

MULTIMEDIA DOCUMENT TYPE DIFF

Jozef Lang

Master Degree Programme (2) FIT BUT

E-mail: xlangj01@stud.fit.vutbr.cz

Supervised by: Petr Chmelař

E-mail: chmelarp@fit.vutbr.cz

Abstract: There is a lot of tools for the comparison of text files but only a few for the comparison of video files. This paper explores basic techniques of content-based video comparison and presents a MediaDIFF module for comparing video files using low level features. The presented module uses low level features in a shot-boundary detection algorithm.

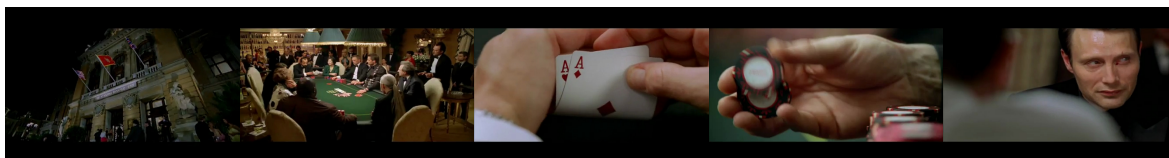
Keywords: Computer vision, video, content-based comparison, features extraction, distance metrics, shot, shot-boundary detection, dynamic time warping

1 ÚVOD

Zvyšovanie rýchlostí pripojenia k internetu a jeho penetrácia do bežného života v posledných rokoch spôsobila zvýšenie množstva multimediálneho obsahu nachádzajúceho sa na internete a to predovšetkým videa. Nástroje na porovnávanie textových dát existujú už niekoľko desaťročí ale o nástroje určené na porovnávanie videa je stále núdzna pričom nájdu uplatnenie aj v profesionálnej sfére. Táto práca predstavuje návrh modulu VideoDiff nástroja MediaDIFF schopného porovnávať videá na základe obsahu využitím nízkourovňových rysov.

2 VIDEO

Video sa skladá zo snímok. Snímky medzi ktorými existuje časová závislosť tvoria zábery a zo záberov, ktoré majú rovnakú sémantiku sa tvoria scény. Prechody medzi jednotlivými zábermi môžu byť uskutočnené strihom, rozplynutím alebo rozotretím. Pri strihu sa jedná o ostrý prechod medzi zábermi ale pri rozplynutí a rozotretí dochádza k prechodu postupne. V sekvencii snímok dochádza k priestorovým redundanciám [1]. Kľúčový snímok, anglicky nazývaný *key-frame*, je snímok, ktorý je reprezentatívny pre určitú sekvenciu snímok, Kľúčové snímky detekuje aj MPEG, ale často je ich počet nízky. Porovnávanie videa je založené na porovnávaní obsahu jednotlivých snímok a preto sa nazýva ako porovnávanie na základe obsahu.



Obrázek 1: Ukážka kľúčových snímok.

3 POROVNÁVANIE NA ZÁKLADE OBSAHU

Porovnávanie jednotlivých snímok je avšak veľmi neefektívne nakoľko dochádza k priestorovým redundanciám a preto je výhodné získať z porovnávaných snímok dáta popisujúce ich obsah – metadata.

Pri porovnávaní multimediálnych dát na základe obsahu je možné použiť rysy. Rys je charakteristika multimediálnych dát a proces ich získavania sa nazýva extrakcia rysov. Rysy je možné rozdeliť do dvoch skupín:

Nízkoúrovňové rysy sú rysy späté z obsahom a konkrétnym typom multimediálnych dát – rysy spojené s farbou a priestorovým rozložením. Je ich možné získavať automatizovane. Štandardom pre popis nízkoúrovňových rysov je norma MPEG-7 [3].

Vysokoúrovňové rysy sú sémantické rysy majúce význam pre koncového používateľa. Pri ich získavaní je potrebná spolupráca s používateľom.

4 METRIKY

Metriky používané pri porovnávaní na základe obsahu sú založené na porovnávaní získaných rysov. Porovnaniu výkonnosti rôznych typov vzdialenostných metrík sa venujú Hampapur a Bolle v práci [2].

Porovnávanie rysov je založené na transformácii vektorov na body v n-dimenzionálnom priestore a následnom použití vzdialenostných metrík. Existuje viacero metrík, napríklad Euklidovská metrika (1) alebo chi-square metrika (2).

$$d_E(a, b) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_n - b_n)^2} = \sqrt{\sum_{i=1}^n (a_i - b_i)^2} \quad (1)$$

$$d_{\chi}(a, b) = \sum_{i=1}^n \frac{(a_i - b_i)^2}{a_i + b_i} \quad (2)$$

5 MEDIADIFF

MediaDIFF je open source nástroj na porovnávanie rôznych druhov dokumentov a formátov na základe ich obsahu. Cieľom je vytvoriť modulárny nástroj schopný porovnávať viacero druhov dokumentov a formátov. MediaDIFF je vyvíjaný na fakulte informačných technológií Vysokého Účenia Technického v Brne v spolupráci s firmou Red Hat. Je vyvíjaný v programovacom jazyku Python a C++. V súčasnosti je schopný porovnávať textové, zvukové súbory a statické obrázky.

6 MODUL VIDEODIFF

Navrhovaný modul implementuje algoritmus pozostávajúci z troch krokov – získania množiny kľúčových snímok porovnávaných videí, získania ich rysov a určenia rozdielnosti použitím Dynamic Time Warping algoritmu.

Proces získavania a reprezentatívnych snímok sa skladá z určenia hraníc jednotlivých záberov. Pre každý záber sa určia kľúčové snímky tak, že pre dva rovnaké zábery sú kľúčové snímky vždy zhodné. Hranice záberov sa určia pomocou podobnosti susedných snímok normalizovanej do intervalu $< 0; 1 >$. Miesto strihu – hranica záberu sa nachádza v bode minima plávajúceho okna o šírke 10 snímok a hodnota minima je nižšia ako hraničná hodnota strihu. Hraničná hodnota strihu je definovaná ako priemerná hodnota podobnosti všetkých snímok znížená o k-násobok smerodajnej odchylky. Hraničnú hodnotu strihu a mohutnosť množiny získaných kľúčových snímok z každého záberu je možné užívateľsky meniť.

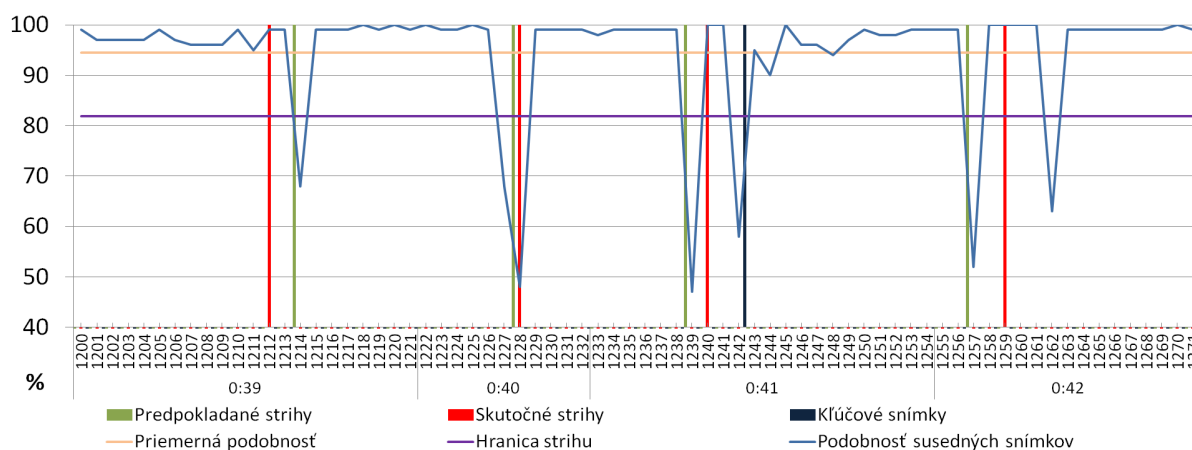
Získané kľúčové snímky zo všetkých porovnávaných videí sú následne použité k určeniu podobnosti pomocou Dynamic Time Warping (DTW) algoritmu. Použitie DTW algoritmu redukuje riešenie problému na problém najdlhšej spoločnej postupnosti.

7 VÝSLEDKY

Výsledky boli získané použitím nízkoúrovňových rysov *Color-layout* definovaných normou MPEG-7. K určovaniu vzdialenosti bola použitá Euklidova metrika. V tabuľke 1 je prehľad nameraných hodnôt a na obrázku 2 je zobrazený výsledný priebeh podobnosti susedných snímkov. K určení hranice strihu bol použitý 1-násobok smerodajnej odchylky. Jednotlivé skutočné strihy boli anotované ručne a je možné vidieť, že MPEG automaticky označil snímok 1242 ako kľúčový. Zároveň je v ukážke možné zreteľne vidieť aplikovanie podmienky plávajúceho okna. Snímky 1242 a 1263 majú nižšiu podobnosť ako je hraničná hodnota strihu ale aj napriek tomu neboli označené ako predpokladaný strih.

Šírka okna	10 snímkov
Priemerná hodnota podobnosti	94,43%
Smerodajná odchylka	12,52%
Hraničná hodnota predpokladaného strihu	81,91%

Tabuľka 1: Namerané hodnoty.



Obrázek 2: Graf podobnosti susedných snímkov s nameranými hodnotami.

8 ZÁVER

Tento príspevok sa snažil čitateľovi priblížiť problematiku porovnávania videa na základe obsahu pomocou nízkoúrovňových rysov. V príspevku bol navrhnutý modul VideoDiff nástroja MediaDIFF, ktorý je schopný porovnávať video na základe obsahu. Príspevok prezentuje výsledky použitého algoritmu na detekciu záberov. Výsledky boli získané z krátkej pasáže voľne dostupnej upútavky k celovečernému filmu. Navrhnutý modul VideoDiff nájde uplatnenie či už v súkromnej alebo profesijnej sfére.

REFERENCE

- [1] Blanken, Henk M.: *Multimedia retrieval*. Springer, 2007, ISBN 978-3540728948.
- [2] Hampapur, A.; Bolle, R.: Comparison of distance measures for video copy detection. In *Multimedia and Expo, 2001. ICME 2001. IEEE International Conference on*, 2001, s. 737 – 740.
- [3] INTERNATIONAL ORGANISATION FOR STANDARDISATION: MPEG-7 Overview (version 10) [online]. <http://mpeg.chiariglione.org/standards/mpeg-7/mpeg-7.htm>, [cit. 2012-03-02].