

骨格推定および呼吸検知を用いたヨガのポーズ評価手法の提案

宗像 修平† 村木 祐太† 小堀 研一†

大阪工業大学 情報科学研究科 情報科学専攻†

1. はじめに

近年、気軽に行える運動としてヨガを習う人口が増加している。特に、自宅でも楽しめるヨガ教室として、オンライン会議アプリを利用したオンラインヨガが注目されている。しかし、ヨガ教室への参加は、時間に制限がかかってしまう問題点や、人数が多いためインストラクターから直接指導を受けることが困難であるという問題点が存在する。また、スポーツにおける姿勢評価の従来研究^[1]では、Kinectなどのモーション認識デバイスにより姿勢の評価を行う手法を提案していた。しかし、ヨガでは柔軟性の違いなどから、ポーズの類似性よりも適切なタイミングでの呼吸が重要とされている。そのため、従来研究と同様に姿勢を評価するだけでは、ヨガのポーズ評価としては不十分となってしまう。そこで本研究では、骨格推定と呼吸検知を活用し、ヨガのポーズを総合的に評価する手法を提案する。提案手法では、骨格推定結果に対し、前後フレームを活用した座標補間を行い、正確な三次元骨格を生成する。また、ポーズ完成時の呼吸回数と時間を計測することで、ヨガ教室における指導に即した評価を行う。

2. 提案手法

提案手法の処理手順を図1に示す。同図左側は動画、右側は音声に対する処理である。

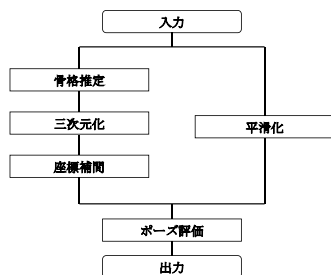


図1 提案手法の処理手順

提案手法では、利用者の正面と側面の2方向から撮影した動画とヘッドセットなどにより録音した呼吸音を入力とする。入力動画に対して、OpenPoseを用いた骨格推定により各視点での骨格情報を取得する。三次元化では、側面からの入力を奥行き情報とし、三次元座標を求め、座標補間では、OpenPoseにより取得できなかった関節や誤った座標値の関節に対し、前後のフレームを利用し補間を行う。入力音声に対して平滑化を行うことで、呼吸における吸・吐きを明確化する。最後に、ポーズ評価では、求めた三次元骨格からポーズの完成と終了を検出し、その間における呼吸時間と回数を評価する。

2.1 骨格推定

骨格推定では、人の骨格情報を取得できるライブラリであるOpenPoseを用いる。OpenPoseでは、人の肘や膝といった関節点を取得することが可能である。OpenPoseを用いて、各視点の入力動画に対し1フレーム毎に骨格推定を行い、関節の座標値を取得する。入力動画に対して骨格推定を行った様子を図2に示す。

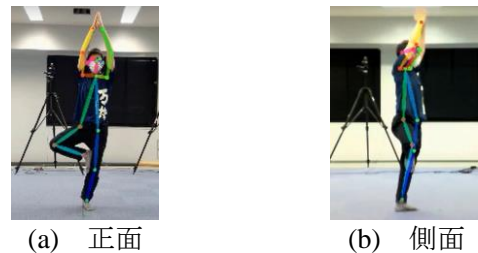


図2 骨格推定

2.2 三次元化

取得した2視点の座標値を用いて、二次元情報のみであるOpenPoseの骨格情報の三次元座標を取得する。正面からの入力動画を基準とし、側面からの入力動画における各関節のX座標を奥行き情報として保存する。骨格情報を三次元化した結果を図3に示す。

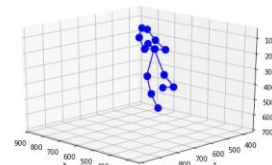


図3 三次元出力

2.3 座標補間

座標補間では、山川らの手法^[2]をもとにOpenPoseにより正しく取得できなかった座標値に対して補間を行う。提案手法では、ポーズの種類により関節の座標値が大きく変化する場合があります。そのため、山川らの手法と異なり、前後各2フレームずつと比較を行う。そして、座標値の差が閾値以上の関節に対して補間を行う。また、補間に用いる値は前後のフレームで座標値が同じである場合、その値と同じ値で補間する。座標値が異なる場合、前後各2フレーム、計4フレームの中央値を用いて補間する。

また、座標補間後に全フレームに対して直前の2フレームを使用したローパスフィルタを用いることで、補間対象とならない骨格推定時に発生するノイズを除去する。

2.4 平滑化

入力音声に対して Savitzky-Golay filter を用い

Proposal of Yoga Pose Evaluation Method Using Skeletal Estimation and Breathing Detection

Shuhei Munakata† Yuta Muraki† Kenichi Kobori†

Osaka Institute of Technology Graduate School of Information Science and Technology†

て平滑化を行う。平滑化により、不安定な形状の音声波形を滑らかにすることで、呼吸における吸う・吐く音量を明確化する。平滑化した結果を図4に示す。同図において青色波形は平滑化前であり、赤色波形は平滑化後である。また、緑色の直線は音量の平均値である。

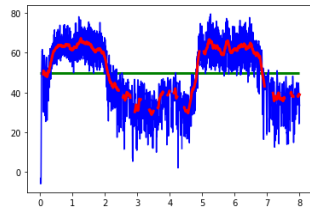


図4 波形の平滑化

2.5 ポーズ評価

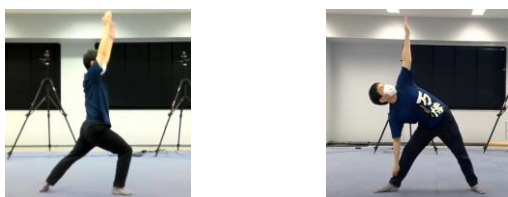
提案手法では、取得した三次元骨格と平滑化した入力音声をもとに、ポーズを行っている際の呼吸の時間と回数を評価する。時間と回数の評価は、ポーズ完成後から終了までの間で行う。この際のポーズ完成は、各フレームにおいて、最も大きい変化量の座標値を全関節点から1点ずつ算出する。そして、その変化量が閾値以下の状態が4秒続いた時、ポーズが完成したと判断する。ポーズ終了は、ポーズ完成と同様に変化量の値を算出し、閾値以上の変化が4秒続いた時、ポーズが終了したと判断する。

呼吸の時間は、入力音声の音量をもとに、音量の平均より大きい値が続いている時を「吸い」の時間として計測する。そして、「吸い」と次の「吸い」の間を「吐き」として計測する。それぞれ時間を計測し、平均を求めることで「吸い」と「吐き」の時間が適切であるか評価を行う。また、同時に「吸い」と「吐き」を1セットとして、回数も評価を行う。

3. 実験と考察

3.1 実験

提案手法の有効性を確認するために、図5に示す2種類のポーズに対して実験を行った。また、表1にそれぞれのポーズにおける正しい呼吸時間と回数、実験での呼吸時間と回数を示す。



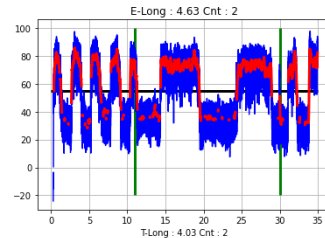
(a) ポーズ1 (b) ポーズ2

図5 対象ポーズ

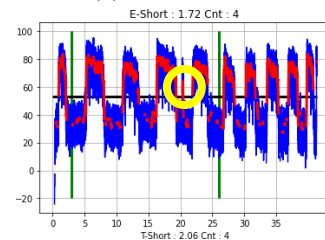
表1 ポーズ詳細

	正しい呼吸	実験での呼吸
ポーズ1	吸い: 3秒	吸い: 約5秒
	吐き: 3秒	吐き: 約5秒
	回数: 3回	回数: 2回
ポーズ2	吸い: 3秒	吸い: 約2秒
	吐き: 3秒	吐き: 約2秒
	回数: 3回	回数: 4回

実験結果を図6に示す。同図において横軸は時間、縦軸は音量である。また、青色波形は平滑化前であり、赤色波形は平滑化後、緑色の直線はポーズの完成と終了を示す。さらに、「E」は「吐き」、「T」は「吸い」であり、それぞれ Long, Short, Good で評価を提示している。



(a) ポーズ1



(b) ポーズ2

図6 実験結果

3.2 考察

図6(a)と表1より、ポーズ1に対して正しく呼吸の回数と時間を評価できている。また、同図(b)と同表より、ポーズ2に対して正しく呼吸の回数と時間を評価できている。しかし、同図(b)に示す黄枠内のように平滑化により取り除けないノイズが存在する。今回、「吐き」と「吸い」の基準として、開始と終了の間隔が1秒未満の時は呼吸ではないとした。そのため、小さなノイズの影響を受けなかったと考える。しかし、1秒に満たない短い呼吸も存在すると考えられるため、音量以外での呼吸回数の計測方法の検討も必要であると考えられる。

4. おわりに

本研究では、骨格情報と呼吸検知を用いたヨガポーズの評価手法を提案した。提案手法では、骨格推定と座標補間により骨格情報の三次元出力を行った。これにより、呼吸検知と組み合わせたヨガインストラクターに近い評価が可能となった。

参考文献

- [1] 星野直紀, et al. “Kinectを用いた弓道訓練システムの提案”, 情報処理学会第78回全国大会論文集, Vol.78, No.4, pp.239-240(2016).
- [2] 山川敦也, et al. “時系列相関性を用いた姿勢推定モデルの骨格推定精度改善”, 情報処理学会第82回全国大会講演論文集, Vol.82, No.1, pp.249-250(2020).