

H-047

スポーツニュース映像をクエリとした野球放送映像からの同一シーン検索 Identical Scene Retrieval from Baseball Broadcast Video using Sports News Video as Query

望月 貴裕† 藤井 真人† 酒井 善則‡
Takahiro MOCHIZUKI Mahito FUJII Yoshinori SAKAI

1. まえがき

春に行われた WBC の大変な盛り上がりが見えるように、様々なスポーツの中で、野球は恒久的に人気の高いコンテンツのひとつである。本来ならば、臨場感を感じながら野球中継をリアルタイムで視聴するのが理想であるが、ライブスタイルの多様化により、その時間帯にテレビを見られる状況にない視聴者も増えている。そのような人々にとって、主に深夜に放送されるスポーツニュースは非常に有用な情報となる。

しかし、スポーツニュースにおける大リーグのダイジェスト映像は、日本人選手の成績を伝える点に強く重きが置かれて非常に短く編集されているため、選手の出塁や投手の奪三振後の試合展開など、個人の成績だけでなく試合状況に関する情報を求める視聴者が、情報不足を感じる場合も少なくない。

そこで我々は、スポーツニュースにおける大リーグ野球映像の各シーン（打席単位）から、「試合放送映像における対応シーンを見つけ出しその時点から再生する」といったサービスを前提とした検索技術を提案する。この技術は、上記のような「野球が好きでじっくり見たいが時間が無い、しかしニュースの内容だけでは物足りない」という視聴者にとって有用なものになると期待できる。

2. スポーツニュース映像の編集傾向

ニュースにおける野球試合結果のダイジェスト映像は、基本的に、A：試合放送映像、B：放送以外の映像（取材クルーが撮影）の2つの映像を素材として編集されている。

大リーグについては、スタジアムで撮影できるテレビ局独自の取材クルーの人数に制限があるため、入手できる B の素材量は少ない。つまり、大リーグのニュース映像は、A を中心に編集されているケースが多い。その結果、大リーグについては、ニュース映像と試合映像において、全く同じ内容の映像区間が高い確率で現れる。

さらに、ニュース映像の構成を詳しく検証してみると、図 1 に示すように、ニュース映像の打席単位のワンシーンにおけるショット（映像の切り替わりで区切られる単位）系列の構成を見ると、試合映像の「投球ショット」（投球をセンターカメラから撮った映像）と「打球ショット」（打球をカメラが追跡）が非常に高い確率で使用される傾向がある。

3. 提案手法の詳細

前節で述べた傾向を利用し、本稿では、投球ショットと打球ショットから抽出した画像特徴の類似性に基づき、大リーグのスポーツニュースの投球ショットから始まる各シーンに対応した放送映像中の同一シーンの検索手法を提案する。

3.1. 画像特徴ベクトル

本提案では、「ニュース映像のシーンの投球および打球

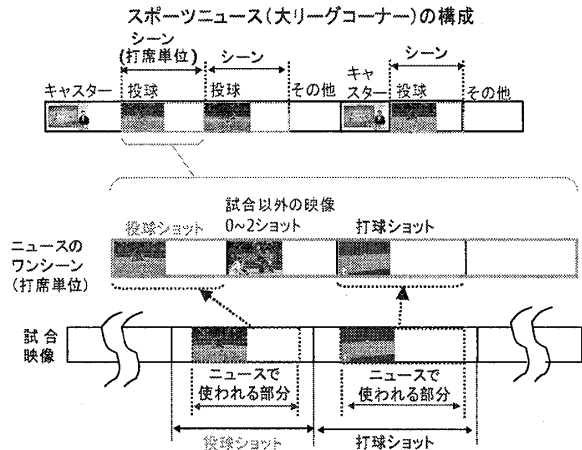


図1 スポーツニュースの構成

ショットと同内容のものが試合映像の中に存在する」という前提があるため、ニュース映像のあるシーンに対応した試合映像を探すには、そのシーン中の1枚のフレームと同じ画像を試合映像の各フレームから探すのみで十分と思われる。しかし、ニュース映像と試合映像は異なる機器およびプロセスで放送映像が作られており、解像度、明るさおよび色の度合いなどに差が生じ、理想的なフレーム画像同士の類似度が得られない。また、投球ショットのように著しく典型的な構図を持つ場合、多くの類似したフレーム画像が存在し、正確なシーン特定が難しくなる。

これらの課題に対処するために、本稿では、複数フレームを用いることとし、「映像中のフレーム番号 t における特徴ベクトル $V(t)$ 」は、 $t \sim t+T$ のフレーム画像から抽出した画像特徴を用いて生成する。まず、映像中の区間 $t \sim t+T$ の各フレーム画像内をブロック分割し（区間長 T およびブロック分割数は予め設定）、全フレームの各ブロックにおいて、色特徴とエッジ量からなる特徴ベクトルを求める。そして、ブロックごとにそれらの特徴ベクトルの $t \sim t+T$ の平均値を計算し、 $V(t)$ を構成する。

3.2. ニュース映像の各シーンからの特徴計算

検索の「クエリ」は、ニュース映像の打席単位の各シーンとした。そこから算出される特徴ベクトルについて述べる。本稿では、投球ショットと打球ショットの類似性に基づいた検索を行うため、これら2つのショットより特徴ベクトルを求めるのが基本方針である。しかし図1にも示したが、ニュース映像のシーンのショット系列は、試合映像とは異なり、投球と打球ショットが連続しているとは限らず、その間に選手アップなどの放送以外の素材からの映像が0~2ショット挿入される。そこで本稿では、上記を踏まえ、打球ショット用特徴ベクトルは複数のショット（第2~4ショット）から算出しておくこととする。なお、ショット分割および投球ショットの自動検出は、良好な性能を示す[1][2]の手法を用いた。

† NHK 放送技術研究所 ‡ 東京工業大学

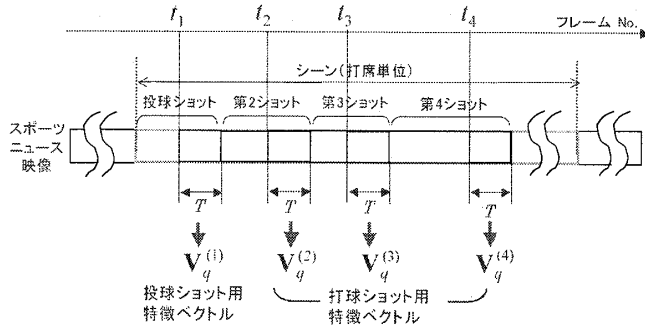


図2 ニュース映像からの特徴ベクトル算出

図2に、ニュース映像の1シーンからの投球ショット用および打球ショット用特徴ベクトル算出の流れを示す。まず、シーンの第1～第4ショットまでのショット終了点から T フレーム遡った時点それぞれ t_k ($k=1,2,3,4$)とし、各 t_k において3.1節の手法で算出した $V(t_k)$ を $V_q^{(k)}$ と記す。そして、 $V_q^{(1)}$ をクエリーにおける投球ショット用の特徴ベクトル、 $V_q^{(2)}$ 、 $V_q^{(3)}$ 、 $V_q^{(4)}$ を打球ショット用特徴ベクトルとする。これらの特徴ベクトルは、各ショットの末尾 T フレームから算出しているが、その理由は、ショットの冒頭よりも末尾のほうがショット毎の画像特徴に関する「分離性」が高いからである。例えば、サードに打球が飛ぶショットとレフトに飛ぶショットでは、ショットの先頭部分では絵柄が類似するケースがある。しかし、ショットの終盤では内野と外野という異なる絵柄となり明確に区別することが可能となる。

3.3. 試合放送映像に対する検索処理

3.2節の手法で求めた $V_q^{(k)}$ ($k=1,2,3,4$)を用いて試合映像から所望の映像区間を検索する。基本的な流れは、まず打球ショットの類似性で候補区間を絞り込み、次に投球ショットの類似性により特定する。最初の絞り込み処理を打球ショットにより行う理由は、投球ショットは構図が酷似しており、類似度に差が生じにくいからである。図3に処理の概要図を示し、その詳細を以下に記述する。

- 試合映像の全ての打球ショット（試合映像では投球ショットに続く第2ショット）について以下の処理を行い、「候補打球ショット」を列挙する。
 - $T_1 \leq t \leq T_2 - T$ の区間 (T_1 : 打球ショット開始フレーム、 T_2 : 終了フレーム)で t をシフトさせながら、 $V(t)$ と $V_q^{(k)}$ とのユークリッド距離に基づいた類似度 $S_k(t)$ を計算し、この区間における $S_k(t)$ の最大値を S_k とする。 ($k=2,3,4$)
 - $S = \max\{S_2, S_3, S_4\}$ とし、 S がある閾値を超えた場合、この打球ショットを「候補打球ショット」に加える。
- 候補打球ショットが0の場合は、ニュース映像の打球ショットが試合映像から切り出したものではないと判断し、やむを得ず投球ショットのみを手がかりとして検索する。そこで、以降の処理のために、全ての打球ショットを候補打球ショットとする。
- 各候補打球ショットの直前の投球ショットについて、 $T'_1 \leq t \leq T'_2 - T$ (T'_1 : 投球ショット開始フレーム、 T'_2 : 終了フレーム)の区間で t をシフトさせながら、 $V(t)$ と $V_q^{(1)}$ との類似度 $S_1(t)$ を計算し、この区間における $S_1(t)$ の最大値 S_1 を求める。

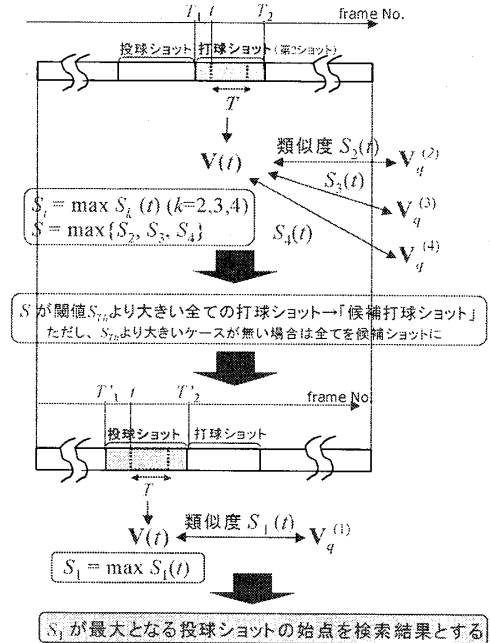


図3 試合放送映像の検索処理

- S_1 が最大となる投球ショットの先頭フレーム番号を検索結果として返す。

4. 実験結果

今回は、ニュースおよび試合映像ともフレーム画像サイズを 320×240 (pixel)、ブロック数を 4×4 とし、 $T=20$ とした。実験で用いたニュース映像は3番組分(ニュース1,2,3とする)で、シーン構成は以下のものであった。

- ニュース1: 本塁打, 遊ゴロ, 中フライ, 三振
- ニュース2: 二塁打, 中前打, 盗塁
- ニュース3: 満塁本塁打

これら3つのニュース映像それぞれについて、検索対象となる試合映像(日付と対戦カードで自動特定可能)に対して検索実験を行ったところ、8つのシーン全てについて、対応する試合本編シーンを見つけることができた。また、検索速度(ニュース、試合とも特徴ベクトルは予め計算済)は各ニュースにつき5～15秒程度であり、実用面でも問題の無い速さと思われる。

5. あとがき

本稿では、スポーツニュースの大リーグダイジェスト映像の各シーンから、投球および打球ショットが高い確率で使用されるなどの映像編集の傾向を利用し、対応する試合放送映像の同一シーンを検索する手法について述べた。

今後は、ニュースのシーンと全く同じ内容の試合映像シーンのみを出力とするだけでなく、絵柄や動きが似たシーン、プレイ内容が近いシーンなども二次的な出力結果とするなどの機能拡張も考えたい。また、野球だけでなく、その他のコンテンツへの応用も検討を進める。

参考文献

- Y.Kawai et al: "Fast detection method for shot boundary including gradual transition using multiple features", IEICE Tech. Rep., CS2007-53, pp. 141-146, Dec. 2007.
- Wen-Nung et al: "Pitching Shot Detection Based on Multiple Feature Analysis and Fuzzy Classification". Proc. of Pacific-Rim Conference on Multimedia (PCM2006), 2006, pp. 852-860.