

スポーツ中継映像におけるRGBヒストグラムとDCTを用いたショット分割

Shot Boundary Detection from Broadcast Sports Video using RGB Histogram and DCT

藤原 雅樹[†]
Masaki Fujiwara

泉 正夫[†]
Masao Izumi

福永 邦雄[†]
Kunio Fukunaga

1. まえがき

映像の保存、解析、インデクシング等における基本的な処理の一つにショット分割がある。ショットとは同一のカメラで撮影された一連の時系列画像を言い、中継カメラ切替え時や編集作業において、別のショットと切替えられる。ショットの切り替わる時点を映像から抽出する作業がショット分割であり、これまで様々な手法が提案されている。それらの手法は大きく2種類に分類できる。1つは分割点において輝度値が大きく変化することを用いる手法 [1] であり、もう一方は周波数領域の変化を用いる手法 [2][3] である。

一方、スポーツ中継映像のインデクシングは大量の映像アーカイブの中から目的のシーンを探索する際に有用であり、その自動化に対する研究 [4] [5] も数多くなされている。これらの手法のほとんどはショットの種類を元に解析を行なっている。すなわち映像の解析を行なうためには、まずショット分割をする必要があるが、スポーツ中継映像は、移動する選手やボールなどを追いかけて撮影することが多く、同一ショット内でも激しく映像が変化するため、従来のショット分割手法では十分な精度を得られない場合がある。また、スポーツ中継映像にはワイプ、ディゾルブ、CGなどの特殊境界が多用され、特に激しく変化するショットの前後に特殊境界が存在する場合、従来のショット分割手法では困難な場合がある。そこで、本稿では輝度値の変化に基づく手法と周波数領域の変化に基づく手法を組合せ、さらに特殊境界においても精度良くショット分割できる手法を提案する。

2. 輝度値の変化に基づくショット分割

RGBやHSI等の色空間での輝度値変化に基づく手法 [1] が数多く研究されており、例えばRGBヒストグラムを用いた境界検出では、異なるショット間には色合いに大きな差が現れると考えられるため、隣接するフレーム間のRGB値を比較することによってショット境界を検出する。各フレームにおいてR、G、Bそれぞれのヒストグラムを作成し、1つ前のフレームとの差分を取る。隣接するフレーム間のヒストグラムの差(式1)を比較し、その値が閾値 T 以上であるならば、現在のフレームと前のフレームでショットが切り替わっていると判定する。

$$\frac{1}{3MN} \sum_{v=R,G,B} \sum_{n=0}^{255} \{h_{v,f}(n) - h_{v,f-1}(n)\} > T \quad (1)$$

ここで、 v はR、G、Bそれぞれの色、 n は色の値域、 f はフレームの番号を表す。 $h_{v,f}(n)$ は f 番目のフレームにおける v 色の値が n となる画素数である。上記の式を満たす場合、 f 番フレームと $f-1$ 番フレームの間にはショット境界が存在すると判定する。

[†]大阪府立大学大学院 工学研究科, Graduate School of Engineering, Osaka Prefecture University

3. 周波数領域の変化に基づくショット分割

DCTクラスタリングによるカット検出手法 [2][3] は、同じショットに属する隣接フレーム同士は画像の構成が類似しているという性質に基づき、連続するフレームのクラスタを形成することによって、形成された2つのクラスタを分割する点をショット境界として検出する手法である。変換式は式(2)で表される。

$$F(u, v) = \frac{4C(u)C(v)}{MN} \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} f(j, k) \cdot \cos \frac{(2j+1)u\pi}{2M} \cos \frac{(2k+1)v\pi}{2N} \quad (2)$$

ただし、 $C(w)$ は $w=0$ のとき $1/\sqrt{2}$ 、 $w \neq 0$ のとき1である。 $f(j, k)$ は画像の各画素の輝度値を表し、DCT成分 $F(u, v)$ は u, v の値が小さいほど低周波成分を表す。DCT成分のうち低周波成分を用いることによって、画面全体の大まかな構造を知ることができる。

ここで、各クラスタが1つのショットに対応するので以下の手順によりクラスタを分割し、境界点を検出する。

- 1枚のフレームのすべての画素の輝度値を求め、その輝度値を用いてDCT成分を求める。ここでは、低い周波数の成分から2つ($F(0, 1)$ と $F(1, 0)$)を用いる。クラスタの正規分布を仮定し、連続する数フレームから平均ベクトル μ と分散共分散行列 Σ を求め、これをクラスタの特徴量とする。
- それ以降の次のフレームに対し、DCT成分におけるクラスタとのマハラノビス距離 $D = \sqrt{(x - \mu)^T \Sigma^{-1} (x - \mu)}$ を求める。
- D が閾値以上であればクラスタから外れるものとし、ショット境界と判定する。そして再び1より処理を繰り返す。
- フレームがクラスタに含まれる場合は、そのフレームを含む5フレームで、クラスタの特徴量を更新し、2より繰り返す。

4. 2段階境界検出

ヒストグラムを用いた手法は色の変化に敏感なため、通常の境界の検出漏れは少ないものの、カメラや画面に写っている物体の動きが激しいシーンなどで誤検出が起こる。また、DCTを用いた境界検出では、動きが激しいシーンでの誤検出は少ないものの、通常の境界の検出漏れが多い。そこで、RGBヒストグラムとDCTを用いた手法を組み合わせることで、効果的なショット分割を行う手法を提案する。RGBヒストグラムを用いた手法では動きが激しいシーンで誤検出が起こりやすい。しかし、この動きが激しいシーンはショット自体は変わっていないため、通常のショット境界ほどはヒストグラム差が大きくないといった特徴がある。そこで、図1で示すように、ヒストグラム差が閾値 A 以上のフレームを検出

し、その中で、周り数フレームに閾値 B 以上のフレームがないフレームをショット境界とする。ヒストグラム差が閾値 B 以上のフレームに対しては、DCT を用いた手法で再検討する。DCT を用いた手法でも境界と検出されていれば、そのフレームをショット境界とする。

RGB Histogram difference

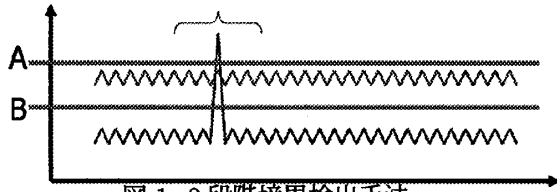


図 1: 2 段階境界検出手法

5. 特殊境界検出

特殊なエフェクトを用いた境界として、主にワイプ、ディゾルブ、CG を用いた境界があげられる。このうち、ワイプを用いた境界と CG を用いた境界の検出法を提案する。

ワイプは左から右に、もしくは右から左にショットが切り替わる特殊エフェクトであり、前ショットと後ショットの間に垂直な境界が現れる。そこでワイプを検出するために、図 2 で示すように中央付近で隣接列の輝度値を比較する。左右端でも同様に調べ、時間的に左 → 中央 → 右、もしくは右 → 中央 → 左の順で閾値以上の輝度値の差が検出されれば、左右端で検出された間をワイプ境界とする。

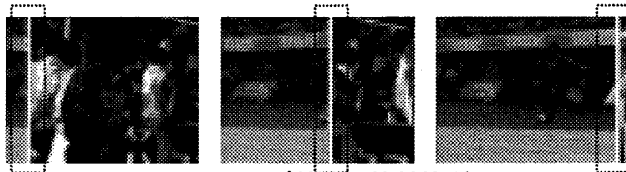


図 2: ワイプ検出の比較箇所

CG 境界は CG の動きが激しく輝度値の変化が大きいものの、一連の動作である。そのため CG 境界の間は、ヒストグラムを用いた手法によって連続的に検出されるといった特徴がある。そこで CG 境界の前後のショットとの境界部分が 2 段階境界検出手法によって検出され、その間の部分がヒストグラムを用いた手法で連続的に検出されれば、CG 境界と判定する。前後の境界のうち片方しか検出されない場合でも、その周りにヒストグラムの手法で連続的に検出されていれば、境界と判定する。

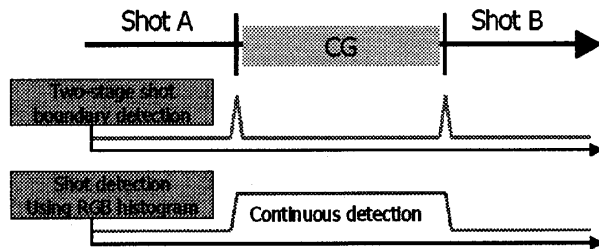


図 3: CG 境界検出

6. 実験と考察

10 分程度のサッカーの映像に対し、ヒストグラムのみを用いた検出法 A. 2 段階境界検出手法を用いた検出法 B. それに特殊境界検出を加えた検出法 C の 3 通りで境

界検出の実験を行なった。用いた閾値は経験的に与えてある。その結果と、同じ映像を用いた従来研究 [1] の結果を表 1 に示す。括弧内の数字は特殊境界の数を表している。検出法 B では DCT を用いることで、検出法 A においてカメラの動きが激しく誤検出された境界をほぼ取り除くことが出来た。検出法 C では全ワイプを検出出来ていて、特殊境界の検出漏れ 3 箇所は全てディゾルブである。

検出法	正検出	誤検出	検出漏れ	適合率	再現率
A	94	20	8	82.4%	83.9%
B	92	3	20	96.8%	82.1%
C	108	3	4	97.2%	96.4%
従来手法	-	-	-	86.7%	93.7%

表 1: サッカー映像に対する実験結果

従来手法は、主に RGB ヒストグラムを用いることでショット境界を検出し、誤検出された境界をショット分類の後、取り除いている。提案手法である検出法 C は 2 段階境界検出手法と特殊境界検出手法を併用することで、適合率 97.2% 再現率 96.4% と、大幅に適合率と再現率を向上することが出来た。

また、サッカー以外の映像に検出法 C を適用した結果を表 2 に示す。全て 10 分程度の映像で、各閾値はサッカーの映像で用いた閾値と同じ値を用いている。アメリカンフットボールで検出漏れとなった特殊境界は CG 境界 2 箇所とディゾルブ 5 箇所、ラグビーではディゾルブ 1 箇所、野球ではディゾルブ 8 箇所であった。

映像	正検出	誤検出	検出漏れ	適合率	再現率
アメフト	69	3	13	95.8%	84.1%
ラグビー	56	0	2	100%	96.5%
野球	44	1	9	97.7%	83.0%

表 2: サッカー以外の映像に対する実験結果

本手法は、ドメイン知識を用いることなく高い適合率と再現率で境界を検出出来ているため、非常に効果的なショット境界検出手法であると言える。

7. むすび

本研究では、RGB ヒストグラムと DCT を用いた 2 段階境界検出手法とワイプと CG 境界に対応する特殊境界検出手法を提案した。そして、本手法が様々な映像に対して、高い適合率と再現率でショット境界検出を行なえることを示した。今後の課題としては、さらに多くの映像に対して実験を行なうことやディゾルブなど他の特殊境界への対応、本研究で得られたデータなどを用いてスポーツの種類自動判別やショット分類を行なうことがあげられる。

参考文献

- [1] 跡部 裕貴, 泉 正夫, 福永 邦雄 “サッカーの放送型映像におけるショット分割および分類,” 電気関係学会関西支部大会, no.G15-1, pp.325 Nov 2005.
- [2] 岩成 英一, 有木 康雄, “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, PRU93-119, pp.23-30, Jan 1994.
- [3] 三浦 宏一, 浜田 玲子, 井出 一郎, 坂井 修一, 田中英彦 “動きに基づく料理映像の自動要約,” 情報処理学会コンピュータビジョンとイメージメディア研究会論文誌, vol.44, no.SIG9, pp.21-29, July 2003.
- [4] N.Uegaki, M.Izumi, K.Fukunaga, “Multimodel Automatic Indexing for Broadcast Soccer Video,” SCIA, pp.808-809, June 2005.
- [5] A.Ekin, M.Tekalp and R.Mehrotra, “Automatic Soccer Video Analysis and Summarization,” IEEE Trans. on Image Processing, VOL.12, NO.7, pp.796-807, July 2003.