

## 投球軌道と捕手の動作特徴に基づく野球の球種識別

高橋 正樹<sup>†</sup> 藤井 真人<sup>†</sup> 八木 伸行<sup>†</sup>

<sup>†</sup> NHK 放送技術研究所 〒157-8510 東京都世田谷区砧 1-10-11

E-mail: <sup>†</sup> {takahashi.m-iu, fujii.m-ii, yagi.n-iy}@nhk.or.jp

あらまし プロ野球ではカウントや球速、コースなどのメタデータが一球ごとに作成され、データ放送やインターネットを通じて配信されている。中でも球種は視聴者の関心を集める情報であるが、自動判定が困難なことから野球の知識・経験のある専門家により手動判定されている。そこでプロ野球の放送映像から得られる情報を用い、球種を自動判定する手法を検討している。オブジェクト追跡により得られるボールの軌跡データに加え、SIFT 特徴量の追跡により得られる捕手の動き、およびテロップ認識により得られる球速値を特徴として Random Forests 識別器で識別したところ、ストレート、シュート、スライダーなど個別球種の識別精度が向上した。

キーワード 球種識別, オブジェクト追跡, SIFT 特徴量, 文字認識, ランダムフォレスト

### Baseball pitch identification based on ball trajectory and catcher motion

Masaki TAKAHASHI<sup>†</sup> Mahito FUJII<sup>†</sup> and Nobuyuki YAGI<sup>†</sup>

<sup>†</sup> NHK Science and Technology Research Laboratories 1-10-11 Kinuta, Setagaya-ku, Tokyo, 157-8510 Japan

E-mail: <sup>†</sup> {takahashi.m-iu, fujii.m-ii, yagi.n-iy}@nhk.or.jp

**Abstract** In professional baseball, metadata such as ball speed and the count are generated with every pitch. That data is delivered by data broadcasting and the Internet. Experts must create the data manually because automatic pitch identification is difficult. We are therefore developing a system that can identify the type of pitch automatically based on features obtained from professional baseball video. The system measures trajectory shape and change of speed. It also measures the motion of a catcher by using SIFT features. It obtains pitch speed by character recognition from broadcast video. The system classifies pitches with high accuracy using these features.

**Keyword** Pitch identification, Object tracking, SIFT features, Character recognition, Random forests

#### 1. はじめに

近年ではデータ放送やインターネットを通じ、スポーツのメタデータがリアルタイム配信されている。特に野球ではカウントやスピード、コースなど多岐に渡るデータが一球ごとに作成されている。

中でも球種情報は打者との勝負を決める大事な要素であり、野球観戦者の関心も高い。しかし球種の自動判定が困難なことから、現在のところ球種メタデータは野球の知識・経験のある専門家により人手で付与されている。

そのため

- ・高精度な判定を行うには習熟が必要
- ・主観が混入する恐れがある

などの問題がある。

そこで、野球放送映像を解析して球筋や球感を判定する様々な研究が行われている[1]-[3]。しかし球種の識別に関しては、直球と変化球の分類程度に留まっており、個別の球種を識別する手法は未だ確立されていない。

また固定カメラを複数台使用し、ボールの3次元位置を求める技術も開発されている[4][5]。この方法はボールの3次元位置計測により高精度な球種識別が期待できるが、システムが大規模となる傾向がある。運用性や即時性を考慮すると、放送映像の情報で球種を判定することが望ましい。

そこで我々は、野球放送映像のみを用いた球種メタデータの自動付与を検討している。これまで、投球軌跡作画装置“B-Motion”[6]から得られる画像座標上のボール位置座標(軌跡データ)を用いて球種を自動識別する手法を考案した[7]。

しかし直球系、曲がる系、落ちる系の3分類では高い精度で識別できたものの、ストレート、シュート、スライダーなど個別球種の識別精度には課題が残った。同系の球種は軌道が似ているため、軌道による識別が困難であることが主な原因である。

そこで今回、球種判定担当者の意見などを参考に、軌道以外にも有効と思われる捕手の動きと球速値を追加して識別実験を行った。捕手の動きは、SIFT 特徴

量を用いて捕手領域を追跡し、定量化した。球速値は、放送映像上にテロップ表示される数値を文字認識により取得した。識別には集団学習アルゴリズムである Random Forests を用いた。その結果、個別の球種に関しても高い識別精度が得られたので報告する。

## 2. 球種の種類

プロ野球には様々な球種が存在するが、スポーツメタデータを配信しているデータスタジアム株式会社（以下 データスタジアム）で付与している球種データは表 1 に示す通りである。

この球種データでは、Type と Kind の 2 種類が存在する。Type データは直球系、曲がる系、落ちる系の 3 種類、Kind データは Type データを細分化した下記の 9 種類である。

Type	直球系		曲がる系			落ちる系			
Kind	ストレート	シュート	カーブ	スライダー	カットボール	フォーク	チェンジアップ	シンカー	特殊球

表 1 球種の種類

## 3. 従来手法

### 3.1. 特徴量

従来手法では、識別特徴量に軌跡データを用いた。軌跡データとは、投球軌跡作画装置“B-Motion”（3.2 参照）から得られるボール位置の時系列データである。図 1 に軌跡データの例を示す。2 次元画像座標上の位置情報ではあるものの、球種によって形状に特徴がみられる。

ただし投球毎のカメラ操作によって被写体の位置が若干変化するため、軌跡データをそのまま特徴量に用いることはできない。そこで撮影方向の影響を受けない以下の 32 次元の特徴量を軌跡データから抽出し、特徴量に用いた。これらの特徴量は、軌道の形状や速度を表している。

- ・軌跡の傾き(1次微係数): 全体, 前半部, 中間部, 後半部
- ・軌跡の曲率(2次微係数): 全体, 前半部, 中間部, 後半部
- ・フレーム間の動きベクトル(水平, 垂直):

前半部, 中間部, 後半部(各 4 フレーム分)



図 1 軌跡データの例

## 3.2. 投球軌跡作画装置“B-Motion”

B-Motion は、映像内からフレーム毎にボール領域を抽出し、その位置に応じた軌跡 CG を実映像上に合成するシステムである。リアルタイム処理が可能であり、投球直後に軌道を表示できる。NHK のプロ野球中継で利用されており、コースや球筋の分析に威力を発揮している。図 2 に B-Motion の出力映像例を示す。



図 2 B-Motion での軌跡作画例

図 3 に B-Motion によるボール抽出・追跡処理の流れを示す。センターカメラは投球時に操作を行わないため、固定撮影映像とみなせる。そこでまずフレーム間差分を取り、ボールの候補となる動オブジェクトを抽出する。得られた候補領域を色、面積、形などの画像特徴フィルタに通し、最もボールに類似度が高い領域を選定する。選んだ領域の重心をボール位置座標として出力する。

ボールの抽出後、カルマンフィルタを用いて次フレームでのボールの位置、存在範囲を予測する。これらの値からボールの探索領域を定め、処理範囲を限定することで高速化、ロバスト化を実現している。

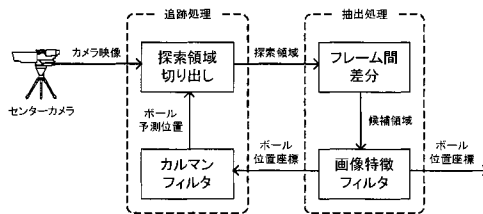


図 3 抽出・追跡処理の流れ

ただし全フレームでボール位置を取得できるとは限らない。左バッターや広告領域など、ボールと輝度・色が近い背景上をボールが通過する際は、抽出に失敗する可能性がある。

そのため B-Motion は予測に基づく再補足機能、およ

び内挿補間機能を備えている。抽出失敗後も予測を継続することで、抽出困難な背景通過後に再補足することが可能である。さらに抽出に失敗した区間のボール位置は内挿補間されるため、途中の抽出結果に関わらず、欠落のない軌跡データを作成できる。

### 3.3. 課題

軌跡データから抽出した特徴量を集団学習アルゴリズムで学習し、識別器を作成した。実験では Type データ 3 分類で高い識別精度が得られたものの、Kind データ 9 分類の識別は低い精度に留まった。投球数の少ない球種はより投球数の多い同系球種に誤判定される場合が多いため、個別球種の再現率向上が課題として残った。

そこで球種判定担当者の意見などを参考に、軌跡データ以外から得られる特徴量を検討した。次章では新たに追加した特徴量について述べる。

## 4. 追加した特徴量

### 4.1. 捕手の動作特徴量

通常、キャッチャーは投球前にサインを出してボールのコースや球種を決定し、狙うべきコースへミットを構える。この動作を分析することで、球種をある程度絞り込むことができる。

例えば低めに構えた場合はフォークの可能性が高く、ピッチャーから見て左に構えた場合はシュートの可能性が低い。そこで捕手領域を追跡し、その動作特徴量を利用することとした。

#### 4.1.1. SIFT 特徴量

特徴点の追跡には SIFT 特徴量を用いた[8][9]。SIFT は対象領域の拡大・縮小、回転および照明変化に対する頑健性を備えており、カメラ操作や輝度調整が生じる本件に適した特徴量である。

SIFT は特徴の抽出に適した点（キーポイント）の検出と、各キーポイントにおけるオリエンテーションの割り当ておよび特徴量記述の 2 段階からなる。各キーポイントにおいて、オリエンテーションとスケールを基準とした  $4 \times 4$  のブロックを配置し、各ブロック内で 8 方向の勾配ヒストグラムを作成する。この計 128 次元の勾配ヒストグラムが特徴量となる（図 4）。スケールとオリエンテーションを考慮しているため、対象領域の拡大・縮小や回転にも頑健である。

図 5 に SIFT 特徴量を図示した例を示す。各矢印の始点がキーポイントの位置、向きがオリエンテーション、長さがスケールを表している。各矢印は図 4 における最大の矢印に相当する。

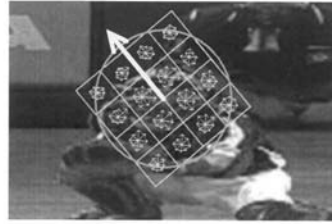


図 4 SIFT 特徴量の概念図



図 5 SIFT 特徴量の表示例

#### 4.1.2. 捕手領域の切り出し

投手や打者の動きが混入すると捕手の動作特徴量の精度が下がるため、図 6 に示す捕手の動作範囲（以下捕手領域）に限定して処理を行うこととした。

ただし投球の合間のカメラ操作により、捕手領域の位置、サイズは投球の度に若干変化する。そこで広告領域内の特徴点を参照し、捕手領域を自動更新する。



図 6 捕手領域（右側矩形）と広告領域（左側矩形）

まず試合前に投球シーンのテンプレート画像を保存し、捕手領域の座標  $\mathbf{P}_1=(x_{P1}, y_{P1})$ 、 $\mathbf{P}_2=(x_{P2}, y_{P2})$  を手動指定しておく。同時に広告領域内の SIFT 特徴量を求め、SIFT 特徴量のテンプレートとして保存する。

続いて各投球シーンにて広告領域内の SIFT 特徴量を求め、テンプレート画像中の特徴点との対応点を求める。時刻  $t$  の現フレーム画像から  $M$  個の特徴点が抽出され、テンプレート画像から  $N$  個の特徴点が抽出されたとすると、2 枚の画像の特徴点は以下ようになる。

$$S_t^M = (S_1^M, S_2^M, \dots, S_{128}^M)$$

$$W_{temp}^N = (W_1^N, W_2^N, \dots, W_{128}^N)$$

現フレーム画像中の特徴点 (m 番目) と、テンプレート画像中の特徴量とのユークリッド距離が最も小さくなる特徴点  $n'$  を下式により求め、対応点とする。

$$n' = \underset{n \in N}{\operatorname{argmin}} \sum_{i=1}^{128} \sqrt{(s_i^m - w_i^n)^2}$$

続いて対応点の位置関係から拡大・縮小、平行移動に関するアフィン行列  $A$  を算出する[10]。求めた行列でベクトル  $\rho_1 = (x_{P1}, y_{P1}, 1)$ ,  $\rho_2 = (x_{P2}, y_{P2}, 1)$  を座標変換し、現投球シーンにおける捕手領域  $P_1', P_2'$  を定める。

$$\rho' = \rho A \quad A = \begin{bmatrix} a & 0 & 0 \\ 0 & d & 0 \\ t_x & t_y & 1 \end{bmatrix}$$

#### 4.1.3. 捕手動作の計測

捕手領域決定後、フレーム毎に時刻  $t$  と  $t-1$  における捕手領域内の SIFT 特徴量を求め、4.1.2 同様、対応点 ( $P_t^i, P_{t-1}^i$ ) を求める。

下式に従い、各対応点の動きベクトルの平均を求め、時刻  $t$  における捕手領域内の平均動きベクトル  $v_t$  を求める。最後に静止状態 (時刻  $t=0$ ) からミットを構えるまで (時刻  $t=m$ ) の動きベクトルの総和を求め、捕手動作を表す動きベクトル  $v$  とする。フレーム毎の動きベクトルの推移を示した例を図 7 に示す。

$$v_t = \frac{\sum_{i=1}^{N_t} (P_t^i - P_{t-1}^i)}{N_t} \quad v = \sum_{t=1}^m v_t$$

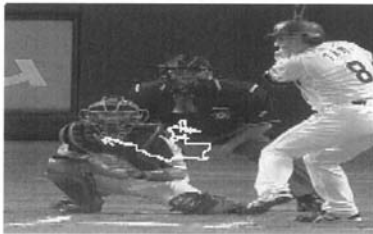


図 7 捕手領域動きベクトルの推移

#### 4.2. 球速値

プロ野球の放送映像には、スピードガンで計測した球速テロップが即座に表示される。球速テロップの例を図 8 に示す。

球種により球速が変化するため、球速値は球種識別に有効であることが予想される。そこで画像中の文字を自動認識し、球種識別に用いることとした。

先行研究でも述べられているが、一般にテロップ領域には以下のような特徴がある[11][12]。

- ・ 強いエッジを有する
- ・ 輝度が高い
- ・ 色が均一
- ・ 一定時間表示される
- ・ 画像座標上の位置が不変
- ・ 文字列が縦または横方向に並ぶ



図 8 球速テロップ

これらの特徴を基に、映像内からテロップ領域を切り出す。その流れを図 9 に示す。まず現フレーム (時刻  $t$ ) から  $m$  フレーム前までの画像それぞれについて、エッジ画像  $E$  を作成する。処理には Prewitt オペレータを用いた。エッジ画像の例を図 10 に示す。続いて隣り合う 2 枚のエッジ画像の差分が小さく、且つ強いエッジを有する画素を真、その他を偽と 2 値化した画像  $S$  を作成する。

$S_t \sim S_{t-(m-1)}$  間で連続して真値を持つ画素を注目画素とする。縦方向および横方向に注目画素をカウントし、両方向でカウント数がしきい値を超えた画素をテロップ画素とし、それらを包含する矩形領域をテロップ領

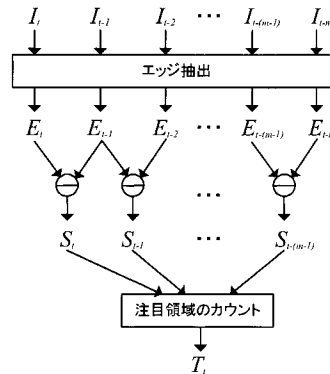


図 9 テロップ領域抽出の流れ

域とする。図 11 にテロップ領域の検出状況を示す。水平、垂直のラインが交差している部分がテロップ画素となる。

最後に、輝度しきい値で 2 値化したテロップ領域を OCR プログラムへ入力する。OCR は Panasonic 製カラー OCR ライブラリー[13]を活用した。OCR により画像内の文字、数値が自動認識され、球速値が得られる。



図 10 エッジ画像（中央下が球速テロップ）



図 11 球速テロップ領域の検出

## 5. 識別器の作成

球種識別器の作成には教師付き集団学習アルゴリズムである Random Forests を用いた[14]。Random Forests は高精度、高速、特徴量の重要度を算出可能、学習データのノイズに頑健、などの特徴を有する。特に軌跡データには誤抽出によってノイズが混入する恐れがあるため、本件に適した学習手法である。

図 12 に Random Forests による学習・識別の流れを示す。まず母集団データから無作為復元抽出によるサブデータセットを多数作成する。続いて各データセッ

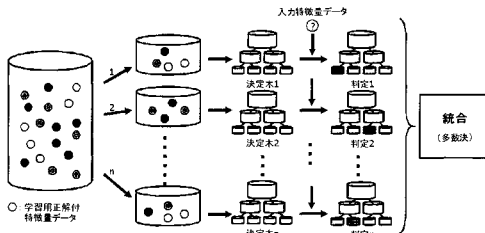


図 12 学習・識別の概要

トにおける決定木を作成する。分岐の各ノードで候補変数をランダムに選ぶことで、互いに異なる決定木が作成される。最終的に各決定木の判定の多数決で球種を識別する。

## 6. 実験

提案手法の有効性を検証するため、NHK プロ野球中継の映像から 5 試合、447 球分の投球シーンを用いて実験を行った。

ピッチャーにより球速、持ち球、変化球の曲がり具合などが異なるため、持ち球が多く、球種の割合が比較的均一なピッチャー 1 人を選定した。対象投手の持ち球はストレート、シュート、カーブ、スライダー、カットボール、フォークの 6 種類である。またカメラ位置によって軌道の形状が変化するため、全て同じ球場での映像を用いた。

1 シーンあたりの特徴量は、従来特徴量 32 次元に捕手動作ベクトル (水平、垂直)、球速値を加えた 35 次元とした。正解データにはデータスタジアムが作成した球種データを用いた。

全データの 2/3 を学習、1/3 をテストデータとしたクロスバリデーションを 5 度行い、作成した識別器を評価した。表 2 に従来手法の特徴量を用いて評価した結果を示す。また表 3 に捕手動作、球速値を含めた提案手法で評価した結果を示す。

提案手法は従来手法に比べ、正答率が 7.5% 向上した。また提案手法では全ての球種において F 値が 50% を超えた。特に従来手法では投球数の少ないスライダーやシュートなどの再現率が低かったが、提案手法にて大きく改善された。これらにより、新たに提案した特徴量の有効性が示された。

	Straight	Shoot	Curve	Slider	Cutball	Fork	Recall (%)
Straight	507	32	0	0	21	5	89.7
Shoot	122	122	0	0	17	14	44.4
Curve	0	0	185	0	0	0	100.0
Slider	0	0	8	12	232	8	4.6
Cutball	40	11	4	53	622	5	84.6
Fork	2	18	0	1	48	146	67.9
Precision (%)	75.6	66.7	93.9	18.2	66.2	82.0	71.3

表 2 球種識別結果 (従来手法)

	Straight	Shoot	Curve	Slider	Cutball	Fork	Recall (%)
Straight	338	38	0	0	24	0	84.5
Shoot	76	120	0	0	26	3	53.3
Curve	0	0	155	0	0	0	100.0
Slider	0	0	4	88	102	16	41.9
Cutball	26	0	0	31	573	0	91.0
Fork	0	0	0	5	34	156	80.0
Precision (%)	76.8	75.9	97.5	71.0	75.5	89.1	78.8

表 3 球種識別結果 (提案手法)

Random Forests による決定木作成の分岐指標は gini 係数を用いたが、分岐による gini 係数の減少量で各特徴量の重要度を評価できる。図 13 に特徴量の重要度を示したグラフを示す。今回用いた特徴量では、球速値、軌道全体の曲率、軌道後半の傾きの順に高い結果が得られた。

目視で球種判定を行う担当者は、投手の持ち球やカウント、捕手動作などで球種の候補を絞り込んだ後、最終的に球速値や軌道の形状で判定を下している。今回得られた重要度は、担当者の視点と概ね一致している。

また軌跡データから得られる軌道の形状や動きベクトルは、B-Motion のボール抽出精度に影響を受ける。ボール追跡中、誤抽出や位置ズレが生じた投球では識別精度が低下する。一方、球速値は抽出精度の影響を受けることなく、安定して取得できる。この点も、球速値の重要度を高めている要因であると考えられる。

捕手動作の重要度は突出した値ではないが、捕手動作を計測できないシーンが多数存在したことが一因である。放送映像では他カメラからセンターカメラ映像へ切り替わるタイミングが捕手動作後になる場合があり、この場合捕手動作特徴量が得られない。ただし目立った重要度ではないものの、捕手動作に特徴が出やすい変化球の識別率向上に貢献している。例えばシュートは捕手が右方向に移動してミットを構える傾向があるが、捕手動作特徴量を加えることで、再現率が 3.1%、適合率が 1.4% 向上した。

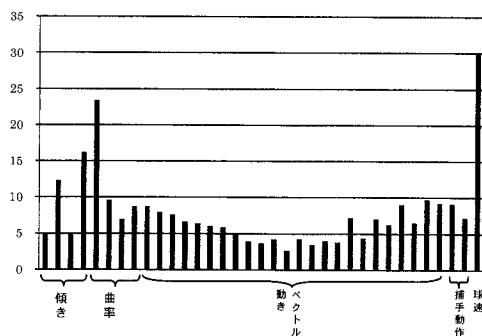


図 13 特徴量の重要度

## 7. まとめ

野球中継における球種の自動識別手法を提案した。これまでは投球軌跡作画装置から得られる軌跡データだけの情報を用いて球種を判定していたが、新たに SIFT 特徴量の追跡による捕手の動作特徴量と、テロップ認識による球速値を用いた。

これらの特徴量を集団学習アルゴリズム Random Forests で学習し、識別器を構築した。実際の中継映像でテストしたところ、78.8%の精度が得られた。各球種においても F 値が全て 50%を超えており、個別の球種でも高い識別精度が得られた。

今回、上手右投げ投手 1 人に限定して実験を行ったが、今後左投げやアンダースローなど、対象投手の数を増やして汎用性を高める予定である。またリアルタイム処理を目標にアルゴリズムを最適化し、プロ野球中継での放送応用を目指す。

## 参考文献

- [1] Hubert Shum, Taku Komura, "A spatiotemporal approach to extract the 3D trajectory of the baseball from a single view video sequence", ICME'04, pp.1583-1586, 2004.
- [2] Hua-Tsung Chen, "A Trajectory-Based Ball Tracking Framework with Enrichment for Broadcast Baseball Videos", ICS2006, pp.1145-1150, 2006.
- [3] Wei-Ta Chu, Chia-Wei Wang and Ja-Ling Wu, "Extraction of Baseball Trajectory and Physics-Based Validation for Single-View Baseball Video Sequences" ICME'06, pp. 1813-1816, 2006.
- [4] P. Rander, "A Multi-Camera Method for 3D Digitization of Dynamic, Real-World Events", doctoral dissertation, tech. report CMU-RI-TR-98-12, Robotics Institute, Carnegie Mellon University, 1998.
- [5] Gueziec, A, "Tracking pitches for broadcast television", IEEE. Computer, Vol.35, No.3, pp.38-43, 2002.
- [6] 高橋他, "画像内の物体抽出技術を用いた高速投球軌跡作画手法", 信学論, Vol.J88-D II, No.8, pp.1672-1680, 2005.
- [7] 高橋他, "B-Motion の軌跡データに基づく野球の球種識別手法", FIT2007, H-63, 2007.
- [8] 都築他, "SIFT 特徴量に基づく Mean-Shift 探索による特徴点追跡", 情処学研報 CVIM 157, pp. 101-108, 2007.
- [9] 市村, "局所不変特徴量に基づく複数広告看板の認識", 情処学研報 CVIM 151, pp. 123-130, 2005.
- [10] 興沼他, "入力映像とパノラマ画像の射影変換を用いた位置合わせパラメータの高速推定手法", 信学技報, PRMU2000-153, pp.47-52, 2001.
- [11] 新井他, "映像中のテロップ表示フレーム検出方法", 信学論, Vol.J83-D II, No.6, pp.1477-1486 (2000)
- [12] 芦田他, "情景画像からの文字抽出", 信学論, Vol.J88-D II, No.9, pp.1817-1824, 2005.
- [13] カラーOCR ライブラリー  
["http://panasonic.co.jp/pps/pstc/products/coloroclib/"](http://panasonic.co.jp/pps/pstc/products/coloroclib/)
- [14] Leo Breiman, "Random Forests", Machine Learning, Vol.45, pp.5-23, 2001.