

## 動きベクトルによるスポーツ映像の競技シーン検出手法の提案

奥川 雄紀<sup>†</sup> 後藤 敏<sup>†</sup> 池永 剛<sup>†</sup>早稲田大学大学院情報生産システム研究科<sup>†</sup>

## 1. まえがき

現在、我々は蓄積技術の向上やネットワークのブロードバンド化の進展により、膨大な映像コンテンツを享受できるが、全てを視聴することは困難であり、見たいシーンのみを時短視聴することのできる環境が求められている。そこで、本研究では、短時間で内容把握が容易なスポーツ映像をターゲットとした早送りアルゴリズムの提案を行う。本手法は、カメラが対象を追うようなスポーツ映像に焦点を当て、カメラワークによって発生する動きベクトルの大きさと、その連続性を用いることで、競技シーンのみを検出することを可能としている。競馬、陸上、水泳映像に対して評価を行い、本研究の有効性を示す。

## 2. スポーツ映像の特徴と研究の目的

本論文では、膨大な映像コンテンツの中で、スポーツ映像に注目する。スポーツ映像は、ドラマ映像やニュース映像などとは異なり、重要シーンとそうでないシーンの区別が明確であり、重要シーンのみを取り出し再生させることで、意味を損なうことなく時短視聴できる効果が大きいと考えられる。しかし、一言でスポーツ映像といっても、その特徴は千差万別であり、全てのスポーツ映像に適用可能な手法の実現は困難である。よって従来より、特定のスポーツに特化した研究が多くなされてきている。なかでも我々に身近なスポーツである野球[1]、サッカー[2]、アメフト[3]などの研究は多数行われており、どの研究もそのスポーツにおいて大きな成果が得られている。これらの研究はカメラワーク、インデクシング、カット点、音声、テロップ、キーワード、クローズドキャプションなどに着目し、長時間連続する競技シーンの中からイベントシーンのみを検出するものである。

一方、競馬、陸上、水泳、相撲といったスポーツは、上記のスポーツカテゴリーとは異なり、実際の競技時間は比較的短時間であり、競技間に多くの時間が費やされるという特徴を有している。このようなスポーツジャンルの映像に対しては、競技中の盛り上がった特定シーン・イベントだけでなく、競技全体を重要シーンとして捉え、再生させることにより、高い時短視聴効果が得られると考えられる。そこで本論文では、競馬等のスポーツジャンルに焦点を当て、その競技シーンを検出する手法を提案する。従来手法で用いられている音声やカット点などでは、競技シーン全体の検出は困難であるため、カメラワークによって発生する映像の特徴の違いに注目し、競技シーンとそれ以外のシーンの切り分けを行う。

## 3. 提案手法

本手法は、スポーツ映像のカメラワークに着目し、カメラワークによって発生する動きベクトルの大きさとその連続性を用いて競技シーンを検出する。スポーツ映像のカメラワークは、スポーツのジャンルによって異なる特徴があることを実際の映像から確認できる。たとえば相撲、野球などは、競技中に比較的にカメラが固定されていることが確認されている[4]。この手法をMPEG形式の動画像に適用することを考える。MPEGは3つのピクチャタイプから構成されI、P、Bピクチャとある。

Iピクチャは、その情報だけから符号化された画像で、フレーム間予測を使わずに生成される。Pピクチャは過去のフレームから一方方向のフレーム間予測を行い、差分を符号化したものである。Bピクチャは、過去と未来の2つのフレームから二方向のフレーム間予測を行い、差分を符号化したものである。本手法では、映像のカメラワークを知るために順方向予測であるPピクチャのみを判定に用いている。Pピクチャの動きベクトルを抽出し、パラメータの調整を行った。パラメータは以下のものを使用する。

- 1フレームごとの動きベクトルの平均の大きさ
- 動きベクトルの連続性
- 種目別の時間的特徴

1フレームごとの動きベクトルの平均の大きさを使用する理由は、競技中のカメラワークの検出が容易であるからである。

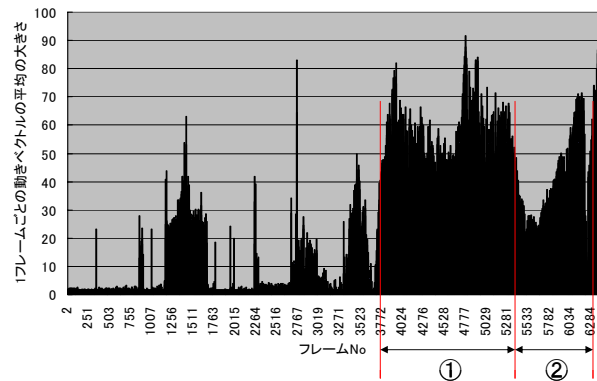


図1：競馬映像の動きベクトルの平均の大きさ

図1の①、②を合わせて1つのレースシーンである。図1から見られるように、レース中には大きな動きが確認できる。①、②の期間は1フレームの動きベクトルの平均の大きさが50以上のものがほとんどである。それと比較して、レース以外の映像であるスタジオ映像、パドック映像などは20～30くらいの大きさとなる。このことから、競技シーンとそれ以外のシーンを1フレームごとの動きベクトルの大きさの平均で切り分けることができる。次に、図1の②について、1フレームごとの動きベクトルの平均の大きさでは検出できないシーンがある。競馬映像の場合では最終コーナーの際、カメラが一時的に固定され、カメラワークが変化する。また、水

The proposal of detection method based on motion vector for the contest scene of sports video  
<sup>†</sup>Yuki Okukawa, <sup>†</sup>Satoshi Goto, <sup>†</sup>Takeshi Ikenaga  
<sup>†</sup>Graduate School of Information, Production and Systems, Waseda University

泳映像に関してはターンシーン、ゴールシーンにおいてカメラが固定される特徴がある。陸上映像もトラック競技では、競馬映像と同様のことがいえる。このようなシーンはレース中の動きベクトルの連続性を用いることにより改善した。図2、3から競技中には大きな動きが連続していることがわかる。このことを利用し、連続して表示条件を満たした場合は、その種目に応じて時間的に追加して表示する改善を加えた。このように、種目によって異なる時間的特徴を利用することで、検出することができなかったシーンを検出することができる。競馬を例にとると、最終コーナーから約30秒でゴールすることは、ほぼ全てのレースにいえる時間的特徴である。このように各々の種目の特徴を利用し、追加して表示することが有効であるといえる。

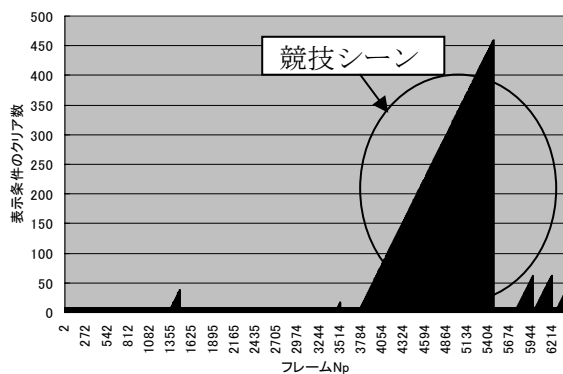


図2：競馬レースに見られる連続性

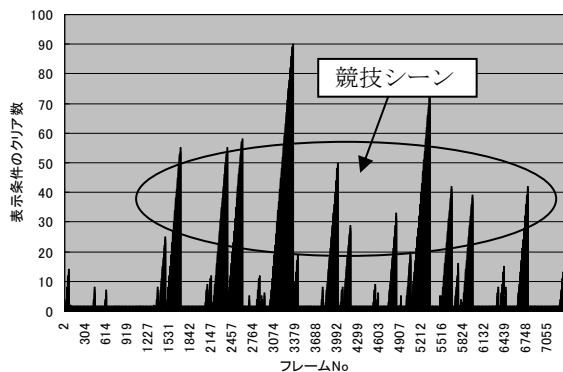


図3：水泳競技に見られる連続性

## 5. 実験結果・考察

表1、2、3に競馬、水泳、陸上の詳細結果を示す。

表1. 競馬映像結果 (3分9秒2.4秒)

	レース時間	レース検出時間	検出率	ミス率
レース1	1分13秒	1分10秒	96%	8%
レース2	1分28秒	1分25秒	97%	2%
レース3	2分15秒	2分05秒	92%	3%
レース4	2分03秒	1分53秒	92%	6%
全レース合計	6分59秒	6分33秒	94%	5%

表2. 水泳映像結果 (1分6分4.1秒)

	レース時間	レース検出時間	検出率	ミス率
レース1	4分10秒	3分54秒	93%	10%
レース2	3分52秒	3分17秒	85%	15%
全レース合計	8分2秒	7分11秒	89%	13%

表3. 陸上映像三段跳び結果 (3分15秒)

	レース時間	レース検出時間	検出率	ミス率
1本目	15秒	13秒	86%	5%
2本目	15秒	12秒	80%	1%
3本目	11秒	10秒	90%	1%

陸上映像トラック競技結果

	レース①	レース検出時間	検出率	ミス率
レース①	2分6秒	1分21秒	64%	19%
レース②	2分0秒	1分45秒	87%	1%
全レース合計	4分47秒	3分41秒	77%	9%

※検出率＝競技検出時間/競技時間

ミス率＝競技外検出時間/(映像時間-競技時間)

表1の競馬映像では、全てのレースにおいて90%以上の高い検出率を確認した。100%の検出ができていない理由は、スタートシーンのカメラ位置によるものである。真横からの撮影の場合100%になり、正面、斜めなど位置によっては数秒検出することができないという結果が得られた。また競馬映像の連続性を利用し、ミス率を大きく削減した。

表2、3の水泳、陸上映像においても約80%の高い検出率を得ることができた。水泳、陸上映像に関しても競馬映像と同様に、スタートシーンの際、カメラが固定されるので、数秒検出が遅れる。また、表2に見られるように水泳映像に関しては、最初の50mまでの間に連続性が見られれば、高い検出率を得ることを確認した。しかし、一部の映像では連続性の判断ができず、前半部分が正しく検出できないシーンが見られる。この原因はカメラワークが上、横、下と切り替わることで、映像によっては連続性を検出できないためであると考えられる。

陸上映像の結果を示した表3では、トラック競技に関して検出率に多少の差が見られた。これは、競馬映像とは異なり、人間の表情に着目することが主な原因である。次に競馬映像とはスピード、カメラワークに若干の違いが生じるためである。表3のミス率が高い理由であるが、前半に行われた三段跳びのハイライトシーンが大きな割合を占めたため、ミス率が多く表れている。

## 6. まとめ

競馬、水泳、陸上映像に適用した結果、比較的単調な競馬映像において90%以上、水泳、陸上映像においても80%以上の検出率を達成し、提案手法の有効性を確認した。ミス率に関しては競馬映像で5%、陸上映像で9%、水泳映像で13%と比較的少ないミス検出で競技シーンの検出を行うことができた。

## 7. 参考文献

- [1]山本拓、佐藤宏介、千原國宏、“野球中継映像における各種プレイシーンの自動検索/編集システム”2000信学総大、情報・システム2、D12-77、P247、2000
- [2]益満健、越後富夫、“映像重要度を用いたパーソナライズ要約映像作成手法”電子情報通信学会論文誌2001/8 Vol. J84-D-II No. 8
- [3]新田直子、馬場口登“放送型スポーツ映像の意味内容獲得のためのストーリー分割法”電子情報通信学会論文誌2003/8 Vol. J84-D-II No. 8
- [4]服部しのぶ、高木真一、小館亮之、富永英義“映像特徴に基づく自動映像分類システムの提案”研究報告オーディオビジュアル複合情報処理Vol. 2002 No. 025