

色特徴量と動き特徴量を用いたショット分割手法

佐藤駿介[†] 青野雅樹[†][†]豊橋技術科学大学情報・知能工学系

1 はじめに

近年, YouTube 等のユーザ投稿型動画サイトの普及により Web 上で動画データが横溢し, 欲する動画を検索する要求が増大している. また, これらのサイトにおける違法アップロードが問題となっており, 動画の内容を検査し, アップロードされた動画が著作権を侵害するものかどうか判断する仕組みの導入要求が増大している.

YouTube に代表される動画サイトでは, 専らテキストを用いた検索が行われ, 検索対象は動画全体である. しかし, 動画は様々な内容を含み, 複数のシーンで構成されるため, ユーザの要求に応えるにはまだ不十分であると考え.

本研究ではテキスト以外の特徴量を用いたショット単位での動画検索システム作成のため, 動画をショットに分割する研究を行う.

2 関連研究

ショット境界検出の研究は, TRECVID という映像に関するワークショップで行われている.

Zhu ら[1]はショット切り替えの種類に応じて 6 つの検出器, 合計 104 素性の特徴量を提案し SVM を用いた学習を行い, ショット境界検出を行った. 検出結果は全ショット数 2317 に対して 95%であったと報告されている.

また, 鈴木ら[2]はフレーム間における動きの方向分布を利用したショット検出を行っている. 検出結果は全ショット数 1656(カットのみ)に対して 95%であったと報告されている.

A Method for Shot Boundary Detection from Videos Using Color and Motion FeaturesSyunsuke SATO[†], Masaki AONO[†]

[†]Dept. of Computer Science and Engineering, Toyohashi University of Technology
441-8580, Toyohashi, Japan

3 提案手法

3.1 色特徴量

色特徴量として隣接フレーム間のカラーヒストグラムの類似度を用いる. ヒストグラムの類似度の算出には Michael らのヒストグラムインタセクション[3]を用いた. ヒストグラムインタセクション D は次式で表される.

$$D = \sum_{i=0}^I \min(H_1[i], H_2[i])$$

H_1, H_2 は隣接するフレームそれぞれのカラーヒストグラムであり, $H[i]$ はヒストグラムの i 番目の要素の値, I はヒストグラムの次元数である. D の値が大きいほど隣接するフレーム同士は類似していることを示す.

本研究ではヒストグラムインタセクションをフレームの RGB 各成分, グレースケール変換したフレームの Gray 成分の 4 色についてそれぞれ計算する. ショット境界は D の値が閾値以上となる色の数がいくつあるかで判断する.

3.2 動き特徴量

動き特徴量としてオプティカルフローの方向分布, 大きさ, ベクトルの始点・終点の位置分布情報を用いる. オプティカルフローはフレーム間の特徴点追跡によって得る.

隣接フレーム間でショット切り替えがない場合, フローは一定の方向分布と大きさを持つ. 逆に隣接フレーム間でショット切り替えがある場合, 特徴点の追跡が不可能になるのでフローは現れないか, 一貫性のないフローが得られると考えられる. この違いを利用してショット分割を行う. また, カメラの前を物体が横切ったりした場合, 色特徴量のみでは誤検出してしまうが, この場合特徴点の並びと動きに一貫性があるはずなので誤検出を防止することも期待できる.

4 評価実験

4.1 実験概要

提案手法の有効性を確認するために評価実験を行った。データとして Hollywood Dataset[4]の内、テストデータ 884 動画を用い、これらの動画に対してショット分割を行った。分割されたショットに対し、正解と照合し Recall-Precision で評価を行った。正解におけるショット総数とショット種別を表 1 に示す。

表 1: データセット

	種類	ショット数
abrupt	Cut	2210
gradual	Dissolve	9
	Fade in/out	3
	Other	3

4.2 実験結果

評価実験の結果を表 2 に示す。αは閾値、βはDが閾値以上となる色の数である。実験結果より、閾値が高いほど再現率は上がるが、適合率は下がることが分かる。また、総合的に見ると閾値が 0.8、D が閾値以上となる色の数が 3 の時、F 値が 80%となり最も良い結果が得られた。

表 2: 実験結果

β	α	Recall	Precision	f-measure
1	0.85	95	58	72
	0.8	81	75	78
2	0.85	97	52	68
	0.8	89	72	80
3	0.85	98	48	65
	0.8	92	70	80
4	0.85	99	41	58
	0.8	95	66	78

ショット境界判別の成功例と失敗例を図 1、図 2 に示す。図 1 のようにフレーム間で色が異なる場合や、色が似ていても特徴点の追跡が行えないものは分割に成功している。しかし、図 2 のようにフレーム内で動く物体が多く、動きの方向も様々であり、色の変化もあるようなものは誤検出されてしまった。



図 1: ショット境界検出に成功した例



図 2: ショット境界検出に失敗した例

5 おわりに

本研究では、色特徴量と動き特徴量を用いたショット境界検出手法を提案した。実験の結果、F 値 80%の精度でショット境界検出が可能であることが分かった。しかし、ショット境界の種類によっては提案した特徴量のみでは十分でないとと言える。

今後の課題として、更なる特徴量の考案や他の色モデルの利用などが挙げられる。また、今後の展望として、分割されたショットに対する検索システムの作成を行いたい。

参考文献

- [1]Zhu Liu, Eric Zavesky, David Gibbon, Behzad Shahraray, Patrick Haffner:”AT&T RESEARCH AT TRECVID 2007”, TRECVID 2007
- [2]鈴木賢一郎, 中島正臣, 牧野鋭, 三部靖夫, 大塚作一:”動き方向ヒストグラム特徴を用いた映像データからのカット点検出法”, 電子情報通信学会論文誌, Vol.J86-D-II, No.4, pp.468-478,2003
- [3]M. J. Swain, D. H. Ballard:”Color indexing”, International Journal of Computer Vision, Vol. 7, No. 1, pp. 11-32, 1991
- [4]M. Marszalek, I. Laptev, C. Schmid:”Actions in Context”, IEEE Conference on Computer Vision & Pattern Recognition”, 2009