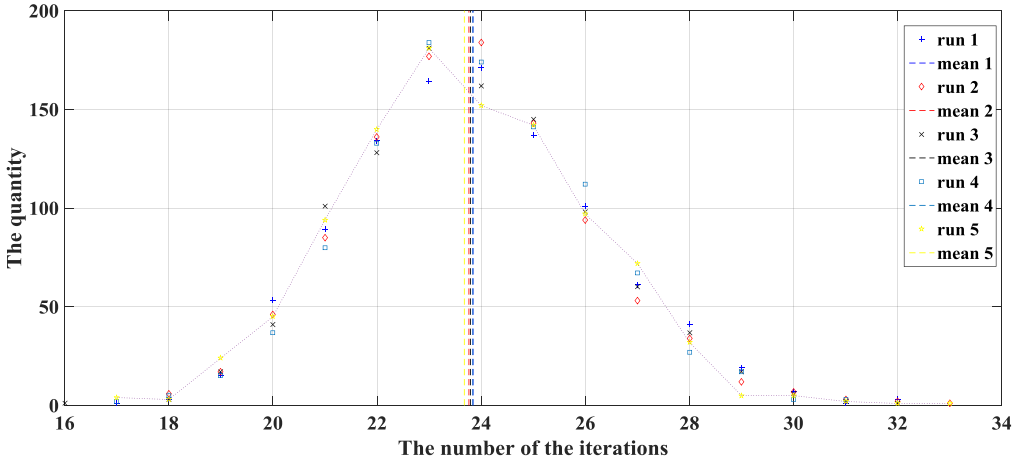


Figure 4: The results for 1 000 repetitions

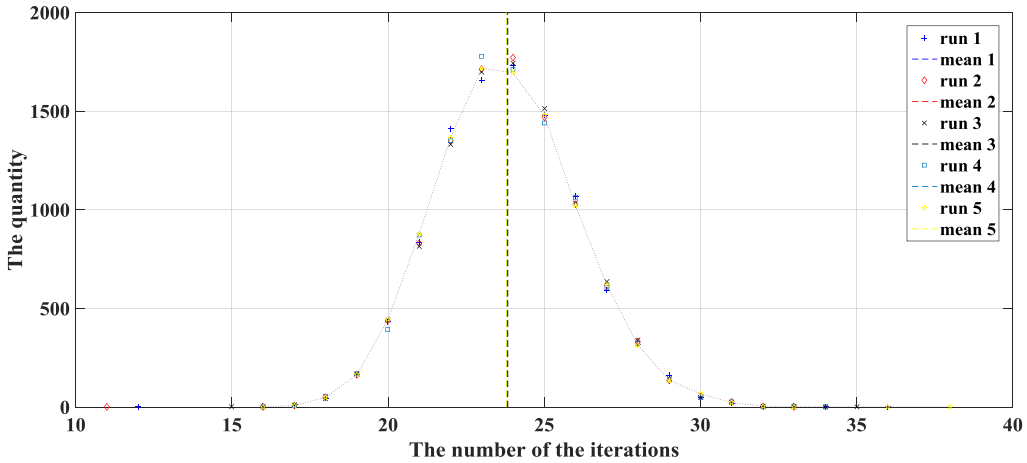


Figure 5: The results for 10 000 repetitions

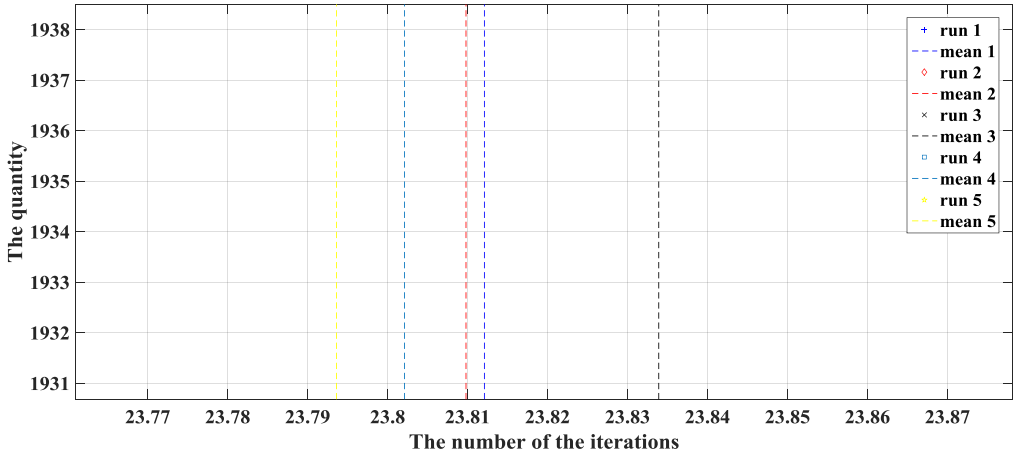


Figure 6: A detailed view of the results for 10 000 repetitions

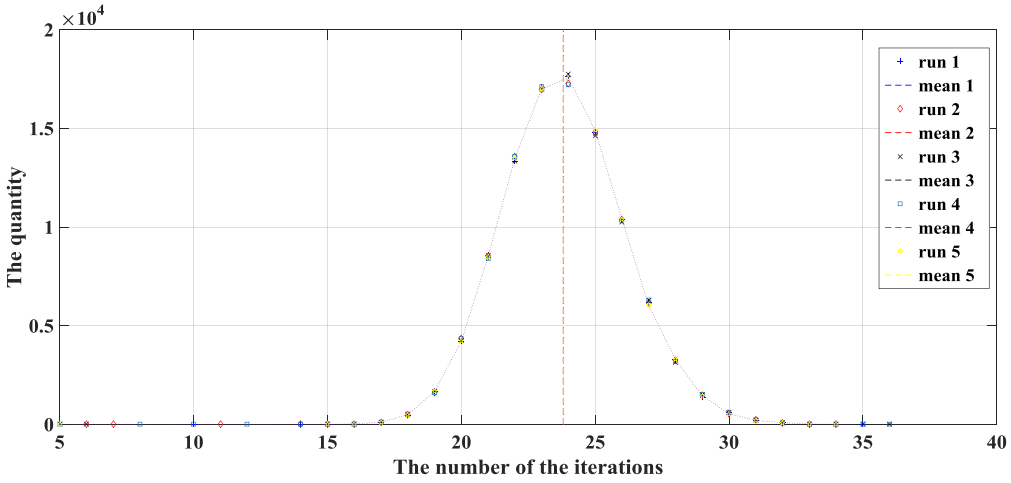


Figure 7: The results for 100 000 repetitions

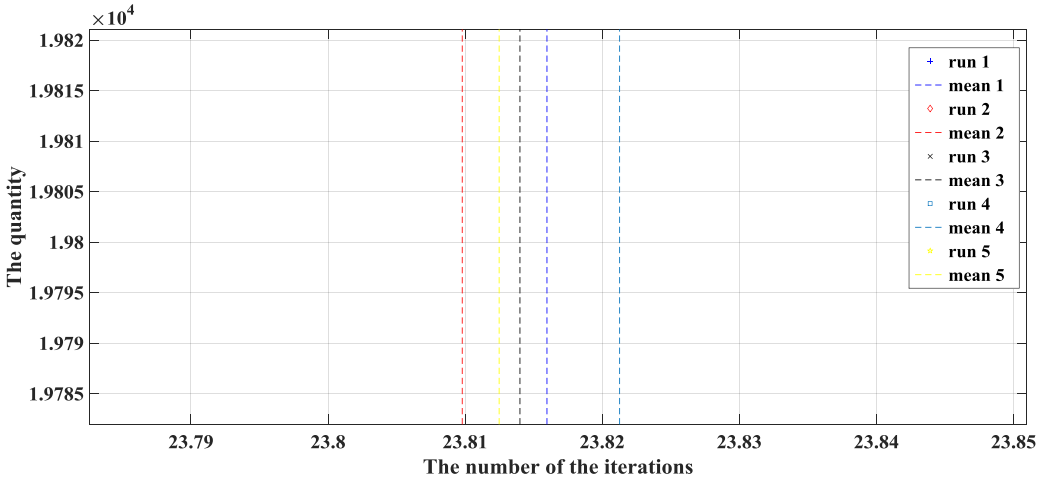


Figure 8: A detailed view of the results for 100 000 repetitions

	10	100	1 000	10 000	100 000
Run 1	22.3	24.02	23.843	23.8121	23.8160
Run 2	24.2	24.11	23.765	23.8098	23.8098
Run 3	24	23.81	23.797	23.8339	23.8140
Run 4	23.8	24.28	23.841	23.8021	23.8213
Run 5	24.5	23.82	23.677	23.7936	23.8125
Extent of variation	2.2	0.47	0.166	0.0403	0.0115

Table 1: The convergence rate mean extent of variations for the particular numbers of the repetitions

As we can see from the results, more repetitions result in a smaller extent of variation. Thus, as expected, a higher number of the repetitions ensures more credible results from the statistical point of view.

5. CONCLUSION

In this paper, the impact of the numbers of the push-sum protocol repetitions on the convergence rate mean extent of variation has been investigated. The experiments were executed on a fully-connected mesh topologies with the size of eight agents. The obtained results show that more repetitions ensure more credible results with a smaller convergence rate mean extent of variation.

ACKNOWLEDGEMENT

Research described in this paper was financed by the National Sustainability Program under grant LO1401. For the research, infrastructure of the SIX Center was used.

REFERENCES

[1] Kempe, David, Alin Dobra, and Johannes Gehrke. "Gossip-based computation of aggregate information." *Foundations of Computer Science, 2003. Proceedings. 44th Annual IEEE Symposium on*. IEEE, 2003.

[2] Steiner, Stefan H., and R. Jock MacKay. *Statistical Engineering: An Algorithm for Reducing Variation in Manufacturing Processes*. Vol. 1. ASQ Quality Press, 2005.

IN-DEPTH ANALYSIS OF SMART CITY IN MODERN AGE

Radek Fujdiak[1] **and Pavel Masek**[2]
Doctoral Degree Programme (3), FEEC BUT
E-mail: {fujdiak, masek}@feec.vutbr.cz

Supervised by: Jiri Misurec[1] and Jiri Hosek[2]
E-mail: {misurec, hosek}@feec.vutbr.cz

Abstract: The current term and concept of Smart City is nowadays not clearly defined. This article provides a robust analysis of Smart City research papers with focus on terminology and definitions. The reviews of current literature and summary of the different definitions of Smart City should help with future researches thanks to the clear definition of this area. The created summary leads to the main and common points, what Smart City is or should be.

Keywords: Smart, Modern, Digital, City, Internet of Things, Analysis

1. INTRODUCTION

The city development was always a hot topic. Improving the quality of life, services in the city, education of the people and many other factors or areas, should be a main topics and tasks of the cities. Nowadays we are seeing many new terms in the area of city research and development as a digital city, smart city, future city, modern city and others. Our analysis shows the exponential growth of interest in the research area of Smart City (see Figure 1). Smart City is nowadays a vision of

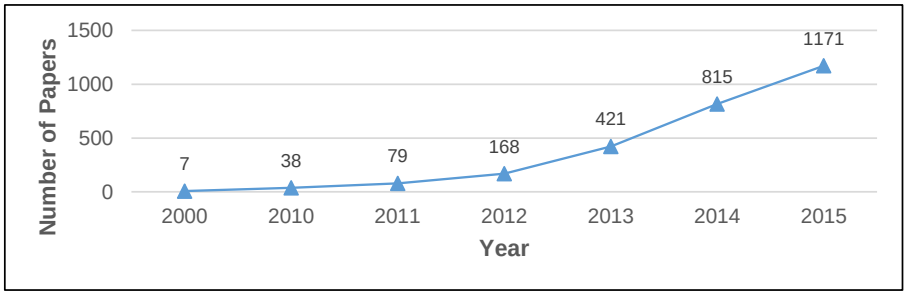


Figure 1: The growing number of papers dealing with Smart City issue merged by years (data source Web of Science database [1])

future city. However, this vision is changing based on the individual opinions on what Smart City is. The individual opinions might be impacted by geographic location, area of interest, social level, development level, education and many other factors. This lead to the fact, that it is no universally accepted definition of a smart city.

This article deals with the terminology and issue of smart city. We provide a summary and reviews of current literature, where we try to find common points in the definitions of Smart City, it's needs or challenges. This should provide a common and basic definitions for future steps in the development and research in Smart City area.

2. SMART CITY DEFINITION

Considering the situation in Europe, The European Union see Smart City issue clear. The city is becoming smarter with the partnership of the ICT, energy and transport sectors integrators. These partnerships aim to implement innovative solutions and ideas to tackle issues as congestion, air pollution, high energy costs and to achieve better mobility, cleaner urban environment and energy efficiency. Smart City should be build in the beneficial way for citizens, the city budget and the environment [2]. This definition is coming from the challenges defined by internal needs or processes in the city. However, if we should generalized the term of Smart City then we should not consider the variables, which are often changing as i.e. needs of the cities.

The second common definition of Smart City is to divide it into a different areas of interest as i.e. smart economy, smart mobility, smart governance, smart environment, smart living or smart people as in [3]. This definition or look to Smart City is very similar as the first definition with same problem. We cannot generalized the term of Smart City if we will consider local needs or areas of interest.

If we look closely to the current research papers, the clear united definition of Smart City is hard to find. We can find different definition of Smart City [4–10] and also different names for Smart City as a wired city [11], virtual city [12], ubiquitous city [13], intelligent city [14], information city [15], digital city [16], knowledge city [17], learning city [18], sustainable city [19], green city [20] and many others [21]. We can see that Smart City does not have only many different definitions, but also many different names. The Figure 2 shows five most frequent Smart City names used in professional research papers over the last years. We will keep the name Smart City, as already well known and

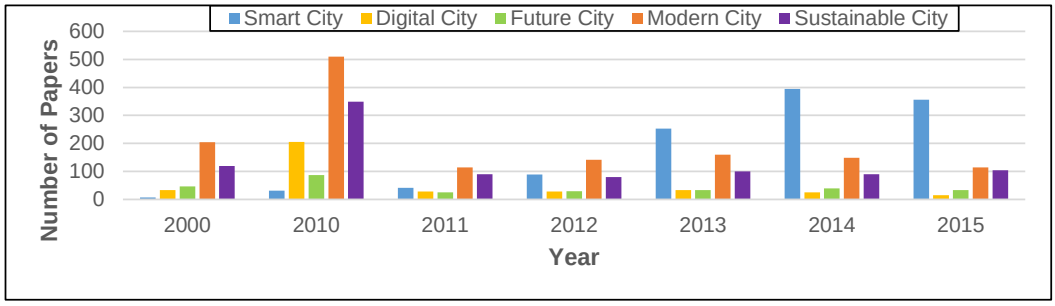


Figure 2: Frequency of different names of Smart City in published research papers merged by years (data source Web of Science database [1])

most frequent name in the professional community, and try to connect it with most common terms used in the professional papers. We analysed 815 articles from Web of Science database, the analytical algorithm is on Figure 3. The most common terms connected with Smart City in research papers

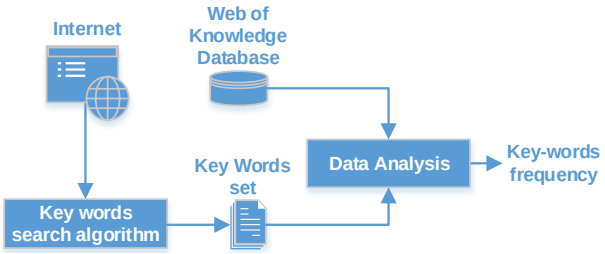


Figure 3: Schematic of our key-words frequency analysis in professional research papers

from Web of Science database are follows (in brackets is frequency of the term): technology (319), data (296), network (289), information (281), development (246), management (232), infrastructure

(185), resources (117), quality (116), knowledge (90), people (82), costs (79), innovation (79), reduce (69), digital (67), security (59), collect (58), connect (54), learn (50), vision(38), health (35), economy (32), green (30), companies (29). This shows that the interest of professional research community in Smart City focus mostly on the technology, data, networking and informations, where we can see the data and informations as same term. Most of the terms are depended on the informations/data and without it is not possible to provide sufficient solutions. This leads to the point that Smart City should be based on the informations.

Nowadays, the main methods for data or information collection are with using the information and communication technologies (ICT) [22]. In the ICT is currently the Internet of Things (IoT) considered as a global infrastructure, which will lead for big data collecting with minimal human intervention [23]. We should then consider the IoT as a basic infrastructure for Smart City, which will lead to the following challenges:

- IoT implementation - The common definition of IoT is still missing, variation of standards, technologies and solutions are considered, but common solution is also missing [23].
- Security - we need to precisely define, which information will be open and how we will authorize the access to the other data [24].
- Data analysis - collected data are unusable if we are not able to analyse them sufficiently [25].

If we will have solved the data collection challenges with it's analysis and distribution, then the other processes could start. The collected informations might be after used for interactive education and learning, citizens security, health protection and care, innovation solutions, quality improvement, environment cleaning (green solutions), urban management, private use (companies and market solutions) and many others areas [25]. Each of this solution or area will bring into the area of Smart City new challenges, but they should be seen as a challenges of internal processes of the city and not as a main challenges for Smart City concept.

This lead to the definition of Smart City based on professional interest. Therefore, Smart City should be a city, where the informations and data are collected from the different kind of sources and with different kind of systems. These data should be effectively collected and distributed only to the authorized individuals. These data are after analysed and used for making the city processes more effective. This should lead to the improvement in the various interest areas of the city. Considering the IoT concept, the several parameters should be considered: power and cost efficiency (millions of devices are considered), security (personal, trade and critical data might transmitted), technology performance (urban areas, closed environment) and many others. Several communication technologies might be used for this task. Following list shows the overview of current technologies which might be used for information/data harvest in city:

- **LPWA - Low Power Wide Area**, it is a new group of wireless telecommunication technologies designed for IoT use-cases, it offers long range, low costs and energy efficiency. Nowadays, the research and development is in radio-communication area, but not limited for other solutions. The concept was developed for areas with millions communicating devices, where is considered only lower needs for security, data-rate speed and in some cases even only one-way transmission is considered. This means areas as a trash-can monitoring, parking monitoring, other sensors applications, remote long-range metering and other applications, where we need simple low-bit rate transmission with low energy and cost parameters. This technologies are i.e. LoRaWAN, LoRa, SIGFOX, Weightless or NWave.
- **NB** - Narrow-Band or also Narrow-Band IoT are group of IoT technologies, which are considered also for long-range use-cases. Compared with the LPWAN technologies, they offer similar range, licensed band, competitive speed, higher security, higher reliability and QoS. However,

they have also higher power needs, costs and lower penetration. This technologies have different use-cases, we should use them there, where we need secure and reliable communication and when we do not need low-power or low-cost solutions. This technologies are coming mostly from 3GPP standard - NB-IoT, LTE-MTC (LTE-M), NB-LTE-IoT.

- **SRN - Short-Range Networks**, it is a group of short-range wireless networks technologies. This solutions have its place in end-customers applications, as i.e. Smart Home solutions. They provide from cost-effective solutions, energy efficiency solutions to big-data rate solutions and they are mostly limited only with the short-range use. This technologies are i.e. ZigBee, WiFi, Bluetooth, Z-Wave or Thread.
- **Others** - There are many others technologies as i.e. PLC (Power-line communication), OP (Optic-Fibre), other radio or wireless technologies, other wire technologies. These technologies have very specific use-cases and they should be considered for specific areas of communication in Smart City (infrastructure connectivity, smart grid solutions, etc.), but these technologies have for use-case of millions of devices too many limitations, because they were not developed for main IoT concept of high amount of devices.

3. CONCLUSION

We provide the robust analysis of current terminology in the area of Smart City. We analysed hundreds of professional research papers to find out the common terms. We defined Smart City term in three contexts - Europe Parliament, Western Country look and Professional look. The main definition of Smart City is based on the key-words frequency and it comes from the main principals and connections in Smart City terminology. We introduce also the main needs for Smart City as a future city vision, together with challenges and its possible solutions. The Article should be used as a guide for future researches in the area of communication technologies for Smart City.

The analyses show that the informations and data are the basic tool for the city improvements and Smart City realization, where the Internet of Things should be considered as a main infrastructure for collection these data or informations. The data are the key tool for making all the city processes more efficient. We are not able without informations or data analyse the situations, challenges and create more effective solutions. This show that the realization of the ICT infrastructure - Internet of Things, is a key challenge for Smart City concept, from the data/information point of view.

We see Smart City as a concept of city, where data are wisely collected, analysed and shared. These data should be used for improvements, bigger efficiency, higher life quality and also for dealing with everyday challenges in the city. However, Smart City should be considered as a vision for people and not for technologies. This means to always give a significant value to their opinion and their needs. We should always start with right question “What we really need” and not with “What we can”, i.e. in a context of data harvest, where the citizens lost their freedom thanks to the over sensored city.

ACKNOWLEDGEMENT

The research described in this paper was financed by the National Sustainability Program under grant LO1401. For the research, infrastructure of the SIX centre was used.

REFERENCES

[1] Web of Science, Online Subscription-based scientific citation indexing service. [Online, Accessed 15.3.2016] Available on: http://www.webofknowledge.com

[2] European Innovative Partnership, Smart Cities & Communities, Smart City. [Online, Accessed 15.3.2016] Available on: https://ec.europa.eu/digital-single-market/en/about-smart-cities

[3] Technische Universitat Wien, EuropeanSmartCities project. [Online, Accessed 15.3.2016] Available on: http://www.smart-cities.eu/

[4] ANTHOPOULOS, Leonidas; FITSILIS, Panos. From digital to ubiquitous cities: Defining a common architecture for urban development. In: Intelligent Environments (IE), 2010 Sixth International Conference on. IEEE, 2010. p. 301-306.

[5] KOMNINOS, Nicos. The architecture of intelligent clities: Integrating human, collective and artificial intelligence to enhance knowledge and innovation. In: Intelligent Environments, 2006. IE 06. 2nd IET International Conference on. IET, 2006. p. 13-20.

[6] COUCLELIS, Helen. The construction of the digital city. Environment and Planning B: Planning and design, 2004, 31.1: 5-19.

[7] ERGAZAKIS, Kostas; METAXIOTIS, Kostas; PSARRAS, John. Towards knowledge cities: conceptual analysis and success stories. Journal of knowledge management, 2004, 8.5: 5-15.

[8] LARSEN, Kurt. Learning cities: the new recipe in regional development. Organisation for Economic Cooperation and Development. The OECD Observer, 1999, 217/218: 73.

[9] BĂTĂGAN, Lorena. Smart cities and sustainability models. Informatica Economică, 2011.

[10] OECD, Green Cities Programme. [Online, Accessed 15.3.2016] Available on: http://www.oecd.org/gov/regional-policy/49318965.pdf. 2010.

[11] DEAKIN, Mark; AL WAER, Husam. From intelligent to smart cities. Intelligent Buildings International, 2011, 3.3: 140-152.

[12] Deakin, Mark. From intelligent to smart cities. In Deakin, Mark. Smart Cities: Governing, Modelling and Analysing the Transition. Taylor and Francis. p. 15. 2013.

[13] Giffinger, Rudolf; Christian Fertner; Hans Kramar; Robert Kalasek; Nataša Pichler-Milanovic; Evert Meijers. Smart cities: Ranking of European medium-sized cities. Centre of Regional Science, Vienna University of Technology, 2007.

[14] Smart Cities Council. Definition and overviews. [Online, Accessed 15.3.2016] Available on: http://smartcitiescouncil.com/smart-cities-information-center/definitions-and-overviews

[15] CARAGLIU, Andrea; DEL BO, Chiara; NIJKAMP, Peter. Smart cities in Europe. Journal of urban technology, 2011, 18.2: 65-82.

[16] SINGH, Sarwant. Smart cities: a $1.5 trillion market opportunity. Forbes. Retrieved, 2014, 4.

[17] IEEE Smart Cities, About. [Online, Accessed 15.3.2016] Available on: http://smartcities.ieee.org/about.html

[18] Business Dictionary. Smart City - Definition. [Online, Accessed 15.3.2016] Available on: http://www.businessdictionary.com/definition/smart-city.html

[19] Government of India - Ministry of Urban Development, Concept on Smart City Scheme. 2014.

[20] HOLLANDS, Robert G. Will the real smart city please stand up? Intelligent, progressive or entrepreneurial? City, 2008, 12.3: 303-320.

[21] SCHULER, Doug. Digital cities and digital citizens. In: Digital Cities II: Computational and Sociological Approaches. Springer Berlin Heidelberg, 2001. p. 71-85.

[22] DUTTA, Soumitra; GEIGER, Thierry; LANVIN, Bruno. The global information technology report 2015. In: World Economic Forum. 2015.

[23] ROSE, Karen; ELDRIDGE, Scott; CHAPIN, Lyman. The internet of things: An overview. The Internet Society (ISOC), 2015.

[24] SHIPLEY, A. J. Security in the Internet of Things: Lessons from the Past for the Connected Future. Security Solutions, Wind River, White Paper, 2013.

[25] VERMESAN, Ovidiu; FRIESS, Peter (ed.). Internet of Things-From research and innovation to Market Deployment. River Publishers, 2014.

PULSE GENERATOR FOR Φ–OTDR IN VPIPHOTONICS

Milan Cucka
Doctoral Degree Programme (2), FEEC BUT
E-mail: xcucka00@stud.feec.vutbr.cz

Supervised by: Miloslav Filka
E-mail: filka@feec.vutbr.cz

Abstract: The paper focuses on ϕ–OTDR (Optical Time Domain Reflectometry) systems which need short time duration and high powered pulses for proper function, because the backscattered signal is weak for proper detection. The article includes a simulation of the signal generator for ϕ–OTDR sensoric systems in simulation software. Designed generator allows to try pulses, with different time of duration, pulse repetition frequency and power. The article also includes scientific notation of backscattered signal and a short description of the used systems.

Keywords: Optical fiber sensors, Pulse generator, Simulations, VPIphotonics, Φ–OTDR.

1 INTRODUCTION

Nowadays, optical sensors are widely used for temperature sensing, preassure sensing or vibration sensing due to their physical properties. Optical sensors are resistant to high temperature, and also accuracy of the sensing vibrations and preassure is very good. These dynamic processes involve time–varying signals such as motor vibration, shift of the building structures and walking person. It is necessary to monitor not only the position but also to accurately distinguish recevied signal. That means processing the frequency of the signal from a few Hz to hundreds kHz, or tens of MHz. Classification of individual frequencies must be very precise. Commonly used systems are not able to classify vibration without complicated postprocessing and measurement cannot be real time [1].

The disadvantage of optical sensors is their high price. Thanks to new developments the price has been constantly decreasing during last years. In recent years, most optical sensors have used a combination of several sensing principles. Most commonly used are backscattered systems based on Rayleigh or Brillouin scattering combined with any interferometric method (Mach–Zehnder interferometer, Michelson interferometer) [2], [3], [4].

Huang et al. [3] describe Mach–Zehnder interferometer combined with ϕ–OTDR. System in the article can detect and localize vibration with spatial resolution of 100 m on 8 km long optical fiber. The system operates at a frequency starting from 100 Hz to 2 kHz and uses phase comparison and monitoring changes in the amplitude of the received signal. To improve spatial resolution ϕ–OTDR system is used. The received signal is processed with FFT (Fast Fourier Transform) analysis and high–speed data processing. Thanks to the complex process measurement can be real time. The main point of this article is mathematically described received signal. A description of transmitted pulses and their duration and power which uses ϕ–OTDR are missing in the paper.

The authors in [4] also describe combination of MZI (Mach–Zehnder Interferometer) and ϕ–OTDR. This sensor device requires a wide frequency response. For proper function two narrow lasers which have slightly different wavelengths are needed. The first laser is used for interferometer and the second one for ϕ–OTDR. Laser beams are separated by FBG (Fiber Bragg grating) and heterodyne detection is used. This system operates at a frequency starting from 1 Hz to 50 MHz, unfortunately there is a limited sampling rate for data capture. The spatial resolution of the sensoric system is 20 m

and the system can operate on 2.5 km. Generation of the transmitted signal is better described in this article, but a detailed explanation of the system is missing.

The article describes the simulation of pulse signal generator for ϕ–OTDR. For simulation the VPIphotonics software was used. To simulate optical sensors no appropriate software have been developed yet. VPIphotonics is mainly used for simulation of data networks. Many optical elements are prepared as optical fibers, lasers, modulators. But any pulse generator for sensoric systems. The advantage of our simulation is the ability to set any pulse power level together with different duration and pulse repetition rate.

2 Φ–OTDR SENSOR

Devices which use this type of fiber sensor are applied for vibration sensing or mechanical tense sensing. The extent of these vibrations that could be compared to the acoustic wave is from 100 Hz to 17 kHz. These systems most often use Rayleigh and Brillouin scatterings. Phase sensitive OTDR sensor commonly works with a single mode optical fiber G.652.D, which is standardly used for data transmission.

Repetition rate can be calculated as follows:

$$f \leq = \frac{1}{\tau}[Hz], \tag{1}$$

explanation of time [6]:

$$\tau = \frac{2 \cdot s}{v}[s], \tag{2}$$

where we use s as the total fiber length and v is the velocity of light in an optical fiber [6]:

$$v = \frac{c}{n} = 2 \cdot 10^8 [m/s], \tag{3}$$

in this equation c is the velocity of light in the vacuum and n is the refractive index.

The pulse repetition rate is dependent on the maximal length of the optical fiber, pulse repetion rate can be expressed for simulation system as [5]:

$$\tau = \frac{2 \cdot s}{v} = \frac{2 \cdot 100 \cdot 10^3}{2 \cdot 10^8} = 1 \cdot 10^{-3}(s), \tag{4}$$

where repetition rate can be calculated:

$$f = \frac{1}{\tau} = \frac{1}{1 \cdot 10^{-3}} = 1 \ (kHz). \tag{5}$$

From these equations we are able to calculate the spatial resolution for the pulse width for example of 2 μs [6]:

$$\Delta z = \frac{v \cdot T_p}{2} = \frac{2 \cdot 10^8 \cdot 2 \cdot 10^{-6}}{2} = 200 \ m. \tag{6}$$

Optical pulses coupled into the fiber must be short time and high powered, which can be propagated in optical fiber. Then the part of light is continuously backscattered because of Rayleigh or Brillouin

scattering. This beam is process by processing the scattered signal. Capturing and processing of the signal is the most used technique. Phase sensitive OTDR needs pulses with short time duration and high powered because the backscattered signal is very low and can be described as [5]:

$$P_R(t) = \frac{P_i \cdot S_{\gamma R} \cdot W_0 \cdot \upsilon_g \cdot exp^{-\gamma \cdot v_g \cdot t}}{2} = P_R(0) \cdot exp^{\gamma \cdot v_g \cdot t}, \tag{7}$$

where γ is the attenuation coefficient and is equal to $1/20^{th}$ of the gradient of a plot of Rayleigh backscatter attenuation in dB, P_i is the input power, γR is the Rayleigh scattering coefficient, S is the fraction of captured optical power which is defined for step index fibers as [5]:

$$S = \frac{NA^2}{4n^2}, \tag{8}$$

where NA is the numerical aperture and n is the fiber refractive index.

The pulse in our simulation model has inputed power 150 mW and pulse duration 2 μs. If we use equation $P_R(0)$ as is described in [5] it can be seen that backscatter power is 52 dB lower than the input power.

3 SYSTEM DESCRIPTION

The simulation is performed in the program VPIphotonics, which is a professional simulation software supports active or passive photonic integrated circuits, application of optical fiber, optical transmission systems, network applications and their subsequent analysis.

Simulation in VPIphotonics software consists of two different pulses. Every pulse has same power but different repetition rate and pulse duration. The sensoric part which is depicted in Fig. 1 uses block Rect which can generate pulses of the required parameters. Block PckBlockE simply sets the repetition rate of the generator. Other blocks (1D, Signal Analyzer) are used to track the signal. This system is in Fig. 1. Block Fork in Fig. 1 is used for proper function of simulation. This block is not nessesary to understand.

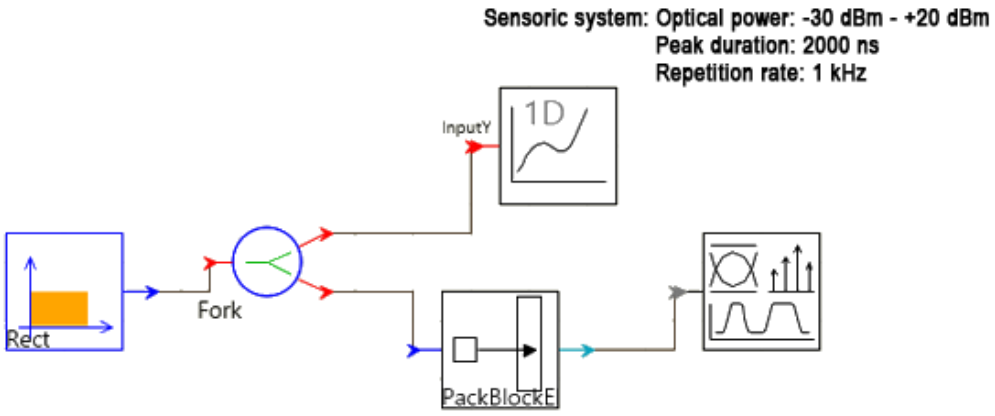


Figure 1: Simulation of pulse generator for ϕ–OTDR.

4 RESULTS AND DISCUSSION

Simulation consists of two different pulses. The first pulse has duration of 250 ns and repetition rate 0.124 ms that mean 10 kHz. The pulse have 25 m spatial resolution and can be used on optical fiber of 12.4 km long. This signal is in Fig. 2. The second pulse with duration of 2 μs and repetition rate of 1 kHz can be used with 124 km long optical fiber. The spatial resolution is 200 m. A graph in Fig. 3 shows the signal from simulation.

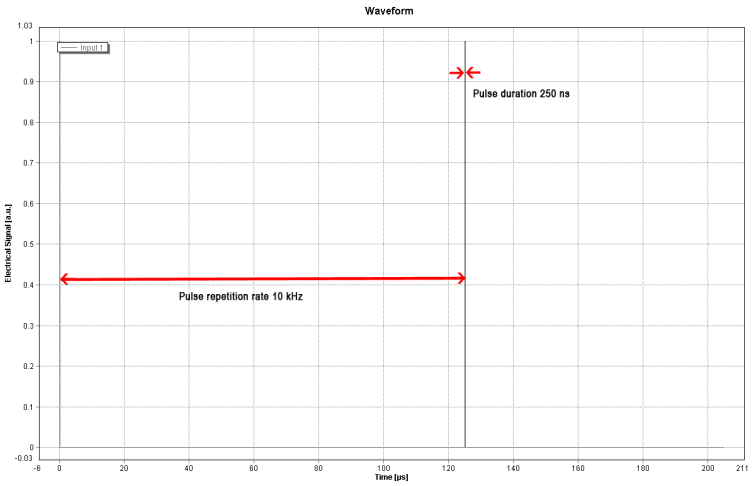


Figure 2: Signal with pulse duration of 250 ns and repetition rate of 10 kHz

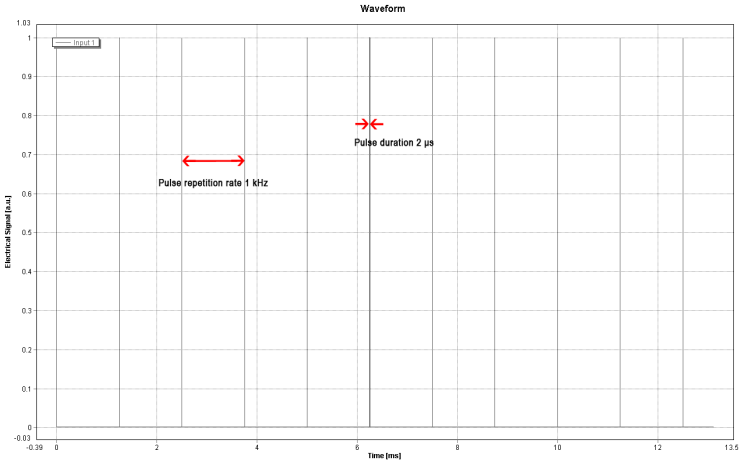


Figure 3: Signal with pulse duration of 2000 ns and repetition rate of 1 kHz

5 CONCLUSION AND FUTURE WORK

The article showed simulation setup in VPIphotonics software which describes pulse generator for ϕ–OTDR. For verification two sample situations are described. The first simulation use pulse for 12.4 km long optical fiber with spatial resolution 25 m. The second one for 124 km long optical fiber with spatial resolution 200 m. The article also includes equations which describe the calculation of pulse repetition rate frequency, pulse duration and backscattered power. To confirm the simulation it is nessesary to create pulse generator in another simulation software, for example in Simulink. For the proper results it is better to compare those two simulation softwares. The future measurements confirm the simulation.

REFERENCES

[1] T. Zhu, Q. He, X. Xiao, X. Bao, "Modulated pulses based distributed vibration sensing with high frequency response and spatial resolution," *Optics Express*, vol. 21, no 3, pp. 29–53. 2013.

[2] Y. Muanenda, C. Oton, S. Faralli, F. D. Pasquale, "A Cost-Effective Distributed Acoustic Sensor Using a Commercial Off-the-Shelf DFB Laser and Direct Detection Phase–OTDR, rq' *IEEE Photonics Journal*, vol. 8, no. 1, pp. 1–10, 2016.

[3] S. Huang, W. Lin, M. Tsai, M. Chen, "Fiber optic in-line distributed sensor for detection and localization of the pipeline leaks," *Sensors and Actuators*, vol. 135, no 2, pp. 570–579. 2006.

[4] Y. Shi, H. Feng, "Distributed fiber sensing system with wide frequency response and accurate location," *IEEE Sensors Journal*, vol. 77, no. 3, pp. 219–224. August 2015.

[5] P. Munster, J. Vojtech, P. Sysel, R. Sifta, V. Novotny, T. Horvath, S. Sima, M. Filka, "ϕ–OTDR signal amplification," *Optical Sensors*, vol. 9506, pp. 950606-1, 2015.

[6] R. Sifta, P. Munster, P. Sysel, T. Horvath, V. Novotny, O. Krajsa, M. Filka. "Distributed Fiber-Optic Sensor for Detection and Localization of Acoustic Vibrations," *Metrology and Measurement Systems*, vol. 22, no. 1, 2015

xcucka00@stud.feec.vutbr.cz
filka@feec.vutbr.cz

RELATIVE PERMITTIVITY MEASUREMENT OF 3D-TEXTILE

Jan Vélim
Doctoral Degree Programme (3), FEEC BUT
E-mail: xvelim00@stud.feec.vutbr.cz

Supervised by: Zbyněk Raida
E-mail: raida@feec.vutbr.cz

Abstract: The paper deals with the relative permittivity determination of the 3D-textile by a transmission line method (TLM). Measuring points are shifted to avoid inaccuracies caused by transitions between a coaxial cable and a microstrip line. A substrate where the microstrip line is designed is composed of three layers where one layer (3D-textile) has unknown permittivity. The relative permittivity of the 3D-textile is calculated from effective permittivity by a conformal mapping method.

Keywords: 3D-textile, permittivity, microstrip line

1 INTRODUCTION

Nowadays, still higher emphasis is on reduction of vehicles fuel consumption. One of the ways to reduce the fuel consumption is to reduce weight of the vehicles. Because today's vehicles contain plenty cable harnesses which significantly increase the weight, a trend of car makers is to replace the cables by wireless communication links. Unfortunately the wireless links are sensitive to movement of passengers, seats etc. An alternative solution could be exploitation of construction parts of the vehicles for a guiding of the electromagnetic waves which means that such phenomenon could be used to communicate instead of wires. This phenomenon is going to be used in a communication link along inner side of the roof where an upholstery serves as a waveguide. The upholstery is made of a material known as 3D-textile. Because the 3D-textile serves as a substrate on which the waveguide is going to be designed it is necessary to know properties like the relative permittivity and the propagation losses. The 3D-textile is anisotropic material composed of three layers (Figure 1). The top and bottom layers of the 3D-textile look like common textile separated by plenty threads which are perpendicular to these layers. The whole structure has thickness equal to 3.4 millimetres. The material composition of the measured sample is: 17 % of the polyester silk, 63 % of the polyester monofilament and 20 % of the cotton.

Several articles have been published about measuring of the permittivity by a resonance method [1], a free space method [2] and the transmission line method [3, 4]. The transmission line method (TLM) is the most suitable from a frequency band point of view.

The TLM is based on the knowledge of the propagation of the electromagnetic waves in a transmission

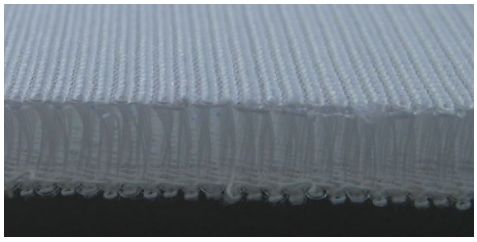


Figure 1: The detailed view of the 3D-textile.

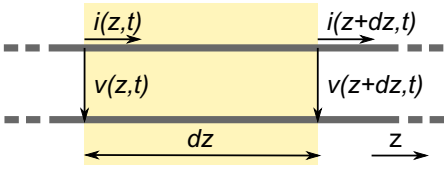


Figure 2: Infinitesimal segment of the general transmission line, voltage and current definitions.

line segment. The book [5] proves that the wave equations (1, 2) for an infinitesimal segment of the transmission line segment is possible to derive from the Maxwell equations and the equivalent circuit. The wave equations describe the distribution of voltage and current in the line segment (Figure 2).

$$\frac{d^2V(z)}{dz^2} - \gamma^2 V(z) = 0, \tag{1}$$

$$\frac{d^2I(z)}{dz^2} - \gamma^2 I(z) = 0, \tag{2}$$

where γ is the propagation constant. It is possible to find a solution of the equation (1) in shape:

$$V(z+dz) = V(z)e^{-\gamma dz}. \tag{3}$$

The equation (3) is valid also in the case that the infinitesimal length dz is replaced by a real length l. However small modification of the equation (3) brings the general formula of the transmission coefficient T for the transmission line segment with the length l:

$$T = e^{-\gamma l}. \tag{4}$$

2 PRECISE MEASUREMENT OF THE TRANSMISSION COEFFICIENT

Determination of a transmission characteristic of the transmission line segment is carried out by a vector network analyser (VNA). Transitions between the VNA and the measured transmission line is realized by connectors. Calibration is performed usually at the connectors by precisely defined loads. However this procedure brings values of scattering parameters of the transmission line with the connectors $[S_L]$ (Figure 3 a). From practical experience there are serious reasons to assume that the connectors bring significant inaccuracies in the measurement. Therefore the whole structure is divided into three parts with appropriately determined lengths $l_{CON} = 60$ mm and $l_{LINE} = 15$ mm. The lengths are chosen so that only the quasi transverse electromagnetic (Q-TEM) mode is present in the middle section. In order to calculate a transmission matrix only of the middle segment $[T]$, the measurement of the THRU calibration substrate (Figure 3b) has to be done.

Because of difficulty mathematical operations with the scattering matrices $[S_L], [S_T]$, the matrices are converted to transmission matrices$[T_L], [T_T]$ [5],

$$T_{11} = \frac{(1+S_{11})(1-S_{22}) + S_{12}S_{21}}{2S_{21}}, \tag{5}$$

$$T_{12} = Z_0 \frac{(1+S_{11})(1+S_{22}) - S_{12}S_{21}}{2S_{21}}, \tag{6}$$

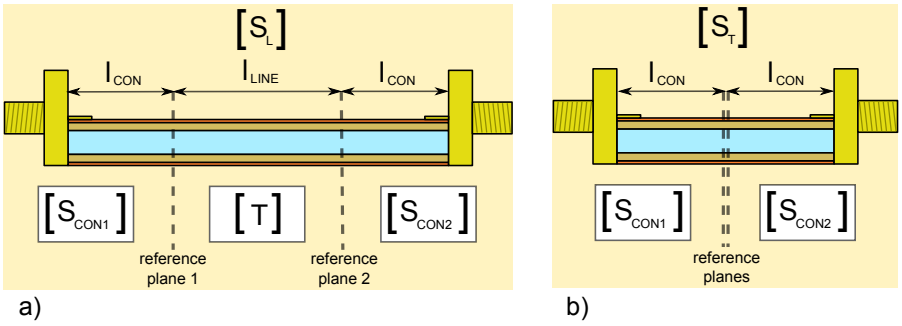


Figure 3: The transmission line section with connectors sections (a), the connectors sections only (b).

$$T_{21} = \frac{1}{Z_0}\frac{(1-S_{11})(1-S_{22})-S_{12}S_{21}}{2S_{21}}, \tag{7}$$

$$T_{22} = \frac{(1-S_{11})(1+S_{22})+S_{12}S_{21}}{2S_{21}}, \tag{8}$$

where Z_0 is the characteristic impedance of the connectors.

Rolain in his article [4] proved that the transmission matrix of the middle segment is possible to express by *eig* function. The transmission matrix of the transmission line with connectors and inverted transmission matrix without connectors are input arguments of the function. Eigen values are outputs of the function which correspond with the transmission coefficient of the middle segment.

$$[T] = eig(T_L T_T^{-1}) = \begin{pmatrix} e^{-\gamma l_{LINE}} \\ e^{\gamma l_{LINE}} \end{pmatrix}. \tag{9}$$

3 CALCULATION OF THE RELATIVE PERMITTIVITY

We know several kinds of transmission lines which are used in the cm-wave frequency band (eg. coaxial cables, microstrip lines, coplanar waveguides, grounded coplanar waveguides, rectangular waveguides, etc.). With regard to easy manufacturing, two samples of the microstrip lines with the different lengths are used in this investigations (Figure 4). The dielectric substrate is composed of three layers where only the middle layer (3D-textile) has unknown permittivity. The top and the bottom layers are made from Arlon 25N of the thickness 0.762 mm, the relative permittivity 3.38 and the propagation loss 0.0025.

To ascertain the values of the effective relative permittivity of the layered substrate, the propagation constant must be expressed from the equation (9):

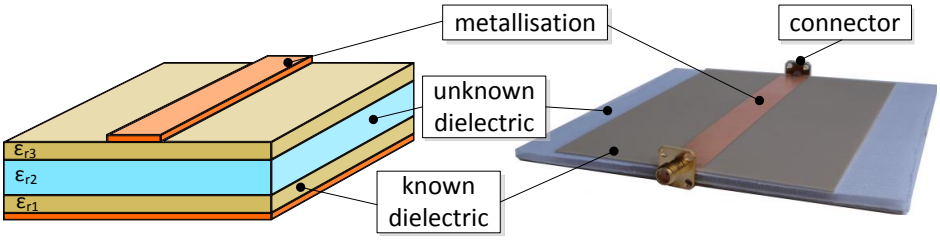


Figure 4: The layered substrate of the microstrip line (left) and the measured sample (right).

$$\gamma = -\frac{ln(T)}{l_{LINE}}. \tag{10}$$

The propagation constant is a complex number $\gamma = \alpha + j\beta$, where α is the attenuation constant and β is the phase constant. As mentioned, the mode Q-TEM is propagating in the required section therefore a relationship between the phase constant and the effective relative permittivity is possible to express by [6]:

$$\varepsilon_{r_{eff}} = \left(\frac{c_0}{\omega}\beta\right)^2, \tag{11}$$

where c_0 denotes the speed of light in vacuum and ω the angular frequency.

The last step involves calculating the relative permittivity from the effective relative permittivity. This is accomplished using the conformal mapping method [7]:

$$\varepsilon_{r_2} = \frac{q_2(1-q_1-q_2-q_3-\varepsilon_{r_{eff}})}{\frac{q_1\varepsilon_{r_{eff}}}{\varepsilon_{r_1}} + \frac{q_3\varepsilon_{r_{eff}}}{\varepsilon_{r_3}} - \frac{q_1(1-q_1-q_2-q_3)}{\varepsilon_{r_1}} - \frac{q_3(1-q_1-q_2-q_3)}{\varepsilon_{r_3}} - (1-q_1-q_2-q_3)^2}, \tag{12}$$

where q_1, q_2, q_3 denote filling factors and $\varepsilon_{r_1}, \varepsilon_{r_2}, \varepsilon_{r_3}$ denote the relative permittivity of the layers numbered from the bottom. Calculation of the filling factors depends on real dimensions of the structure.

In Figure 5 the final values of the relative permittivity of the 3D-textile are depicted in the whole measured frequency band. Decline of the values of the relative permittivity in the range above the 8 GHz is caused by the emergence of higher modes. For the measuring at the higher frequencies, a different transmission line has to be used. Because of using of imperfect cables for the measuring, the value of the relative permittivity is oscillating around the real value therefore the interpolation technique is used to display the real value.

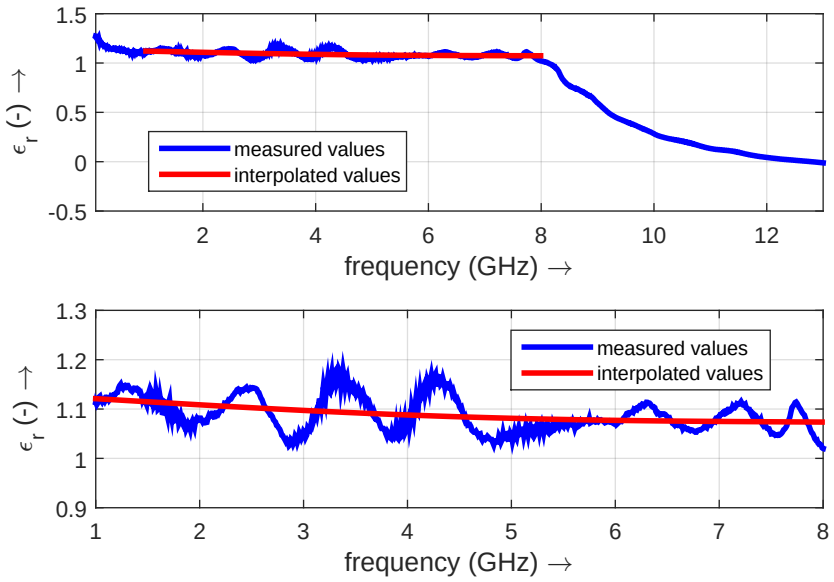


Figure 5: The relative permittivity in the frequency band 0.1 up to 13 GHz (top), detailed view of the relative permittivity in the frequency band 1 up to 8 GHz (bottom).

4 CONCLUSION

The easy straightforward method of the relative permittivity measurement in the wide frequency band is described. The value of the relative permittivity of the 3D-textile is about 1.1 in the frequency band from 0.1 up to 8 GHz (Figure 5). The accuracy is possible to improve by using a new cables and by multiple repetition of the measurement.

ACKNOWLEDGEMENT

Research described in this paper was financially supported by the Czech Science Foundation under grant no. P102/12/1274. Research is a part of the COST Action IC1301 supported by the grant LD14057 of the Czech Ministry of Education. The presented research was supported by the Internal Grant Agency of Brno University of Technology project no. FEKT-s-14-2483. For research, infrastructure of the SIX Center was used.

REFERENCES

[1] S. Sankaralingam and B. Gupta, "Determination of Dielectric Constant of Fabric Materials and Their Use as Substrates for Design and Development of Antennas for Wearable Applications," in IEEE Transactions on Instrumentation and Measurement, vol. 59, no. 12, pp. 3122-3130, Dec. 2010.

[2] N. Tamyis, A. Ramli and D. K. Ghodgaonkar, "Free space measurement of complex permittivity and complex permeability of magnetic materials using open circuit and short circuit method at microwave frequencies," Research and Development, 2002. SCOReD 2002. Student Conference on, 2002, pp. 394-398.

[3] S. F. Barot and J. T. Bernhard, "Permittivity measurement of layered media using a microstrip test bed," Antennas and Propagation Society International Symposium, 2008. AP-S 2008. IEEE, San Diego, CA, 2008, pp. 1-4.

[4] Y. Rolain, A. Barel, F. Gubbels and W. van Moer, "Microwave permittivity measurement of polymers by deposition on a coplanar wave guide," Instrumentation and Measurement Technology Conference, 2004. IMTC 04. Proceedings of the 21st IEEE, 2004, pp. 333-338 Vol.1.

[5] Pozar David M. Microwave engineering. Wiley, 2012.

[6] P. Seiler, B. Klein and D. Plettemeier, "Broadband characterization of planar transmission line substrate permittivity up to 67 GHz," Antennas and Propagation (ISAP), 2014 International Symposium on, Kaohsiung, 2014, pp. 373-374.

[7] SVAČINA, J.: Řešení mikrovlnných planárních struktur metodou konformního zobrazení. Akademické nakladatelství CERN, 2006

STRONG UNILATERAL AUTHENTICATION OF LOW-COST DEVICES INVOLVED IN INTERNET OF THINGS IN SMART HOMES

Vlastimil Clupek
Doctoral Degree Programme (4), FEEC BUT
E-mail: clupek@phd.feec.vutbr.cz

Supervised by: Vaclav Zeman
E-mail: zeman@feec.vutbr.cz

Abstract: In this paper we focused on authentication of low-cost devices involved in Internet of Things (IoT) in smart homes. We proposed a unilateral authentication protocol for authentication of low-cost devices involved in IoT with a Control Unit (CU) of a smart house. The protocol allows secure transmission of secret information between communication entities. Only cryptographic primitives as Hash functions, XOR operations and Physical unclonable functions (PUFs) are necessary for execution of the protocol. Security of the proposed protocol is described.

Keywords: Authentication, Physical unclonable functions, Low-cost devices, IoT, Smart homes.

1 INTRODUCTION

Nowadays, Information technology (IT) has become an inseparable part of human lives. The fastest growing part of IT is Internet. Internet is used by many people on the Earth everyday. The form of Internet is still in an evolution. Internet directs to its further phase now. The most important role in this phase will play IoT. According to [1] it is expected that IoT, which excludes PCs, tablets and smartphones, will grow to 26 billion units installed in 2020. This is a huge number and will play an important role in human lives. IoT usually uses low-cost devices (which represent computationally and resource constrained devices) to collect and exchange data from sensors and another equipment by Wireless sensor network (WSN), RFID technology, 3G and 4G connections, Wi-Fi, WiMAX and so on. It is presumed that these devices will be everywhere around us and will help us in our day lives. These devices will be implemented in smart homes, smart cities, hospital environment, industry monitoring applications and so on. Since data from these devices are crucial for specialized applications, it must be secured.

The most important terms in cryptographic security are authentication, confidentiality, integrity and non-repudiation. With using authentication protocols it is possible to verify authenticity of entities during electronic communications. Confidentiality ensures that only authorized entities can know the secret information. Integrity ensures that data were not modified during transmission between communication entities. Non-repudiation ensures that any communication entity cannot deny some fact which was done before. Authentication between communication entities can be unilateral or mutual. In case of unilateral authentication or one-way authentication, the first entity is authenticated to the second entity or the second entity is authenticated to the first entity. In case of mutual authentication communication entities are authenticated to each other. Mutual authentication is typically used to ensure an extra level of security, for example, in financial transactions between organizations or in some cases of access control.

2 PHYSICAL UNCLONABLE FUNCTIONS

Many PUFs concepts have been presented over the past thirty years. The first studies cannot be called PUFs by their authors, but they may have some PUF properties. This is due to the fact that the PUF concept was systematically written only by Ravikanth Pappu in 2001. Pappu described the concept of PUF in his dissertation thesis [2]. Pappu called PUFs Physical one-way functions (POWFs) and defined them as functions which are easy to compute but hard to invert. Most concepts of PUFs are based on heterogeneities and the differences between physical components of a device, which are caused by manufacturing differences. Heterogeneities of the produced components are random and they cannot be controlled by a manufacturing process. For this reason PUFs cannot be cloned. PUFs represent an alternative to classic store of secret keys, which are stored in nonvolatile memories. In PUFs secret keys are generated during an authentication process with using unique properties of the concrete device. These unique properties depend on variations in the manufacturing process. In case a secret key is stored in a nonvolatile memory, an attacker gets it. After that the attacker can perform a clonning attack. A protection against this attack can be ensured by encryption of secret data if a system is switched off. If the system is switched on, secret data are decrypted. PUFs can be used as an alternative to a protection against clonning attacks. Most PUFs exploit the differences from chips, which are caused by the differences in a manufacturing process, to generation unique keys by direct outputs from system circuits of chips without explicitly storing them. PUFs can be used for authentication (identification of chip) and generation of the secret keys. PUFs are functions with an internal random character, which mostly exploit manufacturing variabilities to generation unpredictable responses. Therefore, responses of PUFs are compared to fingerprints. An input of PUF is called *challenge* and an output of PUF is called *response*. PUFs are not really mathematical functions since these functions are able to produce several output values from one input value or from several input values they can produce one output value. $Challenge-response\ pair$ (CRP) is created by mapping of *challenges* to corresponding *responses* $(response = PUF(challenge))$. Authentication by PUFs consists of two phases. The first is an initialization phase in which a PUF is inquired and CRPs are stored in a database. In the second phase a *challenge* from the database is applied to a PUF and an obtained *response* is compared with the corresponding *response* which is stored in the database. If they are equal, the device will be authenticated. This phase is called an identification phase or a verification phase.

PUFs can be created by many ways. Many of them can be used to authentication on low-cost devices. Optical PUFs are represented in [3] and [4]. Other papers use a ring oscillator (RO) to build PUFs. In [5] a lightweight RO-PUF is described. In this work [6] are described two methods to extract secret keys from a ring oscillator. The papers [7] and [8] represent delay-based PUFs. The works [9] and [10] use SRAM memory to build PUFs. PUFs can be used to authenticate, generate secret keys and solve the "Night-shift problem". PUFs bring more security and simultaneously reduction of costs. PUFs do not require any permanent storage to save a unique secret information. PUFs are able to generate the unique secret information "on-the-fly". The main disadvantage of PUFs is their output noise. All PUFs produce unstable responses, some of them are very sensitive to the environmental conditions. These errors of outputs may have a random nature or a deterministic nature. Deterministic generated errors are caused by differences of environmental conditions, differences of supply power, ageing of components and so on. Random errors are caused by a peripheral noise. The noisy output of a PUF is a problem for authentication. An error correction code (ECC) can be used to solve this problem. The disadvantage is that an ECC requires a non-volatile memory. However, the PUF for authentication can be implemented without an ECC. In this case the parameters of one PUF must be different enough from the another PUF. Even a very noisy output can be identified correctly.

3 OUR PROPOSAL OF AUTHENTICATION PROTOCOL WITH SECURE TRANSMISSION OF INFORMATION FOR SMART HOMES

Many protocols for authentication of devices involved in IoT were proposed. [11] uses secret sharing scheme to authentication. The authors of [12] use Elliptic Curve Cryptography (ECC) to authentication. [13] uses Fermat Number Transform (FNT) and Chinese Remainder Theorem (CRT) to authenticated communication. The authors of [14] use some implicit certificates to authentication. [15] shows a comparative study on various authentication protocols in WSN. We have proposed a strong unilateral authentication protocol with secure transmission of information between a low-cost device involved in IoT and CU of a smart house. The proposed protocol uses only Hash functions, XOR operations and PUFs. Figure 1 shows the principle of the proposed protocol. On the left side of Figure 1 a concept of wireless communication between low-cost devices involved in IoT and a CU of a smart house is shown. The blue numbers represent the steps in the Figure 1 on the right side. On the right side of Figure 1 an exchange of messages between a low-cost device and the CU during in an authentication process can be seen.

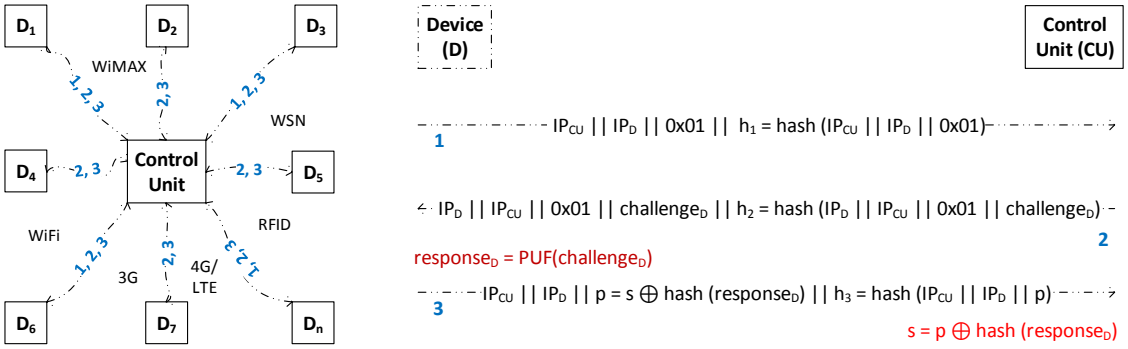


Figure 1: Principle of proposed protocol.

There are low-cost devices $(D_1 - D_n)$ and CU in the proposed protocol. The CU has a database of CRPs of PUFs implemented in low-cost devices $(D_1 - D_n)$. In case a low-cost device wants to send an authenticated secret information to the CU, it will act as follows. The device in the step 1 sends $IP_{CU}||IP_D||0x01||h_1$ to the CU. IP_{CU} represents the IP address of the CU. IP_D represents the IP address of a low-cost device. 0x01 represents a type of a transmitted secret information, which the device will send to the CU. The CU compares the received hash h_1 with its own hash $h_1^{'}$, which is computed by a Hash function with using received $IP_{CU}||IP_D||0x01$. If they are equal, integrity of transmitted data will be guaranteed. Subsequently, the CU in the step 2 sends $IP_D||IP_{CU}||0x01||challenge_D||h_2$ to the device. The device compares the received hash h_2 with its own hash $h_2^{'}$, which is computed by a Hash function using received $IP_D||IP_{CU}||0x01||challenge_D$. If they are equal, integrity of transmitted data will be guaranteed. The device D computes a $response_D = PUF(challenge_D)$ and a $hash(response_D)$. Now the device performs an operation XOR with a secret information and the $hash(response_D)$. The size of the transmitted secret information must be equal to the size of the $hash(response_D)$, in order to the conditions of the Vernam cipher [16] will be satisfied. The size of the output hash function can be equal for example 256 bits. The result of the operation XOR represents a public value $p = s \oplus hash(response_D)$. The device in the step 3 sends $IP_{CU}||IP_D||p||h_3$ to the CU. The CU compares the received hash h_3 with his own hash $h_3^{'}$, which is computed by a Hash function using received $IP_{CU}||IP_D||p$. If they are equal, integrity of transmitted data will be guaranteed. Subsequently, the CU computes $hash(response_D)$ from the corresponding $response_D$ from its database to the $challenge_D$, which CU sent to the device. The CU computes the secret information $s = p \oplus hash(response_D)$ using

the received value p and the $hash(response_D)$. The secret information can represent for example a temperature from a gas boiler. In case the CU starts protocol, the step 1 will be skipped. In the proposed protocol a light-weight variant of the Hash function SHA-3 or the light-weight hash function LOCHA for WSN [17] can be used. The PUF suitable for implementation on low-cost devices is presented for example in [5]. A suitable Error correcting code (ECC) must be included on outputs of the used PUFs. An ECC for output of a PUF is necessary to secure transmission of the secret information. In this case the Hamming distance between the save $response_D$ in the database and the computed $response_D$ must be zero, in order to decryption of the secret information will be successful.

Authentication of the device D to the CU is achieved by CRPs of the PUF, which is implemented in the low-cost device. Integrity of the received data is ensured by Hash functions (hashes $h_1 - h_3$). On inputs of Hash functions all transmitted values are inserted. Randomness of messages is achieved by CRPs, which are different for each PUFs. Unrepeatability is achieved by an assumption that each CRP can be used only once. Non-repudiation of authentication of the device and receiving secret information is achieved by CRPs. The corresponding $response_D$ to the specific $challenge_D$ can create only one specific PUF, which is implemented in one specific low-cost device. Confidentiality of the transmitted secret information is achieved by a Hash function, XOR operation and a PUF. The secret information is XORed with the $response_D$, which is known only to the CU and to the device D, which has implemented the corresponding PUF, which can be used to generate the corresponding $response_D$. The proposed protocol can be implemented in the constrained application protocol (CoAP) [18] or in the 6LowPAN [19].

4 CONCLUSION

In this paper a strong unilateral authentication protocol with secure transmission of information between a low-cost device involved in IoT and Control Unit of a smart house was presented. The protocol uses only PUFs, Hash functions and XOR operations. Our protocol is easy to implement on both ASIC and FPGA. The proposed protocol will be resistant to attacks coming from quantum computers, due does not use PKI, Elliptic curves, hyperelliptic curves, class groups and so on. In our feature work we will implement our protocol on suitable low-cost devices and we will test stability of PUFs outputs against local environmental and voltage fluctuations.

ACKNOWLEDGEMENT

Research described in this paper was financed by the National Sustainability Program under grant LO1401 and by the Czech Science Foundation under grant no. 102/12/1274. For the research, infrastructure of the SIX Center was used.

REFERENCES

[1] Janessa Rivera and Rob van der Meulen. Gartner says the internet of things installed base will grow to 26 billion units by 2020. *Stamford, conn., December*, 12, 2013.

[2] Pappu Srinivasa Ravikanth. *Physical one-way functions*. PhD thesis, Massachusetts Institute of Technology, 2001.

[3] Malcolm Spain, Benjamin Fuller, Kyle Ingols, and Robert Cunningham. Robust keys from physical unclonable functions. In *Hardware-Oriented Security and Trust (HOST), 2014 IEEE International Symposium on*, pages 88–92. IEEE, 2014.

[4] Ulrich Rührmair, Christian Hilgers, Sebastian Urban, Agnes Weiershäuser, Elias Dinter, Brigitte Forster, and Christian Jirauschek. Revisiting optical physical unclonable functions. *IACR Cryptology ePrint Archive*, 2013:215, 2013.

[5] Chaohui Du and Guoqiang Bai. A novel relative frequency based ring oscillator physical unclonable function. In *Computational Science and Engineering (CSE), 2014 IEEE 17th International Conference on*, pages 569–575. IEEE, 2014.

[6] Onur Gunlu, Onurcan Iscan, and Gerhard Kramer. Reliable secret key generation from physical unclonable functions under varying environmental conditions. In *Information Forensics and Security (WIFS), 2015 IEEE International Workshop on*, pages 1–6. IEEE, 2015.

[7] Teng Xu and Miodrag Potkonjak. Stable and secure delay-based physical unclonable functions using device aging. In *Circuits and Systems (ISCAS), 2015 IEEE International Symposium on*, pages 33–36. IEEE, 2015.

[8] Zhi-Hua Xia, Xing-Ming Sun, and Xing-Wei Wang. Techniques for design and implementation of an fpga-specific physical unclonable function. 2016.

[9] Susumu Okumura, Shusuke Yoshimoto, Hitoshi Kawaguchi, and Masahiko Yoshimoto. A physical unclonable function chip exploiting load transistors' variation in sram bitcells. In *Design Automation Conference (ASP-DAC), 2013 18th Asia and South Pacific*, pages 79–80. IEEE, 2013.

[10] Yu Zheng, Maryam S Hashemian, and Swarup Bhunia. Resp: a robust physical unclonable function retrofitted into embedded sram array. In *Design Automation Conference (DAC), 2013 50th ACM/EDAC/IEEE*, pages 1–9. IEEE, 2013.

[11] Omaimah Omar Bamasag and Kamal Youcef-Toumi. Towards continuous authentication in internet of things based on secret sharing scheme. In *Proceedings of the WESS'15: Workshop on Embedded Systems Security*, page 1. ACM, 2015.

[12] Ning Ye, Yan Zhu, Ru-Chuan Wang, Reza Malekian, and Lin Qiao-min. An efficient authentication and access control scheme for perception layer of internet of things. *Applied Mathematics & Information Sciences*, 8(4):1617, 2014.

[13] Manali D Shah, Shrenik N Gala, and Narendra M Shekokar. Lightweight authentication protocol used in wireless sensor network. In *Circuits, Systems, Communication and Information Technology Applications (CSCITA), 2014 International Conference on*, pages 138–143. IEEE, 2014.

[14] Pawani Porambage, Corinna Schmitt, Pranaw Kumar, Andrei Gurtov, and Mika Ylianttila. Two-phase authentication protocol for wireless sensor networks in distributed iot applications. In *Wireless Communications and Networking Conference (WCNC), 2014 IEEE*, pages 2728–2733. IEEE, 2014.

[15] S Raja Rajeswari and V Seenivasagam. Comparative study on various authentication protocols in wireless sensor networks. *The Scientific World Journal*, 2016, 2016.

[16] Gilbert S Vernam. Secret signaling system, July 22 1919. US Patent 1,310,719.

[17] Amrita Roy Chowdhury, Tanusree Chatterjee, and Sipra DasBit. Locha: A light-weight one-way cryptographic hash algorithm for wireless sensor network. *Procedia Computer Science*, 32:497–504, 2014.

[18] Zach Shelby, Klaus Hartke, and Carsten Bormann. The constrained application protocol (coap). 2014.

[19] Zach Shelby and Carsten Bormann. *6LoWPAN: The wireless embedded Internet*, volume 43. John Wiley & Sons, 2011.

Doktorské projekty

Mikroelektronika a technologie

ANALYSIS OF WHIRLING OF GAS FLOW WITHIN THE SPECIMEN CHAMBER OF ESEM

Robert Bayer

Doctoral Degree Programme (1), FEEC BUT

E-mail: xbayer02@stud.feec.vutbr.cz

Supervised by: Jiří Maxa

E-mail: maxa@feec.vutbr.cz

Abstract: The usage of an Environmental scanning electron microscope (ESEM) brings the possibility of studying many kinds of materials and chemical reactions never studied before. In this article there are compared and analyzed several simulation results of gas flow in the specimen chamber of ESEM with high gas pressure throughout the whole volume except on its optical axis, where low gas pressure is needed.

Keywords: ANSYS Fluent, Electron microscopy, Environmental scanning electron microscope, ESEM, Gas whirling, SolidWorks, Specimen chamber, Von Karman vortex

1. INTRODUCTION

Because of high gas pressure in the specimen chamber using Environmental scanning electron microscope (ESEM) we are able to study that kind of samples, we have never been able to before. Without ESEM we had to dry up samples containing water or even air moisture by sublimation process (e.g. biological samples that of course had to be dead) to avoid the water evaporation that could damage the samples or we had to metallize their surface (e.g. insulation materials with high electrical capacity) due to its charging causing electron beam deflection. Using ESEM we are able to study samples containing water, liquid samples, dielectric materials or even chemically active samples, chemical reactions or biological processes with some modifications. [1]

Achieving and keeping such a high gas pressure within the specimen chamber (up to 3 kPa) separate from low gas pressure areas around the electron source (from 1 mPa to 1 nPa, according the electron source type) demands some specific changes in the construction of the microscope and its pumping system itself. The microscope has to be split up by at least two apertures into three or more parts with different gas pressure levels. The apertures have holes in themselves on the optical axis each ten to hundred micrometers large allowing the electron beam to pass through them and to keep high pressure differences between adjacent sections of the microscope with lowest gas pressure on its optical axis. [2]

2. SIMULATION SETTINGS

Whole model of the microscope was built in the SolidWorks program and then imported into ANSYS Fluent, which provides more accurate simulation results then SolidWorks. The block diagram of the microscope including the boundary conditions is in figure 1. In the figure 1 the DP is the diffusion pump, RP 1, RP 2 and RP 3 are rotatory pumps, V1 and V2 are common valves, NV is the needle valve and PLA 1 and PLA 2 are pressure limiting apertures delimiting the differentially pumped chamber as an area between them.

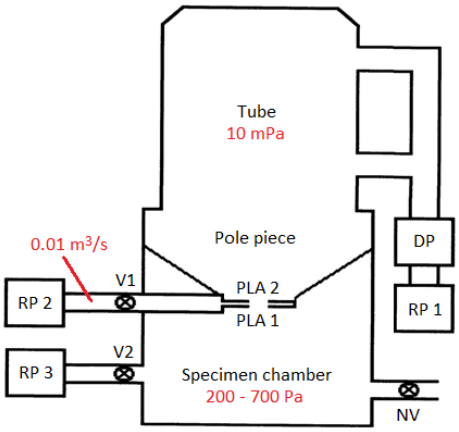


Figure 1: Block diagram of ESEM with boundary conditions [1]

2.1. SOLVER SETTINGS

All calculations were made in Density based mode, which couples the calculations for more accurate results, and with discretization of the second order due to the complexity of calculations of low pressures. The turbulent model used for the simulations was "k-omega SST". For the calculations it was the one most suitable from all the models available in ANSYS Fluent because of its compromise between the calculation time length and its results accuracy.

2.2. BOUNDARY CONDITIONS

In all the simulations used for the analysis the gas pressure within the tube of the microscope was set to 10 mPa and the output flow rate (speed of pumping) of gas from within the differentially pumped chamber was set to 0.01 m^3/s. The only boundary condition that varied in all the simulations was the pressure under which the gas flowed into the specimen chamber and it was set to 200, 300, 400, 500 and 700 Pa.

3. ANALYSIS OF GAS FLOW SIMULATIONS

To speed the simulation process up only the specimen chamber and the first differentially pumped chamber above it were used from the whole microscope model allowing to study the whirling of the gas within the specimen chamber while flowing into the differentially pumped chamber.

Figures 2 to 4 represent gas pressure distribution within the specimen chamber with gas flowing in from the bottom of the chamber under the pressure from 200 Pa to 400 Pa. With such a low entering pressure there is almost no whirling within the gas flow, so it flows around the pedestal with the specimen symmetrically.

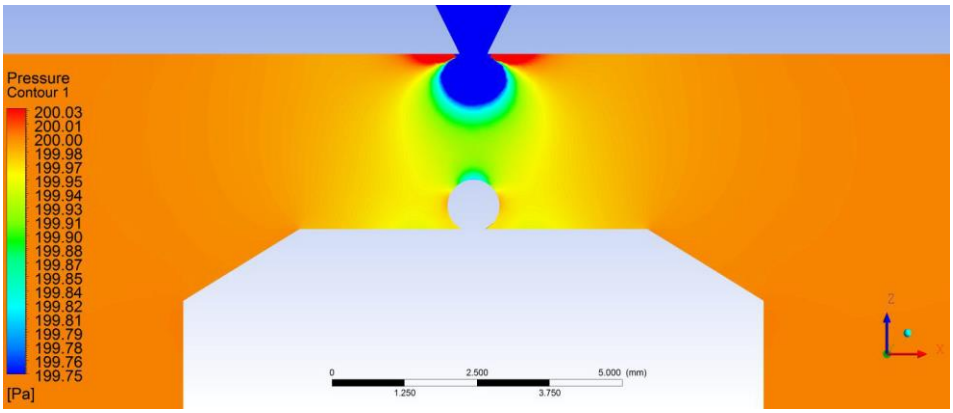


Figure 2: Gas pressure distribution within the specimen chamber at 200 Pa input

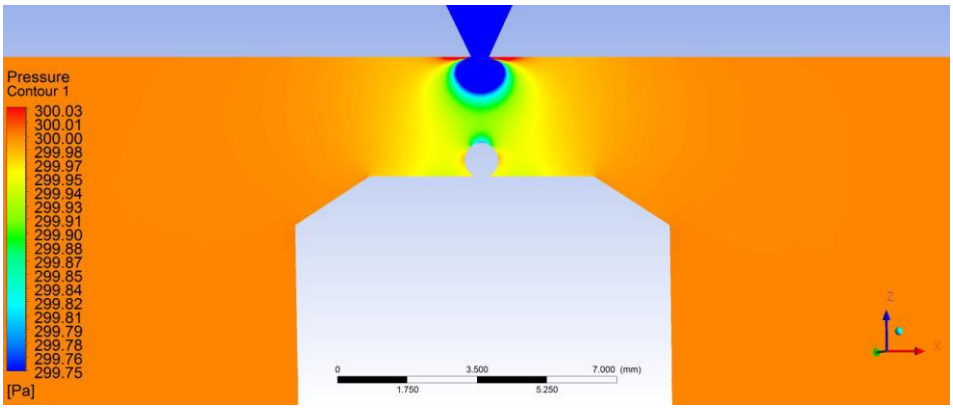


Figure 3: Gas pressure distribution within the specimen chamber at 300 Pa input

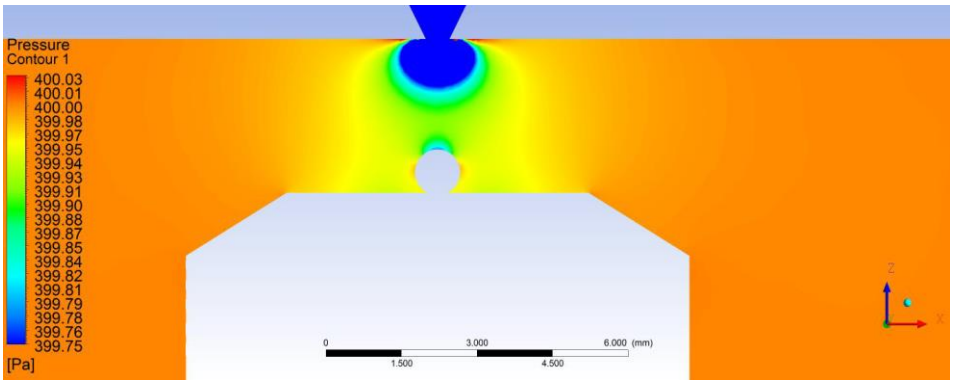


Figure 4: Gas pressure distribution within the specimen chamber at 400 Pa input

However when the pressure of the entering gas rises above 500 Pa little whirling will appear around the sample as shown on figures 5 and 6.

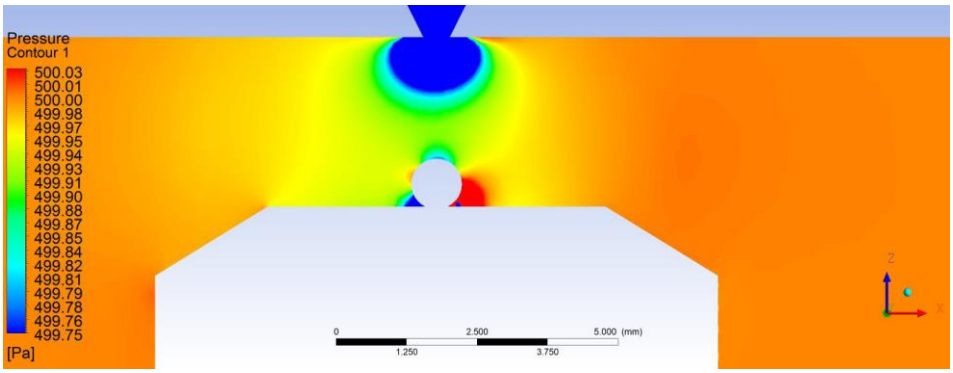


Figure 5: Gas pressure distribution within the specimen chamber at 500 Pa input

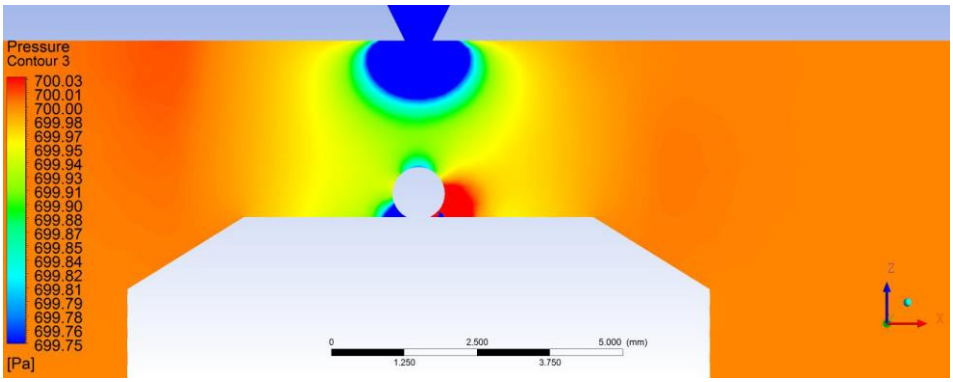


Figure 6: Gas pressure distribution within the specimen chamber at 700 Pa input

These whirling is caused by a chain reaction which starts the moment when the velocity of the gas flow above the pad near the sample rises enough to cause Von Karman vortexes to appear. When these vortexes appear, their path crosses the original path of the gas flow at the other side of the sample right above the edge of the sample pad, which makes a bigger vortex to appear above the pad around the whole sample as shown on figures 7 to 9.

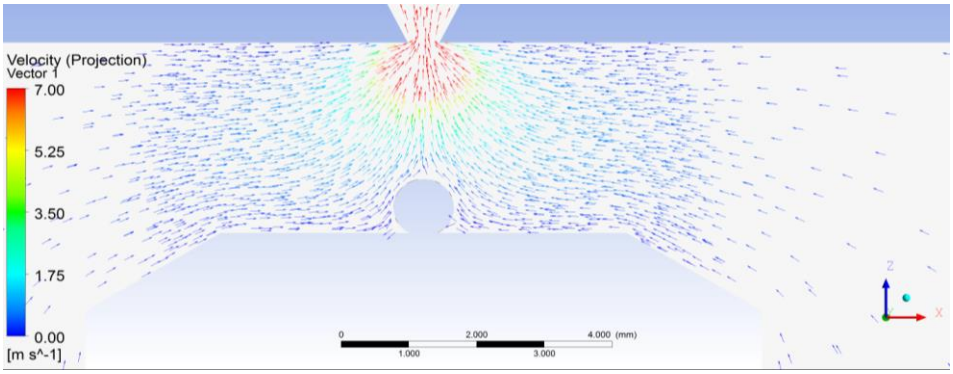


Figure 7: Velocity vectors of the gas flow at 400 Pa input

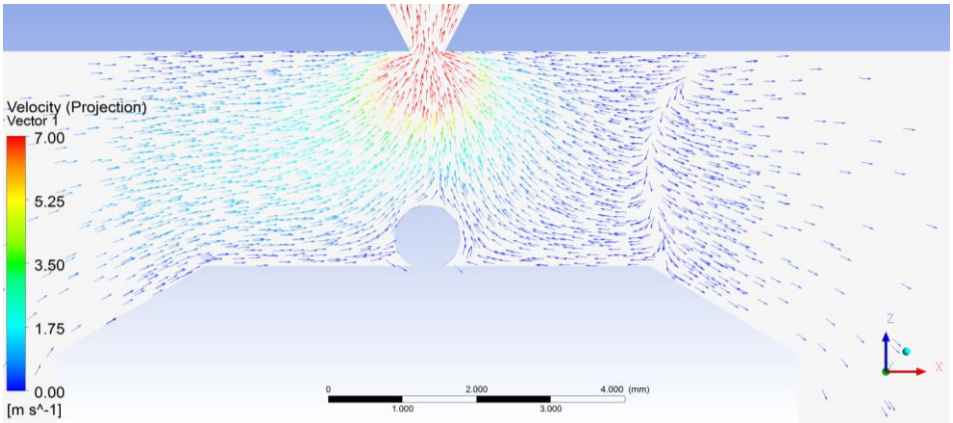


Figure 8: Velocity vectors of the gas flow at 500 Pa input

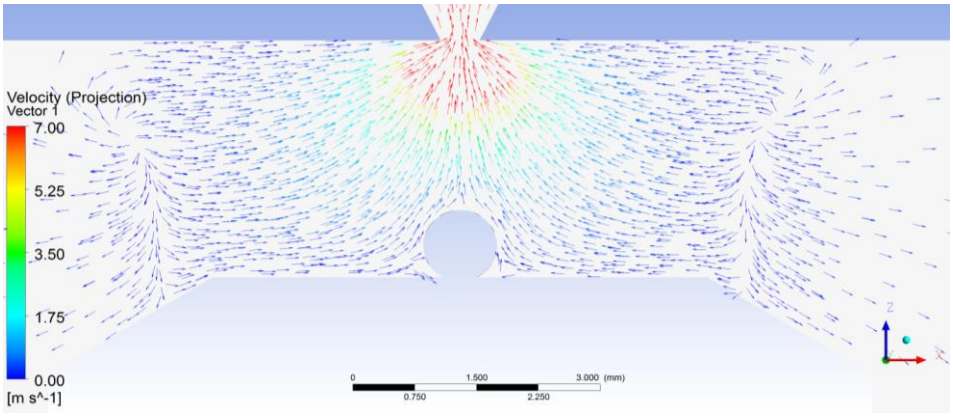


Figure 9: Velocity vectors of the gas flow at 700 Pa input

4. CONCLUSION

The Karman vortexes combined with the big vortex above the pad cause small deflections of gas flow from the optical axis of the microscope that may move the sample on the pad or even to disperse it completely. They may even cause the electron beam to deflect from its original optical axis because of asymmetrical distribution of charged gas particles distorting the observation results.

The process of creation of the vortexes could be suppressed by adding eligibly shaped barriers around the sample, that will give the gas flow the right direction.

ACKNOWLEDGEMENT

This article was created with the help of an internal grant FEKT-S-14-2293 „Materials and technology for electrical engineering II“.

REFERENCES

[1] Neděla, V., Autrata, R. Environmentální rastrovací elektronová mikroskopie, Československý časopis pro fyziku: Czechoslovak journal of physics. Praha: Fyzikální ústav AV ČR, 2005, 2005(3), ^^^sv. ISBN 0009-0700. ISSN československý časopis pro fyziku.

[2] Bayer, R.; Maxa, J.; Vyroubal, P. Analýza proudění plynu v komoře vzorku EREM simulační metodou. In Elektrotechnika a informatika 2015. Plzeň, Tribun EU s.r.o. 2015. p. 3 - 264. ISBN 978-80-261-0514-5.

COMPARISON OF RESULTS OBTAINED USING MONTE CARLO AND ANSYS FLUENT IN ANALYSIS DIFFERENTIALLY PUMPED CHAMBER

Michal Bílek, Pavla Hlavatá
Doctoral Degree Programme (1),FEEC BUT
E-mail: xbilek07@stud.feec.vutbr.cz

Supervised by: Jiří Maxa
E-mail: maxa@feec.vutbr.cz

Abstract: The goal of this thesis is the analysis of a gas flow running through the drain channels within the apertures of differentially pumped chamber of the Environmental Scanning Electron Microscope (ESEM). This thesis contains a verification of current simulation results of gas flow within the differentially pumped chamber published by D. Danilatos using Monte Carlo method in comparison with simulation results achieved by using simulation program ANSYS Fluent, which uses the mechanism of continuum for its calculations.

Keywords: ANYSY Fluent, Environmental Scanning Electron Microscope, ESEM, Monte Carlo

1 INTRODUCTION

This thesis analyses the gas flow within the differentially pumped chamber, which is a part of the Environmental Scanning Electron Microscope (ESEM). This type of microscope is designed for studying samples, which naturally contain water. The purpose of the differentially pumped chamber is to separate the high vacuum area (the microscope's column) from the high pressure area (the sample chamber). This separation is being achieved by two apertures PLA1 and PLA2, preventing the fast equalization of gas pressure throughout the microscope at the time the column is being drained. The microscope column is pre-drained with a rotatory pump and then the gas pressure is stabilized by a diffusion pump. Another rotatory pump is placed into in to the differentially pumped chamber stabilizing the gas pressure at the required figure with continual pumping.

The design and the boundary conditions used for the simulations are the same as G. D. Danilatos [3] published in his calculations, in which he used Monte Carlo method in order to achieve the best pumping results and the minimal loss of the electron beam possible. The Monte Carlo method is based on series of randomly runned expreriments using the model of the system where their evaluation shows the possibility of the specific effect. In this thesis the results obtained using Monte Carlo are compared with the results obtained using the method which deals with the mechanics of the continuum which is also required for calculations of the simulations by the ANSYS Fluent program. The mathematical principles of this method are further described in chapter two. The output of this work is the comparison of the two methods mentioned above and their results for the purpose of deciding, which one is more accurate and could be used for further simulation calculations for the purpose of additional improvements of the construction of the ESEM type microscope.

2 MATHEMATICAL INTERPRETATION OF THE ANSYSFLUENT

The kinetics of the fluids can be described in two ways. The first one is using the Lagrange method which studies the motion of the fluids as individual particles and the second one is using the Euler method which describes the fields of kinetic quantities.

All the calculations were solved with the ANSYS Fluent program. This program solves the system of three partial differential equations, which are extended by an equation of state and the continuity equation. These five equations describe a three dimensional flow of compressible, viscous fluid considering the heat transfer. The Euler method describes the state of the flowing fluid in a specific point. The vector of this point and the time are called the Euler variables. Another term, the "streamline", is connected with the Euler method description of the fluid flow. The streamline is described as a group of points in which the vectors of the fluid flow velocity are tangential at the same moment. This equation can be expressed as (1) :

$$\frac{dx_1}{u_1(x,t)} = \frac{dx_2}{u_2(x,t)} = \frac{dx_3}{u_3(x,t)} \tag{1}$$

where x_i and u_i are components of the position vector which means the velocity in specific point and time, where the time is constant. If the functions are different from zero, they are definite and con-tinuous, so are their first derivations and through every single point of the vector field goes only one single streamline. The streamlines represent the flow in a specific time whereas the trajectories describe the movement of particles in a time interval. [2]

The simulation of a fluid sticks to three basic principles:

- the principle of maintaining the mass (the continuity equation),
- the principle of maintaining the momentum (the Euler equation),
- the principle of maintaining the energy (the Bernoulli equation).

The Continuity equation The continuity equation describes the physical principle of maintaining the mass in the field of mechanics of fluids. The equation states that during the change of velocity, diameter and density the mass of the fluid remains unchanged.

The Euler equation of Hydrodynamics The Euler equation of hydrodynamics describes the equilibrium of forces of mass affecting the fluid from the outside, pressure forces taking effect within the fluid and inertial forces of moving particles of the fluid.

The equation of state The equation of state describes the correction between the pressure, density and temperature of the fluid.

The Bernoulli equation The Bernoulli equation describes the principle of maintaining the mechanical energy of stabilized flow of an ideal fluid.

The Navier – Stokes equations The equilibrium of forces affecting the flow of a real fluid is described by Navier - Stokes equations.

$$\underbrace{\frac{\delta(\rho v_i)}{\delta t} + v_j \frac{\delta(\rho v_i)}{\delta x_j}}_{1} = \underbrace{-\frac{\delta p}{\delta x_i}}_{2} + \underbrace{\eta \frac{\delta^2 v_i}{\delta x_j^2} + \eta \frac{1}{3} \frac{\delta}{\delta x_i} \left(\frac{\delta v_j}{\delta x_j} \right)}_{3} + \underbrace{\rho a_i}_{4} + \underbrace{F_i}_{5} \tag{2}$$

The forces affecting the elementary volume in real fluid flow are:

1. Inertial forces; 2. Pressure forces; 3. Viscous forces; 4. Gravitational forces; 5. Centrifugal forces, electromagnetic forces and others

Where:
ρ – is fluid density $[kgm^{-3}]$; t – is time $[s]$; v – is velocity $[ms^{-1}]$; p – is pressure $[Pa]$; x – is trajectory $[m]$; η – is dynamical viscosity $[Nsm^{-2}]$; a – is acceleration $[ms^{-2}]$; F – is force $[N]$; indexes $_i$ and $_j$ are parts of individual directions of axes (x,y,z) [1]

Atomic Number Density The number of atoms or molecules(n) in a mass (m) of a pure material having atomic or molecular weight (M) is easily computed from the following equation using Avogadro's number ($N_A = 6.022 * 10^{23}$ atoms or molecules per gram-mole):

$$n = \frac{mN_A}{M} \tag{3}$$

In some situations, the atomic number density (N), which is the concentration of atoms or molecules per unit volume (V), is an easier quantity to find when the material density (ρ) is given[4].

$$N = \frac{n}{V} = \frac{\rho N_A}{M} \tag{4}$$

3 CALCULATION SETTINGS

3.1 BOUNDARY CONDITIONS SETTINGS

In the figure 1 is shown the part of the differentially pumped chamber in which is plotted the dimensions of the apertures PLA1 and PLA2. There are also described boundary conditions as the pressure in the specimen chamber, the pressure in the differentially pumped chamber or the pressure in the microscope column.

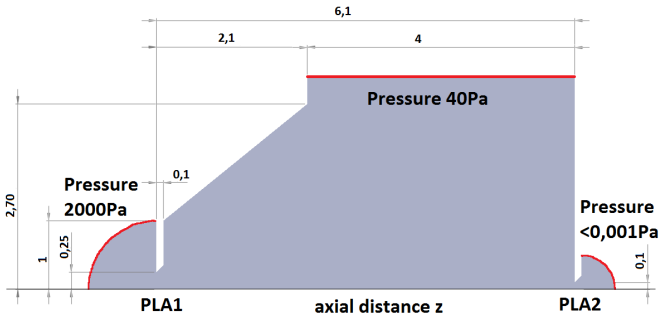


Figure 1: Boundary conditions settings

3.2 SOLVER SETTINGS

It was used Density based settings, which carried out by a coupled calculation pressure, speed, energy and so on. Additionally were used the settings Flux type AUSM, which is better suited for supersonic flow than type Reo-FDS. AUSM stands for Advection Upstream Splitting Method. It is developed as a numerical inviscid flux function for solving a general system of conservation equations. The AUSM first recognizes that the inviscid flux consist in two physically distinct parts, i.e. convective and pressure fluxes. The former is associated with the flow (advection) speed, while the latter with the acoustic speed; or respectively classified as the linear and nonlinear fields. Currently, the convective

and pressure fluxes are formulated using the eigenvalues of the flux Jacobian matrices. Discretization was done in a second order with respect to the complexity of calculation in the area of low pressure and supersonic flow. This required careful preparation of the mash and calculation had to be steering at low values Curant the number of decreased relaxation factors.

4 RESULTS

The comparison of simulation results from Monte Carlo method with the results obtained using the ANSYS Fluent shows both methods having similar results. Comparing the graphical expression of the distribution of Number density shown on figure 2 in the left part, where the results published by D. Danilatos are shown, with the distribution shown on figure 2 in the right part, which were obtained using the system ANSYS Fluent, shows almost identical results including characteristics gradients caused by supersonic flow. The comparisons of values of velocity (on the graph 3 in the left part) and temperature (on the graph 3 in the right part) on the trajectory between the two pressure limiting apertures PLA1 and PLA2, which are also almost overlapping, is also worth attention. The only small difference can be found at the temperature curve showing the area near the aperture PLA2, where the value obtained by the ANSYS Fluent system changes slower. The sharp transition between the values published D. Danilaos may be caused by a singular point. In this case the resulting curve obtained by the ANSYS Fluent system is from its nature physically more.

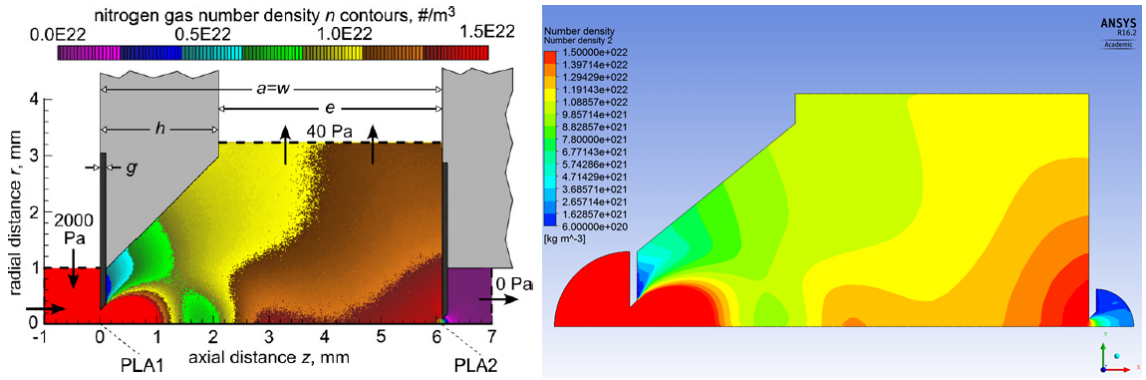


Figure 2: In the left part of inter-aperture flow field with plane geometry PLA2 calculated using the Monte Carlo [3] in the right part calculated using the AnsysFluent

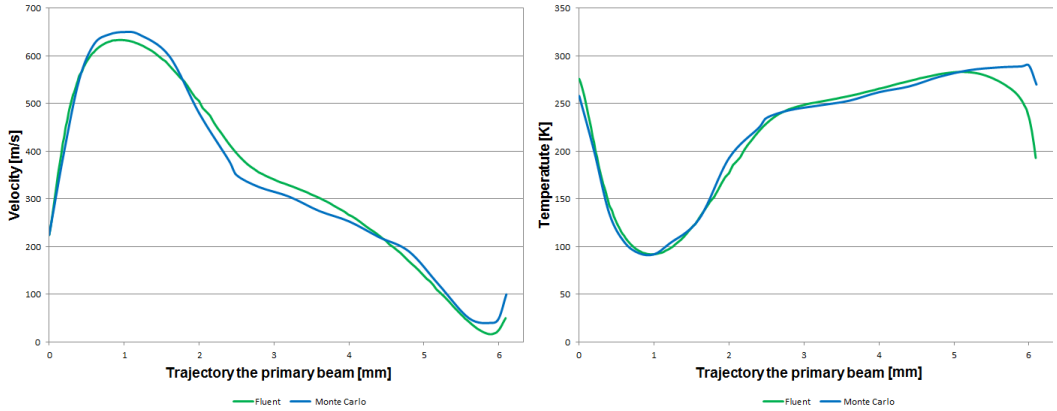


Figure 3: In the left part velocity on the trajectory of the primary beam in the right part temperature on the trajectory of the primary beam

5 CONCLUSIONS

The results published by D. Danilatos, who used for his calculations the Monte Carlo method, are practically identical with the results obtained using the ASYS Fluent system, which uses the method of finite volumes meaning the mechanics of continuum.

Method Monte Carlo does not require the connections between individual subdomains and thus it can describe the effects that are impossible to describe with continuous function, but on the other hand it can also give us wrong results, because it is not conditioned by the continuity condition. The Monte Carlo method is suitable only for the cases, where it is about solving the movement the lone molecules, also called the particulate movement, which has a stochastic character. So it is there, where Navier-Stokes equations are no longer valid. At the opposite, in the cases of a continuous environment is the usage of Navier-Stokes equations definitely an advantage because of its exact description of continuity, because the Navier-Stokes equations are derived from the forces affecting the individual parts of the fluid: gravitation, pressure, friction between adjacent parts of the fluid, formation of the turbulence. The liquids state is described by its velocity and pressures at all points forming the liquid.

Each element of the fluid in a continuum is affected by two forces: the volume one (simple) and the surface one (complicated) which is caused by neighboring elements. What's more the surface force consists of two parts: the first one is related to the pressure and the second one which is more complicated to describe, is describes the shear of the surfaces of individual elements of fluid between each other, which is the viscosity (the internal friction). This complex process of continuum can be precisely evaluated by the Navier-Stokes equations, which compared to the Monte Carlo method do not solve the problem stochastically, but with accurate description of the complex behavior of the fluids. The Navier-Stokes equation is an emblematic example of very complex partial differential equation.

ACKNOWLEDGEMENT

This article was created with the help of internal grant FEKT-S-14-2293 (Materials and technology for electrical engineering II).

REFERENCES

[1] JANALÍK, Jaroslav. *HYDRODYNAMIKA A HYDRODYNAMICKÉ STROJE* [online]. Ostrava, 2008 [cit. 2014-12-02]. 190 p. Available from: http://www.338.vsb.cz/PDF/Janalik-HYDRODYNAMIKAAHYDRODYNAMICKESTROJE.pdf Skripta. VŠB - TU Ostrava, Faculty of Mechanical Engineering.

[2] URUBA, Václav. *Turbulence. 2.*, revised edition In prague: Czech Technical University in Prague, 2014. ISBN 978-80-01-05600-4.

[3] DANILATOS, GD. *Velocity and ejector-jet assisted differential pumping: Novel design stages for environmental SEM.* Micron, 2012, vol. 43, no. 5, p. 600-611.

[4] HOLBERT, K.E. *ATOMIC NUMBER DENSITY* [online]. EEE460-Handout. 2014 [cit. 2016-03-08]. Available from: http://holbert.faculty.asu.edu/eee460/NumberDensity.pdf

ENERGY HARVESTING USING NANOGENERATORS IN IOT

Michal Kerndl

Doctoral Degree Programme (1), FEEC BUT

E-mail: xkernd00@stud.feec.vutbr.cz

Supervised by: Pavel Šteffan

E-mail: steffan@feec.vutbr.cz

Abstract: This paper is a summary and comparison of different types of energy harvesting devices known as nanogenerators which are suitable for wireless sensor network and IoT applications. There is brief introduction of common types of energy harvesting devices in chapter two, than there are actual types of selected nanogenerators described in third chapter. There is a table with comparison of key features of selected nanogenerators in conclusion.

Keywords: EEICT, Energy harvesting, Internet of Things, IoT, Wireless sensor network, Nanogenerator, Piezoelectric nanogenerator, Triboelectric nanogenerator, Pyroelectric nanogenerator

1. INTRODUCTION

The Internet of Things (IoT) is the network of objects embedded with electronics, sensors and software with network connection that allows these objects to collect and exchange data remotely via existing network infrastructure [1]. A wireless sensor network (WSN) is a network formed by nodes equipped with sensor to detect physical phenomena such as light, heat, pressure etc. [2] WSN are autonomous sensors that cooperatively pass their data through the network to a main location [3]. Communication system formed by WSNs improves reliability, efficiency and flexibility of infrastructure systems compared to wired solution. Wireless sensor networks becomes the key technology for IoT [2].

The powering the WSNs becomes common problem due to high flexibility of these systems. The possible solution could be energy harvesting. Energy harvesting is the process when energy is derived from external source such as solar power, thermal energy, wind energy, kinetic energy etc. [4] The device using energy harvesting itself should be low-energy designed to save power, because energy harvesting provides relatively small amount of power. The energy source for harvesters is presented as ambient background and it is free [4]. As the natural reserves are limited, the modern electronic devices should focus on renewable energy sources.

2. COMMON TYPES OF ENERGY HARVESTING DEVICES

The most common types of energy harvesting devices can be divided into two main categories. First are devices that uses ambient sources such as photovoltaic, ambient-radiation, magnetic induction etc. Second category are devices that converts mechanical energy into electricity [5]. These harvesters can use small-scale physical changes and are suitable for example in smart wearables applications and one of these devices are nanogenerators.

3. NANOGENERATOR

Nanogenerator has three typical approaches: piezoelectric, triboelectric and pyroelectric nanogenerators. Both piezoelectric and triboelectric nanogenerators can convert mechanical energy into electricity. However, the pyroelectric can be used to harvest thermal energy from a time-dependent temperature fluctuation [5].

3.1. PIEZOELECTRIC NANOGENERATORS

Piezoelectric nanogenerators are energy harvesting devices that converts external kinetic energy into an electrical energy based on conversion by nano-structured piezoelectic material [5]. It is source of electricity with high effective efficiency using low-cost, non semiconducting, organic nanomaterials are attractive for many applications, including IoT. The article proposed by group led by Prof. Zhong Lin Wang shows first self-powered system with data transition [6]. The structure of nanogenerator can be seen on Fig. 1 and it was made of free cantilever beam that consisted of a five-layer structure: a flexible polymer substrate, ZnO nanowire textured films on its top and bottom surfaces, and electrodes on the surfaces. The measured output voltage reached 10 V, and the output current exceeded 0,6 μA corresponding to power density 10 mW/cm^3. This system transmit wireless signal for distances of 5-10m [6].

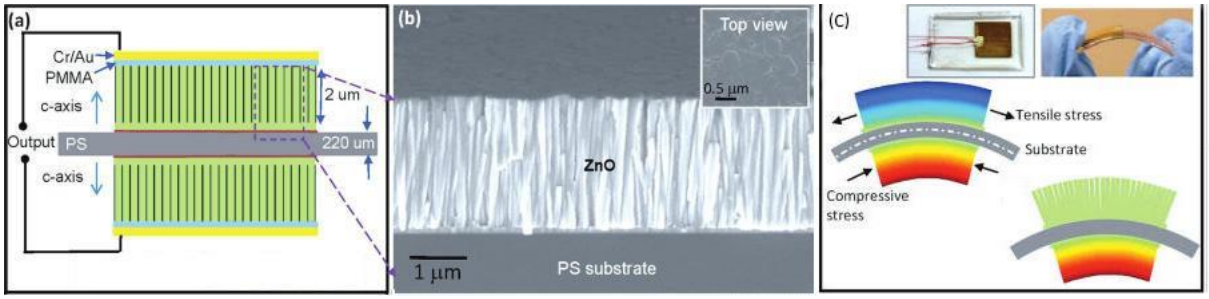


Figure 1: Nanogenerator fabrication and its working mechanism. (a) Fabrication process of the nanogenerator. (b) Cross section SEM image of the as-grown nanowire textured film on the substrate. The insert is the top view of the nanowire film. (c) Simulation of tensile stress and photos of a fabricated nanogenerator after packaging. The bending of the nanogenerator shows its good mechanical flexibility. (Image source: [6])

The piezoelectric nanogenerators have wide variety of usage, it can be used in biomedical applications such as harvester in a pacemaker or in a knee-joint replacement. There is also usage of these nanogenerators in self-powered nano-devices that can transmit data wirelessly. This will enable these tiny devices to operate without battery [6]. There is sample usage in shoe that can be seen on Fig. 2. As user walks the nanogenerator is generating power for wireless sensor that can transmit data without battery.

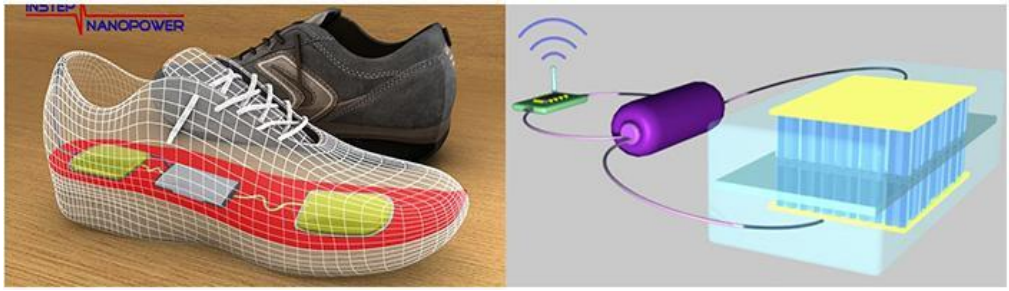


Figure 2: In-shoe device that harvests energy created by walking (on left). The nanodevice consists of a sensor and transmitter (left), a capacitor (middle), and a nanogenerator (right) (Image source: Georgia Tech and [6]) .

3.2. TRIBOELECTRIC NANOGENERATORS

A triboelectric nanogenerator (TENG) is energy harvesting device that converts external mechanical energy into electricity by conjunction of triboelectric effect and electrostatic induction [5]. The potential is created by the triboelectric effect due to charge transfer between two thin organic/inorganic films that exhibit opposite tribo-polarity. Since the most useful material for TENG are organic, it is also named organic nanogenerator, which is first of using organic materials for harvesting mechanical energy [5].

The new type of high-output triboelectric nanogenerator using frequency-multiplication have been proposed by Prof. Zhong Lin Wang's group at Peking University in China. There is novel type of sandwich-shape TENG in their article, to convert low-frequency mechanical energy to electric energy with double frequency [7]. As seen on Fig.3, there is an aluminum film that is placed between two polydimethylsiloxane (PDMS) membranes for frequency multiplication by twice contact electrifications within one cycle of external force. The output peak voltage, current density and energy volume density with this structure was achieved 465 V, 13.4 μA/cm^2 and 53.4 mW/cm^3 [7].

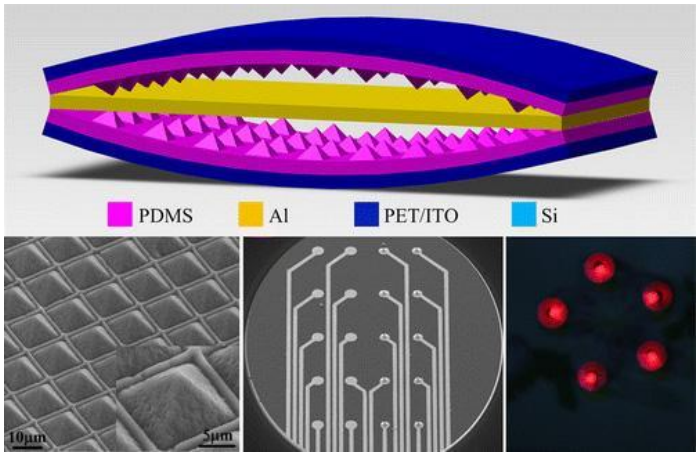


Figure 3: Frequency-Multiplication High-Output Triboelectric Nanogenerator for Sustainably Powering Biomedical Microsystems (Image source: [7])

There is also a silk-fibrion based transparent triboelectric generator which is suitable for autonomous sensor network proposed in article by Xiao-Sheng Zhang, J. Brugger and B. Kim. The novel material for this TENG is silk fibron that possesses the outstanding ability of losing electrons easily [8]. For instance, an oxygen plasma process is employed to improve interfacial energy between silk fibron film and substrate by chemical modification and nanoscale roughness increase, which allows for the reliable switching mechanism needed to generate the electric power [8].

The proposed TENG exhibits stable high-output performance with maximum voltage, current, and power density of 268 V, 5.78 μA/cm^2, and 193.6 μW/cm^2 [8]. There is one outstanding advantage of this TENG, it is higher than 90% transmittance in the visible spectrum which can be suitable for high variety of biomedical applications as well as fully integrated transducers that are needed for autonomous sensor networks [8].

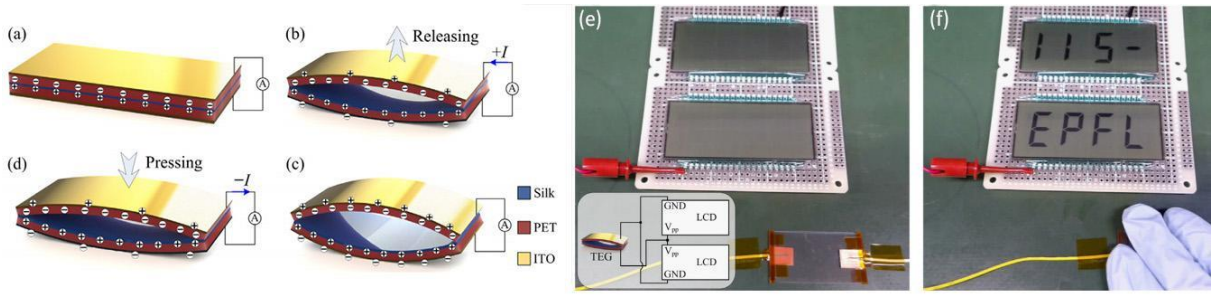


Figure 5: Schematic view of the operating principle of the silk-based TENG: (a) the triboelectric pairs of silk fibroin film and PET film make contact with each other under the external force and charges are generated; (a)-(b)-(c) on removing the external force, the TENG reverts to the original arch shape under the mechanical restoring force - a process called the releasing step; (c)-(d)-(a) following application of the external force, the TENG deforms from the original arch shape to the contact state - a process called the pressing step. (e) Photograph of the silk-based TENG, and (f) two 4-bit liquid crystal displays driven by the silk based TENG without any external circuit (Image: [8])

3.3. PYROELECTRIC NANOGENERATORS

A pyroelectric nanogenerator (PENG) is an energy harvesting device converting the external thermal energy into an electrical energy by using nano-structured pyroelectric materials. Usually, harvesting thermoelectric energy mainly relies on the Seebeck effect, however, in an enviroment that the temperature is spatially uniform without a gradient, the pyroelectric effect has to be choice due to spontaneous polarization [5]. The latest article by Prof. Zhong Lin Wang's group demonstrates a PENG based on lead zirconate titanate (PZT) film, which provides the output open-circuit voltage and short-circuit current density of 22 V and 171 nA/cm^2, respectively corresponding to maximum power density of 0,215 mW/ cm^3 [9]. This PENG was used to drive liquid crystal display for longer than 60 s, using a Li-Ion battery that was charged by this PENG at a different working frequencies. This PENG and Li-Ion battery have been used to power wireless temperature sensor successfully with working distance larger than 50 m [9], as seen on Fig. 6.

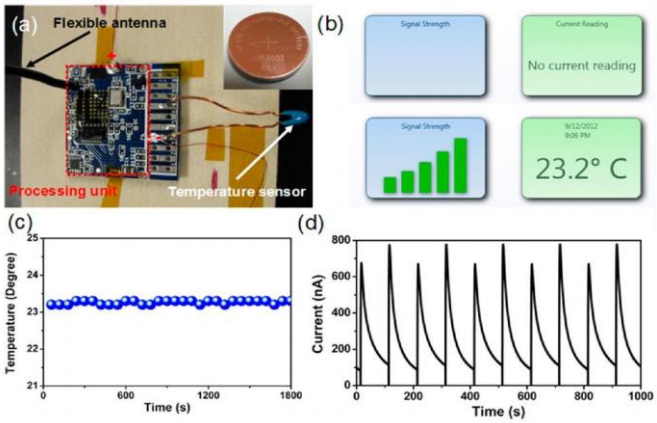


Figure 6: Wireless temperature sensor. (b) The obtained temperature signals when the wireless temperature sensor was turned off and on. (c) The obtained temperature data in half an hour, where the temperature value was recorded in every 1 min. (d) Output current of the PENG. (Image source: [9])

4. CONCLUSIONS

There are three basic categories of nanogenerators described in this paper. Each of this category have its specific advantages and are suitable for specific applications. Piezoelectric nanogenerators have wide variety of usage for example in medical, IoT or environmental applications. There is vertical bending model described in this article, which is one of three major models. The other models are lateral stretching and lateral blending nanogenerators. Second category is triboelectric nanogenerators which have the most outstanding output performance, especially the frequency-multiplication sandwich structure. Second described TENG based on silk fibron have outstanding transmittance which is 90% in visible spectrum. The key fact that its highly transparent make this nanogenerator suitable for implantable biomedical applications and future self-powered wearable microsystems. Last category is pyroelectric nanogenerators which is mainly used as active sensors.

Type	V_{out} [V]	Current density [A/m^2]	Power density [W/m^3]
ZnO nanowire	10	$0,6*10^{-6}$	$10*10^{-3}$
TENG - PDMS based	465	$13,4*10^{-6}$	$53,4*10^{-3}$
TENG - silk fibron	268	$5,78*10^{-6}$	$193,6*10^{-6}$
PENG - PZT film	22	$171*10^{-9}$	$215*10^{-6}$

Table 1: Comparison of max. output open-circuit voltage, max. short-circuit current density and max. power density of nanogenerators described in this article. (Data source: [6], [7], [8], [9])

ACKNOWLEDGEMENT

The work was supported by the Brno University of Technology project no. FEKT-S-14-2300: "New types of electronic circuits and sensors for specific applications".

REFERENCES

[1] *Internet of Things*, [online]. 2016. San Francisco (CA): Wikimedia Foundation. [Accessed 10 March 2016]. Retrieved from: https://en.wikipedia.org/wiki/Internet_of_Things

[2] YINBIAO, Shu a Wang QIN. Internet of Things: Wireless Sensor Networks. *International Electrotechnical Commission.* Geneva, Switzerland, 1-78. ISSN 978-2-8322-1834-1.

[3] *Wireless sensor network*, [online], 2016. San Francisco (CA): Wikimedia Foundation. [Accessed 10 March 2016]. Retrieved from: https://en.wikipedia.org/wiki/Wireless_sensor_network

[4] *Energy harvesting*, [online], 2016. San Francisco (CA): Wikimedia Foundation [Accessed 10 March 2016]. Retrieved from: https://en.wikipedia.org/wiki/Energy_harvesting

[5] *Nanogenerator*, [online], 2016. San Francisco (CA): Wikimedia Foundation [Accessed 10 March 2016]. Retrieved from: https://en.wikipedia.org/wiki/Nanogenerator

[6] HU, Youfan, ZHANG, Yan, XU, Chen, LIN, Long, SNYDER, Robert L. and WANG, Zhong Lin, 2011, Self-Powered System with Wireless Data Transmission. *Nano Letters* [online]. 08 June 2011. Vol. 11, no. 6p. 2572-2577. DOI 10.1021/nl201505c. Retrieved from: http://pubs.acs.org/doi/abs/10.1021/nl201505c

[7] ZHANG, Xiao-Sheng, HAN, Meng-Di, WANG, Ren-Xin, ZHU, Fu-Yun, LI, Zhi-Hong, WANG, Wei and ZHANG, Hai-Xia, 2013, Frequency-Multiplication High-Output Triboelectric Nanogenerator for Sustainably Powering Biomedical Microsystems. *Nano Letters* [online]. 13 March 2013. Vol. 13, no. 3p. 1168-1172. DOI 10.1021/nl3045684. Retrieved from: http://pubs.acs.org/doi/abs/10.1021/nl3045684

[8] ZZHANG, Xiao-Sheng, BRUGGER, Jürgen and KIM, Beomjoon, 2016, A silk-fibroin-based transparent triboelectric generator suitable for autonomous sensor network. Nano *Energy* [online]. 2016. Vol. 20, p. 37-47. DOI 10.1016/j.nanoen.2015.11.036. Retrieved from: http://linkinghub.elsevier.com/retrieve/pii/S2211285515004723

[9] YANG, Ya, WANG, Sihong, ZHANG, Yan and WANG, Zhong Lin, 2012, *Nano Letters* [online]. 12 December 2012. Vol. 12, no. 12. DOI 10.1021/nl303755m. Retrieved from: http://pubs.acs.org/doi/abs/10.1021/nl303755m

ENHANCED METHOD OF THROUGH-HOLE COPPER PLATING

Alexandr Otáhal, Adam Crha, Václav Šimek

Doctoral Degree Programme (4), FEEC BUT

E-mail: xotaha00@stud.feec.vutbr.cz

Supervised by: Ivan Szendiuch, Richard Růžička

E-mail: szend@feec.vutbr.cz, ruzicka@fit.vutbr.cz

Abstract: This paper is dealing with an introduction of novel approach conceived towards the optimization of widespread through-hole copper plating process. Main objective in this context is to achieve significantly increased process reliability and overall fabrication yield also for the small-scale production runs or prototyping purposes of printed circuit boards. Essential aspect of the procedure outlined in this paper is inherently connected with the activation of printed circuit board surface, which serves as the basis for the subsequent step that involves the conformal copper plating of holes drilled through the substrate compound. An innovative part of the whole process can be recognized in activation of drilled holes surface by means of employing combination of vacuum (low pressure environment) and ultrasound within the dedicated setup. The application of the suggested procedure helps to considerably reduce the failure ratio (i.e. the presence of holes with faulty surface plating) even in case of drilled holes at considerably small dimensions and, as a direct consequence, improves the overall process reliability.

Keywords: Through-hole technology, copper plating, PCB

1. INTRODUCTION

Nowadays, there exists a large variety of methods that are suitable for the purpose of through-hole copper plating in case of the high-volume fabrication flow, as well as in the prototyping domain. In fact, the actual copper plating procedure variant is typically selected with respect to a given application scenario, and it may involve e.g. direct current plating, pulse plating or reverse pulse plating. In case of lower production volumes and smaller holes diameters pulse or reverse pulse plating [1, 2] is typically chosen. Switchable reverse pulse plating technique is considered to deliver more uniform copper plating for difficult aspect ratios and smaller holes. Reverse pulse plating method is also very useful for small holes on high-density PCBs and improves production reliability during the soldering phase as well. Reverse pulse plating uses precisely controlled reverse pulses to remove excess material during the plating process [3].

The combined influence of diverse parameters makes the through-hole plating process rather complex task, which is more than symptomatic for low-volume or prototyping conditions. In that way especially the problem of variability behind the commonly used processes comes into attention as one of the key hurdles preventing satisfactory results and stable yield. One of the significant issues is clearly associated with the preparation of holes surface before the actual copper plating takes place. Activation of surface can be performed by means of using a few special chemical solutions (for example solution of copper sulfate, calcium hypophosphite and ammonium hydroxide [4]), but the principle is straightforward - homogenously activate and finally cover surface of drilled holes.

Main goal behind this paper is focused on possible mitigation of at least some of the aforementioned challenges through the introduction of an innovative method for copper plating, where the attention is predominantly given to the aspects of sample surface preparation for the following process stages. The proposed method is based on using reverse pulse plating feature at its core while

the main innovation dwells in usage of lower ambient pressure during an activation of the surface, which should help to expel the formation of air bubbles from the holes openings that prevent the activation liquid from even penetration into the confined holes volume. It is obvious indeed that the careful surface treatment and preparation of the drilled holes plays an important role within the effort to ensure the reliability of copper plating technique with the aid of an electrical energy. Furthermore, this process step is crucial in terms of the formation and uniformity control of the conformal layers that allow adhesion and succeeding growth of the copper layer [3].

2. METHODOLOGY

The purpose of this chapter is to provide closer insight into the methods, equipment and materials, which were used throughout the implementation and subsequent verification of the innovative approach to the widespread through-hole plating technique.

2.1. INNOVATIVE METHODS IN THROUGH-HOLE PLATING

As this observation was already mentioned within the introductory chapter of this paper, the introduction of additional techniques may bring significant benefits for the activation of non-conducting surface (drilled holes). One of the innovations behind the standard process is the intentional addition of lower pressure for better propagation of activator liquid into drilled holes of smaller dimensions. It is well known fact that already low vacuum helps to notably decrease the surface tension of liquids and causes expansion of gasses. From a technical perspective it simply allows the air confined into the typically small volume of holes to go outside and lets the activator solution to penetrate inside along the whole drill path [5].

Other promising technological concept is building upon the exploitation of ultrasound transfer to the test vehicle. An ultrasonic wave is reflected when it strikes an interface between materials with different speeds of sound (acoustic impedance). Furthermore, an interface between materials with a larger difference in acoustic impedance reflects ultrasonic waves with a higher force and that with a smaller difference in acoustic impedance reflects them less strongly and lets part of them travel through [6].

2.2. TEST VEHICLE

Test vehicle used for the experimentation can be seen on the Fig. 1. Dimensions of test coupon are approximately 70 x 13 mm. Through holes are represented by green circles. Test pads are fabricated on the both sides of the PCB and they are connected to the VIAs. Holes diameter is in the range 0,2 – 1,5 mm. Holes with different diameters are positioned uniformly for better control of plating process. Two rows of holes are mirrored in sequence to each other. FR4 substrate was used as a base material with both sides plated with 18 μm of copper.

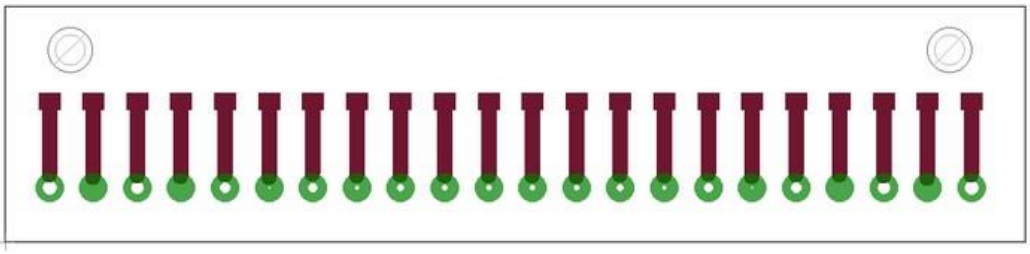


Figure 1: Test vehicle

2.3. EQUIPMENT

Copper plating facility was available as a LPKF MiniContac RS unit with closed circulation of individual baths content (Fig. 2). This device is using the principle of reverse pulse method for cop-

per plating or deposition. That method, reverse pulse copper plating, provides sufficient parameters for conformal coating of copper during plating.



Figure 2: MiniContac RS

Substantial innovation in this standard process is represented by the creation of vacuum chamber (Fig. 3) for activator solution, which enables reliable activation of holes with smaller diameter than before. Chamber was made from special chemical resistant plastic compound. This chamber was also used for the activation during the simultaneous exploitation of ultrasonic waves generator.

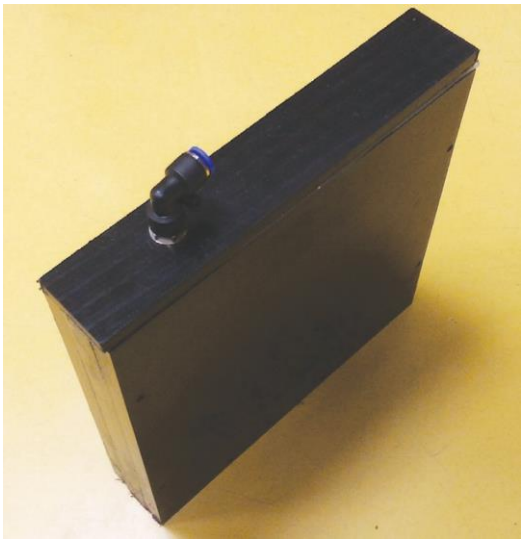


Figure 3: Additional equipment – vacuum chamber

For investigation of the plated holes reliability the current overload test method was chosen. Plated holes were be tested for their structural integrity once exposed to the current reaching the level of 1 amp. On the Fig. 4 can be seen the fixture for current load measurement. Pictures on the right top corner are providing closer view on the measurement pins. It can be seen that two pins are placed from top side and two from bottom. This particular arrangement is due to the selected usage of pins for four-point method (two pins for current, two pins for voltage). Between top and bottom pins there are also placed the measured pattern of test vehicle. During the test procedure the individual measurements were focused on the evaluation of voltage/resistance of the plated holes.

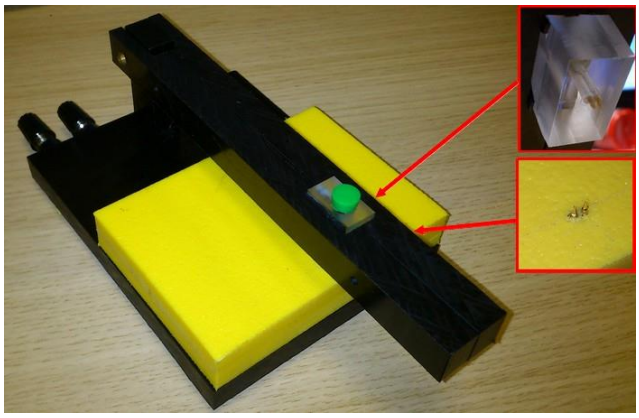


Figure 4: Fixture for current load measurement

3. RESULTS

Samples were fabricated by means of using the standard method (reference sample) and optimized method based on utilization of vacuum chamber (sample fabricated with using vacuum).

First analysis comprised electrical testing of resistivity. There were seen two cases of resistivity, resistivity under 0,01 Ω (plated holes) and very high resistivity (unplated holes). Defects were identified mainly for 0,2 mm diameter holes. Approximately 30 % of these small holes were unplated.

Second analysis was conducted with the assistance of X-Ray inspection equipment. Results are presented on the figure 5 and figure 6. The former of them is showing unplated through-hole of reference sample. On the contrary, figure 6 depicts plated hole of sample fabricated with the usage of vakuum-equipped chamber.

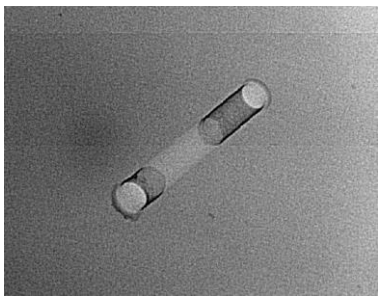


Figure 5: Reference sample – unplated hole

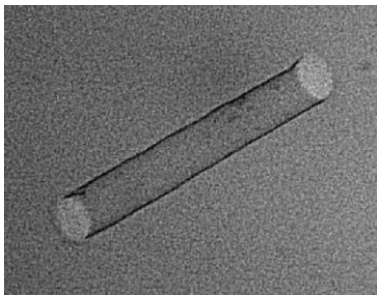


Figure 6: Vacuum activated sample –plated hole

Microsections show the same behavior as X-ray figures. On the figure 7 is shown microsection of reference sample and on the figure 8 is shown microsection of vacuum activated sample.



Figure 7: Microsection - reference sample



Figure 8: Microsection - vacuum activated sample

4. CONCLUSION

Copper plating methods are well-known and field-proven technologies across the electronic fabrication industry. Settings of the process parameters are depends on quality/purity of solvents and knowledge/responsibility of operator. Therefore using of vacuum or ultrasound during activation process can bring new enhanced process of through-hole plating. Reliability of copper plated holes activated with the parallel assistance of vacuum and ultrasound has shown notable benefits in comparison with the reference sample, which was fabricated using the conventional approach.

In summary, this paper contains description of the first results obtained during the avaluation of the innovated process of copper plating. Further research and development activities will be focused on the reliability (stress test, microsections, etc.) improvement and overall process refinement.

ACKNOWLEDGEMENT

Funding for this research work was supported through a grant project FEKT/FIT-J-15-2832. Thanks to Sanmina – SCI Czech Republic for opportunity use their X-Ray inspection system.

REFERENCES

[1] Fang-Yu Shen, Wei-Ping Dow, An-Hong Liu, Jing-Yuan Lin, Ping-He Chang, and Su-Mei Huang. Periodic Pulse Reverse Cu Plating for Through-Hole Filling. ECS Electrochem. Hawaii, Meeting of the Society, October 7–12, 2012. doi:10.1149/2.003305eel. Web: http://eel.ecsdl.org/content/2/5/D23.abstract

[2] Baker Technology Associates Inc. Why Pulse Plating? [online]. [cit. 2015-10-06]. Web: http://www.bakertech.com/pulse.htm

[3] LPKF Laser&Electronics. [online]. [cit. 2015-10-06]. Web: http://www.lpkfusa.com/RapidPCB/ThroughHolePlating/minicontac_rs.htm

[4] instructables. Industrial-level quality PCB through hole plating [online]. [cit. 2015-10-06]. Web: http://www.instructables.com/id/Inexpensive-method-of-industrial-level-quality-PCB/

[5] John F. O'Hanlon. A user's guide to vacuum technology. 1980. Wiley-Interscience p.536. ISBN978-0471270522

[6] NIHON DEMPA KOGYO CO., LTD. Ultrasound and its properties [online]. [cit. 2015-10-06]. © 1997- 2015. Web: http://www.ndk.com/en/sensor/ultrasonic/basic01.html

THE INFLUENCE OF THE BINDER ON THE FINAL CAPACITY OF THE NEGATIVE ELECTRODE MATERIAL CARBON CR5995 FOR LITHIUM-ION BATTERIES

Tomáš Gottwald

Doctoral Degree Programme (2.), FEEC BUT

E-mail: xgottw03@stud.feec.vutbr.cz

Supervised by: Marie Sedlaříková

E-mail: sedlara@feec.vutbr.cz

Abstract: This article deals with the issues of using binders for lithium-ion electrodes and optimizing the technological process for their production. The work focuses on optimizing the amount of binder used in order to achieve the most capacity. The methodology consists of creating an appropriate experimental electrode and subsequent measuring its associated capacity with BioLogic. This approach leads to creating graphs which show the dependencies of capacity on the amount of binder used.

Keywords: Binder, carbon, electrode, lithium battery, PVDF

1. INTRODUCTION

Electrode materials are normally supplied in powder form but for the use in the cell they must be coated on the electrode current collector foils, usually aluminium and / or copper, in a homogenous layer to form a conductive mass with a precisely controlled thickness. This is the job of the binder materials. They are non-conductive materials dissolved in a temporary solvent to form a carrier in which the electrode powder is mixed to form a slurry, which is spread over the foil current collectors enabling the electrode powder to adhere to the foil. The viscosity of the slurry must be tightly controlled to ensure consistent and event coating of the electrode layer. Subsequent passage of the coated foils through heating oven drives off the solvents leaving a solid coating of the electrode material on the foils with the binder providing a conductive path between the electrode particles. Changing the composition of the binders can affect the viscosity of the slurry, the adhesion of the electrode to the foil, the conductivity of the cells as well as the self discharge. [1]

The solvents used in the binders, non-active solvents may also be necessary to carry the electrolyte. In order to get the optimum performance from these solvents it may be necessary to resort to the use of toxic materials. N-methyl pyrrolidine (NMP), which has been shown to cause birth defects when pregnant women have been exposed to do it, is one such solvent used by the majority battery manufactures in applying a coating on the electrodes of lithium batteries. [1]

Polyvinylidene fluoride, or polyvinylidene difluoride (PVDF) is a high non-reactive and pure thermoplastic fluoropolymer produced by the polymerization of vinylidene difluoride. PVDF is the standard binder material used in the production of composite electrodes for lithium ion batteries. 1–2 % weight solution of PVDF dissolved in N-methyl-2-pyrrolidone (NMP) is mixed with an active lithium storage material such as graphite, silicon, tin, $LiCoO_2$, $LiMn_2O_4$, or $LiFePO_4$ and a conductive additive such as carbon black or carbon nanofibers. This slurry is cast onto a metallic current collector and the NMP is evaporated to form a composite or paste electrode. PVDF is used because it is chemically inert both to electrolyte and lithium metal. [2]

2. EXPERIMENTAL

Carbon CR5995 electrode material thin layers with different weight of PVDF binder on copper foil were prepared. These electrodes were measured against the metal lithium and assembled in EL-CELL®. Was used 1 M electrolyte $LiPF_6$ dissolved in 1:1 EC:DMC solvent was used.



Figure 1: Schema of assembled EL-CELL®. The first layer is thin copper foil with carbon and binder (electrode material), the second layer is separator and the third layer is Lithium metal form. The EL-CELL® is made of steel.

2.1. TECHNOLOGY PROCESS OF ELECTRODE PREPARATION

160 mg of graphite carbon CR5995 was put in to the glass vial together with binder polyvinylidene fluoride (PVDF). This electrode material was dissolved with organic compound N-Methyl-2-pyrrolidone (NMP). Glass vial was placed on electromagnetic stirrer and electrode material was stirred for 24 hours at room temperature for good homogenization of carbon electrode particles with the binder. Subsequently, thin layer (200 μm) of prepared electrode material was deposited on thin copper foil and inserted into the dryer overnight at 50 °C (122 °F). We need to evaporate the water and the solvent. Electrodes were cut out with a diameter of 18 mm and used as working (negative) electrode. The electrode was dried in vacuum overnight at room temperature.

2.2. MORPHOLOGY OBSERVATION BY SEM

Were investigated two electrodes of carbon CR5995 + 6 % PVDF. Morphology after the cycling (b) has shown slightly larger gaps between the particles of carbon. This may be the reason for the loss of capacity during cycling of the electrode material.

a **b**

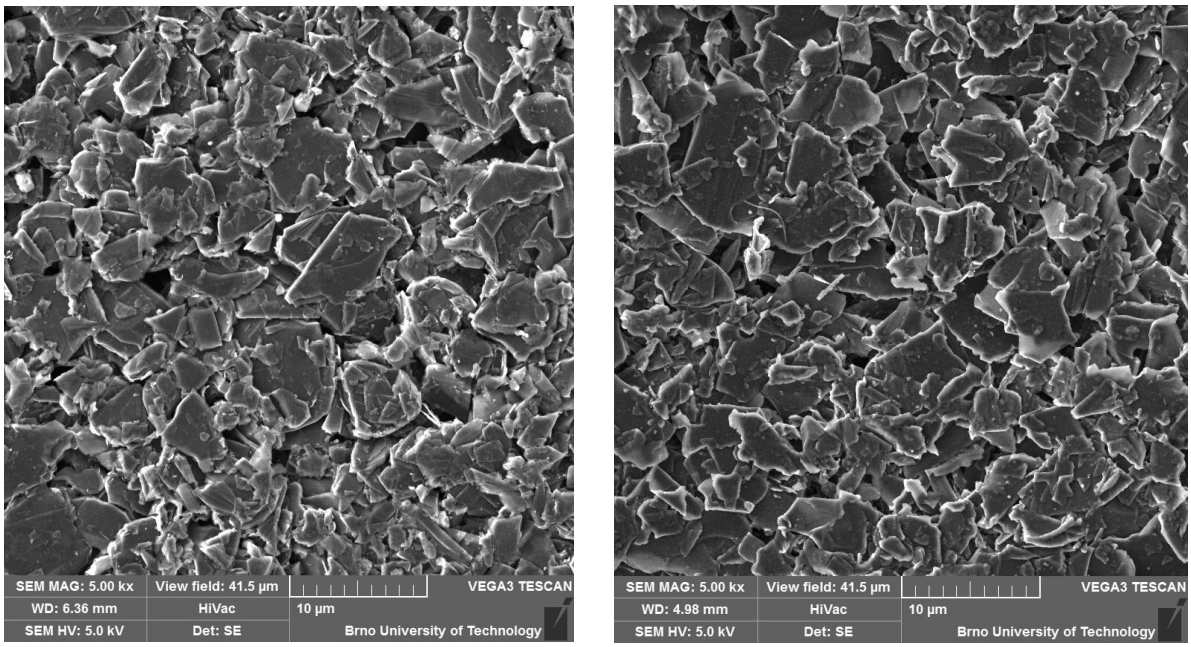


Figure 2: High-resolution SEM analysis of carbon CR5995 + PVDF binder. High-resolution SEM images showing layered structure for (**a**) CR5995 + 6 % PVDF before cycling and (**b**) CR5995 + 6 % PVDF after cycling.

2.3. RESULTS

Electrode material CR5995 + 4 % PVDF surface area was 2,550 cm^2 and characteristic mass was 3,1 mg. Specific capacity for the second cycle was 307 $mAh.g^{-1}$ and cycling stability was excellent. Electrode material CR5995 + 6 % PVDF surface area was 2,550 cm^2 and characteristic mass was 2,1 mg. Specific capacity for the second cycle was 342 $mAh.g^{-1}$ and cycling stability stability was worse from the sixth cycle. Electrode material CR5995 + 10 % PVDF surface area was 2,550 cm^2 and characteristic mass was 2,1 mg. Specific capacity for the second cycle was 251 $mAh.g^{-1}$ and cycling stability stability was worse from the beginning of cycling.

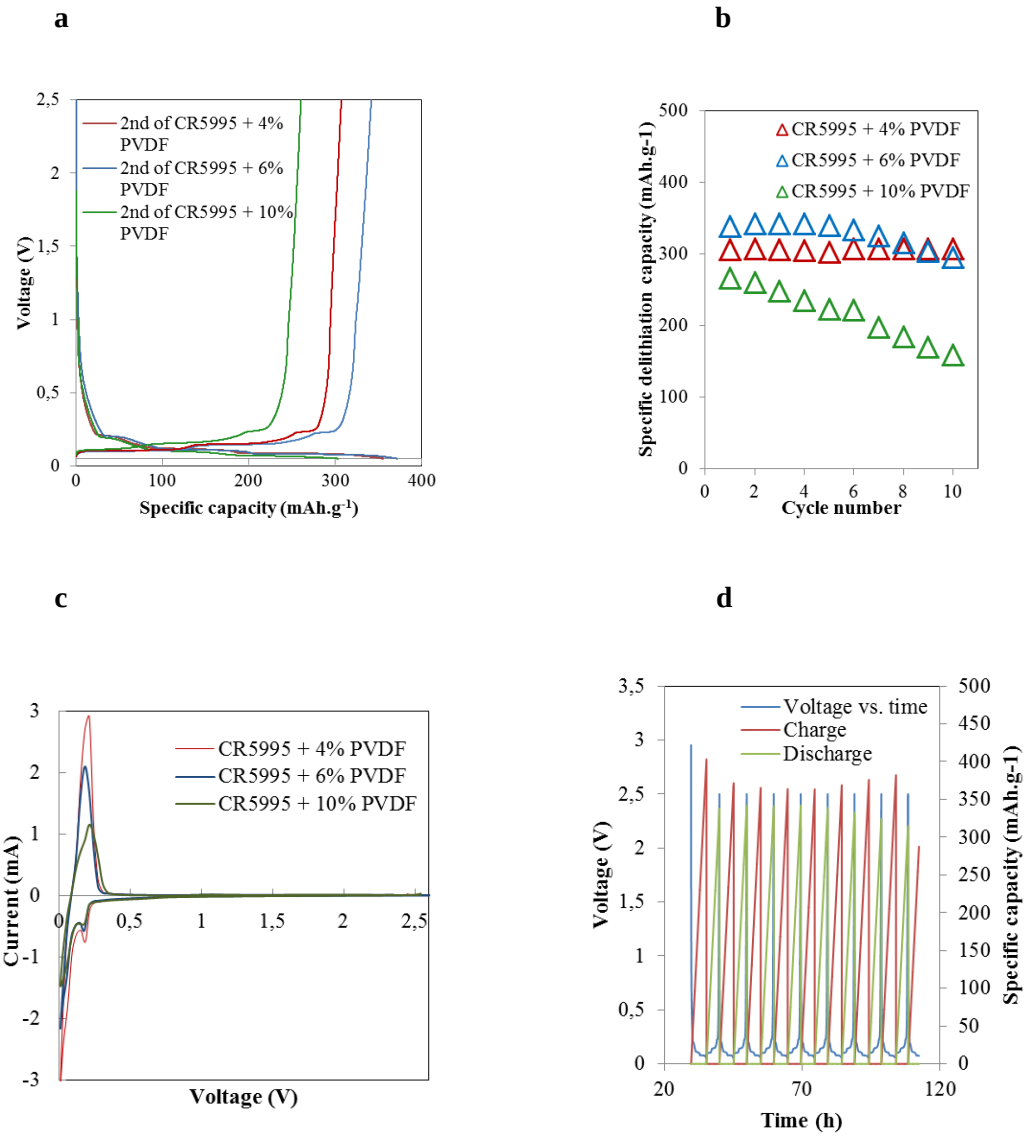


Figure 3: Electrochemical performances. (a) Charge/discharge curves for the second cycles of CR5995 + PVDF. (**b**) Short-term cycling stability test for CR5995 + PVDF. (**c**) Cyclic voltametry of CR5995 + PVDF at a scan rate 0,1 $mV.s^{-1}$. (**d**) Galvanostatic cycling with potential limitation for CR5995 + 4 % PVDF, the current was 159 µA.

3. CONCLUSION

The short-term cycling stability of CR5995 was investigated with a 4 % PVDF for the first 10 cycles to achieve a full lithiation capacity 307 $mAh.g^{-1}$. CR5995 + 4 % PVDF demonstrate excellent cycling stability for 10 cycles with a very low capacity decay rate of ~ 0.003 % per cycle from the 1st cycle to 10th cycle. CR5995 + 6 % PVDF provides a capacity ~ 342 $mAh.g^{-1}$ for the second cycle, as shown in Fig. 3a. Even though the capacity of CR5995 + 4 % PVDF is lower then the reported carbon CR5995 + 6 % PVDF, the cycling stability of CR5995 + 4 % PVDF is much better.

ACKNOWLEDGEMENT

This work was supported by the grant FEKT-S-14-2293 "Materiály a technologie pro elektrotechniku II" and Centre for Research and Utilization of Renewable Energy under project No. LO1210 - "Energy for Sustainable Development (EN-PUR)".

REFERENCES

[1] *Batteries and Energy Storage Technology magazine: Breakthroughs don't always make better batteries.* Barrie LAWSON. Worthing: Energy Storage Publishing, Ltd., Autumn 2012, No. 38, ISSN 1741-8666

[2] Polyvinylidene fluoride. (2016, March 11). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:33, March 12, 2016, from https://en.wikipedia.org/w/index.php?title=Polyvinylidene_fluoride&oldid=709606137

[3] MAKOVIČKA, J. *Negative Electrode of Lithium Secondary Cells* – Doctoral's Thesis. Brno: Brno university of technology, Faculty of electrical engineering and communication technology, 2008. Supervised by: Doc. Ing. Marie Sedlaříková, CSc.

[4] GOTTWALD, T. *Binders for supercapacitors electrodes* – Bechelor's Thesis. Brno: Brno university of technology, Faculty of electrical engineering and communication technology, 2008. 56 s. Supervised by: doc. Ing. Marie Sedlaříková, CSc.

[5] LIBICH, J. *Negative Electrode for Lithium-Ion Batteries*– Doctoral's Thesis. Brno: Brno university of technology, Faculty of electrical engineering and communication technology, 2016. Supervised by: doc. Ing. Marie Sedlaříková, CSc.

GEL POLYMER ELECTROLYTES FOR USE IN LITHIUM-ION BATTERIES

Michal Jahn

Doctoral Degree Programme (2), FEEC BUT

E-mail: xjahnm00@stud.feec.vutbr.cz

Supervised by: Marie Sedlaříková

E-mail: sedlara@feec.vutbr.cz

Abstract: Gel polymer electrolytes are perspective electrolytes for new types of Li-ion batteries today. The properties respectively advantages of these gels can be used also in other fields of electrochemistry. The gel polymer electrolyte is composed of a conductive component (liquid electrolyte) and the polymer component (methacrylate).

In this paper I present fundamental characteristics of gel polymer electrolytes. This paper presents a principle of gel polymer electrolytes, fundamental measurement and their evaluation. I present the techniques of measurement and preparation of this gel polymer materials. Potentiostatic impedance, linear sweep voltammetry, cyclic voltammetry and galvanostatic charge/discharge techniques were used for characterization of the samples.

Keywords: Gel, polymer, electrolytes, batteries, cell, methyl methacrylate, UV light, conductivity.

1. INTRODUCTION

Lithium cells respectively batteries are the most used and most widely electrochemical power sources at present. These cells use a graphite negative electrode, which receives released lithium ions during charging from the positive electrode. This process is called intercalation and is sufficiently reversible. Currently Li-ion batteries have a long lifetime that is up to several thousand cycles [2,3,5,8].

Much attention is focused on the research and development of new materials and technologies for use in advanced power sources (batteries), especially with regard to their efficiency, reliability, ecological production, low financial costs incurred to obtain materials that have these resources for the best properties and equipment that produce these chemicals resources are used. Among the perspective materials in the manufacture of chemical current sources belong gel polymer electrolytes, which now replace obsolete liquid electrolytes. These materials often have properties that can be used in other industries such as the construction of electrochemical sensors, super-capacitors, or electrochromic devices. Polymer gels based on lithium salts and other equivalents have many advantages compared to liquid electrolytes. The main advantages are the elimination leakage of electrolyte, better resistance to vibration and mechanical damage, or better resistance to volume changes of the electrodes. Another advantage is the adhesion of the gel to the electrode surface, this effect reduces the formation of aggregates on the surface and thus provides protection against internal short circuit and reduction capacity. Advanced gel polymer electrolytes having the same ionic conductivity as liquid electrolytes [2,3,5,8].

2. GEL POLYMER ELECTROLYTES

Elektrolyte materials separate anode and cathode and play the significant role of transmitting electrons and lithium ions during charging and discharging processes. Also, electrolyte is one of the

key components that define the battery's performance: charging/discharging capacity, safety, cycling

performance, and current density [4,6,9]. Figure 1 shows the design of the Li-ion battery with a gel polymer electrolyte.

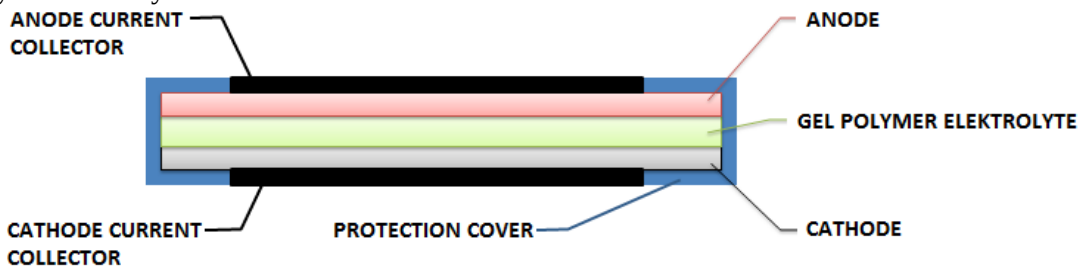


Figure 1: Schematic cross-section structure of the Li-ion battery with gel polymer electrolyte [6]

The basic desired qualities of gel polymer electrolytes are follows [4,6,7,9]:

- High ionic conductivity at wide range of temperatures: increase the lithium ions diffusion and resist polarization during charging/discharging
- Good thermal stability: ensures the battery operation under appropriate temperature
- Wide electrochemical window: prevents side reactions between electrodes and electrolyte
- Good mechanical property: ease of manufacturing and enhanced safety
- Low cost
- Safety: high flashing point
- Non-toxic: environmental friendly

2.1. CHEMICAL COMPOSITION

The gel polymer electrolyte is composed of a conductive component and the polymer component. The basis of conductivity component is a conventional liquid electrolyte. The polymeric part is based on methyl methacrylate and other necessary substances to form the gel polymer electrolyte. Other chemicals may be added to improve chemical, electro chemical and mechanical properties of the gel polymer electrolyte [2,3,5,8]. Individual components and some selected chemicals to form the gel polymer electrolyte are shown in Table 1.

Conductivity component		**Polymeric component**		**Others chemicals**	
Solvent	*Propylene carbonate, ethylene carbonate, dimethyl carbonate, sulfolane, etc.*	Polymer	*Methyl methacrylate, poly(methyl methacrylate, 2-ethoxyethyl methacrylate, 3-(trimethoxysilyl) propylmethacrylate, etc.*	Substances to improve electrical, electrochemical and chemical properties	*Ionic liquids, flame retardants, nanoparticles (Al_2O_3, TiO_2), etc.*
		Polymeric catalyst	*Benzoin ethyl ether (UV light)*		
Electrolyte salt	*Lithium perchlorate, sodium perchlorate, lithium tetrafluoroborate, etc.*	Networking agent	*Ethylen glycol dimethacrylate, diethylene glycol dimethacrylate, triethylene glycol dimethacrylate, tetraethylene glycol dimethacrylate, etc.*		

Table 1: Chemical components of the gel polymer electrolyte [2,3,4,5,6,7,8,9]

2.2. PREPARATION AND POLYMERIZATION

All experiments were based on theoretical knowledge and previous experiments. Sample resp. gel polymer electrolyte is circular gel with a diameter 16 mm and a defined thickness, which was cut out from the total area of formed (polymerized) gel. All handling of the gel was performed in a glove box with an argon atmosphere. To form a gel polymer electrolyte is necessary to follow a precise technological process. This technological process is very effective for the production of gel polymer electrolytes [2,3,4,8].

Technological process consists of several following steps:

1. Weigh all solid state components (electrolyte salt and polymeric catalyst) on the analytical balance.
2. The addition all liquid chemicals (first solvent, second polymer and last networking agent). Chemicals are precisely measured using a pipette.
3. All chemicals are mixed in a vial with a magnetic stirrer. The time required for complete dissolution of all chemicals is about 20 minutes. Must be checked visually.
4. Using a syringe, the mixed solution is injected into the prepared form. Form for the gel is a square shape measuring 10 x 10 cm and is composed of several layers. These layers are in the following order: glass, transparent film, spacer silicone (determines the shape and thickness of the gel), teflon and glass. This form is shown in figure 2. All the layers are using clips and clamp fixed to each other. Large fixation is required the drain of the gel in liquid form.

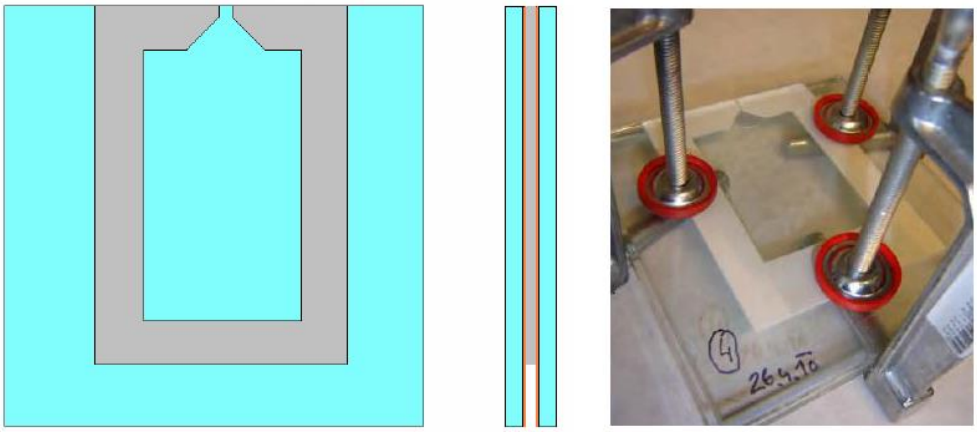


Figure 2: Schematic structure of form for gel (front, side, real photo) [8]

5. This form with liquid gel is placed under a UV lamp – polymerization of gel. The time required for the polymerization is dependent on the particular chemical composition of the gel (from minutes to hours).
6. Complete decomposition of this form ant cut out sample (gel polymer electrolytes) from gel.
7. Documentation of the state sample (record of the state, possibly capture).
8. Measuring the thickness of the sample using a micrometer.
9. The sample is inserted into the el-cell and followed by measurement of the electrical properties (PEIS, LSV and also CV a GCPL for selected samples).

The Potentiostatic Impedance (PEIS) experiment performs impedance measurements into potentiostatic mode by applying a sinus around a potential E that can be set to a fixed value or relatively to the cell equilibrium potential [1].

Cyclic Voltammetry (CV) is the most widely used technique for acquiring qualitative information about electrochemical reactions. CV provides information on redox processes, heterogenous elec-

tron-transfer reactions and adsorption processes. It offers a rapid location of redox potential of the electro active species. [1].

2.3. MEASUREMENT AND EVALUATION

In electrolytes apply Ohm's law, as in other conductors, because all conditions are fulfilled. Especially it is the proportionality between the mean velocity of the ions and electric field strength. In the field of electrochemistry for determine the properties of the electrolyte is advantageous to calculate the conductances *G* and specific conductivity resp. conductivity γ:

$$\gamma = \frac{1}{R} \times \frac{d}{S} \quad [mS/cm], \tag{1}$$

where *R* is the resistance, *d* is the thickness and *S* is the area [2,3,5].

Measurement and evaluation was performed for a range of samples, the following table lists only some of them. The table shows chemical composition, time of polymerization and the calculated conductivity (equation 1).

	Chemical composition and polymerization parameters						
Sample [code]	***propylene carbonate* [ml]**	***lithium perchlorate* [g]**	***methyl methacrylate* [ml]**	***benzoin ethyl ether* [g]**	***ethylen glycol dimethacrylate* [µl]**	***time of polymerization* [hour]**	***conductivity* [mS/cm]**
16/02/02	1,00	0,053	1,75	0,039	9,75	*1*	**0,02129**
16/02/15	1,00	0,053	1,75	0,039	9,75	*1*	**0,02258**
16/02/22	1,00	0,053	1,75	0,039	9,75	*2*	**0,04868**
16/02/23	1,00	0,053	1,75	0,039	9,75	*0,75*	**0,01734**
16/03/01	1,50	0,080	1,75	0,039	9,75	*1,5*	**0,18610**
16/03/07	1,50	0,080	1,75	0,039	9,75	*2,5*	**0,03306**
16/03/09	1,50	0,080	1,75	0,039	9,75	*3,5*	**0,10866**

Table 2: Some selected samples of the gel polymer electrolyte

The following figure shows a measured nyquist impedance of sample 16/03/01. Measurements are made using a potentiostat BioLogic and related software EC-Lab, using software to calculate the value of resistance *R*. The resistance value used to calculate the specific conductivity γ (equation 1). Measurement range is from 1 MHz to 0,1 Hz.

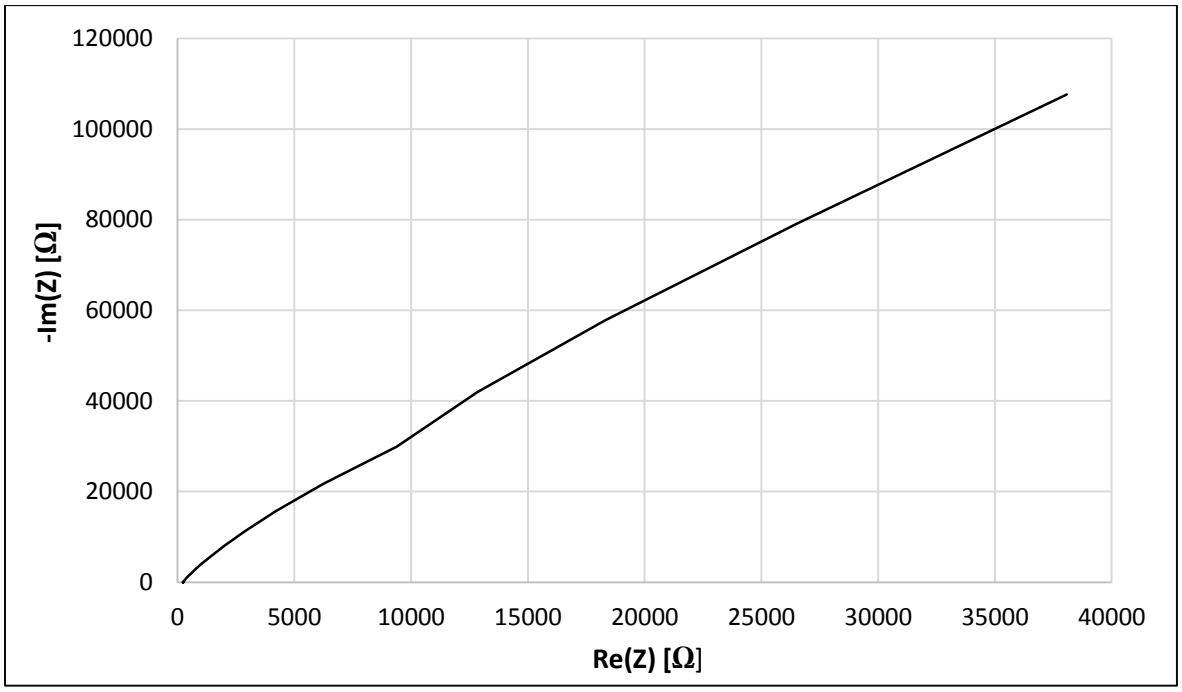


Figure 3: Nyquist impedance of sample 16/03/01

3. CONCLUSION

Current development is focused on improving the parameters of polymer electrolytes: increase ion conductivity, extension in potential window and to improve long-term chemical and electrochemical stability of the polymer and solvent. Very promising combination can be electrochemically stable polymers and ionic liquids, which are substances that consist of an organic cation and an inorganic or organic anion, with a melting point below 100 ° C. Main advantages of ionic liquids are high ionic conductivity, flammability and excellent electrochemical stability. These ionic liquids may be suitable for new lithium-ion batteries and supercapacitors, especially where the required higher thermal stability.

The main goal is to create a gel polymer electrolyte for use in lithium ion batteries with the same or better performance than conventional liquid electrolytes.

ACKNOWLEDGMENTS

This work was supported by the grant FEKT-S-14-2293 "Materiály a technologie pro elektrotechniku II", grant FCH/FEKT-J-16-3390 "Gelové polymerní elektrolyty pro Li-ion akumulátory" and Centre for Research and Utilization of Renewable Energy under project No. LO1210 - "Energy for Sustainable Development (EN-PUR)".

REFERENCES

[1] BIOLOGIC SCIENCE INSTRUMENTS. *EC-Lab Software Techniques and Applications*. 10.1x. Bio-Logic SAS, 1, rue de I´Europe, F-38640 Claix, France, 2011. Available form: www.bio-logic.info

[2] CAFOUREK, Martin. *Nové gelové polymerní elektrolyty*. Brno, 2010. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Vedoucí práce Doc. Ing. Marie Sedlaříková, CSc.

[3] KRATOCHVÍL, Martin. *Technologie přípravy aprotických gelových polymerních elektrolytů na bázi PMMA*. Brno, 2008. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Vedoucí práce Doc. Ing. Marie Sedlaříková, CSc.

[4] KREJZA, Ondřej. *Gel polymer electrolytes for electrochromic device*. Brno, 2009. Dizertační práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Vedoucí práce Prof. Ing. Jiří Vondrák, DrSc.

[5] KUČERA, Vladimír. *Fotopolymerace*. Brno, 2012. Bakalářská práce. Vysoké učení technické v Brně, Fakulta chemická. Vedoucí práce Mgr. František Kučera, Ph.D.

[6] SEDLAŘÍKOVÁ, Marie. *Polymerní gelové elektrolyty a jejich využití v elektrotechnice*. Brno, 2003. Habilitační práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií.

[7] STEPHAN, A. Manuel. Review on gel polymer electrolytes for lithium batteries. *European Polymer Journal* [online]. Science Direct, 2006(42), 21-42 [cit. 2016-03-06]. Dostupné z: http://krc.cecri.res.in/ro_2006/034-2006.pdf

[8] ZÍTKA, Jan. *Studium vlastností gelových polymerních elektrolytů pro lithno-iontové akumulátory*. Brno, 2013. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Vedoucí práce Ing. Michal Musil.

[9] ZHANG, Ruisi. *Advanced gel polymer electrolyte for lithium-ion polymer batteries*. Iowa, 2013. Graduate theses and Dissertations. Iowa State University. Vedoucí práce Professor Reza Montazami.

HIGH PRECISION REFERENCE POSITION GENERATOR FOR MOTOR CONTROL IN FPGA

Vojtech Dvorak

Doctoral Degree Programme (3), FEEC BUT

E-mail: xdvora99@stud.feec.vutbr.cz

Supervised by: Lukas Fujcik

E-mail: fujcik@feec.vutbr.cz

Abstract: The paper deals with an architecture of a high precision reference position generator for a motor controller implemented in FPGA. The architecture was developed to achieve high accuracy and tunability with concern to resource efficient implementation in FPGA. Simulations of the proposed generator as well as synthesis results are presented in the paper.

Keywords: FPGA, Motor Controller, Synchronization

1. INTRODUCTION

A motor driving is wide group of task for electronic design. The most common approach is to utilize microcontroller due to low cost and well understand implementation of motor controller there. However, in some cases utilizing FPGA is desired, especially when other monitoring and control functions are required to implement beside motor controller. The structure of the motor controller in FPGA (in Figure 1) was developed to fulfill requirements to rotate motor with constant speed and reach commanded position with high precision in a time.

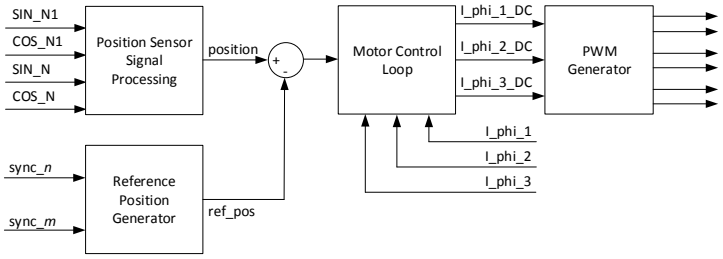


Figure 1: Block diagram of motor controller

The current position of the motor is computed in the *Position Sensor Signal Processing* from a position sensor with two tracks with different number of sinewaves per one revolution (Vernier principle). The difference between the current position and the reference position is input for the *Motor Control Loop* that is implemented as a cascaded structure of PI regulators regulating position, speed and current through motor windings. The output of the *Motor Control Loop* is voltage to be applied on the motor and is transferred to the *PWM Generator* that generates PWM modulated voltage for a motor driver. In this case, the reference position is a ramp computed from periodic pulses generated by a superior system. The computation of the reference position is performed by the *Reference Position Generator* block. The accuracy of motor driving directly depends on the accuracy of the computed current position of the motor and the accuracy of the reference position.

The FPGA vendors provide dedicated IP core resources for motor driving. However, the IP cores deal mostly with the motor control loop PID regulators [1][2].The algorithm to process data from position sensor is implemented according algorithms described in [3][4]. In this paper, an architecture of the *Reference Position Generator* is described.

2. REFERENCE POSITION GENERATOR

The reference position generator generates the ramp representing required position of the motor in time. There are two synchronization pulses to indicate required position in time. When rising edge of the pulse arrives, it is expected to be in the position defined in Figure 2. However, motor has to rotate with constant speed between rising edges of the pulses. This is achieved by interpolation of the required position with high resolution ramp.

The synchronization pulse *sync_m* indicates one revolution of the motor. The second synchronization pulse *sync_n* is used for accurate synchronization during rotation. The timing diagram of pulses together with the required position of the motor is in Figure 2.

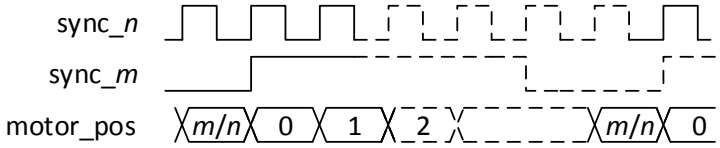


Figure 2: Timing diagram of synchronization pulses

The period of synchronization pulses can vary in a time and thus can be used to change speed of rotation. There is only one requirement for speed variability of the motor, the ratio *m*/*n* has to be a constant (i.e. the number of *sync_n* pulses in one *sync_m* period has to be still same when periods of synchronization pulses are changed).

The developed *Reference Position Generator* contains two counters (the *Reference Counter* and the *Ramp Counter* in Figure 3). *The Reference Counter* generates a time-accurate low-resolution ramp similar to the signal *motor_pos* in Figure 2. The low-resolution ramp is used as a reference for the high-resolution Ramp Counter that generates reference position. Due to required tunability, the additional correction logic is implemented to control the slope of the generated ramp to be the same as the slope of the output of the Reference Counter.

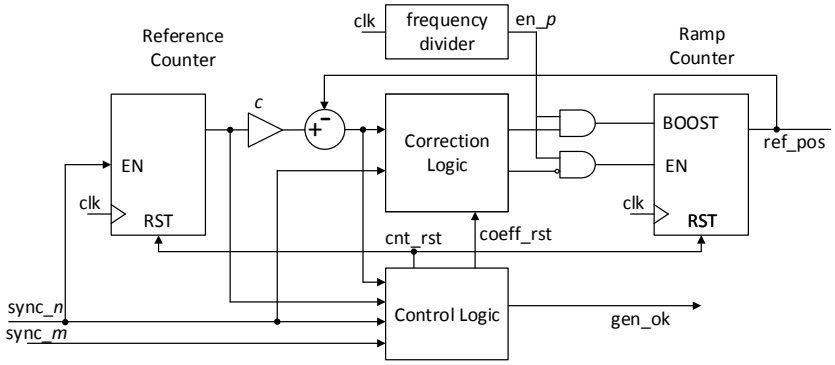


Figure 3: Block diagram of Reference Position Generator

The block schema of the developed Reference Position Generator is in Figure 3. Beside the counters, there is a correction logic, a control logic and a frequency divider. The frequency divider is optional and was implemented to reduce size of the Ramp Counter and to decrease sensitivity to signal jitter on the synchronization link.

The Control Logic ensures correct sequence of synchronization after power on. It also detects the loss of the synchronization link and provides sequence to resynchronize after synchronization link is established again.

The correction of the slope of the *Ramp Counter* is driven by the *Correction Logic* with control signals. The control signals can disable the *Ramp Counter* (to slow down counting) or boost counting of the *Ramp Counter*. When the boost is enabled, the *Ramp Counter* increments its output by 2.

A difference between both counters is used as an input for both the *Correction Logic* and the *Control Logic*. Because the *Reference Counter* has smaller resolution than the *Ramp Counter*, the value of the *Reference Counter* needs to be multiplied by a constant *c* computed according the (1).

$$c = (m / n) / p \tag{1}$$

Where:

m is a period of synchronization pulse *sync_m* (i.e. period of one motor revolution)

n is a period of synchronization pulse *sync_n*

p is a period of divided frequency of a system clock (if the frequency divider is omitted, *p* is period of system clock).

The proposed reference position generator was developed to achieve tunability ±10% with respect to the nominal speed of rotation. However, the tunability can be improved without any change in the architecture of the generator.

2.1. CORRECTION LOGIC

The *Correction Logic* was implemented to allow tunability of the slope of the generated reference position. The block diagram of the *Correction Logic* is in Figure 4. There are two parts of the logic, the Correction Coefficient Computation block computes correction coefficient to correct slope of the *Ramp Counter*. The second part applies the computed correction coefficient to the *Ramp Counter* with two signals – boost and disable.

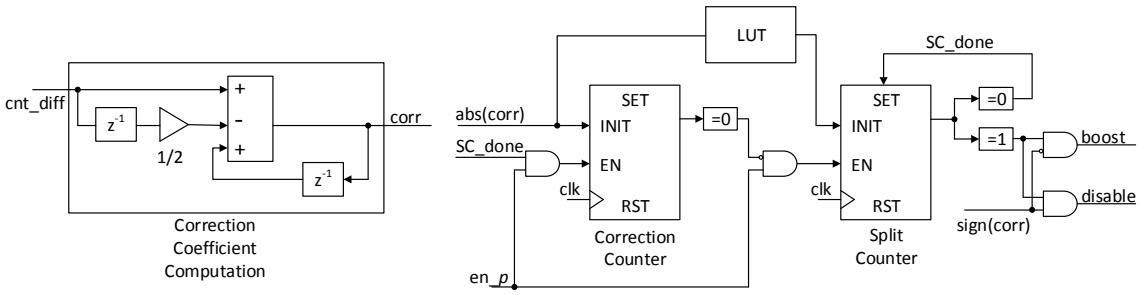


Figure 4: Block Schema of the Correction Logic

The correction coefficient is computed according (2).

$$corr = cnt\_diff - \frac{1}{2} cnt\_diff * z^{-1} + corr * z^{-1} \tag{2}$$

An application of the computed correction coefficient is split over the whole period of the *sync_n* pulse. This approach was selected to generate ramp with constant slope for whole period. Two counters were utilized to generate boost or disable signal for the *Ramp Counter*. The *Correction Counter* is first counter and counts number of required corrections. It is a down counter which is set to the value of the correction coefficient with rising edge of the *sync_n* and decrease its value by 1 when the second counter is finished. The second counter is the *Split Counter* that is used to split period of *sync_n* pulse to smaller periods according to the size of the current correction coefficient.

2.2. CONTROL LOGIC

The *Control Logic* starts synchronization of the *Reference Counter* and the *Ramp Counter* after power on and checks whether the generated reference position is valid. The *Control Logic* is implemented as a finite state machine according the state diagram in Figure 5.

The *Control Logic* generates sequence of resets for the counters and for the *Correction Logic* after power on. Both counters and the *Correction Logic* are hold in reset until the first synchronization pulse *sync_n* arrives. Consecutively, counters and the *Correction Logic* are enabled to determine the period of synchronization pulse *sync_n* and to compute corresponding correction coefficient. Until the first sync_*m* pulse arrives, the generated reference position is not valid (indicated by signal *gen_ok*). A rising edge on the *sync_m* is the trigger to reset both counters and start counting from zero with the correction coefficient computed in the previous states.

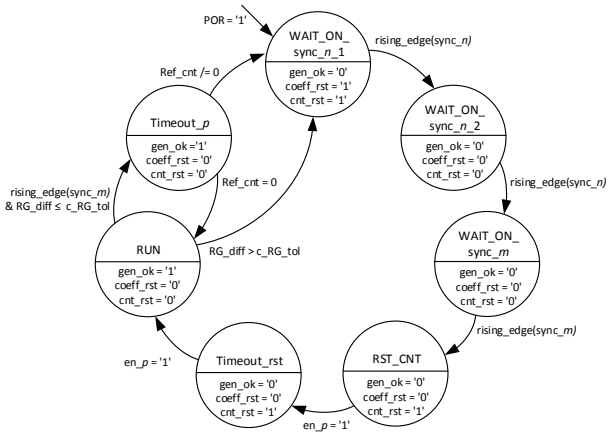


Figure 5: State diagram of Control Logic

The detection of the errors on the synchronization link is performed only in the RUN state of the *Control Logic*. There are two possible errors on the synchronization link. First error is the loss of the link due to disconnection. To detect this error, difference between the *Reference Counter* and the *Ramp Counter* is observed and once the difference exceeds limit defined by the variable *c_RG_tol*, the *Control Logic* goes to the WAIT_ON_sync_n_1 state and waits for the new synchronization pulse. The second error is caused by error in the superior system generating pulses. The number of *sync_n* pulses in one *sync_m* pulse is incorrect. The output of the Reference Counter is observed and checked whether is in zero after *sync_m* arrives. A timeout is utilized to avoid false error detection due to jitter on the synchronization link.

3. SIMULATION AND SYNTHESIS RESULTS

The parameters of the *Reference Position Generator* was set according Table 1 to perform simulation and synthesis to FPGA of the proposed architecture.

Parameter	Value
n	25 ms
m	5.5 s
p	5 µs

Table 1: Parameters of the Reference Position Generator

The functional simulation of the proposed architecture was performed with QuestaSim 10.2b digital simulator. The synchronization pulses were generated in verification environment coded in SystemVerilog. Simulation waveform can be found in Figure 6.

There were 8 tests performed to examine the behavior of the *Reference Position Generator* under different conditions. Operation with the nominal period of synchronization pulses was tested in Test 1, Test 2, Test 6 and Test 8. All the tests were performed to prove that the *Reference Position Generator* is able to work correctly after power on or after previous different periods of the synchronization pulses. Tunability of ±10% to the nominal period

was tested in Test 3 and Test 4. Test 5 and Test 7 cover situation when the period of the synchronization pulses is out of allowed range (greater than ±10%).

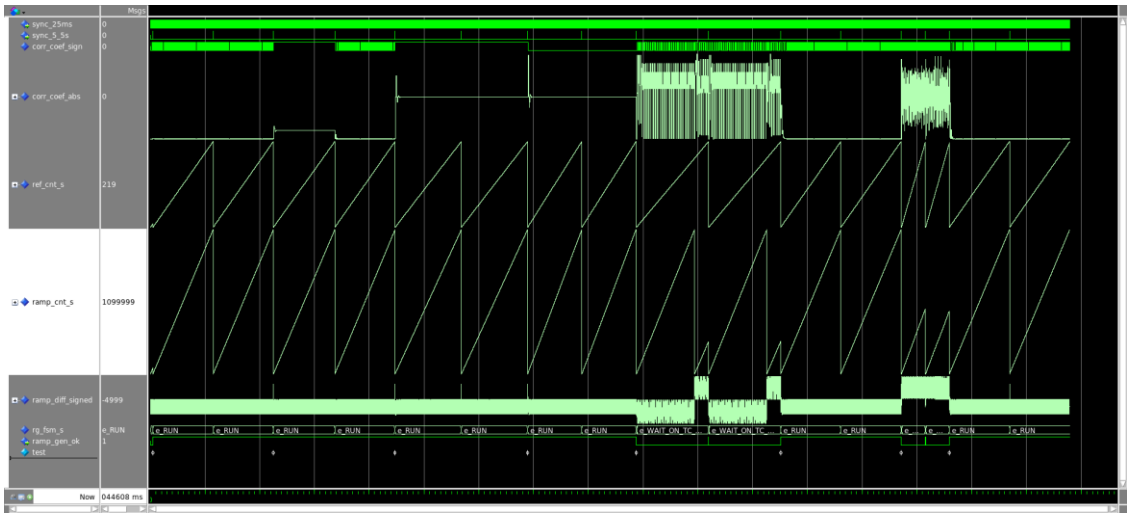


Figure 6: Simulation Waveform of the Reference Position Generator

The logic synthesis was performed to estimate resource utilization in FPGA. Microsemi RTAX 1000S FPGA was selected as a target FPGA. Results of the synthesis for the parameters defined in Table 1 are summarized in Table 2.

	Used	Total	Utilization
LUT [-]	564	12096	5%
DFF [-]	169	6048	3%
f_{max} [MHz]	75.5	-	-

Table 2: RTAX1000S FPGA Resource Utilization

4. CONCLUSION

The architecture of a high precision reference position generator was presented in the paper. The proposed generator is able to generate high precision reference ramp for motor controller to rotate motor with constant speed and reach commanded position indicated by events on synchronization pulses. Simulation of the proposed architecture was performed to prove functionality of the concept. Synthesis for target FPGA was performed to provide resource utilization of the generator in hardware.

ACKNOWLEDGEMENT

Research was supported from the project no. FEKT-S-14-2300 "A new types of electronic circuits and sensors for specific applications".

REFERENCES

[1] Microsemi: Dual-Axis Motor Control on a Single SoC FPGA, Microsemi, 2015

[2] Altera: Motor Control IP Suite Components for Drive-on-Chip Reference Designs, Altera, 2014

[3] iC Haus: High-Precision Sine/Cosine Interpolation, iC-House GmbH, 2014

[4] GEMAC: 2-Channel Interpolation Circuit with Nonius Calculation, GEMAC, 2013

ORGANIC–INORGANIC HYBRID SOLAR CELLS

Dávid Strachala, Josef Hylský, Michal Kadlec

Doctoral Degree Programme (2.), FEEC BUT

E-mail: xstrac07@stud.feec.vutbr.cz, xhylsk00@stud.feec.vutbr.cz

Supervised by: Jiří Vaněk

E-mail: vanekji@feec.vutbr.cz

Abstract: The purpose of the presented work is to find an optimal way for the preparation of the organic-inorganic hybrid solar cells based on perovskite structure $CH_3NH_3PbI_{3-2}Cl_2$ /$PC_{61}BM$. The main idea of the work is to create an eight solar cell samples with the highest possible conversion efficiency and stable structure. Conversion efficiency of created samples reached approximately η_{MAX} = 2, 5 %. Samples were futher analyzed and based on that were implemented other techniques for their improvement.

Keywords: Organic – inorganic hybrid solar cells, perovskite solar cells, conversion efficiency.

1. INTRODUCTION

Predicted reduction of the fossil energy sources and increasing environmental needs, leads to an intensive research of the new renewable energy sources - especially the methods of conversion the solar energy via alternative photovoltaic cells. The attention was focused on the development of solar cells based on perovskite in the recent years. Well-known scientist Michael Grätzel called perovskite as "unusually cheap" material that could operate successfully in solar cells and combine the seemingly incompatible - low acquisition cost and high efficiency of energy conversion of sunlight into electricity [1]. Advantages of perovskite solar cells lies in their potential use in a broad application such as windowpanes, facade systems and much lower production costs (also availability of materials) in comparison to a traditional silicon solar cells with a relatively easier methods of its preparation. Perovskite solar cells can be considered as promising power source worthy attention.

2. ORGANIC-ANORGANIC HYBRID SOLAR CELLS

2.1. PEROVSKITE SOLAR CELL

Perovskite in general is a name for a group of substances with the same crystal structure as $CaTiO_3$. In photovoltaic cells, modified compounds of $CaTiO_3$ are used for absorbing of sunlight. A semiconductor is created as an organic-inorganic hybrid composition as $CH_3NH_3PbX_3$ (X = Cl, I, Br). Solar cells based on perovskite have been continuously developed in the last five years. The first time of publication of this material in the solar cells was in 2009 - energy conversion efficiency of photovoltaic solar cells with liquid electrolyte was 3.13% and 3.81% ($CH_3NH_3PbBr_3$ and $CH_3NH_3PbI_3$) [2]. In the following years, due to optimizations of various parameters of the photovoltaic cells there has been a rapid increase in energy conversion efficiency in a relatively short time [2].

Presented work deals with a hybrid peroskvit solar cell structure of $CH_3NH_3PbI_{3-2}Cl_2/PC_{61}BM$. Fig. 1 shows the general device structure of the perovskite solar cells.

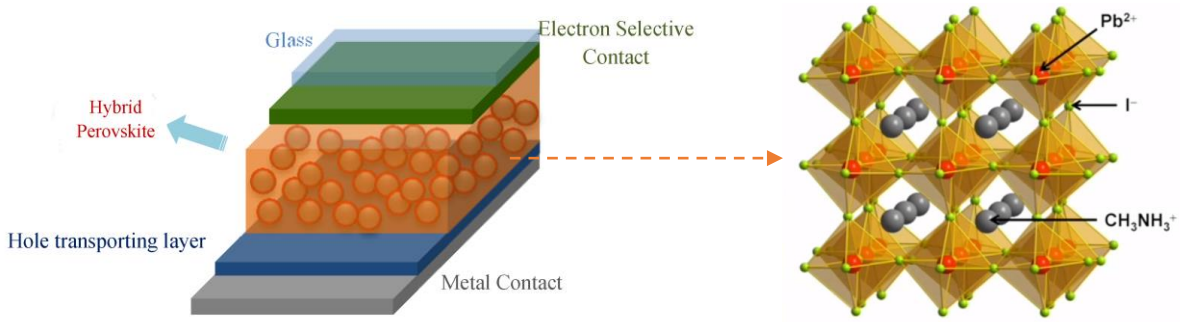


Figure 1: a) General structure of perovskite solar cells [1]

2.2. MANUFACTURING OF PEROVSKITE SOLAR CELL

An important step of work was to choose an appropriate process of manufacturing of examined peroskvite solar cell. Solar cells were fabricated on PEDOT coated indium tin oxide glass (ITO) with emphasis on the optimal way of heat treatment. The ITO substrates were before this process cleaned in an ultrasonic bath. $CH_3NH_3PbI_{3-2}Cl_2$ film was formed through a spincoating of a mixture of CH_3NH_3PbI, PbI_2 and $PbCl_2$. After verifying that the precursor solutions were successfully synthesized, the organic semiconductor $PC_{61}BM$ was deposited. To prevent the creation of interdiffusion between CH_3NH_3I and $PC_{61}BM$, high temperature wasn´t applied. Finally, aluminum cathode was deposited through a mask on the active layer. A detailed description of the preparation is given in the following text.

Preparation of the carrier substrates with an ITO layer was proceeded according to the following steps: the substrates were firstly put into the beaker with a 5-10% of NaOH solution. Subsequently, after a period of 3-5 minutes, were left in the ultrasonic bath, then rinsed with MilliQ water and finally cleaned with an isopropylalcohol.

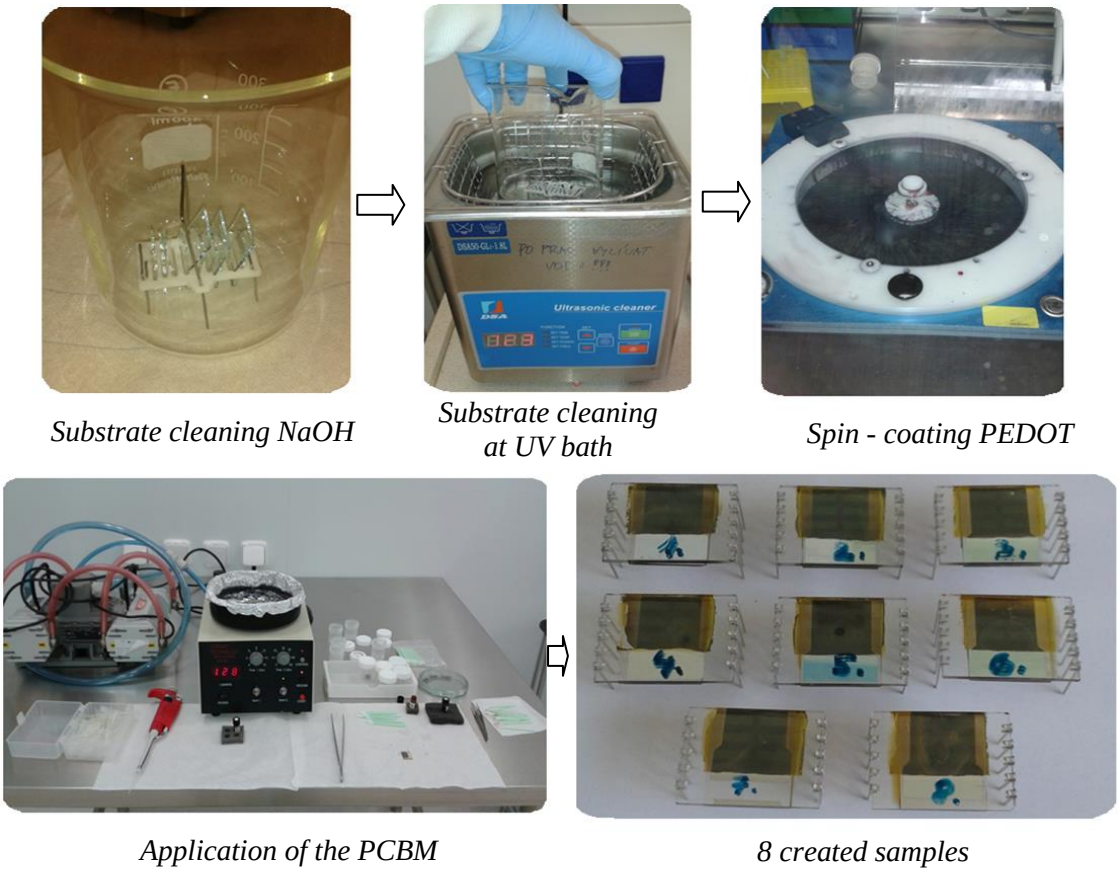
PEDOT: PSS was before applying filtered through a 45 μm PVDF filter. A 55 ml of solution was used. Spin-coating was adjusted to a rotational speed of 3000 rpm (with the acceleration speed of 500 rpm) for 30 seconds. The cathode strip was after this process purified by MilliQ water. The final layer thickness after annealing (120 ° C for 30 sec.) was 30 - 40nm.

Perovskite layer was prepared from molar solution of the mass fraction of 1:1:4 with a total concentration of 9 percentage by weight. In numbers it means: of CH_3NH_3PbI (0.19561g), PbI_2 (0.141822g) and $PbCl_2$ (0.08555g). The solution was prepared within the laboratory conditions for 12 hours at the temperature of 60 °C with the stirring speed of 650 rpm. The solution was then filtered with a 45μm filter and applied to a preheated glass slides (125 °C for a 30 sec) by spin-coating in an amount of 35μl with the rotational speed of 3000 rpm with a speed acceleration of 500 rpm. After deposition was on the slides cleaned anode strip. After this step were slides transferred to a hot plate with an 100 °C temperature for a period of 120 minutes. This process wasn´t unfortunately successful - created samples has shown a low efficiency in the light spectrum of 280 - 350 nm. It could be due to burn-through of layers which was indicated by the higher amount of carbon contained in the samples (see chapter Microscopic Analysis of Defective Sample). The procedure was therefore adjusted to heating for only 30 seconds and preheating of the glass was reduced from 100 ° C to 70 ° C. After this steps did not occur burn-through of layers anymore.

The layer of $PC_{61}BM$ was applied similarly. The solution was prepared from 10 mg / ml of $PC_{61}BM$ in a mixture of anhydrous solvents of chloroform-chlorobenzene (1: 1). Deposition was performed through a spin-coating at 1200 rpm for 60 sec. in an amount of 50 ml, but this time without a subsequent annealing. Cathode strip was refined with a chloroform.

In the end an aluminum electrodes were vapor-deposited with a thickness of 100 nm followed by encapsulation of the samples with the epoxy resin. Cover glass with contacts were also attached.

The result of the manufacturing process were 8 created samples. Samples no. 7 was damaged due to incorrect technological procedure. This sample was used for an microscopic inspection. Shortened process of realization of samples is for demostration purposes shown in Fig. 2.



Substrate cleaning NaOH *Substrate cleaning at UV bath* *Spin - coating PEDOT*

Application of the PCBM *8 created samples*

Figure 2: Manufacturing Process of the Perovskite Solar Cell

3. MEASURED DATA

In the structure of the peroskvite solar cells was necessary to monitor various parameters that influenced the resulting device functionality. One of the parameters investigated in the work was the preparation of optimized layer $CH_3NH_3PbI_{3-2}Cl_2$ / $PC_{61}BM$, which was used to form functional and stable reference solar cell with the best electrochemical and physical properties. The structure of these films were scanned by several methods using environmental scanning electron microscope and an atomic force microscope. The chemical composition of the layers was then examined via X-ray diffraction spectroscopy for a suitable setup of deposition processes.

3.1. MICROSKOPIC ANALYSIS OF DEFECTIVE SAMPLE

Microscopic scanning of the cross section of sample number 7 has shown that the solar cell did not have required structure. This can be caused by burn-through of layers during the manufacturing process which could be indicated by a very high amount of carbon contained in the sample. This assumption was also confirmed by images from the SEM microscope Vega3 Tescan and X ray spectroscopy analysis measured at Brno University of Technology. Based on this observation were the manufacturing process of other samples changed - a temperature of preheating during application perovskite material was reduced from 100°C to 70°C, as was mentioned before.

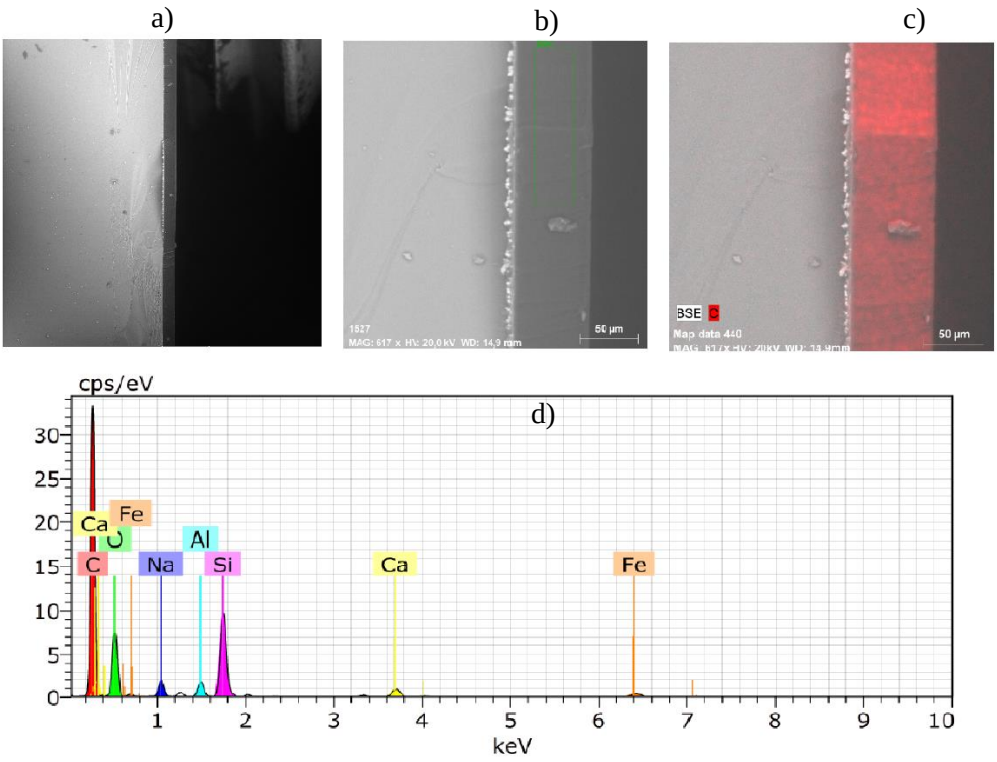


Figure 3: a) Microskopic analysis of the deffected sample 7 by SEM Vega3 Tescan -magnitude 150x, HV 20kV; b) X ray spectroscopy analysis - magnitude 617x, HV 20kV, c) X ray spectroscopy analysis- BSE C magnitude 617x, HV 20kV d) increased incidence of carbon cauesed probably by burn-through of applied layers

3.2. REACHED CONVERSION EFFICIENCY AND CURRENT-VOLTAGE CHARACTERISTIC

Initial measurement has shown efficiency of the created perovskite solar cells around 2 percent. The highest conversion efficiency was achieved in sample no. 1 (η_{MAX} = 2.46%); no.2 (η_{MAX} = 2.85%) and no.4 (η_{MAX} = 2.56%). Measurements were done through a solar simulator with source of light energy of 1000 W/m^2. Investigated solar cells generated values in average shown in Table I.

Parameters	I_{SC} [mA]	U_{OC} [mV]	I_{MAX} [mA	U_{MAX} [mV]	P_{MAX} [mW]	*FF* [%]	η [%]
Average value	0.439	570.228	0.289	366.565	0.109	41.749	1.817

Table 1: Averague value of parameterers of measured samples

For an example are in fig. 4 shown current-voltage characteristic of samples no.4, no.5 and no.6.

With a measurement during illumination the efficiency has increased to a certain value, then again has dropped. The reason is probably in the temperature of the sample during measurements (samples were rapidly heating). Samples were also measured consecutively in short intervals. The gradual degradation of the material, well-known phenomenon of perovskite solar cells, has occurred later. The conversion efficiency of the cells has decreased significantly one week after the measurements. Degradation of the perovskite solar cells is currently one of the biggest problem of these materials [3]. Subsequent research will focus mainly on preventing of the degradation by applying the appropriate interlayers between perovskite and its metal contacts. The Cr_2O_3 interlayers will be studied in the context of resistance of the perovskite material against environmental influences [3]. The Cr_2O_3 layer will be applied through magnetron sputtering equipment in the Department of Electrical and Electronic Technology of Brno University of Technology.

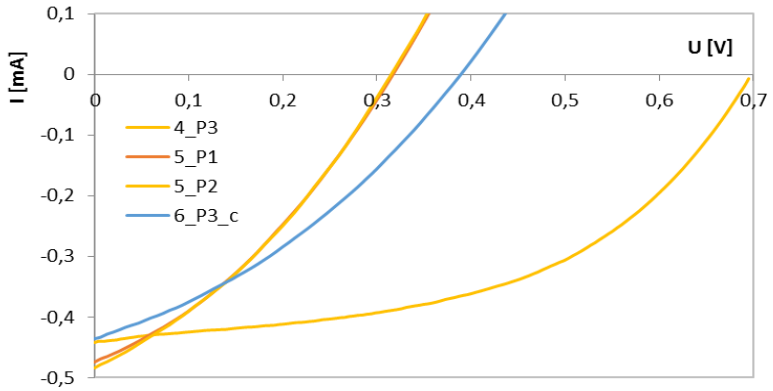


Figure 4: Current-voltage characteristic of solar cells

4. CONCLUSION

This work is aimed at optimizing the perovskite solar cells with synthehis of $CH_3NH_3PbI_{3-2}Cl_2$ through a mixture of CH_3NH_3PbI, PbI_2 and $PbCl_2$. In the perovskite solar cell was examined the influence of reactant ratio, time and annealing temperature during its preparation with an emphasis on the absorbation and stability of the used layers.

The first measurements of realized samples according to chosen procedure has shown an efficiency of peroskvite solar cells in the units of percent. The highest conversion efficiency was achieved in sample no. 1 (η_{MAX} = 2.46%); no.2 (η_{MAX} = 2.85%) and no.4 (η_{MAX} = 2.56%). Microscopic scanning of the sample number 7 has shown burn-through of layers during the manufacturing process (indicated by a very high amount of carbon contained in the sample).

Originally, there were perovskite solar cells prepared in combination with a liquid electrolyte, but there occurred rapid degradation due to dissolution of crystals of perovskite. Presented work aims to find the optimal method for the synthesis and preparation of peroskvit solar cell based on $CH_3NH_3PbI_{3-2}Cl_2$ /$PC_{61}BM$. The main idea of the following work is to achieve the highest quality and most stable structures of synthesized materials and to build the long-term stable solar cells with high efficiency of photovoltaic conversion.

ACKNOWLEDGEMENT

This research work has been carried out in the Centre for Research and Utilization of Renewable Energy (CVVOZE). Authors gratefully acknowledge financial support from the Ministry of Education, Youth and Sports of the Czech Republic under NPU I programme (project No. LO1210) and BUT specific research programme (project No. FEKT-S-14-2293).

REFERENCES

[1] Kim, H.-S.; Lee, C.-R.; Im, J.-H.; Lee, K.-B.; Moehl, T.; Marchioro, A.; Moon, S.-J.; Humphry-Baker, R.; Yum, J.-H.; Moser, J. E.; Grätzel, M.; Park, N.-G. Nat. Sci. Rep. 2012, 2,591

[2] Heo, J.H.; Im, S.H.; Noh, J.H.; Mandal, T.N.; Lim, C.-S.; Chang, J.A.; Lee, Y.H.; Kim, H.-j.; Sarkar, A.; NazeeruddinMd, K.; Gratzel, M.; Seok, S.I., Efficient inorganic-organic hybrid heterojunction solar cells containing perovskite compound and polymeric hole conductors. Nat Photon, 2013, 7, (6), 486-491.

[3] KALTENBRUNNER, Martin, Getachew ADAM, Eric Daniel GŁOWACKI, et al. 2015-8-24. Flexible high power-per-weight perovskite solar cells with chromium oxide–metal

INFLUENCE OF ANODIZATION PARAMETERS ON MORPHOLOGY OF TITANIA NANOTUBES

Kateřina Přikrylová

Microelectronics and Technology (2nd year), FEEC BUT

E-mail: xprikr26@stud.feec.vutbr.cz

Supervised by: Jana Drbohlavová

E-mail: drbohla@feec.vutbr.cz

Abstract: Highly ordered titania nanotubes (TNTs) were fabricated via electrochemical anodization of titanium thin film in fluoride and water containing organic electrolyte. Titanium film was sputter-deposited on silicon wafer with thermal silicon dioxide layer. The effects of applied anodization potential, electrolyte composition, anodization time on the formation of TNTs were discussed. The influence of anodization parameters on morphology of TNTs was observed by scanning electron microscopy (SEM). The anodization time mainly influenced TNTs length, while applied potential and electrolyte composition strongly affected the morphology of TNTs during the anodization process where the removing of initial barrier layer plays a significant role.

Keywords: Titanium dioxide, anodic oxidation, nanotubes, SEM.

1. INTRODUCTION

Titanium dioxide (TiO_2) is known to be environmentally friendly, and corrosion-resistant having its characteristic morphologies, unique physical and chemical properties. Therefore TiO_2 is a versatile material used in many different research and application fields including paints [1], biomedicine [2], photocatalysis [3], photo-splitting of water [4], dye sensitized solar cells [5] and sensing [6]. For many of these applications, it is crucial to maximize the surface area to achieve a maximum overall efficiency, and therefore nanoparticulated forms of TiO_2 are widely used. Over the last few years an extensive research effort has been dedicated to the development of novel nanomaterials, including nanostructured titania prepared via anodization [3,7], sol-gel [8], hydrothermal synthesis [9], inject printing [10], and vapour deposition techniques [11].

The easiest way for obtaining self-organized TiO_2 nanotubes or nanopores can be an anodization process of titanium foil or thin coatings [12]. Nanoporous and tubular titania layers are formed in fluoride containing electrolytes when a suitable anodization potential is applied. The titania layers growth perpendicular to the metal substrate which can be self-organized under controlled conditions [13].

The mechanism of TNTs growth is shown in **Figure 1**.During the first few seconds of anodization the Ti started to be oxidized to Ti^{4+} ions at the surface while releasing electrons into electrolyte solution (Eq.1). The effect of electric field leads to the dissociation of water into OH^- and O^{2-} anions. Both of these ions recombine with Ti^{4+} and create either titanium hydroxide (Eq.2) or TiO_2 (Eq.3) in simple single process. Titania is further produced when the titanium hydroxide releases water by condensation reaction (Eq.4) [13,14].

$$Ti \rightarrow Ti^{4+} + 4e^{-} \quad (1)$$

$$Ti^{4+} + 4OH^{-} \rightarrow Ti(OH)_4 \quad (2)$$

$$Ti^{4+} + 2O^{2-} \rightarrow TiO_2 \quad (3)$$

$$Ti(OH)_4 \rightarrow TiO_2 + 2H_2O \quad (4)$$

In addition to oxidation, it is important to define the process of field-assisted dissolution of oxide preferentially at the tube base (bottom), where the electric field is stronger. Titanium dioxide are easily dissolved in fluoride based electrolytes. Fluoride anions attack the barrier oxide layer under the influence of the electric field generating the small dimples irregularly over the whole surface. This step results in the formation of F^- rich layer at the metal-oxide interface. In a fluoride containing electrolyte the oxide dissolves forming titanium hexafluoride complex [TiF_6^{2-}] (Eq.5), which is stable in water. The applied electric field may also weaken the bond between Ti and O in the oxide, effectively facilitating chemical dissolution of TiO_2 and therefore the formation of [TiF_6^{2-}] (Eq.6) [13,15].

$$Ti^{4+} + 6F^{-} \rightarrow TiF_6^{2-} \quad (5)$$

$$TiO_2 + 6F^{-} + 4H^{+} \rightarrow TiF_6^{2-} + H_2O \quad (6)$$

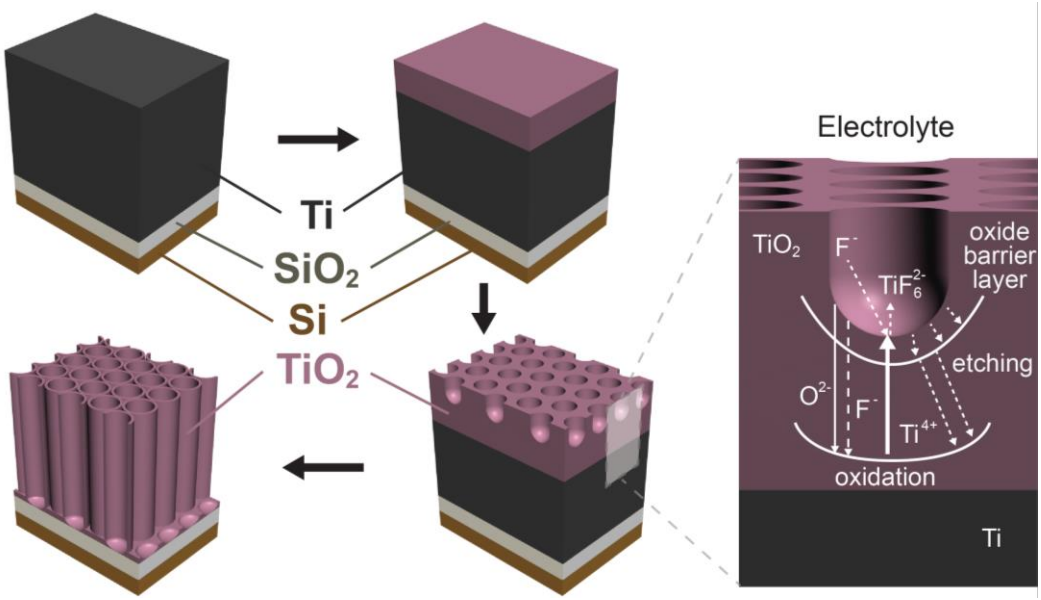


Figure 1: Schematic illustration of TNTs growth [15].

2. FABRICATION OF NANOTUBES

2.1. MATERIAL

Titanium (99.99%, Porexi, CZ), ethylene glycol ($C_2H_6O_2$, p.a., Penta, CZ) and ammonium fluoride (NH_4F, Reidel-de Haen, DE). All chemicals were used as purchased without any purification. Deionized water was obtained from Millipore RG system MilliQ (Milipore Corp., USA).

2.2. EXPERIMENTAL

The three-grid radio frequency-induced coupled plasma Kaufman ion beam source (Kaufman & Robinson–KRI®, Inc., USA) was employed to deposit titanium thin films with thickness of 500 nm

on 4-inch p-type silicon wafer (100) covered by 1 μm of thermal silicon dioxide. . The energy of Ar bombarding ions was set to 300 eV with corresponding deposition rate of 0.16 Å s^{-1}. The process pressure was 1.2 x 10^{-2} Pa.

Before anodic oxidation, this substrate was consecutively immersed into acetone, isopropanol, and deionized water to remove the impurities and then dried by compressed nitrogen. The anodization was realized in the two-electrode configuration and using power supply controlled by LabView program. The substrate operated as a working electrode and the stainless steel mesh was used as a counter electrode. The potential was kept constant during the whole time and different values in the range of 10–20 V were tested. The electrolyte solution was prepared from ethylene glycol containing ammonium fluoride (0.7–1 wt%) and deionized water (0–1 vol %). The anodized titania was rinsed with deionized water and dried by compressed nitrogen.

All TNTs samples were examined using scanning electron microscopy (FE Tescan Mira II LMU) under following conditions: work distance of 2.5 nm, high vacuum mode (10^{-3} Pa), voltage of 15 kV, and spot size of 2.4 nm.

3. RESULTS AND DISCUSSION

3.1. THE INFLUENCE OF ANODIZATION VOLTAGE

The oxidation potential is an important factor affecting the formation of TNTs. The highly ordered TiO_2 can be formed using suitable potential range, which is strongly given by used electrolyte. Every electrolyte has so called potential window. When the oxidation potential is lower than the potential window, only nanoporous structure is formed, however when the potential is too high, TNTs with residual oxide layer are formed (**Figure 2**b). For the selected electrolyte composed of 0.9 wt% NH_4F and 0.5 vol% H_2O, the optimal value of potential was 10 V , see **Figure 2**a.

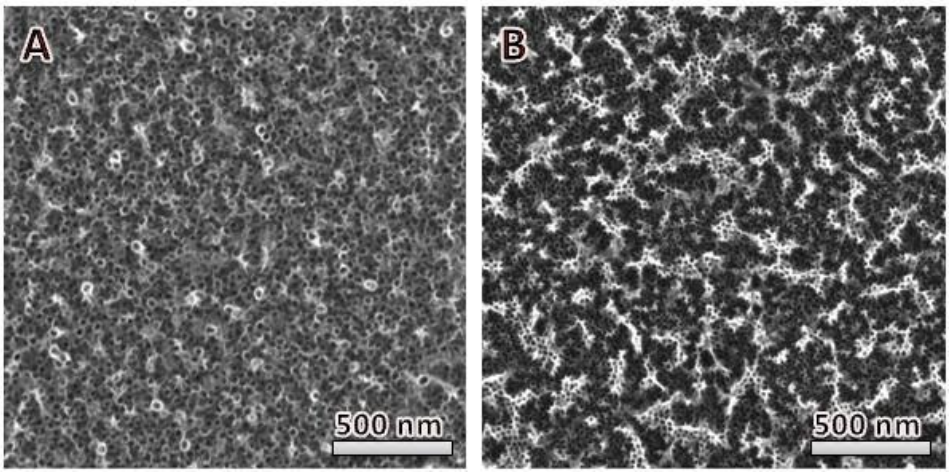


Figure 2: SEM images of TNTs created at voltages a) 10V and b) 20V.

3.2. THE INFLUENCE OF ELECTROLYTE COMPOSITION

The electrolyte composition determines the final morphology of TNTs. Namely, the balance between oxidation of Ti (water content) and dissolution of TiO_2 (amount of F^- anions) is very important. If the oxidation dominates in the process, water content is very high. This leads to disordered porous oxide layer (**Figure 3**a). If the higher amount of F^- anions is presented, the dissolution rate is faster than oxidation and resulting structure is only irregularly etched surface (**Figure 3**b). The electrolyte containing 0.9 wt% NH_4F and 0.5 vol% H_2O in ethylene glycol was used for anodization of 500 nm Ti thin film.

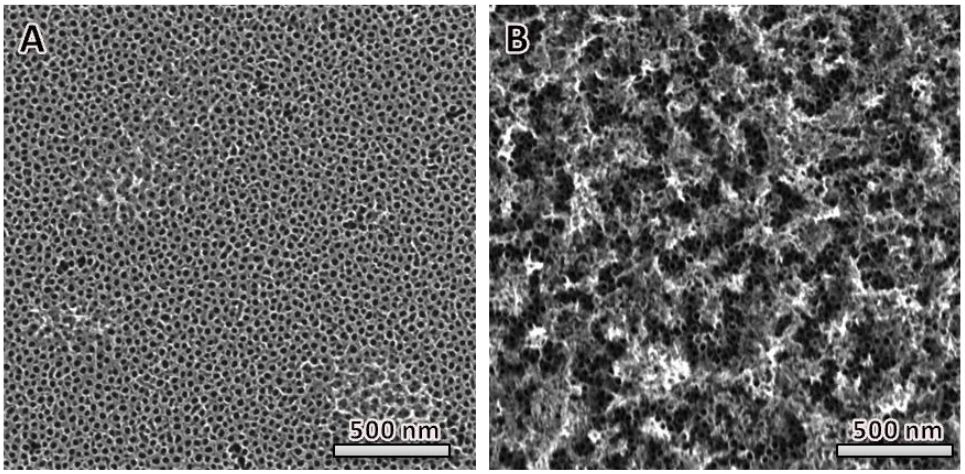


Figure 3: SEM images of TNTs created in electrolyte contaning high amount of a) water (1 vol%) and b) fluoride anions (1 wt%).

3.3. THE INFLUENCE OF ANODIZATION TIME

Another important factor for the nanotube formation is anodization time. If the time for oxidation is too short (tens of seconds), the electrochemical reaction has not reach the phase, when the fluoride anions diffuse into the pores and create F^- rich layer at the metal-oxide interface. Only few nanometers of oxide layer can be created (**Figure 4**a). If the anodization is stopped between 100–300 s, the TNTs were created but with initial oxide in tube's top. The anodization time also influenced the TNTs length (**Figure 4**b). According to our results, the appropriate anodization time for 500 nm layer was 500 s.

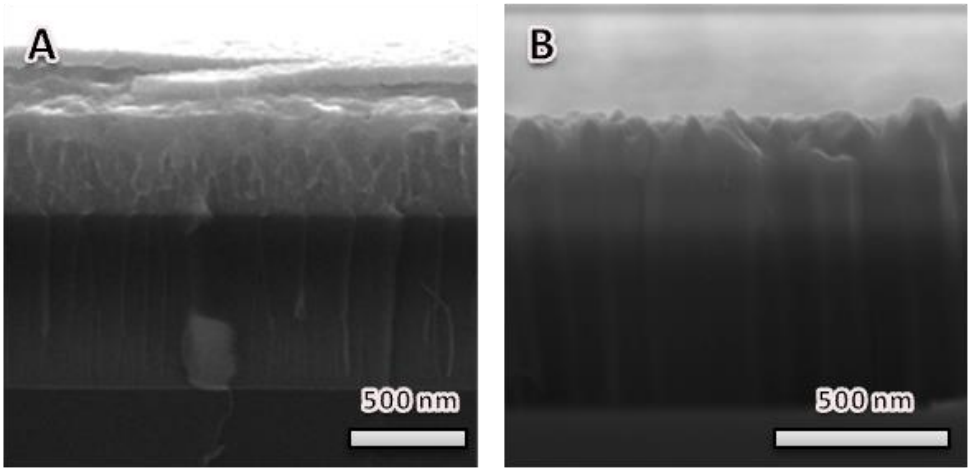


Figure 4: SEM images of anodization time about a) 50 s and b) 300 s.

4. CONCLUSION

The present work demonstrates a method for direct anodic transformation of Ti thin layer deposited on silicon wafer into nanotubular TiO_2 array. Simple anodization in electrolyte containing ethylene glycol and NH_4F allows the controlled fabrication of ordered and highly regular TiO_2 NTs. The explanation of anodization parameters, and SEM characterization of resulting nanostructures enables us to better understand how the NTs creation can be controlled. Fabricated TiO_2 NTs have very good potential for future sensing, biomedical and especially photocatalytic applications, mainly in advanced degradation of persistent organic pollutants in drinking water, etc.

ACKNOWLEDGEMENT

The research was supported by project no. FEKT-S-14-2300 A new types of electronic circuits and sensors for specific applications also National Sustainability Program under grant LO1401 is highly acknowledged. For the research, infrastructure of the SIX Center was used.

REFERENCES

[1] COLOMBO, A. et al. Highly transparent poly(2-ethyl-2-oxazoline)-TiO2 nanocomposite coatings for the conservation of matte painted artworks. *Rsc Advances*, 2015, vol. 5, no. 103, pp. 84879-84888.

[2] HUO, K. et al. Fabrication, modification, and biomedical applications of anodized TiO2 nanotube arrays. *Rsc Advances*, 2014, vol. 4, no. 33, pp. 17300-17324.

[3] QIN, L. et al. Effect of Anodization Parameters on Morphology and Photocatalysis Properties of TiO2 Nanotube Arrays. *Journal of Materials Science & Technology*, 2015, vol. 31, no. 10, pp. 1059-1064. ISSN 1005-0302.

[4] PIHOSH, Y. et al. Ubiquitous element approach to plasmonic enhanced photocatalytic water splitting: the case of Ti@TiO2 core-shell nanostructure. *Nanotechnology*, 2014, vol. 25, no. 31, pp. 8. ISSN 0957-4484.

[5] ZHANG, Q. et al. Anodic Oxidation Synthesis of One-Dimensional TiO2 Nanostructures for Photocatalytic and Field Emission Properties. *Journal of Nanomaterials*, 2014, pp. 14. ISSN 1687-4110.

[6] SOLOVEI, D. et al. Chemical Sensor Platform for Non-Invasive Monitoring of Activity and Dehydration. *Sensors*, 2015, vol. 15, no. 1, pp. 1479-1495. ISSN 1424-8220.

[7] DRBOHLAVOVA, J. et al. Self-ordered TiO2 quantum dot array prepared via anodic oxidation. *Nanoscale Research Letters*, 2012, vol. 7, pp. 4. ISSN 1931-7573.

[8] PINJARI, D. V. et al. Synthesis of titanium dioxide by ultrasound assisted sol–gel technique: Effect of calcination and sonication time. *Ultrasonics Sonochemistry*, 2015, vol. 23, no. 0, pp. 185-191. ISSN 1350-4177.

[9] LIN, X. X. a FU, D. G. Facile one-pot hydrothermal synthesis of B/N-codoped TiO2 hollow spheres with enhanced visible-light photocatalytic activity and photoelectrochemical property. *Solid State Sciences*, 2014, vol. 34, pp. 73-77. ISSN 1293-2558.

[10] MOROZOVA, M. et al. Electrochemical Properties of TiO2 Electrode Prepared by Various Methods. *Procedia Engineering*, 2012, vol. 42, pp. 573-580. ISSN 1877-7058.

[11] BESSERGENEV, V. G. et al. Preparation of TiO2 films by CVD method and its electrical, structural and optical properties. *Vacuum*, 2002, vol. 64, no. 3–4, pp. 275-279. ISSN 0042-207X.

[12] MAZZAROLO, A. et al. Anodic TiO2 nanotubes: Influence of top morphology on their photocatalytic performance. *Electrochemistry Communications*, 2012, vol. 22, pp. 162-165. ISSN 1388-2481.

[13] ROY, P. et al. TiO2 Nanotubes: Synthesis and Applications. *Angewandte Chemie International Edition*, 2011, vol. 50, no. 13, pp. 2904-2939. ISSN 1521-3773.

[14] FISCHER, K. et al. Nanoneedle and nanotubular titanium dioxide – PES mixed matrix membrane for photocatalysis. *Applied Catalysis B: Environmental*, 2014, vol. 160–161, no. 0, pp. 456-464. ISSN 0926-3373.

[15] PRIKRYLOVA, K. et al. Fabrication of highly ordered short free-standing titania nanotubes. *Monatshefte für Chemie - Chemical Monthly*, 2016.

MODELLING OF GAS DIFFUSION IN THE 3-D STRUCTURED METAL OXIDE NANOFILM

Zdeněk Pytlíček

Doctoral Degree Programme (5), FEEC BUT

E-mail: xpytli02@stud.feec.vutbr.cz

Supervised by: Alexander Mozalev; Jan Prášek

E-mail: alexander.mozalev@ceitec.vutbr.cz; prasek@feec.vutbr.cz

Abstract: Nanostructures of various materials and arrangements have been shown to be highly effective in the detection of chemical and biological matters. In this paper, a novel self-assembled 3-D nanofilm is described. We present a new computer-aided modelling approach applying a latest scientific knowledge in the field of microfluidics to improve a complex gas sensor design. The right questions are: What kind of 3-D structure is the best and how can we reach the best sensing parameters? Presented computer-aided modelling of gas diffusion in the 3-D nanofilm reveals an opportunity to substantially enhance sensor performance by the appropriate top electrode design.

Keywords: Gas sensor, 3-D nanostructure, Modelling, Diffusion, Microfluidics

1. INTRODUCTION

The ability to detect small concentrations of specific chemical substances is fundamental for a variety of industrial and scientific processes as well as for medical applications and environmental monitoring [1,2]. Semiconducting metal oxides can exhibit a conductivity changes due to the adsorption or reactions of molecules from the gas phase with the surface [3,4]. Over the last decades, scientists are ardently investigating new possibilities how to enhance all sensing parameters by using available nanotechnology, nanomaterials and nanostructures in general [5,6].

Nanostructured gas sensor developed by our research group is formed mainly by anodic oxidation of a valve metal and electrodeposition of a noble metal within anodized self-assembled nanoporous alumina template forming an array of spatially-ordered upright-standing Nb_2O_5 nanorods with gold 'nanocaps' (see Fig.1). These nanorods are placed in between two parallel patterned electrodes and work as long semiconducting nanochannels for chemisorption reactions in a gas atmosphere (Fig.1). Particular fabrications steps and the sensing properties have been described in recent publications [7-9].

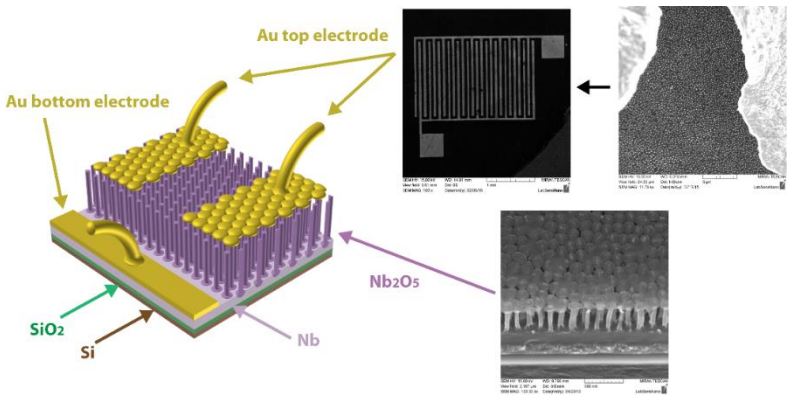


Figure 1: A 3-dimensional (3-D) nanostructured gas sensor with a Nb_2O_5 active layer [7-9].

During research and development of the new 3-D nanostructured gas sensor, basic questions about speed and magnitude of gas diffusion under the patterned top electrode have been introduced. A microfluidics theory has been included into our design process as a strong tool at the present scientific level.

2. A TINY MICROFLUIDICS BACKGROUND

Theoretical microfluidics deals with the theory of flow of fluids and of gases in micro- and nano-sized systems influenced by external conditions. Recent development of microfluidics has been enhanced mainly thanks to the lab-on-a-chip (LOC) systems and microelectromechanical systems (MEMS) [10-11].

2.1. SCALING LAWS IN MICROFLUIDICS

Scaling law indicates that when scaling down to the microscale in LOC or MEMS, the volume forces, which are very common in our daily life, become rely unimportant. Instead, the surface forces become dominant, and as a result, we must rebuild our imagination and be prepared for some surprises on the way [10-11]. The physics of gas interaction with gas sensor differs from the continuum flow problems. At the nanometer scale, the mean free path of gas molecules becomes comparable with the system size, and gas begins to act as rarefied. As quick method to quantify proper flow regime, the Knudsen number can be computed. The Knudsen number is a dimensionless number defined as

$$Kn = \frac{\lambda}{L} \quad [-], \tag{1}$$

where λ is the mean free path of gas molecules, and L is the representative physical length scale of the system. This number indicates what type of flow is considerable as it is shown in Table 1 [10-11].

(Kn < 0.01)	Continuum flow
(0.01 < Kn < 0.1)	Slip flow
(0.1 < Kn < 10)	Transitional flow
(Kn > 10)	Free molecular flow

Table 1: Flow regimes according to Knudsen number [10-11].

3. EXPERIMENTAL

According to the equation (1), the encountered flow regime for our gas sensor (see Fig. 1) is Free Molecular Flow where the gas molecules interact with the surfaces only, and where no collision between the molecules themselves occurs. For modeling purpose, the Free Molecular Flow module in COMSOL Multiphysics 5.0 software (COMSOL, Inc., US) was used. The module employs finite element method (FEM). An issue is a modeling setup by specifying the appropriate boundary and initial conditions. A various top electrode designs and prepared mashes for the solver are shown in Figure 2.



Figure 2: Two-dimensional meshes created for finite element analysis: (a) a solid top electrode, b) a perforated top electrode, (c) a porous material.

The used model requires the Particle Tracing Module. Particle trajectories are computed using Newton's law of motion

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(m\frac{\mathrm{d}q}{\mathrm{d}t}\right)=0\,, \tag{2}$$

where t is the time, m is the mass of the molecule and q is the kinetic energy of the molecule. The particle velocity is given by

$$v=\frac{\mathrm{d}q}{\mathrm{d}t}\,. \tag{3}$$

The transmission probability for the molecular flow interface is given by the following equation:

$$\chi=\frac{\int_{outlet}(J-G)\mathrm{d}l}{\int_{inlet}(J-G)\mathrm{d}l}\,, \tag{4}$$

where G is the molecular flux striking a surface, J is the outgoing flux from a surface and l is the path.

The speed constants for adsorption and desorption of molecules on the surface [12-14] are also accounted for evaluation in this models, as the delay constants play marginal role in the diffusion process.

4. RESULTS AND DISCUSSION

4.1. PARTICLE TRACING

Particle tracing provides a description of a problem by solving ordinary differential equations using Newton's law of motion (2). This model requires specification of the particle mass, and all forces acting on the particle. The results from direct simulation Monte Carlo (DSMC) with randomly spread CO_2 molecules over a 3-D structured sensor covered by the top electrode are shown in Figure 3. The mean speed of molecules was calculated for a pressure of 100 kPa and a temperature of 300 K. The distribution of particle trajectories and their speeds are depicted in Figure 3a. The relative magnitude of impacted surface is shown in Figure 3b. This simulation may reveal most exposed surface, which plays a crucial role in the response rising in the beginning of the gas detection.

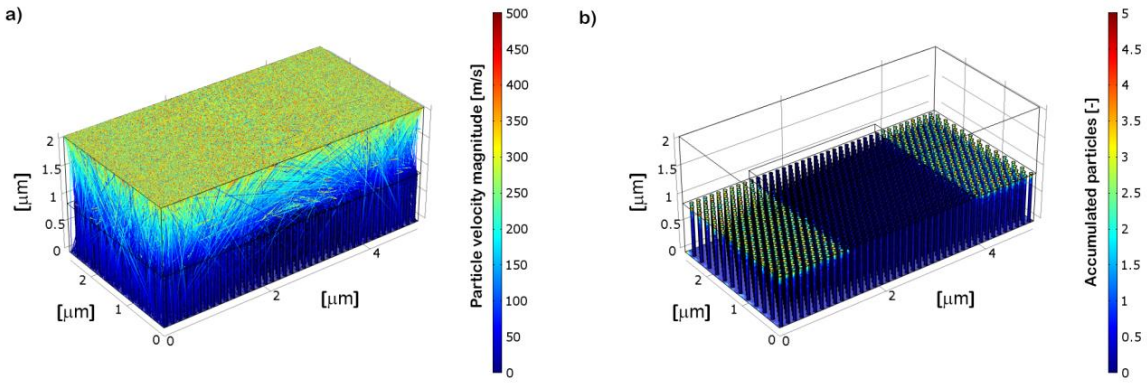


Figure 3: MonteCarlo 3-D CO_2 particle trajectories tracking (pressure 100 kPa, temp. 300 K).

4.2. FREE MOLECULAR FLOW IN THE 3-D NANOFILM

The models were developed for two types of the top electrode layout and for the porous material with the equivalent pore diameter of 150 nm. The simulated gas propagation over the complete 50 μm wide solid lines, as it was formerly used in our sensor design, is depicted in Figure 4a. The simulated gas propagation for the perforated (hole-having) line is shown in Figure 4b. From the results it is clear that, after about 15-20 min a small number of the columns is still affected by gas of a relatively high concentration. Evacuation of the target gas should lead to even slower outflow and resulting recovery time. This may be the reason for a longer recovery time of our 3-D nanostructured gas sensor in comparison with the faster response time. A time propagation of diffusion for three types of top electrode structure are shown in Figure 5.

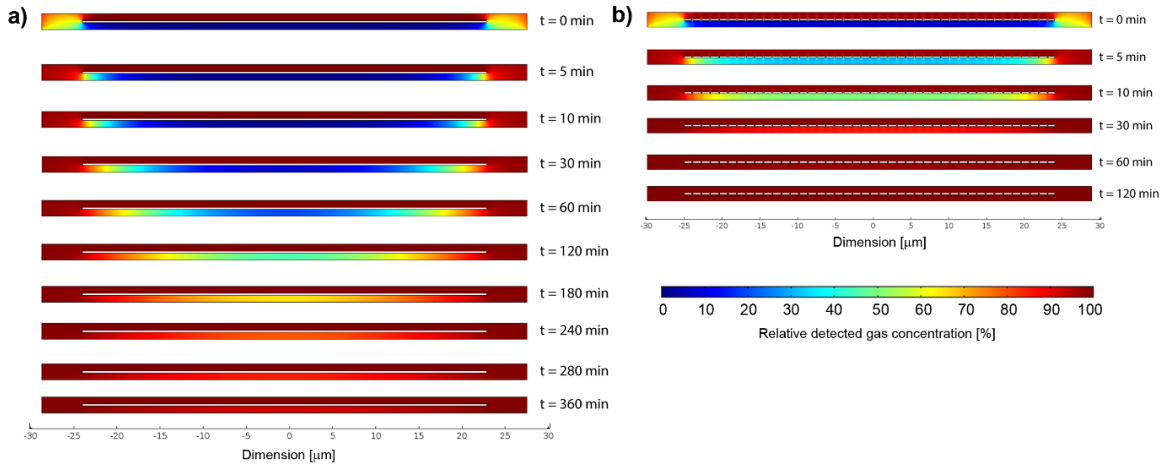


Figure 4: Gas propagation in time under (a) the solid top electrode, (b) the perforated electrode (pressure 100 kPa, temperature 300 K).

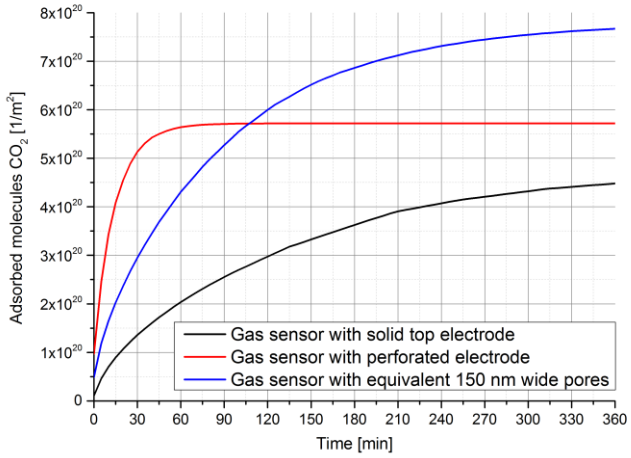


Figure 5: Gas propagation in time for three different type of the nanostructured gas sensors (pressure 100 kPa, temperature 300 K).

5. CONCLUSION

The simulation study of gas diffusion in the 3-D structured metal oxide nanofilm has been carried out. This study reveals how the diffusion under the top electrode affects the penetration of gas to the sensor active layer. Our results would allow to establish advanced design rules for the nanostructured gas sensors. Computer-aided modelling and simulation is still in progress including more variables for a better interpretation of gas behavior on the surface [12-14] to enhance the

performance of the nanostructured gas sensors. Samples with an adequate design of the top electrode are prepared to validate this simulation results. COMSOL software is also possible to be used to study how the changes of surface charges may contribute to the electric conductivity.

ACKNOWLEDGEMENT

The article was supported by project no. FEKT-S-14-2300 A - new types of electronic circuits and sensors for specific applications.

REFERENCES

[1] McGrath, M., Scanaill, C. N., Sensor technologies: healthcare, wellness, and environmental applications, New York, Springer Science Business Media 2014, ISBN 14-302-6013-0.

[2] Fleischer, M., Lehmann, M., Solid State Gas Sensors – Industrial Application, Brlin, Springer Berlin Heidelberg 2012, ISBN 978-3-642-28092-4.

[3] Korotcenkov, G., Metal oxides for solid-state gas sensors: What determines our choice?, Materials Science and Engineering B-Solid State Materials for Advanced Technology, 2007, vol. 139, DOI 10.1016/j.mseb.2007.01.044.

[4] Comini, E., Faglia, G., Sberveglieri, G., Solid State Gas Sensing, New York, Springer 2009, ISBN 978-0-387-09664-3.

[5] In, H. J., Field, C. R., Pehrsson, P. E., Periodically porous top electrodes on vertical nanowire arrays for highly sensitive gas detection, Nanotechnology, 2011, vol. 22, DOI 10.1088/0957-4484/22/35/35501.

[6] Sun, Y. F., Liu, S. B., Meng, F. L., et al., Metal Oxide Nanostructures and Their Gas Sensing Properties: A Review, Sensors, 2012, vol. 12, p. 2610-2631, DOI 10.3390/S120302610.

[7] Mozalev, A., Pytlíček, Z., Bendová, M., Vazquez, R. M., Llobet, E., Hubálek, J.: 3-D nanostructured tungsten-oxide gas-sensing film via anodizing sputter-depositied Al/W metal layers, IEEE Sensors 2015, p. 170-173, DOI 10.1109/ICSENS.2015.7370205.

[8] Mozalev, A., Vazquez, R. M., Bendová, M., Pytlíček, Z., Llobet, E., Hubálek, J.: Porous-alumina-Assisted Formation of 3-D Nanostructured Niobium Oxide Films for Advanced Sensing Applications, Procedia Engineering, 2015, vol. 120, p. 435-438, DOI 10.1016/j.proeng.2015.08.660.

[9] Mozalev, A., Bendová, M., Vazquez, R. M., Pytlíček, Z., Llobet, E., Hubálek, J.: Formation and gas-sensing properties of a porous-alumina-assisted 3D niobium-oxide nanofilm, Sensors and Actuators B: Chemical, 2016, vol. 229, p. 587-598, DOI 10.1016/j.snb.2016.02.024.

[10] Bruus, H., Theoretical microfluidics, Oxford, Oxford University Press 2008, ISBN 978-0-19-923508-7.

[11] Karniadakis, G., Beskök, A., Aluru, N. R., Microflows and nanoflows: fundamentals and simulation, New York, Springer 2005, ISBN 978-0387-22197-7.

[12] Dabrowski, A., Adsorption – from theory to practice, Advances in Colloid and Interface Science, 2001, vol. 93, p. 135-224, DOI 10.1016/S0001-8686(00)00082-8.

[13] Ducere, J. M., Hemeryck, A., Esteve, A., Rouhani, M. D., et al., A computational chemist approach to gas sensors: Modeling the response of SnO2 to CO, O2, and H2O Gases, Jurnal of Computational Chemistry, 2012, vol. 33, p. 247-258, DOI 10.1002/Jcc.21959.

[14] Roque-Malherbe, R. M. A., Adsorption and diffusion in nanoporous materials, Boca Raton, CRC Press 2007, ISBN 978-1-4200-4675-5.

LOW VOLTAGE CONVERTER FOR ORGANIC TRANDUCERS IN ENERGY HARVESTING

Laila Znbill

Doctoral Degree Programme (2), FEEC BUT

E-mail: xznbil01@stud.feec.vutbr.cz

Supervised by: Jaroslav Boušek

E-mail: bousek@feec.vutbr.cz

Abstract: The paper describes the possibilities of energy harvesting using organic semiconductors. It is shown that organic semiconductors enable a large variability and some of their properties may be better than by inorganic semiconductor materials. Simple and reliable JFET driven DC to DC converter was designed for low voltage applications as thermoelectric generators and single-cell solar panels. It is able to operate from the voltage of few tenths of Volt and the output voltage could be in the level of several Volts. The efficiency was found close to 50%. Material and production costs are significantly lower than in case of standard low power integrated circuits.

Keywords: Organic semiconductors, Energy harvester, low power DC/DC converter.

1. INTRODUCTION

Energy harvesting is becoming more feasible today because of the increased efficiency of devices used to capture, store, and produce electrical energy. Nevertheless, many energy harvesting transducers provide a low output voltage not sufficient for electronic devices. For DC sources, like solar cells and thermal generators, there is therefore a need to slowly boost the voltage to a level sufficient for opera- tion of regular converters or low voltage circuits.

2. ORGANIC TRANSDUCERS FOR ENERGY HARVESTING

For electronic properties of organic materials there is crucial the existence of conjugated bond system where carbon atoms covalently bond with alternating single and double bonds. Hydrocarbon electrons delocalize and form a delocalized bonding π orbital with a π* antibonding orbital [1].

The delocalized π orbital is the highest occupied molecular orbital (HOMO), and the π* orbital is the lowest unoccupied molecular orbital (LUMO). In organic semiconductor physics, the HOMO takes the role of the valence band while the LUMO serves as the conduction band. The energy separation be- tween the HOMO and LUMO energy levels is considered the band gap of organic electronic materials and is typically in the range of 1eV to 4 eV [2].

2.1. ORGANIC PHOTOVOLTAIC CELLS

When organic material absorbs a photon, an excited state is created and confined to a molecule or to a region of polymer chain. The excited state can be then regarded as an exciton, or an electron-hole pair bound together by electrostatic interactions.

In organic solar cells (OSC) excitons are broken up into free electron-hole pairs by effective fields. The effective fields are set up by creating a heterojunction between two dissimilar materials. In this process the electron falls from the conduction band of the absorber to the

conduction band of the acceptor material. It is necessary that the acceptor material has conduction band edge that is lower than that of the absorber material [3].

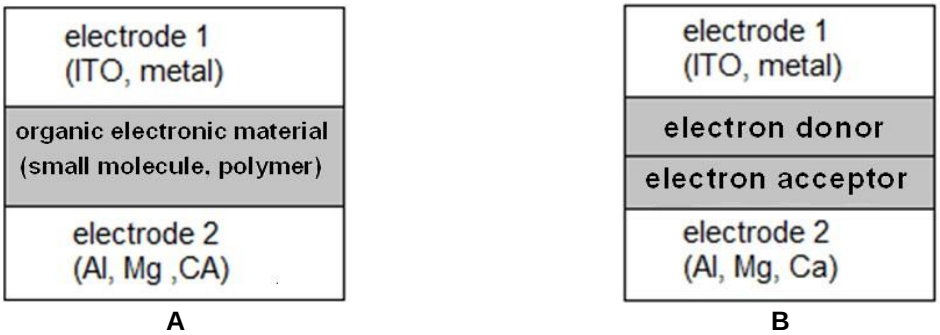


Figure 1: A) Single layer organic photovoltaic cell; B) Bilayer organic photovoltaic cell.

Single layer OSC. Single layer organic photovoltaic cells are the simplest form. These cells are made by sandwiching a layer of organic electronic materials between two metallic conductors, typically a layer of indium tin oxide (ITO) with high work function and layer with low work function metal such as Aluminum, Magnesium or Calcium. The difference of work function between the two conductors sets up an electric field in the organic layer. When the organic layer absorbs light, electrons will be excited to the LUMO and leave holes in the HOMO, thereby forming excitons. The potential created by the different work functions helps to split the exciton pairs, pulling electrons to the positive electrode and holes to the negative electrode. See Figure. 1A).

Bilayer OSC. Bilayer organic solar cells contain two layers in between the conductive electrodes. The two layers have different electron affinity and ionization energies, therefore electrostatic forces are generated at the interface between these two layers. The materials are chosen to make the potential differences large enough to split excitons much more efficiently than in case of single layer photovoltaic cells. The layer with higher electron affinity and ionization potential is the electron acceptor, and the other layer is the electron donor. This structure is also called a planar donor-acceptor heterojunction. See Figure. 1B).

2.2. ORGANIC THERMOGENERATORS

Thermoelectric modules work on two basic principles which are related one to another:

Seebeck effect. This effect creates potential difference across the connection of two materials module by heating one side of the module and cooling the opposite.

Peltier effect. When current passes through the junction of different materials one side is cooling and the other side is heating. Peltier cells are to time routinely used for cooling. Consequently, when we will deliver thermal energy to this junction we may produce electricity.

2.2.1. Thermoelectric power conversion

There are two critical factors for efficient thermoelectric power conversion. Firstly, the amount of heat flux must be successfully moved to the module. Secondly, the thermal conductivity of module material between hot and cold side must be as low as possible. Low thermal conductivity is a problem by almost all inorganic semi-conductive materials. Organic semiconductors could have the thermal conductivity almost one order lower. In any case the basic parameters there are the difference in work functions and Seebeck and Peltier coefficients [4]. See the Figure 2 for reference.

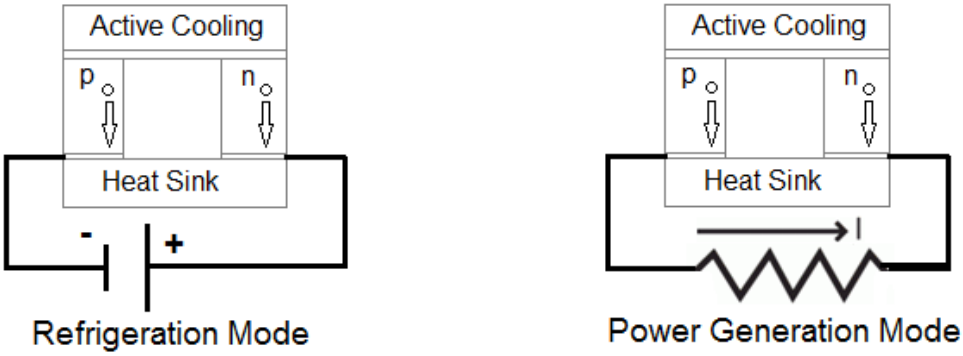


Figure 2: Peltier cell in power generation and thermoelectric cooling

2.3. LOW VOLTAGE CONVERTERS AVAILABLE AT THE MARKET

For utilization of organic thermo-generators it is necessary to consider that the output voltage of one thermoelectric element is very small. Even with the integration of many thermoelectric elements in series is not possible to increase the voltage to a level applicable for powering conventional electronic circuits. A similar situation can occur with single-cell solar panels applications. For optimal use of energy harvesters is therefore necessary to use a voltage converter with extremely low power supply voltage. Low voltage converters therefore implement boost converter that requires only microwatts of power to begin the operation.

The boost converter implemented in integrated circuit BQ25504 from Texas Instruments can start with input voltage as low as 330 mV. Once started, it can operate down to input voltage 80 mV.

Other types for example LTC3107, LTC3108 and LTC3109 from Linear technology, or MAX17710 from Maxim take power supply from output voltage. Therefore, to start up they usually need the external battery. Advantage of this design is that the input voltage could be extremely low. Basic properties of these circuits are listed in Table 1.

TYPE	Start Voltage	Lowest Voltage	Need power supply	Efficiency / Standby	
BQ25504	330 mV	80 mV	NO	> 50%	0,33 µA
LTC3108	-	20 mV	YES	> 50%	6 µA
MAX17710	-	750 mV	YES	> 50%	1,3 µA

Table 1: The overview of low voltage converters on the market

3. DESIGN OF J-FET LOW VOLTAGE CONVERTER

The biggest advantage of Organic Energy Harvesters (OEH) is their low price. The price of the necessary DC/DC converter may then be higher than the price of OEH alone.

The unit price of low voltage integrated circuits is around 50 CZK (US $ 2) and the price of other components included SMT PCB can be estimated at around 25 CZK (US $ 1). Simple and reliable JFET driven converter suggested here has material and production costs significantly lower.

3.1. PRINCIPLE OPERATION

Block diagram of converter is on Figure 3. The transformer (2) is used to increase the voltage at the inverter output. This transformer is also used to control the switching element (1). The output voltage is taken between the capacitors (3) and (4). The energy harvester (5) supplies a voltage less than 1 V. The voltage level to operate the device (6) at the inverter output is several volts.

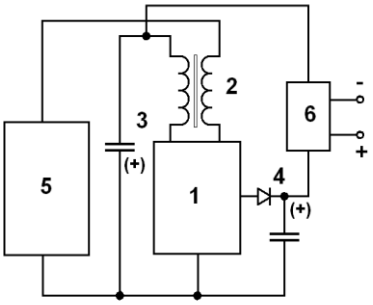


Figure 3: Block diagram of converter

Fundamental scheme of the inverter is shown in Figure 4 A). The transistor JFET operates as a switching element. As soon as the transistor turns on the secondary winding generates a voltage pulse. Capacitor C_1 is being charged by means of this pulse. The junction between the Gate electrode and Source of the transistor in this case operates as a rectifying diode. Because in this state the GateSource junction of JFET is in forward polarization the width of this junction is very small and there is very high conductivity across the transistor – it means between the electrodes S and D. Once the transformer core becomes saturated the voltage on the secondary winding starts to drop. Due to the positive feedback given with actual polarity of primary and secondary windings the transistor closes. With a negative voltage on the capacitor C_1, JFET is maintained in a closed state until the next part of the cycle where it passes into the on-state and consequently the whole process is repeated. The voltage on the capacitor C1 is at the same time output voltage of the converter as a whole.

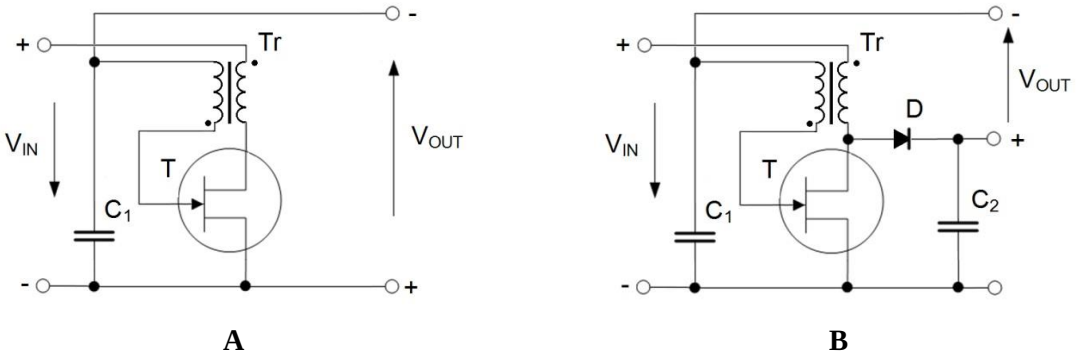


Figure 4: A) Basic scheme of the converter B) Converter with increased output voltage

The modified circuit shown in Figure 4 B) utilizes the energy stored in the transformer during the switch-on state. As soon as the switch-off starts a voltage pulse appears on the primary winding, which charges the capacitor C_2. This is due to the drop of magnetizing current and subsequent collapse of the magnetic field in the transformer core.

With optimal settings of the circuit the positive voltage on capacitor C_2 has approximately the same size as negative voltage on capacitor C_1. The output voltage can therefore be approximately twice as large. At the same time, there will be a slight increase in efficiency. At high current consumption, however, the output voltage decreases rapidly.

It was verified that the circuit can operate in a large range of frequencies. The working frequency of the circuit is determined by the transformer. To ensure high efficiency, it is necessary that the trans- former should minimize leakage inductance.

At low frequencies, the switching losses are small but the transformer is bulky and it is a problem to achieve a small leakage inductance. When using a toroidal transformer with a diameter of 8 mm the operating frequency of the circuit was 700 kHz. In this case, the efficiency has been limited by dynamic losses on the switch (JFET) and on the diode D.

4. RESULTS AND DISCUSSION

In Table 1 there are results of testing of inverter operated at frequency 700 kHz for both modifications of the inverter according to Figure 4. The threshold voltage V_T of JFET was $V_T = 2$ V. The transformer secondary to primary ratio was $N_S:N_P = 6:1$. For the converter, which utilizes energy stored in the transformer core, the efficiency is slightly higher. The output voltage is here greater however the cur- rent capacity is smaller.

MOD	Start Voltage	Stop Voltage	Output Current / Voltage	Power / Efficiency	
A	0,5 V	0,3 V	3V / 100 μA	300 μW	45 %
B	0,6 V	0,3 V	4 V/ 80 μA	320 μW	50 %

Table 1: Modifications of the circuit A) B); measured parameters.

4.1. CIRCUIT EFFICIENCY AND POSSIBILITIES OF FURTHER IMPROVING

To ensure high efficiency the transformer must have minimal leakage inductance. To limit the impact of dynamic losses the operating frequency of 100 kHz should not be exceeded. Here it is necessary to find a compromise between the size of the inverter and its effectiveness. Start up voltage could be much lower by higher transformer ratio and/or when using the transistor with lower threshold voltage. In both cases however the energetic efficiency of the circuit drops.

5. CONCLUSION

The paper describes the possibilities of organic energy harvesters. It is shown that organic semiconductors enable a large variability of properties and some of their properties may be better than by standard semiconductor materials. This applies especially for organic semiconductors for use in thermo-generators. To achieve long-term reliability of such devices, however, further development is needed.

DC to DC converter was designed for low voltage applications as thermoelectric generators and single-cell solar panels. It is able to operate from the voltage of few tenths of Volt and the output voltage could be in the level of several Volts. Achieved efficiency of the apparatus is approximately 50%, but it could be possible to increase it in further optimization, particularly in the design of the transformer.

ACKNOWLEDGEMENT

The article was supported by project no. FEKT-S-14-2300 "A new types of electronic circuits and sensors for specific applications".

REFERENCES

[1] Scharber M.C. et al.: "Design Rules for Donors in Bulk-Heterojunction Solar Cells – Towards 10 % Energy-Conversion Efficiency", Adv. Mater. 2006, 18 , pp 789–794

[2] I.G. Hill at al.: "Charge-separation energy in films of -conjugated organic molecules", Chemical Physics Letters 327, 2000, pp181–188

[3] Harald Hoppe et al.: "Organic solar cells: An overview" Journal of Materials Research, 2011, Volume 19, Issue 07, pp 1924-1945

[4] Bubnova O. Crispin X.: "Towards polymer-based organic thermoelectric generators", Energy & Environmental Science, 2012, ISSN 1754-5692, Vol. 5, no 11, pp 9345-9362

ELECTROCEMICAL DEPOSITION OF GOLD NANORODS VIA AAO TEMPLATE FOR VOCS SENSING

Milena Šetka

Doctoral Degree Programme (1), CEITEC VUT E-mail: milena.setka@ceitec.vutbr.cz

Supervised by: Jana Drbohlavová

E-mail: jana.drbohlavova@ceitec.vutbr.cz

Abstract: The use of gold nanostructures for the creation of sensing devices for volatile organic compounds (VOCs) is particularly important because of their unique electronic, optical, and catalytic properties. In this paper, the synthesis of gold nanorods was described. The preparation of nanoporous alumina template was done via electrochemical anodization. The fabrication of gold nanorods was provided via electrochemical deposition into the nanoporous alumina template. Working conditions during the anodization had a significant influence on the pore diameter and thus on size of gold nanorods. The detailed topography of the nanostructured surfaces was characterized by scanning electron microscopy (SEM).

Keywords: Gold nanorods, electrochemical anodization, electrochemical deposition, anodization voltage, pore geometry, VOCs sensors

1. INTRODUCTION

Gold nanostructures have attracted great attention due to their potential application in chemical and biochemical sensing, medical diagnostics and therapeutics, and biological imaging because of their unique optical and electrical properties [1]. Gas sensors based on gold nanostructures have high sensitivity and selectivity, good reproducibility and long-term stability [2]. Geometrical parameters and size of nanostructures play an important role in performances of a gas sensor [3]. Among various shaped gold nanoparticles, nanorods and nanofibers have received the most attention, especially in the field of optical sensors. When gold nanorods are exposed to laser light resonant with their surface plasmon oscillation, they can strongly absorb the light and rapidly convert it into heat via a series of photophysical processes, what is suitable for practical application. The intensity and wavelength of these surface plasmon resonances is highly shape and size dependent [4].

Another important characteristic of gold is that it can bind thiols with high affinity and it does not undergo any unusual reactions with them, e.g., the formation of a substitutional sulphide interphase [5]. This characteristic of gold is important for gas sensing. Gold nanoparticles covered with thiols are used like sensitive layer in chemiresistive sensors for detection of VOCs. VOCs have been found as significant biomarkers in exhaled breath, which reflect the biochemical alterations related to metabolic changes. Therefore, they are extremely helpful in diagnosis and monitoring of various diseases [6]. Those kind of sensors have a very high sensitivity, they are able to give response with a wide range of concentrations from tens of ppb to hundreds of ppm, with a detection limit of 1–5 ppb. Thiols help making a surface of sensitive layer hydrophobic. The sensor shows a low sensitivity to water and it is particularly suitable for breath testing, which is an advantage since the exhaled breath contains ~ 80% relative humidity [7].

A combination of nanostructures with conducting polymers could improve characteristic of gas sensors, by giving them properties such as faster response, long-term stability and high sensitivity [8]. Lee et al. developed VOCs gas sensors based on metal-organic framework of gold–polypyrrole (Au–Ppy) nanorods, see Figure 1 [9]. This gas sensor operates on the principle of localized surface plasmon resonance. The authors used an electrochemical deposition method for fabrication of Au–

Ppy nanorods via anodic aluminum oxide template with 200 nm diameter of pores. This kind of sensor shows a good stability and high sensitivity up to 10 ppm of target VOCs gases [9].

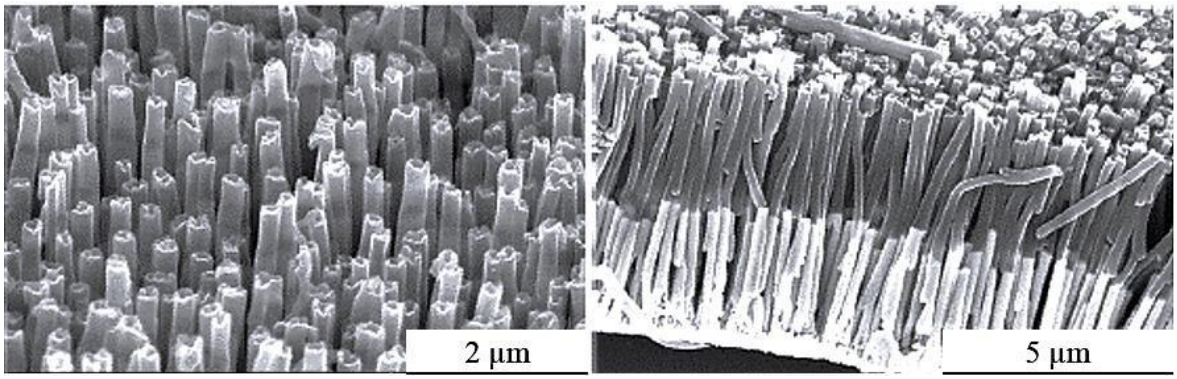


Figure 1: SEM images of the Au-Ppy nanorods [9]

Nanostructures can be synthesized by wet-chemical synthesis or fabricated by templates such as nanoporous alumina template. It can be prepared by electrochemical anodization of aluminum, which allows to produce uniform pores with the diameter as small as 5 nm [10]. An acidic solution is usually used as an electrolyte. A length of nanopores is dependent on the anodization time. An electrochemical deposition of metal into pores leads to the growth of metal nanorods and nanowires [11]. The fabrication using anodic alumina oxide is less complicated, less expensive, and less time–consuming when compared to traditional lithographic approach or epitaxial growth of highly ordered arrays of nanostructures [12].

This work describes a developing of gold nanorods by electrochemical anodization and electrochemical deposition. It shows a different working conditions during the anodization, which have significant role for the size of the gold nanorods. The anodization voltage together with the current density and the electrolyte type are recognized as the most critical parameters to control the pore geometry [13].

2. EXPERIMENTAL

2.1. MATERIALS

Titanium (99.99%), tungsten (99.999%), purchased from Porexi, Czech Republic, and aluminium (99.999%), purchased from Goodfellow, United Kingdom, were used for the PVD deposition of metallic layers. For electrochemical fabrication of gold nanostructures were used: oxalic acid, sulphuric acid (96%), boric acid, chromium trioxide, phosphoric acid (98%), sodium dihydrogen phosphate dihydrate (99%), obtained from Penta, Czech Republic, potassium dicyanoaurate (gold 68.3%, purity 99.9%), obtained from Safina, Czech Republic, and disodium hydrogen phosphate dihydrate (98%), obtained from Fluka, Czech Republic. Deionized water (18.2 MΩ) was bought from Millipore RG system MilliQ (Millipore Corp., USA).

2.2. PREPARATION OF THIN LAYERS FOR FABRICATION OF NANORODS

The silicon wafer was covered by SiO_2 layer with thickness of 1 μm to isolate the nanostructures from n–doped silicon and thus to minimize junction effects at the interface between Si and other metal layers [12]. Titanium layer with a thickness of 20 nm was sputtered on a silicon wafer. Titanium layer serves as an adhesive layer and does not change during the electrochemical fabrication of gold nanorods. The 150 nm tungsten layer was deposited by the ion beam sputtering method. The aluminium layer with a thickness of 300 nm was obtained by thermal evaporation.

2.3. CHARACTERISTIC OF PREPARED GOLD NANORODS

The morphology of nanostructures, homogeneity of surface coverage, and the nanostructures size were studied by SEM (Tescan Mira II, Tescan, Czech Republic).

2.4. ELECTROCHEMICAL FABRICATION OF NANOSTRUCTURED SURFACES

In the first step, the aluminium layer was transformed by anodic oxidation to nanoporous alumina template (Al_2O_3). The anodization process was done as described in the following paper[14], under constant voltage, either at 53 V in 0.3 M oxalic acid or at 20 V in 2 M sulphuric acid. In both cases, the temperature of acid electrolyte was kept at 10 °C. In the second step, when the anodization of aluminium was finished, the bottom tungsten layer was oxidized into the tungsten trioxide nanostructures. In the third step, the WO_3 nanostructures were selectively etched in the phosphate buffer solution at pH= 7. The temperatures of the etching solution were: 25 °C, 40 °C and 60 °C. After etching of WO_3, tungsten nano-dimpled structure was created on the surface. Those nanodimples are a base for gold nanorods and they help to improve their stability. In the fourth step, the gold was electrodeposited from potassium dicyanoaurate solution into the alumina oxide template with the nanodimpled bottom. The conditions were set up during this process as follows: the constant current of 1 mA, the pulse length of 400 ms, the period between pulses of 2 s, the amplitude of 5 V, and the number of pulses 45. In the fifth step, the alumina template was selectively etched in the mixture of chromium trioxide and phosphoric acid at 60 °C for 10 min. All those steps are schematically shown in Figure 2 [12].

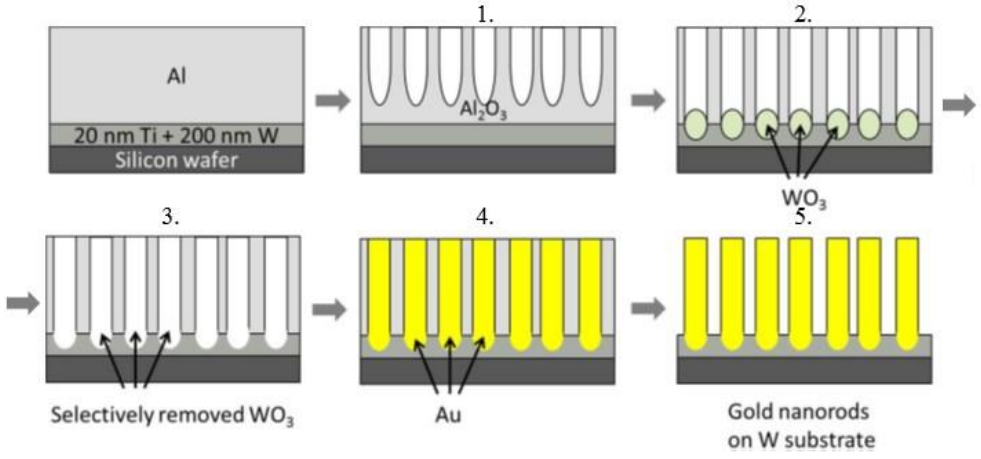


Figure 2: The fabrication process of gold nanorods [10]

3. RESULTS AND DISCUSSION

The analysis confirmed that the anodization of aluminium layer in different acid electrolytes gives the different pore geometry of the alumina template. In 0.3 M oxalic acid and at applied voltage of 53 V was obtained the nanopores with the pore diameter of about 50–60 nm. The pore length depends on the thickness of the alumina layer, it was ~ 300 nm. The interpore distance is similar to the pore diameter, i.e. ~ 50–60 nm. In 2 M sulphuric acid and at applied voltage of 20 V, the nanopores have the size of pores of around 25 nm and similar interpore distance, which is two times smaller than in the previous case. These results confirmed the very well–known theory, the anodization voltage and the electrolyte type are the most important parameters to control the pore geometry of the alumina nanoporous template. The Figure 3 shows SEM images of template after etching of Al_2O_3 and WO_3 nanostructures, anodized in oxalic acid (left) and sulphuric acid (right). On the left, half of WO_3 nanostuctures was not etched, because the phosphate buffer solution did not reach that part. The influence of etching is clearly seen.

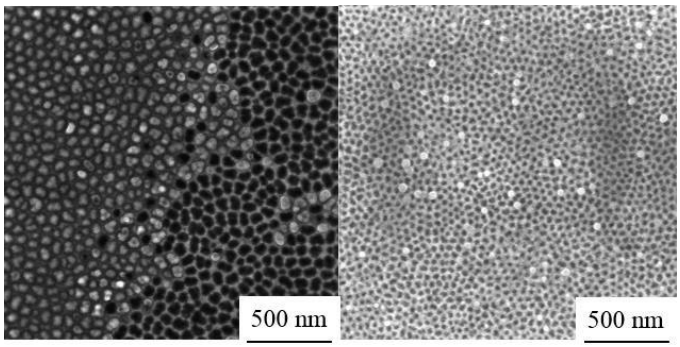


Figure 3: SEM images of template after etching of WO_3 and Al_2O_3 nanostructures, anodized in: 3 M oxalic acid at applied voltage of 53 V (left) and 2 M sulphuric acid at applied voltage of 20 V (right);

The results of selective etching of tungsten oxide nanostructures can be seen in Figure 4. The time of etching process was 60 min when temperature of etching solution was 25 ℃ (left) and 40 ℃ (middle), and 45 min when temperature was 60 ℃ (right). The electrochemical anodization of those samples was done in the sulphuric acid.

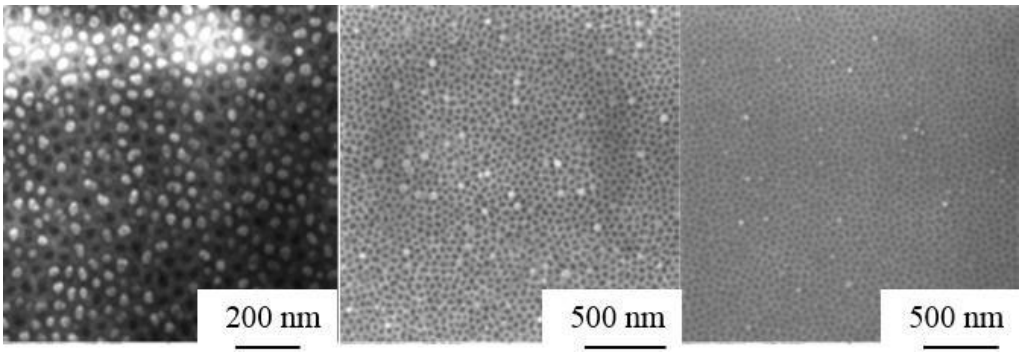


Figure 4: SEM images of etched WO_3 nanostructures in the phosphate buffer solution, pH= 7; A) etching temperature (T_e)= 25 °C, etching time (t_e)= 60 min; B) (T_e)= 40 °C, te= 60 min; C) (T_e)= 60°C, t_e= 45 min;

The SEM characterization shows that the increasing temperature strongly affected the dissolution of WO3 nanostructures. The best dissolution of tungsten oxide nanostructures was reached when temperature was 60 °C. The increase in temperature successfully decreased the etching time to 45 min at 60 °C. The final dimpled surface was nicely homogenous and can serve for electrochemical deposition of gold nanorods. The electrochemical deposition of gold into nanoporous alumina template was successfully done under above mentioned conditions. It was possible to obtain gold nanorods with the uniform size, the diameter of ~ 50–60 nm. The homogeneous surface of gold nanorods is illustrated in Figure 5.

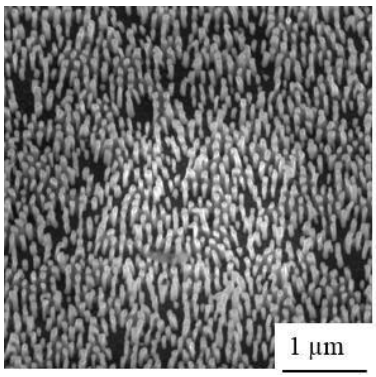


Figure 5: SEM image of gold nanorods prepared via electrochemical deposition

4. CONCLUSIONS

In summary, this work has demonstrated the preparation of nanoporous alumina template via electrochemical anodization of aluminum, and successfully synthesized gold nanorods via gold pulse deposition into the template. Sulphuric acid comparing to oxalic acid greatly decreased the pore diameter of nanoporous alumina template. In future work, gold nanorods will be obtained in the nanoporous alumina template created by sulphuric acid. Higher temperature of etching solution for WO_3 nanostructures perfectly improves the dissolving of WO_3 structures. This helps to create the nanodimpled surface and enables to successfully deposit gold into dimples via alumina template. With the rapid developments in nanorods preparation, surface modification, and assembly, the emergence of novel applications of gold nanorods are likely to impact the areas of VOCs sensing.

REFERENCES

[1] LEE, K.-S. a EL-SAYED, M. A. Gold and silver nanoparticles in sensing and imaging: Sensitivity of plasmon response to size, shape, and metal composition. *Journal of Physical Chemistry B*, 2006, vol. 110, no. 39, pp. 19220-19225. ISSN 1520-6106.

[2] ZANOLI, L. M. et al. Functionalized gold nanoparticles for ultrasensitive DNA detection. *Analytical and Bioanalytical Chemistry*, 2012, vol. 402, no. 5, pp. 1759-1771. ISSN 1618-2642.

[3] WANG, Y. et al. Synthesis of Ag Nanocubes 18-32 nm in Edge Length: The Effects of Polyol on Reduction Kinetics, Size Control, and Reproducibility. *Journal of the American Chemical Society*, 2013, vol. 135, no. 5, pp. 1941-1951. ISSN 0002-7863.

[4] HUANG, X. et al. Gold Nanorods: From Synthesis and Properties to Biological and Biomedical Applications. *Advanced Materials*, 2009, vol. 21, no. 48, pp. 4880-4910. ISSN 0935-9648.

[5] LOVE, J. C. et al. Self-assembled monolayers of thiolates on metals as a form of nanotechnology. *Chemical Reviews*, 2005, vol. 105, no. 4, pp. 1103-1169. ISSN 0009-2665.

[6] MAZZATENTA, A. et al. Volatile organic compounds (VOCs) fingerprint of Alzheimer'sdisease. *Respiratory Physiology & Neurobiology*, 2015, vol. 209, pp. 81-84. ISSN 1569-9048. [7] PENG, G. et al. Diagnosing lung cancer in exhaled breath using gold nanoparticles. *Nat Nano*,2009, vol. 4, no. 10, pp. 669-673. ISSN 1748-3387.

[8] LE VIET, T. et al. On-chip fabrication of SnO2-nanowire gas sensor: The effect of growth time on sensor performance. *Sensors and Actuators B-Chemical*, 2010, vol. 146, no. 1, pp. 361-367. ISSN 0925-4005.

[9] LEE, J.-S. et al. Au-Polypyrrole Framework Nanostructures for Improved Localized Surface Plasmon Resonance Volatile Organic Compounds Gas Sensing. *Journal of Nanoscience and Nanotechnology*, 2015, vol. 15, no. 10, pp. 7738-7742. ISSN 1533-4880.

[10] MARTIN, C. R. Membrane-based synthesis of nanomaterials. *Chemistry of Materials*, 1996, vol. 8, no. 8, pp. 1739-1746. ISSN 0897-4756.

[11] VIGDERMAN, L. et al. Functional Gold Nanorods: Synthesis, Self-Assembly, and Sensing Applications. *Advanced Materials*, 2012, vol. 24, no. 36, pp. 4811-4841. ISSN 0935-9648.

[12] DRBOHLAVOVA, J. et al. Gold Nanostructured Surface for Electrochemical Sensing and Biosensing: Does Shape Matter? *Analytical Letters*, 2016, vol. 49, no. 1, pp. 135-151. ISSN 0003-2719.

[13] SANTOS, A. et al. Nanoporous Anodic Alumina: A Versatile Platform for Optical Biosensors. *Materials*, 2014, vol. 7, no. 6, pp. 4297-4320. ISSN 1996-1944.

[14] HRDY, R. et al. Electrochemical Impedance Spectroscopy Behaviour of Guanine on Nanostructured Planar Electrode. *International Journal of Electrochemical Science*, 2013, vol. 8, no. 4, pp. 4384-4396. ISSN 1452-3981.

SENSITIVE ANALYSIS OF IR HEAT EMITTER FOR SOLDERING

Josef Skácel, Alexandr Otáhal

Doctoral Degree Programme (1), FEEC BUT

E-mail: xskace09@stud.feec.vutbr.cz

Supervised by: Ivan Szendiuch

E-mail: szend@feec.vutbr.cz

Abstract: This paper deals with the possibility of simulation in ANSYS. Research is focused on the sensitive analysis of heat emitters. The first part is about basic properties of infrared heat emitter. For soldering is necessary to use heat emitter with the homogenous thermal field. In this case was used two types of wire arrangement. Classics spiral and type with two winding for the better homogenous thermal field. Next part deals with steady-state thermal simulation and transient simulation of heat emitter in ANSYS Workbench.

Keywords: ANSYS, heat emitter, sensitive analysis, reflow

1. INTRODUCTION

Reflow soldering is one of many methods of soldering [1]. For the reflow process is used heat emitter with different shapes. In figure 1a can be seen the main shape of heat emitter. The most commonly used are square and rectangular shape. The materials used to the heater is ceramic and resistance wire (compound of chrome, aluminium, iron, cobalt) but, in this case, was used the base from copper due to easier realisation. The heat emitter consists from the copper base, high-temperature cement, resistance wire and copper cover, see Fig. 1b.



a) b)

Figure 1: a) Types of heat emitter, b) internal layout of heat emitter

The main requirement of heat emitter is a homogenous thermal field. Another requirement is the high coefficient of emissivity. The first it´s given by the arrangement and material of resistance wire. The second is given by the heat emitter and substrate surfaces, see Tab. 1. It can be seen that copper as the material is not a good choice for heat emitter but is potentiality way to modify the surface. The first way is perform oxidated. Classic oxidate with 0,78 coefficient of emissivity or black oxidated with 0,88 coefficient of emissivity. Another way to improve emissivity is coating. For example, the high-temperature black coating has 0,97 coefficient of emissivity [2].

Dimensions of emitter: Copper base (100 x 100 x 10) mm, the thickness of the copper base is 2 mm, thickness of copper backplate is same 2 mm, the diameter of resistance wire is 1mm.

Dimensions of high temperature cement are (90 x 90 x 0,6) mm. The classic winding is 1070 mm long and new type of winding is two 73 cm long winding.

Table 1: Table of emissivities

Material	Coefficient of emissivity
Copper oxidated	0,78
Copper black oxidated	0,88
Copper polished	0,01
Copper plated	0,03
Alumina 96%	0,7
Gold polished	0,02
Black coating	0,97
Black el. component	0,9

2. SIMULATION OF HEAT EMITTER

For the purposes of this paper was used ANSYS Workbench. Steady-state and transient thermal is classic FEM programme [3]. For simulation purpose was used 4 materials. Substrate material was Alumina 96%. For base and plate was used copper. Resistance wire is Kanthal A-1 [4]. For high-temperature cement was selected OmegaBond 600 [5] shows table 2.

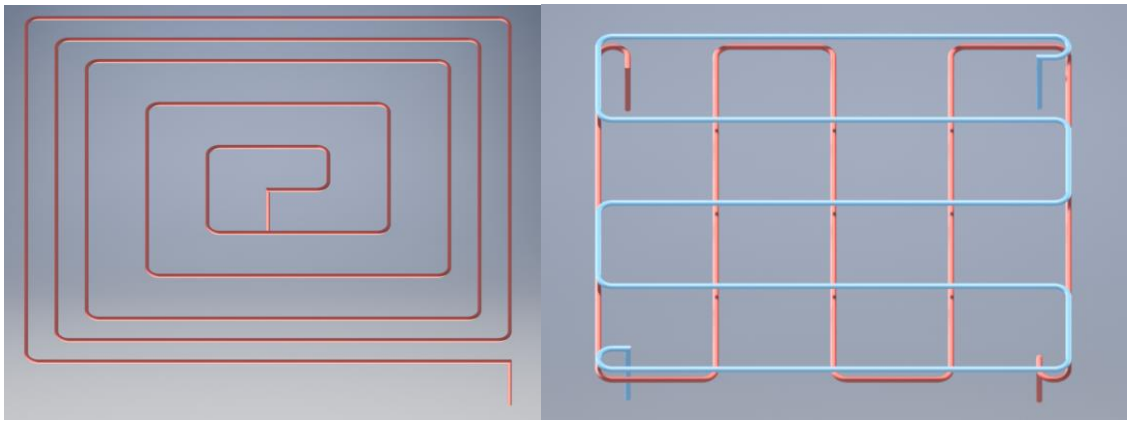
The sensitive analysis contains five power loads: (600, 650, 700, 750, 800) W. Convection was set at 20 $W/m^2 \cdot C$ for all simulations. Internal heat generation depends on power load. Emissivity was set 0,97 for black coated copper base and 0,7 for Alumina substrate. These coefficients are for the best situation.

Table 2: Materials and basic simulation settings

Material	Density [kg/m^{-3}]	T. Conductivity [$W/m^{-1} \cdot C^{-1}$]	Specific Heat [$J/kg^{-1} \cdot C^{-1}$]	
Alumina 96%	3800	25	880	
Copper	8300	401	385	
Kanthal A-1	7100	20 (600 °C)	750 (600 °C)	
Omegabond600	2563	17,28	1550	
Simulation	**Convection [$W/m^2 \cdot C$]**	**Internal Heat Generation [W/mm^3]**	**Radiation Emitter / Alumina**	
600 W	20	0,75	0,97	0,70
650 W	20	0,8125	0,97	0,70
700 W	20	0,875	0,97	0,70
750 W	20	0,9375	0,97	0,70
800 W	20	1	0,97	0,70

Conventional solutions of heat emitter have the same pitch between wire. The figure 2a shows the classic arrangement of wire winding with improvements. It can be seen that pitch between wire is not same. It´s due to get more homogenous thermal field than classic heat emitter sees at figure 1a. Thermal energy is more accumulate in the center of heat emitter so it´s necessary to arrangement wire with the closer pitch at the edge of the emitter.

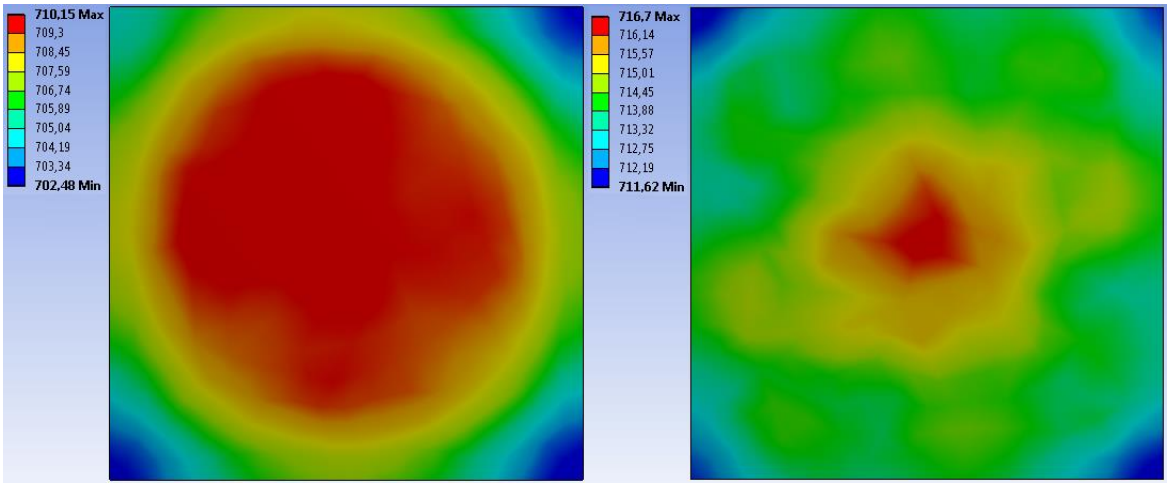
The figure 2b shows another solution of wire arrangement. In this case is used two windings for higher heat flux and more homogenous field than classic heat emitter. The wire starts with a special bend in order to obtain more thermal energy in the corner than the center of the emitter.



a) b)

Figure 2: a) Classic winding, b) new type of winding

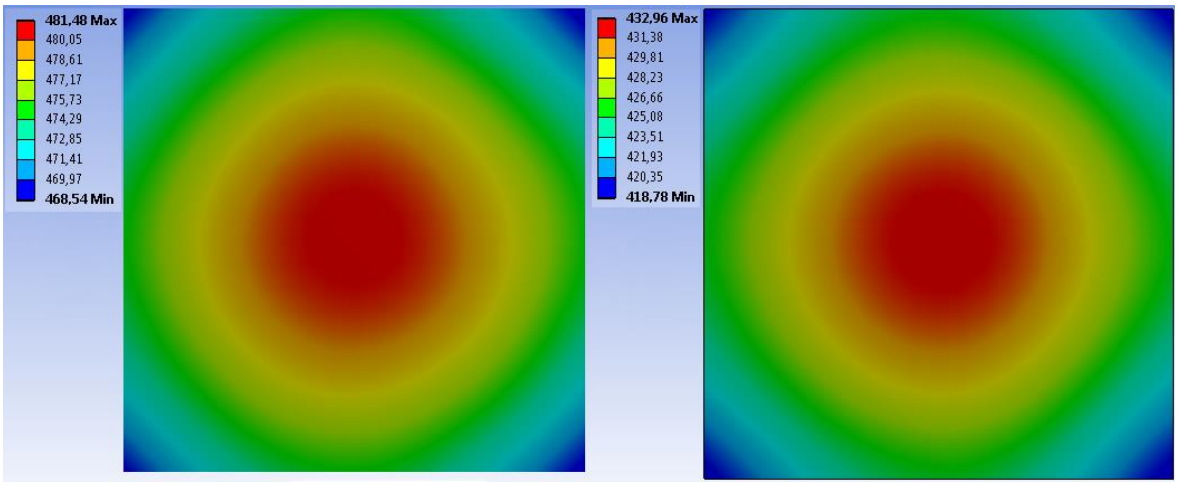
From result can be seen that thermal field is different in this two types of winding. Figure 3a shows the thermal field of classic spiral winding. Distribution of thermal field is concentrated in the center of the emitter. Temperature different on the surface of the copper base is 7,67 °C. The thermal field of the new type of winding shows better behavior than the first type. The thermal field is more homogenous and the temperature difference is lower. Different is 5,08 °C.



a) b)

Figure 3: a) Thermal field of classic winding, b) New type of winding

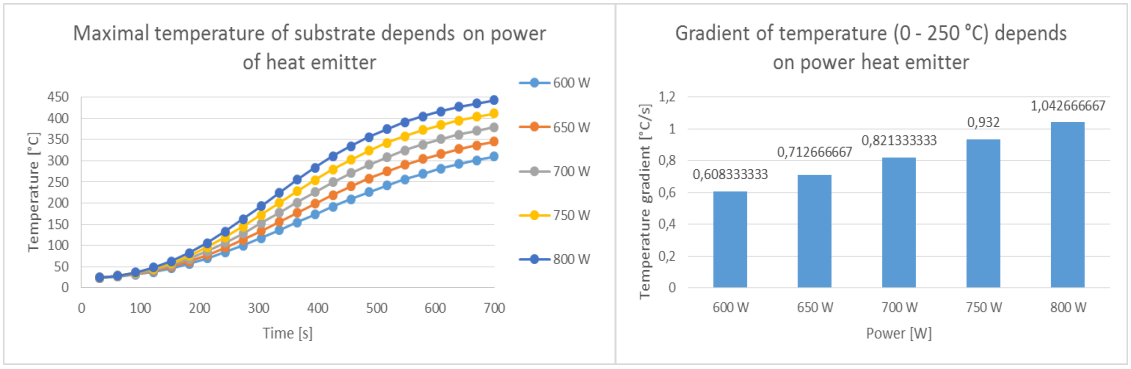
For soldering purpose is a necessary temperature gradient. In Fig. 5a can be seen temperature gradient for different power loads. The heat emitter with 600 W power load reaches the needed temperature 250 °C for 550s but with 800 W power load reach 250 °C for 350s. In Fig 5b can be seen gradient temperature for all power loads. Gradient for 600 W is 0,608 °C/s. Gradient for 800 W is 1,04 °C/s.



a) b)

Figure 4: a) New type of winding (substrate), b) Classic winding (substrate)

In the figure, 4b can be seen a thermal field of the substrate with the classic winding. Temperature different is 14,18 °C. The maximal temperature at the emitter is 710 °C but the maximal temperature of the substrate is 432 °C. A new type of winding shows better results than classic winding. The temperature at heat emitter is 716 °C it´s only 6 °C more than the classic emitter. But the maximal temperature of the substrate is 481 °C and different between maximal and the minimal temperature is 12,94 °C.



a) b)

Figure 5: a) Maximal temperature depends on power load, b) Gradient of temperature (0-250 °C) depends on power load

3. CONCLUSION

This paper deals with the sensitive analysis of IR heat emitter for soldering and utilization Ansys Workbench for this problem. The results show that is possible use transient thermal simulation for find the best solution for power load of the emitter. Simulation in ANSYS can find a better type of arrangement of wire winding for the more homogenous thermal field before construction start and is possible save time. These simulations show the way to better heat emitter with the more homogenous thermal field. But it´s necessary to confirm this result by real experiments. In this case, the experiment wasn´t done.

ACKNOWLEDGEMENT

Funding for this research work was supported through of the Czech Ministry of Education for Brno University of Technology FEKT-S-14-2168.

REFERENCES

[1] KEMET. *Reflow Soldering Process Considerations for Surface Mount Application* [online]. 2000 [cit. 2016-02-15]. Online at: http://www.kemet.com/Lists/TechnicalArticles/Attachments/28/f2102a.pdf

[2] INFRARED THERMOGRAPHY. *Emissivity Values for Common Materials* [online]. 2000 [cit. 2016-02-15]. Online at: http://www.infrared-thermography.com/material.htm

[3] ANSYS. *Thermal Transient* [online]. 2015 [cit. 2016-02-15]. Online at: http://resource. ansys.com/Resource+Library/Thermal+Transient

[4] KANTHAL. *Kanthal A-1* [online]. 2012 [cit. 2016-02-15]. Online at: http://kanthal.com/en/products/material-datasheets/strip/kanthal-a-1/

[5] OMEGA. *High Temperature Cements* [online]. 2012 [cit. 2016-02-15]. Online at: http://www.omega.com/pptst/OB_BOND_CHEM_SET.html

SEPARATION AND DETECTION OF BIOCONJUGATED QUANTUM DOTS USING ON A CHIP ELECTROPHORESIS

Jelena Pejović

Doctoral Degree Programme (3), FEEC BUT

E-mail: jelena.pejovic@ceitec.vutbr.cz

Supervised by: Jaromír Hubálek

E-mail: hubalek@feec.vutbr.cz

Abstract: Semiconductor nanocrystals, quantum dots (QDs), are nanoscale particles that have been attract a lot of attention due to their unique optic and electronic properties. Due to very diverse and numerous applications, it is really important to make a tool for precise and controlled detection of QDs. In this study, we investigate on a chip detection and separation of QDs bioconjugated with BSA (bovine serum albumin) using homemade equipment based on principle of capillary electrophoresis (CE) and optical detection. Quenching effect of different concentration of BSA on fluorescence intensity of QDs was monitored. It was found that with increasing concentration of BSA fluorescence intensity of QDs is decreasing. This research can lead to a better understanding of interaction between different size QDs and biomolecules.

Keywords: Quantum dots, bovine serum albumin, capillary electrophoresis, bioconjugation

1. INTRODUCTION

QDs belong to family of inorganic nanoparticles and they are defined as semiconductors nanocrystals [1]. Their unique optical and electronic properties can be placed between those of bulk materials and isolated molecules of atoms [2]. QDs are nanoscale particles that consist of semiconductor core, such as CdTe and have tunable emission spectrum, from ultraviolet to infrared, by simple changing of the size of cores [3].

CE with high separation efficiency based on the size-to-charge ratio of analytes, holds promise for efficient size determination of nanoparticles (NPs) in aqueous samples [4]. In recent 10 years, CE has emerged as a powerful tool for characterizing and separating QDs and other NPs [5]. On a chip electrophoresis has some advantages over conventional one since it requires low sample volume, fast separation time, high separation efficiencies and additionally microchip platform provides the possibility to integrate sampling, separation, and detection on a chip [6].

Due to their unique optical properties, photostability and resistance to chemical degradation in last decade QDs find great application in medicine and biology usage, for luminescence tagging and molecular imagining, drug delivery, as tumor biomarkers and other [7]. QDs modified by BSA have been applied as ion sensors, fluorescence resonance energy transfer and chemiluminescence resonance energy transfer [8]. Capillary electrophoresis with laser-induced fluorescence detection (CE-LIF) has been used for characterization of the surface modifications of QDs and evolution of the particle charge [9], [3], [10]**.** Hung et al. demonstrated how simple optimization of pH of buffer can substantially improve the separation of CdTe QDs and their conjugates with an antibody [9]. Application of QDs as fluorescence label in immunoassay was reported by Feng et al, where QDs were first conjugate with antibody then separated by CE form free antibody and antibody-antigen complex [3]. However, the core of QDs created with inorganic elements is toxic for living systems and cells. Nowadays, concerning environment issues and toxic of QDs lot of improve environment-friendly technologies are used for synthesis of QDs as well as modification of their surface to make

it less toxic for usage in medicine [11]. Even, these modified QDs bring doubts [12], and more research need to be done to investigate QDs and their bioconjugation.

Taking in consideration everything above, aim of this work is to use fast and simple method for investigate bonding between QDs and biomolecules. In this work conjugation of QDs with bovine serum albumin (BSA) was achieved via covalent coupling using different crosslinkers: carbolyldiimidazole (CDI), 1-ethyl-3-(3-dimethylaminopropyl), carboiimide hydrochloride (EDC), N-hydroxysulfosuccinimide (NHS/EDC). QDs and QDs-protein conjugates were successfully separated using on a chip CE. Fluorescence detection was used to study quenching effect of protein and QDs. It was found out that higher concentration of BSA led to a stronger quenching of fluorescence emission, which could be explained by the covalent interaction between the protein and QDs Drbohlavova, Chomoucka [13]. In this work detection and separation of bioconjugated QDs was done in microfluidic system, using on chip electrophoresis and optical detection. By investigating interaction between different sized QDs and biomolecules this research can lead to a better understanding of how QDs behave in cell environment, which is extremely important for their biological application.

2. MATERIALS AND METHODS

2.1. REAGENTS AND CHEMICALS

The following chemicals were used for preparation of QDs and their conjugates with BSA: Sodium borohydride ($NaBH_4$, 99%), BSA (96%), CDI (97%), Sodium citrate dihydrate (99%), cadmium chloride ($CdCl_2$, 99%), sodium tellurite (Na_2TeO_3, 99%) were purchased from Sigma Aldrich. Sodium hydroxide (NaOH, 98%) were purchased from Penta. MPA (98%) were purchased from Merck. Isopropyl alcohol (99.7%) was purchased from Lach-Ner. The following chemicals were used for fabrication of silica mold and polydimethylsiloxane (PDMS) chip: SU-8 developer were purchased from MicroChem (Newton, MA,USA). The poly dimethylsiloxane prepolymer (Sylgard 184) and a curing agent were purchased from Dow Corning.

2.2. SYNTHESIS OF CDTE QDS CAPPED WITH MPA

To obtain CdTe-MPA QDs, $CdCl_2$ solution (91.6 mg) was diluted to 50 ml in one-necked flask and sodium citrate dihydrate (200 mg) was added followed by addition of MPA (52 µl). The pH of the solution was adjusted to 10.5 using NaOH (1 mol/l), followed by addition of Na_2TeO_3 (22.15 mg) and $NaBH_4$ (50 mg) under vigorous stirring. Solution was then refluxed at 95°C for 4 h [14].

2.3. PREPARATION OF QDS-BSA CONJUGATES VIA CDI, EDC, EDC/NHS

Briefly, to the solution of CdTe QDs (1 mg/ml), CDI (10 µl, 10 mmol/l) containing PBS buffer (100 mmol/l, pH=7.4) was added and the solution was then incubated at room temperature for 30 min in order to activate carboxyl groups [14]. Then, BSA (200 µl; 0, 0.002, 0.005, 0.01, 0.05, 0.075, 0.1 mg/ml) was added to the solution and incubated at 32°C for 2 h while shaking [14].

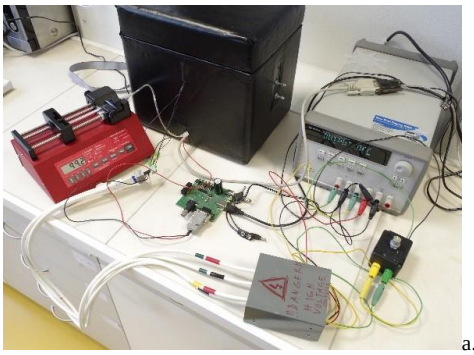
Briefly, to the solution of CdTe QDs (200 µl, 0.1 mg/ml) EDC (200 µl, 50 mmol/l) and NHS (200 µl, 5 mmol/l) were added and solution was then incubated at 32°C for 30 min. Then, BSA (200 µl; 0, 0.002, 0.005, 0.01, 0.05, 0.075, 0.1 mg/ml) was added to the solution and incubated at 32°C for 2 h while shaking [14].

Briefly, to the solution of CdTe QDs (250 µl, 0.1 mg/ml) BSA (250 µl; 0, 0.05, 0.5, 1 a 1.5 mg/ml) and EDC (57 µl, 10 mg/ml) were added. The solution was then incubated at room temperature for 2 h [14].

2.4. APPARATUS

Capillary electrophoresis of QDs and QDs-protein conjugates was carried out using home built system “black box” (**Figure 1**.). The black box consists of the three parts: light source, sample

holder and light detector. Excitation light is generated by ultraviolet light-emitting diode (UV LEDs) and filtered by optical 380 nm low pass filter. This light excites fluorescence of the sample and emitted light is selected by 560 nm high pass optical filter. This emitted light is detected by photodetector including photomultiplier.


a.

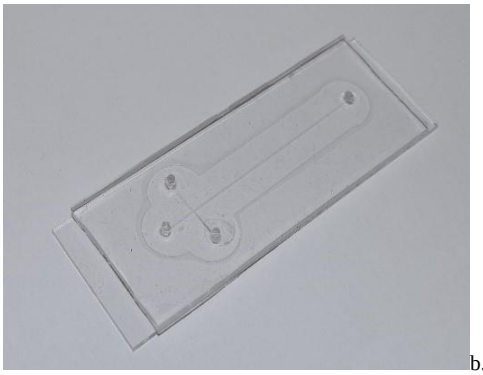

b.

Figure 1a: Experimental set up: "black box", syringe pump and power supplies, **b:** PDMS chip

2.5. FABRICATION OF A CHIP

Rapid prototyping of chip starts with creating design for a device in a computer/aided design program. A high resolution printing is used to print design on a chrome transparency mask. This transparency serves as a photomask in photolithography to produce a positive relief of photoresist on silica wafer. Silica mold with channels of 50μm width was made and used for rapid molding of PDMS. A mixture of the PDMS, prepolymer and curing agent at a ratio of 10:1 (w/w) was completely degassed and poured onto the silica mold, which was cured for 2 h at 65 ^{O}C, and peel off. After air drying of channels PDMS is punched to make holes (2 mm) as fluid reservoirs. The patterned side of the PDMS was treated with oxygen plasma (200 W, 15 min) and bonded permanently with a plasma-treated glass substrate to form a closed fluidic system (Figure 1b.).

2.6. CE PROCEDURE

After making the chip, microchannels are first filled with buffer using syringe pump. After filing the sample reservoir, platinum electrodes are placed in each fluid reservoir and contacted to power supplies. The applied voltage for the electrophoretic separation was 3 kV. Initially a new chip was flushed with 0.1M NaOH (10 min), water (10 min), and running buffer (10 min). A 10mM Borate Buffered Saline (BBS) aqueous solution pH 9.14 was used as a BGE. All solutions were filtrated through a 0.45μm membrane filter before use. In the slight alkaline solution, CdTe possess negative charges due to dissociation of carboxylic groups on their surface and therefore migrate to the anode in the electric field.

3. RESULTS

3.1. SEPARATION OF CONJUGATED CDTE- MPA QDS BY ON A CHIP CE

Measuring the changes in the electrophoretic mobility overtime of bioconjugated QDs by CE can determined binding of QDs with biomolecules, since interaction with these molecules will result in changes in the hydrodynamic size of QDs. The QDs used in this studies were conjugated to BSA (bovine serum albumin). First, CdTe QDs capped with MPA (CdTe-MPA) were prepared in aqueous solution phase. Crosslinkers CDI, was used to bioconjugate QDs with BSA. **Figure 2** shows electropherogram of free CdTe-MPA QDs, QDs modified by CDI, and QDs bioconjugated with BSA (QD-BSA). From results we can see that free QDs are well separated from bioconjugates by CE on a chip within 10 minutes.

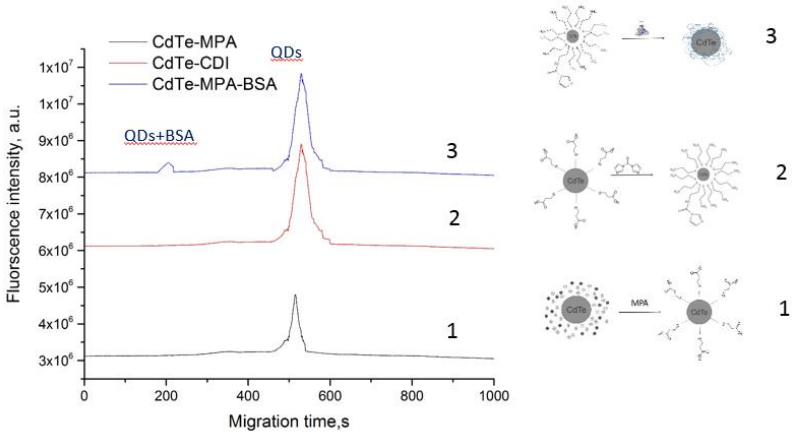


Figure 2: Electropherogram of the QDs CdTe-MPA bioconjugated with BSA

3.2. STUDDING OF FLUORESCENCE QUENCHING OF QDS EFFECT CAUSED BY COVALENT CONJUGATION WITH BSA

QDs and BSA were conjugated via covalent coupling using carboxyl groups of crosslinkers (CDI, EDC, and EDC/NHS). With increasing of concentration of BSA, fluorescence intensity of QDs is decreasing. The fluorescence quenching of QDs by BSA can be described by the linear Stern-Volmer equation:

$$I_0/I = 1 + K_{sv}[Q] \quad (1)$$

where I_0 and I are the steady-state fluorescence intensities of QDs in the absence and presence of BSA, respectively. K_{SV} is the Stern-Volmer quenching constant, and [Q] is the concentration of BSA. The I_0/I were calculated and plotted against BSA concentration as described in (1).

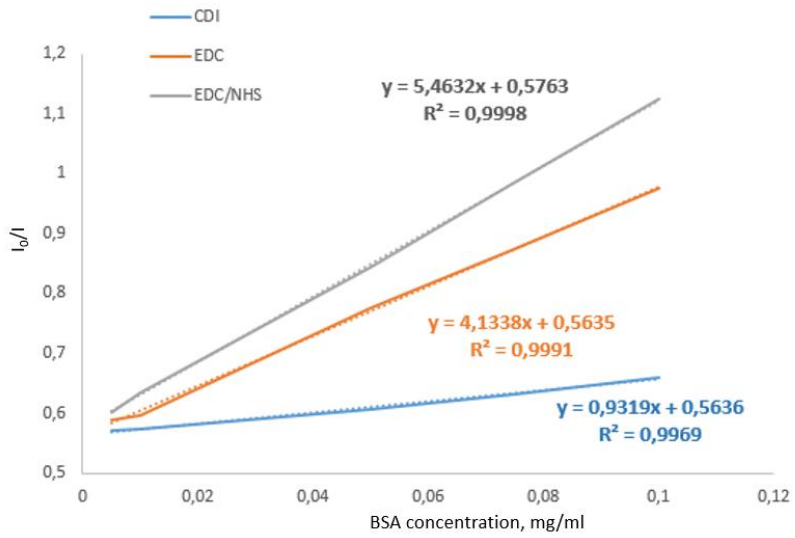


Figure 3: Stern- Volmer plot of QDs quenching effect caused by CdTe QDs electrostatic conjugation with BSA

The slope of the curve represents the Stren-Volmer quenching constant (K_{sv}) and the higher K_{sv}, the higher the quenching effect is. Stern- Volmer plot of all three crosslinkers quenching properties is shown exhibiting linear trend. The results show that the quenching constant is different for different types of bioconjugated QDs. From Figure 3 it is obvious that in case where EDC/NHS is

used as a crosslinker the fluorescence intensity is decreasing significantly, so the highest quenching effect was notable.

4. CONCLUSIONS

Water soluble CdTe QDs were prepared using a simple one step method. In this study, we successfully separated CdTe QDs bioconjugated with BSA using on a chip electrophoresis. What characterize this method are simplicity, short analysis time, small sample and reagent requirements, high separation efficient. The present of BSA is leading to a strong quenching effect of fluorescence emission of QDs. Quenching effect depends of a cross linker we are using to conjugate protein and it was shown to be the highest one when using EDC/NHS. Overall, this simple, fast and not expensive method can be used as powerful tool to better understand interaction between biomolecules and QDs.

REFERENCES

1. Mattoussi, H., G. Palui, and H.B. Na, *Luminescent quantum dots as platforms for probing in vitro and in vivo biological processes.* Advanced Drug Delivery Reviews, 2012. **64**(2): p. 138-166.
2. Tomczak, N., et al., *Designer polymer-quantum dot architectures.* Progress in Polymer Science, 2009. **34**(5): p. 393-430.
3. Feng, H.T., et al., *Immunoassay by capillary electrophoresis with quantum dots.* Journal of Chromatography A, 2007. **1156**(1-2): p. 75-79.
4. Stewart, D.T.R., et al., *Potential use of capillary zone electrophoresis in size characterization of quantum dots for environmental studies.* Trac-Trends in Analytical Chemistry, 2011. **30**(1): p. 113-122.
5. Song, X.T., et al., *Highly efficient size separation of CdTe quantum dots by capillary gel electrophoresis using polymer solution as sieving medium.* Electrophoresis, 2006. **27**(7): p. 1341-1346.
6. Nandi, P. and S.M. Lunte, *Recent trends in microdialysis sampling integrated with conventional and microanalytical systems for monitoring biological events: A review.* Analytica Chimica Acta, 2009. **651**(1): p. 1-14.
7. Byers, R.J. and E.R. Hitchman, *Quantum Dots Brighten Biological Imaging.* Progress in Histochemistry and Cytochemistry, 2011. **45**(4): p. 201-237.
8. Liang, J., Y. Cheng, and H. Han, *Study on the interaction between bovine serum albumin and CdTe quantum dots with spectroscopic techniques.* Journal of Molecular Structure, 2008. **892**(1–3): p. 116-120.
9. Huang, X.Y., et al., *Characterization of quantum dot bioconjugates by capillary electrophoresis with laser-induced fluorescent detection.* Journal of Chromatography A, 2006. **1113**(1-2): p. 251-254.
10. Pathak, S., et al., *Hydroxylated quantum dots as luminescent probes for in situ hybridization.* Journal of the American Chemical Society, 2001. **123**(17): p. 4103-4104.
11. Jha, R.K., et al., *An emerging interface between life science and nanotechnology: present status and prospects of reproductive healthcare aided by nano-biotechnology.* 2014, 2014.
12. Morelli, E., et al., *Chemical stability of CdSe quantum dots in seawater and their effects on a marine microalga.* Aquatic Toxicology, 2012. **122**: p. 153-162.
13. Drbohlavova, J., et al., *Effect of Nucleic Acid and Albumin on Luminescence Properties of Deposited TiO2 Quantum Dots.* International Journal of Electrochemical Science, 2012. **7**(2): p. 1424-1432.
14. Mihajlovic, A., *Synthesis of core/shell quantum dots for diagnostics.* Brno: Brno University of Technology, The Faculty of Electrical Engineering and Com-munication, 2014: p. 88.

SHAPE CONTROLLED SYNTHESIS OF SILVER NANOPARTICLES WITH ANTIMICROBIAL EFFECT

Evelína Polievková

Advanced Nanotechnologies and Microtechnologies (2nd level), CEITEC BUT

E-mail: Evelina.Polievkova@ceitec.vutbr.cz

Supervised by: Jana Drbohlavová, Jan Křivčík

E-mail: drbohla@feec.vutbr.cz

Jan.Krivcik@membrain.cz

Abstract: This paper deals with shape controlled synthesis of silver nanoparticles with antimicrobial effect. Silver nanoparticles were prepared by UV-assisted chemical reduction with two different wavelengths. Influence of UV radiation on final shape and size of nanoparticles was investigated.

Keywords: UV radiation, chemical reduction, shape, size, photoinduced reaction

1. INTRODUCTION

Nowadays, nanochemistry frequently uses metallic nanoparticles with unique properties compared to their macro scaled counterparts coming from their high surface to volume ratio [1,2]. Silver nanoparticles (further denoted as Ag NPs) can be utilized thanks to their remarkably distinct physical, biological and chemical properties in various research areas (physics – e.g. sensing and optics, chemistry – catalysis, biology, pharmacy, and medicine). These features are very significantly affected by their shape or size and also their monodisperzity. The ideal Ag NPs evince small particle dimension, high surface area, quantum confinement or no agglomeration.

Recently Ag NPs have been very used for antimicrobial purposes [1,2]. For instance, antibacterial activity is very famous not just in research or medicine area but also in commercial products such as clothes with nanosilver exploited (not only) in army [3]. Ag NPs are parts of various antibacterial agents and of cosmetics products. This antibacterial effect is very closely related with Ag NPs size; general postulate is that the smaller is the silver nuclei, the higher is antibacterial activity, see Figure 1 which shows the mechanism of antibacterial activity [4]. Control of size or shape of Ag NPs can be reached by various preparation techniques using reducing agents or stabilizers [5,6]. It is also well-known that silver-based materials are significantly toxic to microorganisms; exactly to 16 major species of bacteria including *E.coli*. Silver is usually used in its nitrate form to originate antimicrobial effect but in the case of using Ag NPs the available surface area exposed to microbes increases very rapidly [7,8].

UV-induced synthesis seems to be very efficient route how to prepare Ag NPs with exact physical properties for concrete application [9-11]. This paper presents an innovative synthesis of Ag NPs via UV-assisted chemical reduction. Utilization of different wavelengths can significantly affect size and shape distribution of colloidal Ag NPs. These Ag NPs will be used as antibacterial substances for commercial polymeric membranes used in food industry for filtration of milk in cooperation with company MemBrain, ltd. These membranes suffer during filtration process from contamination with bacteria, namely above mentioned strain *E.coli,* which damage and congest membrane pores and subsequently debase not only membranes but also filtration process.

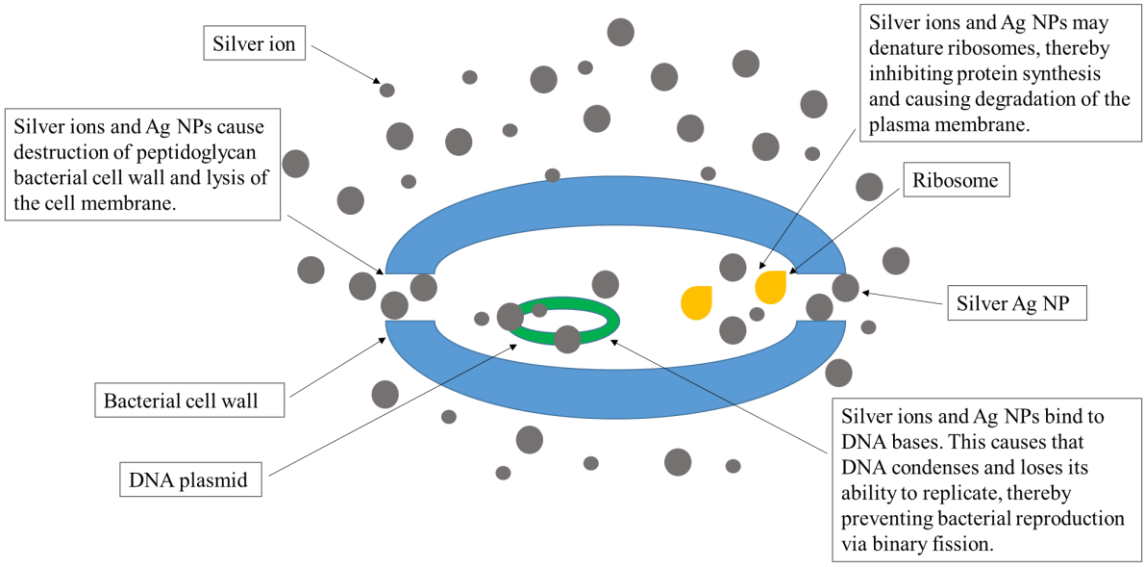


Figure 1: Mechanism of antibacterial activity of Ag NPs

2. SYNTHESIS OF SILVER NANOPARTICLES

Materials: sodium citrate dihydrate (Sigma-Aldrich), hydrazine hydrate (Aldrich) and silver nitrate (Penta). All materials were used as purchased without further purification. All chemicals are used in chemical purity grade p.a. (per analysis).

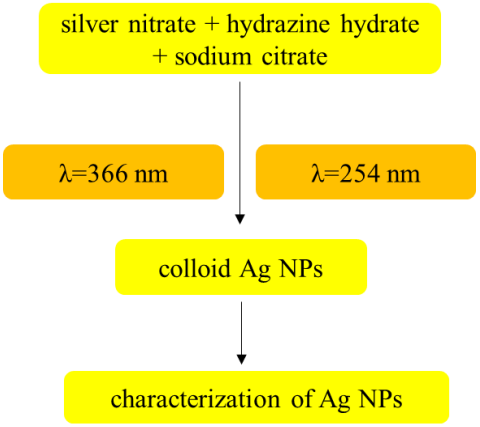


Figure 2: Preparation process of Ag NPs

Ag NPs were prepared via chemical reduction according to Guzmán et al. with modified volumes of reagents and conditions [5]. Scheme of this process is shown in **Chyba! Nenalezen zdroj odkazů.**. 1 mM aqueous solution of silver nitrate was used as precursor and 1mM aqueous solution of hydrazine hydrate and 2 mM aqueous solution of sodium citrate as reducing agents. Sodium citrate also played a stabilizing role of Ag NPs. Deionized water (resistivity of 18.2 MΩ·cm) was used to prepare all aqueous substances of reaction. Reaction time was 1 h and preparation process was carried out at temperature of 22–24 °C employing UV radiation with two different wavelengths of λ_1= 254 nm and λ_2= 366 nm. UV lamp was placed above a beaker in the distance of 1.5 cm from the top of the beaker. The beaker with aqueous solution of silver nitrate was placed on magnetic stirrer under 200 rpm. After that the aqueous solutions of reducing agents were added to silver nitrate at the same time. The colour of reaction solution changed after 1 min after adding reducing substances into light orange which indicates the formation of Ag NPs in solution. Colloid samples of Ag NPs were purified 3 times with deionized water via centrifugation (centrifuge EBA 20) under

14 500 rpm for 30 min. Ag NPs were dispersed in ultrasound water bath (power of 100%, temperature of 25 °C) for 1 min between every purification and also finally after purification process. Colloidal Ag NPs were stored in dark bottles in fridge.

3. CHARACTERIZATION OF SILVER NANOPARTICLES

Shape, size and stability of Ag NPs were characterized by transmission electron microscopy (TEM) using JEM 2100F (Schottky cathode, energy of electron beam of 200 kV) at Institute of Physics of Materials at Academy of Sciences of the Czech Republic. Absorbance of Ag NPs was measured by UV-VIS spectrophotometer using Spectronic Helios Alfa at Institute of Physical and Applied Chemistry at Faculty of Chemistry.

4. RESULTS AND DISCUSION

4.1. TRANSMISSION ELECTRON MICROSCOPY ANALYSIS

Ag NPs prepared employing UV radiation with wavelength of 254 nm (Sample Ag 254) can be seen in Figure 3 middle and right. These Ag NPs have mostly spherical shape with average size of about 20–30 nm. Figure 3 left shows the detail of Ag NPs grown into bigger objects. Thus Ag NPs of this sample are not very stable in colloidal solution and create aggregates in short period after preparation process.

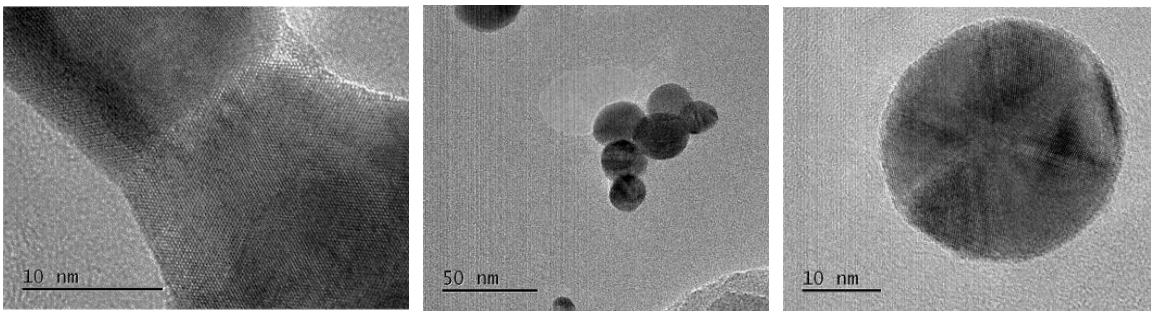


Figure 3: TEM images of Sample Ag 254 (left - grown Ag NPs, middle - Ag NPs cluster, right - single Ag NP)

Figure 4 shows Ag NPs prepared under UV radiation with wavelength of 366 nm (Sample Ag 366). In this case Ag NPs have anisotropic shape distribution; triangulars with truncated tops, spherical particles and rods. A smooth amorphous layer that covers the surface of Ag NPs is clearly visible. This cover protects particles from growing together in comparison with Sample Ag 254 where this stabilization layers is missing. Thus it can be assumed that Ag NPs of this sample are stable and without aggregation or grown particles. This layer is probably formed from sodium citrate which plays also a role of stabilization of Ag NPs. The size of these colloidal Ag NPs varies from 10 to 50 nm. The stability of the sample was observed up to 2 weeks after the synthesis when aggregates could be seen by naked eye at the bottom of the flask.

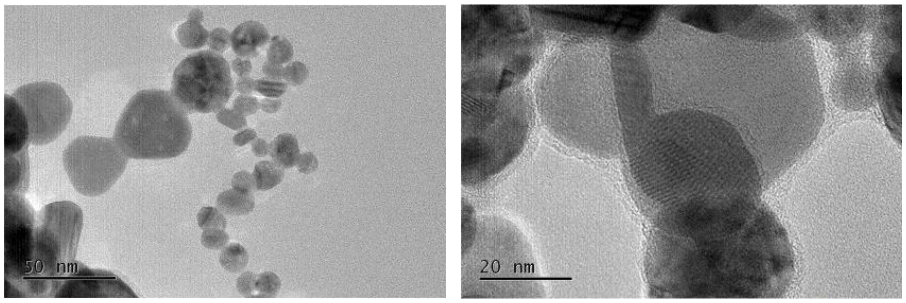


Figure 4: TEM images of Sample Ag 366 (left - Ag NPs with different shapes, right - the detail of stabilization layer covering Ag NPs)

4.2. ANALYSIS OF OPTICAL PROPERTIES

UV-VIS spectroscopy in the range of 300 to 700 nm was performed. This method can help to determine not only optical properties of sample but also aggregation, size or real presence of Ag NPs in the solution. Ag NPs show typical absorption peak of surface plasmon resonance (further SPR) between 400 and 450 nm.

Figure 5 shows absorption maxima of both samples approximately at wavelength of 420 nm, thereby it can be declared that Ag NPs are really present in both colloidal solutions. The absorption maximum of SPR of Sample Ag 366 shows higher intensity in comparison with Sample Ag 254. Intensity of absorption maxima is related with size and aggregation of Ag NPs [12]. Both samples have another peak visible at shorter wavelengths. In the case of Sample Ag 254 the presence of the second peak can be caused by grown particles into bigger objects which was confirmed by TEM analysis as well. In the case of Sample Ag 366 the secondary peak can be caused by uneven size and shape of Ag NPs; primary peak can belong to smaller particles while secondary peak to bigger or anisotropic particles. Second peak at shorter wavelengths of both samples can be also caused by free ions which have not been reduced by reducing agents during reaction. If Ag NPs were aggregated, another peak at longer wavelengths would appear between 600 and 700 nm. This peak is missing at both samples so they does not contain any aggregates.

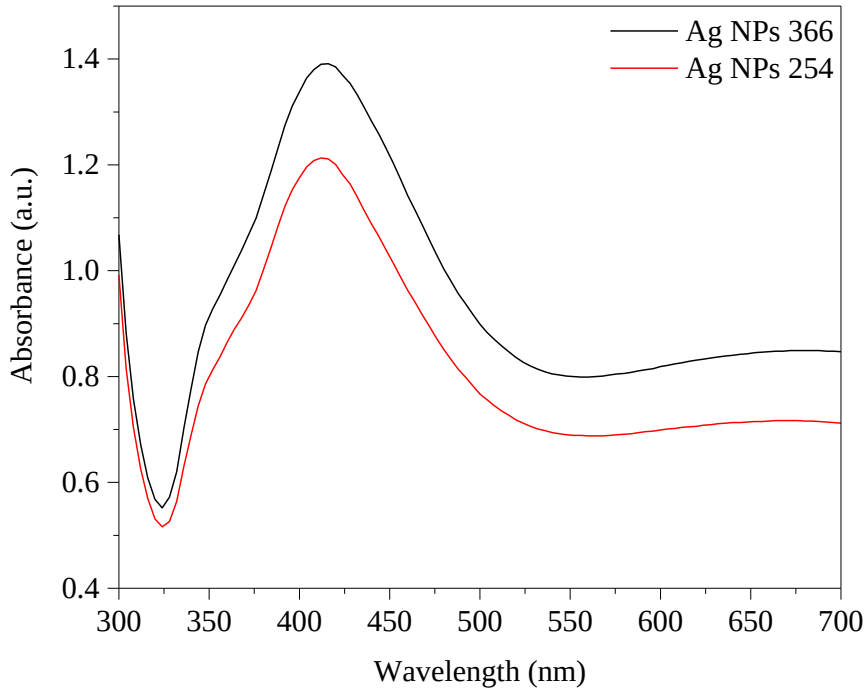


Figure 5: Absorption spectra of Sample Ag 366 and Sample Ag 254

5. CONCLUSION

Ag NPs were prepared via UV-assisted chemical reduction. It was demonstrated that different wavelengths of UV radiation significantly affected both the shape and the size of final Ag NPs in colloidal solution. When employing wavelength of 254 nm, the final Ag NPs dominated in spherical shape with average diameter of 20 nm. Size and shape distribution of this Sample Ag 254 was very narrow. Unfortunately TEM analysis confirmed that this sample was not stabilized by sodium citrate, thus particles grown into bigger clusters. This fact also affects their stability; it can be assumed that these Ag NPs will aggregate fast after the synthesis. In the case of using 366 nm wavelength Ag NPs had various shapes, namely triangulars, rods and spherical particles. TEM analysis showed that surface stabilization with sodium citrate was successful and Ag NPs have smooth stabilization layer. Thus, Ag NPs of this sample are not growing together and they are more stable

against aggregation than Ag NPs of Sample Ag 254. Photoinduced synthesis of Ag NPs seems to be very easy and efficient way how to tune the size and the shape of Ag NPs in colloidal solution according to desired application. Very important facts that can affect final properties of Ag NPs are duration of reaction, wavelength, and intensity of UV source and last but not least concentration of used substances. General postulate is that the shorter is the wavelength, the more spherical shape of nanoparticles can be obtained. The longer is the wavelength, the more anisotropic shape of nanoparticles can be observed. This is advantage enables to tailor not only optical and antimicrobial properties of Ag NPs depending on their size and shape distribution.

ACKNOWLEDGEMENT

This paper was supported by IGA FEKT/STI-J-16-3625 research grant.

REFERENCES

[1] PRABHU, S. a POULOSE, E. Silver nanoparticles: mechanism of antimicrobial action, synthesis, medical applications, and toxicity effects. *International Nano Letters*, 2012, vol. 2, no. 1, pp. 1-10. ISSN 2008-9295.

[2] TRAN, Q. H. a LE, A.-T. Silver nanoparticles: synthesis, properties, toxicology, applications and perspectives. *Advances in Natural Sciences: Nanoscience and Nanotechnology*, 2013, vol. 4, no. 3, pp. 033001. ISSN 2043-6262.

[3] RAI, M. et al. Silver nanoparticles as a new generation of antimicrobials. *Biotechnology Advances*, 2009, vol. 27, no. 1, pp. 76-83. ISSN 0734-9750.

[4] KIM, J. S. et al. Antimicrobial effects of silver nanoparticles (vol 1, pg 95, 2007). *Nanomedicine-Nanotechnology Biology and Medicine*, 2014, vol. 10, no. 5, pp. 1119-1119. ISSN 1549-9634.

[5] GUZMÁN, M. G. et al. Synthesis of silver nanoparticles by chemical reduction method and their antibacterial activity. *Int J Chem Biomol Eng*, 2009, vol. 2, no. 3, pp. 104-111.

[6] SHAMELI, K. et al. Investigation of antibacterial properties silver nanoparticles prepared via green method. *Chemistry Central Journal*, 2012, vol. 6, no. 1, pp. 73. ISSN 1752-153X.

[7] SHARMA, V. K. et al. Silver nanoparticles: Green synthesis and their antimicrobial activities. *Advances in Colloid and Interface Science*, 2009, vol. 145, no. 1-2, pp. 83-96. ISSN 0001-8686.

[8] ZAIN, N. M. et al. Green synthesis of silver and copper nanoparticles using ascorbic acid and chitosan for antimicrobial applications. *Carbohydrate Polymers*, 2014, vol. 112, no. 0, pp. 195-202. ISSN 0144-8617.

[9] CALLEGARI, A. et al. Photochemically grown silver nanoparticles with wavelength-controlled size and shape. *Nano Letters*, 2003, vol. 3, no. 11, pp. 1565-1568. ISSN 1530-6984.

[10] JIN, R. et al. Controlling anisotropic nanoparticle growth through plasmon excitation. *Nature*, 2003, vol. 425, no. 6957, pp. 487-490. ISSN 0028-0836.

[11] LEE, S.-W. et al. Effect of Temperature on the Growth of Silver Nanoparticles Using Plasmon-Mediated Method under the Irradiation of Green LEDs. *Materials*, 2014, vol. 7, no. 12, pp. 7781. ISSN 1996-1944.

[12] OLDENBURG, S. J. Silver Nanoparticles: Properties and Applications. [online]. 15.3.2016 [cit. 15.32016 2016]. Available from: <http://www.sigmaaldrich.com/materials-science/nanomaterials/silver-nanoparticles.html#top>.

SOLID-STATE SYNTHESIS OF GOLD NANOPARTICLES ON DIMPLED ALUMINUM SURFACES AND THEIR TRANSFER TO TRANSPARENT SUBSTRATES

Tomáš Lednický
Advanced Materials and Nanosciences,
CEITEC - Central European Institute of Technology,
Brno University of Technology
E-mail: tomas.lednicky@ceitec.vutbr.cz

Supervised by: Mária Bendová
E-mail: maria.bendova@ceitec.vutbr.cz

Abstract: In this work, it is proposed a novel and cost-effective fabrication method for forming gold nanoparticle films on transparent substrates. The method involves fabrication of templated aluminium substrate by porous-type anodic oxidation of aluminium foils. Furthermore, a thin gold layer is deposited over and subsequently annealed which results into the gold nanoparticles formation. Additionally, it is proposed a transfer onto transparent substrates such as PDMS or PMMA. The goal of this research is to build a localized surface plasmon resonance (LSPR) sensor element with a high performance and low fabrication costs.

Keywords: gold, nanoparticles, aluminium, LSPR, SPR, PDMS

1. INTRODUCTION

Surface plasmons (SPs) are coherent oscillations of free electrons at the boundaries between a metal and a dielectric. They are often categorized into two classes: propagating surface plasmons (PSPs) and localized surface plasmons (LSPs) shown in Fig. 1. PSPs can be excited on metallic films, using several approaches involving the use of a prism coupler, an optical waveguides coupler, a diffraction gratin, and others. On the other hand, LSPs are easily excited on metallic nanoparticles (NPs) without a complex setup.

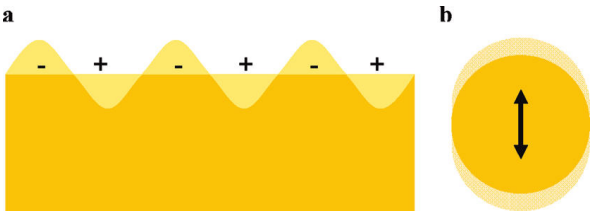


Figure 1: Schematic illustrations of (a) PSPs and (b) a LSP. [1]

Surface plasmons have attracted worldwide attention of scientists thanks to their ability to induce a strong enhancement of electromagnetic field in the near-field region, leading to an extensive application in surface-enhanced Raman scattering (SERS), fluorescence spectroscopy, refractive index measurements, biomolecular interaction detection, and many others. Thus, SPs have found application in research areas such as electronic data storage or light generation, but mostly in the detection of biological and chemical species, environmental monitoring, and medical diagnostics [2].

The wide spread of SP use into various fields in the past two decades was significantly stimulated by rapid developments in nanotechnology, which led to a new understanding of tuning the various

properties of nanomaterials, which can support SPs for specific applications. For example, surface plasmon resonance (SPR) is one of the widely used key optical phenomenon for the development of fast, sensitive, and economical devices [3].

Nowadays, the classical setup based on utilizing SPR on thin gold films using PSPs is a great success story, and besides research purposes it is widely adopted in medical industry (especially by pharmaceutical companies) for highly sensitive and high-throughput optical measurements. On the other hand, there is a beneficial advantage of localized surface plasmon resonance (LSPR) compared to SPR based on a transmission optical setup instead of a reflective one (SPR), which enables its easy integration into small hand-held monitoring devices. The sensitivity and detection limit of LSPR sensor elements are strongly correlated with the utilized nanostructures. In recent years, much attention was given to noble metal nanoparticles (NPs) with sizes well below 100 nm, being the active LSP element. This size parameter of the NPs is critical for LSPR detection in near infrared or even visible region.

However, nowadays, several technologies for fabrication of NPs are capable of developing extremely sensitive sensors, but the costs are unacceptably high, while simpler technologies result in sensors having poor sensitivity [1]. Therefore, the urgent challenge is to find optimal fabrication technologies combining realistic manufacture costs with high sensor sensitivities.

In this work is present a novel technological approach for synthesis of gold NP surfaces. This technology is based on a combination of two well-studied processes, one of them being the porous-type anodic oxidation of aluminum [4, 5] and the second one being the dewetting effect of thin gold layers [6, 7, 8]. The fundamental advantage is that both processes are **versatile**, could be performed in **large-scale**, and are of **low-cost**, being favorable as compared to other methods. The prepared NP films are characteristic by **well hexagonal alignment**, **high density**, and a relatively **narrow size distribution** of NPs as well as **organic-free composition**.

2. METHODOLOGY

The essential starting point is the fabrication of **nanostructured substrate**, which will be subsequently used as the template for **controlled dewetting of thin gold layer**. In this work, we present the use of dimpled aluminium substrate, being a secondary product of anodic oxidation of aluminium into porous anodic alumina (PAA), as shown in Fig. 2.

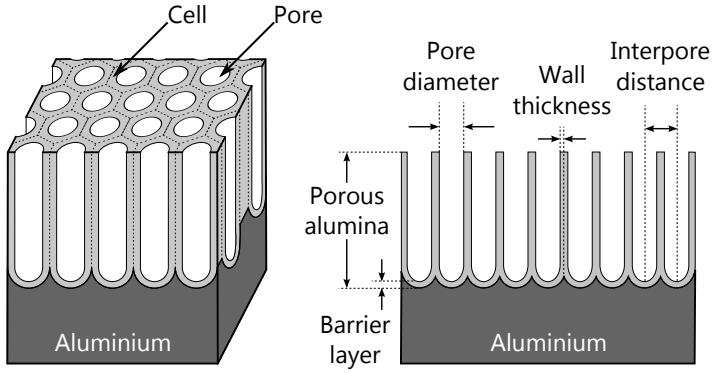


Figure 2: Schematic illustration of an idealized PAA film grown on aluminum substrate.

Typically, the formation of a PAA film is performed by anodizing of aluminium thin films or foils in an acid solution (e.g. sulphuric, oxalic, phosphoric acid). By controlling the electrochemical conditions (anodizing voltage, electrolyte composition, processing time, etc.), a variety of PAA parameters (cell size, pore population, pore order, etc.) can be tailored [5]. Since the aluminium template inversely

replicates the PAA bottom structure, these parameters play a crucial role for the final size, alignment, and homogeneity of gold NP layers. To reveal the Al template, it is necessary to selectively remove the formed PAA film. These processes of Al template fabrication are shown schematically in Fig. 3.

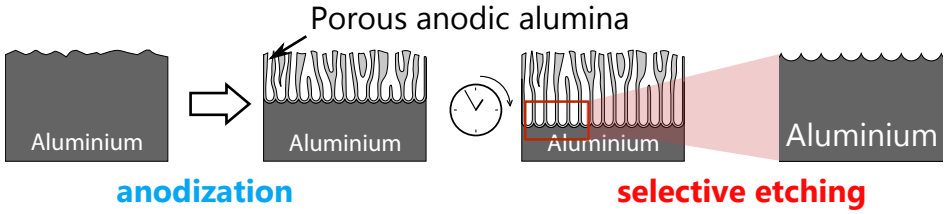


Figure 3: Scheme of aluminium template fabrication by formation and etching of PAA layer.

Generally, two starting materials are in consideration: a sheet of pure aluminium or a thin film of aluminium deposited over e.g. silicon wafer. The Al sheet has an advantage of its thickness that is necessary for forming a well-ordered PAA structure. However, the economic site and implementation in current electronic devices can be a potential issue.

The gold NP film is obtained by deposition of a thin gold layer over the Al template, followed by an annealing in a vacuum or even ambient atmosphere. This technique, called controlled dewetting, is well studied especially in the case of gold layers [7, 8]. After the NP layer is created, it can be transferred to a more suitable (transparent) substrate by its application over the layer and subsequent Al dissolution. This process is illustrated in Fig. 4.

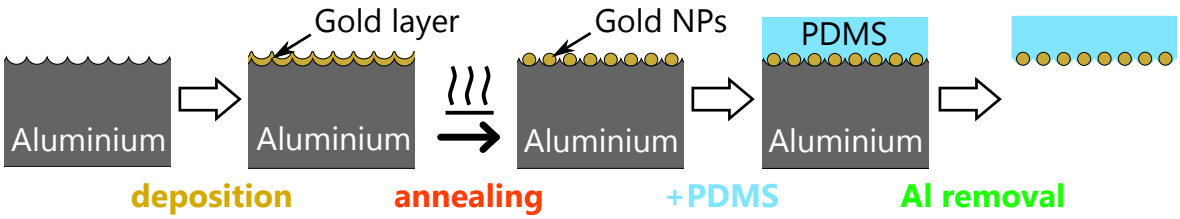


Figure 4: Schematic illustration of gold NP layer preparation on the structured aluminium template and their transfer to another substrate.

3. EXPERIMENTAL AND RESULTS

An aluminium foil of 99.999 % purity and 250 μm thickness (Goodfellow) was used as the starting material. The aluminium foil was annealed in a vacuum at 550 °C for 5 h to allow the material to relax and recrystallize. Subsequently, the foil was mechanically and electrochemically polished. The mechanical polishing was done under ethanol by an ultra fine P3000 sandpaper. The electrochemical polishing was performed in 1:4 (*v:v*) perchloric acid ($HClO_4$) and ethanol solution at 3 °C, without solution stirring for 1 min at 20 V (about 100 mA/cm^2). The Al foil was set as the anode and a stainless steel mesh was used as the cathode.

After polishing the foil was anodized in galvanostatic mode at 40 V in 0.3 M oxalic acid ($H_2C_2O_4$) at 7 °C for 18 hours. The long anodizing was set to obtain the best order of pores at the aluminium/alumina interface [9]. The nanostructured aluminium surface was revealed by selective etching of PAA in 0.42 M phosphoric acid (H_3PO_4) and 0.2 M chromium trioxide (CrO_3) solution at 60 °C for 1 h. An example of such dimpled aluminium template is shown in Fig. 5A. Furthermore, over the template was deposited by sputtering 7 nm thin gold layer (shown in Fig. 6A). This sample was further annealed in the ambient atmosphere at 350 °C for 2 h. Resulting NP surface is shown in Fig. 5B.

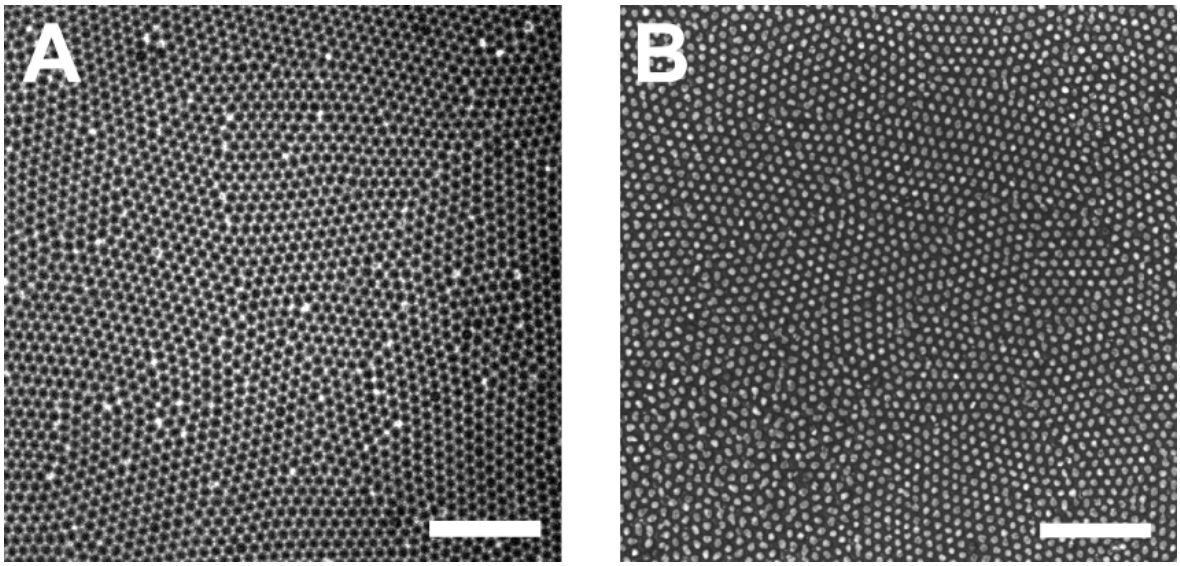


Figure 5: SEM image of dimpled aluminium surface after removal of PAA (A), and with gold NPs formed over the surface (B). The mean size of NPs is 75 nm. The scale bar is 1 µm.

The NP surface was easily transfered to a PDMS substrate by putting it face-down on gel-like solution and polymerized it on a hotplate at 60 °C for 2 h. When PDMS was hardened (shown in Fig. 6B), samples were immerse in 5 M hydrochloric acid (HCl) and 2 M copper(II) chloride ($CuCl_2$) to completely remove Al substrate. After the Al removal, the PDMS is covered by layer of NPs (shown in Fig. 6C), which in our case (NPs sizes) absorbs in red and infrared region, thus it appears blue or violet.

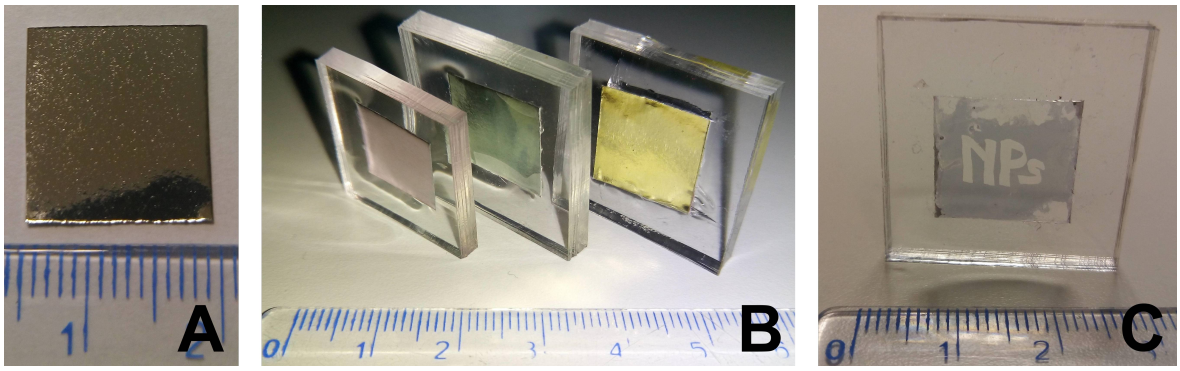


Figure 6: A photo of deposited thin gold layer over dimpled aluminium substrate (A), NP layer on Al substrates within PDMS matrices (B), and a separated NP layer on a PDMS matrix (C).

4. CONCLUSION

In this work, we have developed and demonstrated an interesting technological approach for synthesis of ordered gold nanoparticle layers. We have demonstrated that this technique could be performed in large-scale, and thus it can be unpredictably economical. Additionally, this technique compared to others, produce NPs which are well align, with a high density and organic-free. Furthermore, this technique can be extraordinarily versatile since NPs' sizes and distribution depends mainly on the template and the thickness of deposited layer, where both are undeniably tunable.

In the last decade, the demand for NPs, especially on layers with a high density, has increased im-

mensely due to rapid progress in fields of optics, especially the plasmonics. Nowadays, there are a vast number of scientific ideas as well as commercial devices utilizing SPR or LSPR for detection and recognition of various bio-molecules [10, 11, 12]. Therefore, it is greatly encourage to further develop and optimize NPs fabrication techniques such as this.

ACKNOWLEDGEMENT

I would like to acknowledge my supervisor Dr. Mária Bendová for giving me a space for development and research of fun ideas like this one, and Mr. Eiichiro Oda for his great contribution to the world.

REFERENCES

[1] Kathryn M. Mayer and Jason H. Hafner. Localized surface plasmon resonance sensors. *Chemical Reviews*, 111(6):3828–3857, 2011.

[2] Yong Chen and Hai Ming. Review of surface plasmon resonance and localized surface plasmon resonance sensor. *Photonic Sensors*, 2(1):37–49, 2012.

[3] Eleonora Petryayeva and Ulrich J. Krull. Localized surface plasmon resonance: Nanostructures, bioassays and biosensing – a review. *Analytica Chimica Acta*, 706(1):8–24, 2011.

[4] J. W. Diggle, Thomas C. Downie, and C. W. Goulding. Anodic oxide films on aluminum. *Chemical Reviews*, 69(3):365–405, 1969.

[5] Grzegorz D. Sulka. *Highly Ordered Anodic Porous Alumina Formation by Self-Organized Anodizing*, pages 1–116. Wiley-VCH Verlag GmbH & Co. KGaA, 2008.

[6] Carl V. Thompson. Solid-state dewetting of thin films. *Annual Review of Materials Research*, 42(1):399–434, 2012.

[7] Claudia Manuela Muller, Flavio Carlo Filippo Mornaghini, and Ralph Spolenak. Ordered arrays of faceted gold nanoparticles obtained by dewetting and nanosphere lithography. *Nanotechnology*, 19(48):485306, 2008.

[8] Minhee Kang, Sang-Gil Park, and Ki-Hun Jeong. Repeated Solid-state Dewetting of Thin Gold Films for Nanogap-rich Plasmonic Nanoislands. *Scientific Reports*, 5:14790, 2015.

[9] Hideki Masuda. *Ordered Porous Nanostructures and Applications*, chapter Highly Ordered Nanohole Arrays in Anodic Porous Alumina, pages 37–55. Springer US, Boston, MA, 2005.

[10] Eleonora Petryayeva and Ulrich J. Krull. Localized surface plasmon resonance: Nanostructures, bioassays and biosensing – a review. *Analytica Chimica Acta*, 706(1):8 – 24, 2011.

[11] Jacqueline Jatschka, André Dathe, Andrea Csáki, Wolfgang Fritzsche, and Ondrej Stranik. Propagating and localized surface plasmon resonance sensing – a critical comparison based on measurements and theory. *Sensing and Bio-Sensing Research*, 7:62 – 70, 2016.

[12] Jiří Homola and Marek Piliarik. *Surface Plasmon Resonance Based Sensors*, chapter Surface Plasmon Resonance (SPR) Sensors, pages 45–67. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

STUDY OF ELECTROCHEMICAL BEHAVIOUR OF GOLD NANOSTRUCTURED SURFACES

Hana Kynclová

Doctoral Degree Programme (4), FEEC BUT

E-mail: xkyncl02@stud.feec.vutbr.cz

Supervised by: Jan Prášek

E-mail: prasek@feec.vutbr.cz

Abstract: Recently, nanostructured surfaces have been extensively studied due to their possibility to use them in construction of electrochemical sensors and biosensors. We have prepared tungsten working electrode modified with gold nanocolumns via galvanic deposition of gold ions through anodic alumina nanoporous template. This nanostructured surface was subsequently electrochemically characterized by cyclic voltammetry using $K_3[Fe(CN)_6]/K_4[Fe(CN)_6]$ probe and compared with the results measured on a flat gold electrode. The electroactive areas for both types of electrodes were evaluated and compared. It was found that the nanostructured area affected the response on lower scan rates (up to 25 $mV.s^{-1}$) and increased the response at higher scan rates (higher than 50 $mV.s^{-1}$) which is probably related to penetration of electrolyte solution among nanocolumns. Nevertheless, this fact can help in speed of analyses of huge amount of samples.

Keywords: nanostructured electrode, electrochemistry, cyclic voltammetry, gold electrode, nanoporous alumina

1. INTRODUCTION

In recent years, fabrication of various nanostructured surfaces has become very actual topic. Great interest is especially about self-assembled arrays of nanocolumns because of their possible usage in many applications eg. photocatalysis, fuel cells, electrochemical sensors and biosensors [1]. In sensing applications of nanostructured surfaces could increase an active electrochemical area significantly while a small geometrical area is maintained and thereby improve sensitivity of sensors [2]. Electrochemical anodization of aluminium film combined with electrochemical deposition into nanoporous templates is well-established method for fabrication of nanostructured surfaces [1, 3]. Self–ordered anodic alumina oxide (AAO) offers highly ordered and vertically aligned nanoporous templates which are suitable for various material synthesis and alignment of nanosized structures [4]. This method also provides relatively cheap fabrication in comparison with lithographic and others methods.

Nanostructures-modified surfaces are usually characterized by scanning electron microscopy and by electrochemical methods. The electrochemical cleaning of nanostructured electrodes is key process before each electrochemical characterization. Voltammetric cycling in sulphuric acid is generally used for electrochemical cleaning of nanostructures. This cleaning removes impurities from electrode surface and improves wettability of surface. Electrodes are usually studied through the changes of cyclic voltammetry signal of ($K_3[Fe(CN)_6]/K_4[Fe(CN)_6]$) probe [5, 6].

In this study, we prepared and characterized gold nanostructured surfaces on tungsten substrate layer. Gold has many advantages eg. biocompatibility, chemical stability and high affinity of biomolecules (eg. thiols) to the gold. These properties predicted gold nanostructures for utilization in biosensors applications.

2. MATERIALS AND METHODS

2.1. FABRICATION OF NANOSTRUCTURED ELECTRODES

Firstly, the adhesive titanium layer (99.99%, Porexi, CZ) with thickness of 20 nm was sputtered on silicon wafer covered by SiO_2 with thickness of 500 nm in order to insulate the nanostructures from n-doped silicon wafer and thus to minimize junction effects at the interface between Ti and Si. The tungsten (99.999%, Porexi, CZ) layer of 200 nm and aluminium (99.999%, Goodfellow, UK) film of 150 nm were subsequently deposited by ion beam sputtering method (Bestec, Germany). Gold layer with thickness of 100 nm was also prepared for comparison of gold nanostructures impact on the enhancement of electrode active surface area. Self-assembled gold nanostructured surfaces were produced by pulse galvanic deposition of gold material from dicyanoaurate ($K[Au(CN)_2]$, gold content of 68.3% and purity of 99.9%, Safina, CZ) and boric acid solution (H_3BO_3, p., Penta, CZ) into AAO nanoporous template fabricated via electrochemical anodic oxidation method of thin aluminium layer according to our previous papers [5, 7, 8]. The step by step fabrication is schematically illustrated in Figure 1. In the first step, upper aluminium layer was anodically oxidized in 0.3 M oxalic acid ($(COOH)_2$, p.a. Penta, CZ) at 10 °C under the constant voltage of 50 V. During this process, aluminium layer was transformed to nanoporous alumina (Al_2O_3) with hexagonally ordered nanopores perpendicular to the surface. Whereas the anodization was being finished, tungsten oxide nanostructures were created on the bottom as a result of oxidation of tungsten layer below. Then these WO_3 nanostructures were etched in the phosphate buffer of pH 7 at 25 °C, which created tungsten nanodimpled surface with AAO above. These nanodimples enable to strongly anchor the gold nanostructures below the ambient surface and thus improve their stability. Anodic alumina template can be also etched in 5% phosphoric acid for nanopores widening. This treatment causes bigger diameter of resulting nanostructures. Pulse deposition of gold into such AAO with nanodimpled bottom was performed from dicyanoaurate and boric acid solution. The constant temperature of solution was 50 °C. The number of pulses controlled the length of obtained gold nanocolumns in the AAO template. The last step of production is dissolving of alumina template in the solution composed of chromium trioxide (CrO_3, p., Penta, CZ) and phosphoric acid (H_3PO_4, 98%, p.a., Penta, CZ).

Generally, the gold nanostructures shape varied according to the change of fabrication parameters. Namely, the anodization voltage proportionally influenced a diameter of nanopores in AAO template and thus the shape of deposited gold nanostructure the higher the voltage was applied, the wider the nanostructures were obtained. The length of nanostructures is influenced by the thickness of alumina template and deposition setup, namely the number of pulses.

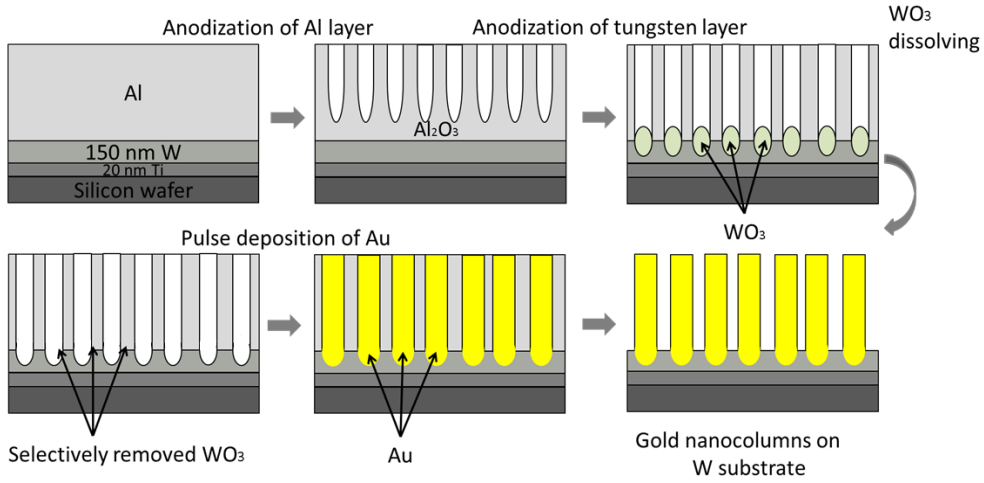


Figure 1: Fabrication process gold nanocolumns modified surface.

2.2. ELECTROCHEMICAL CHARACTERIZATION

All electrochemical measurements of nanostructured as well as flat gold surfaces were performed using μAUTOLAB III/FRA2 in connection with NOVA 1.11 software (Metrohm Autolab, NL). Two-electrode arrangement with Pt pseudoreference electrode (Metrohm AG, CH) was used for all experiments. The measurement was performed in a droplet of 20 μl volume of solution. The geometrical area of electrodes was determined by circle in parafilm with 3 mm in diameter (i.e. 7 mm^2). All fabricated electrodes were electrochemically cleaned by cyclic voltammetry methods performed in 0.1 M sulphuric acid (H_2SO_4, 96% pa, Penta) before each measurement. Cyclic voltammetry (CV) curves were measured in range from -0.5 V to 0.5 V with various scan rates. The measurements were proceeded in 0.1 M potassium chloride (KCl, 99.95%, p.a., Penta, CZ) with additions of 6 mM potassium ferricyanide ($K_3Fe(CN)6$, 99%, Sigma Aldrich, USA) and 6 mM ferrocyanide ($K_4Fe(CN)6$, 99%, Sigma Aldrich, USA). Deionized water (18.2 MΩ) obtained from Millipore Synergy system (Millipore Corp., USA) was used for preparation of all solutions.

3. RESULTS AND DISCUSSION

The tungsten film electrodes modified by gold nanocolumns were successfully prepared by anodization method with subsequent electrochemical deposition of gold material to the prepared alumina templates. The reproducibility of fabrication process of Au nanostructures was excellent. Before electrochemical characterization, the homogeneity of surface coverage with gold nanostructures was verified using SEM (Tescan Mira LMU II, CZ) analysis (Figure 2). According to the SEM analysis, the nanocolumns have about 60 nm in diameter and about 200 nm high.

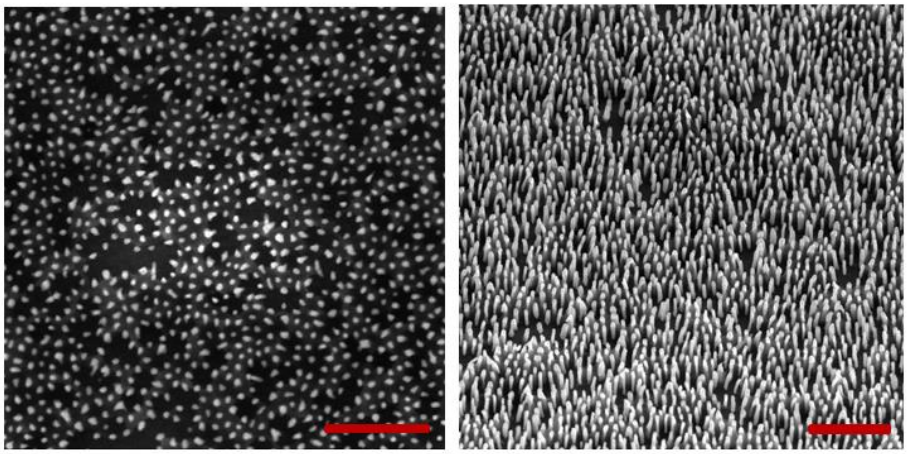


Figure 2: SEM images of prepared nanostructured surfaces. Top view (on the left) and side view (on the right), scale bar 1 μm.

Fabricated nanostructured surfaces were electrochemically characterized by electrochemical method namely cyclic voltammetry (CV). It has been found, key role in electrochemical characterization is electrochemical cleaning which influences the resulted data significantly. In our case, the cleaning was performed by cycling of all electrodes in 0.1 M sulphuric acid.

Cyclic voltammetry performed in redox electrochemical system is useful for evaluation and comparison of electrochemically active surfaces of working electrodes. In this study, CV analysis was performed in 0.1 M potassium chloride with concentration of 6 mM of potassium ferro-ferricyanide redox couple. It is obvious that the electroactive electrode area of electrodes is often very different than geometrical electrode area due to roughness of electrode surface or various surface modifications. Evaluation of CV results can reveal a real penetration of electrolytic solution into the space among the gold nanostructures. The higher current response of CV voltammogram means the higher electroactive area involved in reaction. The peak height is also dependent on the electrolyte concentration. CV voltammograms for flat gold electrode without any modification and for nanostructured surface with gold nanostructures are shown in **Figure 3**. The heights of peaks were determined.

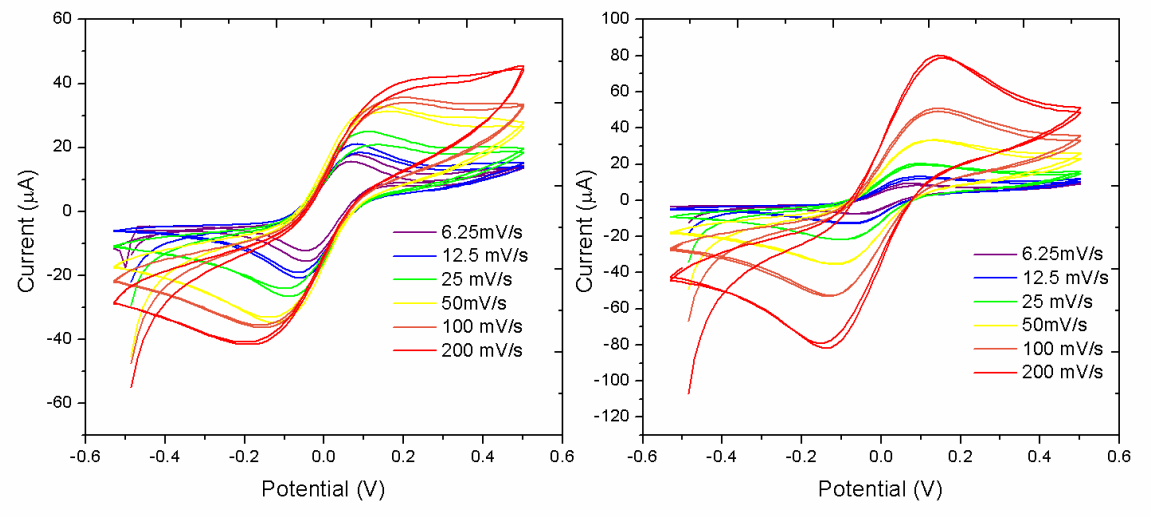


Figure 3: Cyclic voltammograms measured in 6 mM ferro–ferricyanide in 0.1M KCl. Flat gold electrode (on the left) and electrode modified with gold nanostructures (on the right).

In Figure 4, there are dependencies of cathodic peak heights on square root of scan rates for both types of measured electrodes. Both dependencies are linear. We can also observe that flat gold electrode exhibits higher peaks at low scan rates (< 25 $mV.s^{-1}$) and nanostructured surface shows higher peaks at higher scan rates (≥ 50 $mV.s^{-1}$). It is obvious; electroactive area of nanostructured electrode is changed with scan rate. Higher scan rate means higher electroactive area than electroactive area for flat gold electrode and vice versa. Nevertheless, the electroactive area of electrodes is also evaluated as a ratio of the slopes of dependencies of peak heights on square roots of scan rates. According to this fact, nanostuctured electrode exhibits approximately 2.5 times higher electroactive area.

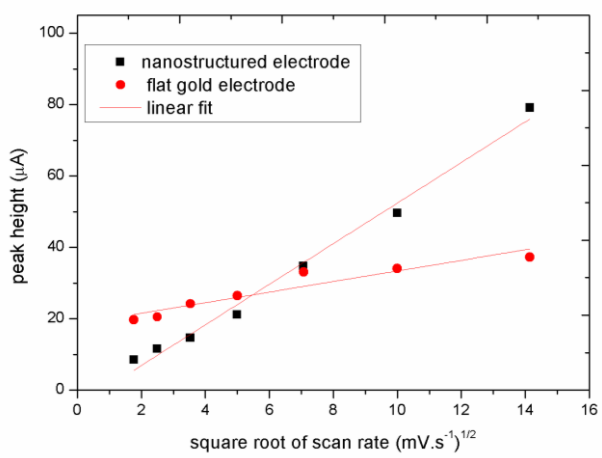


Figure 4: Dependencies of peak heights on square root of scan rates for two types of measured electrodes.

The change of the electroactive area depending on the scan rate is probably related to difficult penetration of electrolyte to the gaps between nanostructures at lower scan rates because the kinetics of mass transport near the interface between electrode and electrolyte plays an important role influencing the current responses of electrochemical sensors [9]. Anandan et all. [10] studied bare gold nanopillar array electrodes at various concentrations of redox $K_4Fe(CN)_6$. Electrodes with different nanopillar heights exhibited almost the same sensitivity response, although the sensitivity of nanopillar array electrodes was much higher than sensitivity of the flat electrode. It was speculated that the electroactive species of the redox probe may had difficulties in its transport to the small spaces between the nanopillars as a result of a low diffusivity or a fast-electron transfer rate constant. When the diffusivity is low, it is difficult for ions to diffuse deep into the small spaces between the nanopillars, and when the electron transfer rate constant is high, most of the species will

get oxidized near the top ends of the nanopillars before it gets diffused deep down the gaps. Under these circumstances, it is possible that only the top ends of the nanopillars are serving their active duty in transferring electrons to oxidize redox probe. They declared the situation for functionalized nanopillar array electrodes is quite different. The sensitivity of these nanopillar array electrodes in glucose detection increased as the height of the nanopillars increased.

4. CONCLUSION

The tungsten electrodes modified with gold nanocolumns were prepared by electrochemical anodization and subsequent deposition of gold into nanoporous alumina templates. The flat gold electrodes were also prepared for comparison with nanostructured surfaces. All electrodes were electrochemically cleaned and then studied by CV in the presence of potassium ferro-ferricyanide probe to determine the electroactive area of nanostructured electrode. It was found, the electroactive area is changing with scan rate. Nanostructured surfaces showed impact at higher scan rates (from $50mV.s^{-1}$). This fact is probably caused by difficult penetration of electrolyte to the spaces between nanostructures at lower scan rates.

ACKNOWLEDGEMENT

This work was supported by project no. FEKT-S-14-2300 A new types of electronic circuits and sensors for specific applications.

REFERENCES

[1] ANANDAN, V., et al. Nanopillar arrays with superior mechanical strength and optimal spacing for high sensitivity biosensors, In 2005 NSTI Nanotechnology Conference and Trade Show 2005, 217-220.

[2] LEE, S.J., et al. Electrochemical fabrication and evaluation of highly sensitive nanorod-modified electrodes for a biotin/avidin system, Biosensors and Bioelectronics, 2008, 23, p. 1117-1124.

[3] BHARATH BHUSHAN, et al. Handbook of Nanomaterials Properties, 2014, Springer Verlag, Heidelberg.

[4] NOH, K., et al. Guided nanostructures using anodized aluminum oxide templates, Nano, 2011, 6, p. 541-555.

[5] DRBOHLAVOVA, J., et al. Gold Nanostructured Surface for Electrochemical Sensing and Biosensing: Does Shape Matter?, Analytical Letters, 2016, 49, p. 135-151.

[6] BAKMAND, T., et al. Fabrication and Characterisation of Membrane-Based Gold Electrodes, Electroanalysis, 2015, 27, p. 217-224.

[7] HRDY, R., et al. Electrochemical Impedance Spectroscopy Behaviour of Guanine on Nanostructured Planar Electrode, International Journal of Electrochemical Science, 2013, 8, p. 4384-4396.

[8] KYNCLOVA, H., et al. Nanostructured Surface Effect of Electrode on Doxorubicin Determination. In BABU, A.S., Mechanical Design and Power Engineering, Pts 1 and 2, 2014, vol. 490-491. Stafa-Zurich: Trans Tech Publications Ltd, pp. 262-266.

[9] ANANDAN, V., et al. Nanopillar array structures for enhancing biosensing performance, International Journal of Nanomedicine, 2006, 1, p. 73-79.

[10] ANANDAN, V., et al. Role of reaction kinetics and mass transport in glucose sensing with nanopillar array electrodes, Journal of Biological Engineering, 2007, 1, p.

SUPERCAPACITOR PARAMETERS DEPENDENCE ON CHARGING DURATION

Tomáš Kuparowitz

Doctoral Degree Programme (3), FEEC BUT

E-mail: xkupar01@stud.feec.vutbr.cz

Supervised by: Petr Sedlák

E-mail: sedlakp@feec.vutbr.cz

Abstract: Supercapacitor (SC) parameters dependence on charging time is analyzed for Nesscap 2.7V / 10F. Charging was performed using constant current in the range from 0.5 A to 5 A. Reciprocal capacitance (C^{-1}) was calculated. It linearly decreases with applied charge. During SC charging, the effective distance of charges on electrode-electrolyte interface decreases, due to the charge distribution being altered by mobile ions inside electric field. Ions are at first moving by drift, and then by diffusion processes. Effective thickness of electrolyte negative charge is determined from reciprocal capacitance. The effective thickness varies between 2.6 nm and 2.1 nm.

Keywords: Supercapacitor, electrolyte effective thickness, constant current charging, supercapacitor charging characteristic evaluation, supercapacitor charging duration, ultracapacitor

1. INTRODUCTION

This paper gives an analysis of the impact of supercapacitor (SC) charging duration on effective thickness of electrolyte negative ions in close proximity to its terminals. There were three samples evaluated for the purpose of this paper, denoted TN-1, TN-2, and TN-3. All of which were of-the-shelf SC NessCap 2.7 V / 10 F.

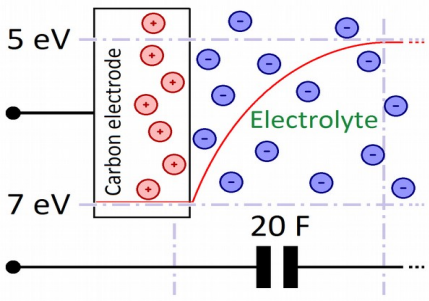


Figure 1: Diagram of single electrode of SC Nesscap 2.7 V / 10 F at steady state

SC structure consists of two porous carbon based electrodes. The electrode effective area of evaluated samples is roughly $A = 200$ m^2. The space between both electrodes is filled with an anion-rich electrolyte, and cellulose separator. The permittivity of the electrolyte turns out to be roughly $\varepsilon = 10^{10}$ F/m. And we suppose its work function to be 7 eV. The diagram of single SC electrode is presented in figure 1. Note that circa 2 eV energy gap exists in the depletion region (between 5 eV carbon electrode and the electrolyte) which is responsible for gradual change in ionic charge density in close proximity to the electrode, as depicted by red curve. An effective capacitance is created on such interface. This capacitance turns out to be roughly 20 F for SC Nesscap 2.7 V / 10 F, as there are two such capacities in series, one at each electrode. [1-2]

SC behaves as a regular capacitor with very high capacitance. But several phenomena, attributed to its irregular structure and composition, are observed. Notably they are voltage relaxation and recovery. Both are well accounted for in the state of the art SC model as described in [1]. Other phenomenon is voltage dependent capacitance, which is described in [2].

Overview of SC charging by constant current 2 A up to its rated voltage 2.7 V is presented in fig. 2. Two things are to be noted. First is the voltage relaxation, taking place since the charging has ended (denoted in black). This is caused by the diffuse capacitance charging. Second is slight bend in the charging curve. Note that in the case of regular capacitor, the charging would be strictly straight, as shown by green line. This phenomenon will be fully addressed later on in this paper.

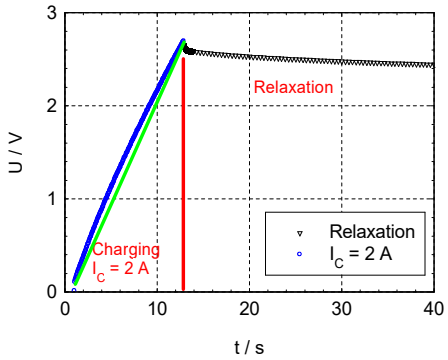


Figure 2: Charging of Nesscap 2.7 V / 10 F sample TN-1 by constant current 2 A (blue) and snippet of subsequent relaxation after charging (black)

The reminder of this article is divided as follows. In next section "Experimental setup", a description of the measurement hardware is presented. Section "Measurement" holds and explains the measured results. Section 4 contains parallel plate capacitor equivalent parameters evaluation. And section "Discussion" attempts to tackle the physics behind the charging phenomenon.

2. EXPERIMENTAL SETUP

The essence of the measurement is to charge SC by constant current, until the voltage on its terminals reaches its rated voltage. For this task a µ-precision programmable current regulator was constructed. The measurement setup is depicted in figure 3.

The measurement is controlled by PC, via a handyscope HS-3 (TiePie engineering) device. This device has 2 separate voltage channels, capable of streaming measurement at 10 kHz, which are used for measured voltage and current values acquirement. And an arbitrary function generator channel, used to set the limiting current.

The regulator, labeled CS 1 in the figure, is capable of supplying precision current in the range from -5 A to 5 A. To achieve negative currents, an H-bridge configuration is used. The current value and direction through the sample is measured using shunt resistor. Its dynamic transition period is greatly bellow 10 ms, even for full current swing. The current set error is in the neighborhood of 0.3% (induced by HS-3 device). And its load regulation is bellow 1 µA. These error values are negligible in the greater picture of transferring over 20 C of charge into the sample.

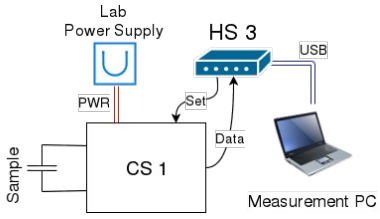


Figure 3: Measurement setup diagram

The current regulator is supplied by standard lab power supply, capable of producing 5 A at 12 V. The power line needs to be decoupled using a π-filter configuration, in order to achieve fast transition. An additional bias power supply, having 5 V high noise lane (cooling) and ±17 V symmetrical supply for powering the on board electronics, is used.

3. MEASUREMENT

Three samples were evaluated for the purpose of this paper. While all the calculations were performed for every sample, only sample TN-2 was chosen as median to be presented graphically.

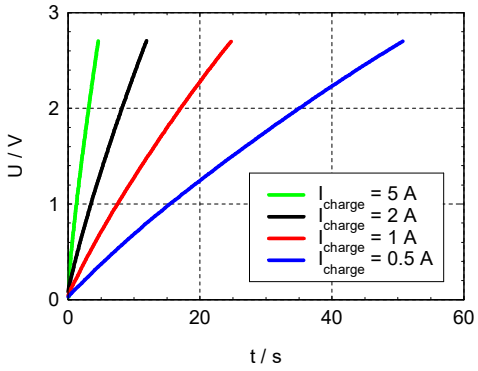


Figure 4: Charging of Nesscap 2.7 V / 10 F sample TN-2 using varying currents in time

Figure 4 shows charging curves of sample TN-2. Four distinct charging currents were used. The duration of charging Δt varies from 4.6 s to 50.8 s, honoring equations (1) and (2) for regular parallel plate capacitor. For transferred charge it holds:

$$Q = It \quad [C] \tag{1}$$

where I is current flowing through the capacitor and t is the time duration for which the current is applied. For capacitor electrodes potential it holds:

$$V = \frac{Q}{C} \quad [V] \tag{2}$$

where Q is charge stored in capacitor and C is its capacity. The non-linear behavior of SC charging curve may be approximated by 2nd order function: [2]

$$Q = C_H V + \frac{1}{2} C_{H1} V^2 \quad [C] \tag{3}$$

where C_H is Helmholtz or immediate capacitance and C_{HI} is constant which characterize voltage dependence of SC capacitance. Parameters for $x = ay^2 + by + c$ quadratic fit of charging characteristics for all samples and currents are given in table 1.

Table 1: Charging characteristic quadratic fit parameters for all samples

Charging current:	0.5 A			1 A			2 A			5 A		
Sample	a	b	c	a	b	c	a	b	c	a	b	c
TN-1	-0.0002	0.0641	0.0607	-0.0010	0.1304	0.0755	-0.0040	0.2632	0.1267	-0.0244	0.6556	0.2675
TN-2	-0.0002	0.0640	0.0591	-0.0010	0.1298	0.0753	-0.0040	0.2631	0.1280	-0.0281	0.6767	0.1764
TN-3	-0.0002	0.0636	0.0595	-0.0010	0.1294	0.0652	-0.0040	0.2644	0.1180	-0.0262	0.6596	0.2525

Note that the parameter a, responsible for the hyperbolic bend of the charging curve, is consistent over all samples at the same charging current. Only exception is charging by 5 A, at which the joule heat generated inside SC is the culprit. Parameter b, the slope of the charging curve is also consistent, as it reflects the constant aggregation of charge in time.

4. RECIPROCAL CAPACITANCE AND EFFECTIVE THICKNESS EVALUATION

Given the charging fit parameters, SC reciprocal capacitance may be calculated from (3). First, the charging curve is to be converted from voltage vs time to voltage vs charge relation, using (1). Then from (2) it follows that $C^{-1} = V / Q$. Calculated reciprocal capacitance characteristics for sample TN-2 are shown in fig. 5 left.

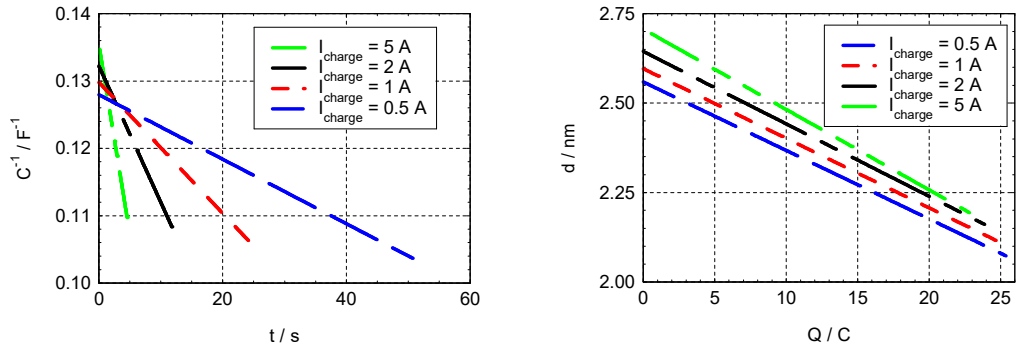


Figure 5: Reciprocal capacitance and effective thickness of SC Nesscap 2.7 V / 10 F sample TN-2 for varying currents

We will model SC behavior through reduction of the problem into system of single plate of parallel plate capacitor, as this fits well the system with very slow charges movement. For distance between electrodes of regular capacitor it holds that:

$$d = \varepsilon A C^{-1} \tag{4}$$

where ε is electrolyte permittivity and A is the effective area of its electrode. Evaluated effective area of sample TN-2 for varying currents can be found in fig. 5 right.

5. DISCUSSION

From fig. 5 it is obvious, that SC charging duration is inversely proportional to the amount of stored charge and increases linearly with time. We propose, that this is caused by anion buildup in the depletion region of anode during charging. Using Gauss's law, the relation of electric intensity over area can be linked to unidirectional negative charge distribution around the electrode by following set of equations:

$$\oint \vec{D}\, d\vec{A} = Q \quad \Rightarrow \quad \nabla \vec{E} = \frac{\rho}{\varepsilon} \quad \Rightarrow \quad E(x) = E_0\left(1 - \frac{x}{d}\right) \quad \text{where} \quad E_0 = \frac{e \cdot n \cdot A \cdot d}{\varepsilon \cdot A} \tag{5}$$

where ρ is charge density, n is negative charges count at distance x from the electrode, and E_0 stands for electric field intensity adjacent to the electrode.

For capacitor, with slowly moving charges, the charge density remains almost constant. The relation of its voltage to its electric field intensity could be inferred from (5):

$$V(x) = \int_0^x E(x)\, dx = E_0 x - \frac{e \cdot n}{2\varepsilon} x^2 \tag{6}$$

The relation of potential distribution between capacitor electrodes is depicted in 6 left. The red curve represents initial electric field intensity vector. By substituting (2) in (6) value of capacitor potential distribution may be obtained, depicted in green for both cases (parallel plate and SC).

The last term of (6) predicts the electric field intensity evolution in depletion region of parallel plate capacitor (fig. 6 middle, brown line). The blue line reflects the constant charge density of parallel plate capacitor.

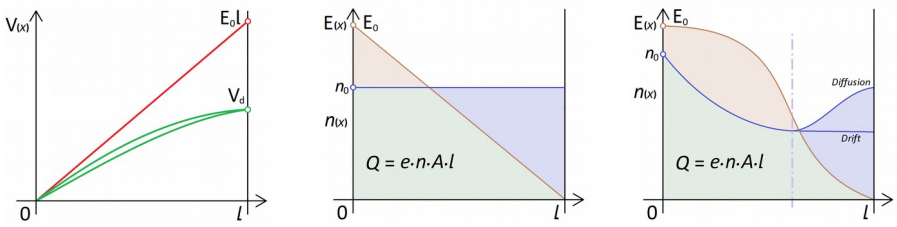


Figure 6: Depletion region potential distribution (left), with consequent electrode voltage marked V_d.. Charge density and electric field evolution in double plate capacitor (middle) and SC (right). Vs electrolyte depletion region fetch.

The relation between second and third term of (6) must be applicable to SC. Therefore it can be assumed, that the charge density changes around electrodes during the charging ever so slightly. This charge buildup (blue curve of fig. 6 right) is caused by two effects. First, charges move freely in electrolyte and are therefore affected by electric field, hence a drift towards the electrode occurs. Second, charge low density region is produced by the drift, which is slowly replenished by diffusion. The latter process is well described in [1]. In the case of discharged SC, situation gradually changes from that described in fig. 6 middle to that in fig. 6 right. During this process an extra charge is slowly accumulated in the area enclosed by *Drift* and *Diffusion* lines of fig. 6 right

The consideration of this new dual principle can account for aforementioned bend in SC charging characteristic. Since the charge density mainly changes by drift, which is in turn dependent on time, the duration of charging plays a role in the amount of charge stored by charging of SC. Note that the value of Q in figure 6 middle and right would not be equal. Also note that the right figure is grossly over exaggerated, in order to pinpoint the electrical intensity inflection point, which is exactly at that distance from electrode where the drift and diffusion lines join (marked by gray).

6. CONCLUSION

Measurements of several samples of SC Nesscap 2.7 V / 10 F were performed. SC charging was studied and its reciprocal capacitance was calculated. From which the effective electrolyte thickness was estimated. The effective thickness varies from 2.6 nm to 2.1 nm and increases linearly with the length of the charging, which is new effect observed in SC.

During the charging negative ions are slowly moved by drift and diffusion processes towards the electrode, as they are affected by electric field, and cause both increase in capacitance and decrease of electrolyte effective thickness. In other words, the longer the charging takes place, the larger distance do ions traverse in electrolyte. This effect increases charge density in close proximity to the electrodes. This new dual effect is in perfect congruence with physics behind Maxwell's equations.

ACKNOWLEDGEMENT

This research activity was also supported by the European Centers of Excellence CEITEC CZ.1.05/1.1.00/02.0068.

REFERENCES

[1] Sedlakova V, Sikula J, Majzner J, Sedlak P, Kuparowitz T, Buergler B and Vasina P 2015 Supercapacitor equivalent electrical circuit model based on charges redistribution by diffusion J. Power Sources 286 58–65

[2] Kuparowitz T, Sedláková V, Szewczyk A, Hasse L and Smulko J 2014 Charge Redistribution and Restoring voltage of Supercapacitors Electroscope 2014 1–7

THE ANALYSIS OF SOLDER PREFORMS IN SURFACE MOUNT ASSEMBLY

Václav Novotný, Radek Vala

Doctoral Degree Programme (2), FEEC BUT

E-mail: novotny.vaclav@azd.cz, radek.vala@sanmina.com

Supervised by: Josef Šandera

E-mail: sandera@feec.vutbr.cz

Abstract: This article is focused on the sphere of solder joints reliability in technology of surface mount assembly and the issue of manufactory of solder joints with solder preforms is introduced closer. It also describes types of solder preforms, their application and different methods of using in surface mount assembly. Text further comprises comparison of the solder joint which is manufactured by printing of solder paste through the cooper stencil and solder joint manufactured by applying of solder preform. The results of these basic processes are compared with method of solder fortification. In conclusion, there are evaluated outputs from experimental manufactory of solder joints with solder preforms and complemented by outputs which are achieved by x-ray machine

Keywords: Solder, Preforms, Reliability, Fortification, Wettability

1. INTRODUCTION

The technology of surface mount assembly is very wide area in these days and new manufacturing processes are included constantly in development and subsequently used in new projects. Surface mount technology involves two main aspects which are SMD components and soldering technology. Technology of soldering further comprises several subchapters where are important mainly method of reflow, soldering profile, methods of mounting of components and finally methods of printing of solder paste. There are two main ways which are used for printing of solder paste. The first method is printing of solder paste through the stencil and it is most often method because it is way of many large companies with electrical manufactory. Production of copper stencils for printing paste is very expensive but the final price of stencil is divided in large series of products. The second method is printing paste by automatic dispenser and it is less frequent and less represented because this method is used mainly by small and medium companies, where production is focused on small or medium sized series of products. The production of stencil for printing is unprofitable in these cases and automatic dispenser is appropriate and cost-effective alternative. This is because there is no need to use stencil for printing and solder paste is printed variably by current requirement.

There also exists an alternative besides these classical methods of printing of solder paste. This alternative is the use of solder preforms. Solder preforms have large range of advantages and disadvantages. It may seem that weaknesses are often bigger than advantages but it is clear that solder preforms are good choice for some specific applications. Solder preforms have large number of different kinds but the main advantage can be identified as mounting of very accurate volume of solder paste to solder joint and another good feature is precise application of flux and this is perfect for testing of reliability of solder joints, where precise of volume of solder paste is very important for final evaluating. Solder preforms are mounted by conventional mounting machine and for some companies isn't necessary to own expensive automatic dispenser for printing of solder paste or machine for stencil printing. For this reason, researching and analyzing of solder preforms can be useful in some applications. [2]

2. APPLICATION AND METHODS OF SOLDER PREFORMS USING

Technology of solder preforms provides very large number of different applications. These preforms are supplied by many manufacturers and it is possible to say that manufacturers of common types of solders provide also solder preforms and these are possible to divide to three basic parts. First part is sphere of solder preforms which can be used for attainment of precise volume in solder joint. It's not critically important in manufactory of large series of products, but it is very important for developmental issues and tests of reliability of solder joints [3]. Solder preforms are typically distributed by manufacturer in usual reel which are compatible with all SMD mounting machines. This method allows mount pieces of solder by the same way as other SMD parts. Similarity with SMD components is also in the fact that various sizes of solder preforms are compatible with sizes of SMD parts. This means that it is possible to find solder preforms with size 0603 and compare with SMD part 0603, both will have same size. There is also need to respect that sizes of solder preforms correspond with sizes of SMD components, but they don't correspond with sizes of footprints, respectively sizes of solder pads. It means, that for each solder pads must be choose the correct size of solder preform. [5]

This can be explained by using of SMD component 2512 for example. SMD component 2512 has typical dimensions 6,3 x 3,1 mm and her solder pads has dimensions 2 x 3,2mm. This means that there is need to achieve volume of solder paste 0,96 mm^3 through cooper stencil 150 µm. Volume of solder paste 0,96 mm^3 for SMD part 2512 can be obtained by application of solder preforms in size of 0603. Dimensions of this preform are 1,6 x 0,85 x 0,45 mm and the total volume of solder paste is 0,612 mm^3 and this result is comparable with volume of solder paste achieved by cooper stencil.

Name	Size (Inches)	Size (Metric)	Shape	Solder weight (each)
-	"	mm	-	g
0201	0.010 x 0.020 x 0.010	0.254 x 0.208 x 0	Rectangle	0.00024
0402	0.020 x 0.040 x 0.019	0.508 x 1.01 x 0.408	Rectangle	0.00182
0603	0.030 x 0.060 x 0.031	0.76 x 1.52 x 0.787	Rectangle	0.00672
0805	0.050 x 0.070 x 0.050	1.27 x 2.03x 1.27	Rectangle	0.0241
1206	0.060 x 0.120 x 0.060	1.52 x 3.04 x 1.52	Rectangle	0.0521

Table 1: Typical names, sizes, shapes and weights of most used solder preforms [5]

Previous solder preforms have one disadvantage in absence of flux in solder. This means that there is requirement to add flux by another way. There are several options to solve this problem for some type of applications and one is using of solder preform which is coated by flux and also is possible to use solder preform in fortification application. Solder fortification is method of using solder preforms in conjunction with solder paste and flux in solder paste is suffice for perfect wetting of solder pads.

Generally, flux is added to the preforms up to 3% by weight; however, much lower amounts of flux rating about 0,5% can be adequate to remove the oxides and promote excellent wetting. The more flux that is used to make the joint, the more residue is either left on the board in a no-clean application or has to be removed if clearing is required. [1]. Solder preforms without flux and solder preforms coated by flux can be recognized by colour difference because there is often added colorant. It is important mainly in case, that are use two types of preforms where are same dimensions and different material composition, for example Pb-Free and Pb-Containing solder. [1]

Finally, it should be mentioned that solder preforms are compatible with solder paste in manufacturing processes. It means that PCB assembly can work with both technologies, solder paste and solder preforms because there can be applied same method of reflow and same thermal profile can be used.

3. TEST OF SAC305 0603 SOLDER PREFORMS ON 2512 SMD COMPONENTS

Solder preforms were experimentally tested on ENIG 70μm FR-4 board with 40 100 Ω SMD 2512 resistors [4] [5]. These resistors with this value were chosen because test boards are ready for cycling in the thermal chamber and evaluation circuit can work only with 300 Ω resistance and any larger resistance will be considered as the crack in solder joint. Testing board includes 20 resistors soldered with Indium SAC305 0603 solder preforms (column A and B on board) and 20 resistors soldered with Indium8.9HF solder paste (column C and D on board) [6]. This is shown on figure (1). Process of manufactory of test board is composed from several steps. Firstly, the flux was printed through the stencil on half of test board and by the same way was printed solder paste through same cooper stencil and to solder paste and flux are mounted SMD 2512 components by SMT Manipulator. Finally, there are mounted 0603 solder preforms on half of test board, where is applied flux. Complete test board as shown on figure (1c) is reflowed by vapors in ASSCON reflow chamber.

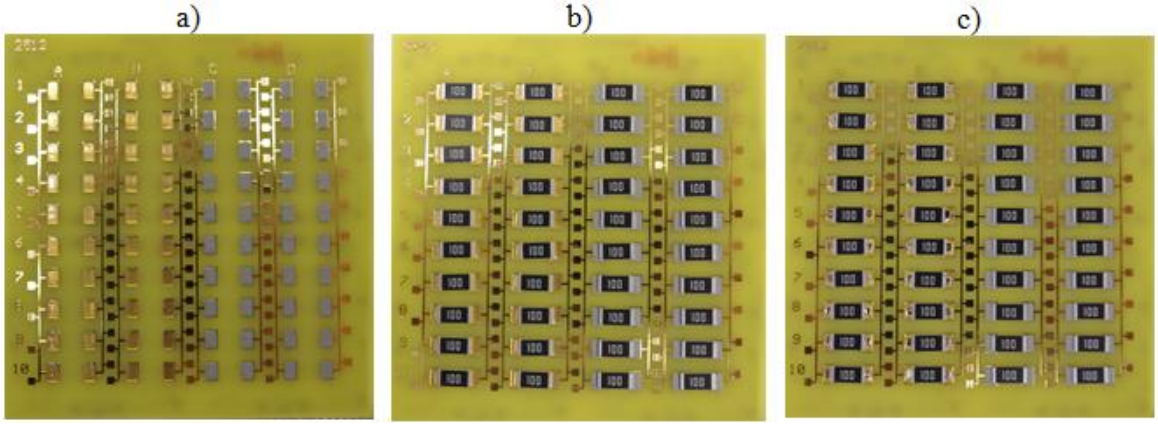


Figure 1: Process of manufactory of solder joint with solder preforms: a) Printed flux and solder paste through the stencil b) Mounted SMD 2512 components to flux and solder paste c) Added solder preforms to flux and SMD 2512 components [4]

Solder joints manufactured by two methods were compared after soldering in vapors and there was confirmed assumption that solder joint with solder preform will have larger volume of solder in solder joint and it can be seen that wettability of solder preforms was better than wettability of solder paste. This is shown on figure (2). Here it is important to add that material of solder preform and solder joint is same, because there is same material composition 96.5Sn/3.0Ag/0.5Cu and manufacturer is also same. Another advantage of solder joint with preform is fact that he has more precise volume of solder and therefore there will ensure the repeatability of manufacture at higher level.

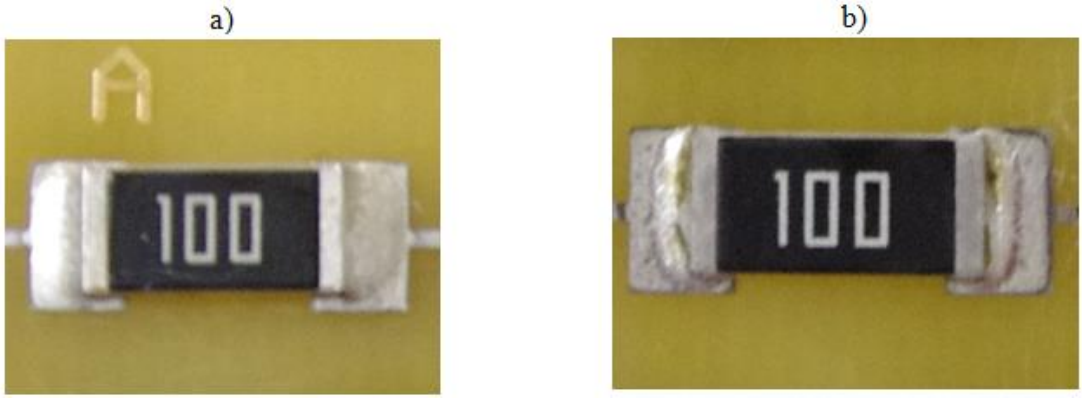


Figure 2: Solder joints soldered by different methods – a) Solder joints with solder preforms in the flux b) Solder joints with solder paste printed through the stencil

The visual analysis of soldered joints with solder preforms and solder paste was further complemented by analysis on x-ray machine. This analysis confirmed that solder joints with preforms have very good wetting of solder and further it is possible to see the integrity of the solder and the amount of voids in solder joints. Solder joints with solder preforms contain amount of fine desirable voids and this is prerequisite for good quality and reliability in operation. This is shown on figure (3a). There is difference compared to solder joints soldered with solder paste where is amount of large voids and in places of these voids can start grow crack of solder joint in operation. This is shown on figure (3b). The difference in the amount of voids is caused by different technologies of solder joints. This means that the flux in the solder paste evaporates during soldering and trying to get to the surface of solder paste while voids grow.

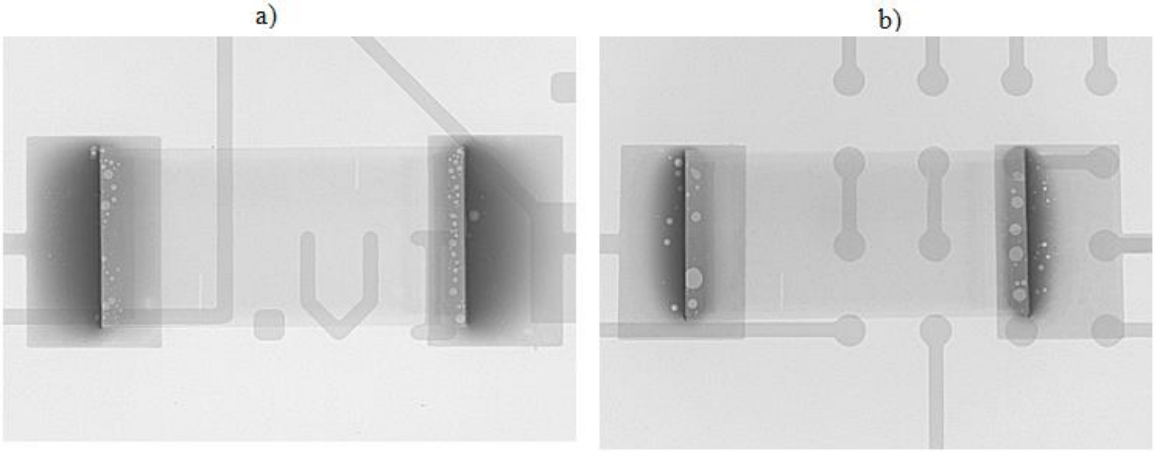


Figure 3: X-ray analysis of solder joints soldered by different methods – a) Solder joints with solder preforms in the flux b) Solder joints with solder paste printed through the stencil

4. TEST OF SOLDER FORTIFICATION METHOD

Solder fortification is method for reinforcement of solder joint for special applications. The main area of application of solder fortification is in THT (through-hole) assembly or in combination SMT (surface mount technology) and THT. There may be problem in stencil printing because cooper stencils for printing of solder paste are very thin due increasing miniaturization and density of components but there are still relatively large components as 2512 SMD parts or chip packages as TQFP-44, SOIC14, SOIC28 and many others [2]. These components need sufficient volume of solder for high-quality and reliable solder joint while most of current stencils have thickness maximally 300 µm or less. It can solve solder preforms in solder fortification configuration and so it is possible to reinforce critical areas quickly and flexibly in PCB. [1]

Solder paste has also disadvantage that she contains only about 50 % volume of allow powder. Process of manufactory of solder joint as fortification is very simple and was also tested on SMD 2512 resistors and same PCB FR-4 board with ENIG 70µm. It was made experimentally only for one 2512 SMD resistor and will be tested further. For this reason, x-ray analysis was not performed and it is not possible to evaluate voids in solder joint now, but very good wettability and shape of joint was confirmed by visual analysis and therefore there is presumption of high reliability of this type of solder joint.

Preparing of solder joint can be seen on figure (4a). Firstly, the solder paste was printed through 150 µm cooper stencil, then was mounted SMD 2512 resistor and finally was mounted two SAC305 0603 solder preforms on each solder pad by SMT Manipulator and complete assembly was soldered in vapors. Result after soldering is shown on figure (4b).

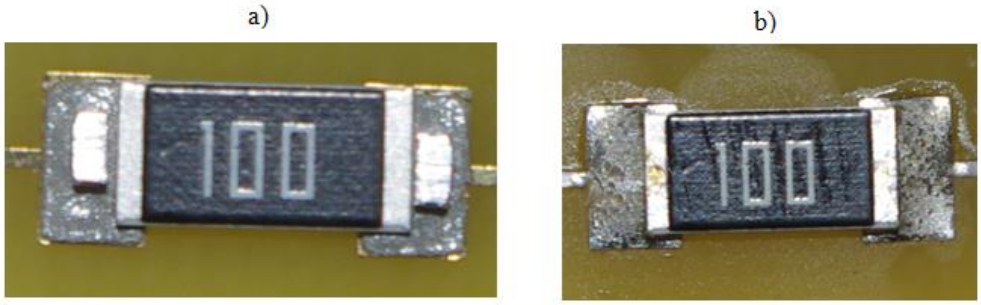


Figure 4: Solder fortification method – (a) Solder preforms mounted to paste before soldering in vapors (b) Solder preforms and solder paste after soldering in vapors

CONCLUSION

Article provides theoretical information about the area of solder preforms and possible reasons for application in surface mount assembly. There was made test board with solder preforms in the experiment and was compared with solder joints manufactured by solder paste. Differences between joints were evaluated and significant facts were identified. It was confirmed better wettability of solder joint with solder preform and very good shape of this solder joint. Then there was favorable size and number of voids and it is prerequisite for high reliability of solder joint. Main advantage of solder preforms is precise volume of solder in solder joint and it was verified on test boards because shapes of solder joints were very accurate and this is good for high level of repeatability of manufacture. Finally, it is possible to say that solder preforms have great importance and firm place in SMT assembly because solder joints with preforms had very good properties compared with solder joints achieved by solder paste. Interesting results were also obtained with solder fortification method. Solder joints with solder preforms will be further investigated in thermal cycling chamber.

ACKNOWLEDGEMENT

Funding for this research was obtained through grant project FEKT-S-14-2168 “Research for advanced and innovative technologies for interconnection and packaging in microelectronics”.

REFERENCES

[1] SOCHA, P. 2009. *Solder preforms basics* [online]. SMT ONLINE 2009 (Indium Corporation), 4p [cit. 2016-03-12]. Available from: http://www.indium.com

[2] ŠANDERA, J. 2010. *Technology and reliability of solder connection electronic modules and compoments [i.e. components] in surface mount assembly: short version of habilitation thesis*. Brno: VUTIUM. ISBN 978-80-214-4221-4.

[3] ENGELMAIER, W. 1989. *Surface Mount Solder Joint Long-term Reliability: Design, Testing, Prediction*. In: . AT&T Bell Laboratories, Whippany, New Jersey, USA: MCB UP Ltd, s. 9. DOI: http://dx.doi.org/10.1108/eb037660. ISSN 0954-0911.

[4] ŠVECOVÁ, O. 2012. *Reliability of lead-free solders and the selected methods to estimate its lifetime*. Brno. Doctoral thesis. Brno University of Technology, Faculty of Electrical Engineering and Communication. Supervisor of doctoral thesis Doc. Ing. Josef Šandera, PhD.

[5] *Solder fortification: Product datasheet* [online]. 2014. Indium Corporation [cit. 2016-03-02]. Available from: http://www.indium.com/

[6] *Indium8.9HF Pb-Free Solder Paste: Product datasheet* [online]. 2014. Indium Corporation [cit. 2016-03-02]. Available from: http://www.indium.com/

IMPACT OF THE SIN_X THICKNESS ON PASSIVATION QUALITY AND CONTACT RESISTIVITY OF SILICON SOLAR CELL

Barbora Mojrová

Doctoral Degree Programme (3), FEEC BUT

E-mail: xmojro00@stud.feec.vutbr.cz

Supervised by: Jaroslav Boušek

E-mail: bousek@feec.vutbr.cz

Abstract: In this work the influence of thickness of Silicon Nitride (SiN_X) layer deposited by Plasma Enhanced Chemical Vapour Deposition (PECVD) on passivation quality and contact resistivity (ρ_C) of *n*-type Passivated Emitter Rear Totally-diffused (*n*-PERT) cell was investigated. The solar cell structure comprises front boron emitter and a phosphorous back surface field (BSF) with SiN_X layers on both sides for surface passivation. Contacts are made by screen printed and fired through metallization using commercial silver (Ag) paste.

Keywords: silicon solar cell, *n*-type, *n*-PERT, SiN_X, passivation, contact resistivity.

1. INTRODUCTION

The solar cell based on *n*-type silicon, as one of the promising alternatives to conventional solar cells, has attracted extensive research attention due to its several advantages. The *n*-type material has relative tolerance to higher minority carrier diffusion lengths compared to *p*-type monocrystalline silicon (mono-Si) substrates with a similar impurity concentration [1]. The minority carrier life time does not suffer from light induced degradation (LID) due to the boron-oxygen related defects which is frequently found in *p*-type material [2]. With conversion efficiencies above 23 %, the potential of *n*-type silicon has been demonstrated at the device level in recent years [3-5].

The solar cell efficiency is directly related to short-circuit current and contact resistivity, which in turn strongly depends on thickness of antireflection (ARC) and passivation layer [5, 6]. It was proved, that SiN_X layer deposited by PECVD provide excellent surface and bulk passivation and their deposition processes can be executed with a high throughput as required by the industry [7].

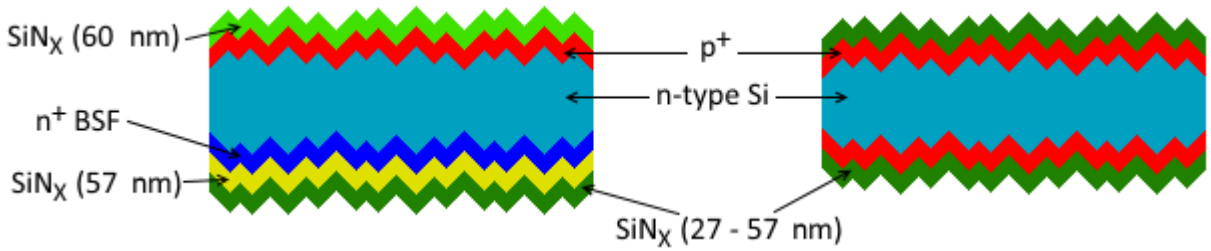
2. EXPERIMENTAL

2.1. SAMPLES PREPARATION

For this experiment the 6'' monocrystalline phosphorous-doped *n*-type mono-Si wafers with average base resistivity ρ_{base} = 2,3 Ωcm were used as a base material. All samples were processed using standard industrially available techniques for mass production such wet chemical alkaline texturization, cleaning by HCl and HF solutions, diffusion in quartz tube furnace containing $POCl_3$ (n^+ BSF, resulting sheet resistance 60 – 65 Ω/sq) or BBr_3 (p^+ emitter, resulting sheet resistance 70 – 80 Ω/sq). The p^+ and n^+ regions were subsequently passivated by SiN_X layer deposited by PECVD for the each side separately.

A schematic cross-section of solar cell precursor (solar cell without metallization) is presented in the Fig. 1 a). The deposition of passivating layer on the rear side of solar cell consisted of two steps. Firstly, the layer with thickness 57 nm was deposited; the SiN_X layer with various thicknesses was deposited subsequently. During the second deposition samples used for

measurement of SiN_X thickness (see Fig. 1 b) were also prepared. The annealing at peak temperature υ_{peak} = 815 °C followed after deposition of all SiN_X layers.



a) b)

Figure 1: a) the structure of cell precursor (solar cell without metallization) and b) sample used for measurement of SiN_X thickness.

In a next step the silver finger grid at the front side and rear side were screen printed. The cell process was finished by co-firing of the metal contacts in a belt furnace in air ambient. The temperature profiles differed from each other in the maximum temperature (υ_{firing}), which was set to the 780 °C, 795 °C, and 815°C.

3. RESULTS AND DISCUSSION

3.1. THICKNESS OF SiN_X LAYER

Method based on measurement of a reflectance spectrum of a wafer was used for the determination of thickness of SiN_X layer (d_{SiNX}) deposited on textured surface. The detector was inclined by 8° perpendicular to the sample. The dependence of reflection on wavelength of incident light for samples A, B, and C is shown in the Fig. 2.

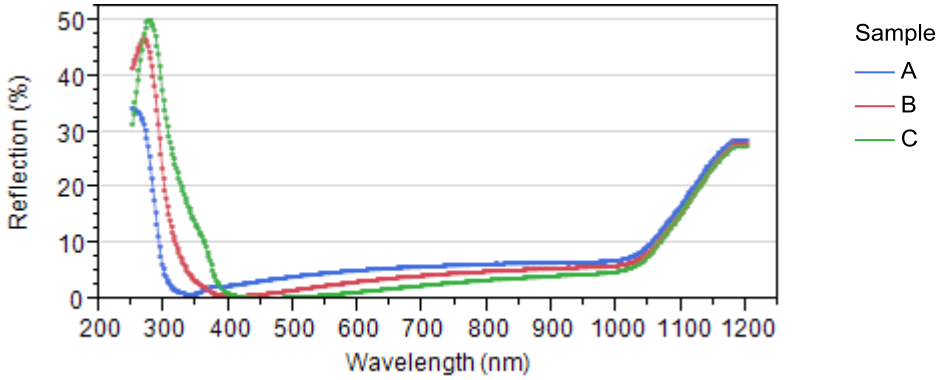


Figure 2: The reflectance of samples A, B, and C with different SiN_X thickness.

The thickness was calculated from the wavelength λ_{min} at the minimum of the reflectance curve according to:

$$d_{SiN_X} = \frac{\lambda_{min}}{4n_1}\sqrt{1 + \frac{{n_0}^2}{{n_1}^2}[\sin(90 - \theta)]^2} \quad (1)$$

where n_0 was the refractive index of air ($n_0 = 1$), n_1 was the refractive index of SiN_X ($n_1 = 2{,}03$), and θ was the angle of incident beam ($\theta = 8°$) [6]. The resulting values of SiN_X thickness are given in the Table 1. In case of samples X and Y the SiN_X thickness was evaluated by extrapolation based on deposition time (see Fig. 3).

Table 1: Calculated of SiN_X thickness for samples.

Sample	$t_{deposition}$ [s]	λ_{min} [nm]	d_{SiNX} [nm]
X	280	_	27,2
Y	380	_	34,8
A	480	341	42,1
B	580	406	50,1
C	680	462	57,0

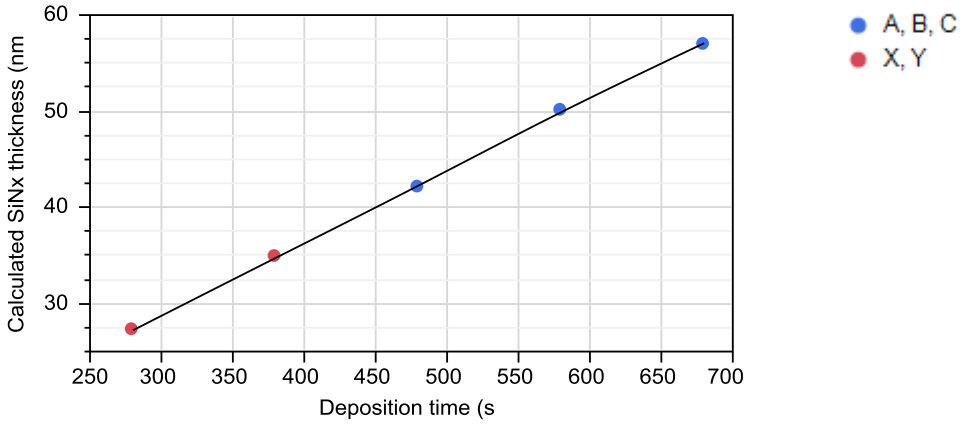


Figure 3: The dependence of SiN_X thickness on deposition time.

3.2. IMPLIED V_{OC} AND J_0

The passivation quality could be determined by the values of implied open-circuit voltage at one sun (iV_{OC}; the V_{OC} value determined from the carrier concentration) and dark saturation current density (J_0), measured by the Quasi-Steady-State Photoconductance technique (QSSPC) [8]. The iV_{OC} and J_0 values were measured after annealing, but before printing the metallization grid. The measurements were done on front side of each sample in five points. In the Fig. 4 there is shown dependence of implied V_{OC} and J_0 on the total thickness of the SiN_X layer on the rear side.

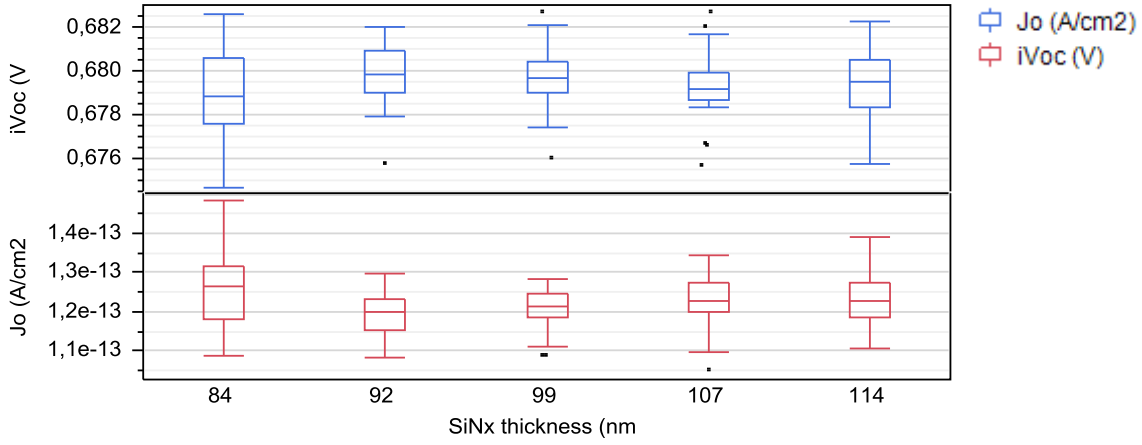


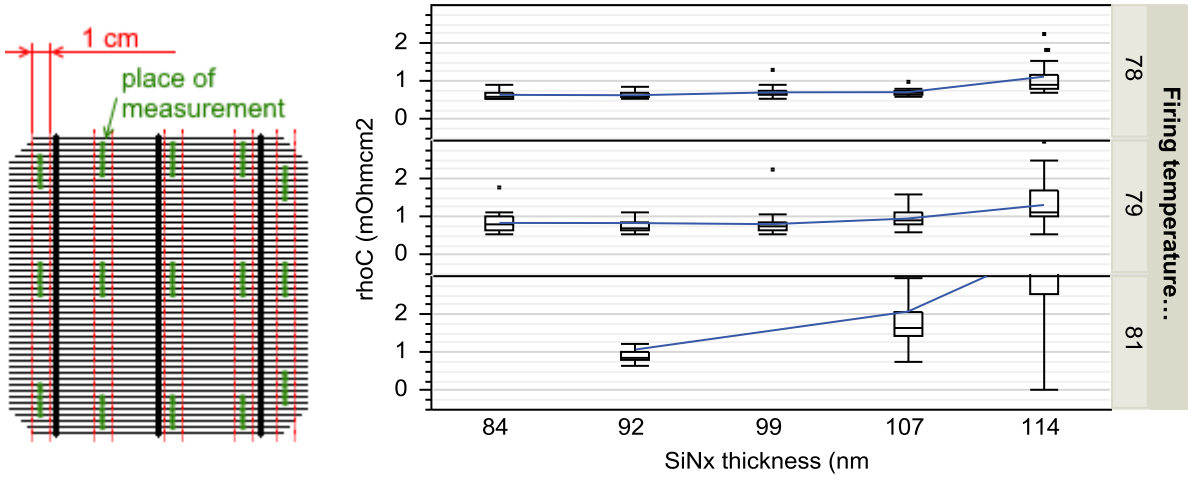
Figure 4: iV_{OC} at 1 sun and J_0 of the cell precursors (before metallization). The J_0 was extracted at an injection level of $\Delta n = 2 \times 10^{15}$ cm^{-3}.

It is obvious that the best passivation quality (high iV_{OC} and low J_0) gives the SiN_X layer with average thickness 92 nm and 99 nm. The mean values of iV_{OC} and J_0 were 679,8 mV and 11,98 fA/cm^2 in case of SiN_X thickness 92 nm, or rather 679,7 mV and 12,10 fA/cm^2 for samples with SiN_X thickness 99 nm.

3.3. CONTACT RESISTIVITY

The contact resistivity (ρ_C; also called specific contact resistance) characterizes the electrical quality of an ohmic contact formed between a metallization grid and an underlying semiconductor. The low value of ρ_C for the metal contacts indicates very good firing-through process, which leads to lower ohmic power loss.

After printing and firing step the dependence of contact resistivity on SiN_X thickness and firing temperature was detected by the Transmission Line Model (TLM) method. Five stripes 1 cm wide were laser-cut out of samples, perpendicularly to fingers (see Fig. 5 a). The contact resistivity of rear side metallization was measured on 3 different positions per stripe; each position includes 6 contact fingers.



a) b)

Figure 5: The influence of SiN_X thickness and firing temperature on resulting contact resistivity.

The dependence of contact resistivity on total SiN_X thickness on rear side of solar cell and on peak firing temperature is shown in the Fig. 5 b). In spite of the fact that the thicker SiN_X layer leads to slightly higher value of ρ_C, the temperature υ_{firing} has a major effect on the resulting contact resistivity. The lower values of ρ_C indicate that the larger Ag crystallites were formed during the firing process. The big variation of ρ_C of overfired cells, i.e. υ_{firing} = 815 °C and cells with d_{SiNx} = 107 nm and 114 nm, can be caused by the high inhomogeneity of junction depth and depletion width.

The lowest contact resistivity (below 1 mΩcm^2) was achieved for combination of low peak firing temperature (780 °C and 795 °C) and SiN_X thickness in the range of 84 nm and 99 nm (or 107 nm for υ_{firing} = 780 °C). The lowest contact resistivity (ρ_C = 0,65 mΩcm^2) was reached for combination of the thickness 84 nm and temperature υ_{firing} = 780 °C.

4. CONCLUSION

The QSSPC measurements for different thickness of SiNX layer showed, that the best passivation quality (high iV_{OC}, and low J_0) gives the SiN_X layer with average thickness 92 nm (679,8 mV; 11,98 fA/cm^2) and 99 µm (679,7 mV; 12,10 fA/cm^2).

The value of contact resistivity depends mainly on the firing profile (in this case υ_{firing}) and much less on the SiN_X thickness. The lowest contact resistivity (ρ_C = 0,65 mΩcm^2) was achieved for combination of low firing temperature (υ_{firing} = 780 °C) and total SiN_X thickness 84 nm.

ACKNOWLEDGEMENT

This work was supported by project no. FEKT-S-14-2300 A new types of electronic circuits and sensors for specific applications.

REFERENCES

[1] MACDONALD, D. and L. J. GEERLIGS. Recombination activity of interstitial iron and other transition metal point defects in p- and n-type crystalline silicon. *Applied Physics Letters* [online]. Melville: American Institute of Physics, 2004, **85**(18) [cit. 2016-02-22]. ISSN 0003-6951.

[2] GLUNZ, S. W., et al. Minority carrier lifetime degradation in boron-doped Czochralski silicon. *Journal of Applied Physics* [online]. United States of America: American Institute of Physics, 2001, **90**(5) [cit. 2016-02-22]. ISSN 0021-8979. Available at: http://scitation.aip.org/content/aip/journal/jap/90/5/10.1063/1.1389076

[3] BENICK, J., et al. High efficiency n-type Si solar cells on Al2O3-passivated boron emitters. *Applied Physics Letters* [online]. 2008, **92**(25) [cit. 2016-02-22]. ISSN 00036951. Available at: http://scitation.aip.org/content/aip/journal/apl/92/25/10.1063/1.2945287

[4] TAGUCHI, M., et al. High-Efficiency HIT Solar Cell on Thin (<100 μm) Silicon Wafer. In: *24th European Photovoltaic Solar Energy Conference and Exhibition Conference 21-25 September 2009*. Munich, Germany: WIP-Renewable Energies, 2009, pp. 1690 - 1693. ISBN 3936338256. Available at: http://www.eupvsec-proceedings.com/proceedings?paper=4112

[5] MIHAILETCHI, V. D., et al. Screen Printed n-Type Silicon Solar Cells for Industrial Application. In: *5th World Conference on Photovoltaic Energy Conversion proceedings*. Munich, Germany: WIP-Renewable Energies, 2010, pp. 1446 - 1448. ISBN 3936338264. Available at: http://www.eupvsec-proceedings.com/proceedings?paper=6722

[6] KRIEG, Alexander, et al. Anti-Reflection-Coating Thickness Measurements on Textured Silicon Surfaces: Evaluation and Accuracy of Different Measurement Techniques. In: *28th EU PVSEC*. München: WIP, 2013, pp. 1820 - 1824. ISBN 3936338337.

[7] DUERINCKX, F. and J. SZLUFCIK. Defect passivation of industrial multicrystalline solar cells based on PECVD silicon nitride. *Solar Energy Materials and Solar Cells* [online]. 2002, **72**(1-4), 231 - 246 [cit. 2016-02-26]. ISSN 09270248. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0927024801001702

[8] SINTON, R.A., et al. Quasi-steady-state photoconductance: A new method for solar cell material and device characterization. In: *Conference Record of the Twenty Fifth IEEE Photovoltaic Specialists Conference - 1996*. Washington DC: IEEE, 1996, pp. 457 - 460. ISBN
0-7803-3166-4.

ULTRA LOW POWER TUNABLE TRANSCONDUCTOR

Salma Bay Abo Dabbous

Doctoral Degree Programme (4), FEEC BUT

E-mail: xbayab00@stud.feec.vutbr.cz

Supervised by: Fabian Khateb

E-mail: khateb@feec.vutbr.cz

Abstract: This paper presents ultra-Low power (ultra-LP) Low voltage (LV) tunable transconductor (Gm) and its application to implement Gm-C filter. The CMOS structure of the Gm is performed using bulk-driven (BD) MOST technique, thus it can operate with extremely low voltage supply of ±0.3 V using 0.18 µm CMOS process. Moreover, it consumes ultra-LP about 4.9 µW. The simple topology, proper operating range, and tunability make this transconductor attractive. The transconductor and the Gm-C filter have been examined using simulation program Pspice.

Keywords: Bulk Driven MOST, low power low voltage, transconductor.

1. INTRODUCTION

Advance in electronics enables designing portable devices and biomedical systems which can be implanted within the body or can be worn by people. These electronic systems may achieve several tasks such as; monitor biological signals and physiological parameters and even may offer some kind of treatment for the patient. Long life, small size and light weight battery is essential need in such devices. Therefore the low-voltage (LV) low-power (LP) systems are highly required. However, analog circuit designers encounter many difficulties to preserve high performance operation with LV LP conditions. Thus, numerous MOS techniques have emerged for designing LV LP analog circuits. Among the most attractive techniques are the unconventional ones i.e. Bulk driven (BD), floating gate (FG), quasi floating gate(QFG), Bulk driven-floating gate (BD-FG) and bulk driven-quasi floating gate (BD-QFG) MOS transistor [1, 2].

The bio-potential signals are characterized by low frequency and low amplitude attributes [3]. Hence, ultra-low power, low frequency filters are crucial in wearable and implantable biomedical systems for enhancing signal quality before processing. Figure 1 shows the approximate operating frequency range of several filters types [4]. It can be observed that switched capacitor, discrete analog active RC, passive LC, and Gm-C filters can be used over the frequencies of the physiological signals. However, switched capacitor based filters are not desired in ultra-LP applications, since it consumes quite amount of power during charging/discharging the capacitors. On the other hand, RC active filters may consume large chip area for resistors or capacitors [5]. In passive LC filter case, a high inductance value is requiered for low frequencies operation, which is not suitable for on-chip integration. However, transconductor capacitor Gm-C filters process the signal contiuously, thus by proper transconductor design the total power consumption will be lower than the switched capacitor topology. Moreover, the Gm-C filters are mainly constructed from transconductors and capacitors, thus they are suitable for system on chip (SoC) integration using proper capacitor values.

This work presents ultra-LP LV CMOS structure of a tunable transconductor. Bulk driven MOS transistor technique has been used for proper LV LP operation. Eventually, second order universal Gm-C filter is introduced as an example of application.

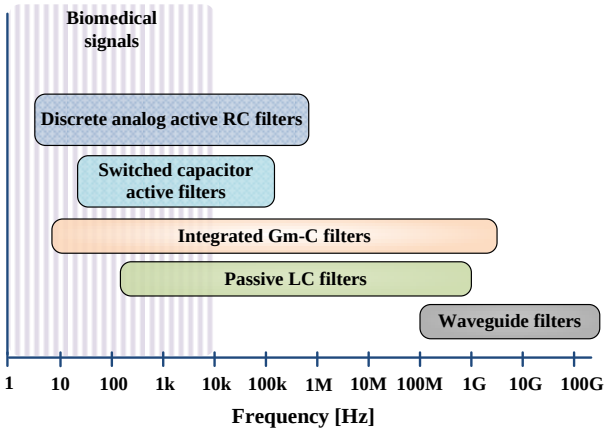


Figure 1: Approximate operating frequency ranges of several filters.

This paper is organized as follows: Sect. 2 presents the internal structure of the ultra-LP LV transconductor based on the BD technique. Universal voltage mode second order Gm-C filter is introduced in Sect. 3. Sect. 4 shows the simulation results, and Sect. 5 is the conclusion.

2. ULTRA-LP LV TUNABLE OTA AND ITS APPLICATION TO IMPLEMENT GM-C FILTER.

Figure 2 shows the symbol of the tunable transconductor. It has two high impedance input terminals (In-, In+), one high impedance output terminals (Out), the external resistor R_{set} is connected to the terminal (set) for tunning the transconductance value externally.

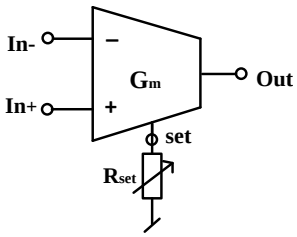


Figure 2: Transconductor schematic symbol.

The internal CMOS structure of the transconductor is shown in Figure 3. A BD differential pairs are implemented via transistors M_1, M_2, M_3 and M_4. The gates of these transistors are tied to the negative supply voltage V_{SS}, and the bulk terminals are used as input terminals. Transistors M_5 and M_6 are the tail current source of the first and second differential input stages, respectively. The current source I_{bias} and the multiple output current mirrors created by transistors M_7, M_8, M_9, M_{10}, M_{11} provide current biasing of the circuit branches. The transistors M_{12}, M_{12c}, M_{13}, M_{13c}, M_{10c} and M_{11c} ensure equal currents flowing through (set) and (Out) terminals. R_c and C_c compose the compensation network. The internal negative feedback enhance the linearity of the transconductance. It can be observed that the cascade structure has been used to build the output stage. Thanks to this structure the output terminal (Out) possess extremely high impedance.

The implementation of the BD MOST enables ultra-LV operation capability. Besides, the differential pair arrangement has very small supply voltage requirement, which can be described by [6]:

$$V_{DD,\min} = V_{GS,M_{6,5}} + V_{DSsat,M_9} \tag{1}$$

Assume that the negative feedback is disconnected and the input terminal of M_1 is grounded, then the self transconductance $g_{m,core}$ can be described by:

$$g_{m,core} = \frac{I_{Out}}{V_{In+} - V_{In-}} \tag{2}$$

By reconnecting the negative feedback and the external resistor R_{set} as shown in Figure 3, the straightforward analysis yields to:

$$V_{In+} - V_{In-} = \frac{I_{Out}}{g_{m,core}} + V_{R_{set}} \Rightarrow \; G_m = \frac{I_{Out}}{V_{In+} - V_{In-}} = \frac{1}{\frac{1}{g_{m,core}} + R_{set}} \approx \frac{1}{R_{set}} \quad \text{if} \quad \frac{1}{g_{m,core}} <<< R_{set} \tag{3}$$

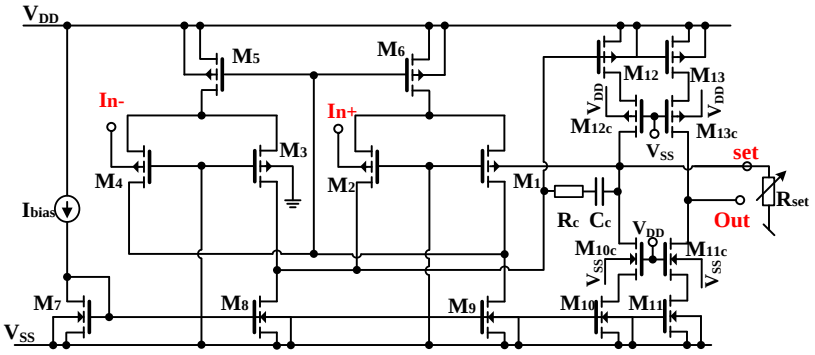


Figure 3: MOS internal structure of the tunable transconductor.

3. SECOND ORDER GM-C UNIVERSAL FILTER BASED ON BD TRANSCONDUCTOR

The second order universal transconductor-capacitor filter (Gm-C filter) is shown in Figure 4 [7]. This filter can provide low pass (LPF), high pass (HP), and band pass (BP) functions. The output voltage of the filter can be expressed by the following equation:

$$V_{out} = \frac{s^2 V_{in-HP} + s(G_{m2}/C_2)V_{in-BP} + ((G_{m1}G_{m2}/C_1C_2)V_{in-LPF}}{s^2 + s(G_{m2}/C_2) + (G_{m1}G_{m2}/C_1C_2)} \tag{4}$$

The pole frequency and the quality factor can be given by:

$$\omega_0 = \sqrt{G_{m1}G_{m2}/C_1C_2} \qquad (5), \qquad Q = \sqrt{G_{m1}C_2/G_{m2}C_1} \tag{6}$$

By applying the condition $G_{m1}=G_{m2}$, the pole frequency can be controlled independently from the quality factor.

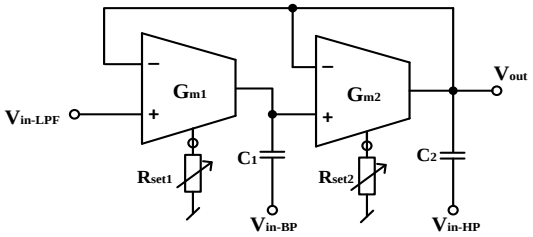


Figure 4: Second order Gm-C filter.

4. SIMULATION RESULTS:

To prove the high performance of the transconductor circuit, the circuit has been simulated using 0.18 µm CMOS process in Pspice environment. The supply voltage is V_{DD}= -V_{SS}= 300 mV, the bias current I_{bias}= 1µA and the total power consumption is 4.9 µW. the transistors aspect ratios W/L are: M_1, M_2, M_3, M_4= 10 µm/0.3 µm, M_5, M_6= 50 µm/0.3 µm, M_7, M_8, M_9= 4 µm /0.3 µm, M_{10}, M_{11}= 8 µm /0.3 µm, M_{12}, M_{13}= 100 µm /0.3 µm, M_{10c}, M_{11c}= 45 µm /2 µm, M_{12c}, M_{13c} = 100 µm /0.3 µm, C_c= 0.1 pF, and R_c= 1 kΩ.

In order to verify the transconductance G_m tunability, The DC curves of the output current I_{Out} versus voltage V_{In+} for $V_{In-} = 0$ are clarified in Figure 5, with stepping R_{set} from 100 kΩ to 1000 kΩ with 100 kΩ step. The obtained linear range is suitable for law amplitude signals such as biomedical ones. The transconductor values have been calculated using the formula (3) for different R_{set} values and are compared with the simulated values in Figure 6. It is notable that the simulated and calculated values are approximately identical.

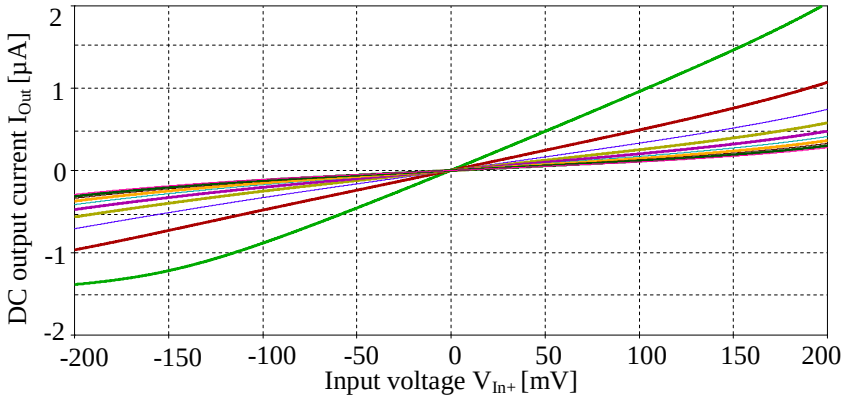


Figure 5: DC curves of the output current versus input voltage for different values of R_{set}.

The simulation results of the second order Gm-C filter are depicted in Figures 7 and 8. The total power dissipation of the filter is 9.8 µW. The frequency responses of the voltage gains of the filter shown in Figure 4 are presented in Figure 7 using C_1=C_2=50 pF with R_{set1}=R_{set2}=1 MΩ. It is obvious that this filter can provide low pass, high pass and band pass functions. The simulated natural frequency is about 3.1 kHz, this value is very close to the calculated value from (5). Monte Carlo simulation has been done to show the effect of threshold voltage, resistors, and capacitors mismatch in filter response. 20% variation of the threshold voltage with 10% tolerance of resistors and capacitors have been applied 100 times and the frequency responses of HP, BP and LPF filters are shown in Figure 8. It is clear that the curves have finite deviation and the effect of the variation of the threshold voltage, resistors and capacitors is limited.

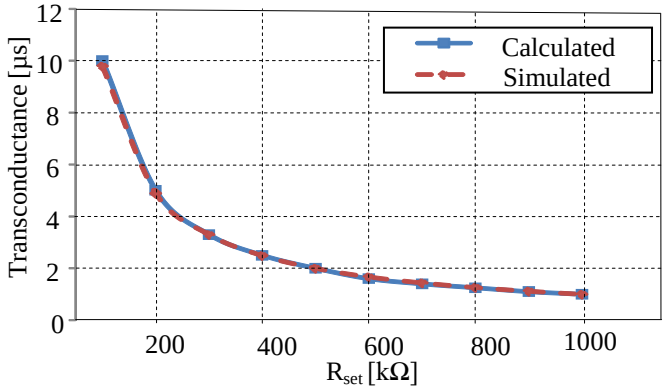


Figure 6: Calculated and simulated values of the transconductance versus R_{set}.

5. CONCLUSION:

This work presents CMOS structure of ultra-LP LV tunable BD transconductor. Utilizing the BD techniques offers high performance operation under low power supply voltage of ±0.3 V. The power consumption of the transconductor is 4.9 µW. Moreover, the circuit has simple structure, extended linear operation range and tunable transconductance. To verify the functionality of the circuit, a multi-function second order Gm-C filter is introduced as an example of application. Thanks to the tunable transconductance the quality factor and the pole frequency of the proposed filter can be adjusted.

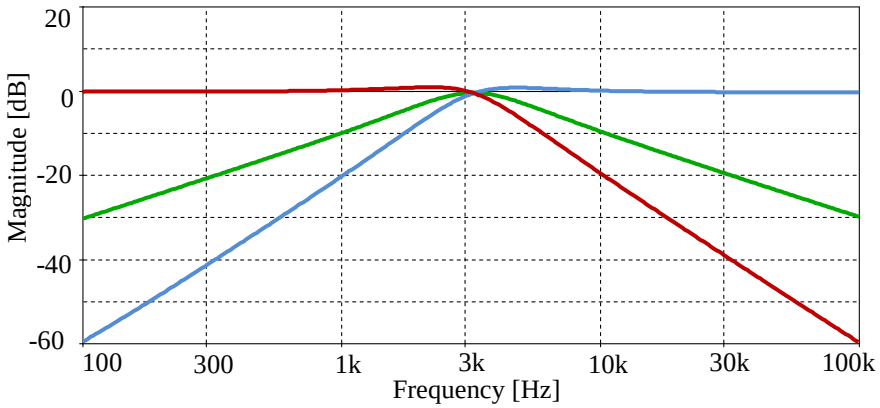


Figure 7: The frequency responses of the second order universal Gm-C filter.

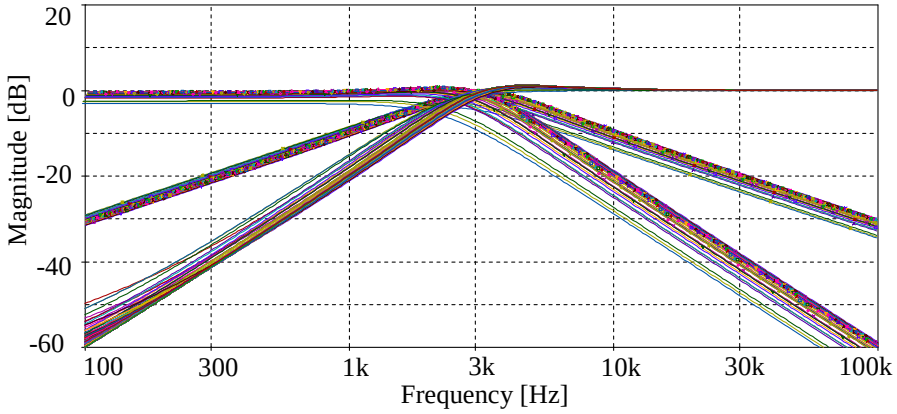


Figure 8: Monte Carlo frequency responses of the universal Gm-C filter.

REFERENCES

[1] Khateb, F., Bay Abo Dabbous, S., and Vlassis, S.: A survey of non-conventional techniques for low-voltage low-power analog circuit design, Radioengineering, pp. 415-427, 2013.

[2] Khateb, F.: Bulk-driven floating-gate and bulk-driven quasi-floating-gate techniques for low-voltage low- power analog circuits design, AEU Electronics and Communications J., vol. 68(1), pp. 64-72, 2014.

[3] Harrison, R. R.: A Versatile Integrated Circuit for the Acquisition of Biopotentials, Proceedings of IEEE Custom Integrated Circuits Conference, pp. 115-122, 2007

[4] Schaumann, R., and Van Valkenburg, Mac E.: Design of Analog Filter. New York: Oxford University Press, Inc, 2001.

[5] Lo, T. Y., and Hung, C. C.: 1V CMOS Gm-C Filters: Design and application, Springer, 2009, ISBN: 978-90-481-2410-7.

[6] Carvajal, R. G., Ramirez-Angulo, J., Lopez-Martin, A. J., Torralba, A., Galan, J. A. G., Carlosena, A., and Chavero, F. M.: The Flipped Voltage Follower: A useful cell for low-voltage low-power circuit design, *IEEE Trans. Circuits and Systems-I: Regular Papers*, vol. 52, no. 7, pp. 1276 - 1291, 2005.

[7] Khateb, F., Khatib, N., Prommee, P., Jaikla, W., and Fujcik, L.: Ultra-low voltage tunable transconductor based on bulk-driven quasi-floating- gate technique, Circuits Systems and Computers Journal, vol. 22(8), 2013.

WIRELESS COMMUNICATION PLATFORM IQRF – A CASE STUDY

Zdeňka Kuchtová

Doctoral Degree Programme (2nd), FEEC BUT

E-mail: klimesova@feec.vutbr.cz

Supervised by: Jaroslav Kadlec

E-mail: kadlecja@feec.vutbr.cz

Abstract: This paper describes wireless communication platform IQRF, history of development, available modules and accessories for development and for production use. Very briefly are described communication gateways and cloud services. Part of the article describes DPA communication framework from Microrisc company and summarizes available communication frameworks for different platforms.

Keywords: IQRF, DPA, IoT, Communication framework

1. INTRODUCTION

Terms like Internet of Things, Smart City, Smart buildings are quite often used in today world. They have at least one common fact, they need communication inside solution and with outside world. For this communication needs wireless or wired solutions are available. Each of them has some advantages and some disadvantages [1, 2, 3]. Within this paper, we will talk about wireless solutions only.

There is a wide range of the wireless solutions on the market. Some of them are standardized some of them are proprietary, and of course, some of them are trying to be a standard. From the standardized solutions, we could select WiFi, BlueTooth, RFID or ZigBee. From proprietary world MiWi, Z-Wave, IQRF and the others. Like always, all of them has some pros and some cons and are better or worse for different applications. In the rest of the article, we will stay with IQRF [4].

The article is organized as follows. The second section contains general information about IQRF wireless communication platform, followed by brief description of development tools and gateways to higher systems. The fifth section describes available communication frameworks for these systems. In the case study section is an example for all described parts. At the end of the article is summary and some future visions, acknowledgement and used references.

2. IQRF WIRELESS COMMUNICATION PLATFORM

IQRF is one of the proprietary wireless communication platforms that work on Sub-1 GHz base. IQRF wireless communication platform is a complete solution for design and development of the wireless applications. It was developed by Czech company Microrisc in a common project with Brno University of Technology.

At the beginning of the development process, only board with a microcontroller, RF part, antenna and some power supply was developed. This first module was only hardware solution with no software support. For communication was used AM modulation. This module was used as a base part for next projects.

After some time, a new, smaller module was designed and developed. This module has similar size like a mini SIM card and uses the same connector. This second developed module version was

composed of a microcontroller, RF part, a connector for an external antenna and was equipped with some LEDs and power supply part. The main change and de facto start of the IQRF wireless communication platform was development of an own operating system for the module. The operating system has functions for transmitting and receiving RF packets. One of a very early version of OS supports star communication structure. It means, in the center is one module called coordinator, rest of the modules are called nodes. For communication with other systems, SPI bus is used. This was a major update because from this time, development of a whole solution was much easier and time to market was much shorter because the implementation of wireless communication was prepared.

In the next versions was updated especially OS. Support of basic MASH was added. In the earlier versions up to 4 hops were supported. The routing table had to be prepared manually by the programmer.

Current versions of IQRF modules exist with or without integrated antenna. They are equipped by LEDs, power supply control, thermometer and external memory. For communication with higher systems SPI, I2C, UART buses are available. For wireless communication FSK and GMSK modulation is used. Block diagram of most used IQRF wireless communication module is shown in figure 1.

The operating system was extended to support MASH network structure. Now the network is created automatically by OS functions. It supports up to 250 routing points in one network. One coordinator is able to control up to 65000 network nodes. The whole network is possible to extend with the node, which works on one network like node and in second as network coordinator. In all versions ad hoc communication is implemented.

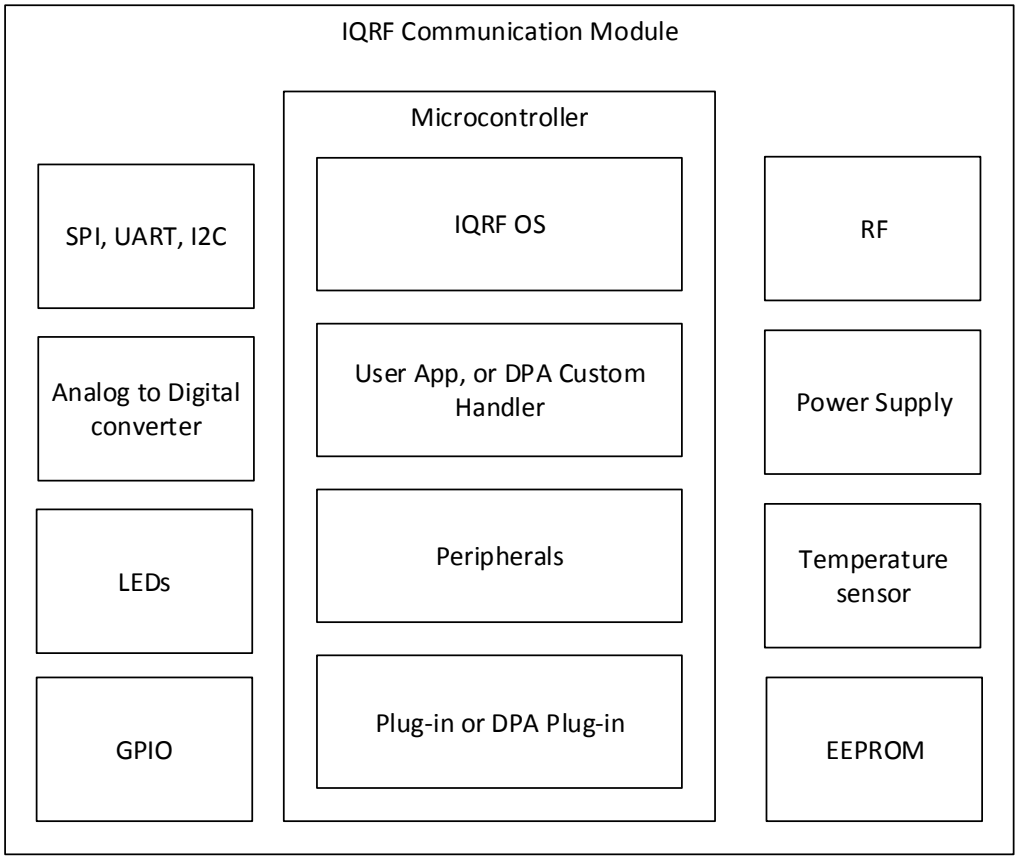


Figure 1: Main parts of IQRF wireless communication module.

Software for the module is a part of operating system. Developer prepares it in C language with BKND compiler. User part of program memory is limited and depends on module and

microcontroller type. One part of the memory is dedicated to plug-ins, which could extend OS functions.

Really interesting part are FRCs - Fast Response Commands. These commands allow reading two bits or one byte from filtered list of nodes by one command. Response time is minimized and depends on network structure.

The most useful OS extension is implementation of Direct Peripheral Addressing communication framework (DPA). With this framework, all available module peripherals can be accessed and controlled without the need to write any part of code on module side. Everything is done via a set of commands sent to coordinator via a local interface or in some special occasions directly via a local interface of the node. In the case of necessity, the developer can extend the functionality of DPA via DPA Custom Handler. In this situation framework provides events for each part of message processing. For example, event before a response is sent, or event after resending and so on. Also, some limited access to interrupt function is available.

This section very briefly described the history and main features of IQRF communication modules and its operating system. But the whole platform is composed of Ethernet gateways, USB gateways, IDE, Development boards and other support for the developer. Some of them are described in next sections.

3. IQRF DEVELOPMENT TOOLS

The main aim of IQRF platform is to minimize time to market. For that reason, development tools were prepared. The main part of development tools is IQRF IDE. This integrated development environment supports the whole development process within IQRF platform.

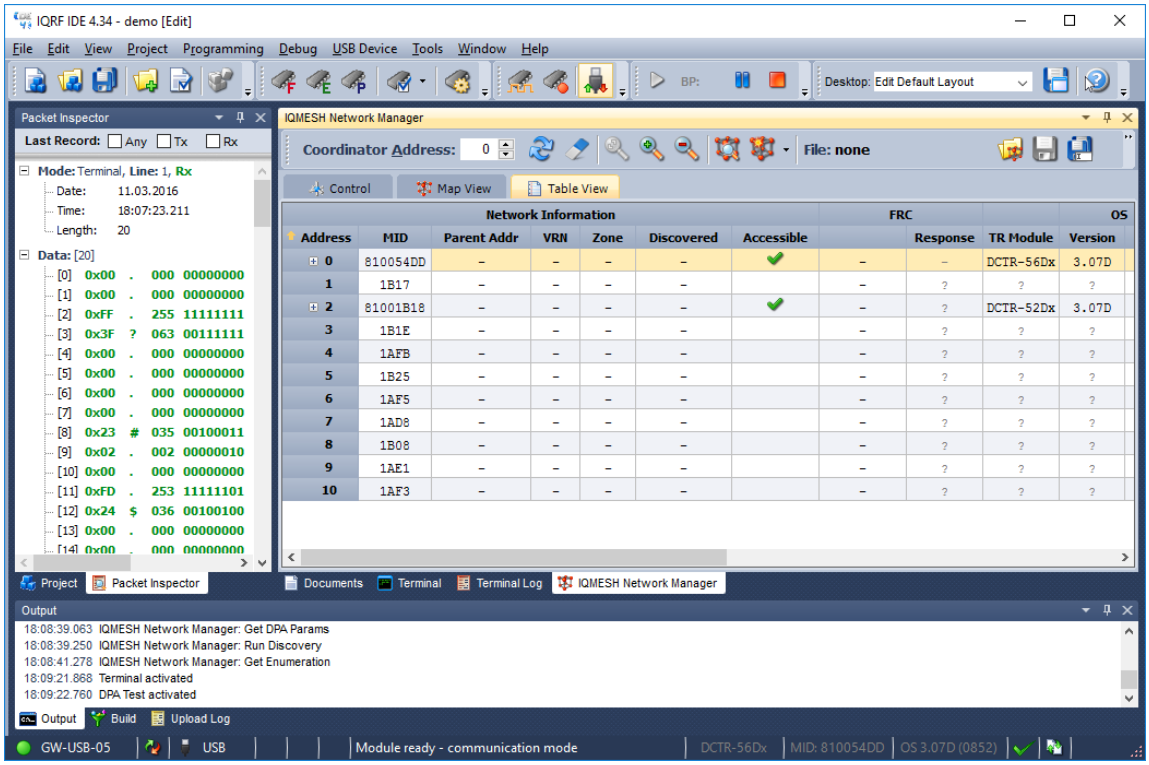


Figure 2: IQRF Integrated Development Environment.

The developer uses IQRF IDE for preparing software in C, compilation and uploading it to the module. There are two main ways how to upload prepared software into IQRF module. One is via programmer - device that could program, update and debug SW in the module. This device is

connected to PC via standard USB bus. During connection with IDE it is possible to set a breakpoint and read all parameters and memories of IQRF module.

The second way is to use wireless programming. In this case, one module connected to the programmer or USB gateway works as a programmer. SW is sent to one or more IQRF modules via IQRF network. In this mode, no debug features are supported.

Communication with module could be tested via integrated terminal. IQRF module connected in programmer works in release mode, sends and receives data on SPI bus. All data are shown in IDE. There is also support for macros definition and repeated request sending.

DPA framework is supported via IQMESH network manager, where is possible to bond modules to the coordinator, set all DPA parameters, analyse network and so on. One part of IQMESH network manager is dedicated to backup and restore the whole module. Network manager shows all information about module connected to the network, like network address, zone, MID, OS version, Hardware profile, all available peripherals and setting for each of them and many other. IDE also supports sending DPA command and analyse response. IQRF IDE window is shown in figure 2.

4. IQRF GATEWAYS AND CLOUDS

IQRF platform offers a wide range of communication gateways. For servers and personal computers are available USB gateways. This type of gateways supports two different communication protocols, CDC and WinUSB. CDC is using a serial port with proprietary commands which is easy to implement in most of the programming languages and in different operating systems. WinUSB is direct support for USB device. There are more features available like device detection etc.

For remote access to IQRF network Ethernet, WiFi and GSM gateways are available. All of them could work in two main modes. The first one is direct communication with user program or communication recorder. This mode is suitable especially for communication within the local network. The second mode is communication with IQRF cloud. Gateways do not need any specific setting for cloud connection, there is a default setting that will work without user attention. Cloud is provided directly from Microrisc company. Another possibility is to use private IQRF cloud installation. In this case, address and password for private cloud have to be set.

Visual gateways are devices with touch screen display, network connectivity, battery and of course IQRF module, that are designed especially for fast visual control panel design. They are based on Microchip visual control library, with prepared board support package.

The last type of gateways is third party devices with IQRF extensions, like Arduino, Gemalto, chipKIT, Raspberry Pi and many others.

5. IQRF DPA COMMUNICATION FRAMEWORKS

DPA is an extension of IQRF operating system. This is a part implemented to communication modules. Each module is composed of different peripherals, standard one, prepared by Microrisc, or users peripherals, prepared by the user. Control of this peripherals in DPA is done by a set of commands, specific for each peripheral, but the structure of command is standardized for whole DPA. Commands are sent in communication packet with defined protocol, wirelessly or directly to module local interface. This is module point of view. However, on the other side is some control device that needs to communicate with modules. This is exactly the place where DPA communication frameworks are used.

DPA communication frameworks prepare packets, send them and parse response for the higher user or, in this case, programmer comfort. Available are three DPA Communication Frameworks. One is prepared in Java and is part of IQRF SDK. This framework is suitable everywhere, where Java is available. It is able to communicate with IQRF gateways in CDC mode directly via SPI on

boards like Raspberry or Gemalto and similar ones. The framework is available with source code via GIT Hub.

Another one is designed for devices with low computing power, without an operating system. It is completely written in C language and available with source code. These twos were prepared by Microrisc company.

The third one is prepared for Microsoft .NET languages by 2KSys company [4]. This framework communicates with USB gateways via WinUSB drivers, or via CDC mode. It could work everywhere, where .NET framework is available. It is suitable for the design of Universal Applications via specific USB gateway, prepared for this kind of applications, or is able to directly communicate with IQRF gateway via the cloud. From the programmer point of view, just an instance of the different class is used. This framework will be used for demo software in the next section.

6. A CASE STUDY

In this use case, we will try to prepare remotely controlled relay and read data from the temperature sensor. Block scheme is shown in figure 3. We would not prepare any hardware, but we will use only devices from IQRF platform. As a communication module are used two pieces of IQRF DCTR module DCTR-72DAT with integrated antenna and temperature sensor. These modules are put into development kit DK-EVAL-04A which has integrated battery and battery charger. One module is extended with DDC-RE-01 relay kit, with two relays. Network coordinator is represented by GW-ETH-02A that has an internet connection.

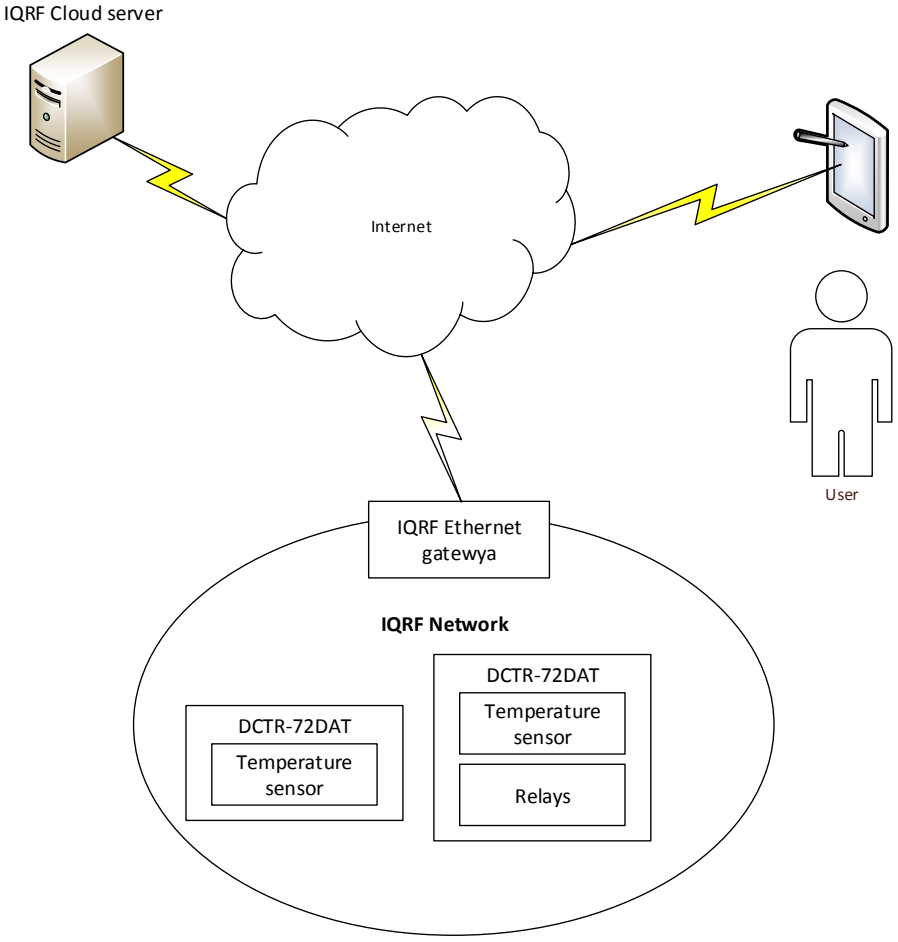


Figure 3: A case study communication chain

We will prepare the main communication part of the user application. This app will have two controls, one will read temperature and second will control relays. After user request, the application will prepare DPA packet and send it to IQRF cloud. The packet is prepared and send to the cloud by 2KSys.NET DPA communication framework. IQRF Ethernet gateway will read the packet from IQRF cloud and send it to network coordinator in the gateway. The packet is than processed by the coordinator or send to IQRF network to appropriate node. The node will handle this packet and send a response to the coordinator. The coordinator will use IQRF gateway to store response packet to IQRF cloud. Finally, user's application will take the response from the cloud and present results to the user.

At the first step, it is necessary to create a network, it means bond nodes to the coordinator. This could be done by IQRF IDE, which is able to connect via a local network to IQRF Ethernet gateway and start the bonding process. Ethernet gateway is connected to the IQRF cloud by default, so no more setting has to be done. Just short note, you have to create an account on Microrisc cloud server and assign a gateway to your account.

Now it is time to prepare software for user application. We will prepare a universal application for Microsoft.NET framework in C# programming language. At the beginning, we prepare an instance of communication framework. We would like to communicate with network coordinator via the cloud.

At first, we need to create cloud communication client. Cloud address is Microrisc cloud server, login and password are your credentials created during gateway registration process, the same is gateway MID and gateway password. We would like to use cloud just like interface to network coordinator. DPA framework needs interface IDpaCommunication for its communication. In the listing bellow is created an instance of this interface in iqrfCloudCommunication property.

With the communication interface we are ready to create network structure and especially network coordinator. In DPA framework is coordinator represented by class DpaCore. The constructor of this class requires one parameter - the communication interface. The instance is stored in property Core, and because quite often will be needed a peripheral coordinator, we also store direct access to this in CoordinatorPeripheral property. By asynchronous command UpdateDataFromCoordinatorAsync, we retrieve from coordinator information about connected nodes and create appropriate instances in DpaCore class. Last things to do is to create an instance of DPA node thermometer and IO peripherals. The whole process is described in DPA Framework online help in details.

```
public async void CreateNetwork()
{
    var client = new IqrfCloudClient
    {
        CloudAddress = new Uri("https://cloud.iqrf.org/api/api.php"),
        UserName = @"cloudusername",
        UserPassword = @"securepassword",
        GatewayMid = @"12345678",
        GatewayPassword = @"iqrf"
    };
    var iqrfCloudCommunication = new IqrfCloudCommunication(
         new IqrfCloudRepository(client));
    Core = new DpaCore(iqrfCloudCommunication);
    CoordinatorPeripheral = (DpaPeripheralCoordinator)Core.Coordinator
        .Peripherals.FindPeripheral(PeripheralTypes.Coordinator);
    await Core.UpdateDataFromCoordinatorAsync();
    Thermometer = (DpaPeripheralThermometer)DpaCore
        .CreatePeripheralInstance(PeripheralTypes.Io, Core.Nodes[1]);
    Io = (DpaPeripheralIo)DpaCore
        .CreatePeripheralInstance(PeripheralTypes.Thermometer, Core.Nodes[2]);
  }
```

Now we need to write some code to read the temperature from the sensor. This is quite easy you need only one line of code:

```
await Thermometer.ReadAsync();
```

A Little bit more complicated is it with relays. They are connected to port C pins 2 and 5. To switch relay 1 on, use something like this:

```
await Io.SetAsync(new Command[] {new Output {
        Port = Output.OutputPort.PortC,
        Mask = 0x04,
       Value = 0x04}});
```

For relay 2 is the same command, just use value 0x20 instead of 0x04. That's all, nothing more and nothing less is needed to create remotely controlled relay and thermometer with IQRF wireless communication platform and related communication frameworks. More examples and details are on [4] and [5].

7. SUMMARY AND FUTURE WORK

This paper introduces IQRF platform, designed and developed by Microrisc company and BUT. In the article, the main features of the platform were described, especially IQRF communication modules, development tools, gateways and cloud and DPA communication frameworks. In the case study, we showed temperature measurement application for PC with IQRF cloud, DPA framework and modules equipped with a temperature sensor.

There is always a place for improvements. IQRF OS is still updated, DPA framework has new functions and features, with the aim to help developers with the implementation of wireless communication. BUT is in direct contact with Microrisc developers and we work together on new projects related to this platform.

8. ACKNOWLEDGEMENT

Research described in this paper was financed by the National Sustainability Program under grant LO1401. For the research, infrastructure of the SIX Center was used.

9. REFERENCES

[1] O. Bello; S. Zeadally, "Intelligent Device-to-Device Communication in the Internet of Things," in *IEEE Systems Journal* , vol.PP, no.99, pp.1-11 doi: 10.1109/JSYST.2014.2298837.

[2] D. Guinard, V. Trifa, F. Mattern, E. Wilde, 'From the Internet of Things to the Web of Things: Resource Oriented Architecture and Best Practices'. In: Dieter Uckelmann, Mark Harrison, Florian Michahelles (Eds.): Architecting the Internet of Things. Springer, pp. 97–129, New York Dordrecht Heidelberg London, 2011.

[3] V. Sulc, R. Kuchta, and R. Vrba, "IQRF smart house—a case study," in Proceedings of the 3rd International Conference on Advances in Mesh Networks (MESH '10), pp. 103–108, July 2010.

[4] Dpa Framework. *2KSys s. r. o. Web* [online]. Brno: 2KSys s. r. o., 2015 [cit. 2016-03-12]. Available: http://www.2ksys.cz/Products/DpaCfDemoFree/cs.

[5] MICRORISC. IQRF - wireless technology. MICRORISC s. r. o., 2015 [cit. 2016-03-12]. Available: www.iqrf.org.

WIRELESS SENSOR TRIGERRING USING BLUETOOTH LOW ENERGY TECHNOLOGY

David Veverka

Doctoral Degree Programme (1), FEEC BUT

E-mail: xvever05@stud.feec.vutbr.cz

Supervised by: Jiří Háze

E-mail: haze@feec.vutbr.cz

Abstract: This work deals with design of stable and fast wireless connection between two endpoints. The Bluetooth technology or its modern alternatives is used to realize this connection because this technology has already been tested in the desired environment. The results of timing analysis of the selected sensors are discussed as well as communication between sensor and a chosen microcontroller and in the end a complete solution will be designed in order to create a functional prototype of the desired application.

Keywords: Pressure sensor, Bluetooth technology, Wireless communication, Gazell, RFduino, Flow sensor

1. INTRODUCTION

The aim of this project is realization of fast and stable wireless connection between a gas flow sensor and a receiver. The use of low cost modules and Bluetooth Low Energy technology was one of the requirements. The technology was chosen thanks to its compatibility with the work conditions in which the modules will be further used. A development of concept based on timing measurements of tested modules and workplace characteristics was the main goal. Several variants of gas flow sensors were evaluated and in the end number of prototypes were realized including a casing for the sensor with mounting mechanism compatible with the desired machine. The prototypes were tested in the chosen work conditions where their functionality and utilization were tested and proved.

2. GAS FLOW SENSORS

There are many types of these sensors on the market. Various sensing mechanisms are used i.e. a turbine rotation detection, pressure difference or calorimetric principle.

As the best solution, a sensor using calorimetric principle was used mainly because of its dimensions and characteristics. There is a small heated element which is cooled by the flow of gas inside of the sensor. Thus the measure of cooling is equal to the increase or decrease of the flow of gas.

The chosen sensor FS7002 is manufactured by company Siargo Ltd. and has following parameters:

- gas flow detection from 0 m/s to 10 m/s,
- supply voltage 5 V,
- output voltage range from 0 to 3,3 V which corresponds to the full scale range,
- time delay 20 ms [1].

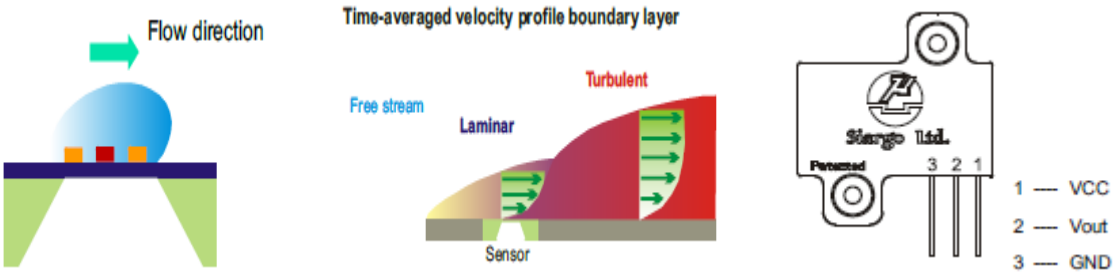


Figure 1: Sensor FS7002 and its functionality explanation [1]

A test unit mounted with a basic control circuit was realized to prove that the sensor housing is tight enough in order not to leak any gas from the valve. It was complicated to manage this tightness because of the characteristic of the used rare gas Argon but the leakage was reduced sufficiently to ensure reliable gas flow detection. The casing was made using a 3D printer and special resin. Figures 2 and 3 show the first and final prototype of the device.

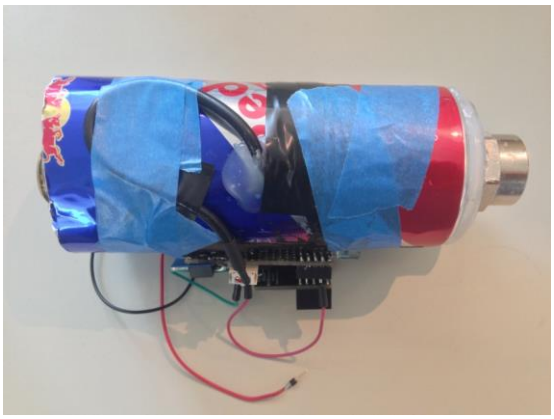


Figure 2: First prototype casing for FS7002 sensor with basic control electronics

The prototype on the figure 2 was further developed, the control electronics were redesigned and the casing was reduced in size. A small housing for the electronics was also designed. The final design can be seen in figure 3.

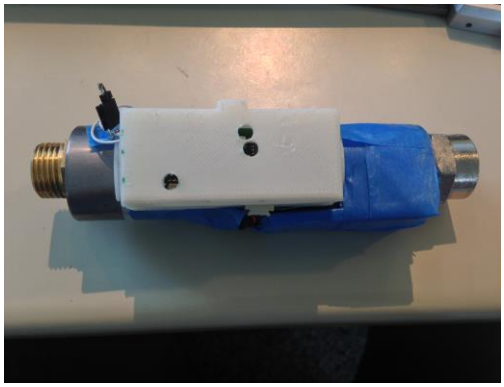


Figure 3: The final prototype of the sensor casing mounted with control electronics

3. COMPARISON AND OVERVIEW OF TESTED MODULES

Choosing of a proper microcontroller for the desired application was another important part. Main focus was placed on low power consumption, small footprint and sufficient amount of I/O ports. The stability of time delay during the streaming of data from sensor to the receiver was the most

important parameter. Several modules from different manufacturers has been tested. Comparison of module parameters is shown in Table 1. In the table 2 there can be seen the result of time parameters testing with the final prototypes in the desired work conditions.

	Delay [ms]	Stability	Power consumption, Vs = 3.3V	Dimensions	Price per piece	Complexity
RFDuino	2 to 7	Very high	18 mA peak, 4 µA in LP mode	(L)15mm x (W)15mm	17$	Low
Bluno	8 to 10	Very high	18.6 mA	Chip only – 6x6 mm	8$	Low
RF Bee	7	Average	34.5 mA peak	(L)33mm x (W)24.4mm	16$	Low
XBee	9	Average	50 mA peak	(L)27.6mmx (W)22mm	23$	Low
BLE Mini	6 to 100	Very low	18.6 mA peak	(L)39mm x (W)18.5mm	30$	Avg.

Table 1: Overview of parameters of tested modules

Rf duino timing measurements	Delay between gas flow and light generation by the appliance
Min. delay: 10.9 ms Max. delay: 16.3 ms	Min. delay: 220 ms, Max. delay 250 ms

Table 2: Overview of timing parameters from final testing

3.1. CHOSEN MICROCONTROLLER

Based on the parameters in table 1 a microcontroller RFDuino was chosen thanks to its simple programmability, sufficient amount of I/O ports, dimensions and preprogrammed functions which allowed for very easy realization of the desired function.



Figure 4: Chosen microcontroller RFDuino [2]

The chip on figure 4 has following parameters:

- supply voltage 3,3V,
- bandwidth 2,4 GHz,
- CPU 16 MHz ARM Cortex-M0,
- dimensions 15mm x 15mm x 3,5mm,
- pin count 19,
- transmit power 4 dBm [2].

4. CONTROL ELECTRONICS

Several variants of electronics were realized which were used to test the proper functionality of the designed circuit. Only the last variant will be shown in this paper which was used in the final prototype.

4.1. CIRCUIT CONTROLLING OF THE SENSOR FUNCTION

Optimal functioning of the sensor and the microcontroller required utilization of voltage regulators in the circuit design. A charging circuit for the Lithium Polymer battery was also added as well as a header allowing programming of the microcontroller on the final PCB, several LED diodes for signalization and circuit for detection of the battery voltage level. The functionality of the sensor side is mainly based on processing the data from the attached flow sensor. The data is collected on the I/O port with an AD converter. When the value of the detected input voltage reaches the defined threshold it inverts the value in a global variable which is then sent over the bluetooth link to the receiver. The code also controls the battery level every half an hour and the signalization showing that the link is established and if the data are being transmitted over the link or not. The schematic can be seen in figure 5.

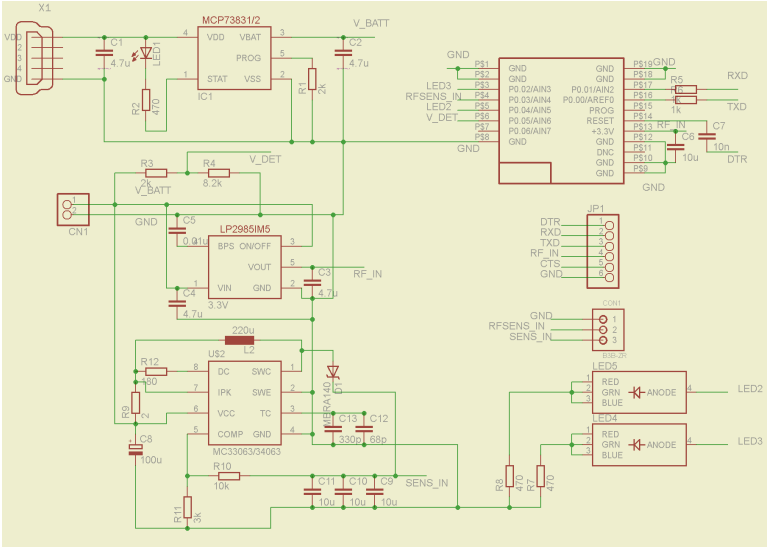


Figure 5: Schematic of control circuit for gas flow sensor

4.2. CIRCUIT CONTROLLING OF THE RECEIVER

The only difference between the previous circuit and this one is the omission of the sensor voltage regulator. One extra output port of the microcontroller was used for triggering the element which is connected to the receiver. The functionality includes the bluetooth connection establishment, pairing of the modules and simple data processing from the sensor which were used for deciding when to trigger the desired element. The battery level detection and LED signalization is the same as on the sensor side. The schematic is shown in the figure 6.

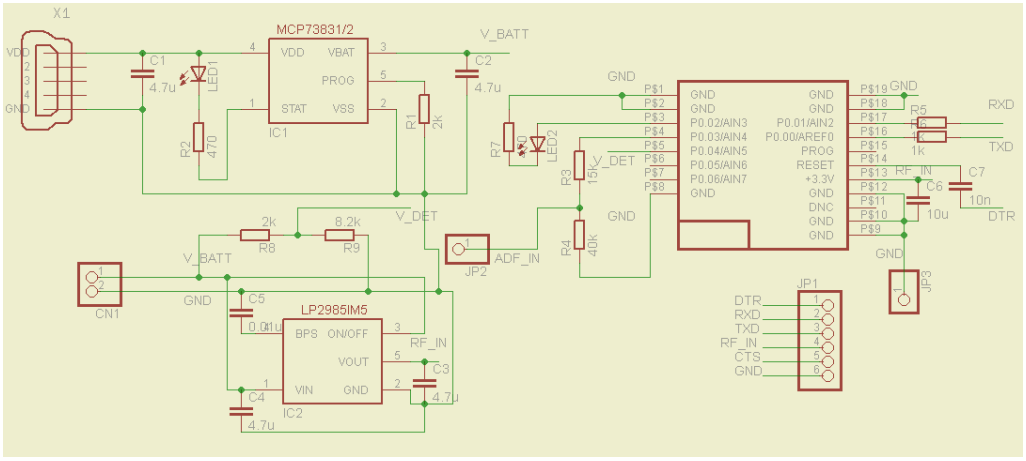


Figure 6: Schematic of the receiver

5. CONCLUSION

This work dealt with the design of a system for transmitting data from gas flow sensor to the receiver. The basic element was a gas flow sensor which measuring mechanism was based on the calorimetric principle. The solution was chosen mainly because of its dimensions because no other suitable sensor with sufficiently small footprint is fabricated. Further a casing for the sensor was designed and realized using 3D printing and special resin encasing. As the last part the control electronics for the sensor and the receiver along with PCBs were designed based upon the requirements of the final application. The final prototype was tested in the desired work conditions. The testing proved the functionality of the whole concept thanks to sufficient time frame for triggering the desired appliance. Thanks to this the concept is currently further developed.

REFERENCES

[1] Apollo Electronics. *www.apollounion.com* [online]. 2004-2013 [cit. 2014-05-01]. Available at: http://www.apollounion.com/en/down.asp?filename=Upload/DownFiles/FS7002-VB0.pdf

[2] RFduino: Rapid Development & Prototyping. *www.rfduino.com* [online]. 2013-2014 [cit. 2014-05-01]. Available at: http://www.rfduino.com/product/rfd22301-rfduino-ble-s

Doktorské projekty

Teoretická elektrotechnika, fyzika a matematika

A COMPARISON OF EDGE-BASED AND REGION-BASED SEGMENTATION PERFORMED WITH PDES

Jiří Sliž

Doctoral Degree Programme (1.), FEEC BUT

E-mail: jiri.sliz@vutbr.cz

Supervised by: Jan Mikulka

E-mail: mikulka@feec.vutbr.cz

Abstract: The paper discusses the level set segmentation method, which employs the solution of the partial differential equation (PDE) describing the multidimensional function whose zero slice in plane *xy* determines the contour of the object. One of the possible approaches to the contour evolution is using the image edges; the edges define the boundaries of the segmented objects. A major disadvantage of this procedure consists in the often unclear or noisy edges, a condition that may lead to incomplete segmentation. Such a drawback related to the curve evolution is solvable through utilizing regions with similar brightness levels; consequently, the actual segmentation cannot be affected by unclear transitions between the objects and the background. The final comparison then shows that both methods find specific application within the segmentation process.

Keywords: Level set, PDE, edge, region, segmentation

1. INTRODUCTION

In general terms, image segmentation consists in separating the objects from one another or from the background; the expected result is then embodied in a description of separated regions corresponding to the individual image objects. A wide range of segmentation methods are available, including, for example, thresholding, region clustering, or watershed; this paper nevertheless focuses on techniques that exploit partial differential equations, and the group of such instruments comprises, among others, the active contours and level set approaches.

Active contours utilize the evolution of an explicitly set parametric curve, an evolution influenced by image forces directing the curve to the object edges. Here, a problem occurs when the topology of the segmented objects is changed, namely, when curve splitting or merging is to be solved. For this reason, it then appears more advantageous to use the level set method, which utilizes the implicit definition of the curve as a zero slice in the *xy* plane, performed via a multidimensional function. This, or the level set, function is then characterized within the present article, where two relevant evolution options are compared; the options consist in the edge- and region-based methods.

2. LEVEL SET SEGMENTATION

The level set technique was proposed by S. Osher a J. A. Sethian in their late 1980s study [1]. The principle of the method relies on forming a level set function, $\phi(x, y, t)$, which has one dimension more than the parametric curve $C(x, y)$. This curve is a zero slice in the *xy* plane of the level set function (zero level set); thus, we have $C(x, y)=\{(x,y)\in \mathbb{R}^2: \phi(x, y, t)=0\}$. By forming the level-set function via utilizing the image energies, we can separate the objects. The formula describing the evolution of the level set function ϕ is comprised in Eq. (1) below:

$$\frac{\partial \phi}{\partial t} = -\nabla \phi \cdot \boldsymbol{\upsilon}\,, \tag{1}$$

where $\boldsymbol{\upsilon}$ expresses the evolution speed of the function ϕ and t is the time. The evolution speed is usually expressed by the force acting in the direction of the normal vector $\boldsymbol{n} = \frac{\nabla \phi}{|\nabla \phi|}$. The discussed equation (1) can then be modified as follows:

$$\frac{\partial \phi}{\partial t} = -\nabla \phi \cdot \upsilon = -\nabla \phi \cdot F \cdot \boldsymbol{n} = -\nabla \phi \cdot F \cdot \frac{\nabla \phi}{|\nabla \phi|} = F \cdot |\nabla \phi|\,. \tag{2}$$

Such a modified form (2) is commonly solved by means of the finite difference method.

2.1. EDGE BASED SEGMENTATION

Caselles et al. [3] propose a method for evolving the level set function via the image gradient (edges); this technique utilizes the edge indication function *g*, which assigns small values to a large gradient in a smoothed image. We then have

$$g = \frac{1}{1 + |\nabla I_\sigma|^2}\,, \tag{3}$$

where I_σ is an image treated with a Gaussian filter. Using the function *g*, the level set function can be written as follows:

$$\frac{\partial \phi}{\partial t} = g|\nabla \phi| div\left(\frac{\nabla \phi}{|\nabla \phi|}\right) + \nabla g \nabla \phi\,, \tag{4}$$

where the expression $\nabla g \nabla \phi$ ensures the direction of the contour *C* (zero level-set) towards the image edges. At this point, let us note that a number of other similar methods with different parameters have been published, but all of these employ edge-based representation of the image gradient ∇I.

2.2. REGION BASED SEGMENTATION

The above-discussed technique to develop the level-set function by means of the image gradient (edges) is satisfactory when the boundaries of the image are clearly defined. In reality, however, a substantial amount of images exhibit fluent transition of brightness levels between the objects in image, and the use of edges for segmentation is thus completely ineffective.

The problem of blurred edge boundaries can be eliminated through the region based approach, a solution proposed by Chan and Vese [4]. In the referenced paper, the authors employ the Mumford-Shah segmentation technique [5], where the initial phase consists in obtaining a model of homogeneous areas in the given image; such a model facilitates the determination of region boundaries even with no sharp edges in the original image. Let us define a curve *C* in the area Ω, which separates the regions Ω_1 and Ω_2. We have that C is a zero slice in the plane *xy* of level set functions ϕ; assume that Ω_1 is an object and Ω_2 the background. Each region is approximately homogeneous and has a certain value u^1 and u^2 for Ω_1 and Ω_2. To determine the relationship to a given area of the function ϕ, the Heaviside function H_ϕ: is used; we then have

$$H_\phi = \begin{cases} 1, & if\ \phi > 0, (\mathrm{x} \in \Omega_1) \\ 0, & if\ \phi \le 0, (\mathrm{x} \in \Omega_2). \end{cases} \tag{5}$$

As the presented function is not derivable due to discontinuity, H_ϕ is, in the region close to zero, replaced with a substituent such as a sigmoid function. Consequently, the derivation H_ϕ is zero

everywhere except at the boundary between the regions Ω_1 and Ω_2. Now we can define the energy functional, whose minimization enables us to derive the development of the level set function.

$$E(\mu_i, \Omega_i) = \int_{\Omega_1} \left(I(\mathbf{x}) - \mu_1\right)^2 d\mathbf{x} + \int_{\Omega_2} \left(I(\mathbf{x}) - \mu_2\right)^2 + \nu\left|\delta\Omega_1\right| = \\ = \int_{\Omega} \left(I(\mathbf{x}) - \mu_1\right)^2 - \left(I(\mathbf{x}) - \mu_2\right)^2 H_\phi + \left(I(\mathbf{x}) - \mu_2\right)^2 d\mathbf{x} + \nu\int_{\Omega} \left|\nabla H_\phi\right| d\mathbf{x}, \tag{6}$$

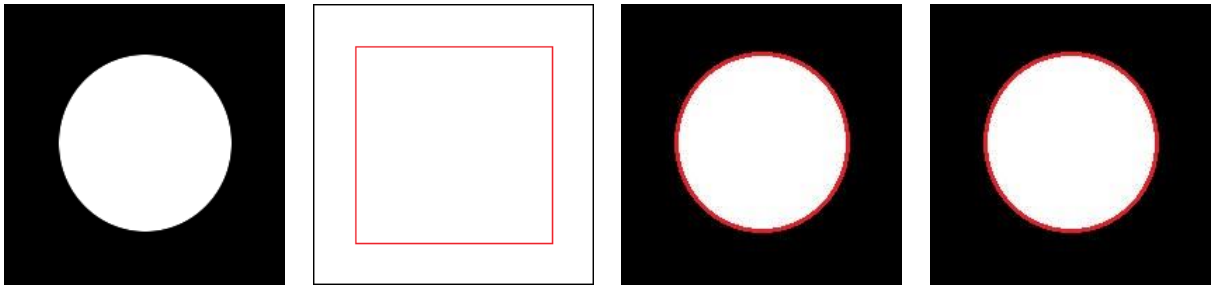
where $\boldsymbol{x}$ is the vector of the coordinates (x, y); $I(\boldsymbol{x})$ is the input image; ν denotes the regularization constant; and μ_1 and μ_2 express the arithmetic mean of the values of the region inside, or outside respectively, the curve C. By minimizing the potential from the above formula (6), we obtain the PDE to solve the level-set function in time. The equation is as follows:

$$\frac{\partial\phi}{\partial t} = \delta(\phi)\left(\nu div\left(\frac{\nabla\phi}{|\nabla\phi|}\right) + \left(I(\mathbf{x}) - \mu_2\right)^2 - \left(I(\mathbf{x}) - \mu_1\right)^2\right), \tag{7}$$

where $\delta(\phi)$ denotes the derivative of the Heaviside function (as mentioned above, this can be, for example, a sigmoid); the constant ν is, as in formula (6), a regularization one; and the meaning of μ_1 and μ_2 is identical with that in the said formula (6).

3. APPLICATION OF THE METHODS

This chapter describes the actual application of the edge- and region-based method to illustrative images, which show the differences between the resulting segmentations. The first test image consists of a white circle on a black background; the related segmentation is shown in Fig. 1. The initial curve was selected to be identical for both techniques (Figure 1 b), and the methods also exhibit the same segmentation results. In the given context, it then follows from this that the procedures yield such identical results if two homogeneous areas are separated by a sharp edge (a large gradient). The times required to run the edge- and region-based techniques on the applied PC[1] corresponded to 0.72 s and 2.24 s, respectively; these times, however, constitute only rough values, which may change depending on a concrete implementation and its optimization. Yet, as can be inferred from the proposed data, the region-based method is generally more time-intensive compared to its counterpart procedure.



a) b) c) d)

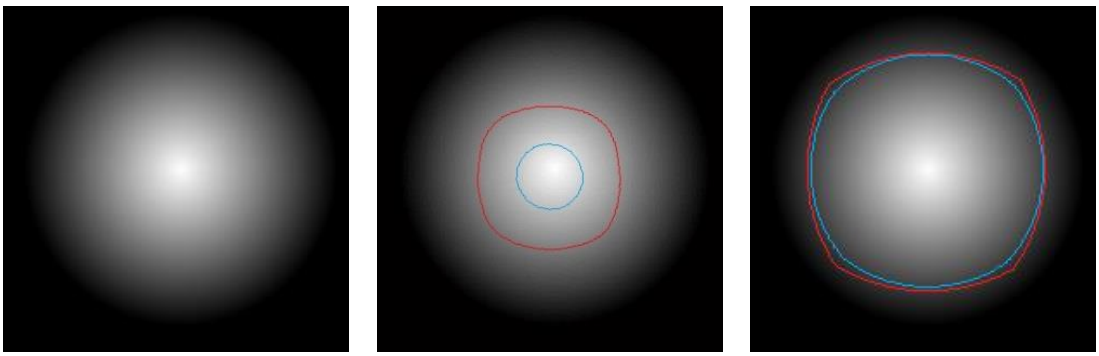
Figure 1: The segmentation of a simple shape: a) The input image; b) the initial contour; c) edge-based segmentation (100 iterations); and d) Chan-Vese segmentation (100 iterations).

Another case is presented in Fig. 2, where the input image again consists in a circle; the boundaries of this shape are nevertheless markedly blurred, thus constituting a fluent transition from the background to the object. In this situation, too, the initial curve is selected to be identical for both techniques (Figure 1 b). The described image configuration is nevertheless characterized by

[1] CPU: i5-4460 3.2 GHz, 8 GB RAM, 64bit Win7, implemented in Matlab 2013b.

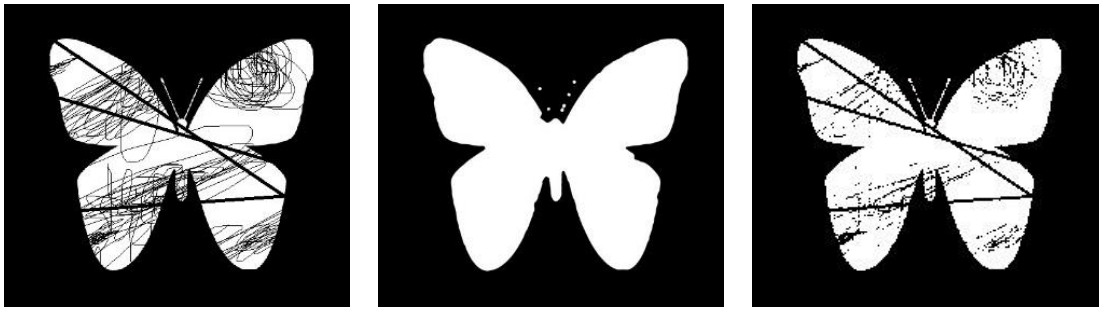
complete failure of the edge-based method. To illustrate the procedure‘s inapplicability, Figure 2 b) shows the curve after 200 (red) and 300 (blue) iterations, indicating that the curve gradually diminishes toward the centre of the image; then, after more than 300 iterations, the curve disappears entirely, and the segmentation finally fails due to the absence of image edges, which otherwise serve as the stop function for the development of the level-set function. Conversely, region-based segmentation remains fully effective, similarly to its performance in the previous image (Fig. 1). Figure 2 displays the results obtained through segmentation via the region-based Chan-Vese method after 300 (red) and 800 (blue) iterations.

The relevant portion of Figure 2 represents an unsuccessful application of the edge-based method. But, in certain conditions, this approach offers major advantages over the other option discussed. In this connection, Figure 3 shows an example of segmenting a damaged object; the related segmentation results for the edge- and region-based procedures are indicated in Figure 3, b) and c). After being treated with the edge-based method, the object assumes a shape almost equal to the form it had before the damage occurred; thus, the actual operation could be most aptly defined as image restoration. As regards region-based segmentation, it is true that individual parts of the object are separated, but such a result in not desirable for the given task.



a) b) c)

Figure 2: The segmentation of a shape with smooth transition of the edges: a) the input image; b) edge-based segmentation (200 iterations/red, 300 iterations/blue); c) Chan-Vese segmentation (300 iterations/ red, 800 iterations/blue).



a) b) c)

Figure 3: The segmentation of a disrupted object: a) the original image; b) edge-based segmentation; c) region- based segmentation.

4. CONCLUSION

In the opening sections, the article discusses the level set segmentation method, whose general framework is introduced to support the subsequent presentation of two relevant solution approaches, namely, the edge-based and region-based (Chan-Vese) procedures.

The capabilities of the techniques are then compared using several test images. According to the results of the related test cycles, the edge-based approach can be used exclusively when the image objects are separated by a detectable edge and, as such, is not applicable at fluent transitions. This technique, however, still performs more than satisfactorily in segmenting disrupted images and is usable for image restoration (Figure 3 b).

The region-based variant of the level-set method finds application also in images that contain unclear edges, creating a model of homogeneous areas based on close brightness levels. The technique can be utilized in separating objects with fluent brightness level transitions at the boundaries; however, compared to the edge-based method, this approach is computationally more intensive.

REFERENCES

[1] OSHER, Stanley a James A SETHIAN. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*. 1988, **79**(1), s. 12-49. DOI: 10.1016/0021-9991(88)90002-2. ISSN 00219991.

[2] SAPIRO, Guillermo. *Geometric partial differential equations and image analysis*. Cambridge: Cambridge University Press, 2001. ISBN 05-217-9075-1.

[3] CASELLES, Vicent, Ron KIMMEL a Guillermo SAPIRO. Geodesic active contours. *Proceedings of IEEE International Conference on Computer Vision*. IEEE Comput. Soc. Press, 1995, , s. 694-699. DOI: 10.1109/ICCV.1995.466871. ISBN 0-8186-7042-8.

[4] CHAN, Tony Fan-Cheong. a Luminita Aura VESE. Active contours without edges. *IEEE Transactions on Image Processing*. **10**(2), s. 266-277. DOI: 10.1109/83.902291. ISSN 10577149.

[5] MUMFORD, David a Jayant SHAH. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*. 1989, **42**(5), s. 577-685. DOI: 10.1002/cpa.3160420503. ISSN 00103640.

AUTOMATED ELECTROCHEMICAL ETCHING PROTOTYPE FOR COLD FIELD EMISSION CATHODES

Michal Jurčík
Doctoral Degree Programme (3), FEEC BUT
E-mail: xjurci03@stud.feec.vutbr.cz

Supervised by: Lubomír Grmela
E-mail: grmela@feec.vutbr.cz

Abstract: This article deals with a setup for automated cold cathode etching, its difficulties and benefits. As the method is well know there are no articles which describes this procedure in detail at least for tungsten wires from which cathodes are made. To provide stable and repeatable process of cathodes manufacturing, the experimental setup was proposed with full hardware control with aim to assure etching circuit disconnection in order of hundreds of nanoseconds.

Keywords: electrochemical etching, field emission cold cathode, tungsten, automation

1 INTRODUCTION

Cold field emission cathodes are widely used in electron microscopy where the small and stable source of electrons is required. The cold field emission cathodes differ from thermionic or Schottky cathodes by presence of strong electric field which modifies shape and height of the potential barrier and allows electrons to secede from the metal surface. Cathodes usually works at high levels of vacuum with relative low potential applied (approximately 200 V) and do not require additional heating like Schottky or thermionic cathodes [1].

Above mentioned cathode can be obatined either from tungsten wire or any other transition metals such as Niobium, Tantalum, Hafnium or Lanthanum. Electrochemical etching is a cathode manufacturing process in which are involved electrolyte (for tugsten usually KOH or NaOH aqueous solution at specified concentration), an anode and a cathode [2]. Tungsten wire acts as an anode and as a negative electrode a chemically resistant stainless steel wire is commonly used [3]. After applying ething potential between cathode and an anode, anode metal is dissolved as cation in the electrolyte and at the same time an equal amount of cation is deposited on the surface of cathode [4]. The rate of dissolution and deposition depends on the applied elecrochemical potential and concentration of the electrolyte solution and also the process of electrochemical reaction is inflouenced by a complex of other electrochemical effects on anode and cathode (e.g. non-uniform concentration of cations in solution, effect of impurities both in electrolyte and on the surfaces of electrodes, local inhomogenities of electric field etc.), but dissolution and deposition are still dominant processes [5]. The main goal of electrochemical etching of cathode is to a obtain sharp metal tip, which a source a stable emission of electron beam with low fluctuations in an electron microscope system. When the tungsten wire dissolves to an electrolyte, it reaches the point where the bottom part of wire drops off to the solution (from which the name of the method is derived) resulting in rapid voltage drop which is indicated by a steep decrease of current flow through the electrolyte. From this reason, there is a crucial requirement for control system of etching circuit to rapidly disconnect etching potential. If not, the wire tip continues to be etched, leading to decrease of cathode sharpness [6].

The proposed circuit for cathode etching is consisted from a DC power source, Microcontroller Unit (MCU) Arduino Mega 2560, a Printed Circuit Board (PCB) with control electronics and a stepper

motor. The MCU sets the parameters of PCB in most of the cases via Serial Peripheral Interface (SPI), such as Digital to Analog Converter (DAC) output voltage and its range, algorithmization of relays switching and selection between AC or DC output. The MCU was chosen because of its well known programming interface, reliability and cost effectivity. The PCB is used as a MCU shield and was designed to control the etching procedure fully via hardware, which eliminates the previous weak point of the process control by Matlab and the GPIB interface - the slow reaction to the drop of the part of the etched wire with subsequently delayed etching potential disconnection. This solution allows the system to simultaneous measurement and amplification of etching voltage and current, motor movement, detection of the etched wire drop, followed with automatic disconnection of the circuit. Because the PCB requires a stabilized power supply with various voltage ranges (+5, +12, +15, -5, -15 V), the power supply with multiple outputs was also designed and manufactured. As the last and the most challenging part of the whole circuit is the motor because as any motor it generates a significant noise background of the circuit due to the Pulse Width Modulation (PWM) which can not be overlooked. This feature is also remarkable during inactivity of the motor because of the effect of the standstill current.

Due to complexity of the circuit this article will be focused on the control PCB, which is the main part of the circuit.

2 EXPERIMENTAL RESULTS

The PCB controls whole etching process and is responsible for applying voltage, right immersion of the wire into the solution and disconnection of the circuit when the etching process is done. Since the PCB the source of etching signal, the DAC converter is utilized. The output of DAC is amplified by the power amplifier in order to obtain required potential between the anode and the cathode. In Fig. 1 the DAC_1 connection diagram as a power source is plotted. The PCB is intentionally designed to govern both oxidation and reduction electrochemical prosesses of further research on cathode manufacturing, i.e. not only etching of the filament, but also for electrochemical deposition of thin layers on the cathode surface with the purpose of the cathode lifecycle enlongation, the increase of its emissivity as a protection from the environmental damage. Beside this, electrodeposition of thin layers on the surface can seriously modify the electronical properties of the cathode, such as the work function of its surface. The desired function of the output is driven by relays K1 or K2. If the relay K2 is switched on, the DAC output will be inverted by the Operational Amplifier (OPA) 4131P.

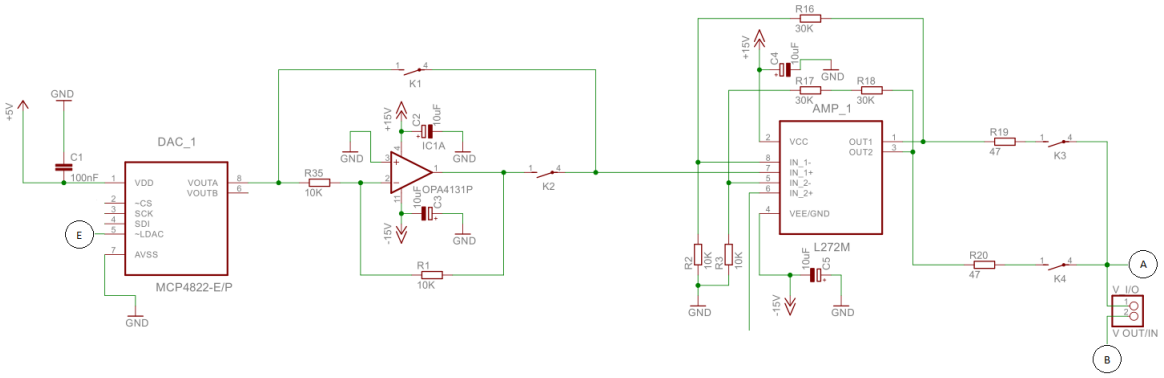


Figure 1: Schematic diagram - DAC power output with inverter (switchable polarity) and amplifier (4 times magnification).

The power amplifier AMP_1 L272M is used to amplify the output of the DAC_1 with the magnification of four. The reason of the signal magnification is the fact that the DAC internal voltage reference

is 2.048V, which is the maximum applicable etching voltage without amplification. Even though the 2.048 V is already to perform cathode etching, one of the requirements on the PCB is to provide higher etching voltage, which allows vaster etching parameters variability, such as control of etching rate by higher etching voltage etc. The second input of AMP_1 is used for generation of harmonic signal (generator not shown in the schema) for cleaning the wire before etching. In this case, the magnification is seven, because the amplitude of harmonic signal generator has its maximum at 1 V. The choice between etching/deposition or cleaning modes of the PCB is realized by appropriate combination of switch of relays K3 and K4. The ouputs of AMP_1 are connected to Analog to Digital Converter (ADC), which is part of the MCU (see Fig. 2) and to the tungsten filament.

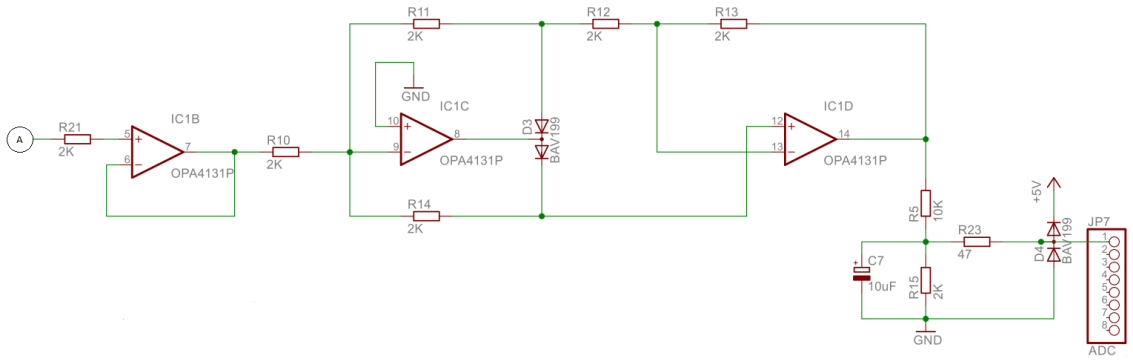


Figure 2: Schematic diagram - full-wave rectifier near ADC converter with voltage divider.

The etching voltage is measured via ADC. Because of the generation of harmonic, positive and negative DC output it has to contain a full-wave rectifier that converts the signal to a positive and prevents the MCU from damage caused by overvoltage as shown in Fig. 2. The rectifier also contains voltage follower composed from OPA to eliminate loading effects. The rectifier is terminated with a voltage divider 1:6 as the input voltage range of ADC is 0–5 V. The pair of diodes D4 strengthens the input protection that the voltage cannot be higher than +5 V and lower then 0 V.

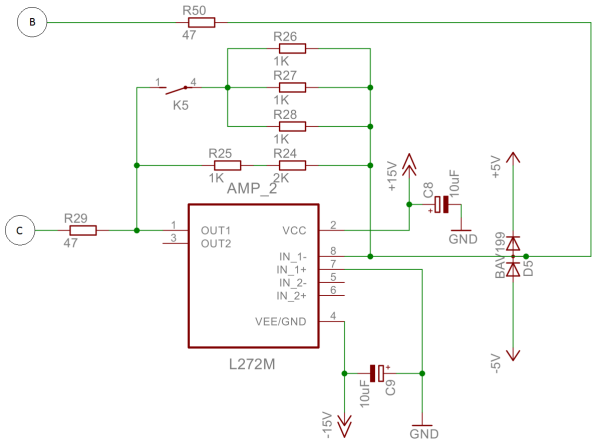


Figure 3: Schematic diagram - current to voltage converter with selectable range.

The cathode is connected to the inverted input of AMP_2 as shown in Fig. 3. This OPA is used as a current to voltage converter with the ability to measure current which flows in the circuit. The range of the amplifier in normal state is 0–10 mA. If the relay K5 is switched on, the range is increased to the value up to 100 mA. For the apllied potential (approximately 5 V), the highest measured current was

20 mA. Such a high current value is caused by biggest etching area of the cathode during initial stage of etching. Thus, the range should satisfy the conditions for even higher etching voltage. The output of the amplifier is devided into two branches where the first one is connected via same configuration as in Fig. 2 to another ADC of the MCU and the second one is connected directly to circuit disconnection logic.

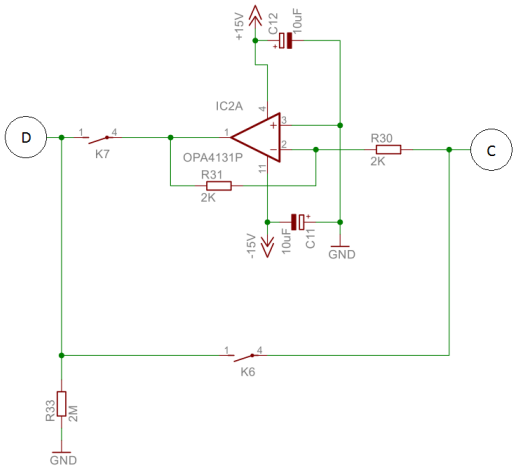


Figure 4: Schematic diagram - inverter near circuit disconnection logic.

If the polarity of output voltage of AMP_2 is negative, the invertor has to be connected into circuit in order to assure the correct function of circuit disconnection logic. This is carried out by switching the relay K7 as shown in Fig. 4. In other cases, i.e. if the output voltage of AMP_2 is either DC positive or harmonic, the inverter must not be used.

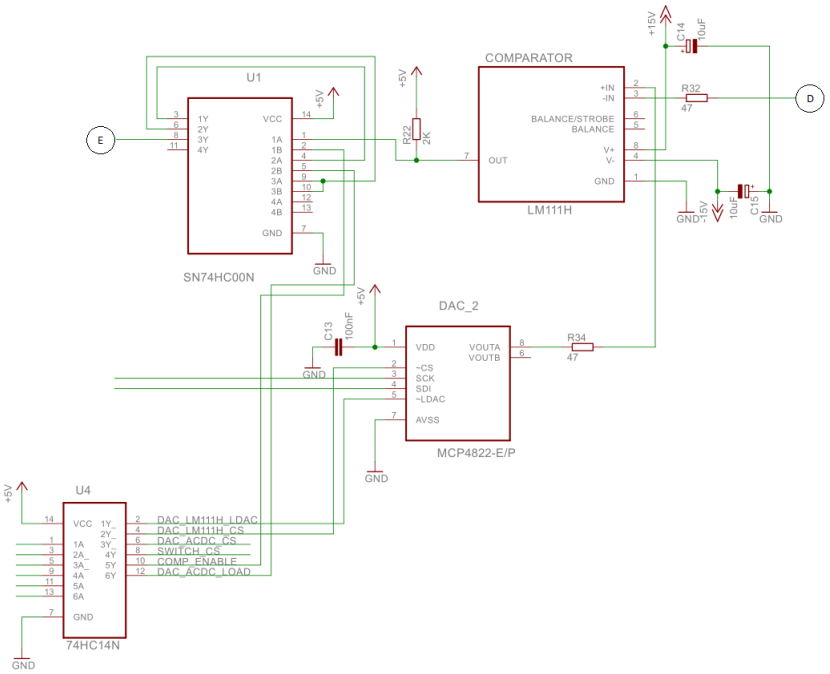


Figure 5: Schematic diagram - circuit disconnection logic.

In Fig. 5 is shown the main logic for circuit disconnection. It includes comparator, DAC_2 for setting

the reference voltage, NAND gate (U1) and inverter gate Schmitt trigger (U4). When the voltage on the comparator inverted input is lower then the reference one, comparator grounds the output and the low (logical 0) signal occurs on 1A input of U1. If output Y5 of U4 is set to high (logical 1) it activates latch registers of DAC_1 which was previously set to 0 V and shuts down the circuit.

3 CONCLUSION

This article showed the prototype of PCB used to cotrol fully automated electrochemical etching process of cold field emission cathode manufacturing. Due to the complexity of the whole etching system, the article contains only description of the main parts of the PCB circuit, which is already functional for application. Compared with the previous system, the reaction times of the system are 6 fold faster (hundreds of ns). One of the issues not described in this paper is the motor and its control. Motor carries out wire immersion and emmerge of a wire. When the motor is stopped on standstill current it generates pulses back to the circuit approximately with frequency of 40 kHz. This can not be ommited because the amplifiers also amplify the noise which may lead to instability of the circuit especially the comparator which is very susceptible to oscillations and may cause premature circuit disconnection.

ACKNOWLEDGEMENT

This work was supported by the Internal Grant Agency of Brno University of Technology, grant No. FEKT-S-14-2240.

REFERENCES

[1] L. W. SWANSON and G. A. SCHWIND. Review of ZrO/W Schottky cathode. *Handbook of charged particle optics*. 2nd ed, edited by Jon Orloff. Boca Raton: CRC Press, 2009, xi, 665 p. [cit. 2016-02-20]. ISBN 978-1-4200-4554-3.

[2] M. RICKART and M. BAUER: An improved tip etch procedure for reproducible sharp STM tips. AG Magnetismus, Universität Kaiserslautern, Annual report (1998), p. 53-55[cit. 2016-02-20].

[3] A. I. OLIVA, A. ROMERO G., J. L. PENA, E. ANGUIANO and M. AGUILAR. Electrochemical preparation of tungsten tips for a scanning tunneling microscope. *Review of Scientific Instruments* [online]. 1996, vol. 67, issue 5, p. 1917- [cit. 2015-02-02]. DOI: 10.140631.1146996. Available from URL:<http://scitation.aip.org/content/aip/journal/rsi/67/5/10.14063/1.1146996>.

[4] J. LINDSAY. A pulse/pulse reverse electrolytic approach to electropolishing and through-mask electroetching. Products Finishing. *Gardner Publications Inc.* [online]. HighBeam Research. 2011 [2016-02-20]. Available from URL:<https://www.highbeam.com>.

[5] O. M. BEHR and M. R. BEHR. Environmentally safe Etching. *Chem Tech* [online]. 1991, vol. 21, issue 4, p. 210-215 [cit. 2016-02-20]. Available from URL:<http://www.electroetch.com/chemtech.htm>.

[6] A. J. MELMED, J. The art and science and other aspects of making sharp tips. *Fifth internationl conference on scanning tunneling microscopy/spectroscopy*. 1991, vol. 9, issue 2, p. 601-608 [cit. 2016-02-20].

COMPUTER-ASSISTED SEARCH OF THE ROOTS OF THE POLYNOMIAL

Bedřich Smetana

Doctoral Degree Programme (1), FEEC BUT

E-mail: bedrich.smetana@unob.cz

Supervised by: Jan Chvalina

E-mail: chvalina@feec.vutbr.cz

Abstract: This article describes finding roots of polynomials, specifically aimed at finding real roots of a polynomial. Computer Support focuses on the use of generally available programs such as spreadsheets, portals to solve mathematical problems and freeware applications. Attention is devoted to a comparison of commercial and open source software.

Keywords: polynomials, Computer Support, freeware, open source.

1. INTRODUCTION

In mathematics, a polynomial is an expression consisting of variables (or indeterminates) and coefficients, that involves only the operations of addition, subtraction, multiplication, and non-negative integer exponents. An example of a polynomial of a single indeterminate (or variable),

x, is $f(x) = x^2 - 4x + 7$, which is a quadratic polynomial.

An example in three variables is

$g(x,y,z) = x^3 + 2xyz^2 + yz + 1$. [1]

Finding roots of a polynomial is a topic that is discussed in Mathematics in their first year at the military five-year study at the Faculty of Military Leadership at the University of Defense, at the civilian branch in Mathematics I. and also in learnings at other universities, like is University of technology.

Polynomials appear in a wide variety of areas of mathematics and science. For example, they are used to form polynomial equations, which encode a wide range of problems, from elementary word problems to complicated problems in the sciences; they are used to define polynomial functions, which appear in settings ranging from basic chemistry and physics to economics and social science; they are used in calculus and numerical analysis to approximate other functions. In advanced mathematics, polynomials are used to construct polynomial rings and algebraic varieties, central concepts in algebra and algebraic geometry. Generally polynomial is an expression of the form:

$$p(x) = \sum_{k=0}^{n} a_k x^k = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n ,$$

where $a_n \neq 0$. The numbers $a_0, a_1, \ldots, a_n$ are called polynomial coefficients. [2]

The number α is called a root of a polynomial p (x), if valid $p(\alpha) = 0$

Polynomial $p(x) = \sum_{k=0}^{n} a_k x^k$ can be written as

$$p(x) = (\ldots((a_n x + a_{n-1})x + a_{n-2})x + \ldots + a_1)x + a_0$$

This notation can be used to calculate polynomial in point of x in the procedure, which is known as Horner.

If you type

$$c_n = a_n,$$

$$c_{n-1} = c_n x + a_{n-1},$$

$$c_{n-2} = c_{n-1} x + a_{n-2},\ldots.$$

$$c_0 = c_1 x + a_0,$$

then last number c_0 is justly the value of a polynomial p (x) at x [1,4].

2. A BRIEF DESCRIPTION OF METHODS OF SOLUTION

Algebraic equation of degree n is called the equation

$$P_n(x) = a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_{n-1} x + a_n = 0,$$

and where $a_0, a_1, \ldots, a_n \in R$, $a_0 \neq 0$, $x \in \mathbf{C}$, $n \in \mathbf{N}$.

Solution (roots) algebraic equation $P_n(x) = 0$ is any number c, which is root of the polynomial $P_n(x)$. Algebraic equation of degree n therefore has exactly n roots (in the field of complex numbers regarding multiplicity).

There is not a formula analogous to the quadratic equation for the equation of degree at least 5, so in a simplified perspective solution can be divided into methods:

- numerical,
- graphics.

Graphically, we can only determine the real roots, and only approximately. $P_n(x) = 0$ adjust the shape of f(x) = g(x). The real roots of the equation $P_n(x) = 0$ They are then equal to the x-coordinates of the intersections curve y = f (x) and y = g (x).

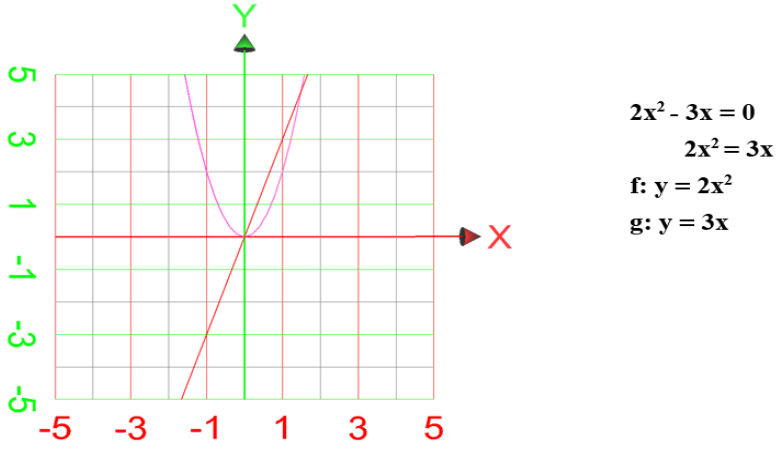


Figure 1: Example of graphical solution

With respect to iteration theory there are possibilities to represent functions using oriented functional (discrete) graphs [2,3].

Numerical Solution (Using formulas to decomposition product of root factors, using the Horner scheme, approximate numerical methods) may provide a more accurate result. We start from a few basic phrases that can be summarized as to cut a long story these basic findings:

- The degree of the polynomial P (x) with the root a1 can be reduced by dividing (x-a1).
- All the roots of the polynomial from set of all integers divides the absolute member. (With divisibility defined only for integer (generally an integral domain), hence an integral domain must be integer integer coefficients of the polynomial and the coefficient of the highest powers must be 1).

The number of positive roots is at most equal to the number of sign changes in the sequence of coefficients or differ by an even number, the number of negative roots is equal at most equal number of hits unsigned or less on an even number. The difference between the positive and negative roots is a number of complex roots- always even.

There is a whole range of terms and theorems associated with the search for roots, such as separation of roots, root approximations, interpolation, least squares methods and methods for special shapes algebraic equations. Given the scope of the article and determined focus on software support is more content devoted to the transfer of Horner's scheme into a spreadsheet, a modification example of freeware solutions and using mathematical portal.

3. A BRIEF DESCRIPTION OF SPREADSHEET SOLUTION

Unlike simple applications on the net, counting with one line objective was to gradually create a table reflecting the basic rules for the number of roots, with an emphasis on real roots. Necessity copying of rows and columns of formula led to a different absolute addressing of cells in the formula with the same result.

Original Coefficients	10	-30	-10	70	0	-40						70	The number of empty cells	4
Greatest common divisor	10	30	10	70	70	40	70	70	70	70	Greatest common divisor	10		
Roots / Coefficients	1	-3	-1	7	0	-4	0	0	0	0		Results		
-1	1	-4	3	4	-4	0	0	0	0	0		True		5
-1	1	-5	8	-4	0	0	0	0	0	0		True		6
2	1	-3	2	0	0	0	0	0	0	0		True		7

Figure 2: Image table of Horner scheme – Libre Office Calc

To simplify the equations used formula for the greatest common divisor -=GCD(B3;C3;D3;E3;F3;G3;H3;I3;J3;K3). Custom shorted coefficients provide condition =IF((COUNTBLANK(C2)=1);A1;C2/M3). The condition, which could be even better for the announcement of the results should shape: =IF(O5<(10-O2);"True";"False"). The formula within a table looks like this: =($A5*B5)+C4. Interesting option would be to automatic testing roots when applying of rules and complement the complex roots. It is clear that this possibility of referring to the fact that the set of real numbers is dense - each subset is the same size as the set of real numbers. On the other hand, for practical reasons, in the teaching are used the equations whose shape is "nice" and roots come "nice". No doubt, because math teachers are kind and welcoming people who love their students. This would allow to create effective application in narrowing down set of roots, usable in exercises and to control students. The condition of a localized version of Excel looks like this: =KDYŽ(B2=0;M2;ABS(B2)). The formula for the greatest common divisor -=GCD(B3;C3;D3;E3;F3;G3;H3;I3;J3;K3) is the same. Table between the two versions can be transferred only as numbers.

	A	B	C	D	E	F	G
1		1	3	3	3	5	
2	-1		-1	-2	-1	-2	
3		1	2	1	2		
4						3	<-zvyšok

Figure 3: Example of MS Excel solution from Internet[8]

Image shows an example of the Horner scheme in Excel on the net. [8]

4. A FREEWARE SOLUTION

From the portal Slunecnice.cz simple application can be downloaded as a zip file. This application just unzip example, using freeware 7-zip. The application can also be used by flashdisc.

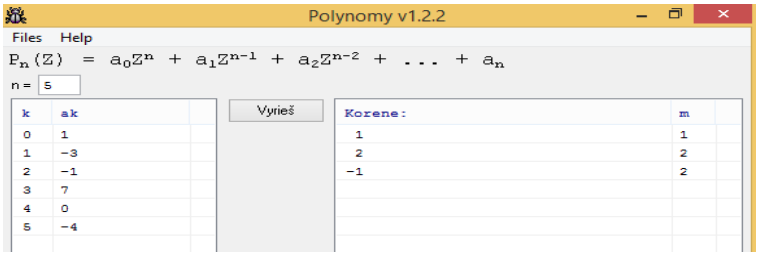


Figure 4: Freeware „Polynomy“ [5]

The polynomials is freeware for calculating the roots of the polynomial equation with real coefficients.

Main features:

- calculate the roots of a polynomial with real coefficients and determine their multiplicity
- the maximum degree of the polynomial 100
- calculation error root | Error | <1E-5 (at least five significant digits).

E.g. I need to find out the roots of the equation $2Z^2 + 8Z - 5 = 0$. This is a polynomial grade 2, so the first step is required in the box next to "n" to write degree polynomial number 2. The next steps should be gradually set values of individual coefficients with respect to the power variable. Coefficients are numbered from 0 index, so the first (A0) is the coefficient of the first term with the highest power of the polynomial (A0 Z^2), so a0 coefficient is the number 2.

The coefficients in this equation are non-zero and have values of 2, 8, -5.

To solve the equation is the click on "Solve". The roots of the equation is displayed in the window, at the same time be copied to the clipboard so that they can be simply by hitting Shift-Insert further arbitrarily used (eg. Notepad).

The application is a useful tool, in theory, can be used from a flash drive. Windows 8. in author computer blocked the use of flash drive.

5. A PORTALS SOLUTION

Embedded image demonstrates how to use the portal http://www.wolframalpha.com in solving algebraic equations. First, select the mathematics of the tabs, then Algebra - solving equations. Now you only have to insert the equation in the prescribed form, may overwrite pattern. As a result, you are not only roots, and also chart polynomial decomposition, where you can see the multiplicity of roots.

For a fee, you can obtain the gradual solution. Using the portal is nice, but it requires Internet access.

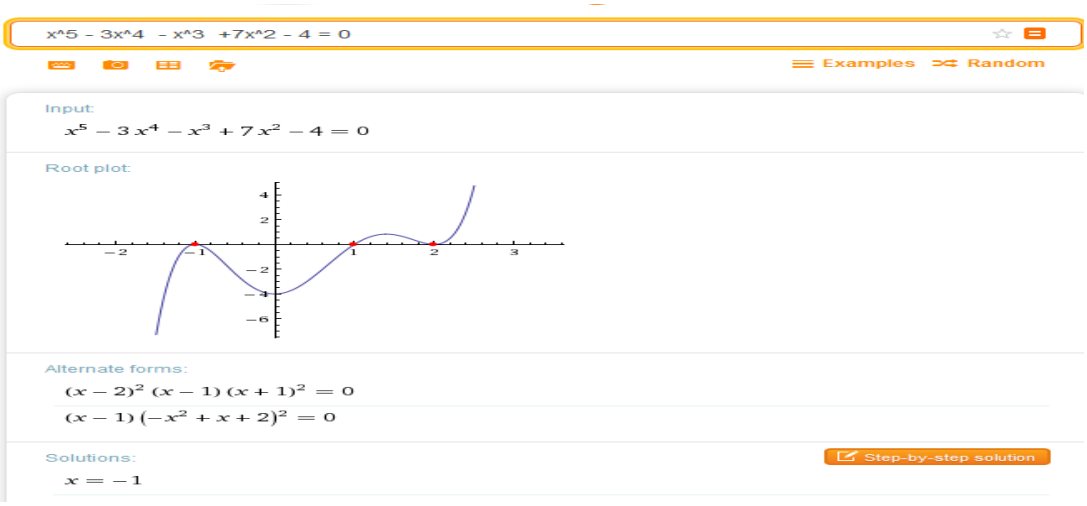


Figure 5: Portal Wolframalpha.com[6,7]

6. CONCLUSIONS

The article describes the possibilities of finding roots of polynomials using a computer. Scope of the article did not allow serious testing of the applications listed above, as well as more serious analysis of all aspects of the search for the roots of a polynomial. On the other hand, there student can find instructions for using a spreadsheet as a powerful tool for this purpose, and further inspiration for the use of various applications and Internet portals. The challenge may also be of use in the construction of mathematical logic conditional functions, allowing to solve not only other than math problems.

REFERENCES

[1] MOUČKA, J., RÁDL, P. *Matematika pro studenty ekonomie*. 2. Praha: Grada Publishing a.s., 2015. ISBN 978-80-247-5406-2.

[2] CHVALINA, J., MATOUŠKOVÁ, K. O aditivních funkcích zaměnitelných s polynomem stupně alespoň druhého. In *Matematika, fyz. tech. X. Sborník prací PdF UJEP*. sv. 93. Brno: PdF UJEP, 1988. s. 65 (s.)

[3] CHVALINA, J., MATOUŠKOVÁ, K. O jedné vlastnosti nejjednodušší kvadratické funkce I. In *Matematika a didaktika matem. Sborník prací PdF UJEP,řada č. ll, sv. 104.* sv. 104. 1989. s. 55 (s.)

[4] HLAVÁČEK, A: *Sbírka řešených příkladů z vyšší matematiky pro přípravu pracujících ke studiu na vysokých školách* (Praha, 1971)

[5] Polynomy. *Http://www.slunecnice.cz* [online]. Praha, 2016 [cit. 2016-02-05]. Dostupné z: http://www.slunecnice.cz/sw/polynomy/

[6] Math. *Www.wolframalpha.com* [online]. Long Hanborough: The Wolfram Centre Lower Road, 2009 [cit. 2016-02-05]. Dostupné z: http://www.wolframalpha.com/examples/Math.html

[7] ŘÍHA, J., LÁTAL, F., ŘÍHOVÁ, V. WolframAlpha ve výuce přírodovědných a ekonomických předmětů. 1. vydání. Olomouc: Univerzita Palackého v Olomouci, 2015, 72 stran. Odborné publikace (Univerzita Palackého v Olomouci). ISBN 978-80-244-4471-0.

[8] STAŇO, M. Tabuľkový procesor vo výučbe. In: *Www.infovek.sk* [online]. Bratislava: Infovek, 2001 [cit. 2016-02-05]. Dostupné z: http://www.infovek.sk/predmety/matem/mater/prip/tpvovyuc.pdf

EPOXY RESIN WITH ALUMINIUM OXIDE NANOFILLER AND ITS DIELECTRIC PROPERTIES

Aneta Lontrasová

Doctoral Degree Programme (1), FEEC BUT

E-mail: xlontr00@stud.feec.vutbr.cz

Supervised by: Karel Liedermann

E-mail: liederm@feec.vutbr.cz

Abstract: Topic of this paper is dielectric properties of epoxy resins, containing nanoparticles of Al_2O_3. Dielectric spectroscopy was used during epoxy nanocomposite diagnostics. These new materials are candidates for use in manufacturing modern insulators, where it can aid in reducing their dimensions. This research also focuses on studying dielectric properties of Al_2O_3 nanofiller over long periods. Large number of nanoparticles and area between them in epoxy matrix means higher number of weak spots, which may be the source of defects and may lead to increased deterioration of nanocomposites compared to composites without nanoparticles.

Keywords: epoxy resin, spectroscopy, relaxation, diagnostics, Al_2O_3 nanofiller

1. INTRODUCTION

Constantly increasing requirements on size reduction of electric devices as well as also improving its parameters makes manufacturers to improve electric insulation of their products. For example, replacing SF_6 insulation system on 66 kV distribution grids or increasing their operational voltage. One of the proposed solutions [1], [2] is based on replacing SF_6 with a combination of vacuum and solid insulation made of nanocomposites containing epoxy and nano- and microparticles. Currently, properties of these particles are subject of intense research, but despite that, their life expectancy and stability are still not well known. This is most unfortunate, because applications of these materials count with 20 – 30 years of operation. Electrical insulating composite with nanoparticles contains, regarding their structure, large amount of interfaces, which in long-term view could be a source of problems. The goal of our research is long-term study of dielectric properties, which can be simulated by accelerated ageing at high temperatures.

2. PRESENT STATE AND BACKGROUND

The subject of our research is epoxy resins containing non-conducting Al_2O_3 nanoparticles. Material samples were acquired from Department of Electrotechnology of Faculty of Electrical Engineering and Communication, Brno University of Technology. Samples were cast in a special mould supplied by ABB Brno. During the process of sample manufacturing, epoxy, hardener, softener and hardening accelerators were mixed together in correct ratio (weight). The resulting mixture was heated and mixed at 60 °C, so the epoxy did not harden and was more viscous, resulting in smoother flow into the mould. Weight of the whole epoxy system was approximately 350 – 500 g and the ratio of nanoparticles was set to 3 %. As soon as the nanoparticles were added, they reacted in the epoxy, creating small clusters accompanied by air bubbles. By this, the filler increases its density and must be degassed. First, nanoparticles are mixed mechanically with the epoxy, followed by ultrasonic mixing for 10 – 30 minutes. Having been degassed, the mixture enters the first hardening stage lasting 3 – 4 hours at 90 – 110 °C. Once the mixture has a rubbery consistency, the mould is disassembled, the samples are removed from the mould and prepared for hardening in the second phase, so they can be used in electrode system. The second stage lasts 10 –

12 hours at 140 °C. Nanoparticles used in the process were supplied by Sigma Aldrich; manufactured using calcination. Purity of the supplied material was 99.2 % and average size of a nanoparticle was 15 nm. The supplier also guarantees that the diameter of nanoparticles will not exceed 50 nm. Dimensions of our nanocomposite samples were 2.5 x 30 x 2 mm. This batch of samples was too thick for dielectric measurements. Another batch was made using a different mould – thickness of these did not exceed 1 mm and they had capacity at least 20 pF. Samples were coated with graphite or silver electrode paste. Several different attachments of samples to measurement equipment were possible. First of all, samples were inserted into Novocontrol holder and analyzed by standard Novocontrol Alfa A analyzer, with frequency range 0,4 μHz – 40 MHz.



Figure1: Novocontrol Alfa-A Analyzer – Example of attachment in Novocontrol Alfa A [6].

In the second case, samples with contacted or sprayed electrodes connected to cold head of cryostat and each sample had thermal contact with the cold head. This mechanical attachment to the head has several advantages – most important being the simplicity of the attachment of different samples in cryostat, so it is not necessary to create contact surfaces using graphite or silver paste. Because the heated sample is subjected to cold of slightly warm nitrogen, it is important that the thermal contact is as good as possible. Terminals from electrode system are connected via BNC cables to Novocontrol Alfa-A analyzer. The control PC managed the whole measurement, including temperature control, so that the influence of the environment on measurement was minimized [4].

3. EXPERIMENT AND RESULTS

Results obtained from Novocontrol Alfa-A analyzer were processed in MS Excel. Figures 2 and 3 show loss factor ε'' as a function of frequency. The graph clearly shows the impact of temperature changes from -125 °C (on the left) up to 125 °C (right).

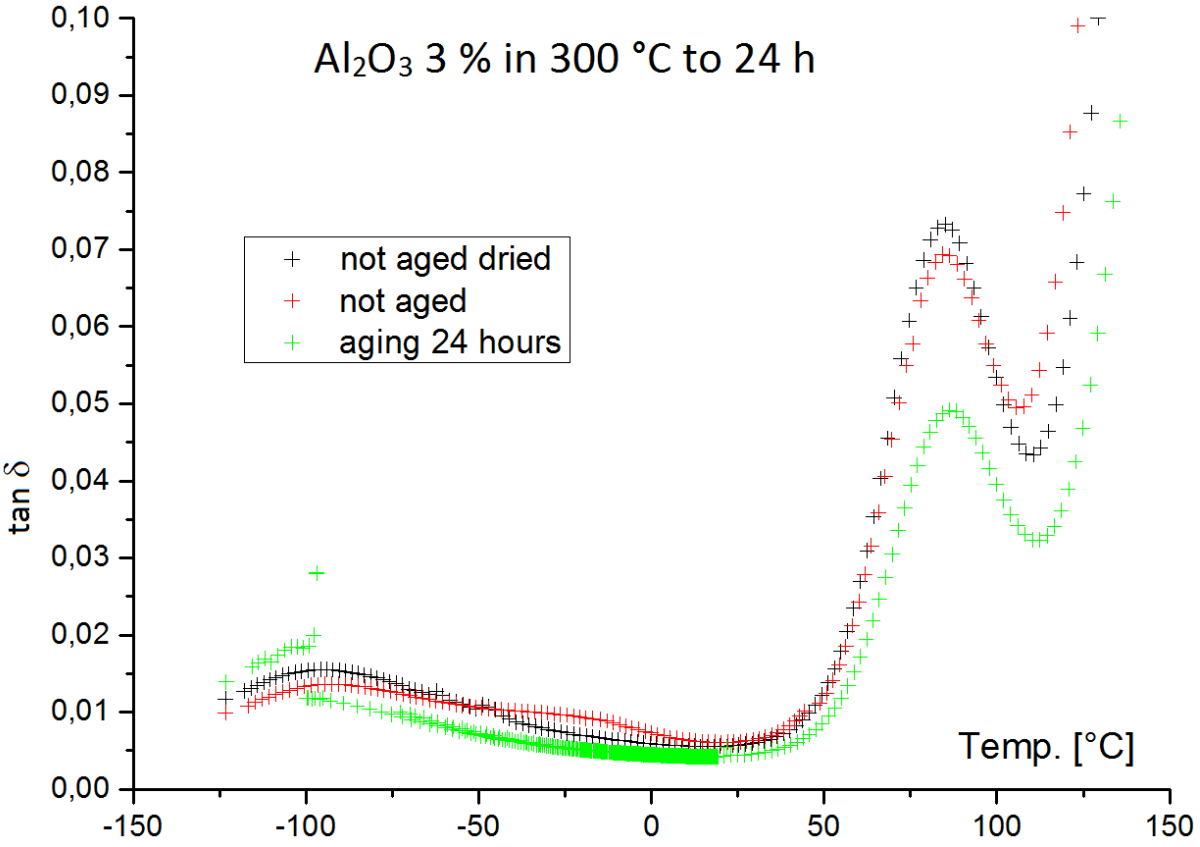


Figure 2: Al_2O_3 aged for 24 hours and drying.

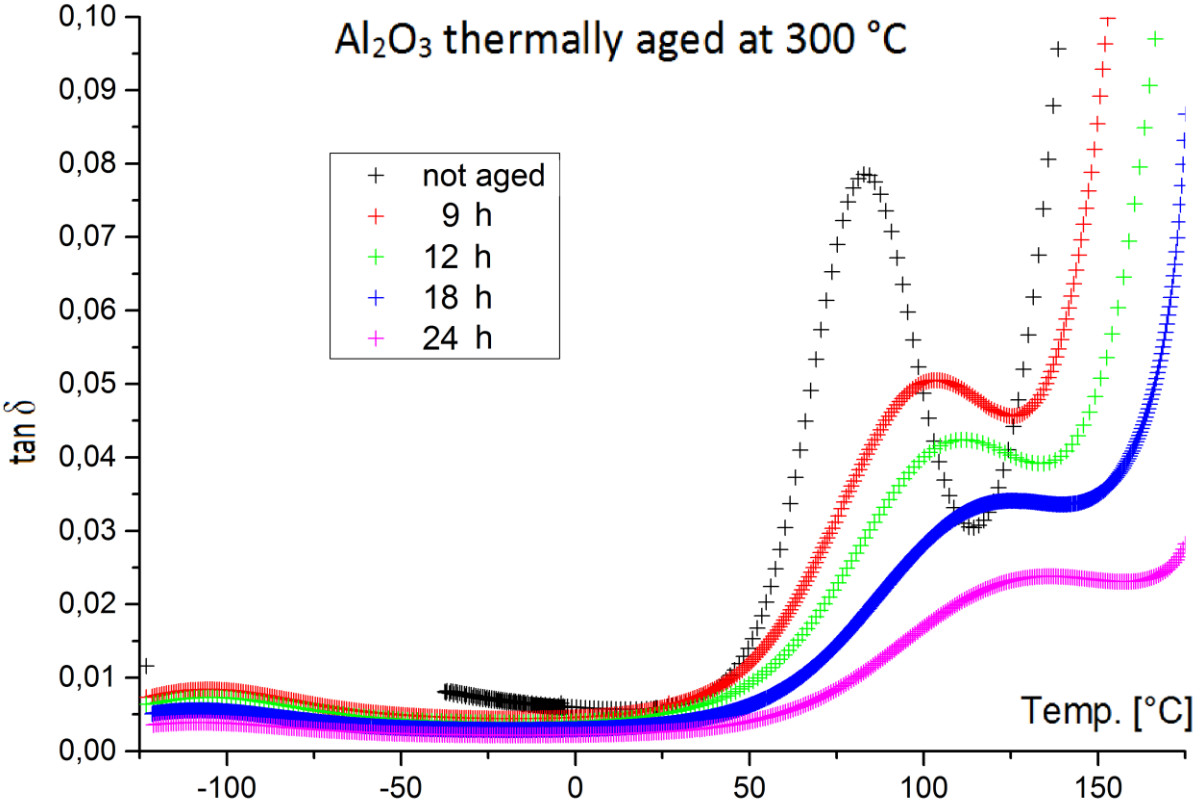


Figure 3: Tan δ dependence on temperature for epoxy resins with 3% Al_2O_3.

The influence of the temperature is clear. In case there are two relaxation mechanisms, the position of the β relaxation maximum is expected in the left part and α relaxation on the right side, before conductivity. The possible explanation might be a combination of several processes. The more probable explanation is incorrect temperature measurement, resulting in incorrect temperature data. Electric contacts of samples were correct, but thermal contact with cold head was insufficient. Thermal sensor was mounted on metallic holder, which was firmly connected to the cold head. Sample was stuck to the cold head using Apiezon H. Acquired data show that the temperature of the sample is not equal with the temperature measured by thermal sensor. This problem was solved, as shown in the following measurement. In similar fashion as the not aged sample in Figure 3, shows thermal dependency of relative dielectric constant (permittivity), which is a region between two relaxations. Discrepancy in thermal data can be observed at lower temperatures.

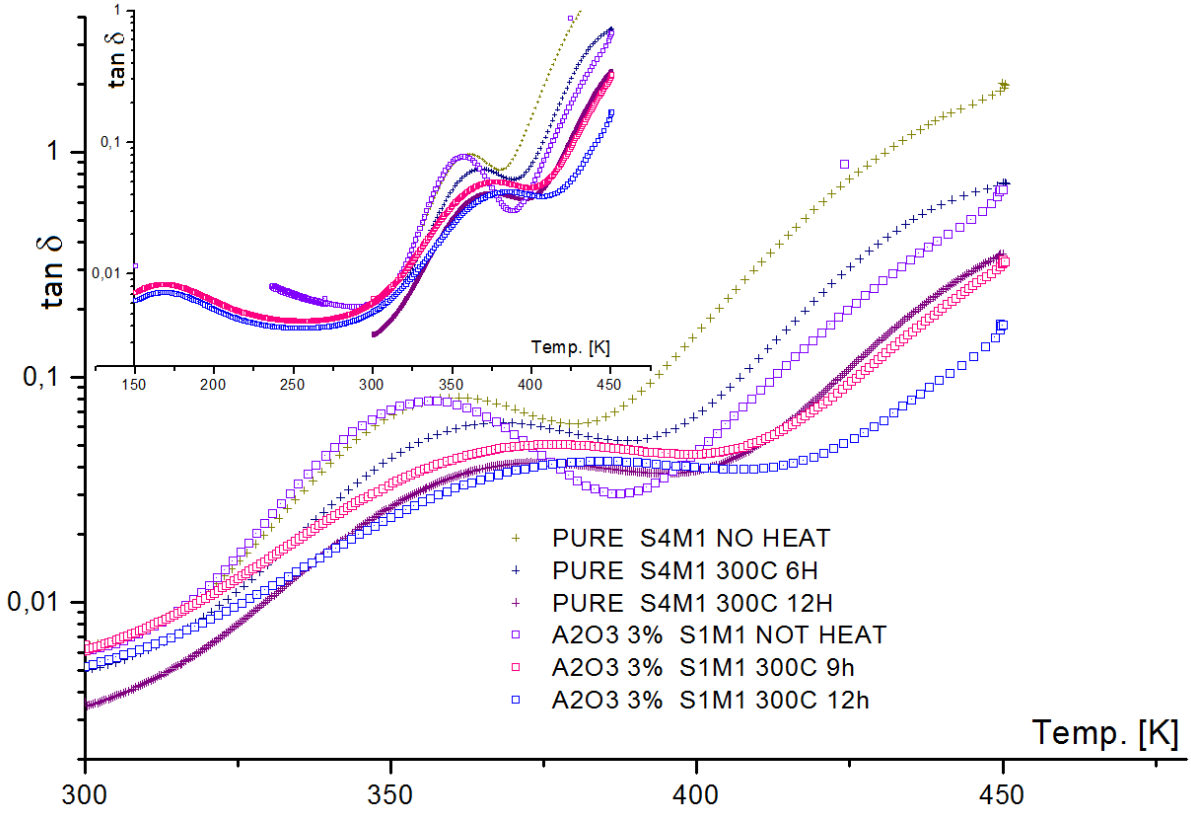


Figure 4: Tan Delta dependence on temp. for epoxy resins with 3% Al_2O_3 compared with pure epoxy.

Figure 4 shows six curves, the larger shows more detailed region of alpha relaxation, where it is also possible to observe changes and impact of the nanoparticles on the material. It can be seen that in the beginning, the maximum associated with the presence of nanofillers is more in the front than pure epoxy, but during the ageing process, the difference is reversed and further magnified. Finally, this maximum is 370 K for pure epoxy and 390 K for Al_2O_3 aged for 12 hours at 300 °C. Pure epoxy samples after ageing also show small maxima on the right in the conductivity area, but they were not subjects of our research. Another thing to mention is that not all samples were measured at subzero temperatures, mostly due to time required to measure and evaluate. The last anomaly is Al_2O_3 sample, which can be seen in region from 300 K and lower, showing considerable increase, which may have been caused by two factors. First, as published in [6], the author claims that there is a water peak which can be observed mostly on samples subjected to air and thus not properly degassed. The second option is a random error in Novocontrol system.

4. CONCLUSION

Observed dielectric spectra exhibit two relaxations and electrical conductivity is present. Thermal dependency of dielectric properties of nanocomposites was studied and measured. From measured data, we can conclude that maximum in the region of the α relaxation maximum, at which we observe shifts towards higher temperatures closer to conductivity, can show itself more clearly during the decrease of the magnitude of conductivity. One of the negative factors contributing to the data scattering might be the absence of uniform nanoparticle distribution in epoxy resin.

ACKNOWLEDGEMENT

This research was carried out under the project CEITEC 2020 (LQ1601) with financial support from the Ministry of Education, Youth and Sports of the Czech Republic under the National Sustainability Programme II.

REFERENCES

[1] N. Tagami, M. Hyuga, Y. Ohki, T. Tanaka, T. Imai, M. Harada, and M. Ochi, Comparison of Dielectric Properties between Epoxy Composites with Nanosized Clay Fillers Modified by Primary Amine and Tertiary Amine, IEEE Transactions on Dielectrics and Electrical Insulation, No. 2, 17 (2010), 214 – 220

[2] N. Tagami, M. Okada, N. Hirai, Y. Ohki, T. Tanaka, T. Imai, M. Harada and M. Ochi, Dielectric Properties of Epoxy/Clay Nanocomposites – Effects of Curing Agent and Clay Dispersion Method, IEEE Transactions on Dielectrics and Electrical Insulation, No. 1, 15 (2008), 24 – 32

[3] R. Pfaendner, Nanocomposites: Industrial opportunity or challenge?, Polymer Degradation and Stability, 95 (2010) 369 – 373

[4] Manual HP http://home.agh.edu.pl/~aguilar/classes/sem1/compmeas/LCR_manual.pdf

[5] M. Klampár – Epoxy resin with titanium dioxide nanofiller and its dielectric properties Proceedings of 9th international conference, Brno, (2011) 69 – 73

[6] M. Klampár – Štúdium nanokompozitov pre elektrické izolácie. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Ústav fyziky, 2015. s. 117.,

GLOBAL SOLUTIONS TO LINEAR DELAYED DIFFERENTIAL EQUATIONS

Gabriela Vážanová
Doctoral Degree Programme (1), FEEC BUT
E-mail: xvincu00@stud.feec.vutbr.cz

Supervised by: Josef Diblík
E-mail: diblik@feec.vutbr.cz

Abstract: In this paper is proved the existence of solutions of a linear delayed differential equation on the whole real line $(-\infty,\infty)$.

Keywords: delayed differential equation, global solution, sufficient conditions

1 INTRODUCTION

Let us consider the equation

$$\dot{y}(t) = -a(t)y(t) - \sum_{i=1}^{m} d_i(t)y\left(t-\delta_i(t)\right) + \omega(t). \tag{1}$$

Throughout the paper we will assume that $a:\mathbb{R}\to(0,\infty)$ is Lipschitz continuous and bounded function on every interval $(-\infty,\theta]$, for $\theta\in\mathbb{R}$, $d_i:\mathbb{R}\to(0,\infty)$, $i=1,\ldots,m$, are Lipschitz continuous and bounded functions on every interval $(-\infty,\theta]$, for $\theta\in\mathbb{R}$, the delays $\delta_i:\mathbb{R}\to(0,r]$, $i=1,\ldots,m$ are Lipschitz continuous functions on every interval $(-\infty,\theta]$, for $\theta\in\mathbb{R}$ and r is a positive constant. Function $\omega(t):\mathbb{R}\to\mathbb{R}$ is assumed to be continuously differentiable and satisfies the following two conditions:

(A) For every positive constant C and every $\theta\in\mathbb{R}$ there exists a constant $N_{\theta,C}$ such that for every $t\in(-\infty,\theta]$ the following estimate holds

$$\left(\left|\omega'(t)\right| + C\left|\omega(t)\right|\right)e^{C|t|} \leqslant N_{\theta,C}.$$

(B) There exists a point $t^*\in\mathbb{R}$ such that

$$\omega(t)\begin{cases} \geqslant 0 & \text{for } t<t^*,\\ = 0 & \text{for } t=t^*,\\ \leqslant 0 & \text{for } t>t^*. \end{cases} \tag{2}$$

The symbol "$\dot{}$" denotes a derivative, which we assume is at least right-sided.

A global solution of the equation (1) is defined as a continuously differentiable function y: $\mathbb{R}\to\mathbb{R}$ satisfying (1) on $\mathbb{R}$. In the paper we give conditions sufficient for the existence of such solutions to equation (1). Unlike of ordinary differential equations, problem of existence of global solutions to (1) is not trivial, because backward continuation of a solution in the class of continuously differentiable functions is often immpossible.

This paper is organized as follows. In section 2 auxiliary inequalities and criteria for existence of global solution of retarded functional differential equations are cited. The result for equation of a type (1) and an example are given in section 3. In the last section the conclusion is formulated.

2 AUXILIARY INEQUALITIES AND LEMMA

Let us consider a retarded functional differential equation

$$\dot{y}(t) = f(t, y_t), \tag{3}$$

where $f : \Omega \to \mathbb{R}$, $\Omega \subseteq \mathbb{R} \times C_r$ is a continuous quasi-bounded map which satisfies a local Lipschitz condition with respect to the second argument and C_r is the Banach space of the continuous functions from $[-r, 0]$ into $\mathbb{R}$. Below we assume (we do not give concrete specification of Ω) that the used operations concerning f are well defined.

Let $\lambda : (-\infty, \theta] \to \mathbb{R}$ be a continuous function where $\theta \in \mathbb{R}$, $t^* \in (-\infty, \theta]$ and $k \in (0, \infty)$. Define

$$I_{(-\infty,\theta]}(k, t^*, \lambda) = k \cdot \exp\left(-\int_{t^*}^{t} \lambda(s)\, ds\right), \quad t \in (-\infty, \theta], \tag{4}$$

and

$$\left(T_{(-\infty,\theta]}\lambda\right)(t) = -k^{-1} \cdot \exp\left(\int_{t^*}^{t} \lambda(s)\, ds\right) \cdot f\left(t, \left(I_{(-\infty,\theta]}(k, t^*, \lambda)\right)_t\right), \quad t \in (-\infty, \theta].$$

The following lemma is a special case of Theorem 2.1 from [1]:

Lemma 1. Let $t^* \in \mathbb{R}$ be fixed. Moreover, we assume:

1. For any $M \geqslant 0$, $\theta \geqslant t^*$ there exists a constant K, such that for all $t,\ t' \in (-\infty, \theta]$ and for any continuous function $\lambda : (-\infty, \theta] \to \mathbb{R}$ with $|\lambda| \leqslant M$

$$\left|\left(T_{(-\infty,\theta]}\lambda\right)(t) - \left(T_{(-\infty,\theta]}\lambda\right)(t')\right| \leqslant K\left|t - t'\right|.$$

2. There are $k \in (0, \infty)$ and continuous functions $\lambda_i : \mathbb{R} \to \mathbb{R}, i = 1, 2$, which are bounded on every interval $(-\infty, \theta]$, for $\theta \in \mathbb{R}$ and which satisfy

$$\lambda_1(t) \leqslant \lambda_2(t),$$

$$\lambda_1(t) \leqslant -k^{-1} \cdot \exp\left(\int_{t^*}^{t} \lambda_1(s)\, ds\right) \cdot f\left(t, \left(I_{(-\infty,\infty)}(k, t^*, \lambda_1)\right)_t\right), \tag{5}$$

$$\lambda_2(t) \geqslant -k^{-1} \cdot \exp\left(\int_{t^*}^{t} \lambda_2(s)\, ds\right) \cdot f\left(t, \left(I_{(-\infty,\infty)}(k, t^*, \lambda_2)\right)_t\right) \tag{6}$$

on $\mathbb{R}$.

3. For any continuous functions $\Lambda_i : \mathbb{R} \to \mathbb{R}, i = 1, 2$, such that $\Lambda_1(t) \leqslant \Lambda_2(t), t \in \mathbb{R}$ we have $T_{(-\infty,\infty)}\Lambda_1 \leqslant T_{(-\infty,\infty)}\Lambda_2$.

Then there exists a global solution $y : \mathbb{R} \to \mathbb{R}$ of (3) satisfying $y(t^*) = k$ and

$$k \cdot \exp\left(-\int_{t^*}^{t} \lambda_1(s)\,ds\right) \leqslant y(t) \leqslant k \cdot \exp\left(-\int_{t^*}^{t} \lambda_2(s)\,ds\right) \quad \text{for} \quad t < t^*, \tag{7}$$

$$k \cdot \exp\left(-\int_{t^*}^{t} \lambda_2(s)\,ds\right) \leqslant y(t) \leqslant k \cdot \exp\left(-\int_{t^*}^{t} \lambda_1(s)\,ds\right) \quad \text{for} \quad t > t^*. \tag{8}$$

3 SUFFICIENT CONDITIONS FOR EXISTENCE OF GLOBAL SOLUTIONS OF (1)

Below we apply Lemma 1 to equation (1). As a result we get not only the existence of global solutions of equation (1) but also their upper and lower bounds on $\mathbb{R}$.

Theorem 1. Suppose that there are $k \in (0,\infty)$ and continuous functions $\lambda_1, \lambda_2 : \mathbb{R} \to \mathbb{R}$ which are bounded on every interval of the form $(-\infty, \theta]$, for $\theta \in \mathbb{R}$ and which satisfy the inequalities

$$\lambda_1(t) \leqslant \lambda_2(t), \tag{9}$$

$$\lambda_1(t) \leqslant a(t) + \sum_{i=1}^{m} d_i(t) \exp\left(\int_{t-\delta_i(t)}^{t} \lambda_1(s)\,ds\right) - \frac{1}{k}\omega(t) \exp\left(\int_{t^*}^{t} \lambda_1(s)\,ds\right), \tag{10}$$

$$\lambda_2(t) \geqslant a(t) + \sum_{i=1}^{m} d_i(t) \exp\left(\int_{t-\delta_i(t)}^{t} \lambda_2(s)\,ds\right) - \frac{1}{k}\omega(t) \exp\left(\int_{t^*}^{t} \lambda_2(s)\,ds\right), \tag{11}$$

where $t^* \leqslant \theta$ and $t \in \mathbb{R}$. Then there exists a global solution y of (1) such that $y(t^*) = k$, moreover,

$$k \cdot \exp\left(-\int_{t^*}^{t} \lambda_1(s)\,ds\right) \leqslant y(t) \leqslant k \cdot \exp\left(-\int_{t^*}^{t} \lambda_2(s)\,ds\right) \quad \text{if} \quad t < t^*, \tag{12}$$

$$k \cdot \exp\left(-\int_{t^*}^{t} \lambda_2(s)\,ds\right) \leqslant y(t) \leqslant k \exp\left(-\int_{t^*}^{t} \lambda_1(s)\,ds\right) \quad \text{if} \quad t > t^*. \tag{13}$$

Proof. If we show, that all the assumptions of Lemma 1 are fullfilled, the proof will be complete.

The right-hand side of the equation (1) is

$$f(t, y_t) = -a(t)y(t) - \sum_{i=1}^{m} d_i(t) y(t - \delta_i(t)) + \omega(t). \tag{14}$$

By substituting (14) and (4) in (5) and (6) we may see that inequalities (9), (10) and (11) are equivalent to condition 2. from Lemma 1.

The operator

$$(T_{(-\infty,\theta]}\lambda)(t) = a(t) + \sum_{i=1}^{m} d_i(t) \exp\left(\int_{t-\delta_i(t)}^{t} \lambda(s)\,ds\right) - \frac{1}{k}\omega(t) \exp\left(\int_{t^*}^{t} \lambda(s)\,ds\right)$$

is nondecreasing because the function ω satisfies (2).

The remaining task is to show, that there exist real numbers K_i, $i = 1,2,3$ such that

$$\left|a(t) - a(t')\right| \leqslant K_1 \left|t - t'\right|, \tag{15}$$

$$\left|\sum_{i=1}^{m} d_i(t) \exp\left(\int_{t-\delta_i(t)}^{t} \lambda(s)\, ds\right) - \sum_{i=1}^{m} d_i(t') \exp\left(\int_{t'-\delta_i(t')}^{t'} \lambda(s)\, ds\right)\right| \leqslant K_2 \left|t - t'\right|, \tag{16}$$

$$\left|\omega(t) \exp\left(\int_{t^*}^{t} \lambda(s)\, ds\right) - \omega(t') \exp\left(\int_{t^*}^{t'} \lambda(s)\, ds\right)\right| \leqslant K_3 \left|t - t'\right| \tag{17}$$

hold for any continuous $\lambda(t)$ and for all $t, t' \in \mathbb{R}$.

Since the function a is assumed to be Lipschitz continuous, (15) is obvious. Inequality (16) can be verified using Lagrange's mean value theorem, Lipschitz continuity of the functions d_i and δ_i and boundedness of λ (For more details see [1]). Now we will look at the inequality (17). Using Lagrange's mean value theorem and the fact that $|\lambda(t)| \leqslant M$ we get

$$\left|\omega(t) \exp\left(\int_{t^*}^{t} \lambda(s)\, ds\right) - \omega(t') \exp\left(\int_{t^*}^{t'} \lambda(s)\, ds\right)\right| \leqslant \left(\left|\omega'(t_\triangle)\right| + \left|\omega(t_\triangle)\right| M\right) \mathrm{e}^{\left(M|t_\triangle - t^*|\right)} \left|t - t'\right|,$$

where $t_\triangle$ is between t and t^*. It can be easily seen, that (17) holds for $K_3 = N_{\theta,M} \cdot \mathrm{e}^{M|t^*|}$.

Upper and lower bounds (12), (13) for solution are derived from inequalities (7) and (8). □

3.1 EXAMPLE.

Consider the equation

$$\dot{y}(t) = -0,001y(t) - \frac{1+0,5\sin t}{35\mathrm{e}} \cdot y(t-1) - \frac{1+0,5\cos t}{30\mathrm{e}} \cdot y(t-5) - t\mathrm{e}^{-t^2} \tag{18}$$

on $(-\infty, \infty)$. This is an equation of the type (1) with $a(t) = -0,001$, $m = 2$, $d_1(t) = (1+0,5\sin t)/(35\mathrm{e})$, $d_2(t) = (1+0,5\cos t)/(30\mathrm{e})$, $\delta_1(t) = 1$, $\delta_2(t) = 5$ and $\omega(t) = -t\mathrm{e}^{-t^2}$. We will search for solutions in the form (4). Let us set $t^* = 0$, k will be specified later. Our task is to find continuous functions $\lambda_1(t)$ and $\lambda_2(t)$ satisfying conditions (9), (10) and (11). For simplicity, we will assume that $\lambda_1(t)$ and $\lambda_2(t)$ are constants, i.e. $\lambda_1(t) \equiv \lambda_1$, $\lambda_2(t) \equiv \lambda_2$. Therefore these conditions may be written as follows

$$\lambda_1 \leqslant 0,001 + \frac{1+0,5\sin t}{35\mathrm{e}} \cdot \mathrm{e}^{\lambda_1} + \frac{1+0,5\cos t}{30\mathrm{e}} \cdot \mathrm{e}^{5\lambda_1} + \frac{t}{k} \cdot \mathrm{e}^{-t^2+\lambda_1 t},$$

$$\lambda_2 \geqslant 0,001 + \frac{1+0,5\sin t}{35\mathrm{e}} \cdot \mathrm{e}^{\lambda_2} + \frac{1+0,5\cos t}{30\mathrm{e}} \cdot \mathrm{e}^{5\lambda_2} + \frac{t}{k} \cdot \mathrm{e}^{-t^2+\lambda_2 t}.$$

Then, according to theorem 1, there will be guaranteed the existence of a global solution $y(t)$ of (18) such that $y(0) = k$ and inequalities (12) and (13) hold.

Upper and lower estimates for addends can be found in following forms

$$\frac{1+0,5\sin t}{35\mathrm{e}}\cdot\mathrm{e}^{\lambda_1}+\frac{1+0,5\cos t}{30\mathrm{e}}\cdot\mathrm{e}^{5\lambda_1}\geqslant\frac{2}{70\mathrm{e}}\cdot\mathrm{e}^{\lambda_1}\geqslant\lambda_1,$$

$$\frac{1+0,5\sin t}{35\mathrm{e}}\cdot\mathrm{e}^{\lambda_2}+\frac{1+0,5\cos t}{30\mathrm{e}}\cdot\mathrm{e}^{5\lambda_2}\leqslant\frac{2}{20\mathrm{e}}\cdot\mathrm{e}^{5\lambda_2}\leqslant\lambda_2.$$

By the Wolfram Alpha software, available online on webpage [3], we can verify that transcendental equation $\lambda=2/(70\mathrm{e})\cdot\mathrm{e}^{\lambda}$ has a root $\lambda^*\cong 0,0106$ and $\lambda=2/(20\mathrm{e})\cdot\mathrm{e}^{5\lambda}$ has a root $\lambda^{**}\cong 0,5357$. (For more information about the existence of the roots of transcendental equations see [2].) Let us set $\lambda_1=0,01$ and $\lambda_2=0,5$. It is true, that $\lambda_1\leqslant\lambda_2$ for $t\in\mathbb{R}$.

By finding extremal points of the function

$$f(t)=t\mathrm{e}^{-t^2+\lambda_i t},\quad i=1,2$$

can be simply shown that $|f(t)|<1$ for every $t\in\mathbb{R}$.

Therefore inequalities

$$0,01\leqslant 0,001+\frac{1+0,5\sin t}{35\mathrm{e}}\cdot\mathrm{e}^{0,01}+\frac{1+0,5\cos t}{30\mathrm{e}}\cdot\mathrm{e}^{0,05}+\frac{t}{k}\cdot\mathrm{e}^{-t^2+0,01t},$$

$$0,5\geqslant 0,001+\frac{1+0,5\sin t}{35\mathrm{e}}\cdot\mathrm{e}^{0,5}+\frac{1+0,5\cos t}{30\mathrm{e}}\cdot\mathrm{e}^{2,5}+\frac{t}{k}\cdot\mathrm{e}^{-t^2+0,5t}.$$

hold for $k\geqslant 500$, theorem 1 holds and the global solution $y(t)$ exists and satisfies the inequalities

$$k\mathrm{e}^{-0,01t}\leqslant y(t)\leqslant k\mathrm{e}^{-0,5t}\quad\text{for}\quad t<0,$$

$$k\mathrm{e}^{-0,5t}\leqslant y(t)\leqslant k\mathrm{e}^{-0,01t}\quad\text{for}\quad t>0.$$

4 CONCLUSION

In this paper we have given the sufficient conditions for existence of global solutions of equation (1). In addition, two-sided inequalities on $\mathbb{R}$ for global solutions are given.

ACKNOWLEDGEMENT

The author was supported by the Grant FEKT-S-14-2200 of Faculty of Electrical Engineering and Communication, BUT.

REFERENCES

[1] Diblík, J., Koksch, N.: Sufficient conditions for the existence of global solutions of delayed differential equations, J. Math. Anal. Appl., vol. 318, no. 1, pp. 611-625, 2006.

[2] Diblík, J., Kúdelčíková, M.: Nonoscillating solutions of the equation $x'(t)=-(a+b/t)x(t-\delta)$. Stud. Univ. Žilina Math. Ser., 15, pp. 11–24, 2002.

[3] Wolfram Alpha (2016) [online software]. Retrieved from <https://wolframalpha.com>.

LOCAL ELECTRON-SAMPLE INTERACTION DURING SCANNING ELECTRON MICROSCOPY ON ORGANIC AND METALLIC OBJECTS

Pavel Kaspar

Doctoral Degree Programme (3), FEEC BUT

E-mail: xkaspa25@stud.feec.vutbr.cz

Supervised by: Pavel Tománek

E-mail: tomanek@feec.vutbr.cz

Abstract: To ascertain how high energy of an electron is needed to acquire a sufficient data yield from organic and metallic sample, a Monte Carlo algorithm is used to compare the behaviour of electrons after contact with the material. Primary electron trajectory, elastic and inelastic scattering and secondary electron generation are described in this paper.

Keywords: SEM, Monte Carlo, Electron scattering, Secondary electron

1. INTRODUCTION

Scanning electron microscopy (SEM) is a method widely used on conductive anorganic materials. Because organic samples do not possess the stability and the ability to carry electric charge, displaying them via SEM requires several preparation steps in order for the method to work, including, but not limiting to, alcohol drying and coating the sample with metal (often gold) or freeze-etching. Because organic materials are prone to degradation due to exposure to high energy particles, in this case electrons, it is critical that the samples are measured with the lowest possible energy, while still getting sufficient amount of information about the object. For the purposes of finding out the differences between electron behaviour with different energies and propagation within, an organic sample is compared with a metallic one.

2. MATERIAL PROPERTIES

The organic sample used in simulation is considered to be muscle tissue dried with set of ethanol solutions of increasing concentration. After such preparation, there is very little water present in the sample [1], if the sample is handled properly, and can be completely neglected. For the purposes of the simulation itself, a muscle tissue sample prepared this way can be substituted by carbon, as it makes up the vast majority of the dried tissues. The remaining compounds within the chemical structure of tissue can be neglected in the case of electron-tissue interaction, because their presence does not make the tissue conductive and their concentration does not influence the density of the material. Majority of the interactions between electrons and the sample will therefore be with carbon atoms.

Gold is often the metal used for coating samples for scanning electron microscopy, and it was thus selected to act as the metallic sample. Preparation of the gold sample is not an issue, since obtaining relatively pure pieces of gold is, compared to the preparation of a biological sample, almost effortless, and can be purchased for the coating purposes from any company supplying electron microscopy materials.

Coating film for SEM is usually around 2-20 nm thick [2], but for better interpretation of the results, the thickness of samples from both materials is considered to be around 100 nm.

2.1. CRUCIAL PAREMETERS FOR SIMULATION

Atomic weight and number of the material play a pivotal role in the calculation of electron scattering, together with material density, and are invariant to the changing of the incident electron energy.

The electron energy variant parameters are elastic and inelastic mean free path, the average distance the particle travels before elastic or inelastic scattering. It is possible to calculate the former by using the equation:

$$\lambda_e = \frac{A}{Na * \rho * \sigma_e}, \tag{1}$$

where A is the atomic weight of the material, Na is Avogadro constant, ρ is material density and σ_e is elastic scattering cross-section. The inelastic mean free path is usually only measured, and cannot be calculated precisely enough. There are databases of inelastic mean free paths of materials for different electron energies and articles regarding calculation and the specific values [3].

	Organic sample (carbon)	Metallic sample (gold)
Atomic weight A	12	197
Material density ρ [kg/m^3]	1230	19320
Inelastic mean free path λ_i [Å] at 15 keV	171.3	116.8
Inelastic mean free path λ_i [Å] at 1500 eV	24.8	18.0
Inelastic mean free path λ_i [Å] at 150 eV	5.1	4.6

Table 1: Material parameters used in the simulation.

3. MONTE CARLO SIMULATION

Standard Monte Carlo method is used to calculate the trajectory and energy levels of the electrons in the simulation. Behaviour patterns of each electron are calculated separately, including energy loss of the primary electron and starting energy and the angle of the secondary electron in the case of inelastic scattering, The calculation itself somewhat simplified, compared to other versions with more electron-matter interaction possibilities [4], but it is suitable for the purpose of data yield, stability and energy level determination.

3.1. ALGORITHM STRUCTURE

Initial coordinates for each electron are set with three dimensional vector in the beginning of the axis – the [0, 0, 0] coordinates, with direction vector in the z axis to represent the incident primary electron beam being perpendicular to the sample. After loading in and calculation of the material parameters and the initial energy of the electrons, the algorithm enters the loop phase, which performs the same calculation for each iteration, gradually moving and adjusting the electron properties along the way.

In each step, the first thing to determine is whether the last step moved the electron outside of the allowed area, and if the energy of the electron is below the set energy threshold, and if so, terminate the loop. After that the scattering itself comes. From the inelastic and elastic mean free paths a frequency of scattering events is determined, and which scattering will happen. If the elastic scattering is selected, a scattering angle is calculated, and the electron is moved along its new direction vector to a new position given by the equation 2, beginning the whole process anew.

$$\begin{matrix} x_{n+1} \\ y_{n+1} \\ z_{n+1} \end{matrix}' = \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} + s_n \begin{pmatrix} \sin\theta_n \cos\phi_n \\ \sin\theta_n \sin\phi_n \\ \cos\theta_n \end{pmatrix}, \tag{2}$$

where θ is the scattering angle, ϕ is the azimuthal angle and n is the number of the steps being performed.

In the case of the inelastic scattering, the calculation is split into three phases. First, the scattering angle is determined and the electron is moved, in a similar manner as with the elastic scattering, then the energy loss is quantified and the electron data is adjusted. The last part is the creating of secondary electron due to ionization. The angle and the energy are dependant on the scattering angle and the energy loss of the primary electron, and they are stored for further processing. The energy of the primary electron after the scattering is given by the following equation:

$$E' = E * \cos^2\theta, \tag{3}$$

while the energy of the created secondary electron can be established by the equation:

$$E'' = E * \sin^2\theta. \tag{4}$$

After the single primary electron trajectories and energies at each step are established, the databank of secondary electrons from this part of the simulation is treated the same way, as if they were primary electrons as well, for the sake of the calculation. Each secondary electron can create more secondary electrons, and the simulation runs, until every entry in the databank has been cleared. After that a new primary electron can enter the algorithm and the whole process repeats itself, until the set number of primary electrons has been simulated.

3.2. SIMULATION RESULTS

Simulations were run with 100 electrons each, 100 times over to achieve dependable results. From the final count of the data yield, in Table 2., we can see that while the median energy of BSE is for both organic and metallic sample very close to the original energy of the primary electron, the mean energy drops in the metallic sample on 1500 eV in contrast to the dried tissue sample, where the BSE mean is still close to 1500 eV. That shows that while most of the backscattered electrons in the gold object still return very soon and largely unaffected, there are more that undergo several inelastic scatterings on this energy level. This is further supported by Figure 1, where the multiple inelastic scatterings and secondary electron generation are clearly visible. Under lower energies the primary electron shows significantly smaller penetration into the gold sample in comparison to the dried muscle tissue. If we gathered information not only from backscattered primary and secondary electrons, but from the ones that pass the sample as well, gold would be much less useful, if we were to disregard conductivity. While secondary electrons help with the material analysis, they carry very little information about topography, due to the level of their energy and the directions they travel. Dried muscle tissue can still provide topographical data, but the secondary electron yield is considerably smaller than in metallic or gold coated samples.

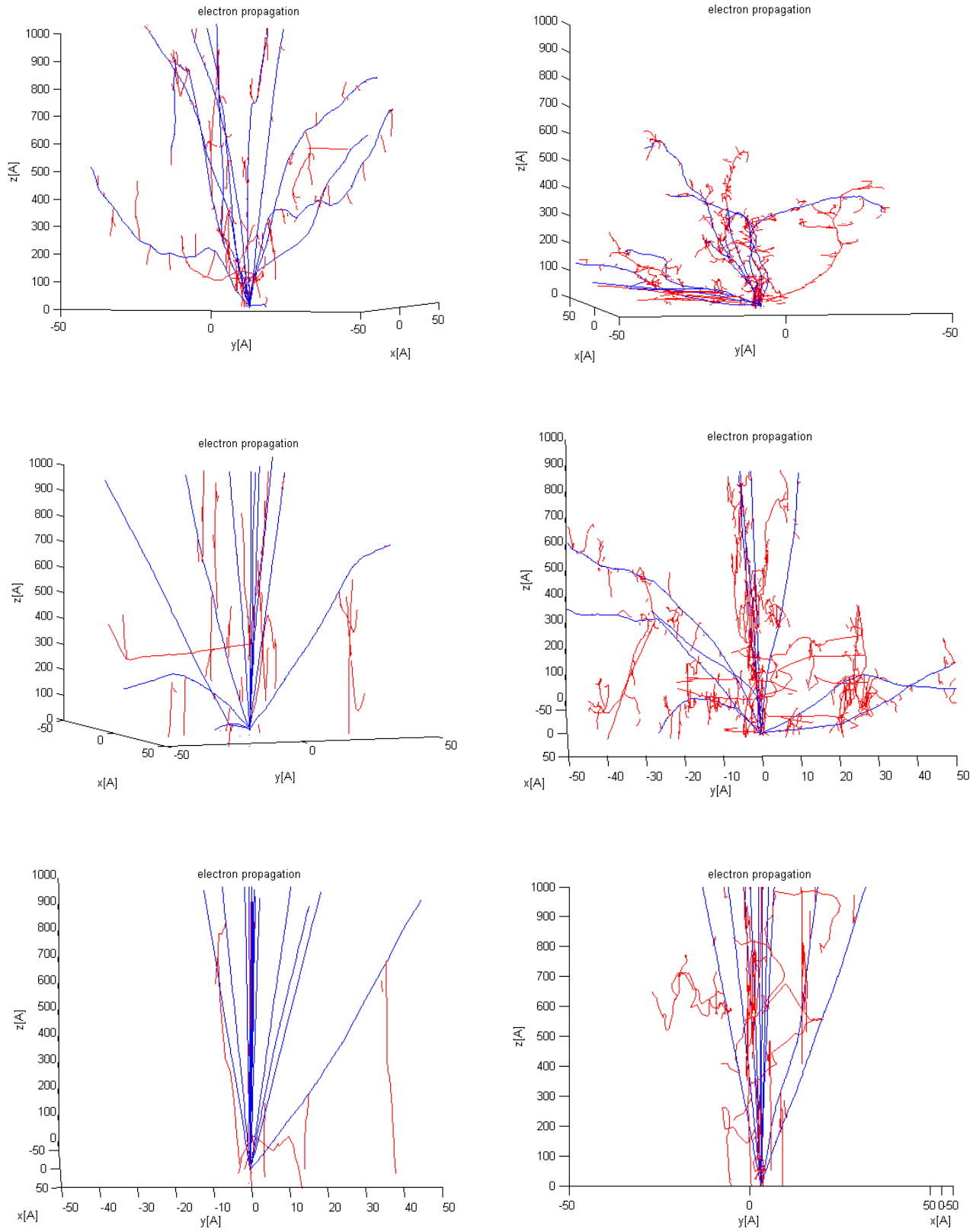


Figure 1: Comparison of propagation of 20 electrons in dried organic tissue (left) and gold (right) with electron energy of 15 keV (first row), 1500 eV (second row) and 150 eV (third row). Primary electrons are depicted in blue, while secondary are in red colour.

	Organic 15 keV	Organic 1500 eV	Organic 150 eV	Metallic 15 keV	Metallic 1500 eV	Metallic 150 eV
Backscattered electrons BSE [%]	48 ± 11	59 ± 13	63 ± 17	48 ± 7	56 ± 10	53 ± 12
Mean energy of BSE [eV]	15000 ± 0.10	1499.97 ± 1.46	141.80 ± 26.82	15000 ± 0.11	1428.90 ± 273.72	121.33 ± 25.65
Median BSE energy [eV]	15000	1500	149.92	15000	1499.83	129.86
Backscattered secondary electrons BSSE	17 ± 5	28 ± 8	22 ± 21	36 ± 7	75 ± 21	64 ± 12
Mean energy of BSSE [eV]	6.34 ± 6.99	5.44 ± 7.92	3.45 ± 6.02	7.44 ± 12.31	11.34 ± 45.97	3.04 ± 4.15
Median energy of BSSE [eV]	3.40	2.23	1.38	2.40	1.97	1.95

Table 2: Data yield on detector from 100 simulations with 100 electrons

4. CONCLUSION

The Monte Carlo model created for this paper proved stable for simulating electron propagation in dried muscle tissue and gold samples. The results demonstrated that in the gold sample far more secondary electrons are created, but cannot be successfully detected due to the speed of their energy dissipation, thus the energy yield of backscattered electrons from the gold sample drops faster with lower energy of the primary electron beam than in the organic sample. Any water molecules in the dried muscle tissue are neglected for the purposes of this simulation.

ACKNOWLEDGEMENT

This work was supported by the Internal Grant Agency of Brno University of Technology, grant No. FEKT-S-14-2240.

REFERENCES

[1] Bozzola, J. J., Russell, L. D.: Electron Microscopy: Principles and Techniques for Biologists, Jones and Bartlett Publishers, Inc. 1992, ISBN 0-7637-0192-0

[2] Goldstein, J. I., Newbury, D. E., Echlin, P., Joy, D. C., Romig, A. D., Lyman, C. E., Fiori, C., Lifshin, E.: Scanning Electron Microscopy and X-Ray Microanalysis, New York, Plenum Press 1992, ISBN13: 978-1-4612-7653-1

[3] Powell, C. J:, Jablonski, A.: Evaluation of Calculated and Measured Electron Inelastic Mean Free Paths Near Solid Surfaces, J. Phys. Chem., 1999, doi 10.1063/1.

[4] Shimizu, R., Ze-Jun, D.: Monte Carlo modeling of electron-solid interactions, Rep. Prog. Phys., Printed in the UK 1992, s. 487-531

MICROSTRUCTURAL DEFECTS OF SOLAR CELLS INVESTIGATED BY A VARIETY DIAGNOSTIC METHODS

Ľubomír Škvarenina
Doctoral Degree Programme (2), FEEC BUT
E-mail: skvarenina@phd.feec.vutbr.cz

Supervised by: Robert Macků
E-mail: macku@feec.vutbr.cz

Abstract: This paper discusses the application of a variety diagnostic methods applicable to the solar cells. More objective results about solar cells quality and reliability are possible to obtain by using a various methods. Diagnostic methods described in this paper are based on a dark and illuminated *J–V* characteristics, a investigation of noise in a wide range of frequency and a radiation detection at a different spectral range, namely by an electroluminescence and a thermography method. These methods are primarily more appropriate for a detection or a localization of microstructure defects when a reverse-bias stress is applied. However, the analysis of a forward-bias conditions is included in an investigation of *J–V* characteristics as well.

Keywords: Solar Cells, *J–V* curve, Noise, Electroluminescence, Thermal Imaging, Defects

1. INTRODUCTION

Although the research and development nowadays are more directed to a thin-film solar cells the majority of a generating production today is still oriented on a silicon wafer-based photovoltaic technology. Despite this mass industrial production success, the manufacturing techniques are largely influenced by the requirements for a very low cost of solar cells that gives a space to creation of the microstructural defects caused by the technological imperfections. Our research is aimed precisely to a detection or a localization of the microstructural defects along with a investigation of their nature. These defects have a significant adverse impact at a required properties of the solar cells leading to a reduction of their durability and efficiency. The solar cell samples used in this investigation are not long term exposed to an ordinary operation, which can be responsible for an emergence of new microscopic cracks and fractures, so therefore the observed defects have the origin only in the process of their production. The solar cells are made from single-crystal silicon of dimensions 125×125 mm (150 mm in diagonal with a surface area of 150 cm^2) and a thickness of 220 µm. The cells are designed for the solar panel fabrication. The samples are usually prepared from the monocrystalline silicon solar cell wafer by breaking the wafer into pieces. The discussion of results is performed for a selected interesting solar cell sample labelled as a 'X'.

2. EXPERIMENTAL

2.1. Current density-voltage characteristics

The dark current density-voltage (*J–V*) characteristics of the solar cells are measured under a reverse and a forward-bias conditions to provide the basic properties about the samples. The samples are an electrically stressed between two aluminium electrodes, where the bottom electrode is equipped with a thermoelectric cooler controlled by the source-meter Keithley 2510 TEC for a temperature stabilization. The *J–V* characteristics are measured by the precision source-meter Keithley 2420 set to the voltage source with a current meter. The measurement of an illuminated *J–V* characteristics is based

on a modular solution from the National Instruments company. The illuminated *J–V* characteristics are measured by the programmable high-power source measure unit card PXI-4130 in the mode of a current sink and voltmeter. The illumination of the samples is realized by a halogen lamp controlled by the data acquisition card PXI-6224. Both of these cards are installed in the hybrid slots of the chassis PXIe-1073 with a integrated controller.

2.2. Electroluminescence and thermal imaging

A low-noise CCD camera is used for a sensing of radiation at the visible and near-infrared spectral range from 0.3 μm to 1.5 μm and a thermal imager is used for the detection of radiation at a far-infrared spectral range from 7 μm to 13 μm. The CCD camera is equipped with a silicon chip with a resolution of 2184×1472 pixels cooled by a dual system of Peltier modules. The imaging system of the IR camera is based on an uncooled focal plane array detector with a resolution of 120×160 pixels. The samples are placed on the aluminium electrode with a thin finger contact equipped with a wrought silver in both of these measurements. The reverse-bias current is realized by the precision current source Keithley 6220 DC.

2.3. Power spectral density characteristics

The samples are placed in the dark environment between two conductive aluminium electrodes for the noise measurements. Two parallel capacitors (a ceramic $C_1 = 1\,\mu\text{F}$ and an electrolytic $C_2 = 470\,\mu\text{F}$) are on the input of the circuit. These capacitors are used for the voltage ripple stabilization and the excess noise elimination of a voltage source Agilent E3631A. The voltage source is set to the current control mode to keep the current density through material constant. The reverse current is set from the 0.5 mA to 3 mA in an increments of 0.5 mA. The noise of the samples is sensed on the output closed by a resistor ($R = 9.966\,\Omega$). The noise signal sensed on the resistor is amplified by a ultra low-noise amplifier Signal Recovery 5184 with a fixed voltage gain of ×1000 (60 dB). The noise is applied to an unbalanced Q input of a baseband analyser Rohde & Schwarz FMU36 with the selected input impedance of 1 MΩ. The noise is finally investigated in the spectral range from 10 Hz to 1 MHz.

3. RESULTS

The illuminated *J–V* characteristics measurement under illumination provides the information for obtaining the basic external electrical parameters of the sample such as an operating point for a maximum power (P_{MAX}), a short-circuit current density (J_{SC}) and an open-circuit voltage (V_{OC}). These parameters are illustrated in the fig. 1a with the obtained maximum power point (MPP). This fourth quadrant of the *J–V* dependence is also supplemented by a power density-voltage (*P–V*) dependence. The power density in this case refers to a power generated per unit area of the sample. The fill factor defined by a relation

$$FF = \frac{J_{\text{MP}} \cdot V_{\text{MP}}}{J_{\text{SC}} \cdot V_{\text{OC}}} = \frac{P_{MAX}}{J_{\text{SC}} \cdot V_{\text{OC}}}, \tag{1}$$

where the J_{MP} current density and the V_{MP} voltage produce the maximal power, refers the ratio of the actual maximum obtained power to the open-circuit voltage and the shot-circuit current [1]. The open-circuit voltage of 0.55 V was measured at open-circuit conditions where the current is equal to a zero. Also the short-circuit current density of $9.15\,\text{mA}\,\text{cm}^{-2}$ was measured at the voltage equal to a zero. The reached fill factor is 0.51 for this obtained parameters by the eq. 1 including the maximum power density of $2.57\,\text{mW}\,\text{cm}^{-2}$. The typical fill factors of the contemporary solar cells are between 0.75 and 0.80 [1]. The inversion of obtained slopes in fig. 1a permits an easy determination of a shunt resistance (R_{SH}) and a series resistance (R_{S}). The origin of the series resistance can be in the current moving through the semiconductor base and an emitter of the *pn* junction, in the top and rear metal contacts and in the interface between the metal contacts and the semiconductor material [2].

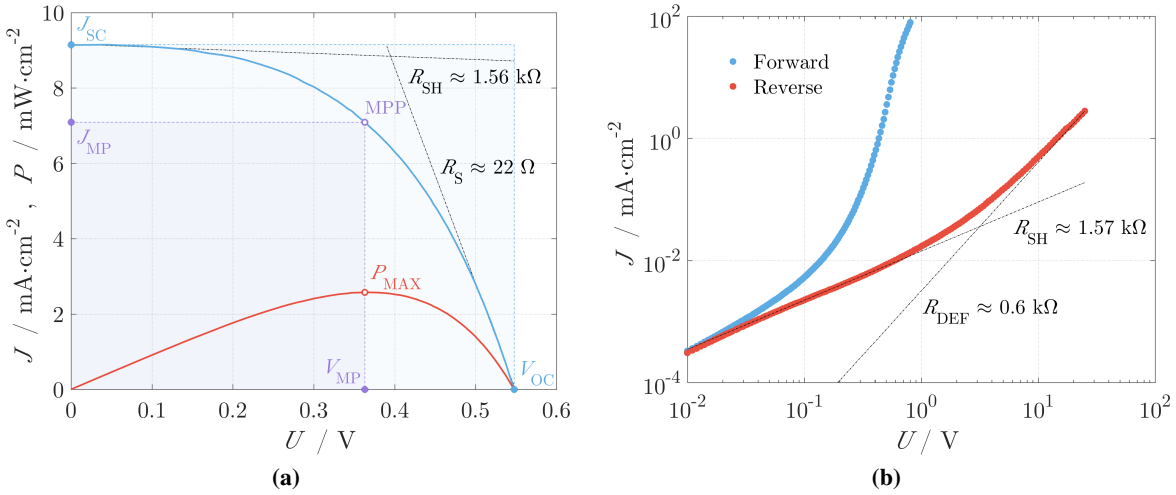


Figure 1: a) The illuminated *J*–*V* characteristic with the obtained *P*–*V* curve, $J_{MP} = 7.09\,\mathrm{mA\,cm^{-2}}$, $V_{MP} = 0.36\,\mathrm{V}$, $P_{MAX} = 2.57\,\mathrm{mW\,cm^{-2}}$. b) The dark *J*-*V* characteristic for the forward and reverse-bias. Solar cell sample X, maintenance temperature $T = 26\,^{\circ}\mathrm{C}$.

The series resistance estimated from the slope is about 22 Ω. The shunt resistance of the illuminated *J*–*V* characteristics is estimated about 1.56 kΩ. It is caused by a microscopic defects which provides the alternative path for the generated photo-current crack through the semiconductor layers or a current path at the edges of the solar cell. To converge to the ideal illuminated *pn* junction should be the shunt resistance large as possible and a series resistance small as possible. The theoretical derivation of these parameters based on the equivalent circuit is beyond the capacity of this paper. The forward-bias curve in the fig. 1b of dark *J*–*V* characteristic shifted by the short-circuit photo-current is actually the illuminated *J*–*V* characteristic in the fig. 1a [3]. The differences between their mutual fitting are probably caused by a heating produced by the halogen lamp [4] and much more robust fitting algorithm for fig. 1b. The dark reverse-bias *J*–*V* characteristic in fig. 1b is more suitable for the detection of the microstructural defects. The estimated shunt resistance in fig. 1b is a nearly identical to the shunt resistance estimated in fig. 1a. The gradual change of the slope at 3 V is caused by the electric field strength activation of a defect with estimated resistance (R_{DEF}) of 0.6 kΩ. The activated defect with a reduced resistance caused the increasing of the reverse current. Basically this may not be only one activated microstructural defect, but a whole range of some parallel microstructural defects activated at the similar time with the resistance of R_{DEF}. However, the electroluminescence investigation with the CCD camera below shows that the only one particular microstructure defect is responsible for the increasing of a reverse current in fig. 1b.

The measurement of electroluminescence and the thermal radiation is provided for the different reverse-bias current of 1 mA, 3 mA and 6 mA. The investigation of the electroluminescence is based on a comparison of the obtained images at the different reverse-bias current. The radiation intensity of the sample is expressed by a logarithm of photons per second (a photon flux Φ_q). The intensity difference of the images at 3 mA and 1 mA shows a presence of the only one defect with a strong intensity near to the edge of sample. Thus, the defect at a lower reverse-bias current is not accompanied by the radiation. This microstructural defect is marked by the dashed-line circle in fig 2a. We assume that, the resistance of this microstructural defect begin to decrease by a higher reverse-bias voltage and causes a steeper increase of the total reverse current (red line) in fig. 1b. The conductive channel is created and the difference of the resistances is almost 1 kΩ. The intensity difference of images at 6 mA and 3 mA shows emergence of two new microstructural defects marked by the dotted-line hexagons in fig. 2a. The negligibly small break disrupting the smoothness of the curve is in the dark *J*–*V* characteristics in fig. 1b at the reverse-bias voltage of 18 V. This can be attributed to the

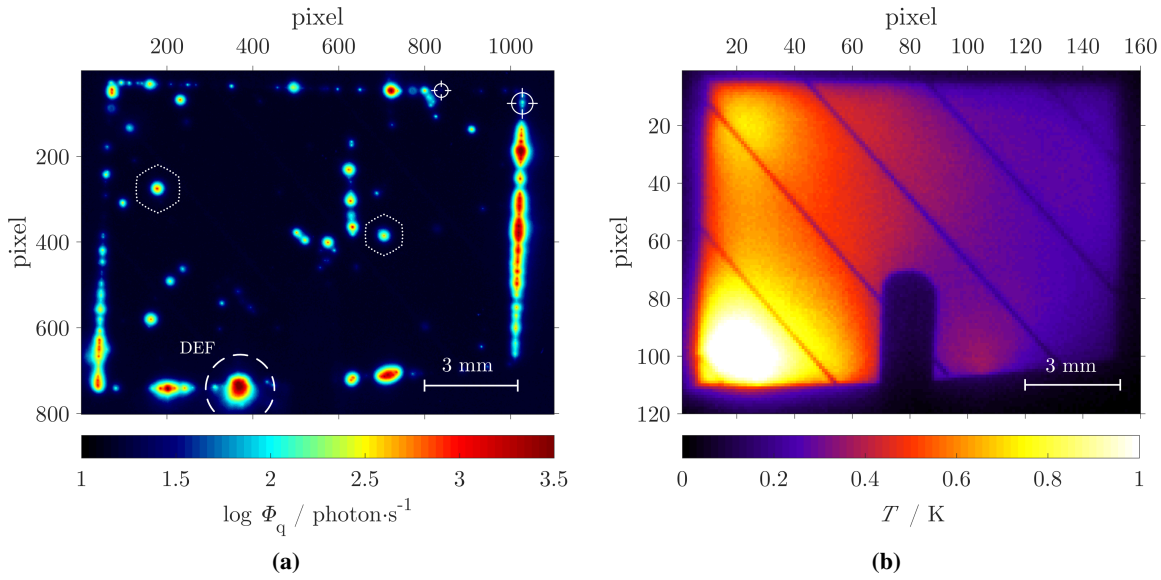


Figure 2: a) The electroluminescence of the solar cell fragment with light defect spots, exposure time 600 s, dark environment. b) The thermal activity of the solar cell fragment. Solar cell sample X, reverse-bias current $I_R = 6\,\text{mA}$, ambient temperature $T = 25\,°\text{C}$, area $126.33\,\text{mm}^2$, perimeter 45 mm.

extinction of an activated microstructural defect due to a small decrease of the reverse-bias current. This behaviour is also observed in fig 2a of the electroluminescence measurement as a decreasing the intensity of two microstructural defects marked by the cross-hairs. We assume these conductive channels caused by the microstructural defects on the left partially disappear and the total resistance of sample increased slightly. The investigation under a different reverse-bias current is also focused on the thermal imaging. The thermal radiation of the sample is shown in fig. 2b. The considerable overheating is observed on the left of the sample with the more significant defective area in the lower left corner comparing to the upper area. This thermal radiation on both cases can not be associated to an any particular light spot of the microstructural defects in fig. 2a. The emergence of the new defects or an extinction of already presented defects is not observed at different reverse-bias current and an increase of the reverse-bias current causes only the temperature increase of the overheating.

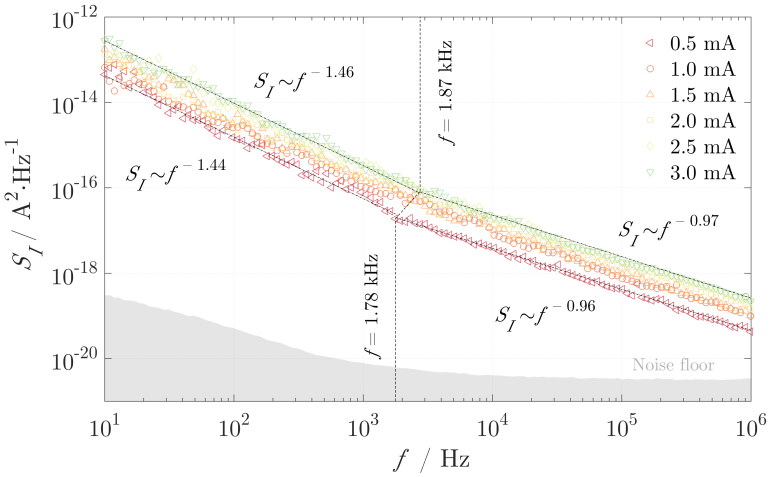


Figure 3: The noise power spectral densities of the current fluctuations for the different bias current, solar cell sample X, dark environment, night-time conditions.

The noise power spectral density (PSD) in fig. 3 shows the inversely proportional dependence on the frequency that correspond to the type of noise commonly know as a flicker noise. The f^{-1} noise interpretation is given by Hooge's formula in the principal form $\frac{S_I(f)}{I^2} = \frac{\alpha_H}{fN}$. There is clearly defined that PSD is inversely proportional to frequency but there exists number of statements that have to be valid. Firstly, the f^{-1} noise originates from fluctuation in the mobility. Secondly, it is believed that carriers provide mutual independent fluctuation. Parameter N is preferable substituted by nV, where n is carrier concentration and V volume [5]. We expect that changes in the slope in fig. 3 could be given by defined statement infraction or voltage dependent changes in the equation parameters.

4. CONLUSION

A variety of the diagnostic methods to obtain the information about the microstructural defects are described with a demonstration in figures. The illuminated J–V characteristic provides the basic electrical parameters under the ordinary operation responsible for the behaviour of the solar cell. The qualitative information about the sample provides also the fill factor which is to a large extent influenced by the shunt and the series resistance. On the other hand, the dark J–V characteristics provides the basis for the estimation of the microstructural defects, which can not be detected with the diagnostics under the normal operation, especially in the reverse-bias. The activated microstructural defect with the estimated resistance R_{DEF} began to affect the dark J–V curve which consequently caused an increase of the reverse current. The measuring of the electroluminescence shows that the decrease of the resistance is mainly caused by the only one particular microstructure defect near to the edge of the sample. The radiation detection by the electroluminescence and the thermal imaging in the different spectral range shows that, the radiation of the microstructural defects in a visible and near infrared range is not also accompanied by the thermal overheating in these affected areas. The investigation of the power noise spectral density $S_I(f)$ shows dependence of f^{-1} noise with changes in the slope. A precisely examination of this changes will be provided in the future. Future research will be focused on the use of these methods for a estimation quality of thin film solar cells, especially for evaluation of the microstructural defects in $\mathrm{Cu(In,Ga)Se_2}$ solar cells. We have a lot of preliminary results now and it will be subject of the future publication.

ACKNOWLEDGEMENT

This research was supported by the Internal Grant Agency of Brno University of Technology, grant No. FEKT-S-14-2240.

REFERENCES

[1] Rauschenbach, H.: Solar Cell Array Design Handbook: The principles and Technology of Photovoltaic Energy Conversion, New York: Van Nostrand Reinhold Co., 1980.

[2] Haschke, J., Mingirulli, N. and Rech, B.: Progress in Point Contacted Rear Silicon Heterojunction Solar Cells, Energy Procedia, vol. 27, 2012, pages 116–121.

[3] Lindholm, A. F.: Application of the Superposition Principle to Solar-Cell Analysis. IEEE Transaction on Electron Devices, vol. 26, no. 3, 1979.

[4] Mártil, I. and Diaz, G.: Determination of the dark and illuminated characteristic parameters of a solar cell from I-V characteristics, Eur. J. Phys, vol. 13, 1992, pages 193–197.

[5] Kiss, L.B. and Kleinpenning, T.G.M: Interpretation of Hooge's $1/f$ formula , Physica B+C, vol. 145, 1987, pages 185–189.

OPTIMIZATION OF LINEAR DIFFERENTIAL SYSTEMS WITH A DELAY BY LYAPUNOV'S DIRECT METHOD

Hanna Demchenko
Doctoral Degree Programme (2. ročník), FEEC BUT
E-mail: xdemch02@stud.feec.vutbr.cz

Supervised by: Josef Diblík
E-mail: diblik@feec.vutbr.cz

Abstract: In the paper linear equations with delays

$$\frac{dx(t)}{dt} = ax(t) + bx(t-\tau) + dx(t-\delta) + cu(t),\ t \geq t_0,$$

(where a, b, c, d are real constants, $\tau > 0$ and $\delta > 0$ are delays, $t_0 \in \mathbb{R}$ and $u(t)$ is a control function) and their systems are considered.

Using Lyapunov's direct method, control functions minimizing quality criteria are constructed and an illustrative example is given.

Keywords: Lyapunov function, optimization problem, integral quality criterion, control function

1 INTRODUCTION

Differential equations are the strong tool for modelling and solving a lot of engineering, mechanical, economic or population problems. But it is well-known that in such problems a delay (e.g. between input in an electrical circuit and its response) arises quite naturally. In general, there is always a delay in the real life processes. So, delay differential equations are the important field of research. As the systems with feedback can be described (under certain conditions) by systems of differential equations with a delay or by difference equations, a wide field of applications is opened.

In this contribution, we will investigate stabilization problem for differential equations with a delay and their systems. We will look for a control function, which will satisfy all desired conditions, like the best possible quality of a transition process and a minimum value of a quality criterion. We describe the general scheme for construction of an optimal control by the second Lyapunov's method (see, e.g. [1, p. 476]).

Consider an arbitrary dynamic process, which can be controlled by means of a function $u = u(t)$ and assume that it can be represented by a system

$$\frac{dx(t)}{dt} = f\left(x(t), x_t, u(t)\right), t \geq t_0, \tag{1}$$

where $f\colon \mathfrak{D}_1 \longrightarrow \mathbb{R}^n$ and $n \geq 1$, $n \in \mathbb{N}$, $t_0 \in \mathbb{R}$, $m \in \mathbb{N}$,

$$\mathfrak{D}_1 := \{(x,y,u) \in \mathbb{R}^n \times C_\tau^n \times \mathbb{R}^m,\ \|x\| < M,\ \|y\|_\tau < M\},$$

where for given $t \in [t_0, \infty)$ we define

$$\|x(t)\|_\tau = \max_{\theta \in [-\tau, 0]} (\|x(t+\theta)\|),\ \ \|x(s)\| = \max_{i=1,\ldots,n} \{|x_i(s)|\},\ s \in [t_0 - \tau, \infty)$$

and $C_\tau^n = C([-\tau,0),\mathbb{R}^n)$ is the space of the continuous mappings from the interval $[-\tau,0)$ into $\mathbb{R}^n$. For each $t \geq t_0$ we define $x_t \in C_\tau^n$ by $x_t(\theta) = x(t+\theta)$.

We will assume that f satisfies following assumptions:

1. $f(\theta_n,\theta_n,\theta_m) = \theta_n$, where θ_n, θ_m are n-dimensional or m-dimensional zero vectors.
2. $f(x,y,u)$ satisfies the local Lipschitz condition in every bounded neighborhood of each point $(x,y,u) \in \mathfrak{D}_1$.

The problem under consideration is following. We need to find a function $u = u(t)$ (called a control function) such that the zero solution $x(t) = \theta_n$, $t \geq t_0$ of the system (1) will be asymptotically stable and for an arbitrary solution $x = x(t)$, $t \geqslant t_0 - \tau$ of the system (1) satisfying $\|x\| < M$, the integral

$$I = \int_{t_0}^{\infty} \omega(x(t), x_t, u(t))\, dt \tag{2}$$

exists and attains minimum value. Here $\omega\colon \mathfrak{D}_1 \to \mathbb{R}$ is a positive definite functional and $t \geq t_0$.

Define an auxiliary functional $B\colon \mathbb{R} \times [t_0,\infty) \times \mathbb{R}^n \times C_\tau^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$,

$$B(V,t,x(t),x_t,u) := \frac{dV(t,x_t)}{dt} + \omega(x(t),x_t,u(t)) \tag{3}$$

where V is a Lyapunov-Krasovskii functional and $dV(t,x_t)/dt$ denotes the derivative of V with respect to t along trajectories of the system (1). The following theorem was motivated by a similar theorem for non-delay systems (see [1], Theorem IV).

Theorem 1. Assume that, for the system of differential equations (1), there exists a Lyapunov-Krasovskii functional $V_0(t,x_t)$ and a function $u_0(t)$ such that

1. The function $\omega(x(t),x_t,u_0(t))$ is positive-definite for every $t \geq t_0$, $\|x\| < M$, $\|x_t\|_\tau < M$.
2. $B(V_0,t,x(t),x_t,u_0(t)) \equiv 0$.
3. $B(V_0,t,x(t),x_t,u(t)) \geq 0$ for any $u(t) \not\equiv u_0(t)$.

Then, the function $u_0(t)$ is a solution of the optimal stabilization problem and

$$\int_{t_0}^{\infty} \omega(x(t),x_t,u_0(t))dt = \min_{u(t)} \left[\int_{t_0}^{\infty} \omega(x(t),x_t,u(t))dt \right] = V_0(t_0,x_{t_0}).$$

2 EQUATIONS WITH TWO DELAYS

Consider linear scalar equations with constant coefficients and two delays

$$\frac{dx(t)}{dt} = ax(t) + bx(t-\tau) + dx(t-\delta) + cu(t),\ t \geq 0, \tag{4}$$

where a, b, c, d are real constants, $\tau > 0$ and $\delta > 0$ are delays and $u(t)$ is a control function.

Together with the equation (4) we will consider a quality criterion (2) with $t_0 = 0$ and

$$\omega(x(t),x_t,u(t)) = \alpha x^2(t) + 2\beta x(t)x(t-\tau) +$$

$$+\gamma x^2(t-\tau) + \varepsilon x^2(t-\delta) + 2\xi x(t)x(t-\delta) + Du^2(t),$$

i.e., (2) being a quadratic criterion

$$I=\int_0^{\infty}[\alpha x^2(t)+2\beta x(t)x(t-\tau)+\gamma x^2(t-\tau)+\varepsilon x^2(t-\delta)+2\xi x(t)x(t-\delta)+Du^2(t)]dt, \tag{5}$$

with $\alpha>0$, $\alpha\gamma-\beta^2>0$, $\alpha\gamma\varepsilon-\gamma\xi^2-\varepsilon\beta^2>0$, $D>0$.

Theorem 2. If

$$\beta b<0,\ \xi d<0 \tag{6}$$

and

$$b(\alpha+\gamma+\varepsilon)=2a\beta,\ d(\alpha+\gamma+\varepsilon)=2a\xi, \tag{7}$$

then an optimal stabilization control function $u_0(t)$ exists and

$$u_0(t)=2\frac{\beta c}{bD}x(t). \tag{8}$$

Proof. We utilize Theorem 1. Define a Lyapunov-Krasovskii functional

$$V(t,x_t)=hx^2(t)+d_1\int_{t-\tau}^{t}x^2(s)ds+d_2\int_{t-\delta}^{t}x^2(s)ds,\ h>0,\ d_1>0,\ d_2>0.$$

Then, in accordance with the condition 2 of Theorem 1 we analyse the expression B given by (3), i.e.,

$$\begin{gathered}B(V,t,x(t),x_t,u(t))=2hx(t)[ax(t)+bx(t-\tau)+dx(t-\delta)+cu(t)]+\\+d_1[x^2(t)-x^2(t-\tau)]+d_2[x^2(t)-x^2(t-\delta)]+\alpha x^2(t)+2\beta x(t)x(t-\tau)\\+\gamma x^2(t-\tau)+\varepsilon x^2(t-\delta)+2\xi x(t)x(t-\delta)+Du^2(t)=0.\end{gathered}$$

Simplifying the last expression, we get

$$\begin{gathered}B(V,t,x(t),x_t,u)=(2ha+d_1+d_2+\alpha)x^2(t)+(\gamma-d_1)x^2(t-\tau)+(\varepsilon-d_2)x^2(t-\delta)+\\+(2hb+2\beta)x(t)x(t-\tau)+(2hd+2\xi)x(t)x(t-\delta)+2hcx(t)u(t)+Du^2(t)=0.\end{gathered}$$

This equation will be satisfied if

$$2ha+d_1+d_2=-\alpha, \tag{9}$$

$$d_1=\gamma, \tag{10}$$

$$d_2=\varepsilon, \tag{11}$$

$$h=-\beta/b=-\xi/d, \tag{12}$$

$$2hcx(t)u(t)=-Du^2(t). \tag{13}$$

Condition (12) is valid because of (6). Substituting (10), (11) and (12) into (9), we get condition (7).

From (13), we obtain an optimal control in the form

$$u_0(t)=-2\frac{hc}{D}x(t)=2\frac{\beta c}{bD}x(t). \quad \square$$

Example 1. Consider the equation (4) with $a=-3$, $b=-1$, $d=-1$, $c=1$, i.e.,

$$\dot{x}(t)=-3x(t)-x(t-\tau)-x(t-\delta)+u(t)$$

together with a quadratic quality criterion (5) with $\alpha=\gamma=\varepsilon=2$, $\beta=\xi=D=1$, (here $\alpha>0$, $\alpha\gamma-\beta^2=3>0$, $\alpha\gamma\varepsilon-\gamma\xi^2-\varepsilon\beta^2=4>0$, $D>0$), i.e.,

$$I=\int_0^{\infty}(2x^2(t)+2x(t)x(t-\tau)+2x^2(t-\tau)+2x^2(t-\delta)+2x(t)x(t-\delta)+u^2(t))dt.$$

Since $\beta b=-1<0$, $\xi d=-1<0$ and $b(\alpha+\gamma+\varepsilon)=2a\beta=-6$, $d(\alpha+\gamma+\varepsilon)=2a\xi=-6$, all assumptions of Theorem 2 are true. By the formula (8), the optimal stabilization control function is

$$u_0(x(t))=2\frac{\beta c}{bD}x(t)=-2x(t).$$

3 SYSTEMS WITH TWO DELAYS

Consider linear systems with constant coefficients with two constant delays

$$\frac{dx(t)}{dt} = A_0x(t) + A_1x(t-\tau) + A_2x(t-\delta) + bu(t),$$

where A_0, A_1, A_2 are $n \times n$ constant matrices, $b \in R^n$, $u(t) \in R$, and a quality criterion (2) with $t_0 = 0$ and

$$\omega(x(t), x_t, u(t)) = \begin{pmatrix} x^T(t), & x^T(t-\tau), & x^T(t-\delta) \end{pmatrix} \begin{pmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{pmatrix} \begin{pmatrix} x(t) \\ x(t-\tau) \\ x(t-\delta) \end{pmatrix} + du^2(t),$$

where

$$C = \begin{pmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{pmatrix}$$

is a positive-definite, symmetric matrix, so C_{11} and C_{33} are $n \times n$ positive-definite, symmetric matrices, matrix $C_{11} + C_{22} + C_{33}$ is a positive-definite, $C_{21} = C_{12}^T$, $C_{31} = C_{13}^T$, C_{23} and C_{32} are zero-matrices and $d > 0$, i.e., (2) is a quadratic criterion

$$I = \int_0^\infty (x^T(t)C_{11}x(t) + x^T(t)C_{12}x(t-\tau) + x^T(t)C_{13}x(t-\delta) +$$

$$+ x^T(t-\tau)C_{21}x(t) + x^T(t-\tau)C_{22}x(t-\tau) +$$
$$+ x^T(t-\delta)C_{31}x(t) + x^T(t-\delta)C_{33}x(t-\delta) + du^2(t))dt.$$

Theorem 3. Assume that there exists a positive definite symmetric matrix H satisfying the Lyapunov matrix equation

$$A_0^T H + HA_0 = -C_{11} - C_{22} - C_{33}. \tag{14}$$

If, moreover,

$$A_1^T H = -C_{21}, A_2^T H = -C_{31}, \tag{15}$$

then the optimal stabilization control function exists and

$$u_0(t) = -\frac{2}{d} b^T H x(t).$$

Proof. We utilize Theorem 1. Define a Lyapunov-Krasovskii functional

$$V(t, x_t) = x^T(t)Hx(t) + \int_{t-\tau}^t x^T(s)G_1x(s)ds + \int_{t-\delta}^t x^T(s)G_2x(s)ds,$$

where H, G_1 and G_2 are $n \times n$ positive-definite matrices. Then, in accordance with the condition 2 of Theorem 1 we analyse the expression B given by (3), i.e.,

$$\begin{aligned} B(V, t, x(t), x_t, u) &= [A_0x(t) + A_1x(t-\tau) + A_2x(t-\delta) + bu(t)]^T Hx(t) \\ &+ x^T(t)H[A_0x(t) + A_1x(t-\tau) + A_2x(t-\delta) + bu(t)] + \\ &+ x^T(t)G_1x(t) - x^T(t-\tau)G_1x(t-\tau) + x^T(t)G_2x(t) - x^T(t-\delta)G_2x(t-\delta) \\ &+ x^T(t)C_{11}x(t) + x^T(t)C_{12}x(t-\tau) + x^T(t)C_{13}x(t-\delta) + x^T(t-\tau)C_{21}x(t) \\ &+ x^T(t-\tau)C_{22}x(t-\tau) + x^T(t-\delta)C_{31}x(t) + x^T(t-\delta)C_{33}x(t-\delta) + du^2(t) \equiv 0. \end{aligned}$$

Simplifying the last expression, we get

$$B(V,t,x(t),x_t,u) = x^T(t)[A_0^T H + HA_0 + G_1 + G_2 + C_{11}]x(t) + x^T(t-\tau)[A_1^T H + C_{21}]x(t)$$
$$+ x^T(t)[HA_1 + C_{12}]x(t-\tau) + x^T(t-\delta)[A_2^T H + C_{31}]x(t) + x^T(t)[HA_2 + C_{13}]x(t-\delta) +$$
$$+ x^T(t-\tau)[C_{22} - G_1]x(t-\tau) + x^T(t-\delta)[C_{33} - G_2]x(t-\delta) +$$
$$+ b^T u(t)Hx(t) + x^T(t)Hbu(t) + du^2(t) \equiv 0.$$

This will hold if

$$A_0^T H + HA_0 = -C_{11} - G_1 - G_2, \tag{16}$$

$$A_1^T H = -C_{21}, HA_1 = -C_{12}, \tag{17}$$

$$A_2^T H = -C_{31}, HA_2 = -C_{13}, \tag{18}$$

$$G_1 = C_{22}, G_2 = C_{33}, \tag{19}$$

$$u(t)[b^T Hx(t) + x^T(t)Hb] = -du^2(t). \tag{20}$$

Equation (16) is valid due to (19) and (14). Equations (17) and (18) hold due to (15).

From (20), we obtain an optimal control in the form

$$u_0(t) = -\frac{2}{d} b^T Hx(t). \quad \square$$

4 CONCLUSION

In the contribution, for solving optimal stabilization problems for differential equations and their systems with two delays, the Lyapunov's direct method (developed by N.N. Krasovskii) was applied. This method permits to find a control function in the form of a feedback such that the zero solution of given equation or system will be asymptotically stable and, simultaneously, an integral quality criterion attains a minimum value. Further research might be focused on similar problems with multiple delays.

ACKNOWLEDGEMENT

The author was supported by the Grant FEKT-S-14-2200 of Faculty of Electrical Engineering and Communication, BUT and by the Norway Grants mechanism under the project "Personality, learning approaches and teaching styles in undergraduate mathematics - PLATSUM", registration number NF-CZ07-ICP-4-324-2016.

REFERENCES

[1] Malkin, I.G.: Theory of Stability of Motion, Second revised edition, (Russian), Moscow, Nauka Publisher, 1966, 530 pp.

[2] Gantmacher, F.R.: The Theory of Matrices, vol. I, II, AMS Chelsea Publishing, Providence, RI, USA, 2002.

ROUGH SETS ON STATE SPACES OF AUTOMATA

David Staněk
Doctoral Degree Programme (1), FEEC BUT
E-mail: xstane41@stud.feec.vutbr.cz

Supervised by: Jan Chvalina
E-mail: chvalina@feec.vutbr.cz

Abstract: This paper discussed a subclass of finite automata, which have ordering on the state sets created by a transition (or next-state) function. Hence, there do not exist cycles of more than one element.
We discuss a relation of equality of upper closure on the systems of all subsets of state systems of quasi-automata, which creates an equivalence.

Keywords: Iterations, Kuratowski-Whyburn equivalence, equivalence, monoid, quasi-automaton, upper closure.

1. INTRODUCTION

Rough sets theory was established in the 1982 by Professor Z. Pawlak. This theory is highly usable when working with vagueness and indiscernibility in an inaccurate data field. Further, theory is applied in fields such as data analysis, determination of recognizing of images theory, electrical (machine) learning, expert systems, using knowledge basis etc.

Original Pawlak rough sets theory was based on equivalence relation concept [9]. This was generalized for arbitrary relations and universum covering, [1, 10]. Further, this theory was developed in the way of topological spaces. Study of rough sets and its connection with algebraic structures and ambiguous mappings and multioperations, has its own importance, [2, 8].

Note that upper and lower approximations of planar field defined by function of one variable, occur in the Riemann definite integral definition.

2. USED CONCEPTS

Let f be a mapping of set a X into itself (transformation of set a X). Then for an arbitrary non negative integer $n \in \mathrm{N}_0$, f^n denotes the $n-th$ *iteration* of the transformation $f : X \to X$. Moreover, $f^0 = \mathrm{id}_X$ is identical transformation of the set X and $f^{n+1} = f \circ f^n$, if $f^n : X \to X, n \in \mathrm{N}$ is defined.

Denote by $\sim_f$ relation on a set X such that $x \sim_f y$ if and only if there exists a pair of non negative integers $m, n \in \mathrm{N}_0$ such that $f^m(x) = f^n(y)$.

The relation $\sim_f$ is obviously reflexive and symmetric on the algebra (X, f). Moreover, if for $x, y, z \in X$ $x \sim_f y$, $y \sim_f z$ holds, then there exist elements $m, n, k, l \in \mathrm{N}_0$, such that $f^m(x) = f^n(y)$, $f^k(y) = f^l(z)$. Hence the $f^{m+k}(x) = f^{n+k}(y) = f^{k+n}(y) = f^{l+n}(z)$. The relation $\sim_f$ is then the equivalence relation on X. Just mentioned relation is called *Kuratowski-Whyburn* equivalence or shortly K-W equivalence.

Blocks $S \in X| \sim_f$ are called $f-$orbits. A subalgebra $(S, f|S)$ (here symbols $f|S$ mean the restriction of the function f on the set S) of the mono-unary algebra (X, f) is said to be a *component* of the algebra (X, f), where the inclusion $f(S) \subset S$ holds.

If $(S_\alpha, f_\alpha); \alpha \in I$ is the system of all components of the algebra (X, f) (here $f_\alpha = f|S_\alpha$), we write

$$(X, f) = \sum_{\alpha \in I} (S_\alpha, f_\alpha)$$

and this sum is called an *orbital decomposition* of the mono-unary algebra (X, f). Notice, that defined concepts are overtaken e.g. from [4, 5, 6, 7].

By *monoid* we mean an ordered pair $(M, \cdot)$, where M is non empty set and $\cdot$ is an associative binary operation with with neutral element.

By a *quasi-automaton* we mean a triad $\mathrm{A} = (S, M, \delta)$, where S is a state set or phase set, M is an input alphabet or a phase semigroup and $\delta : S \times M \to S$ transition function (or next-state function) satisfying axioms:

- $\delta(s, e) = s$ for every element $s \in S$ (neutral element axiom),
- $\delta(\delta(s, m), n) = \delta(s, m \cdot n)$ for every element $s \in S$ and $m, n \in M$ (mixed associativity axiom).

A quasi-automaton can be also considered as a a discrete dynamical system or an action of monoid M on the set S, [3, 6].

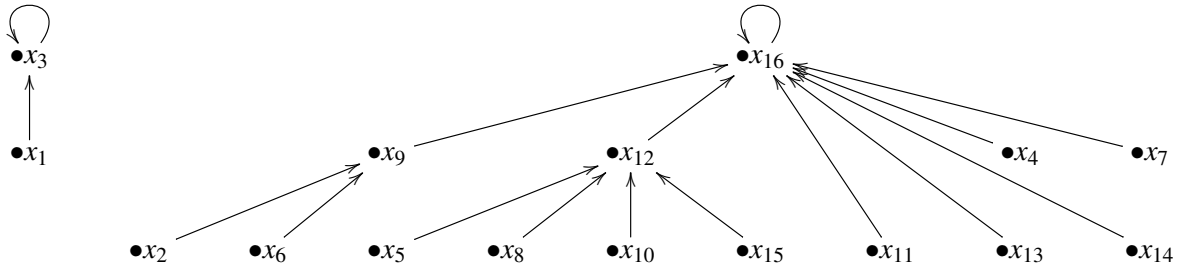
Illustrative example. *Let us consider transformation $f(\vec{x}) : \{0,1\}^4 \to \{0,1\}^4$ defined by the rule $f(\vec{x}) = A\vec{x} + \vec{b}$, where*

$$A = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

and calculations of functional values of function $f(\vec{x})$ are based on Boolean arithmetic. The table below summarizes the sum and the product of elements x, y in Boolean arithmetic.

x	y	$x+y$	$x \cdot y$
0	0	0	0
0	1	1	0
1	0	1	0
1	1	1	1

Orbital structure of function $f(\vec{x})$ with two fixed points x_3, x_{16} is



where x_i are column vectors listed in the table below.

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}
0	*1*	*0*	*0*	*0*	*1*	*1*	*1*	*0*	*0*	*0*	*1*	*0*	*1*	*1*	*1*
0	*0*	*1*	*0*	*0*	*1*	*0*	*0*	*1*	*1*	*0*	*1*	*1*	*0*	*1*	*1*
0	*0*	*0*	*1*	*0*	*0*	*1*	*0*	*1*	*0*	*1*	*1*	*1*	*1*	*0*	*1*
0	*0*	*0*	*0*	*1*	*0*	*0*	*1*	*0*	*1*	*1*	*0*	*1*	*1*	*1*	*1*

Now we show, that for the $n-th$ iteration of function $f(\vec{x})$ we have $f^n(\vec{x}) = A^n \cdot \vec{x} + \vec{b}$. For $n = 1$ the equality is obvious. Induction hypothesis is, that $f^n(\vec{x}) = A^n \cdot \vec{x} + \vec{b}$ holds for all positive integers up to $n \in \mathrm{N}$. Hence

$$f^{n+1}(\vec{x}) = f \circ f^n = f\left(A^n(\vec{x}) + \vec{b}\right) = A\left(A^n(\vec{x}) + \vec{b}\right) + \vec{b},$$

because $A \cdot \vec{b} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, *then*

$$A\left(A^n(\vec{x}) + \vec{b}\right) + \vec{b} = AA^n\vec{x} + \vec{b} = A^{n+1}\vec{x} + \vec{b} = f^{n+1}(\vec{x}) + \vec{b}. \quad \square$$

Let us define a function δ_A as follows:

$$\delta_A(\vec{x}, n) = A^n\vec{x} + \vec{b}$$

for every vector $\vec{x} \in \{0,1\}^4$ and every positive integer $n \in \mathrm{N}_0$. Then the triad $A = \left(\{0,1\}^4, \mathrm{N}_0, \delta_A\right)$ is a quasi-automaton with the state set $\{0,1\}^4$, input alphabet N_0 and transition function δ_A.

With respect to the above considered function $f(\vec{x})$, the elements $s,t \in S$, are in relation $\sim_f$ if and only if for certain $a,b \in \mathrm{N}_0$ we have $\delta(s,a) = \delta(t,b)$, where $\delta(s,a) = f^a(s)$.

Let U be a non empty set - called an universal set - and Θ be an equivalence relation on U. The pair (U,Θ) is called the *approximation space.* Denote $\underline{\mathrm{Ap}}(X) = \{x \in U; [x]_\Theta \subset X\}$, $\overline{\mathrm{Ap}}(X) = \{x \in U; [x]_\Theta \cap X \neq \emptyset\}$. For $x \in X$ a symbol $[x]_\Theta$ denotes a block (class) of equivalence Θ containing the element x, i.e. $[x]_\Theta = \{y \in X; [x,y] \in \Theta\}$.
The set $\underline{\mathrm{Ap}}(X)$ $(\overline{\mathrm{Ap}}(X))$ is called the *lower rough approximation* of X (the *upper rough approximation* of X) in the space (U,Θ). Define a mapping Apr:$P(U) \to P(U) \times P(U)$ by the rule: For $X \in P(U)$ we put $\mathrm{Apr}(X) = (\underline{\mathrm{Ap}}(X), \overline{\mathrm{Ap}}(X))$. A pair $(A,B) \in P(U) \times P(U)$ is called a *rough set* in (U,Θ) if there exists $X \in P(U)$ with the property $(A,B) = \mathrm{Apr}(X)$. (In some papers by a rough set only the upper approximation $\overline{\mathrm{Ap}}(X)$ is meant.)

We define *upper closure* u_δ^+ of set $Q \in P(S)$ as follows

$$u_\delta^+(Q) = \{\delta(q,m) | q \in Q; m \in M\}.$$

3. CHARACTERIZATION OF THE EQUIVALENCE RELATION IN THE UNIVERSAL SET $P(S)$

Now, let us define a binary relation ρ on the set $P(S)$ as it follows: elements (sets) $A,B \in P(S)$ are in relation ρ if and only if $u_\delta^+(A) = u_\delta^+(B)$. It is easy to show, that the given relation is an equivalence, since reflexivity, symmetry and transitivity conditions evidently hold.

With respect to the illustrating example we focus on the class of finite automata, which have ordering on state sets created by transition functions. Hence, for every pair $[s,a],[t,b]$ with property

$\delta(s,a)=[t,b]$, $\delta(t,b)=[s,a]$, we have $[s,a]=[t,b]$. Hence, there do not exist cycles with more than one element. To be more precise, relation $\sigma \subset S\times S$ defined as $s\sigma_\delta t \equiv \exists a\in M$ with the property $\delta(s,a)=t$, is antisymmetric. Moreover, since the relation ρ is reflexive and transitive, it is an quasi-ordering, i.e. (S,σ_δ) is quasi-ordered set.

Theorem. Let $A=(S,M,\delta)$ be a quasi-automaton. For any pair $A,B\subset S$ we have

$A\rho B$ if and only if either $A=B$ or for any state $s\in A\triangle B$ there exists a state $t\in A\cap B$ such that $\delta(t,m)=s$ for suitable element $m\in M$ (Condition **C**).

***Proof.* 1.** *Suppose $A,B\subset S$ such that $A\rho B$, i.e. $u_\delta^+(A)=u_\delta^+(B)$.*
If $A\triangle B=\emptyset$ we have $(A\backslash B)\cup(B\backslash A)=\emptyset$, i.e. $A=B$ and clearly $A\rho B$.
Thus we suppose $A\triangle B\neq\emptyset$ and $s\in A\triangle B$. Assuming $A\cap B=\emptyset$ we obtain that there exists a pair of states $x\in A, y\in B$ with the property $\delta(x,m)=y$, $\delta(y,n)=x$ for a suitable pair of elements $m,n\in M$. In this case states x,y are elements of an at least two-element cycle within S. But this is excluded by the assumption concerning the structure of the quasi-automaton in question. Thus $A\cap B\neq\emptyset$.
If $s\in A\backslash B$, since $s\in u_\delta^+(A)=u_\delta^+(B)$ there exists $t_1\in B$ and $m\in M$ with the property $s=\delta(t_1,m)$. But $t_1\in u_\delta^+(A)$. Hence there exists a sequence $t_2,t_3,\ldots,t_k\in A\cup B$ with the property $t_k\in A\cap B$ (since there are excluded cycles), $\delta(t_i,m_i)=t_{i-1}$ for suitable finite sequence $m_2,m_3,\ldots,m_{k-1}$. Denote $n=m_1\cdot m_2\cdot\cdots\cdot m_{k-1}$. Then $\delta(t_k,n)=t_1$ and $\delta(t_k,n\cdot m)=\delta(\delta(t_k,n),m)=s$. Denoting $t=t_k$ we obtain $\delta(t,n\cdot m)=s$ for $t\in A\cap B$.

*The consideration starting with $s\in B\backslash A$ is quite similar to the just described. Hence the above condition **(C)** is satisfied.*

***2.** Now suppose the condition **(C)** is satisfied. Let $s\in u_\delta^+(A)$ be an arbitrary state. There exist a state $s_1\in A$ and element $m\in M$ such that $s=\delta(s_1,m)$. Two cases are possible:*

(i) $s_1\in A\backslash B$, (ii) $s_1\in A\cap B$.

Consider the case (i): Since

$$A\backslash B\subset(A\backslash B)\cup(B\backslash A)=A\triangle B,$$

*we have $s_1\in A\triangle B$. By the condition **(C)** there exists a pair $[t,n]\in(A\cap B)\times M$ such that $s_1=\delta(t,n)$. Then*

$$t\in B, s=\delta(s_1,m)=\delta(\delta(t,n),m)=\delta(t,n\cdot m),$$

thus $s\in u_\delta^+(B)$.
In the second case (ii) we have also $s_1\in B$ and thus $s\in u_\delta^+(B)$ again, hence $u_\delta^+(A)\subset u_\delta^+(B)$.
Supposing $s\in u_\delta^+(B)$, we obtain in a similar way as above that $s\in u_\delta^+(A)$, thus the opposite inclusion $u_\delta^+(B)\subset u_\delta^+(A)$ holds, as well. Consequently $u_\delta^+(A)=u_\delta^+(B)$, i.e. $A\rho B$. The proof is complete. □

Remark. The above considered necessary and sufficient condition for the relationship $A\rho B$ can be reformulated using the ordering $\leq$ into the form:
For $A,B\subset S, A\rho B$ if and only if either $A=B$ or for any $s\in A\triangle B$ there exists $t\in A\cap B$ such that $t\leq s$.

4. CONCLUSIONS

Rough set theory proposed by Pawlak [9] is an extension of the set theory by study of information systems characterized by inexact and uncertain information. Thus, rough sets theory is an powerful mathematical tool for uncertain data while modelling the problems in computer science, medical science, data analysis and many other diverse fields.
In this contribution there are explained basic terms of the automata theory motivated by an illustrative example. Using closures determined by the input monoid and by the transition function we define

an equivalence relation ρ of the state set of an quasi-automaton. There is characterized a pair of $\rho-$related subsets of the considered state sets, where the equivalence ρ creates an approximation space on $P(S)$.

ACKNOWLEDGEMENT

The author was supported by the Norway Grants mechanism under the project "Mathematical Education Through Modelling Authentic Situations - METMAS", registration number NF-CZ07-ICP-3-201-2015. This support is gratefully acknowledged.

REFERENCES

[1] ABO-TABL, El-Sayed. Rough Sets and Topological Spaces Based on Similarity. *International Journal of Machine Learning and Cybernetics.* 2013, **5**(4), 451-458.

[2] BAGIRMAZ, Nurettin a Abdullah F. OZCAN. Rough Semigroups on Approximation Spaces. *International Journal of Algebra.* 2015, **9**(7), 339-350.

[3] BAVEL, Zamir. The Source as a Tool in Automata. *Information and Control.* 1971, **18**(2), 140-155.

[4] BERÁNEK, Jaroslav a Jan CHVALINA. Extensions of Cascades Created by Certain Functions Systems. *Acta Mathematica Nitriensia.* 2015, **1**(1), 50-56.

[5] CHVALINA, Jan. Diskrétní Orbitální Struktura Zobrazení a Funkcí. In *Sborník VIII. brněnské konference o vyučování matematice.* Brno: MU Brno, 1992, (8),23-28.

[6] CHVALINA, Jan. *Funkcionální Grafy, Kvaziuspořádané Množiny a Komutativní Hypergrupy.* 1. vyd. Brno: Masarykova univerzita, 1995. ISBN 80-210-1148-3.

[7] CHVALINA, Jan. Stars of subsets Within Set Partitions and Isomorphic Closure Operators. *9th International Didactic Conference.* 2015, **9**, 14-18.

[8] NOVÁK, Michal. n-ary Hyperstructures Constructed from Binary Quasi-ordered Semigroups. *Analele Stiintifice Ale Universitatii Ovidius Constanta, Seria Matematica,* 2014, **22**(3), 147-168.

[9] PAWLAK, Zdzisław. Rough Sets. *International Journal of Computer and Information Sciences.* 1982, **2**(5), 341-356.

[10] ZHU, William. Topological Approaches to Covering Rough Sets. *Information Sciences.* 2007, **6**(177), 1499-1508.

SIGNAL FILTRATION FOR DETECTING ELECTROMAGNETIC EMISSION

Aleksandr Podshivalov

Doctoral Degree Programme (3), FEEC BUT

E-mail: xpodsh00@stud.feec.vutbr.cz

Supervised by: Pavel Koktavý

E-mail: koktavy@feec.vutbr.cz

Abstract: There are many methods to describe and study geological processes. This paper deals with a data measured by a loop antenna in a dripstone cave. The measuring system, the sequence of the experiment, and result of signal filtering and spectral analysis were described. Several artificial noise sources in different frequency of signal spectrum bands were found. Signal filtering was carried out. On the basis of the filtered signal central moments during a few days were calculated.

Keywords: electromagnetic emission, central moment, ionosphere, filtration

1. INTRODUCTION

Signals of electromagnetic and acoustic emissions appear during cracks generation when a solid is exposed to mechanical loading (tensile, compressive, shear, torsion, etc.). Generation of electromagnetic emission is related to electric charge redistribution during cracks creation and development and it is in the frequency range from tenths of Hz up to the gamma radiation. Acoustic emission appears due to release of elastic energy during this process and it is in frequency range of ultrasonic waves [1].

Many researchers have noticed a variety of environmental changes occurring prior to local earthquakes. The most remarkable earthquakes are abrupt regional stress changes and deformation. The stress changes reveal as tiny displacements along activated faults and they are represented by two characteristic effects: long-term creep observed in year-long intervals and short-term pulses. Processes of seismic activity may accompanied by electromagnetic emission (EME), which theoretically could be observed for a distance of up to several tens of kilometers from the epicenter. And it is possible to carry out measurements of EME by an antenna. In our research we use a loop antenna for monitoring and study geological processes. This method could be aimed to monitoring seismic activity, detection and prediction of earthquakes.

2. MEASUREMENT SYSTEM

Electromagnetic emission (EME) monitoring was carried out near the active faults by means of the Emission data logger (EDLOG) specifically developed in the framework of the project [6]. EDLOG (fig.1.) is a grid powered standalone multichannel device for long-term measurement of EME signals. It consists of analog and digital parts. Key components of the analog parts are one or more ferrite or shielded loop antennas turned into the specific direction for the EME signals detection. The digital parts are based on special PC components and multichannel data acquisition (DAQ) cards.

The control and pre-processing software was developed in the LabView environment; the signals will be saved to exchangeable data storage with possibility for offline data evaluation. EDLOG is designed to have sufficient ingress protection IP 66 and good EM shielding of key components to avoid interference and satisfy electromagnetic compatibility EMC. Thus the EDLOG is suitable for

use in caves with electric grid where the humidity is high and background EM noise is relatively low.

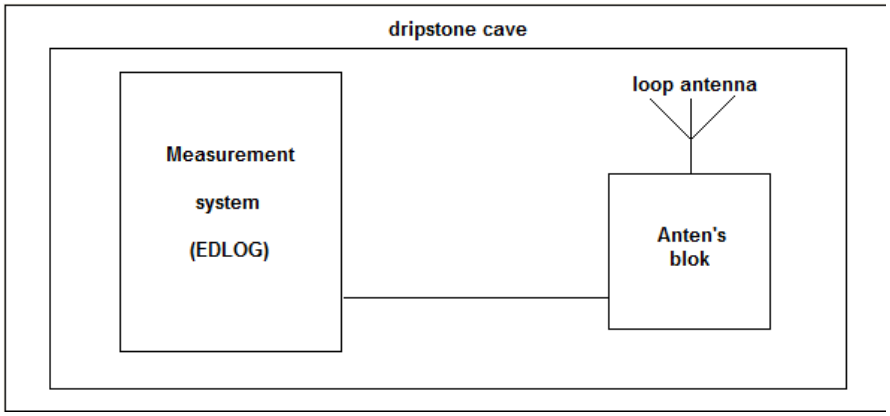


Figure 1: Block scheme of measurement in a cave

These technical approaches will allow us to provide accurate analysis of EME signals with high sensitivity and signal processing demands. The digital part is consist of 6 TB replaceable data storage, 8 GB ECC unbuffered RAM, four simultaneously sampled analog inputs with the 14-16 bit resolution, the sampling rate 2 MS/s for 1-channel, 1 MS/s for multichannel and LabView control software.

2.1. DATA PROCESSING

Results of the signal processing during few days are shown in figures 2, where is shown second central moment. Measurement period was from 05.05.2015 17:29:04 to 14.05.2015 14:15:40.

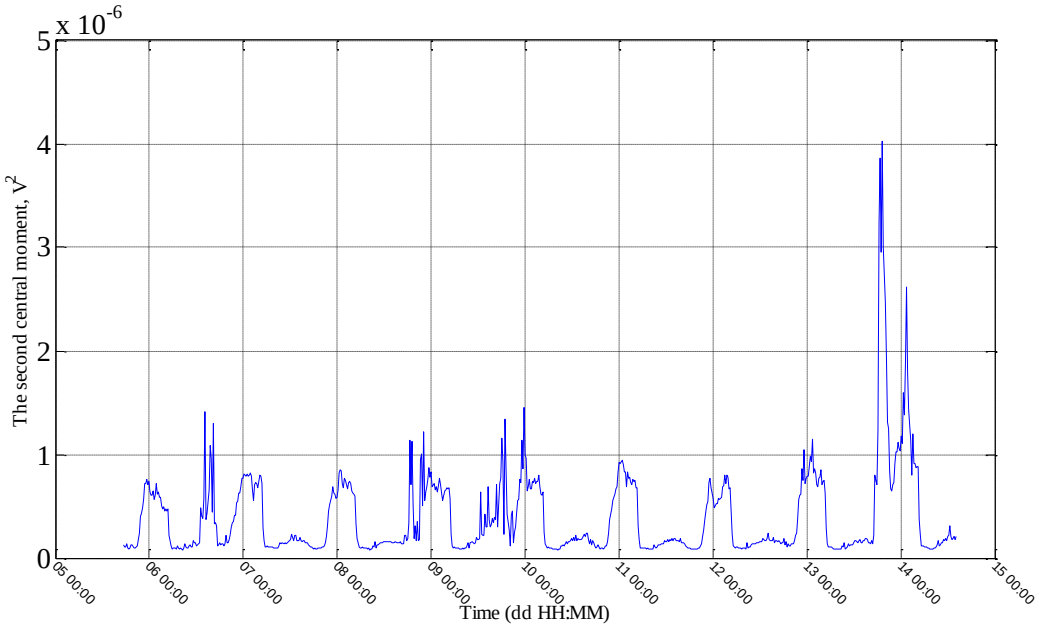


Figure 2: Second central moment during a week of data measured in the cave

The results is quite interesting. The second central moment has the characteristic impulses, which appear every night. It has a periodic nature and likely to be associated with the operation of some device. Also we can see groups of peaks near 6-7, 9-10 and 13-14 days, and this peaks are randomly occur in the daytime and at night time. And likely those anomaly appear due to presence of EME in the measured signal, which can be correlated with seismic activity in the dripstone cave.

2.2. SPECTRAL ANALYSIS RESULTS

Figure 3 shows a part of the signal spectrum in the daytime and at night. Here we can see a part of averaged spectrum (red-nighttime) of 100 data files from 08.05.2015 2:00 to 08.05.2015 4:47 and a part of averaged spectrum (black-daytime) of 100 data files from 08.05.2015 5:12 to 08.05.2015 7:57. The default size of each file is 100 MB. Sampling frequency was 500 kHz. Multiple signals at different frequencies were found. A comparison of spectra at different times of day was carried out. Some of them substantially increase at night. We came to the conclusion that some of interferences come from outside of the cave.

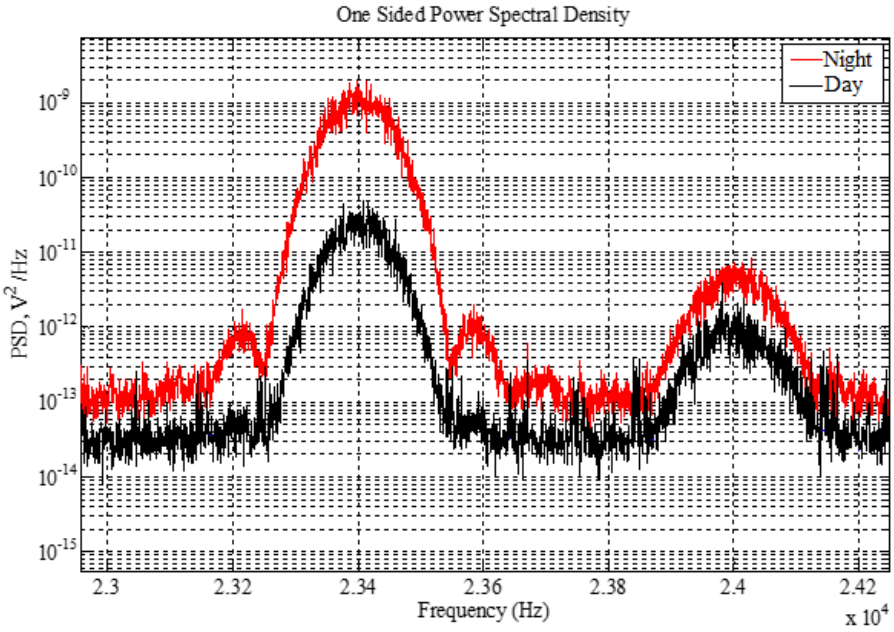


Figure 3: Spectra of nighttime (red) and daytime (black) signals

Table 1 shows some frequency band interference. And 81 % of the power difference at nighttime and daytime is in the frequency range from 23164,27 Hz to 23641,11 Hz.

Table 1: Power difference at nighttime and daytime for the various frequency bands

Power, V^2	Power, %	f mean, Hz	f min, Hz	f max, Hz
4,80E-08	***7,377518***	***20279,41***	***20036,22***	***20522,59***
4,79E-07	***81,24049***	***23402,69***	***23164,27***	***23641,11***
2,51E-09	*0,358295*	*24005,89*	*23869,99*	*24141,79*
1,65E-09	0,178541	25367,26	24199,01	26535,51
1,26E-08	***2,033260***	***26702,40***	***26626,11***	***26778,70***
3,93E-08	**2,541362**	**32090,66**	**26850,22**	**37331,10**
8,26E-09	***1,255274***	***37495,61***	***37359,71***	***37631,51***
1,84E-09	0,199450	38001,06	37879,47	38122,65
1,54E-09	0,222588	40399,55	40277,96	40521,14
3,97E-10	0,048770	42507,17	42390,35	42624,00
2,59E-09	0,406451	44202,33	44078,35	44326,31
2,10E-08	***2,490975***	***45902,25***	***45670,99***	***46133,52***

The findings were confirmed by the data sources of interferences at a given frequencies which was found here [2]. Localizations and channels of sources are represented by figure 4.

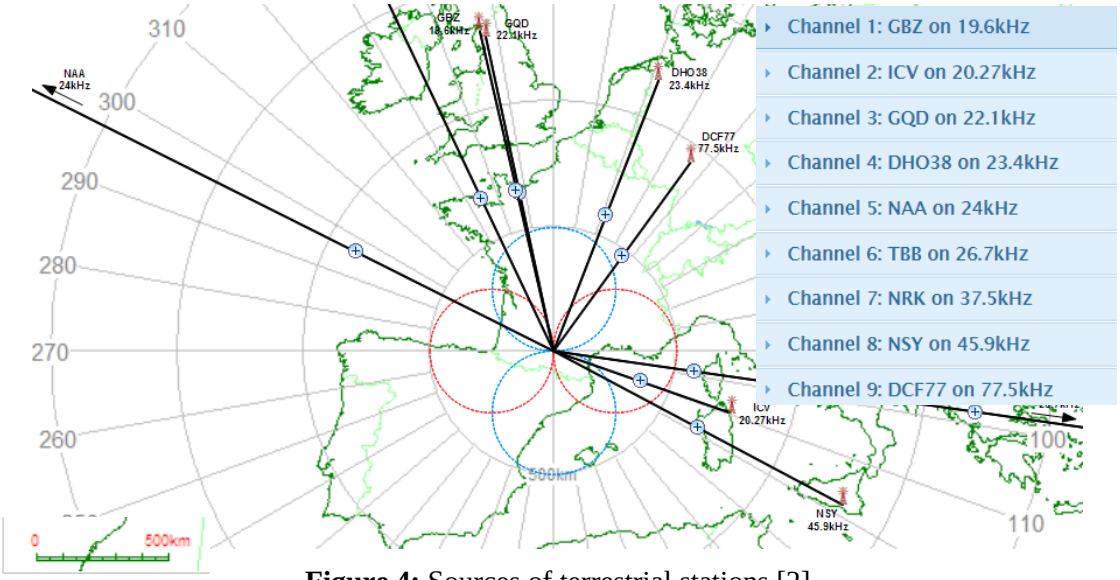


Figure 4: Sources of terrestrial stations [2]

This map shows the location of the monitored VLF transmitters (antenna). The transmitters are from 760km (470 miles) to 5300km (3300 miles) from the receiving station. The azimuthal equidistant projection used for this map renders the signal paths – that follow great circles – as straight lines. The subreflective point icon denotes the location of the subreflective point(s) for each signal path. From one to three points exist depending on the transmitting station distance. They indicate where the signal sky wave bounces off the ionosphere. Finally, the receiving antenna patterns are shown as blue (North-South loop) and red (East-West loop) dotted lines [2].

Therefore, the signals found by us should be filtered. And we get rid of the excess power in the signal. It can eliminate the periodicity of statistical parameters of the signal in the daytime and nighttime. Spectral analysis also showed that the cave does not eliminate all noises from outside, despite depth measurements underground.

2.3. FILTRATION RESULTS

Results of the signal processing during a few days are shown on figure 5, where are shown second central moments before and after filter.

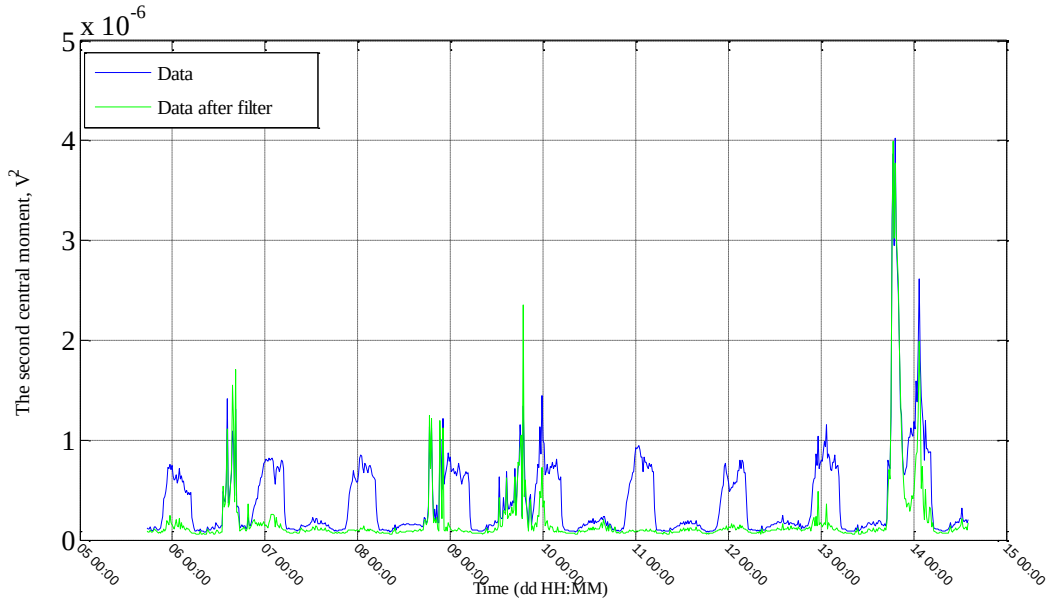


Figure 5: Second central moment during a week before and after filtration

We designed band-stop Butterworth IIR (infinite impulse response) digital filters of third order for a few frequency bands from table 1. Each filter has approximately 50 dB signal suppression for the mean frequency.

On figure 5 we can see some anomalies (before filtration) every night. It depend on ionosphere. At night, the D-layer of ionosphere disappears and the waves (also on frequency bands which were found during spectral analysis) are reflected by the higher E and F layers. The reflection coefficient is higher, thus leading to increased signal strengths at nighttime [2].

After filtration, as can be seen on the figure 5, we were able to get rid of signal power excess. Thus geologic events can be searched in the resulting signal. And further we can carry out correlation between obtained EME anomalies and seismic data.

3. CONCLUSION

Spectral analysis showed that the cave does not eliminate all noises from outside, despite depth measurements underground. The periodicity of statistical parameters of the signal in the daytime and nighttime depends on ionosphere.

We found interferences on different frequencies. We also found evidence that these interferences have artificial origin. These interferences were eliminated by filtration.

Thus, after filtration, we have a signal, where we can search geological events.

ACKNOWLEDGMENTS

This work was supported by the Internal Grant Agency of Brno University of Technology, grant No. FEKT-S-14-2240.

REFERENCES

[1] Koktavý, P.: Experimental study of electromagnetic emission signals generated by crack generation in composite materials. Measurement Science and Technology. 2008. 20(1). p. 0 - 7. ISSN 0957-0233.

[2] http://sidstation.loudet.org/channels-map-en.xhtml

[3] Koktavý, P., Trčka, T., Koktavý, B.: Noise diagnostics of advanced composite materials for structural applications. In 21st International Conference on Noise and Fluctuations ICNF 2011. 1. Toronto, Kanada, IEEE. 2011. p. 88 - 91. ISBN 978-1-4577-0191-7.

[4] Trčka, T. Characteristics of the improved set-up for electromagnetic and acoustic emission signals continual measurement. In Proceedings of the 16th Conference Student EEICT 2010. Volume 595. Brno, 2010, NOVPRESS s.r.o. 2010. p. 139 - 143. ISBN 978-80-214-4079-1.

[5] Podshivalov, A. Method for searching anomalous events in the electromagnetic emission signal using statistical data processing. In MMK 2015, International Masaryk conference for Ph.D. students and young researchers. MAGNANIMITAS, 2015. s. 2446-2450. ISBN: 978-80-87952-12- 2.

[6] ELMA. Electromagnetic emissions of deformed rock due to active tectonics and deep-seated slope failures. 2015.

SOLUTION OF A WEAKLY DELAYED DIFFERENCE SYSTEM

Jan Šafařík
Doctoral Degree Programme (II), FEEC BUT
E-mail: xsafar19@stud.feec.vutbr.cz

Supervised by: Josef Diblík
E-mail: diblik@feec.vutbr.cz

Abstract: The paper solves a weakly delayed difference system

$$x(k+1) = Ax(k) + Bx(k-1)$$

where $k = 0, 1, \ldots$, $A = (a_{ij})_{i,j=1}^{3}$, $B = (b_{ij})_{i,j=1}^{3}$ are constant matrices. An explicit solution is given with a discussion on the number of independent initial data.

Keywords: Discrete system, weak delay, initial problem.

1 INTRODUCTION

We investigate a system of difference equations

$$x(k+1) = Ax(k) + Bx(k-1),\ k = 0, 1, \ldots \tag{1}$$

where A and B are 3 by 3 constant matrices with elements a_{ij} and b_{ij}, $i, j = 1, 2, 3$ respectively and the matrix A has real and mutually different eingenvalues λ_1, λ_2, λ_3 and is given in the Jordan form

$$A = \Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}. \tag{2}$$

We assume that (1) is a weakly delayed system in the sense of the following definition.

Definition 1 *System* (1) *is called weakly delayed if the characteristic equations for* (1) *and for the system without delay*

$$x(k+1) = Ax(k)$$

have identical roots, that is, if, for every $\lambda \in \mathbb{C} \setminus \{0\}$,

$$\det\left(A + \lambda^{-1}B - \lambda I\right) = \det\left(A - \lambda I\right).$$

Applying Definition 1 to system (1) we get conditions under which the system is weakly delayed:

Theorem 1 *System* (1) *is a weakly delayed system if and only if*

$$b_{11} = b_{22} = b_{33} = 0, \tag{3}$$
$$b_{12}b_{23}b_{31} + b_{13}b_{21}b_{32} = 0, \tag{4}$$
$$b_{12}b_{21} + b_{13}b_{31} + b_{23}b_{32} = 0, \tag{5}$$
$$\lambda_3 b_{12}b_{21} + \lambda_2 b_{13}b_{31} + \lambda_1 b_{23}b_{32} = 0. \tag{6}$$

We omit the proof and referring to [1], Theorem 3.

The purpose of the paper is to give an explicit solution of the initial problem

$$x_i(0) = x_{i,0},\ x_i(-1) = x_{i,-1},\ i = 1,2,3 \tag{7}$$

to system (1). If the system considered is weakly delayed, some of the initial data in (7) do not influence the solution's behaviour. We discuss this problem as well.

2 RESULTS

System (1) can be transformed into a system without delay. Below we describe this transformation. Taking into account the above mentioned properties of matrix B, we have

$$B = \begin{pmatrix} 0 & b_{12} & b_{13} \\ b_{21} & 0 & b_{23} \\ b_{31} & b_{32} & 0 \end{pmatrix} \tag{8}$$

and its elements satisfy (3)–(6). From these formulas obviously follows that B is a nilpotent matrix (all its eigenvalues are equal to zero). Below we assume that the geometrical multiplicity of the zero eigenvalue equals 2.

I.e., the system (1) is

$$\begin{array}{rcccccc} x_1(k+1) & = & \lambda_1 x_1(k) & & & +b_{12}x_2(k-1) & +b_{13}x_3(k-1), \\ x_2(k+1) & = & & \lambda_2 x_2(k) & +b_{21}x_1(k-1) & & +b_{23}x_3(k-1), \\ x_3(k+1) & = & & \lambda_3 x_3(k) & +b_{31}x_1(k-1) & +b_{32}x_2(k-1) & \end{array}$$

with initial data

$$\begin{array}{ll} x_1(0) = x_{1,0}, & x_1(-1) = x_{1,-1}, \\ x_2(0) = x_{2,0}, & x_2(-1) = x_{2,-1}, \\ x_3(0) = x_{3,0}, & x_3(-1) = x_{3,-1}. \end{array}$$

Define new dependent functions z_1, z_2 and z_3 by formulas

$$\begin{array}{lcl} z_1(k) = x_1(k-1), & \Rightarrow & z_1(k+1) = x_1(k), \\ z_2(k) = x_2(k-1), & \Rightarrow & z_2(k+1) = x_2(k), \\ z_3(k) = x_3(k-1), & \Rightarrow & z_3(k+1) = x_3(k) \end{array}$$

and, instead of (1), consider a new system without delay

$$\begin{array}{rcccccc} x_1(k+1) & = & \lambda_1 x_1(k) & & & +b_{12}z_2(k) & +b_{13}z_3(k), \\ x_2(k+1) & = & & \lambda_2 x_2(k) & +b_{21}z_1(k) & & +b_{23}z_3(k), \\ x_3(k+1) & = & & & \lambda_3 x_3(k) & +b_{31}z_1(k) & +b_{32}z_2(k), \\ z_1(k+1) & = & x_1(k), & & & & \\ z_2(k+1) & = & & x_2(k), & & & \\ z_2(k+1) & = & & & x_3(k). & & \end{array}$$

Unifying the notation of the dependent variables as

$$y_i(k) = x_i(k),\ i = 1,2,3,\ y_{j+3}(k) = z_j(k),\ j = 1,2,3$$

we have

$$\begin{array}{rcccccc}
y_1(k+1) &=& \lambda_1 y_1(k) & & & +b_{12}y_5(k) & +b_{13}y_6(k), \\
y_2(k+1) &=& & \lambda_2 y_2(k) & +b_{21}y_4(k) & & +b_{23}y_6(k), \\
y_3(k+1) &=& & & \lambda_3 y_3(k) + b_{31}y_4(k) & +b_{32}y_5(k), & \\
y_4(k+1) &=& y_1(k), & & & & \\
y_5(k+1) &=& & y_2(k), & & & \\
y_6(k+1) &=& & & y_3(k). & &
\end{array}$$

The matrix form of the last system is

$$y(k+1) = \mathcal{A}y(k) \tag{9}$$

where

$$\mathcal{A} = \left(\begin{array}{ccc|ccc}
\lambda_1 & 0 & 0 & 0 & b_{12} & b_{13} \\
0 & \lambda_2 & 0 & b_{21} & 0 & b_{23} \\
0 & 0 & \lambda_3 & b_{31} & b_{32} & 0 \\
\hline
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0
\end{array}\right) = \left(\begin{array}{c|c} \Lambda & B \\ \hline E & \Theta \end{array}\right) \tag{10}$$

or

$$y(k+1) = \begin{pmatrix} \Lambda & B \\ E & \Theta \end{pmatrix} \cdot y(k).$$

The initial data for system (9) are

$$\begin{aligned}
y_1(0) &= x_1(0), \\
y_2(0) &= x_2(0), \\
y_3(0) &= x_3(0), \\
y_4(0) &= x_1(-1), \\
y_5(0) &= x_2(-1), \\
y_6(0) &= x_3(-1).
\end{aligned}$$

We will transform system (9) using $y(k) = Sw(k)$ where S is a regular transient matrix and $w(k)$ is a new dependent vector into a system with a matrix of the Jordan form. We get

$$Sw(k+1) = \mathcal{A}Sw(k)$$

or

$$w(k+1) = \gamma w(k) \tag{11}$$

where

$$\gamma = S^{-1}\mathcal{A}S.$$

with the initial data for (11) being

$$w(0) = S^{-1}y(0).$$

It is easy to prove that the eigenvalues of the matrix $\mathcal{A}$ are λ_1, λ_2, λ_3, $\lambda_4 = \lambda_5 = \lambda_6 = 0$. Then the solution of the (11) is

$$w(k) = \gamma^k w(0),\ k = 1,2,3,\ldots$$

where the powers of γ "transient" forms for $k=1$ and $k=2$, i.e.,

$$\gamma = \left(\begin{array}{ccc|ccc} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_3 \end{array}\right) = \left(\begin{array}{c|c} \Lambda_1 & \Theta \\ \hline \Theta & \Lambda \end{array}\right)$$

and

$$\gamma^2 = \left(\begin{array}{ccc|ccc} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & \lambda_1^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_3^2 \end{array}\right) = \left(\begin{array}{c|c} \Lambda_1^2 & \Theta \\ \hline \Theta & \Lambda^2 \end{array}\right).$$

For $k \geq 3$ the powers γ^k are given by

$$\gamma^k = \left(\begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & \lambda_1^k & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2^k & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_3^k \end{array}\right) = \left(\begin{array}{c|c} \Theta & \Theta \\ \hline \Theta & \Lambda^2 \end{array}\right).$$

The solution of (9) is

$$y(k) = Sw(k) = S\gamma^k w(0),\ k = 1,2,3,\ldots .$$

Using an auxiliary matrix

$$Q = \left(\begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right) = (E|\Theta)$$

we can write the solution of the initial system (1) in the form

$$x(k) = QS\gamma^k w(0),\ k = 1,2,3,\ldots \tag{12}$$

where

$$w(0) = S^{-1}x(0). \tag{13}$$

Therefore, the following theorem holds.

Theorem 2 *Let matrix A have the form* (2) *where* λ_1, λ_2, λ_3 *are different real numbers, let matrix B have the form* (8)*, with its elements satisfying* (4)–(6)*, and let the geometrical multiplicity of its zero eigenvalue equal 2. Then the solution of system* (1) *is given by formula* (12) *with* $w(0)$ *given by* (13)*.*

The initial-value problem to system (1) is specified by 6 initial values (7). In the case considered the solution determined by (7) depends, for $k \to \infty$, not on 6 parameters, but only 5 parameters are necessary. This can be seen from system (9) with the matrix $\mathcal{A}$ given by (10). Since in the matrix B only two of its rows are linearly independent, one of the three initial values $y_4(0)$, $y_5(0)$, $y_6(0)$ is unnecessary.

ACKNOWLEDGEMENT

The author was supported by the Grant FEKT-S-14-2200 of Faculty of Electrical Engineering and Communication, BUT.

REFERENCES

[1] Šafařík, J., Diblík, J., Halfarová, H.: *Weakly Delayed Systems of Linear Discrete Equations in* $\mathbb{R}^3$. In *MITAV 2015 (Matematika, informační technologie a aplikované vědy), Post-conference proceedings of extended versions of selected papers*. Brno: Univerzita obrany v Brně, 2015. s. 105-121. ISBN: 978-80-7231-436-2.

[2] Elaydi, S. N.: *An Introduction to Difference Equations*, Third Edition, Springer, 2005.

[3] Khusainov D. Ya., Benditkis D. B., Diblík J.: *Weak delay in systems with an aftereffect.* Functional Differential Equations, 9 (2002), 385–404.

[4] Diblík J., Khusainov D. Ya., Šmarda Z.: *Construction of the general solution of planar linear discrete systems with constant coefficients and weak delay*. Adv. Difference Equ. 2009, Art. ID 784935, 18 pp.

[5] Diblík J., Halfarová H.: *Linear system of discrete equations with a weak delay.* In *Dynamical System Modelling and Stability Investigation*. Kyjev, Ukrajina: University of Kyiv, UA, 2011. s. 26-27. ISBN: 9667652009.

[6] Diblík J., Halfarová H.: *Discrete systems of linear equations with weak delay.* In *XXIX International Colloquium on the Management of Educational Process*. Brno: 2011. s. 1-4. ISBN: 978-80-7231-779- 0.

[7] Diblík J., Halfarová H.: *Explicit general solution of planar linear discrete systems with constant coefficients and weak delays.* Adv. Difference Equ. 2013, Art. number: 50, doi:10.1186/1687-1847-2013-50, 1–29.

[8] Diblík J., Halfarová H.: *General explicit solution of planar weakly delayed linear discrete systems and pasting its solutions.* Abstr. Appl. Anal. 2014, doi:10.1155/2014/627295, 1–37.

STABILITY OF THE ZERO SOLUTION OF STOCHASTIC DIFFERENTIAL SYSTEM

Marie Klimešová
Doctoral Degree Programme (3.), FEEC BUT
E-mail: xklime01@stud.feec.vutbr.cz

Supervised by: Jaromír Baštinec
E-mail: bastinec@feec.vutbr.cz

Abstract: Stability of stochastic differential equations (SDEs) has become a very popular theme of recent research in mathematics and its applications. The method of Lyapunov functions for the analysis of qualitative behavior of SDEs provide some very powerful instruments in the study of stability properties for concrete stochastic dynamical systems, conditions of existence the stationary solutions of SDEs and related problems.

Keywords: Brownian motion, stochastic differential equation, Lyapunov function, stability.

1 INTRODUCTION

Stochastic modeling has an important role in many branches of science and industry. Stochastic model can be used to solve problem which evinces by accident, noise, etc. Definition of probability spaces, Brownian motion, SDE and an existence and uniqueness of solution of these equations were mentioned in Student EEICT 2014 [4]. It was taken from B. Øksendal [8]. The stability theory was introduced by R. Z. Khasminskii [3]. The basic principles of various types of stochastic systems are described by X.Mao [7]. In this paper we will follow up on previous proofs in Student EEICT 2015 [5] and we derived sufficient conditions for general system of the zero solution of the stochastic differential system using Lyapunov function. The results are verified on examples.

2 MAIN RESULTS - SYSTEM WITH TWO-DIMENSIONAL BROWNIAN MOTION

We have a homogenous linear stochastic differential equation

$$\mathrm{d}X_t = AX_t\mathrm{d}t + G\mathrm{d}B_t, \tag{1}$$

where $X_t = \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix}$, $A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}$, $G = \begin{pmatrix} g_1 & g_2 \\ g_3 & g_4 \end{pmatrix}$, $B_t = \begin{pmatrix} B_1(t) \\ B_2(t) \end{pmatrix}$,

$a_1, a_2, a_3, a_4, g_1, g_2, g_3, g_4$ are constants.

Definition 2.1 *Lyapunov quadratic function V is given*

$$V(X_t) = X_t^T Q\, X_t,$$

where $Q = \begin{pmatrix} q_1 & q_2 \\ q_2 & q_1 \end{pmatrix}$ is a symmetric positive-definite matrix, i.e. $q_1 > 0, q_1^2 - q_2^2 > 0$.

Theorem 2.1 *Zero solution of equation (1) is stochastically stable if holds $LV < 0$, where*

$$LV = 2\left[a_1X_1^2(t) + a_4X_2^2(t) + (a_2+a_3)X_1(t)X_2(t) + g_1^2 + g_2^2 + g_3^2 + g_4^2\right].$$

Proof: [6], pp. 12.

Now we can do a discussion under which conditions the system will be stable.
The Euclidean matrix norm on the space R^n can be define as

$$\|A\|_E := \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}^2},$$

where a_{ij} is a matrix element of the i-th line and of the j-th column of the matrix, n is number of matrix raws, m is number of matrix columns. We denote $g_1^2+g_2^2+g_3^2+g_4^2 = \|G\|^2$ and give

$$LV = 2\left[a_1X_1^2(t)+a_4X_2^2(t)+(a_2+a_3)X_1(t)X_2(t)+\|G\|^2\right]. \tag{2}$$

The Lyapunov function LV will be negative definite if and only if

$$a_1X_1^2(t)+a_4X_2^2(t)+(a_2+a_3)X_1(t)X_2(t)+\|G\|^2 \leq 0,$$

because $\|G\|^2 \geq 0$, therefore the matrix A must be sufficiently negative, to obtain a negative definite function. We use the **Sylvester's criterion** which is a necessary and sufficient criterion to determine whether a matrix is positive-definite. [2] In the following consequences we construct solutions to better imagine the stochastic stability.

Example 2.1 *First, we consider a diagonal matrix A and G of equation* (1) *in the form*

$$A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}, G = \begin{pmatrix} \frac{a}{10} & 0 \\ 0 & \frac{a}{10} \end{pmatrix}.$$

Solution: The matrix A will be negative definite under following conditions:

$$\begin{aligned} D_1 &= |a_{11}| = a < 0, \\ D_2 &= a^2 > 0 \end{aligned}$$

if holds D_1 then the condition D_2 is obvious.

Then from (2) follows

$$\begin{aligned} aX_1^2(t)+aX_2^2(t) &\leq -\|G\|^2, \\ a\|X_t\|^2 &\leq -\|G\|^2. \end{aligned}$$

If the variable a is negative and also inequality $a\|X_t\|^2 \leq -\|G\|^2$ is valid, then the system is stochastically stable.
We find a solution of the stochastic system based on eigenvalues. If $a_{12} = a_{21} = 0$, then $\lambda_1 = a_{11}$, $\lambda_2 = a_{22} \Rightarrow \lambda_{1,2} = a$. Because a is negative we make substitution $a = -\alpha, \alpha > 0$. We give a solution of the system

$$\begin{aligned} X_1(t) &= C_1e^{-\alpha t}, \\ X_2(t) &= C_2te^{-\alpha t}, \end{aligned}$$

when C_1, C_2 are constants.
Zero solution of equation (1) with a matrix A is stochastically stable if holds the inequality $a\|X_t\|^2 \leq -\|G\|^2$. We determine stability of solution for $Q = E$

$$\begin{aligned} \mathrm{d}V(X_t) &= 2\left[aX_1^2(t)+aX_2^2(t)+2\left(\frac{a}{10}\right)^2\right]\mathrm{d}t + \frac{aX_1(t)}{5}\mathrm{d}B_1(t) + \frac{aX_2(t)}{5}\mathrm{d}B_2(t), \\ E\{\mathrm{d}V(X_t)\} &= 2\left[aX_1^2(t)+aX_2^2(t)+2\left(\frac{a}{10}\right)^2\right]\mathrm{d}t = LV\,\mathrm{d}t. \end{aligned}$$

There has to hold the inequality $a\|X_t\|^2 \leq -\|G\|^2$, so

$$a^2 + 50a\|X_t\|^2 < 0 \Leftrightarrow a < 0 \vee a < -50\|X_t\|^2.$$

For $X_1(t) = C_1 e^{-\alpha t}, X_2(t) = C_2 t e^{-\alpha t}$ we get

$$a < -50\left(C_1^2 e^{-2\alpha t} + C_2^2 t^2 e^{-2\alpha t}\right).$$

Stochastic differential system is stable for $a < 0$ or $a < -50\left(C_1^2 e^{-2\alpha t} + C_2^2 t^2 e^{-2\alpha t}\right)$.

Example 2.2 *We consider a diagonal matrix A and G of equation* (1) *in the form*

$$A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, G = \begin{pmatrix} \frac{a}{10} & 0 \\ 0 & \frac{b}{10} \end{pmatrix}.$$

Solution: The matrix A will be negative definite under following conditions:

$$\begin{aligned} D_1 &= a < 0, \\ D_2 &= ab > 0 \Rightarrow b < 0. \end{aligned}$$

Then from (2) follows

$$aX_1^2(t) + bX_2^2(t) \leq -\|G\|^2.$$

We find a solution of the stochastic system based on eigenvalues. $\lambda_1 = a, \lambda_2 = b$. We substitute $a = -\alpha, \alpha > 0, b = -\beta, \beta > 0$. We give a solution of the system

$$\begin{aligned} X_1(t) &= C_1 e^{-\alpha t}, \\ X_2(t) &= C_2 t e^{-\beta t}, \end{aligned}$$

C_1, C_2 are constants.
Zero solution of equation (1) with a matrix A is stochastically stable if holds the inequality $aX_1^2(t) + bX_2^2(t) \leq -\|G\|^2$.
We determine stability of solution for $Q = E$

$$\begin{aligned} \mathrm{d}V(X_t) &= 2\left[aX_1^2(t) + bX_2^2(t) + \left(\frac{a}{10}\right)^2 + \left(\frac{b}{10}\right)^2\right]\mathrm{d}t + \frac{aX_1(t)}{5}\mathrm{d}B_1(t) + \frac{bX_2(t)}{5}\mathrm{d}B_2(t), \\ E\{\mathrm{d}V(X_t)\} &= 2\left[aX_1^2(t) + bX_2^2(t) + \left(\frac{a}{10}\right)^2 + \left(\frac{b}{10}\right)^2\right]\mathrm{d}t = LV\mathrm{d}t. \end{aligned}$$

There has to hold the inequality $aX_1^2(t) + bX_2^2(t) \leq -\|G\|^2$, so if for $X_1(t) = C_1 e^{-\alpha t}, X_2(t) = C_2 t e^{-\beta t}$ holds the inequality

$$aC_1^2 e^{-2\alpha t} + bC_2^2 t^2 e^{-2\beta t} \leq -\frac{a^2 + b^2}{100},$$

then the system is stable.

Example 2.3 *We consider a symmetric matrix A and G of equation* (1) *in the form*

$$A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}, G = \begin{pmatrix} \frac{a}{10} & \frac{b}{10} \\ \frac{b}{10} & \frac{a}{10} \end{pmatrix}.$$

Solution: The matrix A will be negative definite under following conditions:

$$\left.\begin{array}{l} D_1 = a < 0, \\ D_2 = a^2 - b^2 > 0 \Rightarrow |a| > |b| \end{array}\right\} \text{ i.e. must be valid } |a| > |b| > 0.$$

Then from (2) follows

$$aX_1^2(t) + aX_2^2(t) + 2bX_1(t)X_2(t) \quad \leq \quad -\|G\|^2 .$$

The variable a must be sufficiently negative and also inequality

$$a\|X(t)\|^2 + 2bX_1(t)X_2(t) \leq -\|G\|^2$$

must be valid, then we can say that the system is stochastically stable.
We find eigenvalues of matrix A as the solution of the characteristic equation

$$det(A - \lambda E) = 0,$$

where E is the unit matrix.

$$\begin{array}{rcl} |A - \lambda E| = (a-\lambda)^2 - b^2 & = & 0, \\ (a-\lambda)^2 & = & b^2, \\ |a-\lambda| & = & |b| . \end{array}$$

Eigenvalues are

$$\begin{array}{r} -a + \lambda_1 = |b| \Rightarrow \lambda_1 = a + |b| , \\ a - \lambda_2 = |b| \Rightarrow \lambda_2 = a - |b| . \end{array}$$

We substitute $a = -\alpha, \alpha > 0, |b| > 0, \alpha < |b|$, i.e. $\lambda_1 = -\alpha + |b|, \lambda_2 = -\alpha - |b|$. For the eigenvalue $\lambda_1 = -\alpha + |b|$ we find the eigenvector $v_1 = (v_{11}, v_{12})$. There is any nonzero vector which fulfills a following relation

$$\begin{array}{rcl} (A - \lambda_1 E)\, v_1 & = & 0, \\ \begin{pmatrix} a - (a+|b|) & b \\ b & a - (a+|b|) \end{pmatrix} v_1 & = & 0. \end{array}$$

For $b > 0$ we choose an arbitrary vector $v_1 = (1,1)^T$, for $b < 0$ we choose $v_1 = (-1,1)^T$. Then

$$\begin{array}{l} \text{for} \quad b > 0 \quad \text{is } X_1(t) = (1,1)^T e^{(-\alpha+b)t}, \\ \text{for} \quad b < 0 \quad \text{is } X_1(t) = (-1,1)^T e^{(-\alpha+b)t}. \end{array}$$

For the eigenvalue $\lambda_1 = -\alpha - |b|$ we find an eigenvector $v_2 = (v_{21}, v_{22})$

$$\begin{array}{l} (A - \lambda_1 E)\, v_2 = 0, \\ \begin{pmatrix} a - (a-|b|) & b \\ b & a - (a-|b|) \end{pmatrix} v_2 = 0. \end{array}$$

For $b > 0$ we choose an arbitrary vector $v_2 = (1,-1)^T$, for $b < 0$ we choose $v_2 = (1,1)^T$. Then

$$\begin{array}{l} \text{for} \quad b < 0 \quad \text{is } X_2(t) = (1,1)^T e^{-(\alpha+b)t}, \\ \text{for} \quad b > 0 \quad \text{is } X_2(t) = (1,-1)^T e^{-(\alpha+b)t}. \end{array}$$

The general solution is given by a linear combination $X_t = C_1X_1(t) + C_2X_2(t)$, with arbitrary constants C_1, C_2.
Zero solution of equation (1) with a matrix A is stochastically stable if holds the inequality $a\|X(t)\|^2 + 2bX_1(t)X_2(t) \leq -\|G\|^2$.
We determine stability of solution for $Q = E$

$$\begin{aligned} \mathrm{d}V(X_t) &= 2\left[a\left(X_1^2(t) + X_2^2(t)\right) + 2bX_1(t)X_2(t) + \frac{a^2}{50} + \frac{b^2}{50}\right]\mathrm{d}t + \frac{aX_1(t) + bX_2(t)}{5}\mathrm{d}B_1(t) \\ &+ \frac{bX_1(t) + aX_2(t)}{5}\mathrm{d}B_2(t), \\ E\{\mathrm{d}V(X_t)\} &= 2\left[a\left(X_1^2(t) + X_2^2(t)\right) + 2bX_1(t)X_2(t) + \frac{a^2 + b^2}{50}\right]\mathrm{d}t = LV\mathrm{d}t. \end{aligned}$$

There has to hold the inequality $a\|X(t)\|^2 + 2bX_1(t)X_2(t) \leq -\|G\|^2$, so if holds the inequality

$$a\|X(t)\|^2 + 2bX_1(t)X_2(t) \leq -\frac{a^2 + b^2}{50},$$

for $b > 0, X_1(t) = (1,1)^T e^{(-\alpha+b)t}, X_2(t) = (1,-1)^T e^{-(\alpha+b)t}$; for $b < 0, X_1(t) = (-1,1)^T e^{(-\alpha+b)t}$, $X_2(t) = (1,1)^T e^{-(\alpha+b)t}$, then the system is stable.

ACKNOWLEDGEMENT

This research was supported by Grant FEKT-S-11-2-921 of Faculty of Electrical Engineering and Communication, Brno University of Technology. The first author was supported by the Norway Grants mechanism under the project METMAS, registration number NF-CZ07-ICP-3-201-2015. This support is gratefully acknowledged.

REFERENCE

[1] DZHALLADOVA, I.A.: *Optimization of stochastic systems*, Kiev, KNEU Press, 2005. (in Ukrainien)

[2] GILBERT, G. T.: Positive Definite Matrices and Sylvester's Criterion. *The American Mathematical Monthly* 98.1 (1991): 44-46. DOI:10.2307/2324036.

[3] KHASMINSKII, R.: *Stochastic stability of differential equations.* 2nd ed. Berlin: Springer, 2012. ISBN 9783642232800.

[4] KLIMEŠOVÁ, M.: Stochastic Differential Equations, *Student EEICT*. Tábor 43a, 61200 Brno: LITERA, 2014. s. 150-154. ISBN: 978-80-214-4924- 4.

[5] KLIMEŠOVÁ, M.: Stability of the Stochastic Differential Equations, *Student EEICT*. Brno: BUT, 2015. s. 526-530. ISBN: 978-80-214-214-5148-3.

[6] KLIMEŠOVÁ, M.; BAŠTINEC, J.: Stability of the Zero Solution of Stochastic Differential Systems with Two-dimensional Brownian motion. *Mathematics, Information Technologies, and Applied Science.* Brno: UNOB, 2015. s. 8-20. ISBN: 978-80-7231-436- 2.

[7] MAO, X. *Stochastic differential equations and applications.* 2nd ed. Chichester: Horwood, 2008, xviii, 422 s. ISBN 978-1-904275-34-3.

[8] ØKSENDAL, B.: *Stochastic Differential Equations*, *An Introduction with Applications*, Springer-Verlag, 1995.

THE INFLUENCE OF THE TYPE OF ELECTRODES ON THE BIOLOGICAL IMPEDANCE MEASUREMENT

Elena Prokopyeva

Doctoral Degree Programme (4), FEEC BUT

E-mail: xproko24@stud.feec.vutbr.cz

Supervised by: Pavel Tománek

E-mail: tomanek@fit.vutbr.cz

Abstract: Electric impedance during meat storage can be affected by various processes: microbial growth, changing of ions mobility, structural changes, etc. Impedance of biological tissues is frequency dependent. The aim of current study is to determine and evaluate the impedance properties of pork meat using two types of electrodes. Measurements were performed during three days in frequency range 100 Hz ÷ 1 MHz.

Keywords: muscle, pork meat, electric impedance, dispersion

1. INTRODUCTION

In recent times, meat production and consumption has increased throughout the world. Therefore, the safety and quality of meat supply is of a great importance for the society. Meat is highly perishable because of the water it contains and abundant nutrients available on the surfaces. There is a strong demand from meat industry for use of nondestructive methods for the meat quality assessing. It is preferable to develop a rapid method that would be also noninvasive, accurate, inexpensive, and capable of automation [1, 2].

The main reason of fresh meat spoilage is microbial growth. During meat storage, the formation of distasteful scent, discoloration, slime and other physicochemical changes affect the consumer's acceptance of meat quality. Hence, meat freshness monitoring during storage is very important.

Measurement of electric impedance has been used for quality control in food industry. It was also shown by various researchers that electric impedance could be used to evaluate meat ageing and meat texture. The impedance of biological tissues and in particular of meat depends on its structure and on ionic conductivity. These two variables vary to some extent as meat ages [3]. The usage of different electrical methods in food industry requires the knowledge of electrical properties of matters and tissues, and effects of their application on products. The electrical properties of muscles have been investigated to determine or predict meat quality. Thus the electrical properties of meat have become an important area of research that should develop a suitable method to ensure quality and safety of meat products, particularly in industry [1].

Meat has a strong electrical anisotropy. Due to that, electrical properties of meat depend on the direction of applied electrical field. This anisotropy is explained by the presence of long aligned muscle fibers filled with an electrolyte and surrounded by a membrane with insulating properties. Muscles have a strongly anisotropic structure organized in three dimensions. A muscle can be considered as a composite network of fibers surrounded by a connective tissue. These fibers contain hundreds of aligned myofibers. All these muscle components form structural levels that have contrasting electrical and dielectric properties. Electrical properties depend on the physical and chemical parameters that determine the concentration and mobility of ions in metabolic fluids [2, 3].

The understanding of behavior of various electrical parameters in meat will help us to create an appropriate technique for assessment the meat ageing.

2. ELECTRICAL CHARACTERISTICS OF BIOLOGICAL MATERIALS

Electrical properties of any material, including biological tissue, can be broadly separated into two categories: conducting and insulating. Electric charges in a conductor move freely in response to an electric field application, whereas in an insulator (dielectric), the charges are fixed and not free to move. If a conductor is placed in an electric field, charges will move within the conductor until the interior field is zero. In the case of a dielectric no free charges exist, net migration of charge does not occur. However, in polar materials positive and negative charge centers of the molecules do not coincide – an electric dipole moment p is said to exist. An applied field E_0 orients the dipoles and produces a field inside the dielectric, E_p, which opposes the applied field. This process is called polarization.

Most materials contain a combination of orientable dipoles and relatively free charges, so that the electric field is reduced in any material relative to its free-space value. The net field inside the material, E, is then

$$E = E_0 - E_p. \tag{1}$$

The net field is lowered by a significant amount relative to the applied field if the material is an insulator and is essentially zero if it is a conductor. This reduction is characterized by a factor ε_r, which is called the relative permittivity or dielectric constant, according to

$$E = E_0 / \varepsilon_r. \tag{2}$$

In practice most materials, including biological tissue, have characteristics of both insulators and conductors because they contain dipoles as well as moving charges. In heterogeneous materials charges may become trapped at interfaces. As positive and negative ions move in opposite directions under the applied field, internal charge separations can then result within the material, producing an effective internal polarization that acts like a very large dipole.

A simple model for a real material, such as tissue, would be a parallel combination of the capacitor and conductor, which is called a Debye-type model. On a macroscopic level, we describe the material as having a permittivity ε and a conductivity σ. The permittivity characterizes the ability of a material to trap or store charge or to rotate molecular dipoles [4]. Electrical conductivity is the ability of a substance to conduct electric current [1]. Electrical impedance is the property of a material to resist electrical current flow, and it consists of resistance, inductive reactance and capacitive reactance [2].

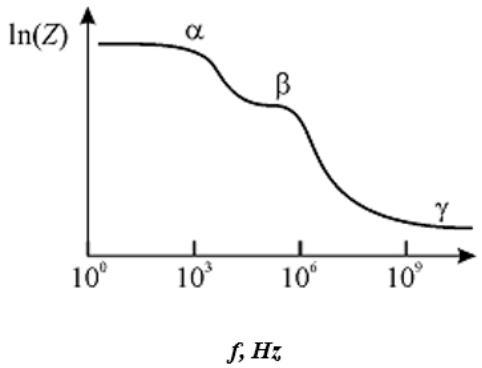


Figure 1: Dispersions in tissues.

In physical terms, we can regard the conductivity of a material as a measure of the ability of its charge to be transported throughout its volume by an applied electric field. Similarly, its permittivity is a measure of the ability of its dipoles to rotate or its charge to be stored by an applied external field. At low frequencies the material will behave like a conductor, but capacitive effects will become more important at higher frequencies. For most materials these properties are not constant, but vary with the frequency of the applied signal. Such a variation is called dispersion. Biological tissues exhibit several different dispersions over a wide range of frequencies (Figure 1).

Dispersions can be understood in terms of the orientation of the dipoles and the motion of the charge carriers. At relatively low frequencies, it is relatively easy for the dipoles to orient in response to the change in the applied field, whereas the charge carriers travel larger distances over which a greater opportunity exists for trapping at a defect or interface. The permittivity is relatively high and the conductivity is relatively low. As the frequency increases, the dipoles are less able to follow the changes of the applied field, and the corresponding polarization disappears. The permittivity decreases and, because trapping becomes less important, the conductivity increases.

Dielectric properties of biological tissues are the result of the interaction between the electromagnetic radiation and the components of cells. These properties can be observed at any frequency of the applied field.

Low frequency (0,1 - 10 kHz) in α-dispersion is caused by the ion diffusion processes in the cell membrane; electric current passing through a material causes the polarization of whole cells, that requires a relatively large amount of time. β-dispersion (hundreds of kHz) is caused by the polarization of cell membranes, where protein macromolecules, and on their borders – globular water-soluble proteins, phospholipids, and tiny subcellular structures take part. Cell membranes act like barriers to the flow of ions between the intracellular and extracellular fluid. γ-dispersion exists in the range of several GHz due to the polarization of water molecules, as well as low-molecular compounds, like saccharides or aminoacids [4].

3. DESCRIPTION OF THE MEASURING TECHNIQUE AND RESULTS

The objective of this work was to investigate the impedance changes during successive degradation of meat using different types of electrodes. The measurement can be generally provided by two- or four-electrode method. Here the two-electrode setup was used to generate the current *I* in the circuit and to measure voltage *V* between the electrodes. The current passing through the sample and the output root mean square (RMS) voltage across meat samples were measured with an impedance analyzer (HP 4284A) through the frequency range 100 Hz – 1 MHz. Stainless steel nails (diameter of 1,5 mm) and an apple cutter were used as electrodes (Fig. 2). The distance between electrodes as well as a depth of their incision was controlled.

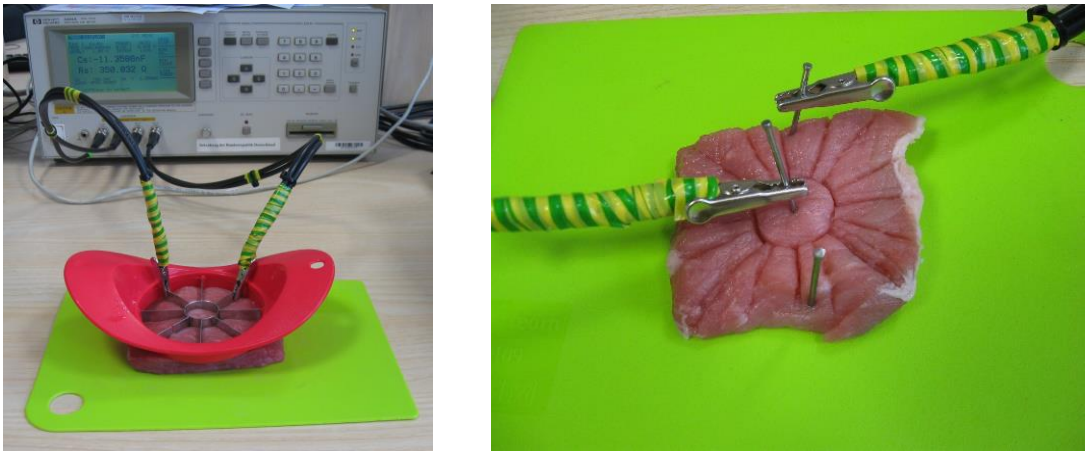


Figure 2: Two-electrode measurement with an apple cutter and steel nails.

Pork chop slices of 12 mm thickness were used as samples. Measurements were provided during 3 days: at the first day fresh meat was measured, then it was kept in the refrigerator at temperature t = +3 °C. The apple cutter was incised into the sample to the depth of 2 mm. The impedance was measured in two inverse directions: between positions 1-4 and 2-3, the distance between those electrodes was 7 cm (Figure 3). Metal parts which were used as electrodes, were sawed up in order to avoid short circuit. The depth of the nails incision was 2 mm. Again the impedance was measured in opposite directions between positions B-C and D-E. The distance between these electrodes was 6 cm.

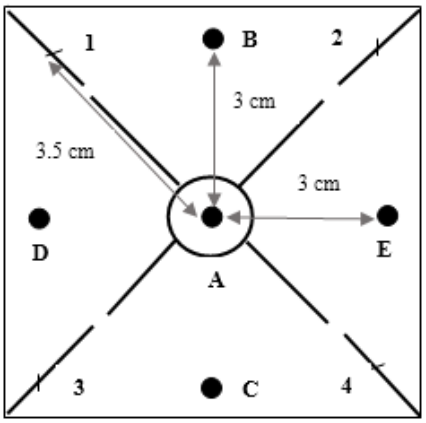


Figure 3: The scheme of electrodes on the slice of meat: 1-4 are positions for the apple cutter, metal parts are sawed up in order to avoid short circuit; A-E are positions for nails.

Firstly, a fresh slice of pork chop was measured by using the apple cutter. Measurements were performed at room temperature three times with an interval 24 hours. Every time the impedance in two inverse directions was measured (curves Z 14 and Z 23). Figure 4 represents the impedance of pork with increasing frequency.

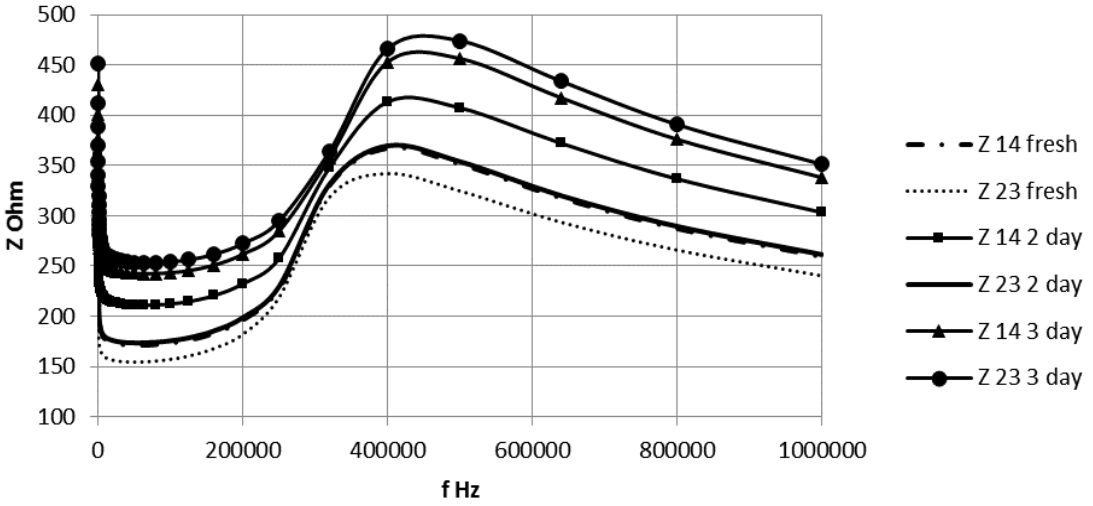


Figure 4: Impedance dependance measured with an apple cutter.

We may correspond the first fall of impedance in the beginning to α-dispersion, the second fall – to β-dispersion. The impedance of fresh meat is about 200 Ohm less than the impedance on the third day. It is noticeable that with meat ageing the peak of β-dispersion is moving to higher frequencies (about 10^5 Hz).

Simultaneously the sample was measured by using steel nails. As we know, muscles are electrically anisotropic. So, these nails were incised to the specified positions in order to measure impedance in transversal and longitudinal directions relatively to muscle fibers (curves Z bc and Z de respectively). Measurements with nails show similar results (Figure 5). Again there are noticeable α- and β-dispersions. The impedance of fresh meat is about 80 Ohm less than the impedance on the last day of measurement. With meat ageing the peak of β-dispersion is moving to higher frequencies (about $5 \cdot 10^4$ Hz).

Values of impedance measured in longitudinal and transversal directions of fresh meat do not differ much, on the second day the impedance measured in transversal direction is about 25 Ohm less than the impedance in longitudinal direction.

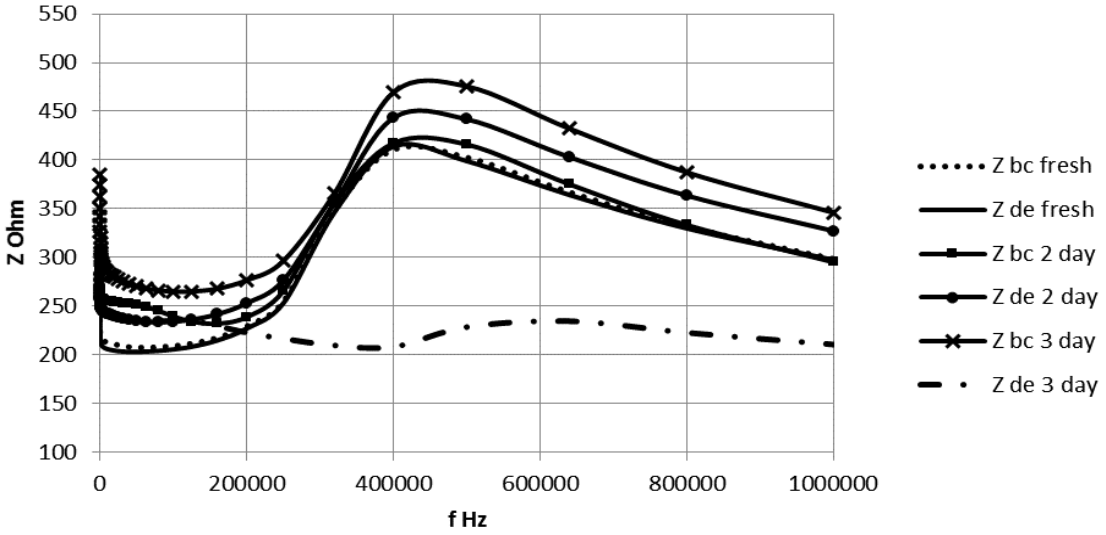


Figure 5: Impedance dependance measured with steel nails.

Myofibers are aligned in the direction of muscle contraction, and electrical conduction along the length of the fiber is significantly easier than conduction across the fibers in the extracellular matrix, because the extracellular matrix is less conductive than the cell. The bottom curve on this graph (Z de 3 day) looks very distinct, it may be explained by a bad contact between electrodes and the sample.

4. CONCLUSION

In this work the electrical impedance of pork chop samples were investigated. Measurements with two types of electrodes show similar results. In measured frequency range it was possible to observe two falls of impedance values, that we may refer to α- and β-dispersions. This effect is connected with the ability of dipoles to orient in response to electrical field changing.

Also impedance depends on the direction of the electric current propagation. In case when the electrodes (nails) were oriented along myofibers, values of impedance were higher.

ACKNOWLEDGEMENT

This work was supported by the Internal Grant Agency of Brno University of Technology, grant No. FEKT-S-14-2240.

REFERENCES

[1] Mahapatra, A. K., Jones, B. L., Nguyen, C. N., Kannan, G.: Experimental Determination of the Electrical Resistivity of Beef. In: Agric Eng Int: CIGR Journal, 2010, 12(3), 124-128

[2] Nguyen, H. B., Nguyen, L. T.: Rapid and Non-invasive Evaluation of Pork Meat Quality during Storage via Impedance Measurement. In: International Journal of Food Science and Technology, 2015, 50, 1718-1725

[3] Damez, J.-L., Clerjon, S., Abouelkaram, S., Lepetit, J.: Beef Meat Electrical Impedance Spectroscopy and Anisotropy Sensing for Non-invasive Early Assessment of Meat Ageing. In: Journal of Food Engineering, 2004, 85, 116-122

[4] Miklavič, D., Pavšelj, N., Hart, F. X.: Electric properties of tissues. Wiley Encyclopedia of Biomedical Engineering, 2006, ISBN: 978-0-471-74036-0