
Extrakce doplňkových materiálů ze záznamů přednášek

Stanislav Sumeč

(Praha 14.11.2006)

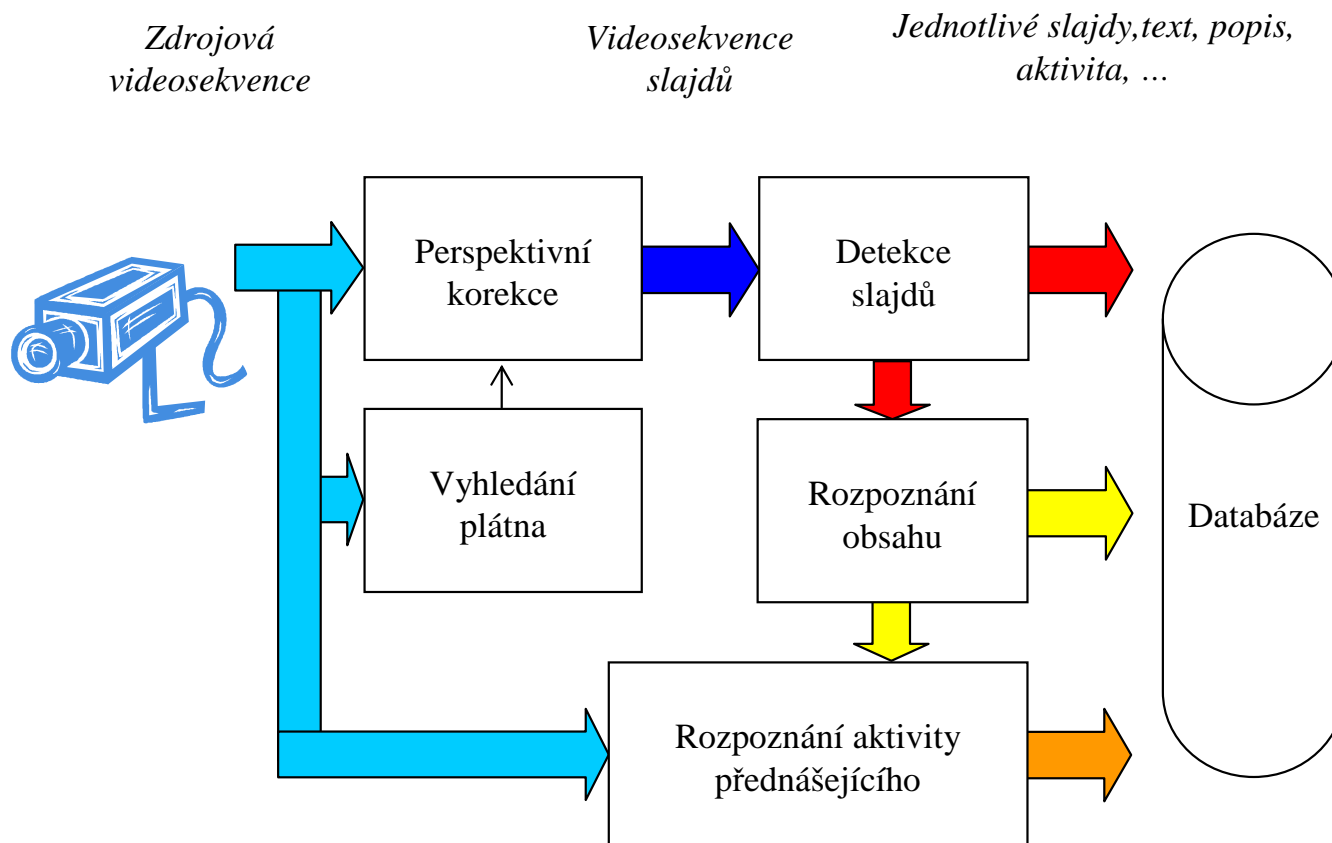
Obsah

- Motivace
 - Návrh systému
 - Vyhledání plátna
 - Korekce perspektivní projekce
 - Detekce slajdů
 - Shrnutí
 - Další práce
 - Odkazy
-

Motivace

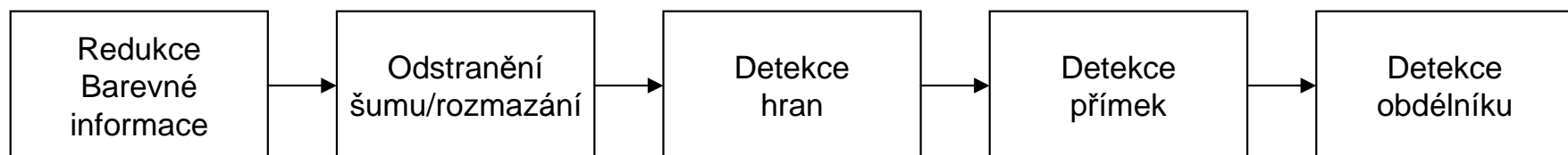
- Pořizování záznamů přednášek v dostatečné kvalitě pro další zpracování.
- Využití záznamu pro automatické získání informací vhodných např. pro indexaci.
- Záznam může obsahovat promítané prezentace – vhodný zdroj dat, není nutné zaznamenávat samostatně.
- Automatizované získání informací o promítané prezentaci.
 - Časová osa zobrazování snímků – vyhledání části přednášky při přehrávání, synchronizace záznamu s originální prezentací, ...
 - Obsah jednotlivých snímků – indexace záznamu, korespondence s výkladem přednášejícího, ...
 - Extrakce snímků ze záznamu – pokud není k dispozici původní prezentace.
- Lze sledovat aktivitu přednášejícího – např. vzhledem k promítané prezentaci, gestikulace, ...

Návrh systému



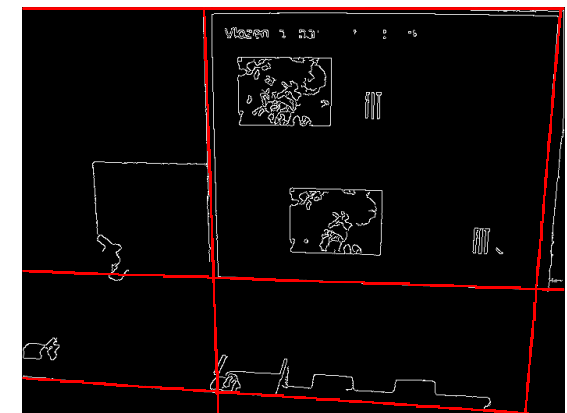
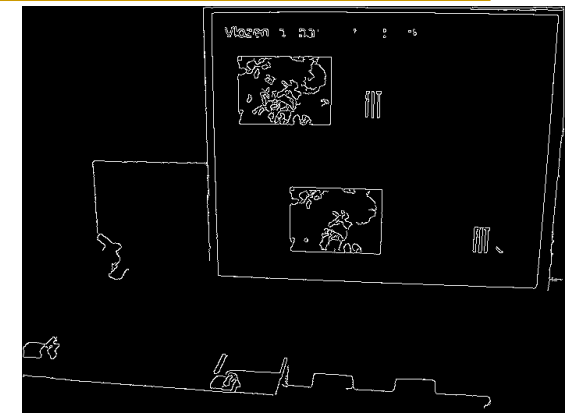
Vyhledání plátna

- Nalezení oblasti obrazu, kde je promítaná prezentace.
- Předpoklady
 - konstantní pozice v celém záznamu
 - výrazná změna jasu na hranici plátna
 - obdélníkový tvar



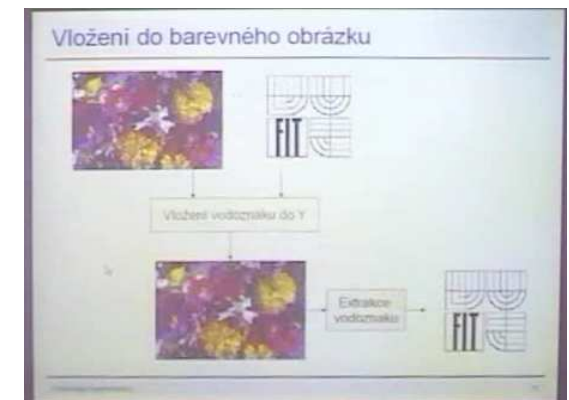
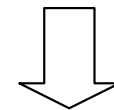
Vyhledání plátna

- Rozmazání obrazu pomocí Gaussian mask.
- Aplikace Cannyho hranového detektoru [5].
- Detekce přímek pomocí Houghovy transformace [6].
- Detekce obdélníku vzhledem k deformaci perspektivní projekcí.
- Opakování hledání v několika snímcích pro zajištění větší spolehlivosti.



Korekce perspektivní projekce

- Přesná pozice projektoru, plátna a kamery není známa.
- Obraz je zkreslen dvěma transformacemi
 - bod (x,y) slajdu je transformován na bod s „neznámou“ pozicí na plátně
 - bod z plátna je transformován na pozici (X, Y) v obrazu kamery
- Cílem korekce je nalézt mapování souřadnic (x,y) a (X, Y) .



Korekce perspektivní projekce

- Řešení navrženo v [7].
- Využívá se znalosti, že všechny body jsou promítány na společnou rovinu.
- Hledá se transformace

$$\begin{pmatrix} xw \\ yw \\ w \end{pmatrix} = \begin{pmatrix} p_1 & p_2 & p_3 \\ p_4 & p_5 & p_6 \\ p_7 & p_8 & p_9 \end{pmatrix} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} \text{ kde } \vec{p} = (p_1 \dots p_9) \quad |\vec{p}| = 1$$

- Dosazením několika „známých“ bodů (např.rohy detekovaného plátna), lze transformaci určit.

Snímek 8

SS1

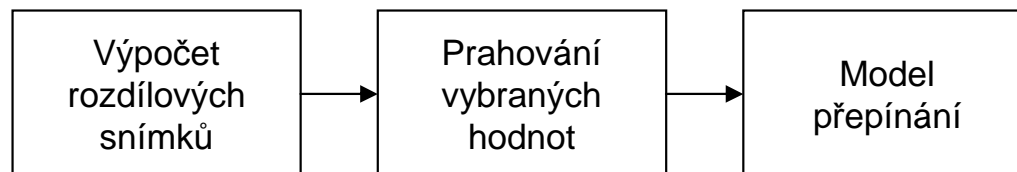
Stanislav Sumec; 14.11.2006

Detekce slajdů

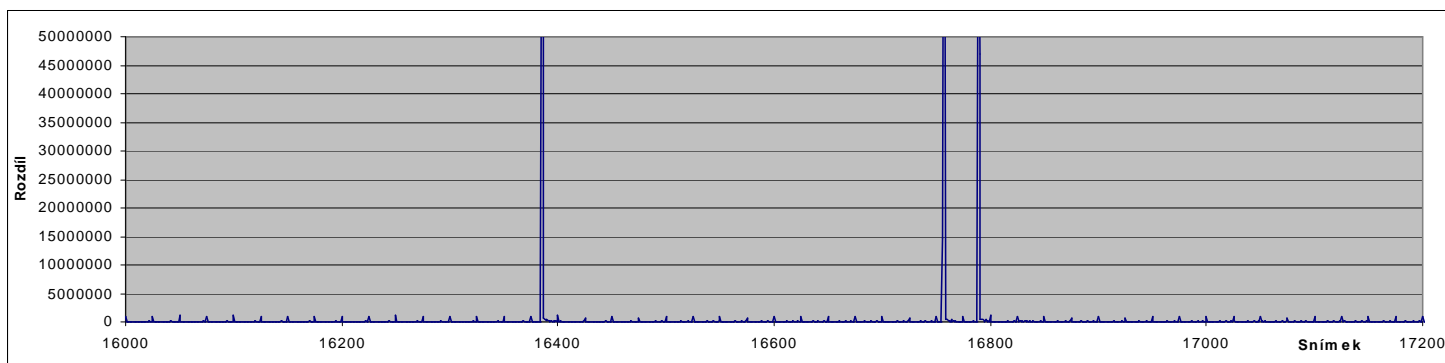
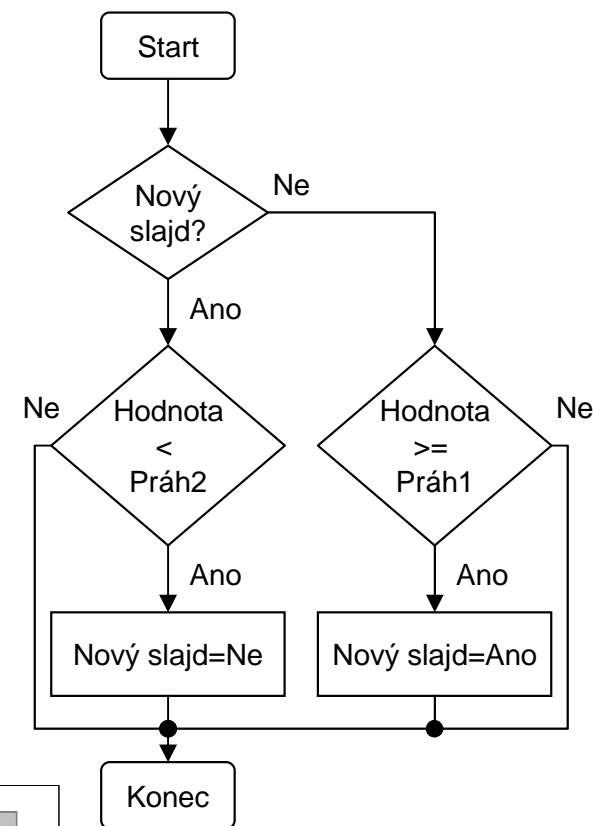
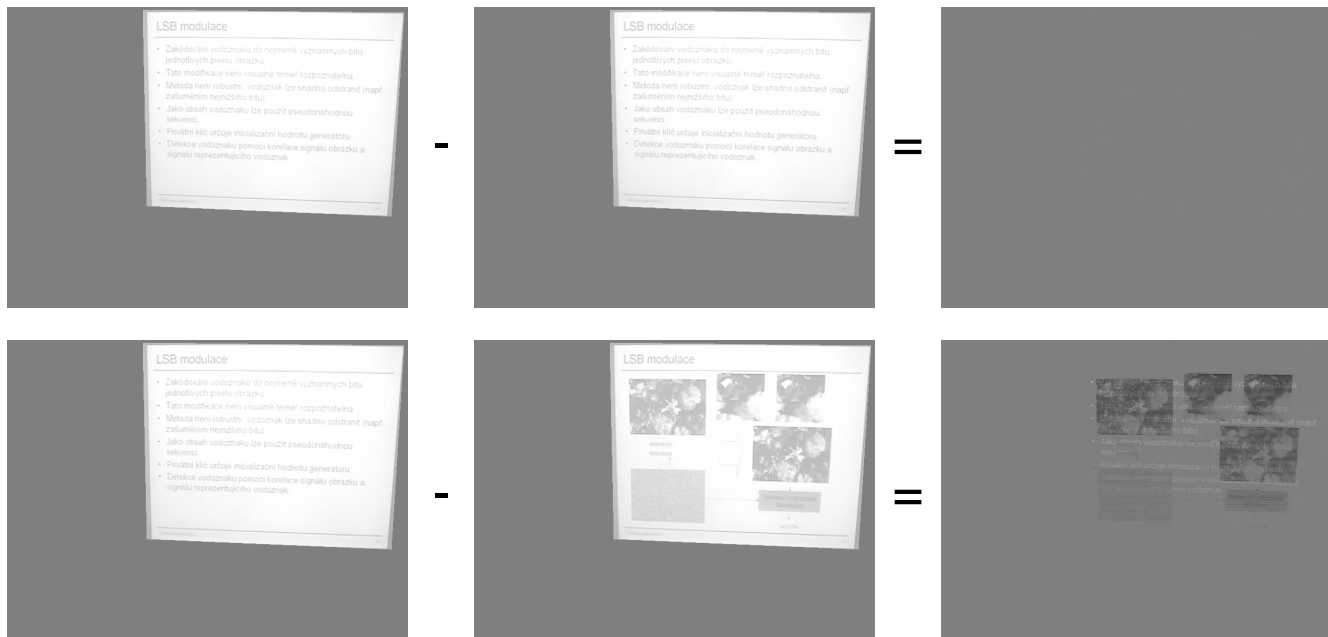
- Různé úlohy podle typu prezentace
 - Identifikace okamžiků, ve kterých dojde ke změně zobrazeného slajdu (přechodové efekty).
 - Identifikace intervalů, kdy dojde ke změně obsahu slajdu (animované zobrazení částí slajdu).
 - Detekce vložených animací a video sekvencí.
 - Nalezení shodných slajdů, odhad pořadí slajdů v prezentaci – sekvenční průchod prezentací.
 - Výsledkem je časová osa promítání snímků a přiřazení skutečného pořadí snímku.
-

Detekce slajdů – statické snímky

- Předpoklady
 - obsah každého snímku je neměnný
 - při přepnutí dojde k významné změně
- Triviální řešení pomocí odečtení dvou nebo několika po sobě jdoucích snímků.
- Vyhodnocení statistických hodnot sekvence rozdílových snímků (minimum, maximum, střední hodnota, suma čtverců).
- Prahování vybraných hodnot konstantním (dynamickým) prahem v daném okně.
- Přejít může být zachycen na několika následujících snímcích => lze stanovit model přepínání.



Detekce slajdů – statické snímky

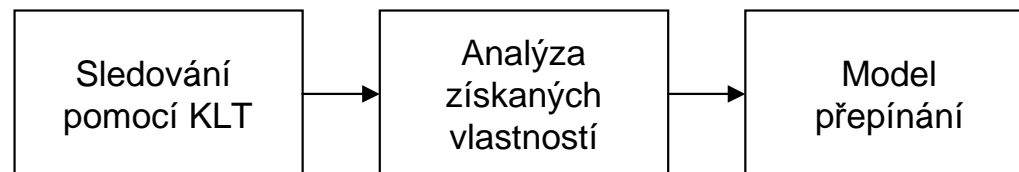


Detekce slajdů – animované přechody

- Snímky jsou zobrazovány postupně, obvykle aplikována animace jednotlivých řádků apod.
- Přechod trvá několik snímků – výsledkem detekce je začátek přechodu, jeho trvání, případně pozice měněné oblasti.
- Rozdílové snímky nelze použít – obtížné stanovení globálního prahu.
- Vhodnější je sledovat lokální změny, aplikace metod pro sledování optického toku, sledování pohybu objektů mezi snímky.

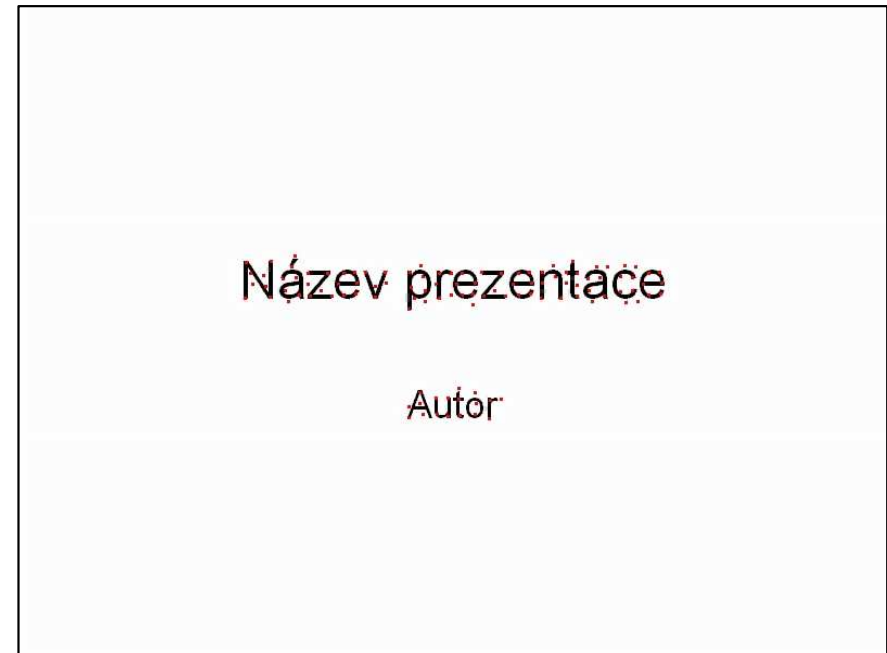
Detekce slajdů – animované přechody

- Testováno použití KLT (Kanade-Lucas-Tomasi) trackeru [4].
- Sledování významných vlastností v obraze (např. rohů).
- Proces sledování je spojen s extrakcí vlastností z textur – „dobrá“ vlastnost je taková, která se „dobře“ sleduje.
- Detekce slajdů je založena na statistickém vyhodnocení sledovaných vlastností.
- Navíc lze využít pozice vlastností pro určení oblasti, ve které dochází k animaci.
- Použitelné také na prezentace bez animací, časově náročnější.



Detekce slajdů – animované přechody

- Vlastnosti při přechodu snímků mohou
 - změnit svou pozici – část slajdu (např. řádek) změnila svou pozici
 - být ztraceny – došlo k významné změně obrazu (např. přepnutí na následující slajd)
 - být vytvořeny nové – objevila se nová část snímku (např. počátek animace řádku)



Detekce slajdů – alternativní řešení

- Přímé napojení na prezentační SW.
- Nutná spolupráce přednášejícího, unifikovaný formát prezentací.
- Lze získat časovou osu zobrazování snímků.
- Analýza video záznamu použitelná pro synchronizaci se záznamem prezentací.
- Microsoft PowerPoint poskytuje napojení prostřednictvím technologie COM (rozhraní IID_IConnectionPointContainer).
- Umožňuje monitorování událostí vzniklých v aplikaci (např. přechod na další slajd, otevření prezentace, ...).

Shrnutí

- Naznačeny možnosti
 - detekce plátna
 - detekce jednotlivých snímků a přechodů
 - perspektivní korekce snímků prezentace
- Výsledky zpracování
 - časová osa zobrazovaných snímků
 - významné oblasti snímků
 - samotné snímky
- Získaná data použitelná
 - při přehrávání záznamu pro snadnější navigaci
 - synchronizace originální prezentace s přehráváním
 - pro tvorbu sestřihů záznamů zohledňujících aktivitu přednášejícího
 - k indexaci záznamů pro vyhledávání
- Možné problémy
 - prezentace promítané „netradičním“ způsobem (např. listování stránek v textovém editoru)
 - nevhodné světelné podmínky
 - ...

Pokračování práce

- Detekce akcí přednášejícího
 - ukazovátko
 - gestikulace
 - ...
 - Evaluace navržených algoritmů
 - stanovení testovací sady dat
 - anotace
 - vyhodnocení
-

Použité nástroje a další odkazy

- [1] Knihovna DigLib (<http://www.fit.vutbr.cz/research/groups/graph/digilib>) – manipulace se snímky, struktura snímku, zobrazování.
 - [2] Knihovna AVFile – práce s video daty (avi, mpeg, ...), transparentní použití na různých platformách.
 - [3] Intel(R) Math Kernel Library (<http://www.intel.com/cd/software/products/asm-na/eng/307757.htm>) – implementace LAPACK.
 - [4] KLT: An Implementation of the Kanade-Lucas-Tomasi Feature Tracker (<http://www.ces.clemson.edu/~stb/klt>).
 - [5] Canny, J.: *A Computational Approach to Edge Detection*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 8, No. 6, Nov 1986.
 - [6] Ballard, D., Brown, C.: *Computer Vision*, Prentice-Hall, 1982, Chap. 4.
 - [7] Sukthankar, R., Stockton, R. G., Mullin, M. D.: *Smarter Presentations: Exploiting Homography in Camera-Projector Systems*, Proceedings of International Conference on Computer Vision, 2001.
-