

動画反転と自動注釈付与によるダンス練習用動画生成

大谷 紀子 研究室

2272010 磯部 彩乃

1. はじめに

近年、SNS でのダンス動画の配信が増加し、投稿されたダンス動画を視聴して振付を習得する学習形態が広く用いられるようになった。従来のダンスの振付習得では対面で指導者から学ぶことが主流であったが、ダンス経験者の間でも動画のみを視聴し振付を覚える機会が増えている。しかし、ダンス動画のみで振付を習得する場合、身体の向きや位置関係を把握しにくい、指導者から口頭での説明が得られないなどの理由から、対面指導と比べて振付習得が困難である。

本研究では、ダンス経験者が動画を見ながらダンスの振付をより簡単に習得できるように支援することを目的とし、動画反転と自動注釈付与によるダンス練習用動画の生成システムを開発する。

2. システム概要

本システムは、覚えたい振付を撮影した動画が入力されると、特徴的な動作に関して注釈を付与した反転動画を出力する。システムの処理の流れを以下に示す。

- ① 動画を読み込む。
- ② 各フレームの映像を左右反転する。
- ③ 各フレームから人物の骨格情報を取得する。
- ④ 体の各部位の動き量を計算する。
- ⑤ 一定時間ごとに動きが最も特徴的な動作を決定する。
- ⑥ 特徴的な動作を表す注釈を PIL を用いて動画下部中央に表示する。
- ⑦ 元動画の音声 MoviePY を用いて、動画に合成して出力する。



図 1:出力動画の画像

③では、機械学習ライブラリ MediaPipe を用いて姿勢推定をする。動画はフレーム単位で処理され、各フレームに対して姿勢推定をし、手・足・耳・肩の x 座標、y 座標、z 座標を取得する。

④では、手・足・頭の移動量と身体の回転量を算出する。現在のフレームと前のフレームにおいて、手・足の座標の差を手・足の移動量、左右の耳を結ぶ直線の傾きの差を頭の移動量、左右の肩の奥行き差を回転量とする。

⑤では、各移動量と回転量から特徴的な動作を決定する。回転、両手、頭、手・足の順に条件を満たしているかを確認し、最初に条件を満たした動きを特徴的な動作とする。最初に、一定時間分の回転量の合計と平均が所定の閾値を超えた場合に回転したと判定し、肩の y 座標差の正負を用いて回転の左右を決定する。次に、左右の手の移動量が同程度である場合、両手が動いていると判定する。第 3 に、頭の移動量が一定以上であり、手・足の移動量を上回っている場合に頭が傾いていると判定し、左右の耳の y 座標差の正負を用いて傾きの左右を決定する。最後に、右手、右足、左手、左足の移動量のうち、移動量が最大の動作を選定

する。

⑥では、判定した動作の部位に応じて、「右手を動かす」「左手を動かす」「両手を動かす」「右足を動かす」「左足を動かす」「頭を右に傾ける」「頭を左に傾ける」「右に回る」「左に回る」という注釈を表示する。動画が反転していることを考慮し、表示される文の左右は利用者の体感と一致するように補正される。出力動画の画像を図1に示す。

3. 評価実験

ダンス経験者4名を被験者として評価実験を実施した。実験では、同一の曲につけられた2種類の異なる振付A、Bについて、それぞれ動画A、Bを用いて被験者に習得させた。動画Aは、振付Aを踊っている様子を撮影した動画を本システムにより変換したものであり、動画Bは、振付Bを踊っている様子を撮影した動画である。

被験者が振付を覚えた後、アンケートを実施し、振付の動きが理解しやすかったか、どの部位を動かすか理解しやすかったか、振付を覚えやすかったか、本システムは振付を覚えるのに役に立つのかをそれぞれ5段階で評価させた。また本システムの良い点、改善してほしい点を自由形式で回答させた。

アンケートの結果、振付全体の動きの理解しやすさについて、生成された動画がやや良いと回答した被験者が3名、どちらともいえないと回答した被験者が1名であった。一方でどの部位を動かすかについて、生成された動画の方が良いと回答した被験者が3名、生成なし動画の方がやや良いと回答した被験者が1名であった。振付の覚えやすさ、振付を覚える際に役に立つかについて、全体的に生成された動画が高く評価された。生成された動画の良い点として、振付の左右が記載されていること、回転の向きが書いてあることが挙げられた。一方、改善してほしい点として、注釈が下にあると視線が足に集中するため頭の上に表示してほしいこと、表示文のサイズを大きくするこ

とが挙げられた。また、表示文が今の振付の表示ではなく、次の振付が表示されると良いとの意見が示された。

4. 考察

評価実験の結果、生成された動画は従来の動画視聴のみの学習と比べて有効であると考えられる。振付の動きの理解しやすさについて、肯定的な意見が多かった一方で、どちらともいえないという回答が示された。ダンス経験者にとって、元動画のみでもある程度全体の流れを把握できること、振付がシンプルであったことにより、生成された動画の効果が表れにくかったと考えられる。しかし、どの部位を動かすかや、振付の覚えやすさ、振付を覚える際に役に立つかについては、生成された動画に対する評価が全体的に高かった。生成された動画が左右反転され、「右手を動かす」「頭を左に傾ける」のような具体的な注釈を自動で表示することにより、被験者がどの部位に注目すべきかを即座に把握できるようになったためであると考えられる。実験結果は、複雑な身体動作の学習において、視覚情報の与え方が学習効率に大きく影響することを示した Bailenson らの研究[1]とも一致している。ダンス経験者であっても、元動画のみでは左右の判断や回転方向の把握により振付を覚える時間が増えるため、注釈による補助情報が有効に機能したと推測される。生成された動画の良い点から、反転表示と注釈表示の組み合わせが、利用者の身体感覚と動画内の情報を一致させる効果を持つことが確認できた。従来の動画学習で問題となっていた、体の向きや位置関係の理解しにくさが軽減されたと考えられる。

参考文献

- [1] J. N. Bailenson, J. Blascovich, A. C. Beall, J. M. Loomis, "Comparing motion capture to video for learning complex movements," *Journal of Computer Assisted Learning*, Vol.24, No.2, pp.75–86, 2008.