

令和 5 年 6 月 14 日現在

機関番号：32689

研究種目：基盤研究(C)（一般）

研究期間：2020～2022

課題番号：20K11344

研究課題名（和文）機械学習による人物運動解析の高精度化

研究課題名（英文）Improving the accuracy of human motion analysis by machine learning

研究代表者

渡辺 裕（WATANABE, HIROSHI）

早稲田大学・理工学術院・教授

研究者番号：10329154

交付決定額（研究期間全体）：（直接経費） 2,800,000円

研究成果の概要（和文）：人物の動作映像から骨格情報を検出し、それらの骨格情報に基づく動作対応付けを自動的に行う手法についての研究に取り組んできた。二つの時系列データの対応付けを行う動的時間伸縮法において、従来法に比べて骨格形状に基づく姿勢の対応付けの精度及び動作の開始と終了点の整合が改善されることを示した。また処理おける演算量削減を可能とした。人物の動作映像から得られる骨格情報は、3次元オブジェクトが2次元に射影された形状となる。そこで、2次元の骨格情報から3次元姿勢を推定する手法について検討を行った。骨格の左右対称性及び骨格の絶対長が変化しないという前提条件の元で、姿勢推定結果が改善されることを示した。

研究成果の学術的意義や社会的意義

本研究では、機械学習を用いて得られる骨格情報をもとに運動動作の解析に必要な処理の高精度化と高速化について検討を行った。特に、動作パターン認識に加えて、推定骨格情報の位置、角度、同期性、一致度の解析が進んだことに意義があると考えられる。機械学習を用いた骨格情報推定アルゴリズムは、非接触なアプローチであることに大きな利点があり、これらの研究成果はスポーツ運動分析や医療介護などの分野に対して幅広く適用できる。今後、スポーツにおけるスキル向上だけでなく、介護におけるリハビリの補助技術として、人類の生活品質の向上に貢献できると考えられる。

研究成果の概要（英文）：We have been working on a method for detecting skeletal information from human motion videos and automatically mapping motions based on the skeletal information. In the Dynamic Time Warping (DTW) method for mapping two time-series data, we have shown that DTW improves posture mapping based on skeletal shape and motion start/end point matching compared to the conventional method. It also made it possible to reduce the amount of DTW operations. The skeletal information obtained from human motion video is a two-dimensional projection of a three-dimensional object. Therefore, we studied a method for estimating 3D posture from 2D skeletal information. We showed that the posture estimation results can be improved under the assumption that the skeletal symmetry and the absolute length of the skeleton do not change.

研究分野：機械学習

キーワード：動作解析 骨格情報 姿勢推定 動作整合

1. 研究開始当初の背景

機械学習を用いた映像からの骨格情報推定手法である OpenPose[1]が米国カーネギーメロン大より提案されて以来、マーカーを用いない人物姿勢推定アルゴリズムが活発に研究されてきた。これらの手法では、映像のみから複数人物のおおよその位置、形態を推定することができる。コンピュータビジョンにおける姿勢推定問題は、人物の動作認識すなわちパターン認識に主眼が置かれている。したがって、推定された骨格の形態及び動作の認識が重要な目的であり、得られるデータの空間的な精度については重視されていなかった。

機械学習を用いた骨格情報推定アルゴリズムは、非接触なアプローチであることに大きな利点がある。したがって、これらのアルゴリズムはスポーツ運動分析などの分野に対して幅広く適用できると考えられる。スポーツ運動分析や介護姿勢推定の分野では、パターン認識に加えて、推定された骨格情報の位置、角度、同期性、一致度の解析が望まれている。

単一のカメラ映像に対して機械学習アルゴリズムを適用して得られる人物の骨格情報は、2次元の関節位置情報である。位置情報の空間的な精度は、画像の解像度、被写体として撮影される人物の画像中の大きさに依存する。したがって絶対値で与えられる位置情報を、直接スポーツの動作解析に用いることは容易ではなかった。

人物が特定の姿勢をとった場合に、推定された骨格情報から、骨格パーツの角度に変換することにより、相対的变化を比較することができる。また骨格パーツ間の長さを相対的に求めることで、個人特徴を抽出することができる。しかし、人物姿勢として推定される骨格の関節位置の2次元情報を、どのように解析目的に応じた相対評価指標に変換すべきかが明確ではなかった。

また、相対指標による動作解析を実行する場合、参照すべき基本動作に対して評価対象となる動作映像の区間抽出が必要となる。しかし、基本動作に一致する区間を抽出するためには評価対象映像の始点と終点を人間の主観により決定する必要があった。そこで、始点と終点を明示的に与えずに、どのように特徴量空間で整合区間を自動的に求めることができるかが、自動的な動作解析の実現に向けた問題となっていた。

2. 研究の目的

本研究では、映像のみから人物の姿勢情報を機械学習により推定し、推定された人物の骨格座標を用いて運動解析や動作解析を行い、動作認識や個人認識の識別率の向上を目指した。機械学習を用いた骨格情報推定アルゴリズムは、非接触なアプローチであることに大きな利点がある。したがって、これらのアルゴリズムはスポーツ運動分析などの分野に対して幅広く適用できると考えた。スポーツ運動分析や介護姿勢推定の分野では、パターン認識に加えて、推定された骨格情報の位置、角度、同期性、一致度の解析が望まれる。これらの解析結果から、運動や動作のパフォーマンスを向上させる動きに関する知識の獲得、トレーニング手法の改善、複数人の同時動作の一致度の評価、怪我の予防、医療介護動作の自動検出などへの応用の可能性を示すことを目指した。

3. 研究の方法

機械学習を用いた骨格情報推定アルゴリズムは、動作解析を目的として開発されたため動画を処理対象としている。しかし、骨格推定自体は静止画像に対して実行される。したがって、連続する画像フレーム間で一貫性のある座標位置が得られる保証がなく、骨格位置情報の欠損やゆらぎが生じる場合がある。そこで、画像フレーム間での動きの連続性に注目し、推定座標を補間あるいは修正する手法について検討を行った。

機械学習を用いた骨格情報推定アルゴリズムの多くは2次元の関節位置情報を与えるものであり、3次元情報が得られる手法は限られている。機械学習により2次元のデータを3次元化するアプローチが提案されているが、得られる座標の精度は不十分である。2次元骨格データを関節間の相対データとなるよう多次元の特徴ベクトルに変換し、特徴ベクトルの同期性や一致度を評価指標とするアプローチを取った。例えば、姿勢把握を目的とする場合には、肩幅と身長比率での正規化を行うような処理が適切であると考えられる。また、複数の人物の腕の振り上げ動作のタイミングを解析するような場合には、正規化された位置情報を元に、複数の推定座標を変換する手法が考えられる。

多次元の特徴ベクトルの生成については、解析対象人物の身長・肩幅など個人差及び位置・向き・撮影角度などの条件に依存しない正規化について検討した。動作整合を目的とした場合、主要なベクトルの選択にはケンドールの順位相関係数を用いることとした。また2次元関節位置情報からの3次元骨格推定については、身体対称性や絶対値を拘束条件とした推定手法を用いた。さらに野球のバッティングフォームの解析などでは、映像中に大きさが既知であるベースが含まれている場合があり、そのような場合には視点映像を透視投影変換により絶対値が得ら

れる映像への変換が可能である。短時間では人間の骨長が変化しないという拘束条件を導入することで、3次元骨格推定値の精度を改善することができる。

また、特徴ベクトルの整合区間を求めるには、動的時間伸縮法(Dynamic Time Warping) が有効であるが、始点及び終点を付与する必要がある、それらの決定が整合の自動化への障害となっていた。そこで、始点及び終点を主観で付与することなく動作する動的時間伸縮法の整合アルゴリズムについて検討した[2]。例えば、動的時間伸縮法によりピッチングフォームの異なる二人の投手のフォームを整合させる場合、異なったフォームでも始点・終点が一致し、フォーム比較が容易となる。特徴ベクトルの次元削減には Kendall の順位相関係数を評価値とし、ablation study と呼ばれる構成要素の一部を削除して性能評価を行う手法をとった。また演算量削減については、動作映像のフレームレートを削減する手法を検討した。

4. 研究成果

本研究成果は、(1)動作を表現する多次元特徴ベクトルの生成に関するものと、(2)動的時間伸縮法の改善に関するものに大別される。

4.1 動作を表現する多次元特徴ベクトルについて

4.1.1 骨格座標の欠損部の修復

OpenPose は 2 次元画像から姿勢推定するため、入力画像の空間的解像度あるいは時間的解像度が低い場合にキーポイントの未検出や誤検出が発生する場合がある。OpenPose の出力結果として得られる姿勢推定結果に対して、提案手法により未検出キーポイント、誤検出キーポイントの座標値を補間及び補正した。その結果、実験に用いた 6 本の動画では姿勢推定結果内の未検出フレームが平均 6.30%、誤検出フレームが平均 0.98%減り、姿勢推定精度が向上した。

4.1.2 骨格座標の異常値の修復

OpenPose を映像に適用する際の問題点として、時系列の前後情報を考慮しない点が挙げられる。本研究で使用した一部の映像に対する OpenPose の推定結果には、選手の腕や足が左右逆に検出されているケースが含まれることを確認し、これらを修復する手法を提案した。

4.1.3 骨格座標からの空間ベクトルの生成

時間的に連続するフレームにおける各身体部位の OpenPose による推定結果に対して、骨格座標の異常値の修復を行った後に、骨格情報を利用してフレーム間の姿勢類似度を算出する。そのため、骨格位置情報から空間ベクトルデータを作成する。OpenPose により検出可能な 25 部位のうち、両肩、両肘、両手首、両腰、両膝、両足首の 12 点の推定結果を使用する。ある関節位置から別の関節位置へのベクトルは、12 点から 2 点を選ぶ組み合わせとなり、66 パターン存在する。任意の 2 点を結ぶベクトル r_{vs} と水平方向からの変位を示す角度 θ_{vs} が得られる。これらを図 4 に示す。

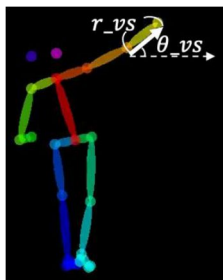


図 4 空間ベクトルの長さや角度

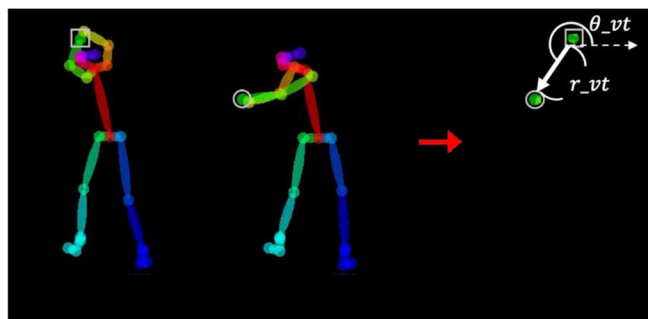


図 5 時間ベクトルの長さ・角度変化の例

4.1.4 骨格座標からの時間ベクトルの生成

時間ベクトルは、あるフレームから数フレーム前と後で、注目する関節位置のベクトル(空間ベクトルの長さや角度)がどれだけ変化したかを要素とする。具体的には 3 フレーム前と後の

r_{vs} と l_{vs} の変化量を要素としている。時間ベクトルの長さ及び角度の例を図5に示す。

4.1.5 動作整合と評価

提案した空間ベクトルと時間ベクトルを用いて、野球の投球動作とゴルフのスイング動作が含まれる映像に対して、動作整合を行い実験によりその評価を行った。その結果、投球動作、スイング動作のいずれにおいても提案手法はより正確な整合を与えることがわかった。

4.1.6 骨格長不変の拘束条件の導入による3次元推定の高精度化

本研究では人間の骨格の絶対的長さが一定で、かつ左右対称であるという条件を付けることで、2次元姿勢推定結果から3次元姿勢を推定する手法を検討した。実験の結果、提案手法では膝や右腕の推定結果がAlphaPoseの姿勢推定結果に比べて正解データの姿勢に近いことが確認できた。

4.1.7 対称性の導入

3次元姿勢推定は、見かけの姿勢でなく実形状を把握できる点で動作解析において優れているが、推定精度が十分でないという問題点がある。動作映像に対する姿勢推定においては、必ずしも連続する映像フレームで骨格長が一定とならないといった問題があった。本研究では、骨格の絶対長がフレーム間で一定であることと、人間の骨格の左右対称性を考慮することで、3次元姿勢推定法の精度向上を図った。

4.2 動的時間伸縮法について

4.2.1 始点終点設定の自動化

従来の始点終点自由な動的時間伸縮法(横井ら)に加えて、4.1.2で提案した骨格座標の異常値の修復手法に加えて、4.1.3及び4.1.4で提案した空間ベクトル及び時間ベクトルを要素として、参照映像と入力映像に対して動作整合を行う実験を実施した。その結果、投球動作およびスイング動作のいずれの場合にも、従来手法と比較して始点・終点の誤差を大幅に削減することができた。

4.2.2 特徴ベクトル削減による処理の拘束化

動的時間伸縮における空間特徴ベクトルの要素は12個の関節の組み合わせからなる66次元のベクトルとなっていた。66個のキーポイントの組み合わせからケンドールの順位相関係数において上位11個の組み合わせを抽出した。実験結果から、ベクトルの各要素単位のケンドール順位相関係数に応じてベクトルサイズを削減することにより、ほぼ性能低下なしに整合処理時間を短縮できることがわかった。

4.2.3 フレームレート削減による処理の高速化

参照・入力動画のフレームを均等に削減することによる動作整合の処理量削減を提案した。従来手法と比較して、2/3削減手法は実行時間を平均9.5%削減し、対応付け性能を平均17.7%向上させる。また、1/2削減手法は実行時間を平均55.2%削減するが、対応付け性能は平均7.4%低下する。実行時間の削減は、従来手法の削減フレーム数が類似フレームの出現頻度に依存するのに対し、提案手法は常に一定数のフレーム削減が可能な点に起因する。また、対応付け性能の結果より、20fpsの均等削減は時系列データの類似性を保つことを示した。

5. 総括

本研究では、人物の姿勢情報を機械学習により推定し、推定された人物の骨格座標を用いて運動解析や動作解析を行い、動作認識や個人認識の識別率の向上を目指した。研究計画時に、研究項目として掲げた、主な6個の研究検討項目はほぼ達成できたと考える。

- 1) 特徴ベクトルの正規化手法と高精度化
- 2) 特徴ベクトルの冗長性除去
- 3) 3次元骨格推定
- 4) 動的時間伸縮法の整合精度からみた特徴ベクトルの最適化
- 5) 特徴ベクトルの次元削減
- 6) 動的時間伸縮法の演算量の削減手法

以上の結果はコロナ禍ではあったが、リモートでの国際会議に4件、国内大会及び研究会に発表できた。

<引用文献>

- [1] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh: "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," IEEE Conference on Computer Vision and Pattern Recognition, pp.7291-7299 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.
- [3] 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合大会, D-12-59, Feb. 2019.
- [4] D. Dwibedi, Y. Aytar, J. Tompson, P. Sermanet and A. Zisserman, "Temporal Cycle-Consistency Learning," IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp.1801-1810, June 2019.
- [5] 大澤遼平, 渡辺裕: "選手の姿勢類似度に基づくスポーツ動作照合", 情報処理学会. 第116

回オーディオビジュアル複合情報処理研究発表会研究報告, No.5, pp.1-3, Feb. 2022.

[6] Hao-Shu Fang, Jiefeng Li, Hongyang Tang, Chao Xu, Haoyi Zhu, Yuliang Xiu, Yong-Lu Li, Cewu Lu: "AlphaPose: Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time", arXiv:2211.03375.

6. 主な発表論文等

〔雑誌論文〕

(なし)

〔国際会議〕(計4件)

[1] Ryohei Osawa, Takaaki Ishikawa, and Hiroshi Watanabe: "Pitching Motion Matching Based on Pose Similarity Using Dynamic Time Warping," IEEE Global Conference on Consumer Electronics (GCCE) 2020, DOI: 10.1109/GCCE50665.2020.9291730, pp.140-144, Sep. 2020.

[2] Atsuya Yamakawa, Takaaki Ishikawa, and Hiroshi Watanabe: "Study on Improvement of Estimation Accuracy in Pose Estimation Model Using Time Series Correlation," IEEE Global Conference on Consumer Electronics (GCCE), DOI: 10.1109/GCCE50665.2020.9291962, pp.570-573, Sep. 2020.

[3] Ryohei Osawa, Shohei Adachi, and Hiroshi Watanabe: "Motion Matching Based on Pose Similarity Using Movement of Body Parts," IEEE Global Conference on Consumer Electronics (GCCE), OS-ADE, DOI: 10.1109/GCCE56475.2022.10014381, pp.83-86, Oct. 2022.

[4] Shohei Adachi, Ryohei Osawa, and Hiroshi Watanabe: "Selective Use of Skeletal Information to Reduce Computational Complexity of Motion Matching," IEEE Global Conference on Consumer Electronics (GCCE), OS-ADE, DOI: 10.1109/GCCE56475.2022.10014090, pp.68-71, Oct. 2022.

〔学会発表〕(計4件)

[5] 大澤遼平, 渡辺裕: "選手の姿勢類似度に基づくスポーツ動作照合", 情報処理学会オーディオビジュアル複合情報処理研究会AVM116, No.5, pp.1-3, Feb. 2022.

[6] 中島聖, 福田大翔, 渡辺裕: "骨格の左右対称性を考慮した3次元姿勢推定法の検討", 情報処理学会第120回オーディオビジュアル複合情報処理研究会, Vol.2023-AVM-120, No.4, pp.1-2 Mar. 2023.

[7] 足立翔平, 渡辺裕: "骨格情報に基づく動作対応付けの処理量削減に向けたフレーム削減手法の検討", 電子情報通信学会総合大会, D-12-42, Mar. 2023.

[8] 福田大翔, 中島聖, 渡辺裕: "骨格絶対長を考慮した三次元姿勢推定法の検討", 電子情報通信学会総合大会, D-12-67, Mar. 2023.

〔図書〕(なし)

〔産業財産権〕

出願状況(なし)

取得状況(なし)

〔その他〕

渡辺研究室論文リスト

<http://www.ams.giti.waseda.ac.jp/publication-new.htm>

7. 研究組織

(1)研究代表者

渡辺 裕 (WATANABE, Hiroshi)

早稲田大学・理工学術院・教授

研究者番号: 10329154

(2)研究分担者

(なし)

(3)連携研究者

(なし)

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 4件 / うち国際共著 0件 / うちオープンアクセス 0件）

| | |
|-----------------------------------------------------------------------------------------|-----------------------|
| 1. 著者名 Ryohei Osawa, Takaaki Ishikawa and Hiroshi Watanabe | 4. 巻 1 |
| 2. 論文標題 Pitching Motion Matching Based on Pose Similarity Using Dynamic Time Warping | 5. 発行年 2020年 |
| 3. 雑誌名 IEEE Global Conference on Consumer Electronics (GCCE) 2020 | 6. 最初と最後の頁 140-144 |
| 掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/GCCE50665.2020.9291730 | 査読の有無 有 |
| オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 | 国際共著 - |

| | |
|---------------------------------------------------------------------------------------------------------------|-----------------------|
| 1. 著者名 Atsuya Yamakawa, Takaaki Ishikawa and Hiroshi Watanabe | 4. 巻 1 |
| 2. 論文標題 Study on Improvement of Estimation Accuracy in Pose Estimation Model Using Time Series Correlation | 5. 発行年 2020年 |
| 3. 雑誌名 IEEE Global Conference on Consumer Electronics (GCCE) 2020 | 6. 最初と最後の頁 409-412 |
| 掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/GCCE50665.2020.9291962 | 査読の有無 有 |
| オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 | 国際共著 - |

| | |
|----------------------------------------------------------------------------------|---------------------|
| 1. 著者名 Ryohei Osawa, Shohei Adachi, and Hiroshi Watanabe | 4. 巻 1 |
| 2. 論文標題 Motion Matching Based on Pose Similarity Using Movement of Body Parts | 5. 発行年 2022年 |
| 3. 雑誌名 IEEE Global Conference on Consumer Electronics (GCCE) 2022 | 6. 最初と最後の頁 83-86 |
| 掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/GCCE56475.2022.10014381 | 査読の有無 有 |
| オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 | 国際共著 - |

| | |
|--------------------------------------------------------------------------------------------------------|---------------------|
| 1. 著者名 Shohei Adachi, Ryohei Osawa, and Hiroshi Watanabe | 4. 巻 1 |
| 2. 論文標題 Selective Use of Skeletal Information to Reduce Computational Complexity of Motion Matching | 5. 発行年 2022年 |
| 3. 雑誌名 IEEE Global Conference on Consumer Electronics (GCCE) 2022 | 6. 最初と最後の頁 68-71 |
| 掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/GCCE56475.2022.10014090 | 査読の有無 有 |
| オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 | 国際共著 - |

〔学会発表〕 計4件（うち招待講演 0件 / うち国際学会 0件）

| |
|--------------------------------------------|
| 1. 発表者名 大澤遼平、渡辺裕 |
| 2. 発表標題 選手の姿勢類似度に基づくスポーツ動作照合 |
| 3. 学会等名 情報処理学会オーディオビジュアル複合情報処理研究会AVM116 |
| 4. 発表年 2022年 |

| |
|--------------------------------------------|
| 1. 発表者名 中島聖, 福田大翔, 渡辺裕 |
| 2. 発表標題 骨格の左右対称性を考慮した3次元姿勢推定法の検討 |
| 3. 学会等名 情報処理学会オーディオビジュアル複合情報処理研究会AVM120 |
| 4. 発表年 2023年 |

| |
|------------------------------------------------|
| 1. 発表者名 足立翔平, 渡辺裕 |
| 2. 発表標題 骨格情報に基づく動作対応付けの処理量削減に向けたフレーム削減手法の検討 |
| 3. 学会等名 電子情報通信学会総合大会, D-12-42 |
| 4. 発表年 2023年 |

| |
|----------------------------------|
| 1. 発表者名 福田大翔, 中島聖, 渡辺裕 |
| 2. 発表標題 骨格絶対長を考慮した三次元姿勢推定法の検討 |
| 3. 学会等名 電子情報通信学会総合大会, D-12-67 |
| 4. 発表年 2023年 |

〔図書〕 計0件

〔産業財産権〕

〔その他〕

ADVANCED MULTIMEDIA SYSTEMS LAB. publication
https://www.ams.giti.waseda.ac.jp/publication/
ADVANCED MULTIMEDIA SYSTEMS LAB. PUBLICATION
https://www.ams.giti.waseda.ac.jp/publication/

6. 研究組織

| | 氏名 (ローマ字氏名) (研究者番号) | 所属研究機関・部局・職 (機関番号) | 備考 |
|--|---------------------------|-----------------------|----|
|--|---------------------------|-----------------------|----|

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

| 共同研究相手国 | 相手方研究機関 |
|---------|---------|
|---------|---------|