

卒業論文概要書

Summary of Bachelor's Thesis

Date of submission: 2/7/2020

専攻名(専門分野) Department	情報理工	氏名 Name	大澤遼平	指導 教員 Advisor	渡辺 裕 ㊞
研究指導名 Research guidance	オーディオビジュアル 情報処理研究	学籍番号 Student ID number	CD 1W162075-8		
研究題目 Title	OpenPose と Dynamic Time Warping を用いた投球動作照合 Pitching Action Matching Using OpenPose and Dynamic Time Warping				

1. まえがき

近年、野球において選手の動作を分析するために試合中や練習中の選手を撮影した映像データが利用されている。映像は繰り返し視聴できるため、選手は一つの映像から動作を細部まで分析することが可能である。

自身の動きを客観的に確認することで、実際の動作と選手自身が理想とする動作の違いを発見できる。また、映像を用いて対戦相手の動作を様々な観点から解析することにより、相手選手の弱点や癖を探し出すことが容易になる。

さらに、複数の映像内の同一動作を比較することで、身体の使い方の相違部分を発見できる。野球の技術が高い選手の動作と自身の動作を比較することで得られる動きの違いから、野球技術の向上のために選手自身が取り組むべき課題や重点的にトレーニングすべき箇所を明確にすることが可能となる。練習時と試合時、試合序盤と試合終盤のように異なる状況での同一選手の動作を比較することで、緊張感や疲労などが動作に与える影響を把握できる。しかし、動作の速度や身体の使い方は選手や状況によって異なるため、複数の動作の細部のタイミングを合わせることは難しい。

本研究では、野球の投球映像から選手の身体座標を取得し、得られた身体座標データに対して Dynamic Time Warping (DTW)を用いることにより、二つの投球動作のタイミングを正確に照合する手法について検討する。

2. 関連研究

2.1. OpenPose

映像内の人物の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 は人間の身体にセンサーなどを装着することなく、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 はモデルごとに検出可能な部位が異なり、15, 18 もしくは 25 点の位置を推定できる。映像内で顔や手が鮮明に映っている場合には、顔の輪郭や指の関節など、より細かい部位を検出することも可能である。この際に、各部位の 2 次元座標値が取得され推定の信頼

度が算出される。

2.2. DTW

Dynamic Time Warping(DTW)は二つの系列の類似度が最も高くなるように系列内の要素を対応付ける手法である。系列内の各要素に対し総当たりで要素間の類似度を求め、類似度の合計が最小になるように対応付けを決定する。DTW は二つの系列の要素数が異なる場合でも適用可能であり、時系列データなどに対して多く用いられる。

DTW は 2 系列のすべての要素を対応付ける。そのため、系列内の各要素はもう一方の系列に類似する要素が存在しない場合でも少なくとも一つの要素と対応付けされる。類似しない要素同士の対応付けにより全体の対応付け精度が低下する可能性がある。よって、DTW を用いて正確に動作を照合するためには、映像内の動作の始点および終点を主観で付与する必要がある。そこで、要素間の類似度に基づき、系列の中から動作の始点および終点を自動で抽出する始点終点自由な DTW[2]が提案されている。

2.3. 投球動作照合

映像から取得した身体座標データに基づく右投手の投球動作照合手法として、OpenPose と始点終点自由な DTW を用いた手法[2]が存在する。OpenPose により映像から推定した投手の身体座標データから投手の姿勢の類似度を算出する。DTW を適用し、姿勢の類似度に基づき二つの映像内の投球動作を対応付ける。しかし、OpenPose による身体座標の誤検出が連続している場合には正しく投球動作を対応付けることが困難であるという問題点が存在する。

3. 提案手法

3.1. フレーム数削減

本手法では、文献[2]で提案されている始点終点自由な DTW を用いる。映像フレーム間の投手の姿勢類似度を算出し、投球動作を対応付ける。DTW は総当たりで要素間の類似度を算出するため、映像のフレーム数に比例して類似度算出の計算量が増加する。そこで、計算量を抑えるために、対応付けの精度に影響を与えな

い範囲で DTW の計算に使用するフレーム数を減らす。
 今回使用した投球動画には、時間的に連続する複数のフレームで投手の姿勢がほとんど変化しないケースが含まれることを確認した。図 1 に例を示す。

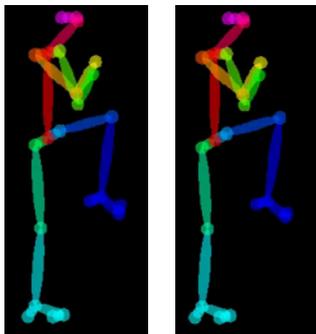


図 1 OpenPose による投手の姿勢推定結果

図 1 に示した 2 枚の画像は、同一映像内の時間的に連続する 2 フレームにそれぞれ映る投手に対し OpenPose を適用して得られた姿勢推定結果である。2 フレームの投球姿勢は酷似しており、DTW によって対応付けられる別の系列内のフレームは同一のものであると予想される。そのため、DTW の計算では、時間的に連続で投球姿勢が似ている複数フレームのうち 1 フレームのみを使用すれば良い。

3.2. ベクトル作成

本研究では野球中継映像から OpenPose により得られる身体座標データのうち、首、両肩、両肘、腰の中央、両腰、両膝、両足首の 12 点を用いる。対象とする 12 点のうち、任意の二つの部位を結ぶベクトルの長さや角度を OpenPose によって得られる各部位の 2 次元座標から計算する。また、使用した 2 部位の信頼度のうち、最小値をベクトルの信頼度とした。

3.3. DTW による投球動作の対応付け

本研究は二つの映像のフレームを対応付けるものである。対応付けの際には始点終点自由な DTW を用いる。本手法では、異なる二つの部位を結ぶ 66 次元のベクトルを用いてフレーム間の姿勢の類似度を算出し、類似度の平均が最小になるように対応付けを決定する。姿勢の類似度は対応するベクトルの差から算出される。ここで、対応するベクトルの差は二つのベクトルの角度差と長さの差の積とする。

姿勢の類似度は OpenPose により得られる各部位の 2 次元座標値から算出されるため、OpenPose の誤検出や検出失敗は不正確な対応付けの要因となる。そこで、二つのベクトルのうち少なくとも片方のベクトルの信頼度が閾値未満である場合は、角度差および長さの差を計算せず、定数値を対応するベクトルの差とする。これにより、OpenPose による誤検出や検出失敗が対応付け精度に与える影響を少なくする。

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投球映像 100 本である。このうち、1 本をモデル系列とし、入力系列はモデル系列を含む全 100 本とする。各映像には 1 球分の投球動作が含まれており、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、終点は蹴り上げた軸足が地面に着く直前とした。入力系列は投球動作の始点と終点が未知の状態を与えた。得られたモデルフレームと入力フレームの対応付けについて成否を目視により判定した、より対応付けに適したフレームが存在する場合は誤った対応付け、存在しない場合は正しい対応付けとした。

4.2. 実験結果

従来手法[2]と提案手法による投球動作の対応付け結果を表 1 および表 2 に示す。

表 1 従来手法による対応付け結果

	正しい対応付け	誤った対応付け
対応付け数	8938	546
割合 [%]	94.24	5.76

表 2 提案手法による対応付け結果

	正しい対応付け	誤った対応付け
対応付け数	10063	523
割合 [%]	95.06	4.94

ベクトルの信頼度が低い場合に姿勢類似度の算出方法を変更することで、従来手法と比較し対応付け精度が向上した。

5. むすび

本研究では、映像から取得した投手の骨格情報を用いて二つの投球動作を照合する手法を提案した。従来手法よりも対応付け精度が 0.82 ポイント向上し、提案手法の有効性を確認した。

参考文献

- [1] Z. Cao, T. Simon, S. Wei, and Y. Sheikh: "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), No.121, pp.1302-1310, July 2017
- [2] 横井真也: "スポーツ映像から取得した身体座標データに対するアラインメント", 早稲田大学大学院基幹理工学研究科情報理工・情報通信専攻修士論文, Feb. 2019

2019 年度 卒業論文

OpenPose と Dynamic Time Warping を用いた投球動作照合

Pitching Action Matching

Using OpenPose and Dynamic Time Warping

指導教員 渡辺 裕 教授

早稲田大学 基幹理工学部

情報理工学科

1W162075-8

大澤 遼平

目次

第1章 序論	1
1.1 研究背景	1
1.2 研究目的	2
1.3 本論文の構成	2
第2章 関連研究	3
2.1 まえがき	3
2.2 OpenPose	3
2.2.1 概要	3
2.2.2 利点	4
2.2.3 問題点	4
2.3 Dynamic Time Warping	5
2.3.1 概要	5
2.3.2 問題点	6
2.3.3 Prefix and Suffix Invariant Dynamic Time Warping	7
2.4 投球動作照合	8
2.5 むすび	9
第3章 提案手法	10
3.1 まえがき	10
3.2 提案手法概要	10
3.3 OpenPose による姿勢情報の取得	10
3.4 フレーム数削減	11
3.4.1 概要	11
3.4.2 OpenPose による検出人数に基づく削減	11
3.4.3 投手の姿勢変化に基づく削減	12
3.5 OpenPose による誤検出部位の座標補正	13
3.5.1 概要	13
3.5.2 足の左右反転検出の修正	13
3.5.3 片足誤検出の修正	17
3.6 ベクトルデータの作成	18
3.7 姿勢類似度の算出	18
3.8 Dynamic Time Warping による投球動作の対応付け	19

3.9	むすび.....	21
第4章	実験結果および考察.....	22
4.1	まえがき.....	22
4.2	実験の概要.....	22
4.3	フレーム数削減の評価.....	23
4.3.1	概要.....	23
4.3.2	精度評価.....	23
4.3.3	実行時間評価.....	24
4.4	誤検出部位の座標補正評価.....	25
4.5	対応付け結果.....	27
4.5.1	従来手法.....	27
4.5.2	提案手法.....	30
4.6	むすび.....	36
第5章	結論と今後の課題.....	37
5.1	結論.....	37
5.2	今後の課題.....	37
	謝辞.....	38
	参考文献.....	39
	図一覧.....	40
	表一覧.....	41
	研究業績.....	42

第1章 序論

1.1 研究背景

近年、野球選手の動作を解析するために試合時や練習時の選手を撮影した映像データが利用されている。映像は繰り返し確認できるため、動作の細部まで分析することが可能である。また、スマートフォンやタブレットなどの普及により、時間や場所によらず映像を視聴することが容易である。そのため、選手は自宅などでも映像データを活用し動作を解析できる。

日本のプロ野球球団である福岡ソフトバンクホークスは、プロ野球界でいち早く映像データの活用を開始した[1]。2009年にiPhoneを、2011年にiPadを所属選手全員に配布し、選手ごとの映像や他球団の選手の映像を常に閲覧できる環境を整備した。文献[1]によると多くの選手が遠征の移動中や練習後のロッカールームなどで動画を視聴し、動作分析に活用している。2000年から2009年までの10年間で福岡ソフトバンクホークスは1度だけ日本一に輝いている。一方で、映像データの活用が進んだ2010年から2019年までの10年間では6度の日本一を達成している。

映像利用による動作解析の目的は大きく三つに分類される。一つ目は、自身の動きを分析することである。実際の動作を客観的に確認することで、選手自身が理想とする動きとの相違部分を発見できる。また、指導者に映像をチェックしてもらい、肉眼では認識し難い細かい修正点や改善すべき点について助言を受けることも可能となる。

二つ目は、対戦相手の研究である。映像を用いて対戦相手の動作を様々な観点から解析することにより、相手選手の弱点や癖を探し出すことが容易になる。発見した動作の特徴は、相手を攻略するための戦略を練る際に役立つ。さらに、試合前に相手選手の動きを繰り返し確認することで、対戦時を想定したイメージトレーニングが可能となる。三上ら[2]はVRを用いたイメージトレーニングシステムの研究開発を進めている。

三つ目は、動作の比較である。複数映像内の同一動作を見比べることで、身体の使い方の相違部分を発見できる。野球の技術が高い選手の動作と自身の動作を比べることで得られる動きの違いから、野球技術の向上のために選手自身が取り組むべき課題や重点的にトレーニングすべき箇所を明確にすることが可能となる。また、練習時と試合時、試合序盤と試合終盤のように異なる状況での同一選手の動作を比較することで、緊張感や疲労などが動作に与える影響を把握できる。

映像を活用し動作を解析することで、肉眼では発見し難い特徴や癖を明らかにすることが可能になる。解析対象の選手に身体データ取得用の装置などを取り付ける必要がないため、選手の身体に負担をかけることなく、動作を撮影・分析できる。また、映像はスマートフォンやタブレットなどで手軽に撮影できるため、高価なカメラは必要ない。そのため、映像利用は、手軽にかつ低コストでの動作解析を可能にする。

1.2 研究目的

映像を利用して動作を比較する場合、複数の映像を見比べる必要がある。この際、動きの細部まで比べるためには複数動作のタイミングを合わせる事が重要となる。例として、投球動作を比較する場合、投球開始時や終了時、ボールがリリースされる瞬間、軸足を蹴り上げる瞬間などのタイミングをすべて対応付ける必要がある。しかし、身体の使い方や動作の速度は選手や状況によって異なるため、複数動作の細部のタイミングを正確に合わせることは難しい。

そこで、本研究では二つの投球動作のタイミングを正確に照合する手法について検討する。野球の投球映像に対して **OpenPose** を適用し、選手の身体座標を取得する。得られた身体座標データに対して **Dynamic Time Warping (DTW)** を用いることで、動作のタイミングを対応付ける。

1.3 本論文の構成

以下に本論文の構成を示す。

第1章は本章であり、本研究の背景および目的を述べている。

第2章では、本研究で用いる **OpenPose**, **Dynamic Time Warping** について説明する。また、関連する投球動作照合手法を示す。

第3章では、二つの投球動作を対応付ける手法を提案する。

第4章では、提案手法における実験および結果、考察を示す。

第5章では、本研究のまとめと今後の課題を述べる。

第2章 関連研究

2.1 まえがき

本章では, 本研究で利用する **OpenPose[3]** と **Dynamic Time Warping** について述べる. また, 投球動作照合に関する従来手法を示す.

2.2 OpenPose

2.2.1 概要

画像中に映る人物の身体部位の位置を推定する手法として Cao らが提案した **OpenPose[3]** がある. **OpenPose** は画像内の複数人の身体部位を検出できる手法である. また, 映像に対しても適用可能であり, 各フレームで独立に関節などの特徴点を検出する. **OpenPose** は使用するモデルごとに検出可能な特徴点の数が異なり, 1 人につき 15, 18, 25 部位の位置を推定できる. 画像内に手や顔が鮮明にかつ十分な大きさに映っている場合には, 指の関節や眉毛, 顔の輪郭など, より細かい部位の検出も可能である. 各特徴点について, 画像上での 2 次元座標値と推定の信頼度が出力される. 信頼度は 0 から 1 の範囲の値であり, 1 に近いほど推定の信頼度が高いことを示す. 特徴点の検出に失敗した場合は, 2 次元座標値および信頼度が 0 となる.

OpenPose の実行時に出力ファイル名などを指定することにより, 検出結果の座標データを JSON 形式や XML 形式などのファイルに出力することができる. 動画に対して適用した場合には, フレームごとに異なるファイルに座標データが出力される. また, 検出された特徴点および特徴点同士を結ぶ骨格を入力画像上に描画した出力結果を取得できる. 入力画像上ではなく, 元画像と同サイズの黒画像上に特徴点と骨格のみを表示した出力結果を得ることも可能である. 25 部位の位置を推定可能なモデルを使用した際の, 出力結果の例を図 2.1 に示す.

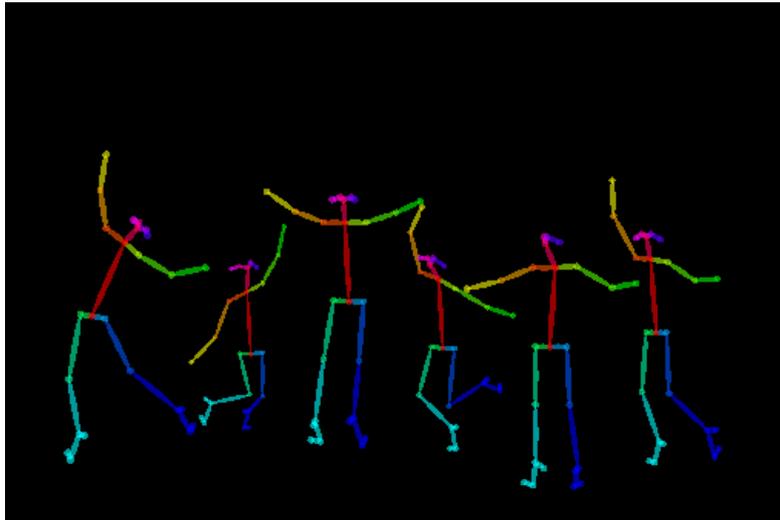


図 2.1 OpenPose による検出結果の例

2.2.2 利点

OpenPose は機械学習を用いて、画像や動画のみから複数人物の身体部位を検出する姿勢推定ライブラリである。関節などの特徴点の位置を画像ごとに推定できるため、スポーツ動作の解析や行動識別のためのツールとして利用することが可能である。

OpenPose の利点として、マーカなどの装置を身体に取り付けることなく、身体部位の位置を推定できることがある。スポーツ選手の試合中の動作をカメラで撮影するだけで、動作の解析が可能になる。また、特殊なセンサなどを使用せず、一台のカメラで撮影した画像や映像のみから身体部位を検出可能であることも OpenPose の利点である。高価な解析器具や高性能な撮影機材を必要としないため、テレビなどで配信されている映像やスマートフォンで撮影した画像に対しても人物の姿勢推定が可能である。

2.2.3 問題点

OpenPose は 1 枚の画像のみから人物の身体部位を検出する。そのため、人物の身体に他の物体が重なり、一部の特徴点が画像上に映らない場合には、検出の精度が低下する。また、画像内で人物同士の距離が近いと、ある人物の身体部位を別の人物の身体部位として誤検出する場合がある。

映像に対して OpenPose を適用する場合は、フレームごとに身体部位の位置を推定する。この際、OpenPose は時系列の前後フレームでの検出結果は利用せずに、各フレームで独立に特徴点を検出する。時系列の前後情報を一切考慮しないため、不正確な検出が生じる可能性がある。図 2.2 に不正確な検出の例を示す。

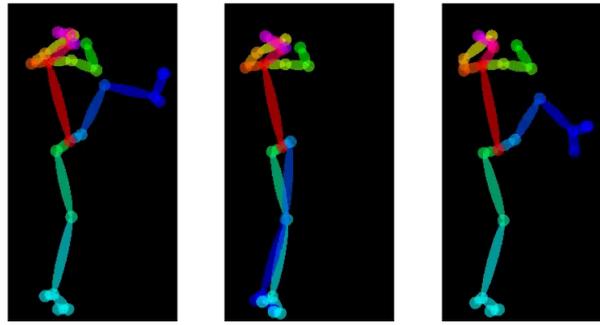


図 2.2 OpenPose による不正確な検出の例

図 2.2 に示した 3 枚の画像は、同一映像内の時間的に連続する 3 フレームに対して OpenPose を適用した結果である。なお、元映像は 30fps の動画であり、フレーム間は約 0.03 秒である。青色の線が OpenPose により左足として検出された骨格を示す。図 2.2 の中央の画像では、青色の骨格の位置が左右の 2 枚の画像とは大きく異なる。元映像と OpenPose による検出結果を比較すると、図 2.2 の左右の 2 枚の画像では正しく左足の部位が検出されていることが確認できた。一方で、中央の画像では、実際には左足が高く上がっているにもかかわらず、左足の位置を誤って推定していることが確認された。図 2.2 のように 3 フレーム、すなわち 0.06 秒の間に左足が一度大きく下がり、再び高い位置に戻ることは人間の動作として現実的ではない。しかし、OpenPose では各フレームで独立に身体部位の位置を推定するため、図 2.2 に示したような誤検出が生じる場合がある。

2.3 Dynamic Time Warping

2.3.1 概要

Dynamic Time Warping (DTW) は二つの系列の類似度が最も高くなるように系列内の要素を対応付ける手法である。系列内の各要素に対して総当たりで要素間の距離を求め、距離の合計が最小になるように対応付けを決定する。要素間の距離が小さいほど、二つの要素が類似していることを表す。DTW では、一方の系列の一つの要素をもう一方の系列の複数の要素に対応付けることが可能である。そのため、二つの系列の長さが異なる場合にも適用可能であり、時系列データなどに対して多く用いられる。要素数が異なる二つの系列に対して DTW を適用した例を図 2.3 に示す。

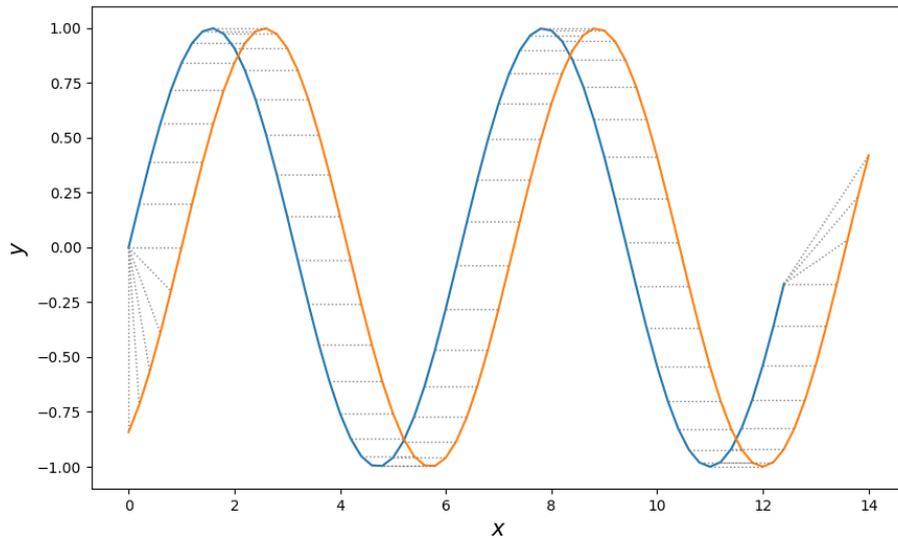


図 2.3 要素数の異なる 2 系列に対する DTW の適用例

図 2.3 中の二つの系列は同じ変化をしているが、青で示された系列は 62 点、オレンジで示された系列は 70 点あり、要素数が異なる。青の系列の点とオレンジの系列の点の y 座標差を総当たりで算出し、 y 座標差の合計が最小になるように対応付けを決定している。対応付けされた 2 点を点線により結んでいるが、 y 座標差が 0 である 2 点を結ぶ横方向の対応付けが多く存在している。また、青の系列の始点および終点がオレンジの系列の複数の要素と対応付けされており、要素数が異なる 2 系列に対しても DTW が適用可能であることが確認できる。

DTW は系列間の類似度を算出可能なため、動作分類や個人識別などに利用されている。DTW を用いた研究の例として、人間の行動認識[4]や歩容データに基づく個人識別[5]が存在する。

2.3.2 問題点

DTW は二つの系列の長さが異なる場合でも、すべての要素を対応付ける。すなわち、一方の系列の各要素は、もう一方の系列内の少なくとも一つの要素と対応付けされる。よって、一方の系列にのみ存在する要素は、もう一方の系列内の対応付けされるべきではない要素と対応付けされる。この際、類似する要素同士の間と比べ、要素間の距離は大きくなる。DTW は距離の合計が最小となるように対応付けを決定する。そのため、一部の対応付けにおいて、他の対応付けと比べて距離が著しく大きい場合、対応付けの精度が低下する可能性がある。

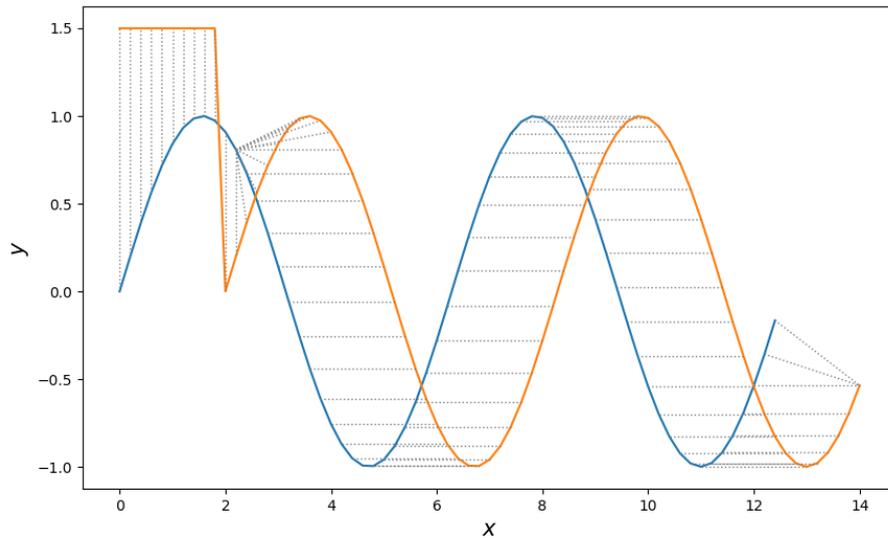


図 2.4 形状の似ている 2 系列に対する DTW の適用例

図 2.4 は形状の似ている二つの系列に対して DTW を適用した例である。ただし、図 2.4 中の青で示された系列の始点付近には存在しない要素が、オレンジで示された系列には含まれている。青で示された系列の点とオレンジで示された系列の点の y 座標差を総当たりで算出し、 y 座標差の合計が最小になるように対応付けを決定している。対応付けされた 2 点を点線により結んでいるが、グラフの左側において y 座標差が大きい 2 点を結ぶ縦方向の対応付けが多く確認できる。青の系列の始点付近の要素は、オレンジの系列内に存在する、 y 座標差が 0 となる要素とは対応付けされておらず、不正確な対応付けが生じている。

また、DTW には、計算量が多いという問題点が存在する。DTW では、二つの系列に含まれる各要素に対して総当たりで要素間の距離を算出する。したがって、二つの系列の長さに比例して計算量が増加する。

2.3.3 Prefix and Suffix Invariant Dynamic Time Warping

Prefix and Suffix Invariant Dynamic Time Warping (PSI-DTW)[6]は、対応付けの始点および終点を固定しない DTW である。ここで、対応付けの始点および終点とは、各系列内の対応付けされた要素のうち、系列自身の始点に最も時間の近い要素および系列自身の終点に最も時間の近い要素を示す。第 2.3.2 項で述べたように、DTW は二つの系列のすべての要素を対応付ける。したがって、DTW では、対応付けの始点および終点は系列自身の始点および終点と常に一致する。一方、PSI-DTW では、対応付けの始点および終点を系列自身の始点および終点に限定せず、特定の範囲から最適な点を選択する。系列の始点から 0 以上 r 以下の時間が離れている要素を対応付けの始点の候補、系列の終点から 0 以上 r 以下の時間が離

れている要素を対応付けの終点の候補として、系列間の類似度が最も高くなるように対応付けを決定する。ここで、 r は二つの系列の長さよりも小さな自然数であり、対応付けの始点および終点の探索範囲を示す。

PSI-DTW では、対応付けの始点よりも前の時間の要素および対応付けの終点よりも後ろの時間の要素はどの要素にも対応付けされない。そのため、系列の始点付近もしくは終点付近に存在する、かつ一方の系列にのみ含まれる要素による不正確な対応付けは減少する。

PSI-DTW の欠点として、要素間の距離の合計が最小になるように対応付けを決定することが挙げられている[7]。PSI-DTW では、二つの系列の要素をすべて対応付ける必要はないため、どちらかの系列の終点候補に含まれる要素が一つ対応付けされた時点で、終了条件を満たす。この際、新たな対応付けを追加すると、距離の合計値は増加する。例外として、新たな対応付けにおいて要素間の距離が 0 の場合には増加しないが、距離の合計値が減少することはあり得ない。そのため、要素間の距離の合計が最小になるように対応付けを決定すると、終点付近に含まれる本来対応付けされるべき要素が対応付けされない可能性がある。

2.4 投球動作照合

映像から取得した身体座標データに基づく右投手の投球動作照合手法として、OpenPose と DTW を用いた手法が存在する[7]。OpenPose により映像から推定した投手の身体座標データから投手の姿勢の類似度を算出する。DTW を適用し、姿勢の類似度に基づき二つの映像内の投球動作を対応付ける。従来手法[7]では、PSI-DTW を改良した DTW を使用している。PSI-DTW は、要素間の距離の合計値が最小になるように対応付けを決定する手法のため、系列の終点付近において本来対応付けされるべき要素が対応付けされない場合がある。そこで、従来手法では要素間の距離の合計値ではなく、要素間の距離の平均値が最小となるように対応付けを決定する手法を提案している。

第 2.3.2 項で述べたように、一方の系列にのみ含まれる要素が存在する場合、DTW による不正確な対応付けが生じる可能性がある。そこで、従来手法では、DTW を実行する前に対象外動作と個人依存動作を除去する。ここで、対象外動作とは投球動作の前後に存在する要素、個人依存動作とは投手や状況により変化する、動作内の個人に依存する動作と定義されている。対象外動作は投球動作ではないため、対応付ける必要がない。また、常に類似する動作であるとは限らないため、対応付けの精度が低下することが考えられる。個人依存動作は一方の系列にのみ存在する要素であるため、もう一方の系列に類似する要素が存在せず、不正確な対応付けの要因となる。そこで、従来手法は、対象外動作と個人依存動作を除去した系列に対して、DTW を適用している。

しかし、個人依存動作を除去することによる問題点も存在する。一方の系列内の除去された個人依存動作は、もう一方の系列内のいずれの要素とも対応付けされない。そのため、対応付けされた投球動作を選手が比較する際に、一部の動作が対応付けされていないことにより違和感を感じる。また、一部の動作が欠けていることにより、選手は正確に投球動作を

比較できない。

さらに、OpenPose による身体座標の誤検出が連続している場合には正しく投球動作を対応付けることが困難であるという課題が存在する。

2.5 むすび

本章では、映像内の人物の姿勢推定手法である OpenPose と、二つの系列の要素を対応付ける手法である Dynamic Time Warping について述べた。また、投球動作照合に関する従来手法を示した。

第3章 提案手法

3.1 まえがき

本章では、映像から取得した投手の姿勢情報に基づく投球動作照合の提案手法について述べる。

3.2 提案手法概要

本手法では、まず野球の投球映像に対して **OpenPose** を適用し、投手の身体座標データを取得する。得られた身体座標データに基づいてフレーム間の投手の姿勢類似度を算出し、DTW を用いて投球動作を対応付ける。

DTW では二つの映像に対して総当たりでフレーム間の姿勢類似度を計算する必要があるため、映像のフレーム数に比例して類似度算出の計算量が多くなる。そこで、**OpenPose** の検出結果に基づき、対応付けの精度に影響を与えない範囲で DTW の計算に使用するフレーム数を減らす。

また、本手法は身体座標データに基づき投球動作を照合するため、**OpenPose** により誤って推定された座標値は対応付けの精度に影響を与える。そこで、正しく推定されていない身体部位の一部について、前後フレームでの **OpenPose** による検出結果を用いて座標値を補正する。また、**OpenPose** より取得される各身体部位の信頼度に応じて、姿勢類似度を算出する際の処理を変更する。

3.3 **OpenPose** による姿勢情報の取得

第 2.2 節で述べた **OpenPose** を用いて、投球映像から人物の身体座標データを取得する。**OpenPose** は使用するモデルごとに検出可能な特徴点の数が異なるが、本研究では、25 部位検出可能なモデルを使用した。各フレームについて、検出された人物ごとに 25 部位それぞれの 2 次元座標値および推定の信頼度が出力される。

本研究で用いた投球映像に対して、**OpenPose** の 25 部位検出可能なモデルを適用した結果の例を図 3.1 に示す。

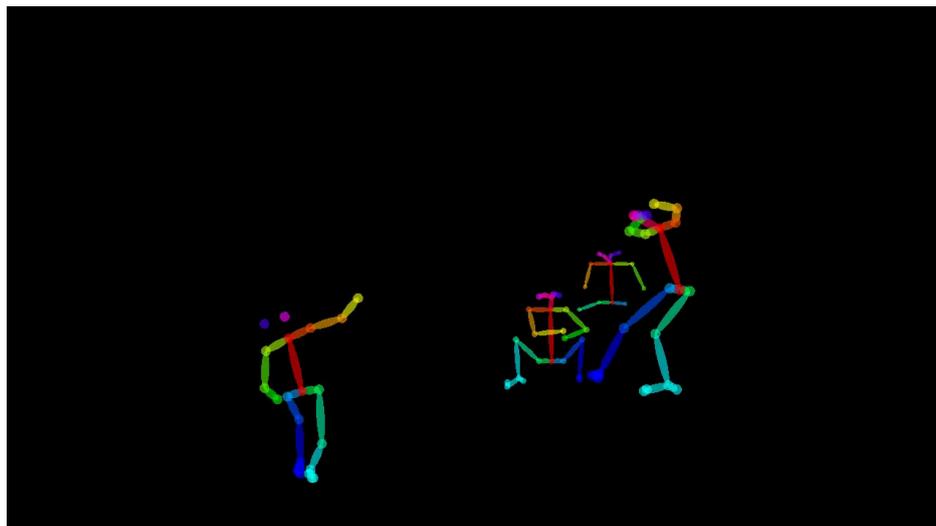


図 3.1 投球映像に対する OpenPose の適用例

図 3.1 のように，投球映像内には投手以外の人物が存在する．本研究では，投手の姿勢情報に基づき投球動作のタイミングを照合する．そのため，OpenPose により検出された人物の中から投手を特定する必要がある．本研究で使用した動画の投球シーンでは，左手前に投手，右奥に捕手や打者，審判などが映る．図 3.1 に示した画像中において，最も左の骨格が投手を表す．各骨格の足首の位置に着目すると，投手の足首が画像上の最も下にあることが確認できる．よって，フレームごとに，検出されたすべての人物の足首の座標値を比べ，足首の位置が最も下である人物を投手とする．

3.4 フレーム数削減

3.4.1 概要

本手法では，DTW を用いて映像フレーム間の投手の姿勢の類似度を算出し，投球動作を対応付ける．DTW は総当たりで要素間の類似度を計算するため，系列の長さに比例して計算量が増加する．そこで，類似度算出の計算量を抑えるために，対応付けの精度に影響を与えない範囲で DTW の計算に使用するフレームを削減する．具体的には，OpenPose による検出人数に基づく削減と投手の姿勢変化に基づく削減である．

3.4.2 OpenPose による検出人数に基づく削減

今回使用した動画の投球シーンでは，投手の他に捕手や打者，審判が映る．よって，画面内に少なくとも 4 人の人物が存在する．そこで，OpenPose により検出された人数に基づき，フレーム数を削減する．ただし，投球シーンにおいて常に 4 人以上が検出されるとは限らない．審判が捕手や打者と重なり，審判の身体の一部が画像上に映らない場合に，OpenPose によって審判が検出されない可能性があることを確認した．投球シーンにおいて審判が検出されているフレームを図 3.2 に，検出されていないフレームを図 3.3 にそれぞれ示す．

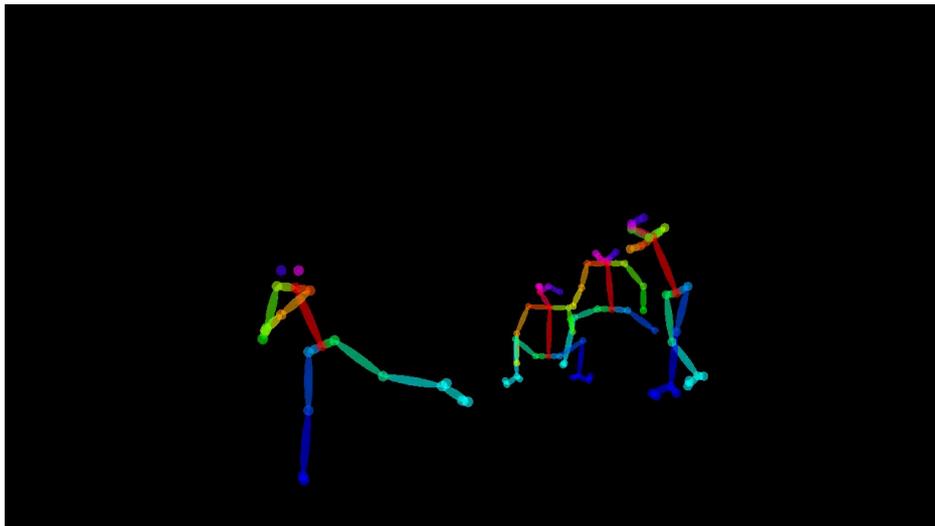


図 3.2 審判が検出されているフレーム

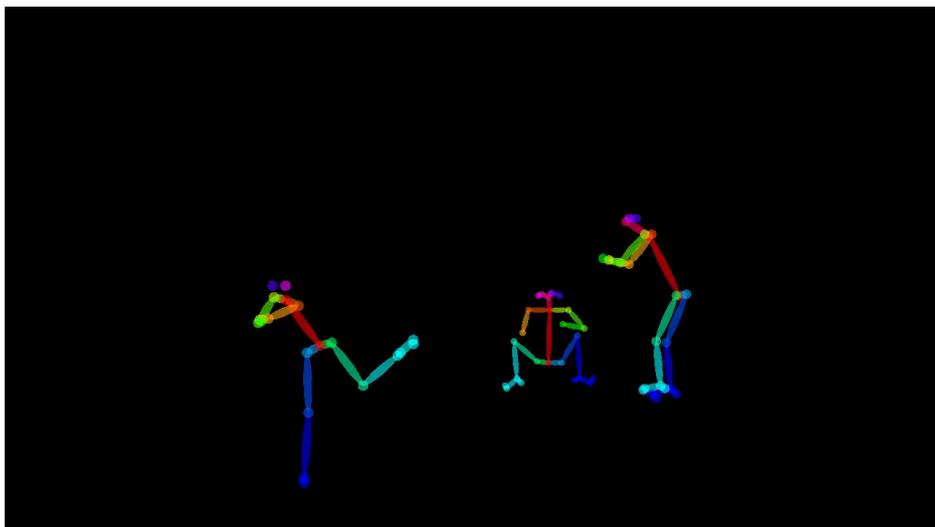


図 3.3 審判が検出されていないフレーム

図 3.2 では 4 人の人物が検出されているが、図 3.3 では 3 人しか検出されていないことが確認できる。本研究で使用したデータセットの投球シーンにおいて、投手、捕手、打者、審判の 4 人のうち 2 人以上検出されないフレームは存在しなかった。そこで、OpenPose により検出された人物が 3 人未満のフレームは投球シーンではないと判断し、DTW の計算に使用しない。

3.4.3 投手の姿勢変化に基づく削減

本研究で使用した投球映像において、時間的に連続する複数フレームで投手の姿勢がほとんど変化しないケースが確認された。図 3.4 に例を示す。

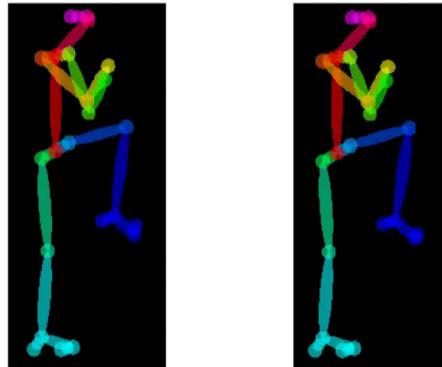


図 3.4 投手の姿勢がほとんど変化しない連続フレーム

図 3.4 に示した 2 枚の画像は、同一映像内の時間的に連続する 2 フレームにそれぞれ映る投手に対する OpenPose による骨格推定結果である。左右の画像を比較すると、投手の姿勢が酷似していることがわかる。本研究では姿勢の類似度に基づき、投球動作のタイミングを合わせる。そのため、図 3.4 の 2 フレームのように、ある系列に含まれる投手の姿勢が類似する複数フレームが DTW によって対応付けされる別の系列内のフレームは同一であると予想される。時間的に連続であり、投手の姿勢が似ている複数フレームのうち、1 フレームの対応付けが DTW により決定されれば、残りのフレームは DTW の類似度計算を行わずに対応付けることが可能となる。よって、同一系列内の時間的に連続する複数フレームにおいて、投手の姿勢がほとんど変化しない場合は、姿勢が似ている複数フレームのうち 1 フレームのみを DTW の計算に使用する。

3.5 OpenPose による誤検出部位の座標補正

3.5.1 概要

本研究では第 2.2 節で述べた OpenPose を用いて、映像から投手の身体座標データを取得する。ただし、オクルージョンなどの影響により、身体部位の位置が誤って推定される場合がある。身体座標データに基づき投球動作を照合するため、誤って推定された座標値は対応付けの精度に影響を与える。そこで、本手法では、正しく推定されていない身体部位の一部について、前後フレームでの検出結果を用いて座標値を補正する。具体的には足が左右逆に検出されている場合と片足のみ誤って検出されている場合の二つについて座標値を修正する。

3.5.2 足の左右反転検出の修正

本研究で使用した投球動画において、投手の足が左右逆に検出されているフレームが存在することを確認した。例を図 3.5 に示す。

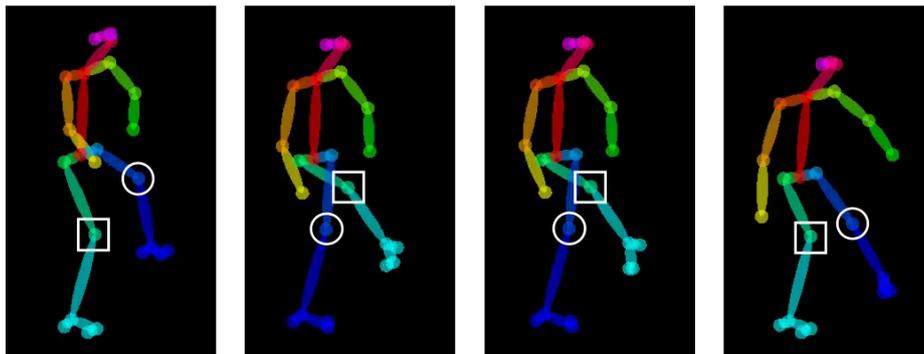


図 3.5 足の左右反転検出例

図 3.5 に示した 4 枚の画像は、同一動画内の時間的に連続する 4 フレームにそれぞれ映る投手に対して OpenPose を適用した結果である。水色の線が OpenPose により右足として検出された骨格、青色の線が OpenPose により左足として検出された骨格をそれぞれ示す。白い四角および白い丸により囲まれている点は、OpenPose によって推定された右膝および左膝の位置である。元映像を確認すると、中央の 2 枚の画像では投手の右膝や右足首が OpenPose により左膝、左足首として、左膝や左足首が右膝、右足首として誤って検出されていた。一方、左右の 2 枚の画像では足の各部位の位置が正しく推定されていた。

左の 2 枚の画像において、右膝同士を比較すると、画像上における位置が大きく変化していることがわかる。また、最も左の画像において白い四角で囲まれた右膝の位置と、左から 2 枚目の画像において白い丸で囲まれた左膝の位置がほとんど同じであることが確認できる。このように、時間的に連続する 2 フレームのうち、どちらか片方のフレームにおいて OpenPose による足の検出が左右逆である場合は、2 枚の画像上において右膝同士の位置よりも右膝と左膝の位置の方が近くなる。

OpenPose では、検出した各部位について画像上の位置を示す 2 次元座標値が出力される。そこで、フレーム間の同一部位の座標差 D_s および、左右が逆である部位の座標差 D_d を算出し、その値に基づき足の検出が左右反転しているかを判定する。 $k-1$ フレームと k フレーム間での同一部位の座標差 $D_{s,k}$ および左右が逆である部位の座標差 $D_{d,k}$ の算出方法を式(3.1)、式(3.2)にそれぞれ示す。なお、フレーム数 M の系列において、 k は 2 以上 M 以下の整数である。

$$D_{s,k} = \sum_{i=1}^6 (d(LP_{k-1,i}, LP_{k,i})) \quad (3.1)$$

$$D_{d,k} = \sum_{i=1}^3 (d(LP_{k-1,i}, LP_{k,i+3})) + \sum_{i=4}^6 (d(LP_{k-1,i}, LP_{k,i-3})) \quad (3.2)$$

ここで、 LP_k は OpenPose によって推定された、 k フレームでの右腰、右膝、右足首、左腰、左膝、左足首の計 6 点の 2 次元座標値の集合である。 $LP_{k,i}$ は 6 点の部位のうち 1 部位の 2 次元座標値であり、 i の値と各部位が対応している。 i の値と部位の対応関係を表 3.1 に示す。

表 3.1 i の値と部位の対応関係

i の値	部位
1	右腰
2	右膝
3	右足首
4	左腰
5	左膝
6	左足首

式(3.1)、式(3.2)において、 d は 2 つの 2 次元座標値の座標差を算出する関数である。同一部位の座標差 $D_{s,k}$ および左右が逆である部位の座標差 $D_{d,k}$ は、表 3.1 に示す 6 点の部位の $k-1$ フレームと k フレームでの 2 次元座標値を用いて算出される。 $k-1$ フレームにおける 6 点の部位と k フレームにおける 6 点の部位を 1 対 1 で対応付け、対応付けされた 2 つの部位の座標差を計算する。 $D_{s,k}$ を算出する際の、1 対 1 の対応付けを表 3.2 および図 3.6 に、 $D_{d,k}$ を算出する際の、1 対 1 の対応付けを表 3.3 および図 3.7 に示す。

表 3.2 $D_{s,k}$ を算出する際の 1 対 1 の対応付け

$k-1$ フレーム	k フレーム
右腰	右腰
右膝	右膝
右足首	右足首
左腰	左腰
左膝	左膝
左足首	左足首

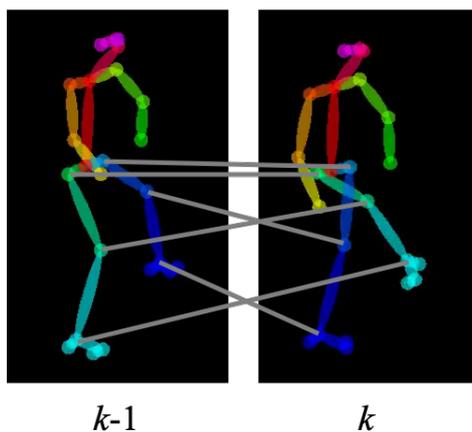


図 3.6 $D_{s,k}$ を算出する際の1対1の対応付け

表 3.3 $D_{d,k}$ を算出する際の1対1の対応付け

$k-1$ フレーム	k フレーム
右腰	左腰
右膝	左膝
右足首	左足首
左腰	右腰
左膝	右膝
左足首	右足首

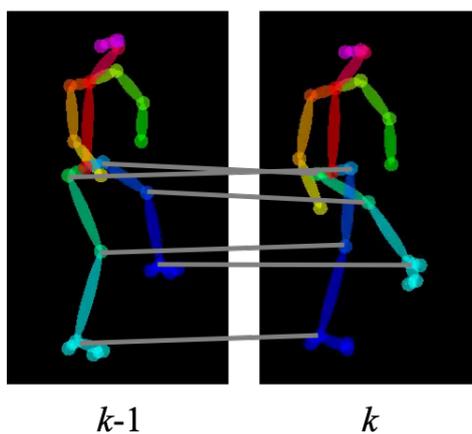


図 3.7 $D_{d,k}$ を算出する際の1対1の対応付け

図 3.6 および図 3.7 においてグレーの直線で結ばれている2つの部位の座標差を計算し、算出された6個の座標差の総和をそれぞれ $D_{s,k}$ 、 $D_{d,k}$ とする。ここで、 D_s から D_d を減算した値を D_{s-d} とおく。時間的に連続する2フレームのうち、どちらか片方のフレームのみ OpenPose

による足の検出が左右逆である場合、画像上における左右の膝や足首の距離が大きいかほど D_s は大きくなり、 D_{s-d} も大きくなる。そこで、各フレームにおいて、 D_{s-d} が一定以上の値となった場合には足の検出が左右反転しているとみなし、右腰と左腰、右膝と左膝、右足首と左足首の2次元座標値をそれぞれ入れ替える。なお、本研究で使用した OpenPose のモデルでは、かかとやつま先の座標値も取得可能であるが、本手法では使用しないため入れ替えない。これは、足首の座標値のみを使用すれば姿勢類似度の算出において十分であると予備実験で確認しているためである。

3.5.3 片足誤検出の修正

本研究で用いた映像の投球シーンにおいて、OpenPose によって片足のみ誤って検出されているフレームを確認した。例を図 3.8 に示す。

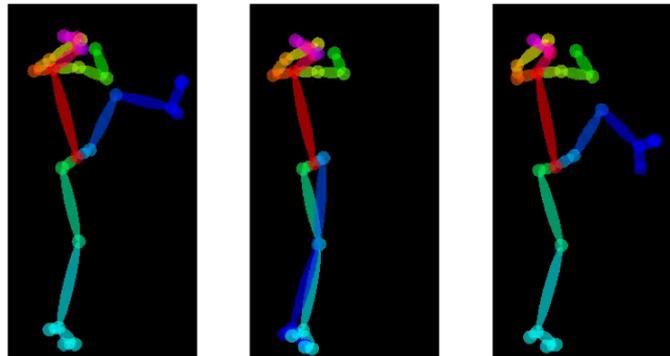


図 3.8 片足誤検出の例

図 3.8 に示した 3 枚の画像は、同一動画内の時間的に連続する 3 フレームにそれぞれ映る投手に対して OpenPose を適用した結果である。中央の画像のみ、OpenPose による左足の誤検出が存在する。水色の線が OpenPose により右足として検出された骨格、青色の線が OpenPose により左足として検出された骨格をそれぞれ示す。3 枚の画像において OpenPose により推定された左足の位置を比較すると、左右の 2 枚の画像では左足の位置が高いが、中央の画像では低いことがわかる。図 3.8 のような誤検出の場合、以下に示す二つの特徴がある。

図 3.8 の左の画像と中央の画像を比較すると、青色の線で示された左足の位置は大きく異なるが、水色の線で示された右足の位置はほとんど変わらないことが確認できる。このように、誤って検出された足の位置は直前のフレームでの足の位置と大きく異なり、もう片方の足の位置は直前のフレームとほとんど同じ位置である。

また、中央の画像では、推定された両膝の位置が近いことが確認できる。片足のみが誤って検出されている場合の多くで、図 3.8 の中央の画像における骨格のように両膝や両足首の OpenPose による推定位置が近いことを本研究で使用した投球動画より確認した。

以上の二つの特徴より、両膝の座標差および両足首の座標差がともに小さく、かつ直前のフレームと比べて片足の位置が大きく異なり、もう片方の足の位置がほとんど変わらないフレームでは、片足のみ誤検出したと判断し、座標値を修正する。なお、座標を修正する部位は、直前のフレームと比較し位置が大きく異なる足の膝および足首である。例として、図 3.8 の場合は、中央の画像における左膝と左足首の座標値を修正する。なお、修正後の座標値は直前のフレームにおける同一部位の座標値とする。

3.6 ベクトルデータの作成

本研究では投球映像に対し、OpenPose の 25 点が検出可能なモデルを適用し映像内の人物の身体座標を取得する。本手法では、検出可能な 25 部位のうち、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、今回使用した映像について、12 点が安定して検出可能であることを確認したためである。対象とする 12 点のうち、任意の二つの部位を結ぶベクトルの長さや角度を OpenPose によって得られる各部位の 2 次元座標値から計算する。なお、二つの部位のうち、少なくとも片方の部位が OpenPose により検出されていない場合は、長さ、角度をともに 0 とする。この際、各ベクトルの長さは選手の身長やカメラ位置等に依存するため、長さについて正規化する。具体的には、投手の左膝と左足首を結ぶベクトルの長さの最大値ですべてのベクトルの長さを除算する。

また、OpenPose によって得られる各部位の信頼度を用いてベクトルの信頼度を求める。二つの部位 p と q からなるベクトルの信頼度 C_{vector_pq} は式(3.3)で与えられる。

$$C_{vector_pq} = \min (C_p, C_q) \quad (3.3)$$

ここで、 C_p 、 C_q はそれぞれ OpenPose から取得した二つの部位 p 、 q の信頼度である。二つの部位のうち、少なくとも片方の部位が OpenPose により検出されていない場合は、信頼度は 0 となる。

映像の各フレームにおいて、OpenPose によって得られた身体座標データから長さ、角度、信頼度を持つ 66 次元のベクトルが作成される。

3.7 姿勢類似度の算出

本研究は DTW を用いて二つの投球映像のフレームを対応付ける。DTW は二つの系列内の要素間の類似度に基づき対応付けを決定する。そのため、二つの映像のフレーム間の類似度を求める必要がある。本手法の目的は、投球動作のタイミングを合わせることであるため、投手の姿勢がどれだけ似ているかに基づきフレーム間の類似度を算出する。第 3.6 節で述べた、長さ、角度、信頼度の 3 要素からなる 66 次元のベクトルを用いて投手の姿勢の類似度を算出する。

まず、対応するベクトルの差を計算する。ここで、対応するベクトルは、照合対象である

2 投手の投球する腕が同一であるか否かによって一部が異なる。ベクトルとは OpenPose により検出される首、両肩、両肘、腰の中央、両腰、両膝、両足首の 12 点のうち、任意の二つの部位の 2 次元座標値を結ぶ直線である。ベクトルを構成する二つの部位の組み合わせはベクトルごとに異なる。右投手同士もしくは左投手同士の投球動作に対してフレーム間の類似度を算出する場合は、対応するベクトルとは二つの部位の組み合わせが同一であるベクトルを指す。一方で、右投手と左投手の投球動作を照合する場合は、対応するベクトルにおいて、常に二つの部位の組み合わせが同一であるとは限らない。右投手と左投手では、ボールをリリースする手や投球時の軸足が左右逆である。そのため、常に二つの部位の組み合わせが同一であるベクトルの差を計算することは、右投手と左投手の投球姿勢の類似度を算出することに適していない。そこで、照合対象となる 2 投手の投球する腕が左右異なる場合は、姿勢の類似度を正確に算出するために、一部について二つの部位の組み合わせが異なるベクトルの差を計算する。

対応するベクトルの差は、二つのベクトルの角度差と長さの差の積とする。ただし、二つのベクトルのうち少なくとも片方のベクトルの信頼度が閾値未満である場合は、角度差および長さの差を計算せず、定数値を対応するベクトルの差とする。これは、OpenPose の検出が失敗した部位が姿勢の類似度に与える影響を減らすためである。式(3.3)に示したように、二つの部位 p と q からなるベクトルの信頼度は、OpenPose により得られる身体部位 p と q の信頼度の最小値である。OpenPose によって取得される身体部位の信頼度が低い場合は、推定された部位の位置が誤っている可能性が高くなる。そこで、二つのベクトルの信頼度がともに閾値以上である場合のみ、二つのベクトルの角度差と長さの差の積を対応するベクトルの差とし、少なくとも片方のベクトルの信頼度が閾値未満である場合は対応するベクトルの差は定数値とする。

対応するベクトルの差をすべて加算した値をフレーム間の類似度とする。

3.8 Dynamic Time Warping による投球動作の対応付け

本研究は、二つの系列内の投球動作のタイミングを対応付けるものである。二つの系列のうち、片方をモデル系列、もう片方を入力系列とする。モデル系列、入力系列ともに 1 球分の投球動作が含まれており、モデル系列に関しては目視により投球動作の始点および終点をあらかじめ決定する。よって、モデル系列には投球以外の動作は含まれない。一方で、入力系列は投球動作の始点および終点が未知の状態を与える。そのため、入力系列に含まれる投球動作の始点および終点を決定するとともに DTW によって投球動作を対応付ける必要がある。

第 2.4 節で述べた投球動作照合の従来手法[7]では、片方の系列内の投球動作の始点および終点を自動で抽出する始点終点自由な DTW が提案されている。本手法では始点終点自由な DTW を用いて二つの投球動作を対応付ける。長さが M のモデル系列と長さが N の入力系列に対する始点終点自由な DTW の概要を以下に示す。

まず、 $(M + 1) \times (N + 1)$ の行列 D および行列 E を用意する。二つの行列の型はともに数値であり、初期状態として1行目の要素の値をすべて0とする。また、行列 D は1行1列成分を除く1列目の要素をすべて無限大とし、行列 E は1行1列成分を除く1列目の要素をすべて1とする。 $M = 5, N = 7$ とした場合の行列 D および行列 E の初期状態を図3.9および図3.10にそれぞれ示す。なお、図3.9、図3.10において白色のマスが行列の要素であり、黄色のマスは行列の添字である。縦方向の数字が行の添字、横方向の数字が列の添字をそれぞれ示し、図3.9における左上の白色のマスは $D(0, 0)$ と表す。

	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	∞							
2	∞							
3	∞							
4	∞							
5	∞							

図3.9 $M = 5, N = 7$ とした場合の行列 D の初期状態

	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	1							
2	1							
3	1							
4	1							
5	1							

図3.10 $M = 5, N = 7$ とした場合の行列 E の初期状態

行列 D について、1行目および1列目を除く要素 $D(m, n)$ ($1 \leq m \leq M, 1 \leq n \leq N$)は式(3.4)に基づき値が算出される。

$$D(m, n) = \min \left\{ \begin{array}{l} \frac{D(m-1, n) \times E(m-1, n) + s(m, n)}{E(m-1, n) + 1} \\ \frac{D(m, n-1) \times E(m, n-1) + s(m, n)}{E(m, n-1) + 1} \\ \frac{D(m-1, n-1) \times E(m-1, n-1) + s(m, n)}{E(m-1, n-1) + 1} \end{array} \right. \quad (3.4)$$

ここで、 $s(m, n)$ はモデル系列の m フレームと入力系列の n フレーム間の姿勢の類似度である。なお、式(3.4)に示した3つの分数のうち、最小値となる分数の分母を $E(m, n)$ とする。

式(3.4)により、行列 D のすべての成分を類似度の平均値に基づいて算出後、対応付けを決定する。 $D(M, n)$ ($1 \leq n \leq N$)の値をすべて比較し、最小値を求める。

上述の手法では、姿勢の類似度の平均値が最小となるように対応付けを決定する。しかし、本手法は平均値を最小にする性質上、最適な対応付けではないが、姿勢が似ている2フレームを多く対応付けする可能性がある。例として、片方の系列の a_1, a_1+1 フレームともう片方の系列の b_1, b_1+1 フレームについて、最適な対応付けが $(a_1, b_1), (a_1+1, b_1+1)$ の二つであるとする。なお、 (a_1, b_1) は a_1 フレームと b_1 フレームの対応付けを表す。この際、最適でない対応付けである (a_1+1, b_1) もしくは (a_1, b_1+1) について2フレーム間の投手の姿勢の類似度が小さい場合、2フレームを対応付けることにより全体の類似度の平均値は小さくなる。そのため、過剰な対応付けが発生することとなる。そこで、このようなケースでは、各対応付けにおける姿勢の類似度に基づき余計な対応付けの発生を防ぐ。

また、モデル系列の終点フレームが入力系列の複数のフレームと過剰に対応付けされることも考えられる。そこで、モデル系列の終点フレームを含む対応付けが複数存在する場合には、モデル系列の終点フレームと最も投手の姿勢が類似している入力系列のフレームを入力系列の投球動作の終点とし、それ以降の対応付けはしない。

始点終点自由なDTWにより、入力系列の投球動作の始点および終点を決定後、DTWの計算に使用されていないフレームを対応付ける。ここで、DTWの計算に使用されていないフレームとは、第3.4.3項で述べた手法により投手の姿勢変化に基づいて削減されたフレームである。同一系列内の時間的に連続する複数フレームにおいて、投手の姿勢がほとんど変化しない場合、姿勢が似ている複数フレームのうち1フレームのみがDTWの計算に使用され、残りのフレームは削減される。系列A内の時間的に連続するかつ投手の姿勢が似ている $f, f+1, \dots, f+l$ フレーム (l は自然数) について、DTWの計算に使用された f フレームが系列Bの g フレームと対応付けされたとする。この場合、DTWの計算に使用されていない $f+1, \dots, f+l$ フレームを系列Bの g フレームと対応付ける。

3.9 むすび

本章では、映像から取得した投手の姿勢情報に基づく投球動作照合の提案手法について述べた。

第4章 実験結果および考察

4.1 まえがき

本章では、提案手法に対する評価実験の結果を示す。まず、第3.4節で述べたフレーム数削減処理および第3.5節で述べた OpenPose による誤検出部位の補正処理に関する評価実験結果を示す。次に、従来手法による投球動作照合と提案手法による投球動作照合の実験を行い、対応付け結果を比較する。また、結果をもとに提案手法の有効性について考察する。

4.2 実験の概要

本研究で用いた実験データセットは、野球中継映像における 101 本の投球映像である。このうち、1 本が左投手のオーバースローの投球映像であり、100 本が右投手のオーバースローの投球映像である。各映像には 1 球分の投球動作が含まれている。右投手のオーバースローの投球映像のうち、1 本をモデル系列 1 とする。また、左投手のオーバースローの投球映像をモデル系列 2 とする。なお、各モデル系列について、元映像から投球動作の始点および終点を目視によりあらかじめ決定し、モデル系列には投球動作のみが含まれるものとする。ここで、投球動作の始点は軸足でない足が地面から離れる瞬間、終点は蹴り上げた軸足が地面に着く直前とした。右投手の場合、軸足は右足となり、左投手の場合、軸足は左足となる。モデル系列 1 の元映像を含む 100 本の右投手のオーバースローの投球映像を入力系列とする。入力系列は投球動作の始点と終点が未知の状態を与える。モデル系列、入力系列ともに、投球している選手が右投手であるか左投手であるかは既知とした。

モデル系列と入力系列に対して本手法を適用し、対応付けされたモデル系列のフレームと入力系列のフレームを左右に並べた画像を作成する。対応付け結果の例を図 4.1 に示す。



図 4.1 対応付け結果の例

図 4.1 において、左がモデル系列のフレーム、右が入力系列のフレームである。また、対応付けされた 2 フレームに映る投手の骨格のみを左右に並べた画像を作成する。図 4.1 内の投手の骨格のみを左右に並べた画像を図 4.2 に示す。

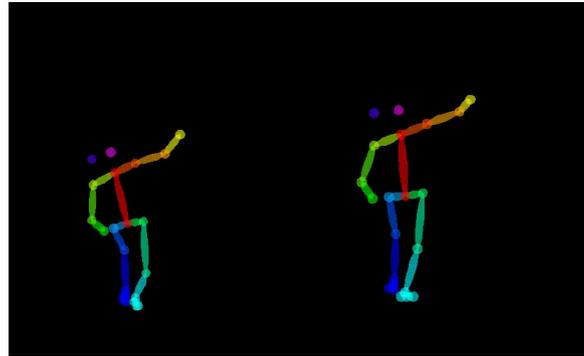


図 4.2 投手の骨格比較画像

得られたモデルフレームと入力フレームの各対応付けについて、より対応付けに適したフレームが存在するかに基づき、目視により対応付けの成否を判断した。

4.3 フレーム数削減の評価

4.3.1 概要

第 3.4 節で述べたフレーム数削減の結果を示す。モデル系列 1 と 100 本の入力系列を使用して、フレーム数を削減する処理の精度およびフレーム数削減による対応付けの実行時間を確認した。なお、モデル系列 1 は 100 フレーム、100 本の入力系列はすべて合わせて 17960 フレームである。

4.3.2 精度評価

第 3.4.2 項で述べた、OpenPose によって検出された人物に基づくフレーム数削減の精度を表 4.1 に示す。削減されたフレームについて、投球フレームであるかを目視により判断し、投球フレームでなければ成功フレーム、投球フレームであれば失敗フレームとした。ここで、投球フレームとは、画像中に投手が存在し、投手が投球動作を行なっているフレームを指す。

表 4.1 検出人数に基づくフレーム数削減の精度

	成功フレーム	失敗フレーム
フレーム数	57	0
割合 [%]	100.0	0.0

削減されたフレームはすべて投球フレームではなかった。また、今回使用した入力系列において、OpenPose による検出人数が 2 人以下のフレームは合計で 57 フレームのみであり、57 フレームすべてが削減されたことを確認した。

第 3.4.3 項で述べた，投手の姿勢変化に基づくフレーム数削減の精度を示す．削減されたフレームのうち，投球フレームについて対応付けの結果を確認した．各対応付けについて，より対応付けに適したフレームが存在する場合は誤った対応付け，存在しない場合は正しい対応付けと判断した．投手の姿勢変化に基づいて削減された投球フレームの対応付け結果を表 4.2 に示す．

表 4.2 投手の姿勢変化に基づき削減された投球フレームの対応付け結果

	正しい対応付け	誤った対応付け
フレーム数	2987	49
割合 [%]	98.39	1.61

多くのフレームについて，正しく対応付けされていた．一方で，誤って対応付けされたフレームについては，削減せずに対応付けた場合でも同様の結果が得られた．今回使用したデータセットにおいて，フレーム数削減によって対応付けの精度が低下したケースは確認されなかった．

4.3.3 実行時間評価

OpenPose によって検出された人物に基づくフレーム数削減の処理を削減処理 1，投手の姿勢変化に基づくフレーム数削減の処理を削減処理 2 とする．削減処理の有効性を確認するために，4 つの手法において実行時間を測定した．4 つの手法の流れを図 4.3 に示す．

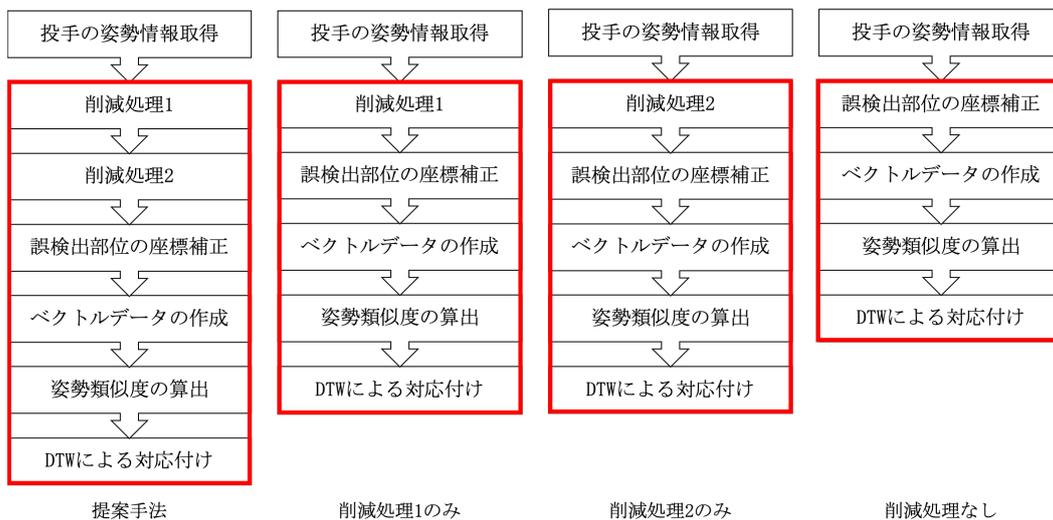


図 4.3 削減処理の有無が異なる 4 手法

図 4.3 に示した 4 つの手法は削減処理の有無のみが異なる．モデル系列 1 と 100 本の入力系列を使用して，各手法の赤枠で囲まれた処理の実行時間を測定した．1 入力系列あたりの平

均実行時間を表 4.3 にまとめる。

表 4.3 4 手法の平均実行時間

	提案手法	削減処理 1 のみ	削減処理 2 のみ	削減処理なし
平均実行時間 [s]	5.22	6.10	5.17	6.09

表 4.3 に示したように、削減処理 1 および削減処理 2 をともに加えた提案手法は、削減処理なしの手法に比べて、実行時間が減少した。一方で、削減処理 1 のみを加えた手法は削減処理なしの手法と比べ、実行時間がほとんど変化しなかった。これは今回使用した投球動画において、OpenPose による検出人数が 2 人以下のフレームがほとんど存在しなかったことが原因である。検出人数が 2 人以下のフレームが 25 フレーム存在する入力系列に対して 4 手法を実行した際の実行時間を表 4.4 に示す。なお、入力系列は 165 フレームからなる動画である。

表 4.4 1 入力系列に対する 4 手法の実行時間

	提案手法	削減処理 1 のみ	削減処理 2 のみ	削減処理なし
実行時間 [s]	4.41	4.66	4.68	5.40

OpenPose による検出人数が 2 人以下のフレームが含まれる系列の場合には、削減処理 1 を加えることで実行時間が減少した。

4.4 誤検出部位の座標補正評価

第 3.5.2 項で述べた、OpenPose による足の左右反転検出の修正について評価実験を行った。座標値を修正したフレームについて、OpenPose による足の検出が左右逆となっているかを確認した。足の検出が左右逆となっていた場合は修正成功フレーム、足の左右が正しく推定されていた場合は修正失敗フレームとした。評価実験の結果を表 4.5 に示す。

表 4.5 足の左右反転検出の修正に関する評価実験結果

	修正成功フレーム	修正失敗フレーム
フレーム数	26	4
割合 [%]	86.67	13.33

座標値を修正したフレームの多くで、実際に OpenPose による足の検出が左右逆となっていることが確認された。ただし、OpenPose による足の検出が正しいフレームにもかかわらず、左右の足の部位の座標値を誤った値に変更したケースも一部存在した。この際、座標値を修正したフレームではなく、直前のフレームにおいて OpenPose による足の検出が左右逆

になっていることが確認された。しかし、直前のフレームは座標値の修正がされておらず、その影響により、足の左右が正しく検出されているフレームにおいて座標値が誤った値に変更されたと考えられる。

上記のように、OpenPose による足の検出が左右逆となっているフレームにもかかわらず、座標値を修正していないフレームが存在した。このようなフレームの多くは、左右の足の位置が近いことが確認された。そのため、左右の足の位置が近い場合でも、OpenPose による足の検出が左右逆であるかを正確に判定可能な手法が必要となる。

第 3.5.3 項で述べた、OpenPose による片足誤検出の修正について評価実験を行った。片足の膝および足首の座標値が修正されたフレームについて、修正前の座標値および修正後の座標値を正しい座標値とそれぞれ比較し、座標差を算出した。ここで、目視により決定した各部位の位置の座標値を正しい座標値とした。片足の膝および足首について、修正後の座標値が修正前の座標値と比べて正しい座標値に近い場合は修正成功フレーム、そうでない場合は修正失敗フレームとした。評価実験の結果を表 4.6 に示す。

表 4.6 片足誤検出の修正に関する評価実験結果

	修正成功フレーム	修正失敗フレーム
フレーム数	32	0
割合 [%]	100.0	0.0

片足の膝および足首の座標値が修正されたフレームはすべて修正成功フレームであった。修正結果の例を図 4.4 に示す。

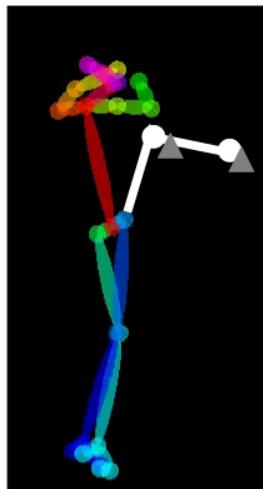


図 4.4 片足修正結果の例

図 4.4 において、白い丸の位置が左膝および左足首の正しい位置であり、白い直線が左足の

骨格である。OpenPose により左足として検出された骨格は青色の直線で示されている。グレーの三角が修正後の座標値に基づく左膝および左足首の位置を示す。なお、二つのグレーの三角のうち右が左足首の位置である。座標値を修正することにより正しい座標値との差を小さくすることが可能となった。

4.5 対応付け結果

4.5.1 従来手法

第 2.4 節で述べた従来手法について、モデル系列 1 と 100 本の入力系列を使用して、右投手の投球動作照合の精度を実験により確認した。得られた各対応付けが正しいかを目視により判定した。より対応付けに適したフレームが存在する場合は誤った対応付けとし、存在しない場合は正しい対応付けとした。従来手法の対応付け結果を表 4.7 に示す。

表 4.7 従来手法による右投手同士の投球動作照合結果

	正しい対応付け	誤った対応付け
対応付け数	8938	546
割合 [%]	94.24	5.76

多くの対応付けが正確であることが確認された。ただし、OpenPose による身体部位の誤検出が入力系列内の複数フレームで連続して存在する場合には不正確な対応付けが増加した。OpenPose の誤検出が連続して発生しているケースの例を図 4.5 に示す。

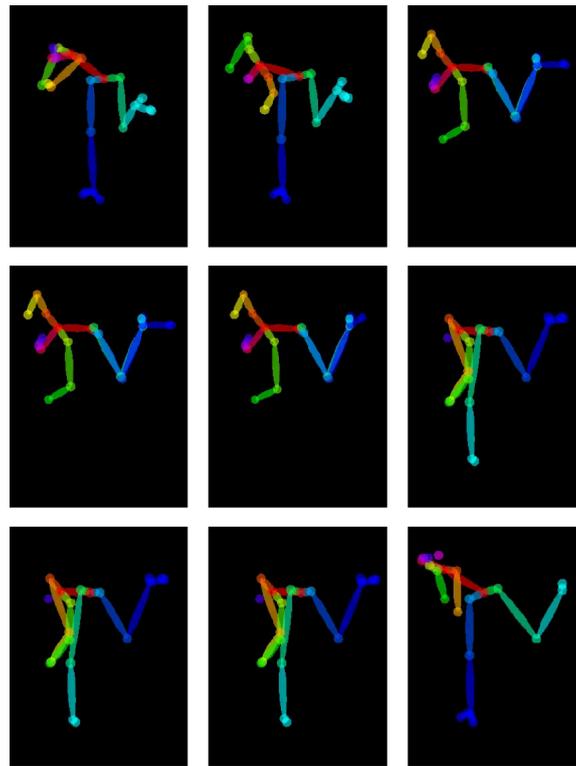
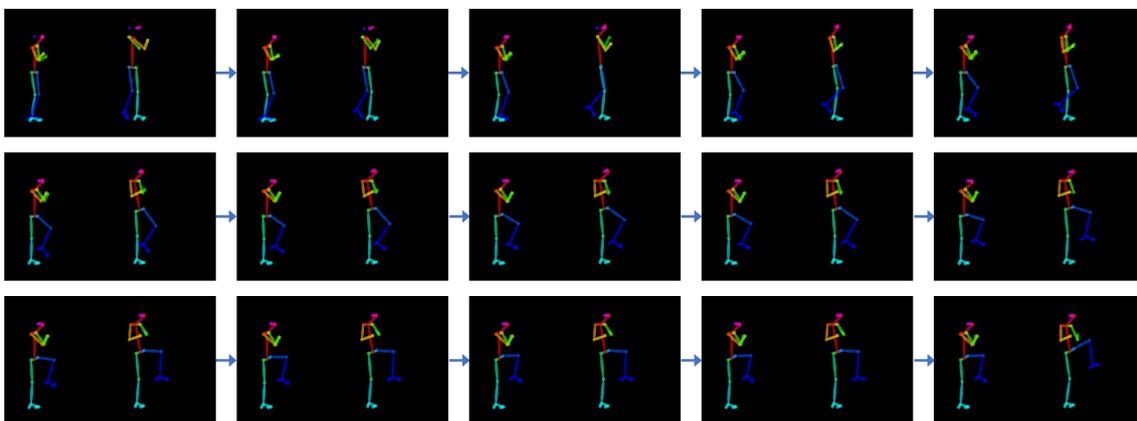


図 4.5 OpenPose による誤検出の連続

図 4.5 の 9 枚の画像は、同一系列内の連続する 9 フレームに OpenPose を適用し得られた投手の骨格推定結果である。左上の画像が 9 フレーム中の最初のフレームであり、各行で右方向に連続している。元映像と比較すると、左上と右下の 2 枚の画像については投手の身体部位が正しく検出されていることが確認できた。一方で、残りの 7 枚については足の左右が逆に推定されているなどの誤検出が存在した。モデル系列 1 と図 4.5 に示した誤検出を含む入力系列に対する従来手法による対応付けの結果を図 4.6 に示す。なお、図 4.6 では左上の画像が最初の対応付け、一番下の行の右の画像が最後の対応付けであり、各行で右方向に連続している。





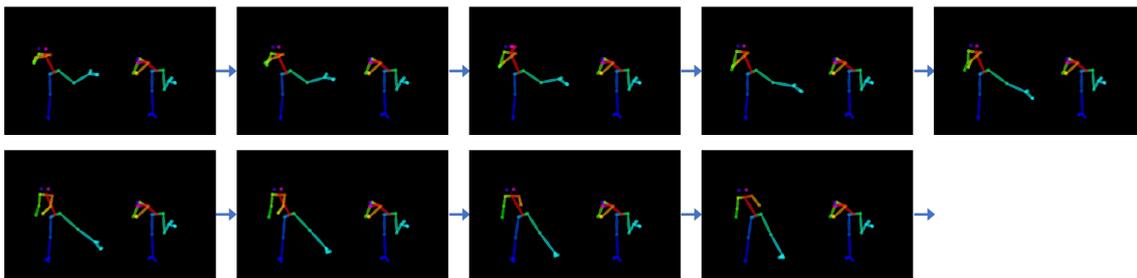


図 4.6 従来手法による右投手同士の対応付け結果

図 4.6 における各画像は、従来手法によって対応付けされたフレームを左右に並べた結果である。ただし、投手の姿勢を比較しやすくするために OpenPose による投手の骨格推定結果のみを表示している。

図 4.6 中の赤枠で囲まれている画像から最後の画像までの 13 枚において、右側の投手の骨格が変化していない。赤枠で囲まれている画像と最後の画像を拡大し図 4.7 に示す。図 4.7 の左が図 4.6 において赤枠で囲まれている画像、右が図 4.6 の最後の画像である。

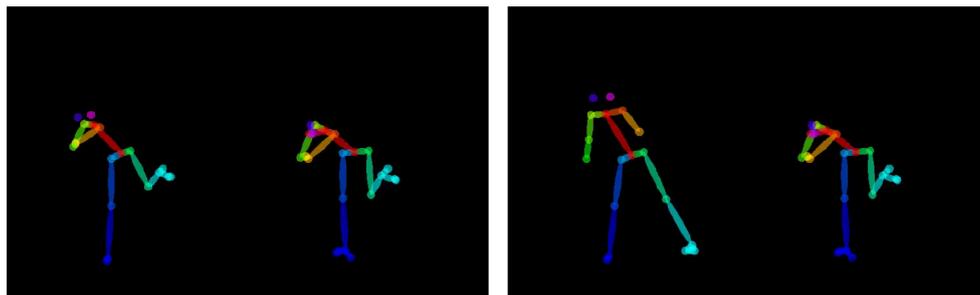


図 4.7 従来手法による対応付けの一部

図 4.7 において各画像の右側の投手の骨格は図 4.5 の左上の画像の骨格と同一である。図 4.7 の左の画像では投手の骨格が類似しているが、右の画像では骨格が大きく異なっており、不正確な対応付けである。これは、図 4.5 に示したように、図 4.7 における各画像の右側のフレーム以降、OpenPose による投手の身体部位の誤検出が連続して発生していることが原因であると推察した。

4.5.2 提案手法

提案手法について、モデル系列 1 と 100 本の入力系列を使用して、右投手の投球動作照合の精度を実験により確認した。得られたモデルフレームと入力フレームの各対応付けについて、より対応付けに適したフレームが存在するかに基づき、目視により対応付けの成否を判断した。提案手法の対応付け結果を表 4.8 に示す。

表 4.8 提案手法による右投手同士の投球動作照合結果

	正しい対応付け	誤った対応付け
対応付け数	10063	523
割合 [%]	95.06	4.94

従来手法に比べ、対応付けの精度が向上した。従来手法において不正確な対応付けが多かったケースについて提案手法を実行し得られた結果を図 4.8 に示す。なお、図 4.8 では左上の画像が最初の対応付け、一番下の行の右の画像が最後の対応付けであり、各行で右方向に連続している。





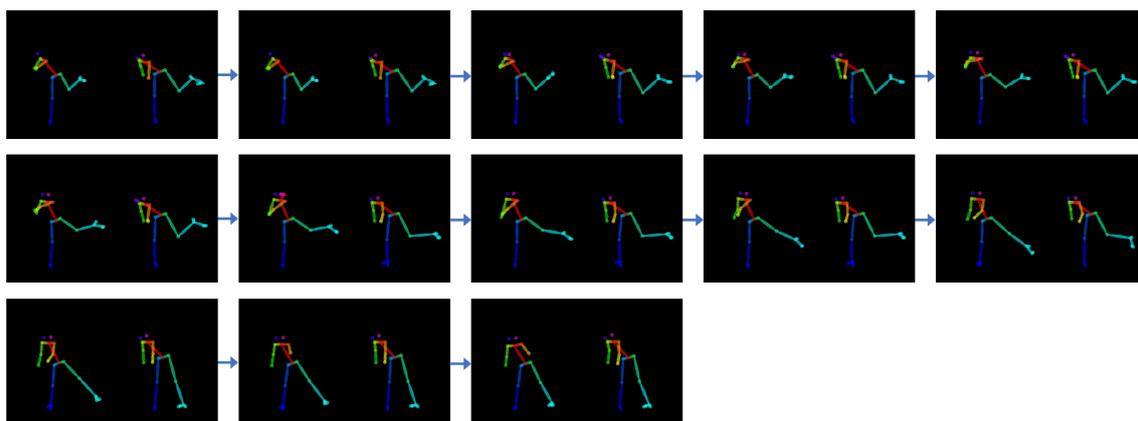


図 4.8 提案手法による右投手同士の対応付け結果

図 4.6 に示したように従来手法による対応付けでは、投球動作の終点付近において、不正確な対応付けが発生していたが、提案手法では終点付近も正確に対応付けされた。図 4.8 に示したケースでは、OpenPose によって足の左右が逆に検出されているフレームも正しい対応付けとなっていた。これは、OpenPose による誤検出部位の座標値を一部修正する処理が本手法に含まれており、修正した座標値を用いて姿勢の類似度を算出することで正しい対応付けとなったと推察した。

ただし、誤検出が含まれるフレームに対して常に正確な対応付けを実現することはできなかった。本手法では、OpenPose による足の検出が左右逆となるなど特定の条件を満たす誤検出のみ座標値を修正した。対応付けの精度をさらに向上させるためには、あらゆる誤検出に対して座標値を修正可能な処理が必要となる。

左投手のオーバースローの投球映像であるモデル系列 2 と 100 本の入力系列を使用して、左投手と右投手に対する投球動作照合の精度を実験により確認した。得られたモデルフレームと入力フレームの各対応付けについて、より対応付けに適したフレームが存在するかに基づき、目視により対応付けの成否を判断した。左投手と右投手の対応付け結果を表 4.9 に示す。

表 4.9 提案手法による左投手と右投手の投球動作照合結果

	正しい対応付け	誤った対応付け
対応付け数	8478	726
割合 [%]	92.11	7.89

左投手と右投手に対して提案手法を実行した結果を図 4.9 に示す。なお、図 4.9 では左上の画像が最初の対応付け、一番下の行の右の画像が最後の対応付けであり、各行で右方向に連続している。



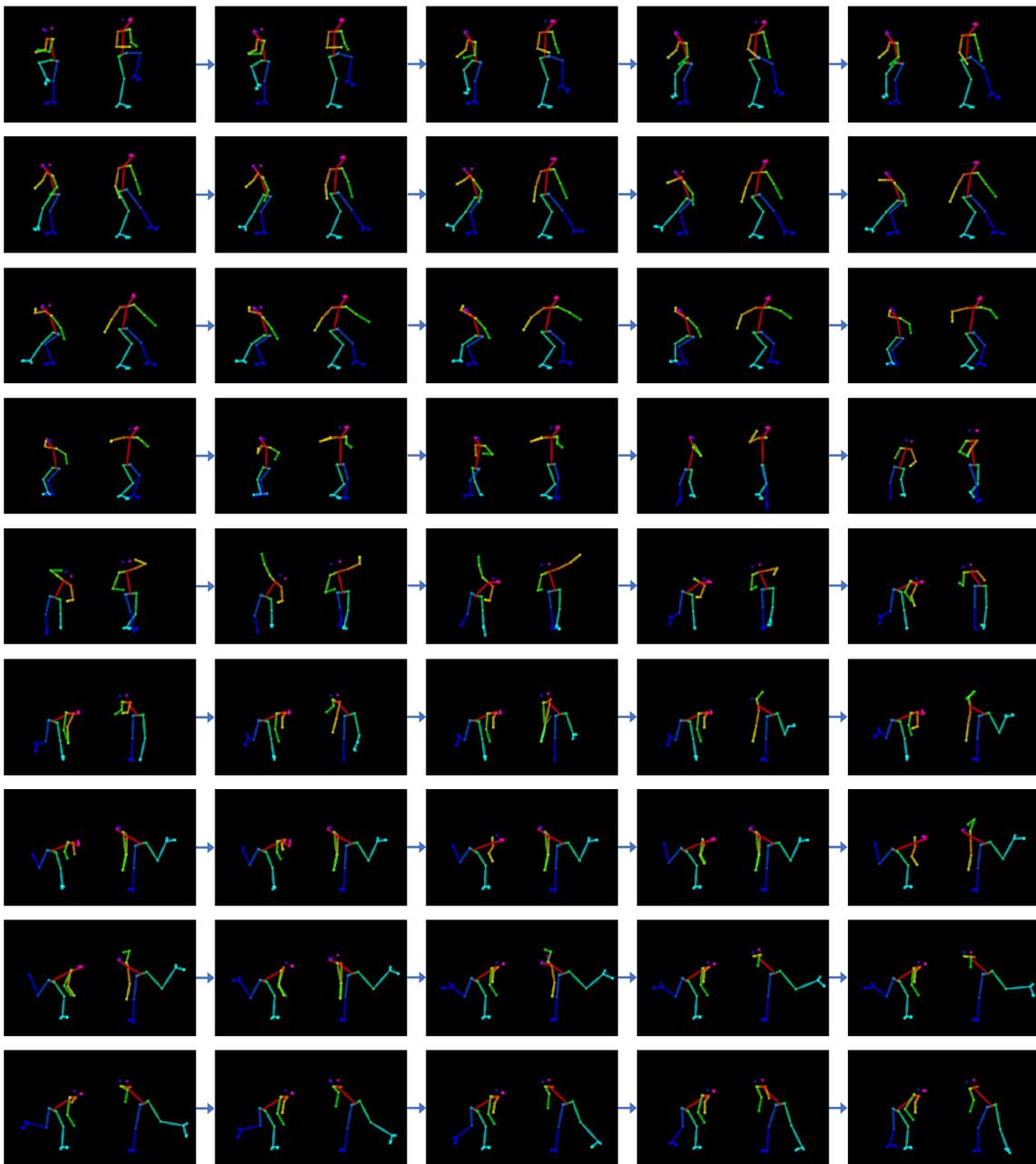


図 4.9 提案手法による左投手と右投手の対応付け結果

図 4.9 の各画像において、左側が左投手の骨格、右側が右投手の骨格を示す。投球する腕が左右異なる場合でも良好な結果が得られた。ただし、右投手同士の対応付け精度に比べ誤った対応付けが増加した。これは、投手とカメラの位置関係が一つの要因であると推察した。今回使用したデータセットにおいては、左投手、右投手の投球映像はともに同じ方向から撮影されている。例を図 4.10 に示す。

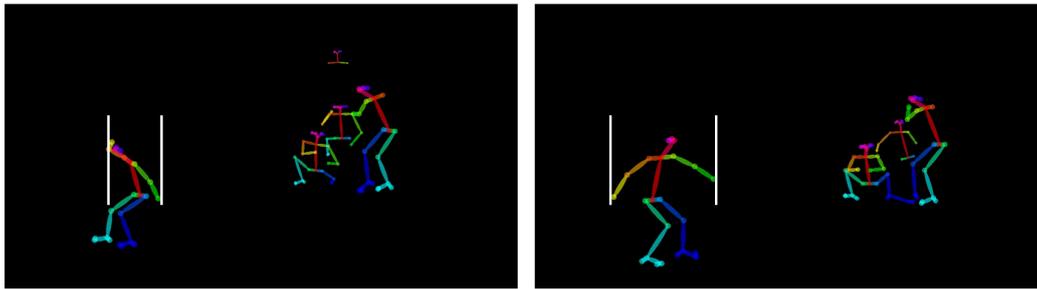


図 4.10 投球映像の比較

図 4.10 の 2 枚の画像では、ともに投手が画像の左側、打者や捕手が右側に存在しており、カメラの位置はほぼ同一である。各画像において白い 2 本の直線は投手の両手首の横方向の位置を表している。図 4.10 に示した 2 枚の画像において投手の姿勢は類似しているが、左の画像に映る左投手と右の画像に映る右投手では両手首の横方向の距離が大きく異なることがわかる。投手とカメラの位置関係が同一であるため、類似する姿勢であっても左投手と右投手では画像上における関節間の距離が大きく異なる場合がある。このようなケースでは姿勢の類似度を正確に求めることが困難となり、不正確な対応付けとなる可能性がある。

4.6 むすび

本章では、提案手法に対する評価実験の結果を示した。まず、フレーム数削減処理および OpenPose による誤検出部位の補正処理に関する評価実験結果を示した。次に、従来手法と提案手法による投球動作照合結果を比較し、提案手法の有効性を確認した。

第5章 結論と今後の課題

5.1 結論

本研究では、映像から取得した投手の骨格情報を用いて二つの投球動作を照合する手法を提案した。OpenPoseによる身体部位の誤検出について、座標値を修正する処理を提案手法に含めることで、従来手法よりも対応付け精度が0.82ポイント向上した。また、DTWの計算に使用するフレーム数を減らすことで、対応付けの精度に影響を与えることなく実行時間を減少させることが可能となった。さらに、左投手と右投手についても多くの場合において正しい対応付けを実現可能であるという結果を得られた。

5.2 今後の課題

提案手法では、OpenPoseによる足の検出が左右逆となるなど特定の条件を満たさず誤検出のみ座標値の修正を図った。そのため、誤検出が含まれるフレームすべてを正確に対応付けることは非常に困難である。対応付け精度をさらに向上させるために、あらゆる誤検出に対して座標値を修正可能な処理が必要となる。

また、左投手と右投手の投球動作を照合する際に、投手とカメラの位置関係により誤った対応付けが増加するという問題点がある。よって、投手とカメラの位置関係を考慮した姿勢類似度の算出手法を検討する必要がある。

謝辞

本研究の実験環境を整えてくださり，研究の方向性や問題点などに関して丁寧かつ熱心にご指導くださった渡辺裕教授に深く感謝申し上げます。

本研究や論文執筆において適切なアドバイスを数多くくださり，丁寧かつ熱心にご指導くださった早稲田大学国際情報通信センターの石川孝明様に深く感謝申し上げます。

日頃から貴重なご意見をくださる研究室の皆様に心より御礼申し上げます。

最後に，ここまで私を育ててくださり，常に支えてくださった家族に深く感謝いたします。

参考文献

- [1] 三宅隆, 滝田勝紀: “〈野球×デジタル〉『福岡ソフトバンクホークス』が iPhone&iPad で試合に勝つ方法【デジタルでスポーツの勝利をつかむ #2】”, MELOS, 2017年8月4日公開 (最終閲覧日: 2019年12月21日) <https://melos.media/hobby/5934/>
- [2] 三上弾, 高橋康輔, 西條直樹, 五十川麻理子, 木村聡貴, 木全英明: “VR イメージトレーニングシステムの実現と野球への適用”, NTT 技術ジャーナル, Vol.30, No.1, pp.22-25, Jan. 2018
- [3] Z. Cao, T. Simon, S. Wei, and Y. Sheikh: “Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), No.121, pp.1302-1310, July 2017
- [4] S. Sempena, N. U. Maulidevi, and P. R. Aryan: “Human action recognition using dynamic time warping”, Proceedings of the 2011 International Conference on Electrical Engineering and Informatics (ICEEI), July 2011
- [5] 森駿文, 菊池浩明: “歩容データの DTW 距離に基づく個人識別手法の提案と外乱に対する評価”, マルチメディア, 分散, 協調とモバイル (DICOMO2018) シンポジウム, pp.672-680, July 2018
- [6] D. F. Silva, G. E. A. P. A. Batista, and E. Keogh: “Prefix and Suffix Invariant Dynamic Time Warping”, IEEE 16th International Conference on Data Mining (ICDM), pp.1209-1214, Dec. 2016
- [7] 横井真也: “スポーツ映像から取得した身体座標データに対するアラインメント”, 早稲田大学大学院基幹理工学研究科情報理工・情報通信専攻修士論文, Feb. 2019

図一覧

図 2.1	OpenPose による検出結果の例	4
図 2.2	OpenPose による不正確な検出の例	5
図 2.3	要素数の異なる 2 系列に対する DTW の適用例.....	6
図 2.4	形状の似ている 2 系列に対する DTW の適用例.....	7
図 3.1	投球映像に対する OpenPose の適用例	11
図 3.2	審判が検出されているフレーム.....	12
図 3.3	審判が検出されていないフレーム.....	12
図 3.4	投手の姿勢がほとんど変化しない連続フレーム.....	13
図 3.5	足の左右反転検出例.....	14
図 3.6	D_s, k を算出する際の 1 対 1 の対応付け	16
図 3.7	D_d, k を算出する際の 1 対 1 の対応付け.....	16
図 3.8	片足誤検出の例.....	17
図 3.9	$M = 5, N = 7$ とした場合の行列 D の初期状態	20
図 3.10	$M = 5, N = 7$ とした場合の行列 E の初期状態.....	20
図 4.1	対応付け結果の例.....	22
図 4.2	投手の骨格比較画像.....	23
図 4.3	削減処理の有無が異なる 4 手法.....	24
図 4.4	片足修正結果の例.....	26
図 4.5	OpenPose による誤検出の連続	28
図 4.6	従来手法による右投手同士の対応付け結果.....	30
図 4.7	従来手法による対応付けの一部.....	30
図 4.8	提案手法による右投手同士の対応付け結果.....	33
図 4.9	提案手法による左投手と右投手の対応付け結果.....	35
図 4.10	投球映像の比較.....	36

表一覧

表 3.1	i の値と部位の対応関係	15
表 3.2	Ds, k を算出する際の1対1の対応付け	15
表 3.3	Dd, k を算出する際の1対1の対応付け	16
表 4.1	検出人数に基づくフレーム数削減の精度	23
表 4.2	投手の姿勢変化に基づき削減された投球フレームの対応付け結果	24
表 4.3	4手法の平均実行時間	25
表 4.4	1入力系列に対する4手法の実行時間	25
表 4.5	足の左右反転検出の修正に関する評価実験結果	25
表 4.6	片足誤検出の修正に関する評価実験結果	26
表 4.7	従来手法による右投手同士の投球動作照合結果	27
表 4.8	提案手法による右投手同士の投球動作照合結果	31
表 4.9	提案手法による左投手と右投手の投球動作照合結果	33

研究業績

- [1] 大澤遼平, 兼近紘之, 武井淳也, 山川敦也, 渡辺裕: “OpenPose を用いたピッチングフォームの解析”, 映像情報メディア学会スポーツ情報処理時限研究会, Feb. 2019
- [2] 大澤遼平, 石川孝明, 渡辺裕: “スポーツシーン照合に関する一検討”, 映像情報メディア学会年次大会, 31C-3, Aug. 2019
- [3] 稲田健太郎, 大澤遼平, 山川敦也: “OpenPose”, 映像情報メディア学会誌「私の研究開発ツール —第 107 回—」, Vol. 73, No. 5, pp.932-934, Sep. 2019
- [4] 大澤遼平, 石川孝明, 渡辺裕: “異なるピッチングスタイルに対するシーン照合の一検討”, 映像情報メディア学会スポーツ情報処理時限研究会・メディア工学研究会 ME2019-117, Vol.43, No.39, pp.13-14, Nov. 2019
- [5] 大澤遼平, 石川孝明, 渡辺裕: “OpenPose と Dynamic Time Warping を用いた投球シーン照合の一検討”, 映像情報メディア学会冬季大会, 11A-3, Dec. 2019
- [6] 大澤遼平, 石川孝明, 渡辺裕: “OpenPose を用いた投球シーン照合の一検討”, 情報処理学会第 82 回全国大会, 2020 年 3 月 5 日発表予定