

# HEAD TRACKING IN MEETING VIDEO

Michal HRADIŠ, Master Degree Programme (1)  
Dept. of Computer Graphics and Multimedia, FIT, BUT  
E-mail: xhradi05@stud.fit.vutbr.cz

Roman JURÁNEK, Master Degree Programme (1)  
Dept. of Computer Graphics and Multimedia, FIT, BUT  
E-mail: xjuran07@stud.fit.vutbr.cz

Supervised by: Ing. Adam Herout, Ph.D.

## ABSTRACT

The aim of our work was to develop a real-time tracker of heads in meeting video data from the AMI project. The AMI project is concerned with the development of meeting browsers and remote meeting assistants for instrumented meeting rooms. Our system is intended to process video data acquired from camera online. This sets bounds in the terms of computational cost of used algorithms. Therefore we decided first to create a fast solution and analyze its results in order to tell if further improvements are needed. We have divided tracking task into two separate parts – head detection and tracking. Tracking results are stored in the XML format for evaluation.

## 1 ÚVOD

Naše práce se zaměřuje konkrétně na sledování lidských hlav ve videozáznamech meetingů, které probíhají v uzavřených prostorách. Díky tomuto můžeme předpokládat konstantní osvětlení, statickou kameru a statické pozadí. K detekci hlav tak můžeme použít znalosti o jejich velikosti, tvaru a také o barvě kůže.

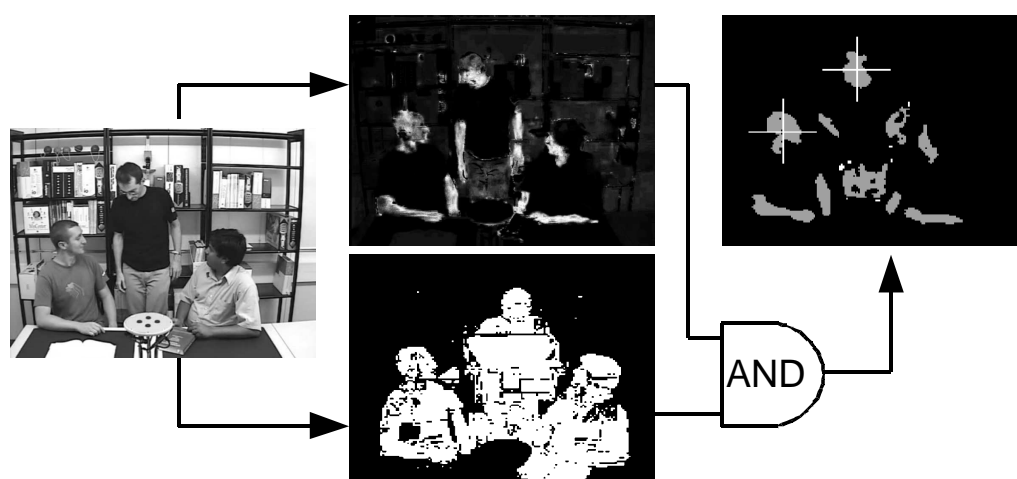
Výsledný systém by měl být schopen zpracovávat online data získaná z jedné či více kamer, což klade značné požadavky na nízkou výpočetní náročnost použitých metod. Proto jsme se rozhodli jako první krok vytvořit jednoduché a rychlé řešení a na základě vyhodnocení jeho výsledků určit, zda je či není nutné použít metody kvalitnějších.

Problém sledování hlav jsme rozdělili na dvě samostatné části. Nejprve je použito barevného modelu kůže a odečtení pozadí k získání regionů s potencionálním výskytem hlav. Zda se skutečně jedná o hlavu je rozhodnuto na základě geometrických vlastností těchto regionů. Detekované objekty jsou následně ve video sekvenci sledovány podobně jako v práci Kölsche a Turka [1]. Pro sledování využíváme upravený KLT feature tracker [3]. Tento postup je výhodný pro sledování velmi deformovatelných objektů, ale také poskytuje odolnost vůči částečnému zakrytí.

## 2 DETEKCE HLAVY

Detekce probíhá v několika krocích. Nejprve je vypočtena maska barvy kůže. V druhé fázi se vytvoří maska popředí. Z obou masek se vypočítá průnik a provede se analýza komponent – hledání oblastí, které tvarově odpovídají hlavě, tedy kompaktní elipsovité tvary. Informace o pozici takovýchto oblastí se předávají na vyšší úroveň do části sledování hlav.

Pro určení pravděpodobnosti barvy kůže používáme model v normalizovaném RG barevném prostoru. V RG barevném prostoru barva kůže většiny lidí vytváří kompaktní shluk, který lze aproximovat Gaussovou funkcí. Tento model je také nezávislý na intenzitě osvětlení. Metoda odstranění pozadí potřebuje pro inicializaci prázdnou scénu k vytvoření počátečního modelu, který se dále průběžně aktualizuje. Při aktualizaci jsou použity jen pixely vyhodnocené jako pozadí. Posledním krokem detekce je analýza topologických vlastností komponent vzniklých průnikem obou masek. V masce se naleznou všechny spojené komponenty pomocí semínkového vyplňování. Z nich se vyberou ty, které mají určitý elipsovitý tvar na základě statistických momentů (kompaktnost a variance) a které zabírají jistou minimální plochu.



**Obr. 1:** Vlevo vstupní obrázek, uprostřed nahoře maska barvy kůže, uprostřed dole maska popředí, vpravo křížkem označené detekované hlavy

## 3 SLEDOVÁNÍ HLAVY

Sledování objektů jsme založili na KLT feature trackeru [3], jenž umožňuje nalézt korepondence bodů v obraze na základě podobnosti jejich malého okolí (např. 7x7 pixelů). Feature tracker jsme doplnili barevným modelem objektu, abychom sledovali jen body odpovídající barvy. K tomu používáme RG barevný model, který může být vytvořen při inicializaci trackeru z barvy objektu, nebo můžeme použít model barvy kůže. Nevýhodou RG barevného modelu je jeho citlivost na osvětlení barevnými zdroji světla, což působí velké problémy například při osvětlení objektu data-projektorem. Barevný model také používáme k detekci ztráty sledovaného objektu – objekt je považován za ztracený pokud procento sledované plochy odpovídající barevnému modelu objektu klesne pod daný práh (např. 40 %).

Sledované body se nesmí vzdálit od společného středu více než o vzdálenost odpovídající předpokládané velikosti hlav v daném obraze, to v kombinaci s barevným modelem téměř

úplně zabraňuje přechodu trackeru na jiný objekt, kromě případu úplného zakrytí. Abychom mohli sledovat i rychle se pohybující předměty, předpovídáme budoucí pozici objektu na základě průměrné rychlosti a zrychlení sledovaných bodů.

## 4 VÝSLEDKY

Výsledná implementace dosahuje průměrné rychlosti 20ms/snímek na procesoru Athlon 64 3500+ při sledování čtyř osob. Zhodnocení výsledků jsme provedli podle standardního schématu používaného v projektu AMI. Bez nutnosti přesného nastavení parametrů jsme dosáhli úspěšnosti 91,5 % při 32 % falešných detekcí. Použití barevného modelu kůže společného pro všechny video sekvence vykazovalo o jednotky procent lepší výsledky než použití speciálního barevného modelu kůže pro každou sekvenci. Odstraňování pozadí mělo kladný vliv pouze ve video sekvencích se statickými objekty v barvě kůže. V kontrolovaných prostředích je však možné takové objekty odstranit. Kvalitnější výsledky při trackingu vykazuje použití konkrétního barevného modelu sledovaného objektu, než použití modelu barvy kůže, což ukazuje, že lze popsaný tracker použít pro sledování i jiných druhů objektů.

## 5 ZÁVĚR

Vytvořili jsme systém, který dokáže správně sledovat hlavy ve více než 90 % testovaných video sekvencí. Počet chybných detekcí je ale stále vysoký – okolo 32 %, protože ruce jsou občas zaměněny za hlavu. Tyto falešné detekce by bylo možné redukovat pomocí dodatečných znalostí o scéně nebo použitím detektoru obličeje. Nevýhodou vytvořeného systému je, že výsledky závisí na množství parametrů, které musí být nastaveny ručně pro různá prostředí. Další nevýhodou je, že barevný model použitý pro detekci kůže je částečně závislý na změnách osvětlení. Výsledek jsme začlenili do systému se snímáním všesměrovým zrcadlem a HDTV kamerou, který pracuje v reálném čase na jednom PC.

Tato práce ověřila, že propojení trackeru s detektorem je vhodné pro sledování objektů v obraze. Také se ukázalo, že pomocí KLT trackeru lze sledovat objekty velikosti hlavy. Budoucí práce na projektu by se mohla zaměřit na použití kvalitnějšího detektoru obličeje – například Viola-Jones a statistického vyhodnocení výstupu KLT trackeru – například pomocí Kalmanova filtru.

## LITERATURA

- [1] Kölsch, M. Turk, M.: Fast 2D Hand Tracking With Flocks and Multi Cue Integration, Department of Computer Science, University of California, 2005
- [2] Potucek, I.: Person Tracking Using Omni Directional Views, Proceedings of Student EEICT 2003, Brno, 2003
- [3] Shi, J., Thomasi, C.: Good Features to Track, IEEE Conference of Computer Vision and Pattern Recognition 1994