

H-009

高信頼動き推定に基づくショット切り替え検出および多種類切り替え効果の認識 Shot Change Detection and Recognition of Large Kinds of Shot Change Effects Based on Reliable Motion Detection

茂呂駿介†
Shunsuke Moro

青木恭太†
Kyota Aoki

1. あらまし

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。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。映像作品では、複数のショットを組み合わせている場合がほとんどで、その際にはショットの切り替えが起こる。切り替え効果の名称として、Windowsムービーメーカー[3]で使用されている名称を用いる。

単純なスイッチは、各種方式で検出されるが、数10フレーム程度継続してショットが切り替わっていくような、より凝った切り替え効果（例として「スライドイン」、「フェード」等）が用いられる場合もある。このような凝った切り替え効果を「継続切り替え効果」と呼ぶことにする。継続切り替え効果が用いられている場合には、切り替え開始点および切り替え終了点を検出し、さらに切り替え効果の方式を識別することが、以降の処理を容易にするために必要である。

高信頼動き推定方式[1]では、動き推定の可否が得られるので、前後のショット間の関連の消失を直接観測することが可能である。本研究では、高信頼動き推定方式を用いることで動き推定不能領域の分布を推定し、この動き推定不能領域の分布からショット切り替えの検出と切り替え効果の認識を行う方式を提案し、実験により多種類の切り替え効果に対する提案方式の能力を検証する。

2. 処理方式の概要

ショットが切り替わるとき、一部の切り替え効果を除き、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できる。この領域は、1ショット内の連続したフレームにおけるオブジェクトによる隠蔽と同様に、前フレームに存在し後フレームには存在しない領域となる。このような連続2フレームに対して一般的な動き推定方式で動き推定を行うと、誤った動きを推定する。一方、高信頼動き推定方式など一部の動き推定方式では、隠蔽などの各種の原因で前後のフレームに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。

映像中のショットが切り替わっている最中の連続2フレームに対し高信頼動き推定を行い、さらにその結果に対して画素単位動き推定[2]を行うと、ショット境界付近に動き推定不能画素や動き量が非常に大きいと推定される画素が多く分布する傾向にある。動き推定不能画素の密度や分布のパターンを調べることで、切り替え位置の検出や切り替え効果の種類の認識を行う。

3. 処理手順

処理手順を図1に示す。提案方式では、まず高信頼動き推定方式を用いて実験対象映像の各連続2フレームに対して動き推定を行い、さらに画素単位動き推定を用いて、画素単位に動き推定の可否を決定する。そして、動き推定不能画素および動き量が非常に大きいと推定された画素を切り替え候補点とする2値画像を作成する。これを切り替え候補点画像と呼ぶ。

切り替え候補点画像を1/6の大きさに縮小して切り替え候補点縮小画像を得る。切り替え候補点縮小画像の縦横各方向の輝度投影分布を作成し、それを時間順序につなぎ合わせることで、切り替え候補点密度分布1次元、時間1次元の2次元画像を作成する。これを切り替え候補点密度時空マップと呼ぶ。

時空マップの各時間の切り替え候補点の密度を調べ、密度が平均よりも十分高い区間を、ショット切り替えが起きている区間と推定する。

時空マップは、適用されている切り替え効果の種類によって特徴的な形を示す。各種切り替え効果において見本となる時空マップを作成し、対象映像から得た時空マップと見本時空マップ群との相関を取り、最も相関の強い切り替え効果を、対象映像で適用されている切り替え効果であると認識する。

4. 実験と評価

実験対象映像として、野球中継の映像から2ショットを切り出し、Windowsムービーメーカーで各種切り替え効果を含む映像を作成した。前ショットでは、カメラワークとしてパンを使用している。また、前ショットと後続ショットには大きく動くオブジェクトが存在する。切り替え効果の種類は、ムービーメーカーで用意されていた全60種類を使用した。前ショットと後続ショット中のフレームを図2に示す。図2において、(a)は前ショット、(b)は後続ショット中の1フレームである。

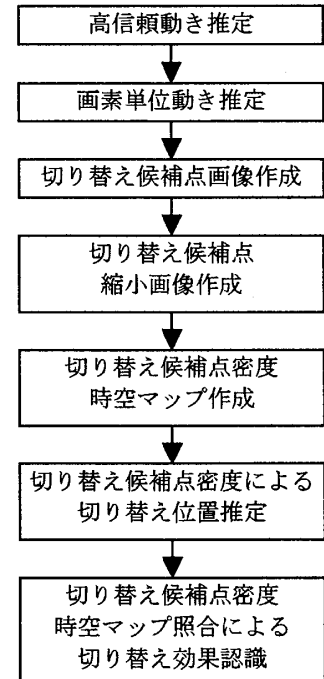


図1：処理手順

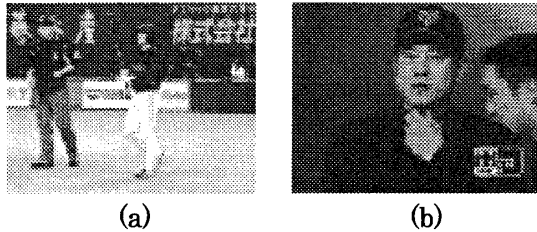


図2: 前ショットと後続ショット

見本時空マップは、平均輝度の異なる2枚のノイズ画像から、ムービーメーカーを用いて各種切り替え効果を含む動画を作成し、対象映像から時空マップを作成するのと同様の手順で処理して作成した。

対象映像中のショット切り替えが起こっている最中のフレームと、切り替え候補点画像の例を図3に示す。図3において、上段が対象映像中のフレーム、下段が切り替え候補点画像である。切り替え候補点画像において、白色の画素が切り替え候補点である。

対象映像から得た切り替え候補点密度時空マップと見本時空マップの例を図4に示す。図4において、上段が対象映像から得た時空マップ、下段が見本時空マップである。

実験結果を表1に示す。60種類の対象映像から得た時空マップのうち、38種類は同種の切り替え効果の見本時空マップとの相関が最も強いという結果が得られ、正しく認識した。また、2〜3番目に相関が強いと認識された対象映像が9種類あった。相関の強さが4番目以下になった対象映像が13種類あり、これらは認識失敗とみなせる。

「フェード」や「ディゾルブ」など、7種類の切り替え効果では、ショット切り替えの際に明確なショットの境界線が現れない。また、「チェッカー」と「円(複数)」では、切り替えの際に映像全体に細かい境界線が多く現れる。これらの切り替え効果では、見本時空マップに明確な形が現れず、相関を取った際に誤った認識がされる場合が多かった。これらの切り替え効果で誤った認識を減らせるような見本時空マップ作成手法を検討する必要がある。

切り替え効果のうち、時空マップを作成した際にそれぞれ似た形の時空マップが得られるものが4組あり、これらの切り替え効果では認識に失敗する場合があった。これらの切り替え効果で誤った認識を減らすために、時空マップをその形から数種類のタイプに分類し、対象映像から得た切り替え候補点密度時空マップと見本時空マップとの相関を取った際に、まずどのタイプに属するかを判断し、その後付加的な識別を行って切り替え効果の種類を正確に認識する方法が考えられる。

6種類の切り替え効果では、対象映像から得た時空マップと見本時空マップの形は、目で見た限りでは類似しているが、認識に失敗した。

表1: 相関を取った結果

最も強い	2番目	3番目	4番目以下
38種類	6種類	3種類	13種類

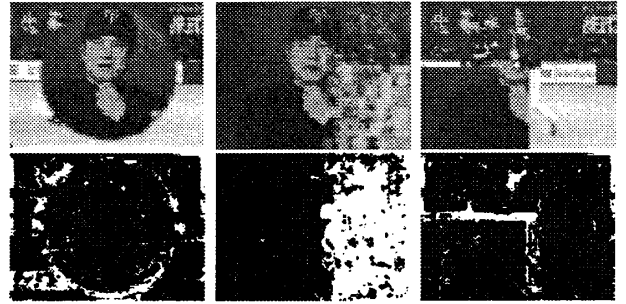


図3: 対象映像と切り替え候補点画像



図4: 時空マップ

5. むすび

本研究では、高信頼動き推定に基づくショット切り替え検出および切り替え効果の認識方式を提案し、実験により提案方式の切り替え効果認識の能力を検証した。

提案方式は、フレーム間で動き推定が可能か否かを基本としており、既存方式と比較してショット内のカメラワークに頑健であると思われる。

今回の実験映像では、切り替え時間は一定であり、切り替え効果継続時間が変化する際には、編集距離などを用いて照合を行う必要があり、本実験の結果がそのまま一般的映像における提案方式の能力を示しているわけではない。

切り替え効果継続時間の異なる映像に対応できる手法の提案や、既存方式との能力比較が今後の課題である。

参考文献

- [1] 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信頼動き推定法”, 進学技報, Vol.106, No.536, pp.95-100, Mar, 2007.
- [2] 野辺昌史, 猪野裕司, 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(20061207), pp.7-12, July, 2006.
- [3] Microsoft Corporation, “Microsoft(R)Windows ムービーメーカー Version 5.1”, (C)1981-2001.