

# 動画像中のサブシーン記述のための動き推定法

4R-11

前田潤治

日本アイ・ビー・エム(株) 東京基礎研究所

## 1 はじめに

動画像データベースや動画像編集システムにおいて、膨大な量のデータを扱うために、動画中のシーンを自動検出し、これを単位として管理する手法が提案されてきた。本研究ではさらにシーン内部の構造に注目し、内容の動きをもとにして「サブシーン」という単位を導入する。動き情報は被写体の動きとカメラ操作に大別できるが、前者の抽出は現状では困難である。したがって、本稿では後者のみを扱う。

現在主流となっているカメラ操作推定法はサブシーン記述という目的に適していない。そこで本稿ではこの目的に適したカメラ操作推定法もあわせて提案する。

## 2 動きをもとにしたシーンの分割

シーン構成は撮影者、編集者の意図を直接反映しているため、これを動画像を管理する単位として用いることは有効であるが、さらにきめ細かな管理を行うために、カメラ操作の情報は内容把握や検索の大きな助けとなる。たとえば、カメラ操作を動画像のインデックスとして利用する方法 [1] などが提案されている。

そこで本稿ではカメラ操作によってシーンを分割し「サブシーン」という構造を導入する。サブシーンとは、シーンの部分集合であり、その内部では同一のカメラ操作が起きているようなものであるとする。図1にサブシーンの例を示す。

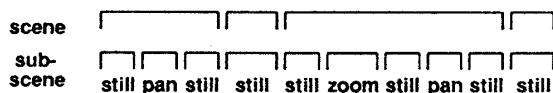


図1: サブシーンの例

## 3 サブシーン記述のためのカメラ操作推定法

現在までに、動画像を解析してその中のオペティカル・フローを求め、その結果からカメラ操作を

“Motion Estimation for Describing Sub-scenes in Video Sequences”

Junji Maeda

IBM Japan, Tokyo Research Laboratory,  
1623-14, Shimotsuruma, Yamato, Kanagawa, 242, Japan

推定する方法は数多く提案されている。それらは基本的に二枚のフレームの間の動きを計算する方法であり、動画像として意味のある情報にするためにはそれらの動き情報をまとめてさらに解析しなければならない。このアプローチは、動きをかなり正確に求められるという利点がある反面、雑音に弱いという欠点がある。雑音の例としては、手ブレによって撮影者の意図していない動きをビデオカメラがする場合や、写真機のフラッシュによってあるフレームだけ前後のフレームと全く違う明度、色あいに写る場合などがある。特に手ブレは三脚等を使わずに撮影されている場合には避け得ないもので、これは解決されるべき重要な課題である。

本研究の眼目は、動画像の中のカメラ操作を雑音の影響を受けないように推定することである。その際、動きの定量的正確さ（どの期間にどの方向に何画素分動いたか、等）は追求しない。本研究が対象としている動画像データベースや動画像編集システムでの応用には定性的な情報（どの期間に右へ動いた、等）で十分であり、定量的な情報はかえって冗長となる。

本稿で提案するカメラ操作推定法の要素技術は、二次元時空間画像に対するテクスチャ解析である。一辺を画面中の線分、他の一辺を時間軸上の線分（ある幅の時間一期間）として作った二次元時空間画像の方向性を調べることで、その画面中の線分の内容がその期間にどう移動しているかを知ることができる。この技術を用いてカメラ操作を推定するには、図2のように画面上に求めたい動きの向きに線分を置き、その線分上の動きを調べればよい。

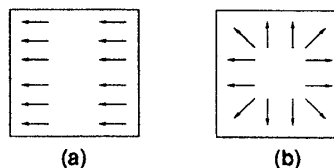


図2: カメラ操作推定の例

たとえば(a)のように水平に置いた線分の上の内容が一様に左へ動いていけば、カメラ操作は「右へのパンニング」、(b)のように放射状に置いた線分の上の内容が一様に外に向かって動いていけば、カメラ操作は「ズームイン」とであると推定できる。

上述の技術を用いて、シーンをサブシーンに分割する手続きは以下のようになる：

0. 動画のシーン切替え点は既存の手法(例えば[2])によって求められているものとする。
1. 画面上の、動きを求めたい位置・方向に適当な数の線分を置く。
2. シーン切替え点から、フレーム列を適当な一定の時間幅に分割する。分割されたそれぞれのフレーム列を「ピリオド」と呼ぶ。
3. 1. で設定された線分が2. の時間幅で作るそれぞれの時空間画像を解析し、各ピリオド中のカメラ操作を推定する(図3)。
4. 一つのピリオドの中に複数の異なったカメラ操作が含まれている場合には、必要に応じてピリオドを分割し、分割されたそれぞれを新しいピリオドとして3. を繰り返す。
5. 隣接する二つのピリオドを比較し、カメラ操作が同じであればそれらをまとめて一つのサブシーンとし、カメラ操作が異なれば別のサブシーンとする。

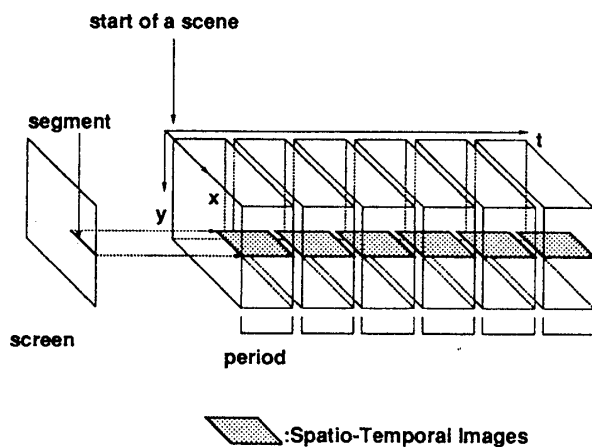


図3: 時空間画像解析

#### 4 実験

二次元時空間画像の方向性を求めるテクスチャ解析手法として、フーリエ変換を使用した。データとして、国立民族学博物館の「アンデス民族

学ビデオテーク」の番組の一部を使用した。所期の目的に対して、

- 手ブレのある静止区間も正確に静止区間と判断できた。
- 写真機のフラッシュがある一フレームに含まれていても正確な推定が行なえた。
- 人間の目で見てもカメラ操作に変化のあった箇所は100%検出できた。
- 画面中に大きな面積を占める被写体の動きなどによって、過検出(実際にはカメラ操作は変化していないのに、変化したと判断してしまうこと)が検出箇所全体の約30%あった。

#### 5 まとめ

動画の柔軟な管理のために、動き情報をもとにした「サブシーン」という構造をシーンの内部に導入した。動きを求める既存の方法はサブシーン記述に不適であるため、この目的に適した技法を提案した。実験により、本技法の有用性が確認された。

実験を通して過検出の問題が挙げたが、これに対して二種類の対策が考えられる。まず、動画全般に対する共通の対策として、画面上の線分をより密に様々な方向に置くことがある。また、スポーツ番組など特定の種類の動画は、その種類に特有の画面構成、カメラ操作がある場合が多く、それに適応した線分の配置をすることも有効であると考えられる。

#### 6 おわりに

本研究は国立民族学博物館(民博)と日本アイ・ビー・エム(株)との共同研究「博物館におけるマルチメディアの有効利用」の一環であり、実験に使用されたデータは民博に所属する。ご指導いただいた民博の杉田繁治教授、データの準備などのご協力をいただいた民博の鈴木明、中川隆、宇治谷恵の諸氏に感謝いたします。

#### 参考文献

- [1] 阿久津他、「動画画像インデクシングを目的としたカメラ操作の規定方法」電子情報通信学会論文誌 D-II, Vol. J75-D-II, No. 2, pp. 226-235, 1992.
- [2] Ioka, M., "Method of detecting scene change in moving picture," IBM Technical Disclosure Bulletin, Vol. 34, No. 10A, Mar. 1992.