

投球の次ショットに重きを置いた シーンのパターン化による野球映像イベント識別

Baseball Video Event Classification based on Patternization of Scenes Focusing on Next Shot of Pitch

望月 貴裕† 藤井 真人† 八木 伸行†
Takahiro MOCHIZUKI Mahito FUJII Nobuyuki YAGI

1. まえがき

近年、スポーツ映像を処理対象とし、映像からハイライトを抽出する技術の研究が行われている。それらの研究においては、イベント内容を自動的に識別するための学習および認識モデルとして、隠れマルコフモデル(HMM)がよく用いられる。

特に、野球映像に関して多くの研究が行われている [1-3]。野球映像は、本塁打、二塁打、四球などのイベントによってカメラワークやスイッチングによる画の移り変わりが比較的異パターン化しており、区別し易いと考えられるためである。

これらの手法は、それぞれ7種から8種程度のイベント識別を試みているが、再現率約70~80%の精度にとどまっている。しかし、実用化を想定して使用者の観点で見ると、イベントの種類は少なくとも、長打や安打などの重要シーンを漏れなく検出することができる手法が実用的である。

そこで本稿では、打球が「遠く、あるいは高く」飛ぶほど重要なシーンであると考え、打球の内容によって分類した3種のイベントについて高精度の識別を行う手法を提案する。

2. 本手法の概要

本手法では、「イベントシーン」を識別の単位とする。イベントシーンとは、イベントの開始から終了までの、複数のショットで構成される区間映像であり、そのイベントが起こる投球ショットの開始点をイベント開始点、以下の定義例で定めた最終ショットの終了点をイベント終了点とする。これらのイベントシーンの映像中からの切り出しは手動で行うこととする。

イベントシーン最終ショットの定義例

- ・ リプレイの直前
- ・ 3アウトチェンジの際にスコアがスーパーされる
- ・ 次投球ショットの直前

また、[3]では7種類のイベントシーンを処理対象としているが、本手法では、これをさらに、打球の内容によるシーンの重要度に基づいて以下の3つのイベントに統合し、識別を行う。

重要度高: 打球が「飛球あるいは遠くへの強いゴロ」

- 本塁打、二塁打、安打、フライアウト

重要度中: 打球が「遠くへ飛ばないゴロ」

- 内野ゴロ

重要度低: 打球が「飛ばない」



Fig.1 パターンデータ例

三振、四球

さらに本手法では推移の単位をショットあるいはそれに準じた区間映像とする。これはフレームを単位とした場合に比べ、投球の遅れなどの時間軸方向ノイズに強いためである。また、各区間映像に対し、[3]で提案されているパターン化処理を施す。パターン化によりデータのバリエーションを抑制することができ、少ない学習データによるシーン識別が可能となる。

本稿では、投球ショットおよびそれに続く第2ショットのみで前述のイベント3種の識別が可能であると考え、イベントシーンの第2ショットまでのパターン化、数値列化に基づいて学習および識別を行う。処理の詳細は以下の章で述べる。

3. 学習処理

3.1. イベントシーンのパターン列化

まず、学習用イベントシーンのショットチェンジ点を検出する。本手法では[4]の手法を採用する。[4]は、観客席等の複雑なテクスチャが激しい動きをすることで誤検出が生じやすいスポーツ映像に対して良好な検出能力を示す

本稿でのイベント識別では、第2ショットが非常に重要な役割を担うと仮定する。打球の方向、野手のキャッチおよび送球という識別のための重要な情報を含むからである。そこで、第2ショットに関する学習データに重きを置くために、第2ショットを T_2 フレーム間隔で分割し、部分第2ショット系列を生成する。

そして、全学習用イベントシーンを構成する全ての投球ショットおよび部分第2ショットを[3]の手法に基づいてパターン化する (Fig2, Step1)。パターン化とは、フレーム画像上に配置された点群の追跡、各点近傍の画像特徴計算、画像特徴に基づく点群のグルーピングなどの処理により、ショットなどの区間映像を複数の矩形データからなるパターンデータ (以下 PD) で表現する手法である。PDを構成する各矩形データは以下の情報を保持する。

頂点位置	矩形の四隅の頂点座標
画像特徴ベクトル	色特徴およびテクスチャ特徴
動きベクトル	矩形中心の大きな動き

†NHK 放送技術研究所

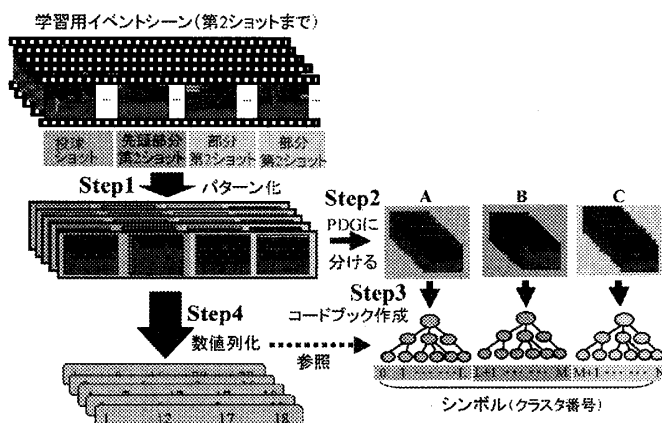


Fig.2 学習データ構築の流れ

PD を視覚的に表現した例を Fig.1 に示す。各矩形を埋める模様は、画像特徴ベクトルに対応したテクスチャであり、矢印は動きベクトルを表す。

3.2. 学習用イベントシーンの数値列化

本処理は、3.1 節で述べたパターン化の結果に基づいて行われる。まず、学習用の全ての PD を以下の 3 つのグループ (以下 PDG) に分ける (Fig2, Step2)。

- A) 投球ショット
- B) 「先頭」部分第 2 ショット (投球ショット直後)
- C) それ以降の部分第 2 ショット

次に、各 PDG に属する PD を以下のように階層的にクラスタリングし、PDG 毎のコードブック (以下 CB) を作成する (Fig2, Step3)。

1. パターン化する前のショットあるいは部分ショットの先頭フレーム画像を切り出し、その画像の類似度に基づいてクラスタリングする。
2. 1 で作成した各クラスターを、PD 中の最大面積の矩形データ (多くは背景に該当) の類似度に基づいてさらに分割する。矩形データの類似度は、位置情報、画像特徴、動きベクトルの類似度に重みを乗じて足し合わせたものである。
3. 2 で作成した各クラスターを、PD 中の中央付近にある矩形データ (多くは人物領域に該当) の類似度 (2 と同様) に基づいてさらに分割する。
4. 3 で作成された各クラスターにシンボル (番号) をランダムに割り当てる。このクラスターと番号との対応関係が CB である。

そして、各ショットおよび部分ショットの PD が、対応する PDG の CB に基づいて数値に変換され、学習用イベントシーンが全て数値列で表現される (Fig2, Step4)。

3.3. HMM のセット

本手法では、一つのイベント種に一つの HMM を割り当てるのではなく、数値列の長さ毎に HMM をセットする。野球映像においてはイベントシーンを構成する区間映像数は、イベント種を識別するための重要な情報の一つであると考えられるからである。

まず、各イベント種において、数値列をその長さで振り分ける。そして長さで振り分けたそれぞれの数値列集

Table.1 実験結果

イベント種	シーン数	再現率	適合率	F 値
重要度高	85	77/85	77/84	0.911
重要度中	50	45/50	45/57	0.841
重要度低	38	31/38	31/32	0.886
total	424	153/173		0.884

重要度高: 打球が飛球、または外野へのゴロ

重要度中: 打球が内野へのゴロ

重要度低: 打球が飛ばない

合に HMM を割り当て、数値列の学習によりそのパラメータを設定する。

4. 識別処理

まず、学習処理と同様に、未知イベントシーンのショットチェンジ点を検出して第 2 ショットを T_2 フレーム間隔で分割し、投球ショットおよび各部分第 2 ショットをパターン化する。次に、学習時に作成した CB に基づき、PD 列を数値列に変換する。さらに、この数値列との長さに対応した学習済 HMM を全てのイベント種から選び、それらの HMM を用いて、この数値列に関する尤度を計算する。そして尤度が最も高い HMM のイベント種を、このシーンのイベント種として決定する。

5. 実験

MLB の放送映像を用いて、2 章で列挙した 3 イベント種の識別実験を行った。実験条件を以下に示す。

映像のサイズ 240×180 (画素)

学習シーン数 約 22 試合分

(重要度高: 500, 重要度中: 403, 重要度低: 673)

第 2 ショットの分割 $T_2=45$ (フレーム)

HMM 離散 HMM, left-to-right モデル, 状態数 4

CB シンボル数 PDG-A: 5, PDG-B: 16, PDG-C: 48

イベント種未知の 173 シーン (約 3 試合分) に対する識別実験結果を Table.1 に示す。全体の F 値 (再現率と適合率の幾何平均) が約 90%, 各イベント種においても F 値が 84%~91% であり、非常に精度の高い識別結果が得られた。

6. あとがき

本稿では、第 2 ショットに重きを置いたシーンのパターン列表現と HMM を用いて、野球映像における 3 種のイベント識別を高精度で行う手法を提案した。

今後の課題として、常に安定した識別結果を得るために学習データを増やす必要がある。また学習データの質を上げるために、パターン化処理および CB の精度を改良していく必要がある。そして本手法を、さらに多種のイベント識別の高精度化へ繋げていきたい。

参考文献

- [1] P.Chang et.al "Extract highlights from baseball game video with hidden markov models", ICIP 2002, Vol.1. pp.609-612
- [2] H. B. Nguyen et.al, "Robust highlight extraction using multi-stream Hidden Markov Models for baseball video", ICIP2005, pp.III-173-176
- [3] T.Mochizuki et.al, "Baseball Video Indexing Using Patternization of Scenes and Hidden Markov Model", ICIP 2005, pp.III-1212-1215
- [4] 望月 他, "フラクタル特徴の変化に基づくカット点検出", 電子情報通信学会総合大会, D-11-134, 2005