Vol.2020-MBL-94 No.32
Vol.2020-UBI-65 No.32
2020/3/2

マルチモーダル社交ダンス認識における従来手法と深層学習の比較評価

松山 仁[1,a)]　廣井 慧[1]　梶 克彦[2]　米澤 拓郎[1]　河口 信夫[1]

概要：社交ダンスの演技は，3 から 4 拍ごとにまとまった動作単位であるフィガーを組み合わせることで構成される．フィガーは基本的な練習項目である一方で，その種類の多さや複雑さゆえ初心者や中級者にとって練習が困難である．そこで我々は行動認識技術を用いてフィガーの識別を行い，フィガーごとの適切な踊り方を指導することで，ダンスフィガーの学習支援を目指す．ダンスの行動認識において一般的なのは画像センサを用いて踊り手の全身の動きを取得する手法だが，社交ダンスは二人一組で多彩な動きを行う為，既存の画像センサを用いた行動認識手法をそのまま適用することは難しい。そこで我々は画像に加えて加速度・角速度を加えたマルチモーダルセンサによるデータ収集を行い，さらに社交ダンスの動作特性を考慮した特徴設計を行うことで，フィガーの複雑さや遮蔽などの課題を解決したダンスフィガー認識手法を実現した．本稿では特にフィガー認識における分類手法を，4 種類の従来型機械学習による手法，さらに深層学習を用いた手法を加えてそれぞれ検証し，社交ダンスフィガー認識における適切な分類手法選択について評価する．

キーワード：行動認識，機械学習，データ収集，ユビキタスコンピューティング

Comparison and Evaluation of Ballroom Dance Figure Type Recognition Using Multi-modal Sensor

Keywords: Activity Recognition, Machine Learning, Data Collection, Ubiquitous Computing

謝辞

本研究の一部は，科研費 JP17H01762 により支援して頂いております。

[1] 名古屋大学 大学院 工学研究科 情報・通信工学専攻
Furo-cho, Chikusa-ku, Nagoya-shi, Aichi 464-0814, Japan
[2] 愛知工業大学 情報科学部 情報科学科
Yakusa-cho, Toyota-shi, Aichi 470-0392, Japan
[a)] hitoshi@ucl.nuee.nagoya-u.ac.jp

ⓒ 2020 Information Processing Society of Japan

Vol.2020-MBL-94 No.32
Vol.2020-UBI-65 No.32
2020/3/2

マルチモーダル社交ダンス認識における従来手法と深層学習の比較評価

松山 仁*, 廣井 慧*, 梶 克彦†, 米澤 拓郎*, 河口 信夫*

* 名古屋大学 工学研究科 情報・通信工学専攻

† 愛知工業大学 情報科学部 情報科学科

背景

- スポーツ学習支援システムの発展
- 社交ダンス学習支援システム
  - フィガー情報の必要性
- 動きの自由度・複雑度，種類が多い
  - **マルチモーダルセンサ**の活用
  - 特徴設計の工夫
  - 分類モデルの検討



Skill Monitor(1)

(1) https://www.rbbtoday.com/article/2016/10/05/145827.html

研究概要

ダンスフィガーの自動認識

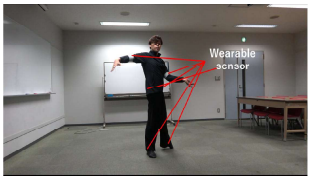
- ダンス技術の客観視
- フィガー(振付)ごとに踊りを評価
  - タイミングや姿勢，足の方向など
- ダンス教示だけでなく他スポーツへの応用を期待

データ収集

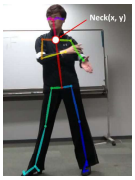
- 1分程度のフィガーを13種類 × 20回 × 7人 (計二時間強)
- サンプリングレート:120 Hz, センサは両肘・両腰・両足首
- 加速度・角速度,動画のマルチモーダルなデータ
- 計二時間超のデータ, 研究者に公開(2)

認識手法

- 加速度・角速度，OpenPoseの骨格位置を利用
- 標準化，移動影響除去，スケーリング処理，入力値生成
  - 従来型, 非回帰型NN: 各フィガーデータの特徴量を計算
  - LSTM: 時系列データを入力
- 従来型，非回帰NN，LSTMそれぞれによりフィガークラス分類



ウェアラブルセンサ

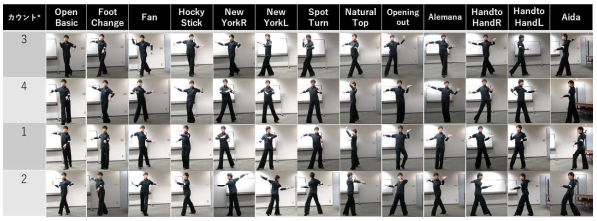


関節位置

(2) http://hasc.jp/

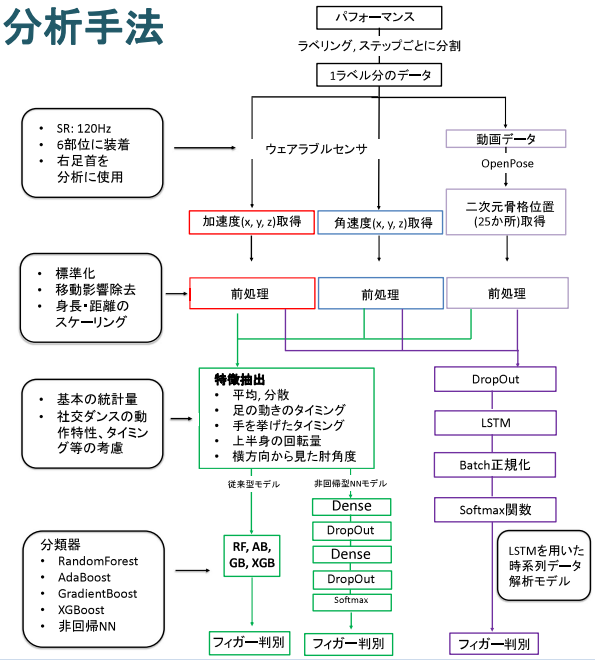
ダンスフィガーデータ

- 各フィガーは「3」の拍から開始すると定義(3-4は予備動作)
- 本データ中の全てのフィガーは4拍で構成され，時間的に同じ長さ
- Fan/AlemanaやNewYork/HandtoHandなど類似動作を含むフィガー



*カウント: 音楽の拍. 四つのカウントで一小節.

分析手法



結果

- "LSTM"以外のコラムは特徴抽出あり
- 従来手法ではAdaBoostが最高精度
- 特徴抽出ありの非回帰NN(Non-RNN)がそれを上回る
- LSTMはデータ不足のためか精度が伸び悩む

分類手法	従来型機械学習モデル				深層学習	
特徴抽出	あり					なし
分類器	RF	AB	GB	XGB	Non-RNN	LSTM
Acc,gyr	0.847	0.836	0.822	0.832	0.860	0.68
Open Pose	0.824	0.845	0.800	0.831	0.831	0.62
Hybrid	0.908	**0.924**	0.892	0.900	**0.931**	0.83

認識結果

今後の課題

- **深層学習を活用するためデータ増強**
  - フィガー数
  - フィガーの種類
  - フィガーの並び方が変わった場合のデータ
  - 2人で踊った場合のデータ
- 音楽的特徴の変化への頑健さ
  - BPMの変化など
- バッチサイズを小さくした認識
- 教示システムの設計
- システムによるダンス技術向上の確認
- 他のスポーツへの応用

ⓒ 2020 Information Processing Society of Japan