

## OpenPose を用いた複数人のダンスの一致度評価

## Coincidence Evaluation of Multiple People's Dance Using OpenPose

稲田 健太郎<sup>†</sup> 石川 孝明<sup>‡</sup> 渡辺 裕<sup>†</sup>Kentaro INADA<sup>†</sup> Takaaki ISHIKAWA<sup>‡</sup> Hiroshi WATANABE<sup>†</sup><sup>†</sup>早稲田大学基幹理工学部情報通信学科 <sup>‡</sup>早稲田大学国際情報通信センター<sup>†</sup>Department of Communications and Computer Engineering, Waseda University<sup>‡</sup>Global Information and Telecommunication Institute, Waseda University

## 1. はじめに

モーションキャプチャやセンサを用いた動作解析では、複数台のカメラやマーカ・センサの準備が必要になる。複数台のカメラやセンサの前でしか動作解析ができない欠点がある。しかし、OpenPose[1] はすでに撮影された動画で解析が行えるため、多人数の同時動作解析に適用できる。また、2012 年度から選択であった中学校での体育授業のダンス科目が必修化された。ダンスがより身近なものとなり、複数人で踊る様子を撮影し公開する人も少なくない。そこで、特別な機具を持たない人が練習で利用できるダンスの評価指標が望まれている。

本研究では、OpenPose の入力として複数人が同様のダンスをしている定点カメラ動画を利用する。ダンスの動作に対して、開始と終了のタイミング一致度の評価をする手法を提案する。

## 2. OpenPose

OpenPose は、単一画像から複数の人間の体や顔のキーポイントをリアルタイムに検出することができる機械学習型動作解析処理である。Convolution Neural Network を用いて、画像に映る人物の肩・肘・目など 18 点の位置推定を行う。動画を入力すると、全フレームに対して人物の姿勢推定を行う。各キーポイントにマークを付け、マーク間を線で結び人体モデルを表示する。また、フレームごとに映るすべて

Coincidence evaluation of multiple people's dance using OpenPose

<sup>†</sup>Kentaro INADA, Department of Communications and Computer Engineering, Waseda University

<sup>†</sup>Hiroshi WATANABE, Department of Communications and Computer Engineering, Waseda University

<sup>‡</sup>Takaaki ISHIKAWA, Global Information and Telecommunication Institute, Waseda University

の人物の各点座標のデータセットを同一ファイルで取得できる。

しかし、人物などが重なり身体の一部が隠れていると誤検出が多くなるため、姿勢推定対象が重なる動画像では得られるデータの信頼度が低下する OpenPose により得られる出力画像の例を図 1 に示す。



図 1 OpenPose による出力画像の例[1]

## 3. 提案手法

本研究では、OpenPose により得られた座標データのうち、首・両肩・両肘・両手首・両尻・両膝・両足首の 13 点を処理対象とする。まず、フレームごとに得られた複数人の座標データセットのうち、1 フレーム前の各人物の首の x 座標と比較して、差が最も小さい座標データセットを同一人物のデータセットとする。

座標の増減から動き出しのタイミングを判断するためピークを求める。フレームに対する座標の変化の波形の中で、極大値と極小値を求める。13 点のキーポイントの x 座標と y 座標の合計 26 個の要素に対して行う。ピーク間が 0.1 秒以上ない場合は、故意的な動きではないとして無視する。また、ピーク間の座標の差が小さいものもダンスの動きではないとして無視する。

あるピークに対して、直前のピークのフレームをピーク開始点とする。

動きの開始と終了のタイミングのずれの許容時間を 0.1 秒と定義する。動画のフレームレートにより 0.1 秒間のフレーム数が異なるが、フレームレートによりずれの許容フレーム数を変化させる。動画内の人物のうちの過半数が 0.1 秒以内に同じピークとピーク開始点があるとき、その動きを正解とする。ピークが 0.1 秒以上ずれている場合と同じピークがある人物が過半数に満たない場合を不正解とする。各座標での各人物の正解数から不正解数を引いたものを点数として一致度を評価する。

#### 4. 実験

本研究では、5 人のアイドルグループがダンスする動画のうち、5 人は移動せず同じ順番で並び同じダンスをする約 9 秒間を使用した。人物を左から A, B, C, D, E とする。この動画の第 1 フレームの人体モデルを図 1 に示す。この動画は 30fps であるため、ずれの許容フレーム数は 3 とした。3 人以上のピークとピーク開始点が 3 フレーム以内にあるときを正解のタイミングとした。

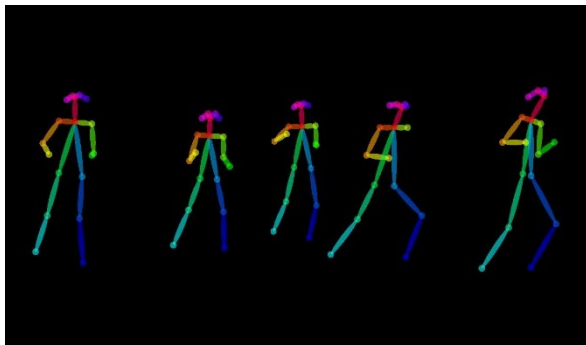


図 2 使用動画の第一フレーム

#### 5. 結果

動画内の 5 人の 13 のキーポイントの正解数から不正解数を引いた点数を表 1 に示す。肩・肘・手首の合計を腕部分と、尻・膝・足首の合計を脚部分と定義して、これらの豪渓点数と全キーポイントの合計点数を表 2 に示す。

表 1 より、各人物に対して動きが周りと一致している身体の部分と一致していない身体の部分が数値として得られた。表 2 より、各点だけでなく腕や脚の全体の動きのタイミングの評価できる。また、合計の点数から B が最もタイミ

ングが一致していることがわかる。

表 1 各キーポイントの点数

	A	B	C	D	E
首	18	25	21	18	12
右肩	1	17	10	19	14
右肘	4	17	-11	4	-4
右手首	-1	24	-5	2	1
左肩	10	28	13	15	10
左肘	1	8	-18	-1	3
左手首	11	10	-1	1	-17
右尻	8	11	23	9	21
右膝	7	3	-5	8	-5
右足首	9	-8	2	-10	-11
左尻	9	15	21	22	8
左膝	9	-19	-6	4	10
左足首	-13	-8	-10	-10	-15

表 2 腕・脚部分と全身の点数

	A	B	C	D	E
右腕部分	4	58	-6	25	11
左腕部分	22	46	-6	15	-4
右脚部分	24	6	20	7	5
左脚部分	5	-12	5	16	3
全身	55	161	46	81	27

#### 6. まとめ

複数人が同じダンスをする定点カメラ動画において OpenPose を利用した動き出しのタイミングの一致度を評価した。しかし、よりダンスを評価するためには動きのタイミングだけでなく動きの大きさからの評価が必要である。

#### 参考文献

- [1] Z.Cao, T.Simon, S-E Wei, Y.Sheikh, "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields", <https://arxiv.org/abs/1611.08050>, 2016.