

OpenPose を用いた複数ダンサーの姿勢評価の基礎検討

稲田 健太郎[†] 石川 孝明[‡] 渡辺 裕^{†‡}

[†]早稲田大学大学院 基幹理工学研究科 基幹理工学研究科 情報理工・情報通信専攻
〒169-0072 東京都新宿区大久保 3-14-9 早大シルマンホール 401

[‡]早稲田大学国際情報通信研究センター 〒169-8555 東京都新宿区大久保 3-4-1

E-mail: [†]{inaken0110@fuji.waseda.jp, hiroshi.watanabe@waseda.jp} [‡]takaxp@ieee.org

あらまし 近年、アイドル等の踊る簡単ダンスが流行となり、動画サイトにダンス動画をアップロードする人が増えている。また、中学校の体育授業において、ダンス科目が必修化された。このように、ダンスを実際に踊り、他人から評価される場面が増えている。複数人のダンス動画から各人物の動作のずれ評価することにより、練習時の一つの指標になる。そこで、画像内の複数人の姿勢推定を行える OpenPose を用いて、ダンサーの動作評価の手法を検討している。本稿では映像内のダンサーが同様のダンスをしている動画を対象とし、映像内のダンサーから基準モデルを生成することによって姿勢のずれを評価する手法を提案する。

キーワード OpenPose, 角度評価, 基準モデル

Basic study of Evaluation of body shape for Multiple Dancers Using OpenPose

Kentaro INADA[†] Takaaki ISHIKAWA[‡] and Hiroshi WATANABE^{†‡}

[†]Graduate School of Fundamental Science and Engineering, Waseda University
#401 Sodai Shillman hall Building 3-14-9 Okubo, Shinjuku-ku, Tokyo, 169-0072 Japan

[‡]Global Information and Telecommunication Institute, Waseda University
3-4-1 Okubo, Shinjuku-ku, Tokyo, 169-8555 Japan

E-mail: [†]{inaken0110@fuji.waseda.jp, hiroshi.watanabe@waseda.jp} [‡]takaxp@ieee.org

Abstract Recently, dance of idol group or singer is popular. Some people upload dance videos to video sites. Also dance course was compulsory in junior high school physical education lesson. Thus, it is useful for training dance, evaluating the action synchronization and body shape of each dancer from videos. We consider the method of evaluation of body shape for multiple dancers using OpenPose that can estimate pose of multiple people in the image. In this paper, we propose a method to evaluate the misalignment of a figure by generating a reference model from a dancer in the image, for a moving image in which dancers are performing same dance.

Keyword OpenPose, Angle Evaluation, Reference model

1. はじめに

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。また、YouTube などの映像サイトにダンス映像をアップロードする人も増えている。その映像の視聴者は、ダンスのキレや可愛さなどから評価している。このように、ダンスというものが、今まではテレビ等でパフォーマーのダンスを見るものから、自分たちが実際に踊るものや他人から評価されるものに変化している。複数人のダンス映像を元に各ダンサーの採点をすることができれば、ダンスの練習時に一つの指標となり得る。

ダンスの評価として、音楽と振り付けの同期評価と振り付けの身体の形状評価がある。本稿では、身体の形状評価に注目し、複数人が同様のダンスをしている映像を対象に OpenPose を用いて姿勢推定を行い、得られる姿勢情報を

用いて複数ダンサーを評価する手法について検討する。

2. 姿勢推定

本稿では Zhe らが提案した OpenPose[1]を用いて複数ダンサーの姿勢推定を行う。OpenPose は、単一画像から複数の人間の体や顔の部位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、画像に映るダンサーの肩・肘・目など 18 点の位置推定を行う。18 点の画像上の座標情報を取得できる。

3. 提案手法

提案手法の概要を図 1 に示す。OpenPose によって得られた座標を用いて、各関節の角度を算出する。そ

それぞれの角度に対して全ダンサーの基準となるモデルを作成する。

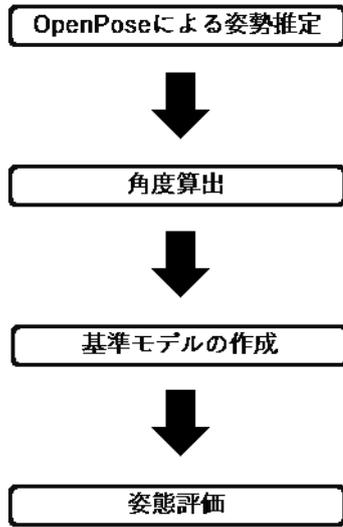


図1 提案手法

3.1. 角度算出

OpenPoseによって取得できる18点の関節のうち、鼻・首・両肩・両肘・両手首・両腰・両膝・両足首の14点を使用する。これらの点から、表1に示す3個の関節A, B, Cを結ぶことによってできる $\angle ABC$ を算出する。本稿では、映像の各フレームおよび各ダンサーの合計12個の角度によって、身体の形状評価をする。

表1 算出する角度

	関節 A	関節 B	関節 C
角度 1	鼻	首	右肩
角度 2	首	右肩	右肘
角度 3	右肩	右肘	右手首
角度 4	鼻	首	左肩
角度 5	首	左肩	左肘
角度 6	左肩	左肘	左手首
角度 7	鼻	首	右腰
角度 8	首	右腰	右膝
角度 9	右腰	右膝	右足首
角度 10	鼻	首	左腰
角度 11	首	左腰	左膝
角度 12	左腰	左膝	左足首

3.2. 基準モデル作成

映像内に映る全ダンサーから、各ダンサーの形状評価の基準となるモデルを作成する。基準モデルの各角度は、映像内の中央値とする。しかし、 $-\pi$ と π は等しくなるように、角度は単純には中央値を算出することはできない。そのため、各角度において全ダンサーの最小値を $0[\text{rad}]$ に移動し、中央値を計算する。基準モデルのフレーム t における表1に示

した角度 k は次の式(1)および式(2)によって与えられる。

$$\theta'_{i,t,k} = \theta_{i,t,k} - \min_{i \in I} \theta_{i,t,k} \quad (-\pi < \theta'_{i,t,k} < \pi) \quad (1)$$

$$\theta_{model,t,k} = \bar{\theta}_{model,t,k} \quad (2)$$

ここで、 I は映像内のダンサー人数、 T は全フレーム数、 $\theta_{i,t,k}$ はダンサー i のフレーム t における角度 k を表す。

3.3. 形状評価

ダンサー i のフレーム t における角度 k の基準モデルとの形状のずれを評価する値 $T_{i,t,k}$ を式(3)のように求める。この $T_{i,t,k}$ は、0から1の範囲の値であり、1に近いほどダンサー i のフレーム t における角度 k は基準モデルと等しい形状をしていることを意味する。

$$T_{i,t,k} = 1 - \frac{|\theta_{model,t,k} - \theta_{i,t,k}|}{\pi} \quad (3)$$

4. 実験

5人が同様のダンスをしている9秒間(24[fps], 228[frame])映像を用いて実験を行った。第1フレームにおけるOpenPoseによる姿勢推定結果のスケルトン画像を図2に示す。ダンサーを左から1, 2, 3, 4, 5とする。また、作成した基準モデルの第1フレームを図3に示す。



図2 使用動画の第1フレーム

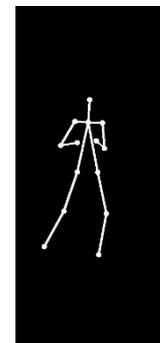


図3 第1フレームの基準モデル

基準モデルが、映像内のダンサー全員の平均的な身体の形状をしていると確認できる。しかし、基準モデルの形状が急激に変化するときがある。

各ダンサーの各角度の評価値 $T_{i,t,k}$ の全フレームでの平均値を図4に示す. また, 評価値 $T_{i,t,k}$ の平均値を表2に示す.

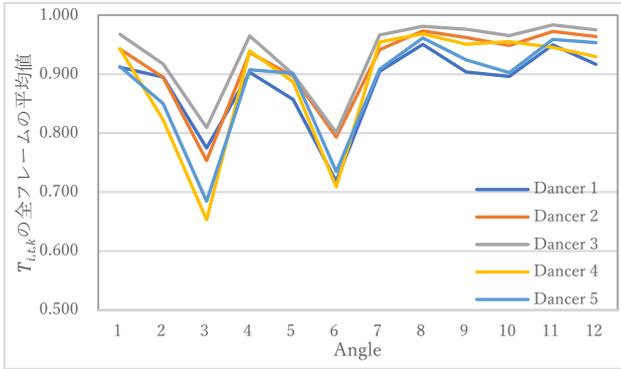


図4 $T_{i,t,k}$ の全フレームの平均値

表2 評価値 $T_{i,t,k}$ の平均値

		Dancer				
		1	2	3	4	5
Angle	1	0.912	0.943	0.968	0.943	0.913
	2	0.895	0.894	0.917	0.823	0.850
	3	0.775	0.753	0.810	0.653	0.685
	4	0.903	0.937	0.965	0.939	0.907
	5	0.857	0.896	0.900	0.887	0.902
	6	0.719	0.793	0.801	0.709	0.735
	7	0.905	0.941	0.966	0.955	0.909
	8	0.950	0.973	0.981	0.968	0.961
	9	0.904	0.962	0.976	0.951	0.924
	10	0.896	0.949	0.965	0.955	0.903
	11	0.950	0.973	0.983	0.945	0.959
	12	0.917	0.964	0.975	0.930	0.953

各ダンサーの基準モデルとの角度差が確認できた. 各ダンサーのずれがある関節がわかり, ダンスの指標となる. しかし, 関節の可動域が異なるため, 式(3)の第2項の分母に係数を付加し, その係数によって重み付けを行う必要がある.

5. まとめ

本研究では, ダンサー全員が同様のダンスしている映像を対象とした, ダンサーの身体の形状のずれを評価する手法を提案した.

課題として, 基準モデルの動きに連続性がないことがあげられる. 今後は, 各関節の角度に閾値を設けることで基準モデルの精度改善を行う. 各関節の可動域の違いによる重み付け係数も用いる必要がある. また, 身体の形状だけでなく

音楽との同期を評価する手法も考えていく.

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