

映像におけるカット変わりの自動検出

6D-7 —フェードイン、フェードアウト、ディゾルブの場合—

村田 充弘†

中村 裕一‡

大田 友一‡

筑波大学

理工学研究科†

電子・情報工学系‡

1 はじめに

現在、種々の分野で動画像(映像)を扱うことが一般的になりつつある。しかし、映像は膨大な情報量をもっているため、蓄積する際に映像を構造化することが必要になる。その基本的な単位となるのが「カット」であり、映像の時間的構造は、カットの時系列を基本として考えることができる。したがって、構造化を行うためには、まず、カット変わりを検出することが必要であり、様々な研究がなされている[1][2][3][4][5]。しかしながら、従来の方法では、複数のフレームにわたって起こるフェードイン、フェードアウト、ディゾルブといったカット変わりを検出することが困難であった。本研究では、これらの時間的変化のモデル化を行い、カット変わりを精度良く検出する手法を開発した。

2 カット変わりのモデル化

映像は、「カット」の組み合わせで構成されている。ここで、「カット」とは、映像として連続したフレームの集まりであり、場面が変わることによって前のフレームとの連続性がなくなるフレーム、つまりカット変わりによって区切られる。主なカット変わりの模式図を図1に、その特徴を表1に示す。

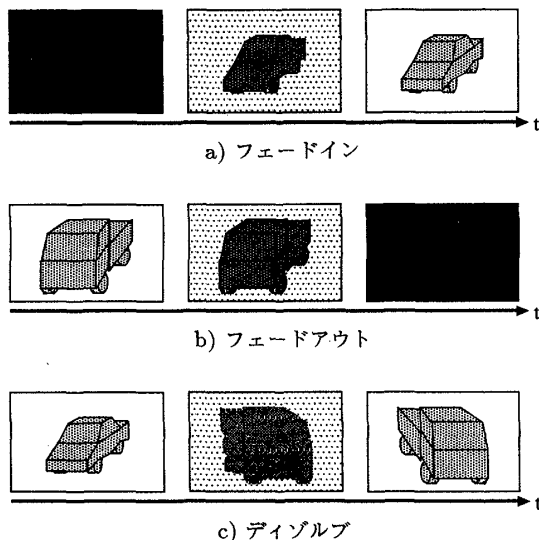


図1: カット変わりの模式図

表1: カット変わりの手法と特徴[6]

手法	特徴
瞬間的切換え	1フレームで画像全体を入れ換える。
フェードイン	暗い画面から画像を徐々に表す。
フェードアウト	画像を徐々に消して暗い画面にする。
ディゾルブ	画像をフェードアウトすると同時に次の画像をフェードインする。

瞬間的切り替えによるカット変わりでは、隣接フレーム間の類似性が低くなるところがカット変わりである。従来の手法では隣接フレームの画素値またはその他の特徴量の差分をとり、その差分が大きくなるところがカット変わりであると判定をしていた。

しかしながら、ディゾルブなどのように複数フレームにわたり徐々にカット変わりが起こる場合は、この手法では検出が困難である。

そこで、まず、ディゾルブのモデル化を行い、検出方法を示し、フェードイン、フェードアウトがディゾルブの特殊な場合として扱うことができることを示す。

2.1 ディゾルブのモデル化

まず、ディゾルブのモデル化を行なう。ディゾルブ前の画像を I_p 、ディゾルブ後の画像を I_n とすると、ディゾルブ中の画像 $I_{dolv}(t)$ (t は時間、 $t_0 \leq t \leq t_1$) は、

$$I_{dolv}(t) = (1-k)I_p + kI_n \quad (t_0 \leq t \leq t_1, k = \frac{t-t_0}{t_1-t_0}) \quad (1)$$

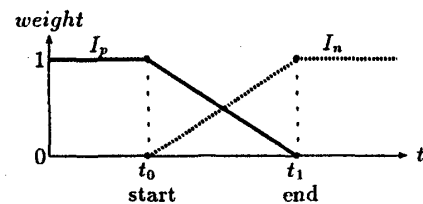


図2: ディゾルブのモデル

とすることができる(図2)。これを微分すると、

$$\frac{d}{dt} I_{dolv}(t) = \frac{I_n - I_p}{t_1 - t_0} \quad (t_0 \leq t \leq t_1) \quad (2)$$

より、時刻によらず一定値をとる。したがって、時刻 T_1 、 T_2 ($t_0 < T_1, T_2 < t_1$) における dI/dt の比の値 r は、常に1となる。

また、カット変わりは画像全体にわたって起こるので、 $r=1$ となる画素の多いフレームがカット変わりが起こっていると判定することができる。

2.2 ディゾルブの検出方法

まず、各フレームで $0.9 \leq r \leq 1.1$ なる画素数 R を求める。 R がピークを形成する部分を探し、ピークの開始点、終了点をそれぞれディゾルブの開始点、終了点とする。ピークとなるのは、ピークの高さを H_t 、ピークの右端の高さを H_r 、ピークの左端の高さを H_l とするとき、次のような値を満たすものとする (図 3)。

$$H_t/H_r > 1.2 \quad (3)$$

$$H_t/H_l > 0.85 \quad (4)$$

$$H_t > 100 \quad (5)$$

$$H_d < H_t/10 \quad (6)$$

また、ノイズの影響などでピークが単調でない場合を考慮し、 H_d 以下の変動は単調であるとみなす。

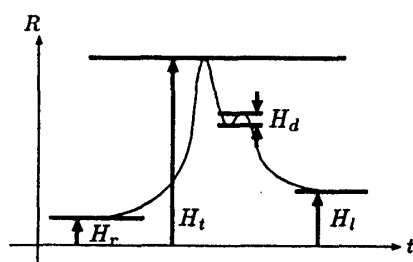


図 3: ピークの概形

2.3 フェードイン、フェードアウトのモデル化および検出方法

式 (1) において、フェードインは I_p の輝度値がすべて 0 の場合、フェードアウトは I_n の輝度値がすべて 0 の場合のディゾルブとすると、このモデルをそのまま適用することができ、ディゾルブの場合と同様の方法で検出することができる。

3 実験

3.1 実験方法

入力データとして、テレビ放映されている CM 映像を用いた。また、輝度値の差分を求める際の間隔は 5 フレームとした。また、ノイズによる影響を取り除くため、比を求める際に、 $\frac{d}{dt}I_{dslv}(T_1)$ 、 $\frac{d}{dt}I_{dslv}(T_2)$ のうちいずれか一つの差分値が小さい (< 10) 画素に対しては、比を計算しないようにした。

3.2 実験結果と考察

R の時間経過および検出されたディゾルブ (網かけ部分) を図 4 に示す。この実験で用いた映像中のディゾルブは全て検出された。また、テロップの表示によって過剰検出 (矢印の部分) がみられたが、これは開始フレームと終了フレームの差分値の和をしきい値処理することにより取り除くことができる。

4 まとめ

従来の方法では困難であったフェードイン、フェードアウト、ディゾルブといった複数のフレームにわたるカット変わりを検出する方法を提案した。実際の映像に対して実験を行い、精度良く抽出できることを示した。

参考文献

- [1] 長坂, 田中: “ビデオ作品の場面変わりの自動検出法”, 情処全大 40(前), 1Q-5, pp.642-643, 1990
- [2] 大辻, 外村, 大庭: “動画カット検出”, 情処論, vol.33, no.4, pp.543-550, 1992
- [3] 宮武, 吉澤, 上田: “フレーム相関係数の変化率に着目したカットの自動検出法”, 信学秋全大, D-299, 1990
- [4] 西尾, 小迫: “エッジ情報に基づく映像シーンの解析”, 信学秋全大, D-197, 1991
- [5] 岩成, 有木: “DCT 成分を用いたシーンのクラスタリングとカット検出”, 信学技報, PRU 93-119, 1994
- [6] 栗田, 高橋, 竹下: “画像のソフトウェア”, テレビジョン学会編, 画像エレクトロニクス講座 8, コロナ社, 1980

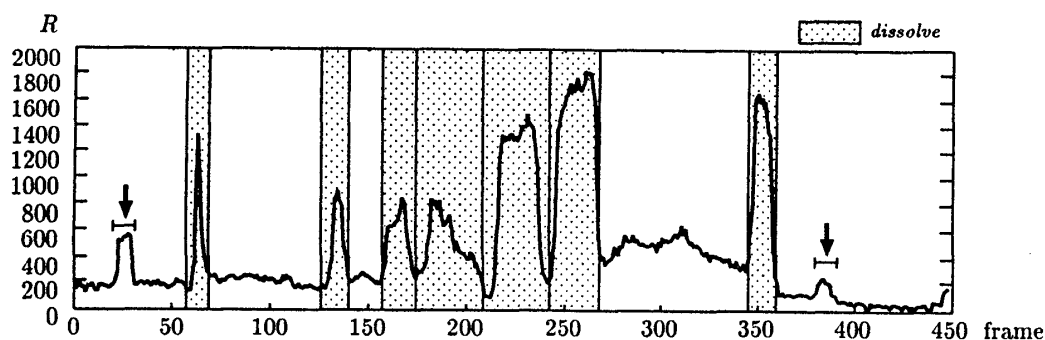


図 4: 実験結果