

令和2年度卒業論文

野球の投球軌跡からの球種判定

宮崎大学工学部 情報システム工学科

小林 稜

指導教員 椋木 雅之

目次

1. はじめに.....	1
2. 球種の自動判定の関連研究.....	2
3. 球種判定手法.....	3
3.1. ボールの座標取得.....	3
3.2. ボール座標の時間合わせ.....	4
3.3. ボール座標の位置合わせ.....	5
3.4. SVM による判定.....	7
4. 実験.....	8
4.1. 実験の設定.....	8
4.2. 実験 1.....	9
4.3. 実験 2.....	12
4.4. 実験 3.....	14
5. 終わりに.....	17
謝辞.....	18
参考文献.....	19

1. はじめに

野球において変化球は戦略的に欠かせない要素であり，視聴している側からも注目される事項である．一方で，高速で飛ぶボールの球種は目視での判断が非常に難しく，専門の人員がいなければすぐに知ることはできない．球種の判定に関する研究には高橋ら[1]が提案した放送カメラ映像から投球の軌跡を抽出するシステムを用いた手法がある．この研究では放送映像から軌跡データを生成し，軌跡データから特徴量を抽出する．そして抽出した特徴量を用いて機械学習を行い球種を識別している．一方で軌跡を用いた研究に松下[2]によるストライク・ボール判定がある．こちらは軌跡情報を時間合わせ、位置合わせした上で直接用いて判定を行っている．本研究は松下の軌跡情報を直接用いる手法が球種判定においてどの程度有効かを評価する．

2. 球種の自動判定の関連研究

高橋らの研究[1]では野球の放送映像から投球の球種判定を行っている。B-Motion と呼ぶシステムを用いて放送映像中の投球の軌跡データを求める。軌跡データから球種判定に有効な特徴量を抽出し、機械学習で識別器を作成する。ここで抽出される特徴量はピッチャーがボールから手を放してからキャッチャーが捕球するまでのフレーム数、軌跡の1次微係数、ホームベースから捕球位置までの距離などの計35次元で人が定めたものである。この研究では7球種での判定は81.4%、直球系、曲がる系、落ちる系に分類した3種では92.3%という結果を出している。

他の野球の投球軌跡を用いた研究として、松下[2]によるストライク・ボール判定がある。この研究では、センターバックスタンドまたはバックネット裏に設置した1台のカメラで撮影した映像から判定を行う。カメラ1台では、ボールの正確な3次元情報は得られないが、このようにカメラを設置することでピッチャーがボールを離してから、キャッチャーが捕るまでの時間情報が奥行きに対応し、疑似的に3次元の情報が得られる。得られた軌跡情報を時間合わせ、位置合わせした上で直接用いて、ストライク、ボール判定を行っている。

3. 球種判定手法

本研究は軌跡を直接用いる松下の手法が球種判定においてどの程度有効かを評価するものであり，基本的に松下の手法を利用している。以下、文献[2]に従って手法を説明する。

球種判定を行う流れは以下の通りである。

- (1) ボールの座標取得
- (2) ボール座標の時間合わせ
- (3) ボール座標の位置合わせ
- (4) SVM による判定

3.1. ボールの座標取得

ピッチャーの投球モーション中，手からボールが離れた瞬間からボールがキャッチャーミットに収まるまでの映像をフレーム単位で分割し，フレーム毎にボールの中心の x ， y 座標を取得する。

3.2. ボール座標の時間合わせ

1 投球から取得できるボール座標の数は，球速，変化球の球種によって増減する．時間情報を奥行き情報とみなし，同じ奥行きでのボール座標の比較により判定が行えるように，判定に利用するボール座標を補間により揃える．ボールがピッチャーの手から離れた瞬間を始点(時刻 t_0)，ボールがミットに収まる瞬間を終点(時刻 t_N)とし，始点から終点までのフレーム数を $N + 1$ フレーム(始点，終点を含む)とする．始点(時刻 0)から終点までの間を P 等分した P 時刻までの各時刻のボール座標を求める．なおフレーム間においてボールは画像上で等速とする．等分した地点の座標は，時刻 i の前後のフレーム $N_{t_i}, N_{t_{i+1}}$ ($i = 0 \dots P$, t_i は i の値によって $0 \leq t_i \leq N$ の範囲をとる整数)におけるボール座標の内分(線形補間)によって求める．フレーム番号 t_i ，内分比 s_i を次式で求める．

$$t_i = \left[\frac{N}{P} i \right] \quad (\lceil \rceil \text{はガウス記号})$$

$$s_i = \frac{N}{P} i - t_i$$

内分比 s_i と t_i 番目のフレームでの x ， y 座標を用い，時刻 i におけるボール座標 (x_i, y_i, z_i) を次式で求める．

$$x_i = x_{t_i} (1 - s_i) + x_{t_{i+1}} s_i$$

$$y_i = y_{t_i} (1 - s_i) + y_{t_{i+1}} s_i$$

$$z_i = i$$

本研究では $P = 9$ とすることで，10点のボール座標を抽出する．

3.3. ボール座標の位置合わせ

映像を撮影するカメラの位置はグラウンドにより異なり，毎回正しく同じ位置に置くことは難しい．そのため，カメラの位置によらずに判定を行うために座標の位置合わせを行う．座標の位置合わせの手順を次に示す．

まず映像中の基準となる4点を定める．本研究では1点目をピッチャープレーターの前方右端，2点目をピッチャープレーターの前方左端，3点目をピッチャーのリリースポイント，4点目をピッチャーの視点から見たホームベースの頂点の最奥の点と定める．ここでリリースポイントとはピッチャーがボールを投げる瞬間のボールの位置のことを指す．リリースポイントは投手によってズレがあるが，本研究では同じであると仮定する．

カメラ1，カメラ2それぞれについて，これら4点の座標を取得する．これら4点の座標を (x, y, z) とすると， x, y はそれらの点の画像座標 (u, v) を用いる．奥行き z は，ピッチャープレーター前方位置を $z = 0$ ，ホームベース最奥の位置を $z = P$ とする．リリースポイントの z の値は， 0.5 と定めた．

次に，カメラ間での位置合わせの変換を求める．ここで，カメラの位置からボールまでの距離が十分遠く，スケール付きの平行投影とみなせるとする．カメラ1のある基準点を (u_1, v_1, z_1) ，対応するカメラ2の基準点を (u_2, v_2, z_2) とすると，回転行列 R ，並進成分 t を用いて

$$\begin{pmatrix} u_2 \\ v_2 \end{pmatrix} = C(R|t) \begin{pmatrix} u_1 \\ v_1 \\ z_1 \\ 1 \end{pmatrix} \quad C \text{は定数の係数}$$

と表現できる.

$$M := C(R \mid t)$$

と置くと, 変換行列 M は, 2×4 の行列となる. カメラ 1 の基準点を齊次座標に変換して並べた行列を A (A は 4×4 行列), カメラ 2 の基準点の x, y 座標を並べた行列を B (B は 2×4 行列)とすると,

$$B = MA$$

が成り立つ. この式から

$$M = BA^{-1}$$

により, 変換行列 M を求める.

求めた M を使って, ボール座標を位置合わせする. 3.2 節で時間合わせしたボール座標について, 時刻の値を疑似的に奥行き z として用いる. 即ち, ボールがピッチャーの手から離れた瞬間(始点, 時刻 0)では $z = 0$, ボールがミットに納まる瞬間(終点, 時刻 P)では, $z = P$ とし, その間の時刻では, 0 から P を等分した値 z_i を取るものとする. 時間合わせしたボール座標 (x_i, y_i, z_i) に対して, 位置合わせ後の座標 (x'_i, y'_i, z'_i) を以下の式で求める.

$$\begin{pmatrix} x'_i \\ y'_i \end{pmatrix} = M \begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix}, z'_i = z_i$$

3.4. SVM による判定

SVM は与えられた学習データを元に識別器を構築する。その後クラスが未知の評価データを与えると、構築された識別器によって、評価データが 2 種類のクラスのどちらに属するか判定を行う。本研究で SVM の実装に用いた `libsvm` は多値分類に対応しており、ここでのクラスは球種を指す。

本研究での SVM による判定手順は以下のようになる。

- (1) 学習データ及び評価データのスケーリング
- (2) グリッドサーチによるパラメータの最適化
- (3) 学習データを用いた SVM の学習
- (4) 学習した SVM による評価データの判定

まず判定精度の向上のため、学習前に(1)(2)の操作を行う。その後 3.1~3.3 節の手順で作成したボール軌道上の 10 点の x , y 座標を特徴量とした学習データにより SVM の学習を行う。最後に学習した SVM によって評価データがどの球種に属するか判定を行う。

4. 実験

4.1. 実験の設定

実験には 2 人の投手による映像を用いる．映像 A，映像 B はそれぞれ異なるカメラ位置，異なる投手による投球 100 球が記録されている．投手はともに右投げ，センターバックスクリーン側からの放送映像である．実験に使用するデータは 3 節の手順により作成されたものを用いる．球種の正解データは MLB.com[3]の Scoreboard 項目の情報を用いた．

4.2. 実験 1

本節では，映像 A，B から作成した学習データと評価データを用いて実験を行う．1 投手の全投球 100 球を評価データ，学習データとして 50 球ずつ 2 分し，SVM によって学習及び判定を行う．この操作をデータを入れ替えてもう一度行い，評価データ合計 100 球の正答率を計算する．結果を表 1 に示す．なお，球種の内訳は表 2，3 のとおりである．また，各判定の詳細を表にしたものを表 4，5 に示す．

表 1 実験 1 の結果(%)

	前半	後半	認識率
映像 A	76.0	70.0	73.0
映像 B	62.0	74.0	68.0

表 2 映像 A の球種

球種	投球数
A フォーシーム	23
B スライダー	37
C スプリット	27
D シンカー	9
E カッター	3
F カーブ	1

表 3 映像 B の球種

球種	投球数
A ツーシーム	23
B スライダー	33
C フォーシーム	33
D カーブ	5
E カッター	5
F チェンジアップ	1

表 4 映像 A の判定詳細

		正解					
		A	B	C	D	E	F
出力	A	20	0	3	6	0	0
	B	1	37	2	1	2	0
	C	1	0	19	2	1	1
	D	1	0	1	0	0	0
	E	1	1	0	0	0	0
	F	0	0	0	0	0	0

表 5 映像 B の判定詳細

		正解					
		A	B	C	D	E	F
出力	A	9	0	7	0	1	0
	B	0	31	0	1	1	0
	C	10	1	22	0	1	1
	D	0	0	0	4	0	0
	E	4	1	4	0	2	0
	F	0	0	0	0	0	0

映像 A において投球数の少ないシンカー、カッター、カーブの認識率は 0%であった。また、映像 B においても投球数の少ないカッター、チェンジアップの認識率が低い。これらはデータ不足で判定が難しい事が原因と考えられる。

また、表 4 より映像 A においてスプリットとシンカーをフォーシームと誤認識している場合が多く、逆のパターンは少ない事がわかる。この投球を目視で確認したところ、ファウルなどでボールがバットに当たっていた場合が多くみられた。これによって本来落ちる球種である 2 球種の座標の最終地点が短縮されたことで本来描く予定だった軌道を描かず、誤認識を起こしたと考えられる。

表 5 より映像 B において 5 投球しかないカーブの認識率が高く、ほかの球種に比べて軌道による認識に向いている事がわかる。カーブは大きく曲がる球種なのが要因と考えられる。ほかには、映像 B においてツーシームとフォーシームはデータ数が多いにもかかわらず誤判定をしていることからかなり似た軌道をしていることが考えられる。

4.3. 実験 2

実験 1 で投球数が少ない球種の判定の認識率が低いという結果があった。本節ではそれぞれ投球数の少ない球種を省いて実験を行う。映像 A はフォーシーム、スライダー、スプリットの 3 球種 87 投球、映像 B はツーシーム、スライダー、フォーシームの 3 球種 89 投球のみを用いる。結果を表 6 に示す。また、各判定の詳細を表 7, 8 に示す。

表 6 実験 2 の結果(%)

	前半	後半	認識率
映像 A	97.7	83.7	90.7
映像 B	77.8	81.8	79.8

表 7 映像 A の判定詳細

		正解		
		A	B	C
出力	A	23	0	3
	B	0	38	1
	C	1	0	21

表 8 映像 B の判定詳細

		正解		
		A	B	C
出力	A	10	0	13
	B	0	33	0
	C	5	0	28

表 6 より，映像 A，映像 B いずれについても実験 1 よりも認識率が向上した．これは実験 1 で考察した通りデータの少ない球種を減らしたためといえる．映像 A において 2 回の判定で大きく結果に差が出た．これはデータの分け方に問題があったと考えられる．

また，表 7 においてスプリットのスライダーとの誤判定について，映像 A を確認したところ，目視ではかなり判別が難しく，撮影の角度で横方向の変化が潰れて、映像上ではほぼ差がなくなっていた．また，表 8 より実験 1 同様映像 B においてフォーシームとツーシームの誤判定が多くみられた．両者の投球軌跡が酷似していると考えられる．

4.4. 実験 3

実験 2 でデータの分け方に問題があったと考察した. 本節ではこの問題を解決するために `leave_one_out` 法を用いた. 具体的にはデータのうち 1 つだけテストに使い, 残りのデータを学習に用いる. これを投球数分だけ繰り返していき, 評価データの正答率を計算する. 実験 1, 実験 2 を `leave_one_out` 法で再度評価した結果を表 9 に, 各判定の詳細を表 10, 11, 12, 13 に示す。

表 9 より, 実験 1, 2 の結果と比較すると大きく認識率が伸びているのがわかる. これよりデータの分割方法に問題があったことがわかる. また, 学習に使ったデータ数が大きく増加している事も認識率の伸びに影響したと考えられる.

表 10 で, 実験 1 では認識率 0%だった D:シンカーの認識率が改善している. シンカーのデータ数が少なかったため, 実験 1 では認識できていなかったが, 本実験では `leave_one_out` 法により, データの少なさをカバーできた.

表 11, 12, 13, では, 認識率は向上しているが, 球種判定の誤り方に実験 1, 2 から大きな変化はない. 実験 1, 2 はデータ数が少なかったことが問題点であり, `leave_one_out` 法によりこの問題点は解消できるといえる。

表 9 実験3の結果(%)

	認識率	
	全球種	3球種
映像 A	83	94.2
映像 B	80	87.6

表 10 映像 A の判定詳細

		正解					
		A	B	C	D	E	F
出力	A	23	1	0	3	0	0
	B	0	33	2	0	2	0
	C	0	2	23	2	1	1
	D	0	0	1	4	0	0
	E	0	0	0	0	0	0
	F	0	1	1	0	0	0

表 11 映像 B の判定詳細

		正解					
		A	B	C	D	E	F
出力	A	15	0	4	0	1	0
	B	0	32	0	0	1	0
	C	7	0	26	0	1	1
	D	0	0	0	5	0	0
	E	1	1	2	0	2	0
	F	0	0	1	0	0	0

表 12 映像 A 3 球種の判定詳細

		正解		
		A	B	C
出力	A	23	1	0
	B	0	35	3
	C	0	1	24

表 13 映像 B 3 球種の判定詳細

		正解		
		A	B	C
出力	A	16	0	4
	B	0	33	0
	C	7	0	29

5. 終わりに

本研究は軌跡を直接用いる松下の手法が球種判定においてどの程度有効かを評価した。実験 2 において実験 1 よりも認識率が向上したが、どれも先行研究と比べ、認識率が高いとは言えなかった。本研究の判定方法の問題は、似た軌道の誤判定率が高いということ、変化が後半に偏る球種がバットに当たった場合にうまく認識できない可能性があることである。今後の課題としては、サンプルを増やし、似た球種の差がどこにあるのか、バットに当たった座標以降の変化の予測を立てる方法などを詳しく調査する必要があると考えられる。また、自身の撮影した映像を用いる場合に関しては、適切な撮影角度を模索する必要があると考えられる。さらに、学習と評価で異なる投手の映像を使った評価も必要である。

謝辞

本論文の執筆にあたり、終始熱心なご指導をいただいた指導教員の椋木雅之教授に感謝いたします。椋木雅之教授は本研究の着想から論文の執筆まで多くのご指導をいただきました。そして日々の活動において様々な支援をくださった椋木研究室の皆様、障がい学生支援室の皆様に心よりお礼を申し上げます。また心身を支え、見守ってくださった家族に深く感謝申し上げます。

参考文献

- [1] 高橋 正樹, 苗村 昌秀, 藤井 真人, 八木 伸行, “B-Motion の軌跡データに基づく野球の球種識別手法”, H-063,FIT,2007
- [2] 松下 功幸, “機械学習を用いた野球映像からのストライク・ボール判定”, 宮崎大学工学部 卒業論文, 2017
- [3]MLB.com, “Scoreboard”, <https://www.mlb.com/scores>