

スポーツシーン照合に関する一検討

A Study on Sport Scene Matching

大澤 遼平[†] 石川 孝明[‡] 渡辺 裕^{†‡}

Ryohei OSAWA[†] Takaaki ISHIKAWA[‡] and Hiroshi WATANABE^{†‡}

[†]早稲田大学 [‡]早稲田大学国際情報通信研究センター
[†]Waseda University [‡]Global Information and Telecommunication
 Institute, Waseda University

Abstract It is useful for finding a difference between two sport motions to compare motions in two different videos. For sport scene matching, we propose the method using body parts coordinates data from sport video. This method enables accurate matching of two sport motions in many cases.

1. はじめに

近年、様々なスポーツにおいて選手の動作を分析するために選手の挙動を撮影した映像データが利用されている。一度撮影した映像は何度でも確認できるため、1つの映像から選手は動作の細部まで解析することが可能となる。一方で、他選手との相違や同一選手間でのわずかな違いを発見するためには選手は複数の映像を比較する必要がある。しかし、比較したい動作の速度や身体の使い方は選手によって異なるため、動作の細部のタイミングを合わせることは難しい。

そこで、本研究では簡易なカメラで撮影された2つのスポーツ映像から選手の身体座標を取得し、その座標データに対して Dynamic Time Warping (DTW)を用いることにより2つの動作の対応付けを行う手法について検討する。

2. 関連研究

2.1. OpenPose

映像内の人物から身体座標を取得する手法として Cao らが提案した OpenPose[1]がある。OpenPose は映像内の複数人の身体部位をリアルタイムに検出可能な手法である。OpenPose はモデル毎に検出可能な部位が異なり、15, 18 もしくは 25 点の位置を推定することができる。この際に、各部位の2次元座標が取得され確度が算出される。

2.2. 始点終点自由な DTW

野球中継映像における身体座標データを用いた投球動作のアラインメントとして始点終点自由な DTW アルゴリズムを用いた手法[2]が存在する。2つの映像でのアラインメントとは、2つの類似度が高くなるように映像内のフレームを対応付けることである。関連手法は正確なアラインメントを実現するために投球動作内の個人依存動作を除去してからアラインメントする。ここで、個人依存動作とは片方の投球動作にのみ存在すると定義されており、OpenPose の検出が失敗し

た部位が存在する動作を含む。関連手法は実際の映像の投球部分から個人依存動作が除去された映像に対してアラインメントする。そのため、選手が動作を比較した際に一部の動作が欠けている場合に違和感を感じる。また、同一の動作を正確に比較できない。

そこで、この欠点を補うために本手法では個人依存動作の除去を行わずに正確なアラインメントを実現する手法を検討する。

3. 提案手法

3.1. ベクトル作成

本研究では、スポーツ映像として野球中継映像を使用する。OpenPose の 25 点が検出可能なモデルを映像に適用すると 25 点の身体座標データが得られる。本論文ではそのうち、首、両肩、両肘、腰の中央、両腰、両膝、両足首の 12 点を用いる。これは、12 点が安定して検出可能なためである。対象とする 12 点の内、任意の 2 つの部位を結ぶベクトルの長さや角度を OpenPose によって得られた 2 次元座標から計算する。この際、各ベクトルの長さは選手の身長やカメラの位置等に依存するため、長さについて正規化する。また、角度は $-\pi$ から π の範囲とする。さらに、各ベクトルについて、使用した 2 部位の確度からベクトルの確度を求める。2 つの部位 p と q からなるベクトルの確度 $C_{\text{vector}_{pq}}$ は式(1)で与えられる。

$$C_{\text{vector}_{pq}} = \min(C_p, C_q) \quad (1)$$

ここで、 C_p, C_q はそれぞれ OpenPose から取得した 2 つの部位 p, q の確度である。

映像の各フレームにおいて身体座標データから長さ、角度、確度を持つ 66 本のベクトルが作成される。

3.2. アラインメント

本研究は 2 つの映像のフレームを対応付けるものである。関連研究で述べた始点終点自由な DTW を用いると、多くの対応付けは正確に行われる。しかし、OpenPose の検出が失敗した部位を含む動作が存在す

る場合、不正確な対応付けを生じる恐れがある。本手法では、対応付けはベクトルを用いて算出されるフレーム間の姿勢の類似度に基づいている。そのため、正確な対応付けを実現するためにフレーム間の姿勢の類似度を算出する際に2つの処理を加える。

まず、ベクトルの確度に応じて処理を分岐する。各フレームは66本のベクトルを持っている。本手法では、対応するベクトルの角度差と長さの差の平均からフレーム間の姿勢の類似度に相当するコストを算出する。この時に常に66本のベクトルの差を計算するのではなく、ベクトルの確度がともに閾値以上の場合のみ差を計算することにより、OpenPoseの検出が失敗した部位によるコストへの影響を少なくする。

また、同一フレームにOpenPoseの検出が失敗した部位が一定数以上存在する場合は検出失敗フレームとしてコストの計算方法を変更する。OpenPoseの検出が失敗した部位が多ければ多いほど、コスト計算に用いるベクトルの数が少なくなるためベクトル1本当たりのコストへの影響は大きくなる。そのため、この場合は直前の対応付けのコストに重み係数をかけることによりコストを決定する。

コスト計算の処理の流れを図1に示す。

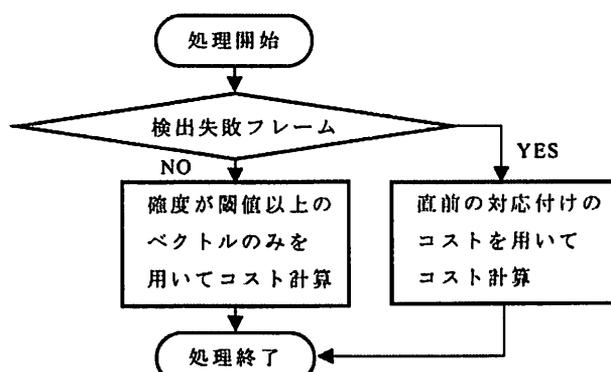


図1 フレーム間のコスト計算の処理の流れ

4. 実験

4.1. 実験手法

本実験で用いたデータセットは野球中継映像における14人の右投手の投球映像89本である。この内、1本をモデル系列とし、入力系列はモデル系列を含む全89本とする。また、モデル系列による違いを確認するために89本の投球映像からモデル系列A、モデル系列Bを選び、それぞれの場合の処理結果を比較する。各映像には1球分の投球シーンが含まれており、2つのモデル系列は目視により投球動作の始点と終点を決定した。なお、投球動作の始点は軸足でない足が地面から離れる瞬間、終点は軸足を蹴り上げた後、その足が地面に着く直前とした。右投手の場合、軸足は右足

となる。入力系列は始点と終点が未知の状態を与え、得られたモデルフレームと入力フレームの対応づけの成否は、より対応付けに適したフレームが存在するかに基づいて目視により判断した。

4.2. 実験結果

各アラインメント結果の全ての対応づけのうち正しい対応づけの割合に応じて89本の結果を分類した。各モデル系列に対する分類結果をそれぞれ表1、表2に示す。

表1 モデル系列Aのアラインメント結果

	95%以上	80%以上	80%未満
系列数	86	3	0
割合[%]	96.6	3.4	0.0

表2 モデル系列Bのアラインメント結果

	95%以上	80%以上	80%未満
系列数	84	5	0
割合[%]	94.4	5.6	0.0

モデル系列A、Bどちらの場合も多くケースにおいて正確なアラインメントが実現できた。対応づけの割合が95%に届かなかったケースの多くは入力系列に10フレーム程度の連続した検出失敗フレームが含まれていることを確認した。これは、フレーム間のベクトルを用いた正確なコスト計算を連続して行えないことが、不正確なマッチングの大きな要因となったと推察した。

5. 結論

本研究では、カメラで撮影された2つの投球映像から投球動作の照合を行う手法を提案した。多くの場合で正確な対応づけを実現できたが、身体座標の検出が失敗した場合のコスト計算方法には改善の余地がある。今後はそのような状況での対応付け精度の改善を図る。

文 献

- [1] Z. Cao, T. Simon, S. Wei, Y. Sheikh: "Realtime Multi-Person 2D Pose Estimation using Part Affinity Field," In CVPR, pp7291-7299 (Jun.2017)
- [2] 横井真也, 石川孝明, 渡辺裕: "スポーツ映像から取得した骨格座標データに対するアラインメント", 電子情報通信学会総大会, D-12-59 (Mar. 2019)

† 早稲田大学 基幹理工学部 情報通信学科 渡辺研究室
〒169-0072 東京都新宿区大久保 3-14-9

早大シルマンホール 401

TEL.03-5286-2509 E-mail: r-osawa@fuji.waseda.jp