

Kinect と OpenPose の統合による姿勢推定性能向上のための関節点の対応関係調査

Joint Point Correspondence Analysis for Integrating Poses from Kinect and OpenPose

森本 隆志[†] 満上 育久[†]
Takashi Morimoto Ikuhisa Mitsugami

1. はじめに

人の全身運動を三次元的に計測するモーションキャプチャシステムは、人の運動の計測や解析などの目的で幅広く利用されているが、多くのシステムでは、特殊なスーツやセンサを装着しなければならないなどの制約がある。また、多くの高額機材を必要とするため、準備の大変さや着用物による違和感などの問題がある。一方、特殊なスーツやセンサを着用する必要のない、Microsoft Kinect (以下、Kinect) や OpenPose を用いる手法も存在する。Kinect では、デプスセンサによって人の三次元姿勢情報が計測できる。また、深層学習ベースの姿勢推定器である OpenPose では、一般的な単眼カメラで撮影した映像に対して姿勢推定が可能である。

両推定器とも、簡単に姿勢推定が行えるものの、問題点もある。Kinect はデプススペースで姿勢推定を行うため、人との相対位置関係によっては自己遮蔽により一部の関節の推定を誤ることがある。また、Kinect はもともとゲーム用のセンサであるため、人と Kinect が向かい合う状況しか想定されていない。そのため、人が Kinect に対して背を向けた場合、人の背面を正面だと認識してしまう。対して OpenPose は、人の向き判定には頑強であるものの、人の姿勢によっては、足などの左右判定を誤ることがある。

このように Kinect による姿勢推定及び OpenPose による姿勢推定には、それぞれに得意・不得意なシーンがあることから、この両方を上手く組み合わせれば、より安定に姿勢を推定できることが期待できる。しかし、図 1 に示す通り、これらが推定する姿勢は、対応する関節がなかったり、関節点の名称は同じでも実際には捉えている位置が微妙に異なることがあり、単純に統合することはできない。そこで本研究では、この問題を解決するため、実際に複数の人物を両姿勢推定機で計測し、両手法における各関節点に対応点として使用可能かを調査するとともに、対応点として用いられない点については、対応点として用いるための関節点補正量を実験的に決定し、対応点として用いることを提案する。評価実験では、複数の人物を Kinect で撮影し 2 種類の姿勢推定器によって推定される姿勢データについて対応の有無の確認および関節点補正パラメータの算出を行い、それが他の人物に関する別の人物の姿勢データについても一貫性を持つことを確認した。

2. 関連研究

Kinect や OpenPose は人の動作分析に関する多くの研究で使用されている[1][2][3][4][5]。しかし、本研究のように、二次元姿勢を用いた三次元姿勢の高精度化を目的とした研究は少ない。Kinect によるデプス画像とカラー画像を併用する研究としては、Nakazawa らによって環境全体のテクスチャ付き三次元形状復元を行う手法が提案されている[6]。



図 1 Kinect と OpenPose の対応関節

3. 関節点の対応関係調査

Kinect と OpenPose による姿勢推定には、それぞれに得意・不得意とするシーンがある。本研究では Kinect のデプス画像から推定される姿勢情報と、カラー画像に OpenFace を適用することで得られる姿勢情報を統合するために、両者の関節点の対応関係を実験的に明らかにする。具体的には、実際に複数の人物を両推定器で計測し、双方における各関節点に対応点として使用可能かどうかを調べる。そのままでは対応点として扱えない点については、実験的に獲得した位置ずれ量をもとに補正を行い、対応点とする。また、正しく補正されたかを確認するため、カメラの内部パラメータを用いて画像平面へ投影し、OpenPose との比較を行う。本手法で使用した関節は図 1 に赤字で示す。以下に詳細を述べる。

適切に Kinect と OpenPose の推定関節の対応付けを行うには、使用する関節の選定だけではなく、必要に応じて補正を行う必要がある。Kinect と OpenPose の関節の推定位置を分析するために 5 人の実験参加者の観測を行った。なお、OpenPose では骨格モデルとして COCO と MPI を選択できるが、本研究では COCO モデルを使用した。

図 1 に示すように Kinect と OpenPose では推定している関節が異なるため、どの関節は対応点として使用でき、どの関節は補正が必要なのかを、図 2 に示す 5 人の実験参加者の姿勢推定結果から確認した。図 2 より、肘、手首関節に関しては両推定機共にほぼ同様の位置を捉えていることが確認できる。それ以外の関節に関しては、まず首に関して、両推定器ともに Neck という関節を推定しているが、図 2 よりそれらは異なる位置を捉えていることが分かる。そこで、OpenPose の Neck とほぼ同位置を捉えている Kinect の SpineShoulder を対応点とする。また、名称からは同一の関節点を捉えていると思われるものの、実際にはわずかに異なる位置が推定されている関節として、肩、腰、膝、足首がある(図 2)。まず、肩の位置に関して、OpenPose は首

と肩をほぼ同じ高さに推定しているのに対し、Kinect は肩をわずかに低い位置に推定する傾向がある。また、腰について、OpenPose のほうが Kinect よりも体の外側寄りに腰関節を推定する傾向がある。膝と足首については、OpenPose よりも Kinect のほうが高い位置に推定する傾向があった。これらの関節については、実験的に獲得した位置ずれ量をもとに、Kinect の姿勢データが OpenPose と対応するように補正を行う必要がある。

Kinect の肩関節については首関節と同程度高さに、腰関節については体の外側方向に、膝と足首関節については下方向に補正を行う。Kinect の各関節の位置ベクトルを p_i 、補正後の位置ベクトルを q_i 、補正パラメータを k としたとき、肩、腰、膝、足首関節の補正処理は以下の式で表せる。

$$q_{ShoulderR} = p_{ShoulderR} + k_1(p_{ShoulderR} - p_{ElbowR})$$

$$q_{ShoulderL} = p_{ShoulderL} + k_1(p_{ShoulderL} - p_{ElbowL})$$

$$q_{HipR} = p_{HipR} + k_2(p_{HipR} - p_{HipL})$$

$$q_{HipL} = p_{HipL} + k_2(p_{HipL} - p_{HipR})$$

$$q_{KneeR} = p_{KneeR} + k_3(p_{KneeR} - p_{HipR})$$

$$q_{KneeL} = p_{KneeL} + k_3(p_{KneeL} - p_{HipL})$$

$$q_{AnkleR} = p_{AnkleR} + k_4(p_{AnkleR} - p_{KneeR})$$

$$q_{AnkleL} = p_{AnkleL} + k_4(p_{AnkleL} - p_{KneeL})$$

各式中の補正パラメータ k の値は実験参加者の観測から決定しており、肩と腰関節の補正時は $k_1 = k_2 = 0.15$ 、膝と足首関節の補正時は $k_3 = k_4 = 0.2$ とした。

最終的に、補正した Kinect の三次元姿勢情報を、カメラの内部パラメータを用いて画像平面へ投影し、OpenPose との比較を行う。

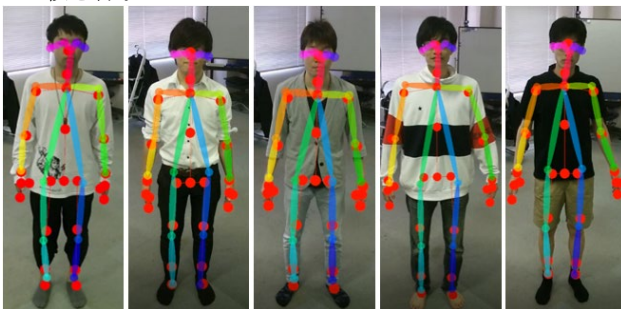


図2 Kinect と OpenPose による姿勢推定の比較

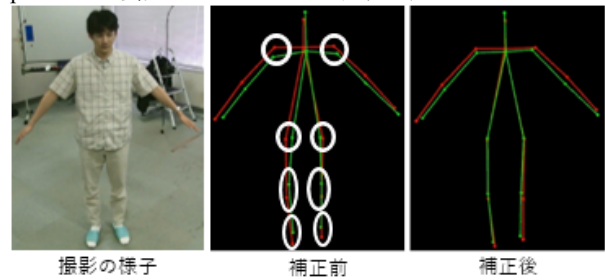
4. 評価

補正パラメータ算出時とは異なる実験参加者の撮影を行い、提案手法により関節の補正を行った結果を図 3 に示す。図 3 中に白丸で囲んでいる関節について補正後の様子を見ると、補正前より一致していることが確認できた。

5. Kinect と OpenPose の統合例

本研究により、Kinect と OpenPose で得られる姿勢の対応付けが可能となる。ここでは、その有用性を確認するための一例を示す。図 4 に人物を前方および側方から Kinect で撮影した際の 2 つの姿勢データを示す。なお、姿勢データの左右の認識が識別できるよう、Kinect による姿勢データについては右足を赤色、左足を青色で表示し、OpenPose による姿勢データは右足が緑色、左足を青色で表示している。(a)では、いずれの姿勢データも人の左右が正しく推定されているのに対して、(b)では OpenPose が足の左右の認識を誤っていることが分かる。Kinect による姿勢データは歩行情報に基づいているため、(b)の状況でも足の左右が正し

く推定されている。この(b)のようなケースでは、Kinect と OpenPose の姿勢データの統合が有効に働くと考えられる。



※ 緑スケルトン：Kinect，赤スケルトン：OpenPose

図3 関節の補正結果

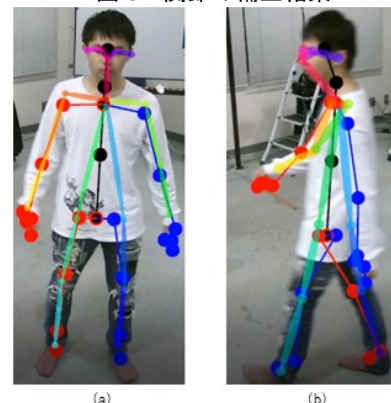


図4 OpenPose による姿勢推定の誤り

6. おわりに

本研究では Kinect のデプス画像から推定される姿勢情報と、カラー画像に OpenPose を適用することによって得られる姿勢情報を統合するために、両者の関節点の対応関係を明らかにした。今後は複数台の Kinect でモーションキャプチャを行い、複数の三次元姿勢情報と二次元姿勢情報を用いることで、人物三次元姿勢の高精度化を考えている。

謝辞

本研究は JSPS 研究費 JP18H03312 の助成を受けたものです。

参考文献

- [1] 紅林秀治, 小林健太, 兼宗進, 「KINECT センサーを用いた簡易動作分析システムの開発」, 情報処理学会研究報告, Vol.2012-CE-118, No.20 (2013).
- [2] J. Murakami, T. Morimoto, I. Mitsugami, "Gaze and Body Capture System under VR Experiences," ACM Symposium on Virtual Reality Software and Technology, No.91 (2018).
- [3] 森本隆志, 満上育久, 「複数 Kinect の姿勢データの時空間統合によるモーションキャプチャシステム」, 情報処理学会研究報告, Vol.2019-CVIM-217, No.25 (2019).
- [4] H. Fani, A. Mirlohi, H. Hosseini, R. Herpers, "SWIM STROKE ANALYTIC: FRONT CRAWL PULLING POSE CLASSIFICATION," 2018 25th IEEE International Conference on Image Processing (2018).
- [5] 榊崎翔太, 田中康, 水上嘉樹, 「歩行動作に基づく人物識別に関する研究」, 研究報告ヒューマンコンピュータインタラクション, Vol.2018-HCI-177, No.31 (2018).
- [6] M. Nakazawa, I. Mitsugami, H. Habe, H. Yamazoe, Y. yagi, "Calibration of multiple Kinects with Little Overlap Regions," IEEJ Transactions on Electrical and Electronic Engineering, Vol.10, No.S1, pp.S108-S115 (2015).