

競技タイムテロップを用いた スポーツ競技番組自動チャプタリングシステム

高山 俊輔[†] 山本 晃司[†]

株式会社 東芝 研究開発センター マルチメディアラボラトリー[†]

1. はじめに

本稿では、テレビのスポーツ番組から競技が行われているシーンを検出することで番組を効率視聴できるようにするシステムを提案する。

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。そのため、映像を解析して自動的にインデックスを付与する技術が注目されている。スポーツ番組は放送時間が長いいため特に有効であり、多くの手法が研究されている。これまでに、音響的な盛り上がりを検出して重要シーンを検出する手法[1]や、番組中で長時間表示されているテロップを用いて番組を試合やラウンドといった区切りで分割する手法[2]などが提案されている。しかしこれらの技術は、検出結果が何のシーンかという意味付けが曖昧なため、ユーザがインデックス情報を用いて見たいシーンを選択するといったインタラクティブな操作には適していなかった。

そこで本稿では、陸上や格闘技などのスポーツ番組で競技中に出現する競技タイムテロップを利用して競技シーンを検出する。これにより、競技の開始・終了時刻を高精度に検出でき、競技の先頭シーンを頭出しすることで興味のない競技をスキップしたり、ゴールシーンだけを視聴することが可能になった。

2. 競技タイムテロップとチャプタリング

競技タイムテロップとは、スポーツの陸上や競泳、格闘技番組などで競技中に表示される経過時間や残り時間を示すテロップである。本手法はこの競技タイムテロップが出現する競技番組や格闘技番組に特化した技術である。これらの番組では、スタジオトークやウォーミングアップの時間が長く、実際に競技が行われているシーンは非常に少ないという特徴がある。そこで、競技タイムテロップの表示区間を競技シーンとして検出して利用できる。競技タイムテロップの出現時刻を競技開始として頭出ししたり、消失直前の時刻をゴールシーンとしてダイジェスト映像の作成などに利用できる。

3. 競技タイムテロップの検出

競技タイムテロップの検出アルゴリズムを示す。検出は、1) テロップ領域識別、2) 競技タイムテロップ領域判定、3) 競技タイムテロップモデル構築、

Video Segmentation for TV Sports Programs using video clock
[†]Shunsuke Takayama, Koji Yamamoto
[†]Multimedia Laboratory, R&D Center, TOSHIBA Corporation

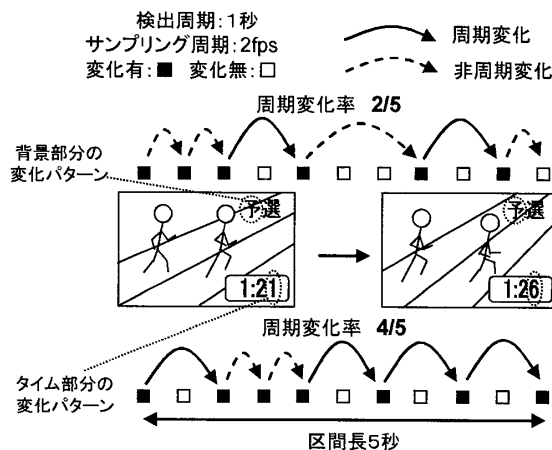


図1 周期変化の計算方法

4) 競技タイムテロップ表示区間検出の4つの部分により構成される。検出は2パス処理で、1) 2) が1パス目、3) 4) が2パス目となっている。1パス目で粗く検出した後、2パス目では検出した競技タイムテロップのモデルを構築し、詳細な検出を行う。以下では順を追って各処理の詳細を説明する。

3.1. テロップ領域識別

始めに、番組内で表示される各テロップを検出し、その出現時刻や画面内での表示領域、色を記憶する。これには[2]のコーナーテロップ検出で提案されている手法を修正して用いる。次に、検出したテロップをその表示領域や色に基づいてグルーピングする。2) ではこのテロップグループを単位として処理する。

3.2. 競技タイムテロップ領域判定

次に、前段で検出した各テロップグループが競技タイムテロップかどうか判定する。競技タイムテロップの判定には、タイムの秒の部分に1秒毎に変化するという特徴を利用する。まず各フレームの画像を輝度値により2値化し、各隣接フレーム間で、同一位置の画素値の変化の有無を検出する。1秒周期で変化する画素の判定は、図1に示す周期変化率を画素毎に求めることで行う。周期変化率は、各テロップグループの表示区間内で1秒周期での変化がどの程度の割合であったかを示すもので、[0, 1]の値をとる。図1は2fpsで比較した場合で、隣接フレーム間で画素値が変化したときが■、しなかったときが□となっている。前回変化があったフレームから2フレーム後に次の変化が観測された場合に周期変化したとみなし、周期変化の回

数をカウントする。図1ではタイム部分は4回の周期変化が観測され、テロップの区間長は5秒なので、周期変化率は4/5となる。一方で背景部分の周期変化率は2/5なので、これを閾値処理することで1秒周期で変化する画素を判定し、そのような画素を含むテロップグループを競技タイムテロップとする。

サンプリングレートが低いとちょうど1秒での変化を観測する際の時間の誤差が大きいため(2fpsだと1秒±0.5)、背景のような頻繁に変化する領域の周期変化率が高くなりやすい。そこで実験では、2fpsと5fpsの2通りで比較を行い、タイム部分の変化と背景変化の識別精度がどの程度変わるか評価した。

3.3. 競技タイムテロップモデル構築

抽出した競技タイムテロップのエッジ形状と色ヒストグラムを抽出してモデルを構築し、検出に用いる。一つの番組中で競技タイムテロップの配置や形状が変わることはほとんどないので、構築したモデルを用いて1パス目で見逃していた競技タイムテロップを検出できる。2)で抽出した競技タイムテロップの表示区間内のフレーム毎にエッジ抽出を行い、区間内でエッジの検出率が高い画素をテロップのエッジ形状として抽出する。また、色ヒストグラムはテロップの表示領域・区間内での平均のものを用いる。

3.4. 競技タイムテロップ表示区間検出

3)で構築したモデルを用いて競技タイムテロップが表示されている区間を検出する。映像中の各フレームについて、競技タイムテロップの表示領域に該当する部分からエッジ形状及び色ヒストグラムを抽出する。これをモデルのエッジ形状、色ヒストグラムとそれぞれ比較し、両方の類似度が高いフレームを、競技タイムテロップの表示区間として出力する。

4. 評価実験

実際のコンテンツを用いて検出精度の評価を行った。また、検出結果を利用して競技シーンを視聴するビューワを作成した。

4.1. 検出精度の評価

陸上、競泳など複数のジャンルから成る22コンテンツのスポーツ番組を用いて競技タイムテロップの検出精度を評価した。評価はテロップ出現時刻の前後5秒以内を正解とし、再現率と適合率により評価した。また、競技タイムテロップ抽出時のサンプリングレートは2fpsと5fpsの2通りで比較した。実験結果を表1に示す。5fpsの平均検出率は再現率93%、適合率86%だった。2fpsと5fpsの比較では、再現率はほぼ同じだが、適合率で5fpsの方が10ポイント程度高い結果となった。これは、フレームレートを増やすことで背景変化とタイム部分の変化の識別性が向上し、過剰検出が抑制されたためである。

4.2. ビューワの作成と考察

図2に検出結果を利用して番組を視聴するための競

表1 競技先頭検出の再現率及び適合率

サンプリングレート	再現率[%]	適合率[%]
2fps	91.6	75.8
5fps	93.2	86.3



図2 競技シーンビューワの外観

技シーンビューワの外観を示す。映像中の右下にある白いテロップが競技タイムテロップである。画面下部のタイムバー中で緑の縞の部分に競技シーンがあり、図の例では4つの競技シーンがある。左右のボタンを押すと、次の競技または前の競技の先頭にジャンプする。ゴールボタンを押すとその競技の終了直前から再生し、ゴールシーンを視聴することができる。

このビューワを用いることで、ユーザは自分の見たい競技シーンを簡単に検索して視聴することができる。例えば、競技開始から数秒だけを視聴して日本人選手が出場していなければスキップする、興味が低い競技はゴールシーンだけ見るといったように、各ユーザの好みに応じた視聴方法でスポーツ番組を効率的に視聴できる。また、タイムバー上の各競技区間の長さにより、100mか400mかといった競技の種類が予測できるという効果もあった。同様に格闘技番組でも試合前のVTRをスキップしたり、競技終了直前のシーンとして各ラウンドの終了やKOシーンを視聴できた。

5. まとめ

スポーツの競技番組や格闘技番組で出現する競技タイムテロップを用いて競技シーンを検出する手法について述べた。本手法を用いることで、高精度に競技タイムテロップを検出でき、競技シーンだけを効率的に視聴することが可能となった。今後の課題としては他のインデクシング技術と組み合わせることで幅広いジャンルのスポーツに対応することが挙げられる。

参考文献

- [1]佐野, 他: “サッカーゲームにおける会場音からの重要場面抽出の検討,” 信学技報, PRMU2001-166, 2001.
- [2]山本, 他: “コーナーテロップを用いたテレビ番組の構造化,” FIT2008 講演論文集, pp287-288