

姿勢推定を用いたダンス教育を支援する評価システムの開発

Development of an evaluation system to support dance education using pose estimation

松本 悠輝^{*1}, 安留 誠吾^{*2}
Yuki MATSUMOTO^{*1}, Seigo YASUTOME^{*2}

^{*1} 大阪工業大学情報科学研究科

^{*1} Graduate School of Information Science, Osaka Institute of Technology

^{*2} 大阪工業大学

^{*2} Osaka Institute of Technology

Email: m1m02530160@oit.ac.jp

あらまし: 学習指導要領改訂により、中学校の体育授業では2012年にダンスが必修となり、適切に評価するために、公平かつ客観的な評価手法の必要性が高まっている。本研究では、姿勢推定技術を用いたダンス動作の評価システムを構築した。お手本動画とカメラで撮影した映像に対して、骨格情報を取得・比較することで、動作の類似度を測定する。さらに、主観的評価との比較実験を通じて、本システムの有効性を検証した。

キーワード: ダンス教育, 姿勢推定, YOLOv8, コサイン類似度, DTW 距離

1. はじめに

学習指導要領⁽¹⁾の改訂により、2012年から中学校の体育授業においてダンスが必修化された。また、各教科において、「知識・技能」、「思考・判断・表現」、「主体的に学習に取り組む態度」の3観点に基づいて評価される。その中で、体育授業におけるダンスでは、「技能」の評価基準において数値的指標がないため、評価の難しさが浮き彫りとなっている。

具体的には、「リズム感」や「変化のある動き」など評価基準が抽象的であることから、教員の主観に基づく採点となり、評価の公正性や客観性が確保されにくいという課題がある。加えて、教員側のダンスに対する専門性のばらつきも、評価の一貫性を損なう要因となっている。このような背景から、本研究では、姿勢推定技術と動作類似度の数値化手法を活用し、生徒のダンス動作を客観的に評価するシステムの構築を目的とする。

2. 関連研究

姿勢推定技術の教育分野への応用は近年注目されており、とりわけOpenPoseやMediaPipeなどを用いた研究が多く見られる。

稲田の研究⁽²⁾では、複数名が同時に踊るダンス映像を対象に、個々のダンサーの動作のタイミングと一致度を定量的に評価するシステムを提案している。評価には、OpenPoseを用いて各ダンサーの骨格情報(キーポイント座標)を抽出し、さらに、それらのデータを用いて動作時系列の類似性を計算している。特徴的なのは、全体の代表的な時系列を基準データとして導出し、各ダンサーとの一致度をこの基礎データと比較する形で算出している点である。

また、動作変化点の検出や、平均類似度の時間的推移の可視化も行っており、これにより演技全体に

わたる協調性や一貫性の評価が可能となっている。この手法は、個々の動作だけでなく、グループ全体の動きの揃い具合やタイミングのズレを客観的に可視化できるという点で、教育や練習支援における活用が期待されている。

3. 研究概要

本研究では、事前に用意したお手本動画と生徒のダンス映像から抽出したキーポイント座標をもとに比較し、コサイン類似度とDTW距離を算出する。

3.1 提案手法

提案手法では、姿勢推定用モデルとしてYOLOv8-pose⁽³⁾を採用し、人物映像から17箇所のキーポイント座標を取得している。これらの情報を用いて、隣接する関節間のベクトルを算出し、正規化を行い、単位ベクトルとした。これにより、個体差(身長やカメラ距離)を排除した動作の「方向性」のみを評価対象としている。取得した単位ベクトルをもとに2種類の評価指標「コサイン類似度」、「DTW距離」を算出し、客観的評価を行う。

3.2 評価指標

コサイン類似度

コサイン類似度は、同一フレームにおけるお手本と生徒の姿勢をベクトルとして比較し、そのなす角のコサイン値を算出することで、動作の「方向の一致度」を評価する指標である。本研究では、図1に示す単位ベクトルの内、両肩-肘、両肘-手首、両腰-膝、両膝-足首間の合計8つについて、それぞれのコサイン類似度を求め、その平均値を1フレームごとのスコアとした。値は-1~1の範囲で変動し、1に近いほど瞬間的な動作方向が一致しているとみなされる。

DTW 距離

DTW (Dynamic Time Warping) 距離は、時間的な

伸縮を許容しながら2つの時系列データの類似度を測るためのDTWアルゴリズム⁽⁴⁾によって得られる数値である。本研究では、図1に示すように関節間の全単位ベクトルに対し、文献(2)の稲田も使用したSoft-DTWアルゴリズム⁽⁵⁾により、得られた数値をDTW距離とした。Soft-DTWアルゴリズムは、従来のDTWアルゴリズムと同様に動作テンポや速度の違いを考慮しつつ、微分可能かつ滑らかなスコアを提供することで、安定性・柔軟性に優れた時系列比較が可能となる手法である。DTW距離は、数値が小さいほどお手本に近いと判断される。

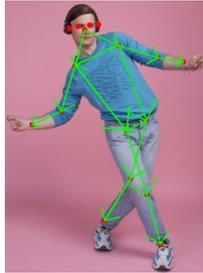


図1 キーポイント間の単位ベクトル

4. 実験・結果

本実験では、提案システムの有効性を検証するため、15秒ほどの簡単な振り付けのダンス動画を用い、被験者1名に3種類のテストケース(1:お手本とほぼ同一, 2:リズムずれ, 3:異なる振り付け)をそれぞれ3回ずつ実施し、計9回分のデータを収集した。評価には、①コサイン類似度(平均値)、②DTW距離、③主観的評価(開発者および被験者自身による10段階評価)を用いた。ここで、②コサイン類似度(平均値)とは、生徒の動作における全フレームに対するコサイン類似度の平均値をとったものである。

実験結果は表1である。DTW距離は主観評価との相関が高く、特に動作全体の再現性に関して有効であることが確認された。一方、コサイン類似度は動作の「一瞬の一致度」を高く評価する傾向があり、動きの流れよりも静止画的な一致に敏感であることが示唆された。また、主観評価と数値評価の乖離が見られたケースでは、振付自体が似ているもののタイミングが大きくずれているパターンが多く、評価の基準に応じて指標を使い分ける必要があることが明らかとなった。

表1 実験結果

テストケース	動作の内容	主観的評価	コサイン類似度(平均値)	DTW距離
1	お手本とほぼ同じ	8.5	0.784	239.482
1		8.5	0.751	280.511
1		7	0.732	351.448
2	踊り出し遅い ペースが早い ペースが遅い	6.5	0.657	291.538
2		4.5	0.618	297.761
2		4	0.769	331.119
3	異なるダンス 歩き回る 直立状態	1	0.691	668.969
3		1	0.576	1278.136
3		1	0.569	1288.245

5. 今後の課題・まとめ

本研究では、被験者1名・お手本動画1種類という限定的な条件下での検証を行ったが、今後は複数名の同時撮影・同時評価や、さまざまなジャンル・テンポのダンスに対応する拡張が求められる。また、現段階ではYOLOv8-poseの推定精度とフレーム単位のマッチングに限界があり、推定精度の向上や時系列の揺らぎへの補正技術の導入も検討すべき課題である。

加えて、教育現場に導入するためには、より簡易な操作性と高い応答性が求められる。軽量なモデルの選定や、Webカメラ・低スペックPCでも動作可能な設計の追求も必要である。また、教員が指導中にシステムを使って生徒にリアルタイムでフィードバックを与えるという活用場面を想定した機能(例:コメント表示、評価記録の自動蓄積)なども、今後の開発課題となる。

最終的には、本システムを教育現場に導入し、主観評価を補完する新たな評価基準として活用できるよう、継続的な実証実験とシステム改善を重ねていく予定である。

実験では、被験者1名が異なる3種類の動作(正確な模倣、リズムずれ、振り付けミス)をそれぞれ3回ずつ行い、計9回の評価を実施した。その結果、DTW距離は主観的評価と高い一致を示し、特に動作全体の再現性に関して有効であることが分かった。一方で、コサイン類似度は瞬間的な一致度を評価するのに適しているが、リズムのずれや部分的な誤りに対しては過大評価される傾向があった。これにより、評価対象に応じて指標を使い分ける必要があることが明らかとなった。

参考文献

- (1) 文部科学省：“新学習指導要領に基づく中学校向け「ダンスリーフレット」”，https://www.mext.go.jp/a_menu/sports/jyujitsu/1306098.htm, (参照 2025.06.03)
- (2) 稲田健太郎：“姿勢情報を用いた複数ダンサーの動作評価”，早稲田大学大学院修士論文(2020)
- (3) Ultralytics チーム：“紹介 Ultralytics YOLOv8”，Ultralytics, <https://www.ultralytics.com/ja/blog/introducing-ultralytics-yolov8>, (参照 2025.05.28)
- (4) H. Sakoe, S. Chiba：“Dynamic programming algorithm optimization for spoken word recognition”, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 26, No.1, pp.43-49 (1978)
- (5) M. Cuturi, M. Blondel：“Soft-DTW: a Differentiable Loss Function for Time-Series”, TheInternational Conference on Machine Learning (ICML), Vol. 70, pp.894-903 (2017)