

時系列分類モデルに基づくハウスダンス振付構成可視化に向けた試作

塚本 健太* 辻 順平*

概要. 近年, ダンス情報処理という分野が発展しており, ダンスのデジタル化や分析が可能になってきている. 本研究では, ハウスダンスの振り分類と振付構成の可視化を目的としたシステムを提案する. 振り分類の手法としては, MediaPipe を用いてダンス動画から骨格情報を抽出した前処理データに対し, LSTM モデルによって 10 種類のハウスダンス基本ステップの分類を行う. 学習済みモデルを用いて振付構成動画の各フレームを分類し, 確信度を時系列グラフで表示することで振付構成の可視化を試みた. 結果としては, 振り分類タスクでは 97.5% の精度を達成した.

1 はじめに

ダンスにおいて振りは一つの技であり, これらを連続的に組み合わせることで振付構成が生まれる. 近年, ダンス情報処理という分野で様々な研究が行われており, ダンスモーションジャンル分類や動画データベース構築などの成果が報告されている. 一方で, 振付構成を自動抽出するための基礎技術である個々の振り分類に関しては, 確立した手法が存在しない.

そこで本研究では, ダンス動画における振りを分類するモデルの構築と振付構成の可視化を目的とする. 対象ジャンルはハウスダンスに限定する. 著者の一人は 6 年間のハウスダンス実践経験を有しており, この知見を振り分類に活用する.

2 関連研究

土田らは, ダンス情報処理研究のための世界初の大規模ストリートダンスの動画データベース「AIST Dance Video Database (AIST Dance DB)」[1] を構築している. このデータベースは, 新規作成されたダンス音楽 60 曲と, それに合わせて独自の振り付けでストリートダンスを踊る 35 名のプロダンサーを撮影した 13,940 本の動画を収録している. 10 種類のダンスモーションジャンルをカバーし, 最大 9 台のカメラによる多角度撮影を実現している.

同研究では, ダンスモーションジャンル分類タスクにおいて 4 つのベースライン手法を提案し, 91.4% の分類精度を達成しているが, 振りそのものの分類をする手法ではない.

3 提案手法

3.1 システム概要

複数種類の振りが収録されたダンス動画を入力として, 収録されている振りの系列を推定し出力するシステムを提案する. その手順は以下の通りである:

1. 入力動画に対して MediaPipe を適用して体の骨格の各パーツの座標の系列を抽出する.
2. 抽出されたデータを入力とし, 振りの種類を出力する LSTM モデルをデータにより学習させることで, 振り分類モデルを構築する.
3. 構築した学習済みモデルに対して複数の振りが含まれる振付構成動画を入力し, 各フレームにおける分類確率の推移を時系列グラフで表示する. さらに, グラフから分類確率が高い振りを抽出することで, 振付構成の可視化を実現する.



図 1. 33 キーポイント



図 2. 全身を代表する
12 キーポイント

3.2 Media Pipe による骨格情報の抽出

Media Pipe とは, Google が開発したリアルタイム機械学習パイプラインフレームワークである.

本研究では, Media Pipe のポーズ推定機能を活用し, 33 のキーポイントから全身を代表する 12 キーポイントに削減して抽出を行う. 抽出されたキーポイントデータは, 各フレームにおける xyz 座標として算出され, CSV ファイルに保存される (図 1・2).

3.3 分類対象

本研究では、ハウスダンス振り分類タスクの第 1 段階として、AIST Dance Video Database [1] にも収録されているハウスダンスの基本動作 10 種類を分類対象とした。具体的には、backskip, farmer, fourstep, heelstep, padobere, peterpaul, shuffle, sidewalk, swirl, train の 10 種類である。

3.4 データセット

全 360 動画のデータセットを用意した。内訳は次の通り：

1. AIST Dance Video Database [1] に収録された動画のうち、ハウスダンス動画のみを抽出した。10 種類の振りに対し各 12 動画（計 120 動画）
2. 著者の一人と協力者（ダンサー）3 名の計 4 名の動画撮影を独自に行った。4 名のダンサーには 10 種類ずつの振りを踊ってもらい、各振りに対し 1 人あたり 6 動画の撮影を行った（計 240 動画）。

さらに、すべての動画を前半と後半に分割することで動画数を 2 倍とし、計 720 動画のハウスダンス用拡張データベースを作成した。

1 動画につき、およそ 100~120 フレームあり、動画内に振りは 3~4 回繰り返し行われている。

3.5 モデル構築

本研究では、動画のフレームごとの時系列データを分類することを目的として、RNN (Recurrent Neural Network) ベースの分類モデルを採用する。特に、長期的な記憶の保持が可能な LSTM (Long Short-Term Memory) を基本分類モデルとして採用する。

分類モデルは入力層・LSTM 層・全結合層・出力層で構成され、入力 N 次元、LSTM のユニット数: M 、出力 K 次元とした。入力する動画のフレーム数を 100 フレームに固定し、学習の際のミニバッチ数は 16 個に設定した（図 3・4）。



図 3. 分類モデル全体図

データセットは、訓練データ、検証データ、テストデータの割合を 4:1:1 で分ける。さらに訓練データには、3 種類の手法（ノイズ加工、時間軸シフト、

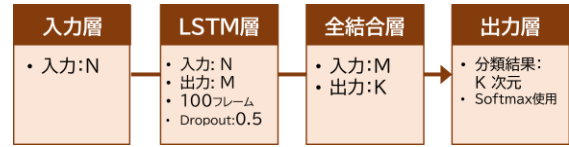


図 4. 分類モデル構成図

時間スケール変更）によるデータ拡張を行い、計 1440 データとした。

3.6 振付構成の可視化

学習済み分類モデルを用いて、複数の振りで構成された動画に対してフレーム単位での分類を実行する。各フレームにおける分類結果の確信度をグラフ化し、時系列で表示することにより、振り構成の変化を可視化する（図 5）。

4 実験

3 節で示した手法に基づき振り分類のプロトタイプモデルを構築し、精度評価実験を行った。

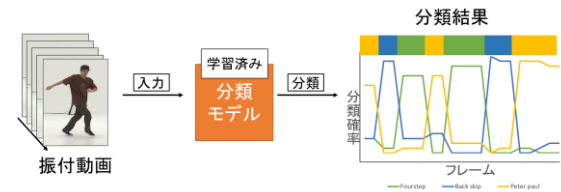


図 5. 振付構成可視化

実験条件として、入力には 3.2 節で示した 12 キーポイントの xyz 座標 3×12 次元、ユニット数は $M = 32$ ユニット、出力は 3.3 節で示した分類対象である $K = 10$ 次元に設定した。評価方法は、事前に分割しておいた 10 種類の振り動画各 12 個をテストデータとし、F1 スコアに基づいて分類精度を算出した。

実験の結果、120 個のテストデータに対して 117 個を正しく分類し、97.5%の精度を達成した。誤分類は 3 個のみであり、この結果からダンス振り動画分類における提案手法の有効性が示された。

5 おわりに

本研究では、MediaPipe を用いた骨格情報の抽出と LSTM モデルによる振り分類手法を提案し、ハウスダンスの基本ステップ 10 種類の分類を試みた。実験の結果、97.5%の分類精度を達成し、提案手法の有効性が示された。一方で、振付構成可視化についてはまだまだ検証実験が必要である。

謝辞

本研究のデータセット作成において、AIST Dance Video Database [1] の一部を使用した。また、3名のハウスダンサー棚山 里奈、太田口 ゆず、田中 美音には、ハウスダンス動画撮影にご協力いただいた。この場を借りて感謝の意を表する。

参考文献

- [1] Shuhei Tsuchida, Satoru Fukayama, Masahiro Hamasaki and Masataka Goto. AIST Dance Video Database : Multi-genre, Multi-dancer, and Multi-camera Database for Dance Information Processing. In Proceedings of the 20th International Society for Music Information Retrieval Conference (ISMIR 2019), (2019).

未来ビジョン

ダンスは、視覚や身体感覚を通して理解されることが多い。しかし、感覚的な理解だけでなく、動きを言語化して理論的に捉えることも重要であると考えた。本システムを活用することで、ダンスの動きを客観的に可視化し、言語化して理解する新たなアプローチの第一歩となること目指した。

また、ダンス文化は古来より人から人へと受け継がれてきた文化である。現代の AI 技術を用いて、こうした継承される技術や表現をデータとして記録することで、過去の文化をより鮮明に保存し、次世代へと引き継ぐこと

が可能になる。これにより、ダンス文化のさらなる進化と発展につながることを期待している。

さらに、本研究をきっかけに、ハウスダンスを対象とした研究がより活発化し、ハウスダンス動画のデータセットの充実や新たな分析手法の開発が進むことを期待している。それに伴い、ハウスダンスというジャンルの魅力や特徴が学術的にも広く認知され、ダンス文化全体の発展に寄与できれば幸いである。