

COMPRESSION, REDUCTION AND ANALYSIS OF DIGITAL AUDIO SIGNALS

Martin ŠEVČÍK, Bachelor Degree Programme (3)
Dept. of Radio Electronics, FEEC, BUT
E-mail: xsevci31@stud.feec.vutbr.cz

Supervised by: Ing. Tomáš Kratochvíl

ABSTRACT

The contribution deals with the comparison of modern audio compression encoders performance characteristics. The fundamentals of basic audio compression principles in time and frequency domains are shortly outlined. Then the modern digital audio compression algorithms (MP3, MP3 Pro, AAC, AAC Plus, OGG, WMA, VQF) features are compared and discussed. The frequency analysis of created test records was made to compare their efficiency of the encoding at low bit-rates.

1 ÚVOD

Cílem projektu bylo srovnání a analýza vlastností moderních komprimačních algoritmů, pro komprimaci zvuku na PC. V příspěvku je vysvětlen základní přístup a problematika při kompresi a redukci digitálního audiosignálu. Pro srovnání byly zvoleny a porovnány formáty MP3, OGG, VQF, MP3 Pro, AAC, AAC Plus a WMA. Pro zadané komprimační algoritmy bude provedena časová a frekvenční analýza spektra vzniklého záznamu. Pro komprimaci a tvorbu zkušebních záznamů a následné srovnání výsledků byly použity programy Nero v6.6 (MP3 Pro, AAC Plus a VQF), dBpowerAMP Music Converter R11 (WMA), iTunes v7 (AAC), OggdropXPdV1.8.6 (OGG), CDex v1.51 (MP3) a Audacity v1.2.3 (frekvenční analýza).

2 MODERNÍ KOMPRIMAČNÍ ALGORITMY

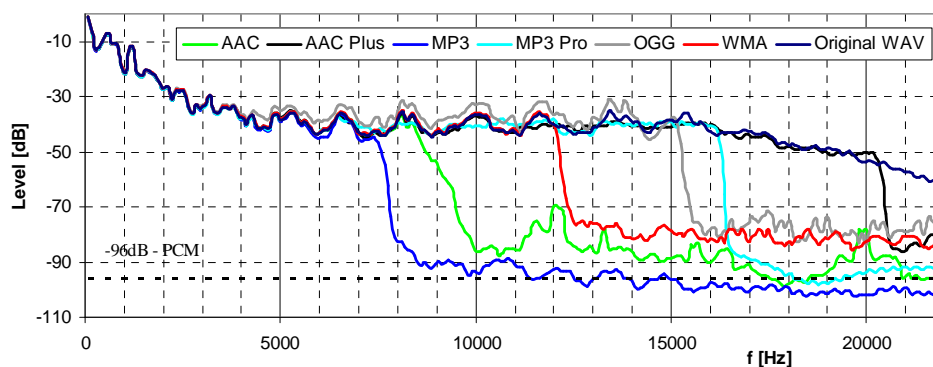
Byly porovnávány moderní komprimační algoritmy, které se používají pro digitální kompresi zvuku. Tyto formáty jsou rozšířeny jak na PC tak některé i ve spotřební elektronice (např. CD/DVD osobní přehrávače). Některé z nich jsou také základem pro ISO standard MPEG-4, který je plánovaný do HDTV. Zde je uveden přehled jednotlivých komprimačních algoritmů a jejich základních vlastností:

- **MP3** (MPEG-1 Layer 3) – byl vytvořen ve Fraunfer IIS, obsahuje psychoakustický model a FFT analýzu, hybridní banky filtrů, MDCT transformaci, nelineární kvantování, Huffmanovo kódování, tvorbu rámců, 2 kanály s konstantní (CBR) nebo proměnnou (VBR) přenosovou rychlostí od 32 do 320 kbit/s.

- **MP3 Pro** – byl vytvořen v Coding Technologies Laboratory, základem je formát MP3 doplněný o technologii SBR (Spectral band Replication) kódující pásmo vysokých kmitočtů. Dekodér je schopný dekódovat tento formát, i když neumí dekódovat SBR část. Podmínkou je podpora používaných vzorkovacích frekvencí.
- **AAC** (Advanced Audio Coding) – byl vytvořen v AT&T, Dolby, Fraunhofer IIS a Sony, ISO standard MPEG-2, obsahuje základní principy komprese MP3, hybridní banku filtrů, volitelnou přenosovou rychlost, predikci signálu a profily, 2 kanály s bitovým tokem od 16 do 320 kbit/s, vícekanálové kódování až do 640 kbit/s.
- **AAC Plus** – podobný formátu MP3 Pro, základ tvoří AAC kódér doplněný o SBR. Navržen pro velmi nízké přenosové rychlosti při uplatnění SBR zhruba do 100 kbit/s.
- **OGG Vorbis** – byl vytvořen v Xiph.org's, schopnost kódování až 255 kanálů, standardní nastavení je VBR (32 až 499,8 kbit/s). Nepodporuje tvoření rámců, synchronizaci ani ochranu proti chybám, dostupný bez licencování.
- **WMA** (Windows Media Audio) – vytvořený firmou Microsoft, obsahuje DRM (Digital Rights Management), CBR a VBR s možností 2-pass kódování. Má čtyři druhy kódování – bezztrátové kódování s poměrem 2:1 nebo 3:1, vícekanálové kódování pro přenosové rychlosti 128 až 768 kbit/s, kódování hlasu s nastavením od 4 do 20 kbit/s a hlavní kodek se ztrátovou kompresí.
- **VQF** (Vector Quantization File) – byl vytvořen v NNT Human Interface Lab a firmou Yamaha, založen na technologii TwinVQ (Transform-domain Weighted Interleave Vector Quantization). Je podobný formátu MP3, ale dosahuje lepších výsledků. Standardní datový tok je 96 kbit/s, je asi o (25-35) % úspornější než MP3 při 128 kbit/s.

3 FREKVENČNÍ ANALÝZA

Vytvořené komprimované zkušební nahrávky byly porovnány s originálním PCM kódovaným digitálním záznamem ve formátu WAV. Analýza byla provedena pro 4 typy skladeb a pro různé datové toky (48 kbit/s, 80 kbit/s, 96 kbit/s, 128 kbit/s) v různých komprimačních programech. Obr. 1 a tab. 1 uvádí výsledky analýzy pro rychlost 48 kbit/s.



Obr. 1: Frekvenční analýza komprimačních algoritmů při 48kbit/s u 2. skladby

1. skladba – Black Sabbath – Paranoid (rocková hudba)
2. skladba – Diana Krall - Stop This World (jazzová hudba)
3. skladba – Genesis - The Brazilian (elektronická hudba)
4. skladba – Mozart - Overture To The Opera The Marriage of Figaro (klasická hudba)

Komprimační algoritmus	Velikost souboru [MB]				Omezení frekvenční charakteristiky na kmitočtu [kHz]			
	1.skladba	2.skladba	3.skladba	4.skladba	1.skladba	2.skladba	3.skladba	4.skladba
AAC	1,058	1,485	1,791	1,524	8,7	8,5	8,4	8,5
AAC Plus	1,023	1,454	1,759	1,492	originál	20,4	20,4	20,3
MP3	1,021	1,435	1,742	1,47	7,5	7,5	7,5	7,3
MP3 Pro	1,022	1,436	1,742	1,471	16,2	16,2	16,2	16,1
OGG	0,943	1,242	1,859	1,185	14,1	15,1	15,1	14,2
VQF	-	-	-	-	-	-	-	-
WMA	1,049	1,472	1,782	1,504	11,9	12	12	12

Tab. 1: Porovnání Komprimačních algoritmů pro čtyři skladby při datovém toku 48 kbit/s.

4 ZÁVĚR

Velikosti jednotlivých souborů jsou srovnatelné pro všechny formáty až na OGG, který má v průměru o 200 kB méně pro zadané přenosové rychlosti, než ostatní komprimační algoritmy. Formát AAC od 128 kbit/s do 80 kbit/s dobře kopíruje průběh originálního spektra a postupné omezení frekvenční charakteristiky se projevuje až od 15 kHz. Naproti tomu při přenosové rychlosti 48 kbit/s dosahuje po MP3 nejhorších vlastností, protože začíná omezovat frekvenční charakteristiku již kolem 8,5 kHz. Formát AAC Plus od 96 kbit/s až do 48 kbit/s dosahuje nejlepších výsledků ve frekvenční charakteristice (pokles až kolem 20,5 kHz). Naopak při datovém toku 128 kbit/s má nejhorší vlastnosti, což je zapříčiněno absencí SBR. Nedostatkem tohoto komprimačního algoritmu je viditelně špatné kopírování originálního průběhu spektra u vysokých frekvencí. I u formátu MP3 Pro lze vyzdvihnout kvality technologie SBR. U datových toků 96 kbit/s a 80 kbit/s je tento formát srovnatelný a u některých skladeb dokonce dosahuje lepších vlastností, než formát AAC Plus. Při nižších datových tocích omezuje frekvenční charakteristiku dříve, ale stále dosahuje velmi dobrých výsledků. Formát VQF používá bitové rychlosti pouze od 192 kbit/s do 80 kbit/s, ale i přesto pro zkoumané datové toky je jeden z nejlepších. Originální průběh spektra kopíruje věrohodněji, než komprimační algoritmy podporující SBR a omezuje frekvenční charakteristiku až kolem 20 kHz. Formát OGG postupně omezuje frekvenční charakteristiky nejhůře do zhruba 15 kHz při datového toku 48 kbit/s, což je velice dobrý výsledek. Výborně kopíruje průběh originálu a při poslechu také zní nejlépe. Formát MP3 je podle předpokladu jeden z nejhorších komprimačních algoritmů při nízkých přenosových rychlostech. Formát WMA nenaplnil předpoklady druhého nejrozšířenější komprimační algoritmu. Špatně kopíruje originální průběh spektra a při nízkých datových tocích omezuje frekvenční charakteristiku již kolem 12 kHz.

PODĚKOVÁNÍ

Tento příspěvek vznikl za podpory rozvojového projektu FRVŠ č. 2508/2005 „*Inovace laboratorní výuky předmětu Nízkofrekvenční elektronika*“ a výzkumného záměru MŠMT č. MSM0021630513 „*Elektronické komunikační systémy a technologie nových generací (ELKOM)*“.

LITERATURA

- [1] Pohlman, K. C.: Principles of Digital Audio. Fourth edition. McGraw Hill, 2000.
- [2] Ševčík, M.: Komprimace, redukce a analýza digitálního audiosignálu. Bakalářská práce, FEKT VUT v Brně, 2006.