

映像の主成分分析に基づくダンスリズムの逐次評価

A Study on Online Rhythm Evaluation of Dance Movement using Principal Component Analysis of Motion Pictures

今野 智明†
Tomoaki Konno徐 建鋒†
Jianfeng Xu高木 幸一†
Koichi Takagi酒澤 茂之†
Shigeyuki Sakazawa

1. はじめに

エアロビクスダンス（以下、ダンス）において、リズム感は重要な要素のひとつである。一般に、ダンスの練習者はお手本（教師）の動きを模倣して運動することが多い。しかしながら、練習者は不慣れさなどから教師の動きに追従できず不正確な動作をしてしまい、両者の間にずれが生じる可能性がある。練習者はこのずれを逐次知覚し修正することで、教師に倣おうとする。したがって、時間セグメント毎に動きの正確さを、練習者に逐次的にフィードバックできれば有用である。これまでに筆者らは、DTW (Dynamic Time Warping) によるセグメント毎のリズム評価手法を検討してきた[1]。しかしながら、評価のためには映像シーケンス全体のデータに対して主成分分析 (Global Principal Component Analysis, G-PCA) を適用することを前提とするため、逐次的に評価結果をフィードバックすることができなかった。一方で、部分的な映像シーケンスに対する短時間主成分分析 (Short Term PCA, ST-PCA) によるリズム抽出の特性についても検討してきたが、逐次的な評価に関しては未検討であった[2]。

そこで、本稿では、映像の ST-PCA により、ダンスのリズムを時間セグメント毎に逐次評価する手法を検討し、その有用性について検証する。

2. 関連研究

人の動作を撮影した映像に基づく動きの評価を目的とした研究が報告されている[3][4]。文献[3]では、応援団の応援の振りを対象として、マーカを付けた練習者の動きを撮影し、お手本と比較することで、そのマーカの位置及び移動方向の時系列情報から動作のタイミングや大きさの正確さを評価する手法を提案している。しかしながら、採点方法の妥当性の検証は十分ではない。文献[4]では、現代ダンスを対象として、映像から人物スケルトンを抽出し、そのハフパラメータに基づく動き特徴量に対し周波数解析を行うことで、リズムを抽出する手法を提案している。リズム特徴の周波数特性の観点からダンスの上級者と初級者との違いを分析しているが、リズムの正確さの時間変化に関しては検討していない。また、文献[3][4]共に撮影した映像シーケンス全体に対するオフライン処理が前提である。

一方、Mocap (Motion Capture) データを利用した高精度な動きの解析に関する研究もなされている[5][6]。文献[5]では、音楽と動きの同期のために、Mocap によって得られた各身体部位の特徴量に対して ST-PCA を適用することで、逐次的にグローバルな動きの特徴（リズム特徴に相当）を抽出し、動きのビートを求める手法を提案している。また、文献[6]では、Mocap によって得られた各部位の位置や角度に基づく特徴の教師データとの差異に基づき動きの評価を

行いそのフィードバックに着目して検討を行っている。各時刻における教師とのずれが最も大きい部位の指摘や正解ポーズを示す CG モデルを提示するフィードバックを行っている。しかしながら、Mocap センサによる動きデータの取得は手間が大きいと、ダンス評価の利用シーンを想定すると望ましくない。

そこで、本稿では、カメラなどによって逐次入力される部分的な映像シーケンスに対して ST-PCA により動きのリズム特徴を抽出し、教師のリズム特徴と比較することでリズムの正確さをオンラインで評価する手法について述べ、その有用性を検証する。

3. ST-PCA を用いたリズム逐次評価

本研究のリズム逐次評価の概要を図 1 に示す。練習者映像シーケンスの一部（所定の分析窓幅）の画素フレームに対して、ST-PCA を適用することでリズム特徴（主成分得点系列）を抽出する[2]。次に、セグメント内での練習者と教師のリズム特徴信号の時間軸上の対応付けを行うとともに微小なずれを吸収するため DP マッチング[7]を行う。動きの遅れと進みを同等に扱うため対称型の DP マッチングとし、さらに信号の極端な伸縮を防ぐため傾斜制限を設ける。時間軸上で対応づけられた両者のリズム特徴信号の類似度を求め、リズムの評価値（類似度、“○”/“×”）をフィードバックする。リズムの正確さは動きの絶対的な大きさに依存しないと考えられるため、両信号ベクトルの方向の違いを表す $\cos$ 類似度を評価尺度とする。

また、これまで PCA の分析窓幅のリズム抽出に対する特性について検討しているが[2]、リズム評価においても同様に検討する。図 2 は、PCA の分析窓幅と評価セグメントの関係を示す。図 2 (a) は、評価セグメントの映像データのみを対象として PCA によりリズム特徴を求めるが、当該評価セグメント中に不正確な動きのみが存在する場合、正しい動きの方向とは異なる射影軸を捉えてしまうことがある。そこで、図 2 (b) のように、評価セグメントの過去の映像も含めて PCA を行うことにより、短時間の不正確な動きの影響を低減させることができる。過去の映像を利用することで、当該評価セグメントの最終フレームを得た時点でそれ以降のフレームを待つことなく、そのセグメントの評価が開始可能となる。

また、PCA で求めた固有ベクトルの方向は正負の 2 パターン存在する。ダンスの動きは一般に滑らかでありその動きの方向を表す固有ベクトルは急激に変化しないと仮定し、スライディング前とスライディング後の分析窓によって求めたそれぞれの固有ベクトルの内積が正となるパターンを選択することで、セグメント毎の主成分得点系列を連結する[5]。なお、先頭のセグメントに関しては、教師の主成分得点系列との類似度が大きくなるように練習者の主成分得点系列の符号を必要に応じて反転する。

† 株式会社 KDDI 研究所, KDDI R&D Laboratories Inc.

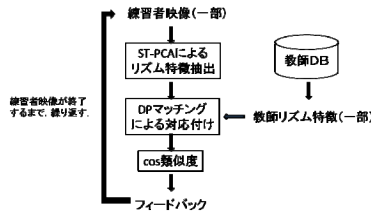


図1 逐次評価の枠組み

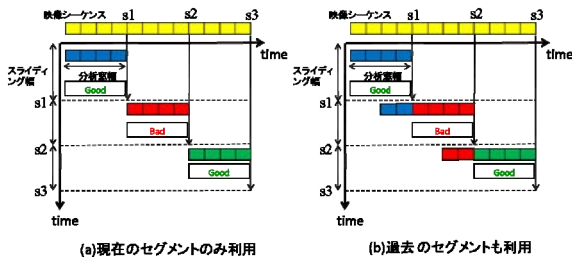


図2 ST-PCAの分析窓幅と評価セグメント

教師映像の動き特徴抽出に関しては、同じ種類の動きをする区間の主成分得点に対してフーリエ変換することにより、ピーク周波数より各動きの周期を算出し、ノイズ除去のためローパスフィルタで平滑化した。

4. 評価実験

本手法によるリズムの客観評価の精度を調べるため、単一の周期的な動きを含む映像に対して、ST-PCAを適用しリズム特徴(第1主成分)を求め、教師特徴との比較により評価する実験を行った。セグメント毎に筆者が付与した練習者映像に対する主観評価(○/×)と、本手法で求めた客観評価(類似度0.8以上を○, 0.8以下を×)の一致率(=評価が一致したセグメント数/全体のセグメント数)を表1に示す。評価に用いた動きはエアロビクスの6種類の動作であり、被験者は計2名(a)~(c)と(d)~(f)であり、それぞれ教師に倣って正しく動作した映像をSet A, 不正確な動作(連続1周期程度)を含んだ映像をSet Bとした。ST-PCAの分析窓幅及び評価セグメント幅は各動きの1周期分のフレーム数とし(教師映像(FR=30fps)は既知のため、教師の動きの周期も既知)、またDPマッチングの傾斜制限は1/2~2倍とした。

表1より、Set Aに関しては、(d)を除いた5種類の動きにおいて主観と一致する結果が得られた。一方、Set Bに関しては、Set Aと比較して全体的に一致率が低く、(d)においては特に低い。動作停止区間に対するST-PCAで微小な動きの方向を捉えてしまったことや固有ベクトルの不要な反転が起こったと考えられる。これは評価が“×”となったセグメントに対するST-PCAを行った際の固有ベクトルを次のセグメントの固有ベクトルの符号反転条件として用いるためである。そのため、前セグメントの評価によって反転するか否かを決定することが考えられる。

表1 主観評価と客観評価の一致率

	(a) Jjack (26fr/seg)	(b) s. touch (56 fr/seg)	(c) walk. (29 fr/seg)	(d) move1 (56 fr/seg)	(e) move 2 (56 fr/seg)	(f) move 3 (56 fr/seg)
Set A	7/7	6/6	10/10	5/8	8/8	4/4
Set B	6/7	6/6	9/10	2/8	7/8	3/4
Total	13/14	12/12	19/20	7/16	15/16	7/8

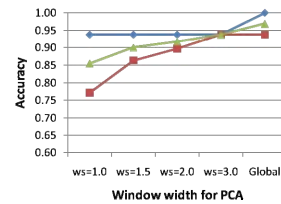


図3 分析窓幅に対する評価の一致率

表2 Set Aの(d)の不一致セグメントの類似度

	ws=1.0	ws=1.5	ws=2.0	ws=3.0	G-PCA
Seg 1	0.71	0.71	0.71	0.71	0.82
Seg 2	0.75	0.76	0.75	0.75	0.85
Seg 6	0.74	0.72	0.77	0.79	0.81
Ave.	0.73	0.73	0.75	0.75	0.83

次にリズム評価時におけるPCAの分析窓幅の評価精度への影響を調べるため、分析窓幅を各動きの1周期分のフレーム数*ws(=1.0~3.0)倍に変化させた場合の各動きの一致率の平均を図3に示す。評価の平均一致率は、ws=1.0の時は0.85と一致率が最も低く、一方、ws=2.0の際に0.92であった。連続する不正確な動作区間が1周期程度と想定すれば、1.5~2周期程度の分析窓幅が有効であることがわかった。これは文献[2]の知見からも推察される。不正確な動きを除けば多くは周期的な動きであるため、分析窓幅が長ければ、正しい動きの方向(射影軸)を求める際の不正確な動きの影響を軽減できるが、その一方で評価のための時間がかかるといった問題がある。

Set Aの(d)に関しては、正しい動きに関わらず、分析窓幅を長くした場合においても、主観評価と一致しないセグメントがあった。この不一致であった3つのセグメントにおける類似度を表2に示す。いずれの分析窓幅においても類似度は0.7程度であり必ずしも低くはない。主観では判断しきれなかった教師とのリズム特徴の差異が生じていると考えられ、主観とより一致する評価尺度が求められるが、今後の課題とする。

5. まとめ

本稿では、映像に対するST-PCAを用いたダンスリズムの逐次的な評価手法について検討した。2周期程度の分析窓幅でST-PCAを行うことにより、G-PCAとほぼ同じ精度で動きのリズムを評価できることがわかった。主観評価との平均一致率は0.92であった。今後は、動きデータを拡充して、さらなる検証を進めていく。

参考文献

- [1] 今野ら, “DTWを用いた動きのリズム評価に関する一検討,” 信学会全国大会講演論文集, 2011
- [2] 今野ら, “ダンスリズムの逐次評価に向けた映像の主成分分析,” 映像情報メディア学会年次大会講演論文集, 2011
- [3] 高橋ら, “振りの練習を支援するインタラクティブシステム,” インタクション論文集, 2004
- [4] M. Naemura et al.: “Extraction of rhythmical factors on dance actions thorough motion analysis,” Proc. of 7th IEEE WACV/MOTION, 2005.
- [5] J. Xu et al., “Beat Induction from Motion Capture Data Using Short-Term Principal Component Analysis” Jour. of ITE, Vol. 64, No. 4, pp.577-583, 2010.
- [6] J. C. P. Chan, : “A Virtual Reality Dance Training System Using Motion Capture Technology,” IEEE Trans. Learn. Tech., 2010
- [7] 内田, “DP マッチング概説: 基本と様々な拡張,” 信学技法 (PRMU), 2006