

## 高信頼動き推定方式に基づくショット切り替え検出と切り替え効果認識 Shot Change Detection and Shot Change Effect Recognition Based on Reliable Motion Detection Method

青木 恭太  
Kyota Aoki

### 1. まえがき

映像処理の基本として、連続した映像からショットへの分割がある。単純な瞬時カットは、各種方式で検出されるが、より凝った切り替え効果が用いられている場合には、切り替え効果の開始点および終了点を検出し、さらに切り替え効果のタイプを識別することが、以降の処理を容易にするために必要である。

スポーツ映像などでは、競技者と競技場所が主な撮影対象となる。複数のカメラを用いて撮影すると同一の競技者を類似の背景で撮影することとなる。また、激しい動きが存在し、それに伴いカメラワークも激しいものとなる。このスポーツ映像から類似のプレイのショットを検索しようとする、色彩やテクスチャの類似した映像をショットに分割することが必要となる。

ショット切り替え検出方式としてよく利用されている方式[5]では、ショット境界においては隣接フレームの色情報が急激に変化する性質を用いる。DVEを含むショット切り替えを認識する方法として、動きベクトルを基本とする方式が提案されている[1]。これは、継続表示されている領域の動きベクトル分布と前後のショットの境界領域の動きベクトル分布が異なることに注目しており、色彩やテクスチャの類似した映像においてもショットへの分割が可能である。しかし、ショット内にカメラワークを含む場合や激しく動く対象物が撮影されている場合には、動きベクトルの変化は多様となり、ショット切り替えの認識が困難となる。

筆者らが提案している高信頼動き推定方式[2]では、動き推定ブロックにおいて動き推定の可否が得られる。動き推定の可否が、カメラのパンやチルト、オブジェクトの移動により受ける影響は、動きベクトルが受ける影響に比較して小さい。本論文では、高信頼動き推定方式を用いることで動き推定不能画素の分布を推定し、動き推定不能画素の分布から激しい動きやカメラワークを含む映像においてもショット切り替え位置の検出と切り替え効果の認識を行う方式を提案し、実験によりその有効性を検証する。

本論文は、2. でショット境界の推定と切り替え効果の認識方式について述べる。さらに、3. で提案方式の有効性を実験により確認し、最後にまとめを述べる。

### 2. ショット境界検出方式

#### 2.1 ショット境界

カメラで連続的に撮影された1連のフレーム系列がショットである。映像作品は、複数のショットを組み合わ

せている場合がほとんどで、ショットとショットの切り替えが起こる。

編集された映像作品では、ショットの境界に1フレーム推移で前ショットから後続ショットに切り替わる瞬時カット以外の各種の切り替え効果が付与される場合がある。本論文では切り替え効果の名称として、MovieMaker[3]で使用されている名称を用いる。

#### 2.2 提案方式

ショットが切り替わる時、fade以外では、連続するフレームにおいて前フレームには存在し、後フレームには存在しない領域が生じる。この領域は、ショット内の連続したフレームにおける動きオブジェクトによる隠蔽と同様に前フレームに対応する領域が後フレームには存在しない。この画像組に対して高信頼動き推定方式[2]を適用すると、対応する領域が存在しない場合には、動き推定不能が得られる。一方、オブジェクトの動きやカメラワークによる変動では、多くの領域で動き推定可能であり、動きベクトルが得られる。

本論文では、上記の動き推定の可否を用いて、ショット切り替えを検出・認識する方式を提案する。提案方式全体の処理の概略の流れを図1に示す。提案方式では、映像から得た連続するグレースケールフレームにおいて高信頼動き推定方式を用いて動き推定を行い、動き推定不能領域を検出する。さらに画素単位動き推定方式[4]を用いて動き推定不能画素を決定する。動き推定不能画素は、ショット切り替えやオブジェクトの動きによる隠蔽領域の変動などにより前フレームの画素に対応する画素が後フレームに存在しないことにより生じる。そこで、この動き推定不能画素を切り替え候補画素とみなす。切り替え基準画像を図2左、同一の切り替え効果を持つ実切り替え画像から得た切り替え候補画素画像を図2右に示す。図2において、切り替え候補画素は、白画素である。各連続フレーム対から得られた切り替え候補画素画像系列から直接ショット切り替え検出と切り替え効果認識を行うことも可能[1]であるが、処理量を削減するために切り替え候補画素画像を $1/n$ に縮小して、切り替え候補画素密度画像を得る。本論文において示す実験では、 $720 \times 480$ の映像を用いており、切り替え効果の継続時間は1S(30フレーム)である。このとき、フレームを上下方向に2分する切り替え効果では、 $480/30/2 = 8$ 画素だけ隣接フレームで切り替え境界が移動する。このとき、 $6 \times 6$ のブロックにおいては、最低でも66%の画素が切り替え境界の移動により前後のフレームで切り替え候補となり、 $1/6$ に縮小後も容易に切り替え候補として認識可能である。 $1/8$ に切り替え候補画素画像を縮小すると最悪の場合、縮小後に1画素となるブロック中の50%の画素しか切り替え候補画素とならず、縮小後の切り替え候補認識が困難となる。

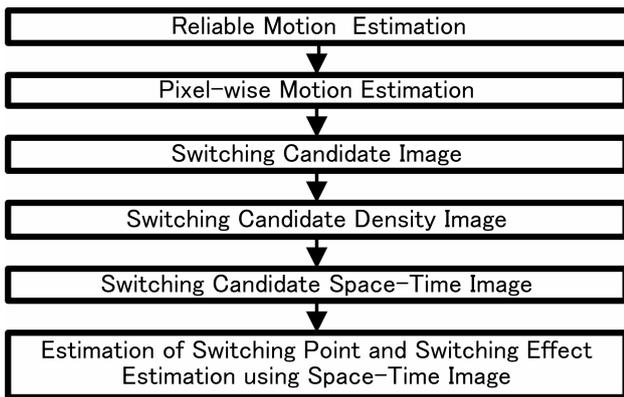


図1 提案方式の処理の流れ

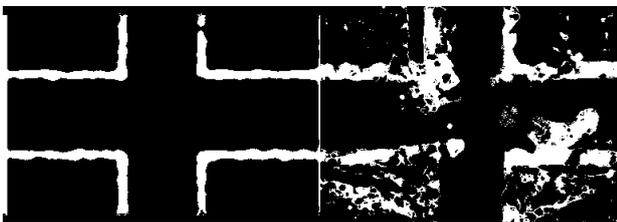


図2 切り替え候補画素画像

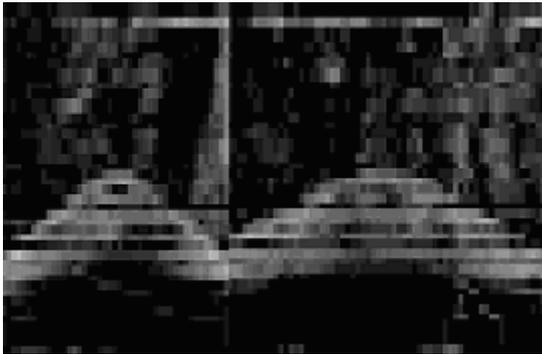


図3 切り替え候補画素時空マップ



図4 切り替え効果 split の切り替え候補画素画像

そこで、本論文で示す実験では、1/6 に縮小する。さらに、切り替え候補画素密度画像の XY 各方向への投影分布を作成し、投影分布の時系列において切り替え点の探索と切り替え効果の認識を行う。連続フレーム対から得られる XY 各方向への投影分布を時間軸方向に並べることにより切り替え候補画素密度投影分布の時系列をえる。これを切り替え候補画素密度時空マップと呼ぶ。切り替え

候補画素密度時空マップは、空間方向 1 次元、時間方向 1 次元の 2 次元データとなり、通常の画像でオブジェクトを探索するのと同様に切り替え点の探索と切り替え効果の認識が可能となる。図 3 に切り替え候補画素密度時空マップ例を示す。ここで、右部分は、X 軸への投影分布を 90 度回転させたもので、右部分は Y 軸への投影分布である。また、上から下へ向かって時間が進む。

ショット切り替えは、かならずフレーム全域を対象とするので、切り替え候補画素密度時空マップを探索する際には、時間方向のみ探索すれば十分である。切り替え効果を認識する際に用いる基準切り替え候補画素密度時空マップは、ランダム輝度静止画 2 枚に切り替え効果を適用した映像から得る。

ショット内でカメラワークによりフレーム全体が移動している場合や、フレーム内で対象となるオブジェクトが移動している場合でも、多くの領域で動き推定可を得るので、ショットの切り替えによる動き推定不能領域の検出によりショット切り替えを検出する提案方式は、誤った切り替えの検出を抑制可能である。

### 3. 実験と評価

#### 3.1 切り替え位置検出

映像 stennis[5]を MovieMaker により 720×480 に拡大した画像の全区間を提案方式で処理し、切り替え候補画素密度を調査した。映像 stennis は、継続時間 10S で瞬時カット 2 個で区切られた 3 ショットより構成されている。この映像は、88 フレームまでズームアウトし、88 フレームと 89 フレームの間で瞬時カットにより切り替わる。この映像を対象として切り替え候補画素密度を求めるとショット切り替え部分を除いた切り替え候補画素密度の平均 4.5%、標準偏差 4.4%が得られた。切り替え区間推定の閾値として平均+3 $\sigma$ を用いると、切り替え位置候補として、4 箇所を発見する。このうち、5 フレーム近傍累積切り替え候補画素密度が 1 を越える位置は、正しい 2 箇所のみである。

切り替え位置検出実験は、eye, split, checker, iris, shutter, wipe, push, slideup, slidein など 12 種の切り替え効果を用いた。映像 stennis の 88-89 フレーム間の瞬時カットを MovieMaker により継続切り替え効果に変更した映像を作成し、前記実験と同一の閾値を用いて 12 映像で切り替え位置を検出すると、12 映像すべてで正しく切り替え位置を推定する。このとき、再現率 100%、誤差率 0%となる。5 フレーム近傍累積切り替え候補画素密度による検査を行わないと、切り替え前ショットのズームアウトのために、切り替え位置以外で切り替え位置と推定される。

前記実験とは別に、野球中継テレビ放送の DVD 録画による 720×480 約 21 分(37405 フレーム)の映像において切り替え点の検出率と継続ショット内のショット継続検出率をほぼ同一となるように閾値を調整して切り替え点推定実験を行った。この映像において、切り替え点は 119 点であった。提案方式は、切り替え点 114 点を検出し、5 点が未検出であった。また、1216 点で切り替え点であると誤検出した。このとき、切り替え点検出の正解率 96.4%で、再現率 95.8%、適合率 8.6%となる。未検出点のうち 3 箇

所は、Fade による切り替えが行われており、提案方式では検出困難である。残り2点のうち1点は、切り替え前ショットと切り替え後ショットが非常に似ており検出困難である。他の1点は、切り替え前ショットにおけるカメラワーク(パン)の影響により、ショット切り替えが検出できなかった。

対照実験として、ショット切り替え検出方式としてよく利用されている文献[7]のショット境界においては隣接フレームの色情報が急激に変化する性質を用いる方式を用いて同一映像に対してショット切り替え検出実験を行った。最良の結果を得る閾値を用いてショット切り替え検出を行うと、切り替え点 110 点を検出し、9 点が未検出であった。また、2297 点で誤って切り替え点と誤検出した。このとき、切り替え点検出の正解率は、93.3%で、再現率は 92.4%、適合率は 4.6%となる。表 1 に実験結果をまとめて示す。

提案方式は、色情報を用いていないにも関わらず、対照実験と比較して、切り替え点の誤検出を抑制することにより約2倍の適合率を得ており、激しいカメラワークをとまなうスポーツ映像では高い性能を示している。

### 3.2 切り替え効果認識

切り替え点検出に用いた継続切り替え効果を付与した映像を用いて、切り替え効果認識実験を行った。実験結果の一部を表 2 に示す。表 2 において、各列は列上部の切り替え効果を施した実験映像から得た切り替え候補画素密度時空マップの切り替え効果付与部分と最左列に記載した各基準切り替え効果候補画素密度時空マップ(基準マップと略記)の画素単位の相関の平均である。切り替え効果継続時間は一定であり 1S である。切り替え効果 split および push のテスト映像以外では、同一切り替え効果基準マップとの相関が他の基準マップと比較して 20% 以上大きく雑音にも頑健であると期待される。

切り替え効果 split は、図 4 に例示した切り替え候補画素画像のように中央から左右に後フレームが出現する。切り替え画素の割合が他の切り替え効果に比較して小さく、雑音が重畳することにより他の切り替え効果に類似する。iris は、図 2 にその切り替え候補画素画像を示した切り替え効果であり、中央から十字状に後フレームが出現する。時空マップにおいては、X 軸方向に投影した部分では、類似している。両者の差異は、時空マップ上では Y 軸方向に投影した部分のみに現れ、相関の差異が小さくなる。

wipe と push は、図 5 に示すように、基準マップが類似しており、本質的に識別困難である。これを識別する際には、提案方式で分類後、切り替え候補画素密度画像の照合により類別することが必要である。

### 4. むすび

本論文のショット境界検出および切り替え効果認識方式は、フレーム間で動き推定が可能か否かを基本としており、色情報に基づく方式や動きベクトルを用いた方式と比較して、ショット内のカメラワークなどの影響に頑健であり、オブジェクトの動きやそれに伴うカメラワークが大きいスポーツ映像などでも利用可能である。

表 1 野球中継映像切り替え点検出実験(%)

	正解率	再現率	適合率
提案方式	96.4(+3.1)	95.8(+3.4)	8.6(+4.0)
既存方式	93.3	92.4	4.6

表 2 切り替え効果認識  
(基準マップと実観測マップの相関)

Test image	eye	split	iris	wipe	push	slidein
eye	0.58	0.38	0.36	0.18	0.19	0.17
split	0.35	0.46	0.33	0.21	0.19	0.18
iris	0.38	0.41	0.65	0.07	0.22	0.16
wipe	0.31	0.42	0.19	0.61	0.39	0.23
push	0.31	0.28	0.29	0.39	0.42	0.27
slidein	0.24	0.22	0.19	0.30	0.29	0.43

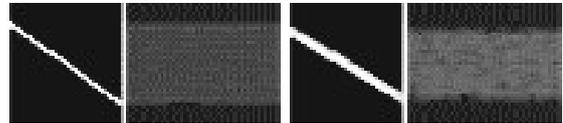


図 5 切り替え効果 push と wipe の切り替え候補画素密度時空マップ

色情報の変動によりショット切り替えを認識する方式と比較すると非常に激しいカメラワークを含む野球中継映像では、適合率が 1.87 倍となり提案方式の大きな動きにも頑健であることを確認した。

今回の実験映像では、切り替え時間は一定である。多くの映像でも切り替え時間は一定である。切り替え時間が増える場合には、編集距離などを用いて照合を行うことで切り替え時間の変化に対応することが必要である。また、多種映像における切り替え効果認識能力の検証は、今後の課題である。

### 文献

- [1] 岩本浩太, 平田恭二, “ワイプ・DVE のパターン非依存な検出手法”, FIT2006, LI-014, pp.211-214, Sep, 2006
- [2] 河合吉彦, 馬場口登, 北橋忠宏, “放送型スポーツ映像におけるデジタルビデオ効果に着目したリプレイシーン検出の一手法”, 進学論, D-II, J84-D-II, 2, pp.432-435, Feb, 2001.
- [3] 青木恭太, “大動き量まで適用可能な圧縮雑音に頑健なブロック単位高速高信頼動き推定法”, 進学技報, Vol.106, No.536, pp. 95-100, Mar, 2007.
- [4] 青木恭太, “平坦領域を考慮した多重動き推定に基づく画素単位動き推定”, 進学技報, Vol.106, No.423, pp. 7-12, July, 2006
- [5] 長坂 晃郎, 田中 譲, “カラービデオ映像における自動索引付け法と物体探索法”, 情処学論, Vol.33, No.4, pp.543-550, Apr, 1992
- [6] MicroSoft Corporation, “Microsoft(R)Windows ムービーメーカーVersion5.1”, c1981-2001.
- [7] <http://sampl.ece.ohio-state.edu/>