

コーナーテロップを用いたテレビ番組の構造化 Video Segmentation for TV Programs using Corner-Subtitles

山本 晃司[†]
Koji Yamamoto

高山 俊輔[†]
Shunsuke Takayama

青木 恒[†]
Hisashi Aoki

1. はじめに

本稿ではテレビ番組で多用されるコーナーテロップを検出・分類することで主にスポーツの個人種目やバラエティ・情報番組の映像を意味的な区間に分割し、見たい場面やシーンだけを頭出しして短時間で視聴することを可能にする構造化手法について述べる。

ハードディスクレコーダ等で録画した大量の番組を短時間で視聴できるように、見たい部分だけを選択して視聴する技術が望まれている。そのような技術の1つとして番組を意味的な区間に分割する技術がある。例えば、ニュース番組において繰り返し登場するアンカーパーソンを検出してトピックごとに分割[1]したり、野球中継番組から投球シーンを検出して要約を作成[2]したりする手法が提案されている。一方で、陸上や格闘技などのスポーツの個人種目やバラエティ・情報番組では必ずしも類似する画面や特徴的な画面が含まれず、前述のような手法では分割が困難であった。

そこで、本稿では近年のテレビ番組で多用されるコーナーテロップに着目し、構造化を行う手法を提案する。本手法の特徴は2つあり、1つ目はコーナーテロップの検出を時空間画像における直線検出に帰着させることで、目立たなく表示されたテロップも安定して検出することを可能としたことである。2つ目は番組の構造をパターン化し、パターンに特徴的なコーナーテロップの分布に基づいて、映像を意味的な区間に分割することである。

2. コーナーテロップと番組構造

コーナーテロップとは番組内の意味的な区間（番組コーナーやショットの集合）に合わせて、画面の隅に長時間、表示されるテロップである。関連するテロップは色や形、表示位置などの特徴が相互に類似し、離れた時刻にある区間でも番組内における位置付けが同じであれば共通した特徴を持つコーナーテロップが用いられる傾向がある。番組の構造に則して表示されるため、コーナーテロップを検出し、特徴に従って分類すれば番組の構造を知ることができる。例えば、ある格闘技番組では試合の単位で変わるテロップとラウンドの単位で変わるテロップが同時に表示される。試合は複数のラウンドで構成されているので、コーナーテロップをもとに番組の階層的な構造を知ることができる。また、あるバラエティ番組では1つの番組コーナーの間、特徴の類似するテロップが繰り返し用いられるため、類似するテロップの分布状況から番組コーナーの構成が分かる。

本手法では次のような特徴を持つテロップをコーナーテロップとして検出する。ただし、試合時間を示す時計

や装飾されたテロップ、ロゴマークなども条件を満たせばコーナーテロップとする。時間的に変化するテロップも条件を満たせば除外しない。

- ・ 1分以上継続して同じ位置に表示される
- ・ 一定以上の輝度があれば、色は限定しない
- ・ 背景が透けた半透明のテロップも含む

一般的にテロップ検出はフレーム単位に画像のエッジ強度を求め、エッジの密度の高い領域を検出することで行われる[3]。しかし、コーナーテロップは長時間表示されても邪魔にならないように、小さく、薄い色で表示されることが多く、縁取りがないものや半透明のものもある。そのため、十分なエッジ強度が得られず、検出漏れが頻繁に発生するという問題がある。本手法ではコーナーテロップが継続して表示される区間を意味的にまとまりのある区間と仮定しているため、途中で検出漏れが起こると不要な分割が生じる原因となる。

また、前述のようにバラエティ番組などでは1つの番組コーナーに何度もコーナーテロップが出現する。検出区間ごとに映像を分割するだけでは過剰な分割が発生するため、関連するコーナーテロップを分類し、番組構造に沿って分割点を決定する必要がある。

3. システム構成

本手法は2つの処理で構成される。初めにコーナーテロップを検出する。次に、検出されたテロップを番組構造のパターンに合わせて分類し、分割点を決定する。

3.1. コーナーテロップの検出

コーナーテロップの検出は画面内の位置の精度よりも時間軸上の検出精度を優先するため、時空間画像における直線として検出する。時空間画像はフレーム単位の画像を時間軸方向に順次並べたもので、時間軸方向の断面はカメラや物体の動きによって特徴的なパターンとなる[4]。本手法では元の映像を2fps、88×60画素にサンプリングしたサムネール画像を用いて時空間画像を構築する。このとき、コーナーテロップは時空間画像において、図1のような直線として表れる。コーナーテロップ以外の部分はカメラや被写体の動き、ショットの切り替えにより連続的な直線とはならない。

直線として検出する画素は初めに輝度が一定以上のものを選択した後、1) 時間軸方向の画素値の類似度、2) x 軸

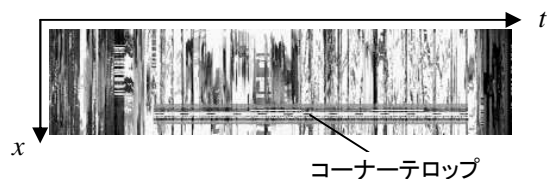


図1 $y = s_i$ における時空間画像の断面

[†] (株) 東芝 研究開発センター
Corporate R&D Center, Toshiba Corporation

方向の隣接画素との差分, 3) 前記 2 の差分の時間軸方向の類似度の 3 指標によりエッジ強度を算出し, 2 値化して抽出する. 第 3 の指標は半透明のテロップに対応するためのもので, 半透明のテロップでは背景の変化に合わせてテロップ部分も変化するが両者の差分は大きく変化しないという特性に因る.

最終的なコーナーテロップの表示区間は y を変化させて複数箇所で作成した断面 ($y = s_i, i = 1 \dots N$) から抽出した直線を統合・整形して決定する. このように複雑な形状を持つテロップの検出を直線の検出に帰着させることで, 時間軸方向に安定した検出が可能となる.

3.2. 番組の分割

スポーツに対しては検出区間をそのまま意味的な区間として先頭を分割点とする. ただし, 2 つ以上の分割点が近くにある場合は統合する. これはスポーツでは選手単位や試合・ラウンド単位でコーナーテロップが切り替わることが多く, バラエティ番組のように頻繁に切り替わることが少ないためである.

一方で, バラエティ・情報番組は複数のコーナーテロップの検出区間をまとめてから分割点を決定する. 初めに番組構造のパターン別に特徴が類似するコーナーテロップを関連テロップとして分類し, 次に関連テロップの分布に従って分割点を決定するという流れで行う. 分類に使う特徴量はテロップ領域の色・画面内の位置・表示区間である. このとき, 色の特徴量は 3.1 節で検出された直線を統合して得られた時空間領域より時間軸方向の変動が小さい成分を選択して算出する. 番組構造のパターンは実際の番組を参考に 3 種類に分けた. 各パターンの特徴を図 2 に示す. 図中のグループとは色・位置が類似するものを示す.

[パターン A] 番組のコーナー分けが明確で, 同一のコーナーでは色や位置が類似するテロップが繰り返される. 色・位置の特徴量によりクラスタリングし, クラスタの重複が小さくなる区間で分割する.

[パターン B] 明確なコーナー分けがないので, 場面転換 (場所の移動など) をコーナーとする. また, グループ間の関連性が少ないので, 色・位置は用いず, 表示区間が重複するテロップをまとめて, 先頭で分割する.

[パターン C] 登場人物や話題がコーナーごとに順次変わりながら, 番組が進行する. スポーツにおける選手や試合の切り替わりと類似しており, スポーツ同様, 意味的な区切りでテロップが切り替わる. 色・位置は用いずに区間の重複するテロップをまとめ, 各テロップの表示開始時刻で映像を分割する.

尚, 番組を通して表示される現在時刻を示す時計などは番組構造に関係がないため, 映像の全体長に対する割合を閾値として除外しておく.

4. 実験

1 時間から 2 時間程度のスポーツ及びバラエティ・情報番組を用いて分割点の有効性を評価した. 後者の番組構造パターンは手動で設定し, 検出された分割点の前後 10 秒以内に番組の構成上, 意味のある区切りが存在する場

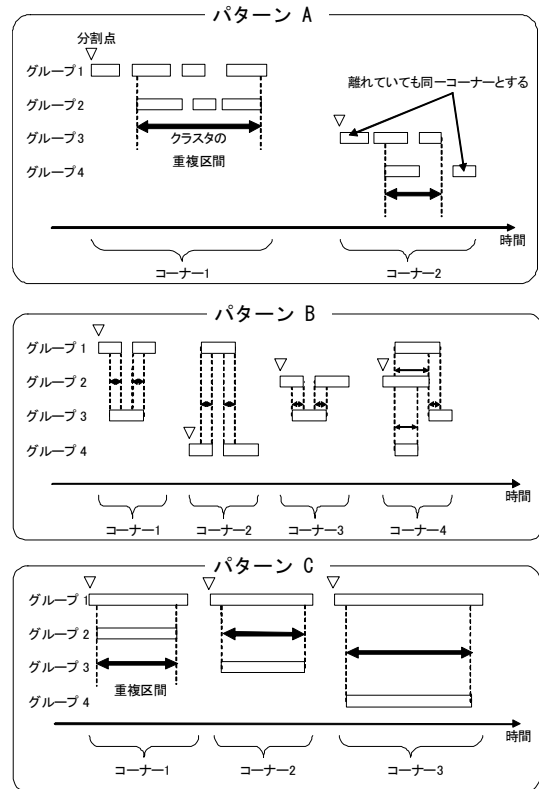


図 2 番組構造のパターン

合に有効 (適合) であると判断した. スポーツ (格闘技中継) は 4 番組の平均で適合率 90% のとき, 試合・ラウンドの開始を 95% 検出した. バラエティ・情報番組ではパターン A において 3 番組の平均で適合率 85% のとき, 全ての番組コーナーを検出した. パターン B および C は同じ処理であるため, まとめて評価を行い, 11 番組の平均で適合率 70% のとき, 番組コーナーの 90% を検出した.

5. まとめ

番組中のコーナーテロップを検出・分類することで, 番組を意味的なシーンの単位に分割する手法について述べた. 本手法を用いることで, 従来の画像や音声の特徴量だけでは効果的なインデクシングが困難であったスポーツの個人種目やバラエティ・情報番組の分割において有効性を確認した. 今後の課題としては現在, 手動で与えている番組パターンの自動推定や他の構造化手法との組み合わせによる分割点の高精度化が挙げられる.

参考文献

- [1] 青木: “映像対話検出によるテレビ番組コーナ構成高速解析システム”, 信学論, Vol.J88-D-II, No.1, pp.17-27 (2005)
- [2] 谷治, 他: “投球シーンの解析とテロップ情報を用いた野球中継ダイジェスト版の自動生成法スポーツのインデクシング”, 信学技報, IE2004-91, Vol.104, No.465, pp. 27-30 (2004)
- [3] 堀: “テロップ認識のための映像からの文字部抽出法”, 情処学研報, CVIM-114-17, Vol.99, No.3, pp.129-136 (1999)
- [4] Chong-Wah Ngo, et al.: “On Clustering and Retrieval of Video Shots Through Temporal Slices Analysis”, IEEE Trans. on Multimedia, Vol.4, No.4, pp.446-458 (2002)