

SHOT BOUNDARY DETECTION

Lukáš Klicnar

Bachelor Degree Programme (3), FIT BUT

E-mail: xklicn00@stud.fit.vutbr.cz

Supervised by: Vítězslav Beran

E-mail: beranv@fit.vutbr.cz

ABSTRACT

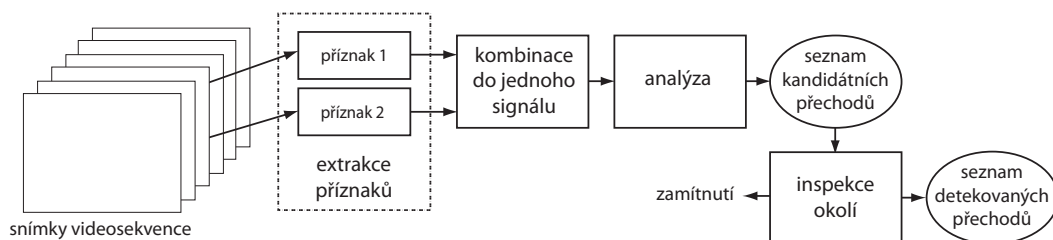
This paper presents a system for automatic detection of shot boundaries in video sequences. Detection is based on combination of visual feature tracking and color histogram comparison. Proposed approach is focused on hard cuts, but can also detect some gradual transitions.

1 ÚVOD

Detekce přechodů scén ve videu je činnost potřebná pro segmentaci videosekvencí na jednotlivé scény. Cílem této práce je navrhnout a implementovat systém realizující právě tuto detekci. Klíčová je však již samotná definice přechodu, protože například různé efekty přinášejí nejednoznačnost, co vůbec za přechod považovat. V tomto dokumentu zjednodušeně předpokládejme, že přechod je rozhraní mezi dvěma scénami a scéna je nepřerušovaná posloupnost snímků natočených jednou kamerou. Podle doby trvání lze rozlišit přechody skokové (stříhy) a postupné, přičemž stříhy jsou výrazně častější. Navržený systém se proto zaměřuje právě na detekci stříhů.

2 BLOKOVÉ SCHÉMA NAVRŽENÉHO SYSTÉMU

Schématický popis systému je zobrazen na obr. 1. Detekce je založena na odhadu míry podobnosti mezi sousedními snímky, ta je v místě přechodu nejmenší. Metrika podobnosti je získána sledováním průběhu hodnot příznaků, které jsou extrahovány z každého snímku videosekvence. Z těchto hodnot je následně vytvořen jediný signál, jehož analýzou je získán seznam kandidátních přechodů. Ty ještě mohou být zamítnuty na základě inspekce okolí.



Obrázek 1: Blokové schéma navrženého systému

3 EXTRAKCE PŘÍZNAKŮ

Z každého snímku jsou extrahovány dva různé příznaky založené na *barevném histogramu a sledování výrazných bodů*. Následuje podrobnější popis jejich získání.

Barevný histogram vyjadřuje rozložení barev v obraze. Použitý barevný prostor se rozdělí na N dílů a hledá se počet bodů, které jednotlivým dílům odpovídají. Pro barevný obraz se jedná o 3-rozměrný histogram. Počet dílů musí být dostatečně malý, aby se obraz zobecnil, experimentováním jsem dospěl k vhodnému počtu 8^3 – 16^3 dílů. Dále jsem ověřil, že detekce se výrazně neliší pro prostory RGB a HSV, ani při použití pouze některých kanálů (HS nebo SV).

Každá scéna je charakteristická množinou výrazných bodů. Dojde-li k přechodu, tyto body zmizí a objeví se jiné, příslušející nové scéně. Pokud budeme body sledovat, můžeme detekovat okamžik zmizení. Vhodné body poskytuje např. Harrisův detektor, pro samotné sledování je využit KLT-tracker. Ten se však nepříliš dobře vyrovnává se situací, kdy zmizí veškeré sledované body a poměrně velké množství z nich se snaží spojit s jinými body v nové scéně. Výrazné zlepšení přineslo použití Kalmanova filtru. Ten pro každý bod predikuje jeho pozici na základě chování v minulosti, čímž poskytuje dobrou výchozí pozici pro hledání, dále je nastaveno omezení vzdálenosti nalezeného bodu od predikce.

4 DETEKCE PŘECHODŮ

K popisu míry podobnosti mezi dvěma snímky je zavedena metrika, detekce přechodů probíhá analýzou jejího průběhu. Metriky jsou definovány pro každý příznak zvlášť a zkombinovány do jedné. Hodnoty jsou normalizovány do intervalu $\langle 0; 1 \rangle$, kde 0 odpovídá přesné shodě. Pro barevný histogram je podle [1] uvažována suma rozdílů odpovídajících dílů histogramů H . Histogramy musí být normalizované, aby výsledek spadal do požadovaného intervalu.

$$d_{\text{histogram}}[n] = \sum_i |(H_n(i) - H_{n-1}(i))| \quad (1)$$

Při sledování bodů je zohledněn počet ztracených bodů N^{out} z minulého snímku, a počet bodů nově objevených N^{in} ve snímku aktuálním. Jejich hodnoty jsou děleny celkovým počtem σ sledovaných bodů v odpovídajících snímcích, jako výsledek se bere maximum:

$$d_{\text{tracker}}[n] = \max(N_{n-1}^{\text{out}}/\sigma_{n-1}, N_n^{\text{in}}/\sigma_n) \quad (2)$$

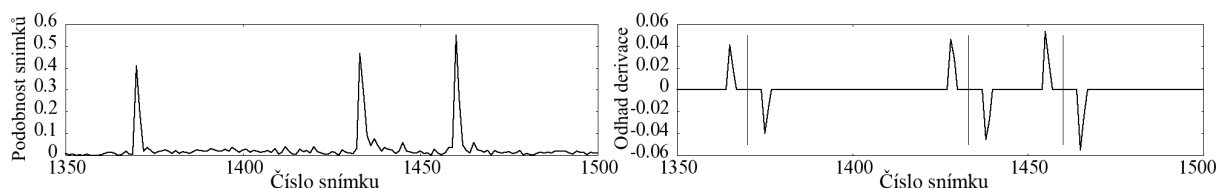
Výsledná metrika je lineární kombinací obou předchozích:

$$d_{\text{final}}[n] = \alpha \cdot d_{\text{tracker}}[n] + (1 - \alpha) \cdot d_{\text{histogram}}[n], \quad \alpha \in \langle 0; 1 \rangle \quad (3)$$

Váha α může být dána staticky, případně se může měnit například v závislosti na odhadu aktuální spolehlivosti detekce pro samostatné příznaky. V dále prezentovaných výsledcích je prozatím použita statická hodnota $\alpha = 0.5$, tedy vliv obou příznaků je shodný.

Průběh výsledné metriky pro jednotlivé snímky tvoří diskrétní signál, ukázka je na obr. 2. Přechody se objevují jako výrazné skokové změny (stříhy), případně plynulé změny trojúhelníkového tvaru (postupné přechody). Tento signál je vyhlazen a je z něj odhadnuta derivace. Přechod se v ní projeví jako dvojice impulsů, v polovině mezi nimi se pak nachází kandidátní přechod.

Barevný histogram způsobuje falešné detekce v případě výrazných jasových změn, např. odpálení blesku fotoaparátu. Kombinace s druhým příznakem tento jev zmírní, ale neodstraní. Proto pro každý kandidátní přechod se porovnávají barevné histogramy snímků vzdálených 2 snímky od jeho pozice v obou směrech. Je-li zjištěna vysoká podobnost, přechod je zamítnut.



Obrázek 2: Ukázka průběhu signálu a odhadu derivace po vyhlazení. Průběh obsahuje tři stříhy, jejich detekované pozice jsou vyznačeny svislými čarami.

5 VÝSLEDKY

Testování proběhlo na vlastní sadě manuálně anotovaných dat. Úspěšnost vyjadřují dvě veličiny: *Přesnost* je poměr počtu správně určených přechodů k počtu všech nalezených (mohou být nalezeny neexistující). *Úplnost* je poměr počtu správně nalezených přechodů k počtu všech obsažených ve videosekvenci. V testovacích datech bylo nalezeno více než 90 % přechodů, spolehlivost detekce je lepší pro stříhy. Shrnutí obsahuje následující tabulka:

Data	Počet přechodů	Bar. histogram		Sledování bodů		Kombinace	
		Úplnost	Přesnost	Úplnost	Přesnost	Úplnost	Přesnost
Zprávy	351 (299/52)	95,7 %	74,5 %	89,0 %	92,3 %	92,9 %	85,6 %
Sport	279 (198/81)	97,5 %	63,9 %	73,5 %	51,9 %	94,6 %	93,6 %
Seriál	412 (390/22)	99,3 %	93,8 %	99,3 %	94,2 %	99,0 %	97,6 %

Tabulka 1: Přehled úspěšnosti detekce pro jednotlivé příznaky a jejich kombinaci na různých testovacích datech. Počet přechodů je uveden jako "celkem (z toho stříhy/postupné přechody)".

6 ZÁVĚR

Byl představen systém pro detekci přechodů ve videu. Barevný histogram i sledování bodů poskytují vhodné příznaky, jejich kombinace zejména snižuje množství chybných detekcí. Nevýhodou je vysoká výpočetní náročnost daná počtem sledovaných bodů. Návrh byl proveden s ohledem na detekci stříhů, systém prokázal schopnost nalézt i některé postupné přechody. Cílem další práce by mohla být právě jejich lepší detekce, případně klasifikace do různých tříd.

PODĚKOVÁNÍ

Tato práce vznikla částečně za podpory grantu VUT FIT, FIT-S-10-2 a specifického výzkumu MSM0021630528.

REFERENCE

- [1] Mas, J., Fernandez, G.: Video Shot Boundary Detection Based on Color Histogram. Notebook Papers TRECVID2003, Gaithersburg, Maryland, NIST, October 2003