

野球のバッティング上達支援のためのスイング解析

濱田 大聖[†] 中島 克人[‡]

東京電機大学 未来科学部 情報メディア学科^{†‡}

1 はじめに

近年のプロ野球界では、高精度のレーダーやセンサーを用いたデータ分析によるプレー上達支援が行われている。一方、部活動や少年野球界では、予算や熟練した指導者の不足等により、上達に向けての効率的かつ適切な指導・練習が十分に行われている事は少ないと思われる。従って、IT 技術を活用した安価な上達支援システムの導入は非常に有用である。バッティングのスイング解析に関しては、既に実用化されているものが幾つかあるが、打者に何らかのセンサーを装着するの必要があり、試合中の解析が不可能である。

そこで我々は、学校の部活動を含めたアマチュア選手のバッティングの上達支援を目的に、近年の深層学習技術等を活用した映像のみによるバッティングのスイング解析手法を提案する。

2 関連研究

映像によるスイング解析研究の1つに、バッティングティー状の打撃目標に対する素振りの精度を可視化する手法 [1] がある。民生用高速度カメラで撮影した動画像から打撃目標位置であるインパクト位置とその前後のバット姿勢を推定し、素振りをしたユーザにその結果を表示するものである。映像のみによるスイング解析という目的は本研究と同様であるが、打撃目標位置にバッティングティー状のものを立てる必要がある他、素振り練習用であり、実際のボールのバッティング解析には用いる事はできない。

3 提案手法

入力は、スマートフォン等の解像度で撮影され、対象となる打者、ボールとバットがそれぞれ1つのみ映っている60 fpsの映像を使用する。映像には、図1に示すように打者の前方約2mの範囲が映っていることを前提とする。また、提案システムの処理の流れを図2に示す。

まず、物体検出器を用いてボールとバットの領域を検出する。検出器には学習済みモデルであるYOLOv7 [2]に追加学習を行って使用する。

3.1 インパクトの推定

インパクトの瞬間と座標を推定するために、横軸を

Baseball Swing Analysis for Improvement Support

[†] Taisei Hamada · Tokyo Denki University

[‡] Katsuto Nakajima · Tokyo Denki University



図1: 入力画像の例

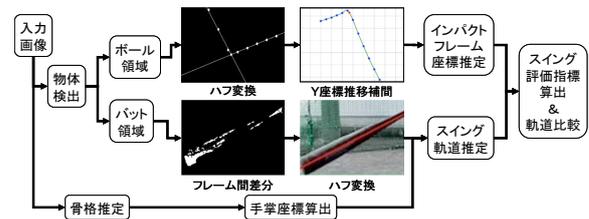


図2: 処理の流れ

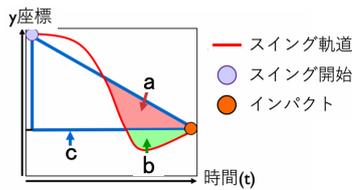
フレーム数、縦軸をx座標とした1枚の2値画像上に検出されたボールの中心座標をプロットする。投球されたボールもバットに当たって飛んでいくボールもインパクト前後では直線軌道であると仮定し、インパクト前後のボール座標をハフ変換し、2本の直線を検出する。そして、その2直線の交点座標からインパクト時のx座標を、また、ボールのy座標推移からインパクト時のy座標を得る。インパクトの瞬間は映像のフレーム番号(以後、インパクトフレームと呼称)で示す。インパクトを捉えた映像フレームが得られるとは限らないため、上記で求めたx座標を元に、インパクト前後の映像フレーム番号の間的小数点数として求める。

3.2 スイング軌道の推定

検出したバット領域におけるフレーム間差分から2値画像を生成し、ハフ変換によってバットの線分を検出する。入力画像に対して骨格推定モデルのOpenPose [3]を用いて手首を検出し、バットの線分端と手首座標の距離からバットの先端と根元を区別する。先端の座標推移をスプライン補完することによってスイング軌道を推定する。

3.3 スイング評価指標の定義と利用

スイング開始からインパクトまでのバット先端のy座標軌道から評価指標を算出する。スイング軌道をダウン/レベル/アッパーの各スイングに分類するための評価指標として、図3内のSを定義する。図3は縦軸がバットの高さ(y座標)、横軸がスイング開始からインパクトまでの時間(t)を示し、赤線がスイング起動である。a, b, cはそれぞれ図3中の面積であり、



$$S = (a + n \cdot b) / c$$

図3: スイング評価指標

a/c はレベルスイングとダウンスイングの違いを定量評価するための指標, b/c はアッパーの度合いの評価指標となり, S はそれらを統合したものである. n は a とのバランスを取るために b の面積を強調するための係数であり, 事前実験により $n=7$ と定めた.

同じ高さのボールに対するバッティングの自他のスイング評価指標やスイング軌道を比較する事で, スイングの効率性や再現性を確認する事ができ, 上達支援に有用であると考ええる.

4 実験と評価

4.1 学習データセット

学習済みモデルである YOLOv7x に以下のデータにより追加学習を行ってバットとボールの検出器とした. バットとボールが 1 つのみ映っている小学校低学年から大人までのスイング画像を 563 枚収集し, それらの画像に水平反転のデータ拡張を行った上で, 訓練画像 920 枚, 検証画像 103 枚, テスト画像 103 枚に分割した. 解像度は 640×640 px である. データには図 4 のような様々な強度のブラーが発生したバットやボールに対してアノテーションが施されている.



(a) ブラー弱

(b) ブラー強

図4: 学習データ例

4.2 物体検出精度評価

追加学習前後のモデルを用いて, テスト画像 103 枚に対するボールとバットの検出精度を表 1 に示す. ボール, バットともに適合率と再現率を大幅に向上できたことが分かる.

表1: 追加学習前(左)と後(右)の精度

評価指標	ボール	バット
適合率(%)	85.7 → 98.8	90.0 → 97.3
再現率(%)	66.7 → 97.4	81.8 → 84.0

4.3 インパクトフレーム・座標推定精度評価

ボールとバットの検出に成功した 48 本のスイング動画に対してインパクトフレーム推定とインパクト座標推定を行った(表 2). フレーム推定精度の評価においては, インパクト直前と直後のフレームを目視

表2: インパクトの推定精度

	フレーム推定	座標推定
動画数(本)	48	48
推定成功数(本)	37	16
推定成功率(%)	77.1	33.3
中央値	0.49 フレーム	2.16 倍
四分位範囲	0.55 フレーム	2.69 倍

により決定し, 両者のフレーム番号の中間を 0.5 として, 3.1 の手法で推定したインパクトフレームとの差の絶対値を推定誤差とし, それが 1.0 以下の場合を推定成功と定義した. インパクト座標の推定精度評価では, 目視によるインパクト座標と, 3.1 の手法での推定座標との距離をボール直径(インパクト直前までの平均値)で割った値を推定誤差とし, 1.0 以下を推定成功とした.

インパクトフレーム推定は実用に十分な精度となった. インパクト座標推定の精度はやや低いが, ボール 2 個分程度以内で推定できていることがわかった.

4.4 スイング軌道比較

バット先端の座標軌道をバット長で正規化することにより, 他者との軌道比較を可能にした. 図 5 のように, 比較対象として指定した軌道とユーザの軌道を重ね合わせて映像上に表示する.

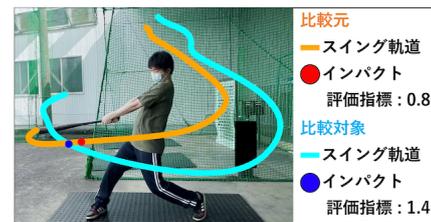


図5: スイング比較例

5 まとめ

本研究では, 物体検出と骨格推定を用いて, スイング映像からのインパクト推定とスイング軌道推定によるスイング解析手法を提案した.

4 章の結果から, 物体検出とインパクトフレーム推定に関しては実用に十分な精度が得られていることを確認できた. インパクト座標推定に関しては, 精度向上のために手法の見直しが必要である.

今回は, バット先端の y 座標軌道をスイング評価指標に用いたが, x 座標軌道を加えたより詳細な評価を行うことも今後の課題である. また, スイング軌道比較は 2 次元平面上で行っているが, 3 次元空間に拡張して解析を可能にすることは今後の目標である.

参考文献

- [1] 新野, 他, “民生用高速度カメラを用いた素振り練習の精度可視化,” エンターテインメントコンピューティングシンポジウム論文集, 2017.
- [2] C. Y. Wang, et al., “YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors,” arXiv preprint arXiv:2207.02696, 2022.
- [3] Z. Cao, et al., “OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields,” CVPR, 2019.