

隠れマルコフモデルを用いた音楽に連動したダンスと表情の自動生成 における違和感のある動作の軽減

糸賀朱音 長名優子

東京工科大学 コンピュータサイエンス学部

1 はじめに

近年、3DCGを使用したエンターテインメントが広く普及してきている。その中の一つとして、3D キャラクタを音楽に合わせて動かすことのできるものがある。MikuMikuDance (MMD) という3Dの映像作成が可能なフリーソフトウェアでは、MMD上で使用できるキャラクタの3Dモデルやアクセサリ、ダンス動作データなどが多く公開されているため、誰でも手軽にアニメーションを作成することができる。しかし、自分で新しく音楽に合わせたダンス動作を作成することは依然困難な作業である。

それに対し、音楽に連動したダンス生成手法 [1] や隠れマルコフモデルを用いたダンス生成手法 [2] も提案されている。文献 [2] の手法では、ダンス動作だけでなく楽曲に連動した表情動作も生成することができる。この手法ではダンス動作は K -means++法でクラスタリングし、分析区間は1小節としている。分析区間どうしは姿勢の差を考慮し、数フレーム分の姿勢データを消去した後、結合される。MMDでは姿勢のデータがないフレームは自動的に線形補間が行われるため、生成されたダンス動作には床を滑って移動しているように見えるなどの違和感が存在する。

そこで、本研究では、文献 [2] の手法において、分析区間を結合する際に調整を加えることでダンス動作の違和感の軽減を目指す。

2 隠れマルコフモデルを用いた音楽に連動したダンスと表情の自動生成

2.1 音楽に連動したダンスと表情の学習

提案手法では、まず、入力された楽曲から音響特徴量 (Mel-Frequency Cepstrum Coefficients : MFCC) を抽出する。ダンス動作や表情動作、MFCCの分析区間は1小節とし、それぞれを K -means++法を用い

てクラスタリングする。クラスラベルをダンス語彙、表情語彙、音響語彙としてそれぞれ割り当て、類似する動作の基本形として扱う。次に、隠れマルコフモデルを用いて音響語彙とダンス語彙、音響語彙と表情語彙の対応関係を学習する。これらの隠れマルコフモデルは、ダンス語彙/表情語彙を隠れ変数、音響語彙を出力としている。

2.2 音楽に連動したダンスと表情の生成

音楽に連動したダンス動作と表情動作を生成は、音響語彙とダンス語彙、音響語彙と表情語彙の対応関係を学習した隠れマルコフモデルを用いて行う。はじめに、入力された楽曲から MFCC を抽出する。抽出された MFCC に対し、各音響語彙とのユークリッド距離を求め、距離が最小であるような音響語彙を割り当てる。割り当てられた音響語彙系列と隠れマルコフモデルから音響語彙系列に対応するダンス語彙系列と表情語彙系列を決定する。続いて、決定したダンス語彙系列/表情語彙系列から、対応するダンス動作/表情動作をランダムに選択する。連続する分析区間の動作を線形補間して結合して出力することで音楽に連動した新たなダンス動作と表情動作が最終的に生成される。

2.2.1 隠れマルコフモデルを用いたダンス語彙/表情語彙の系列の決定

音響語彙とダンス語彙/表情語彙の関係を学習した隠れマルコフモデルは、それぞれ隠れ変数をダンス語彙/表情語彙、出力を音響語彙としている。割り当てられた音響語彙系列 a_1^T から、ビタビ・アルゴリズムを用いてダンス語彙系列 d_1^T と表情語彙系列 e_1^T を決定する。ここで、 T は入力された楽曲の小節数 (分析区間の数) を表している。ビタビ・アルゴリズムは、出力がわかっているときの隠れ変数の尤もらしい遷移を求めるアルゴリズムであり、提案手法では入力された楽曲の MFCC に対応する音響語彙系列に対し、その音響語彙系列が出力される確率の高いダンス語彙/表

Reduction of Uncomfortable Movements in Automatic Generation of Dance and Facial Expressions Linked to Music using Hidden Markov Models
Itoga Akane and Osana Yuko (Tokyo University of Technology, osana@stf.teu.ac.jp)

情語彙の系列を求めることで、ダンス語彙/表情語彙の系列を決定する。

2.2.2 ダンス動作/表情動作の選択

各分析区間で決定されたダンス語彙系列/表情語彙系列から対応するダンス動作/表情動作をクラスタ内からランダムに選択する。

MikuMikuDance では、3D キャラクタを 21 のボーンを各フレームで回転、移動させることでダンスを表現している。提案手法では回転はオイラー角を用いて表現する。分析区間 n におけるダンス動作 $D^{(n)}$ は、位置座標の 3 次元ベクトル $p^{i(n)}$ とボーンの回転を表すオイラー角のパラメータの 63 次元ベクトル $b^{i(n)}$ の 2 つのベクトルを用いて表現する。

2.2.3 動作の出力

選択されたダンス動作/表情動作つなげることで最終的なダンス動作/表情動作が生成され、出力される。

従来の手法 [2] では、2 つの分析区間のダンス動作や表情動作を結合する際、前の分析区間の最後のフレームの姿勢と次の分析区間の最初のフレームの姿勢の差に応じて、数フレーム分の姿勢を削除し、そのうえで線形補間を行っている。そのため、前後の姿勢によって上手くつながる場合もあれば違和感が生まれる場合もある。なお、ダンス動作は、3D キャラクタの位置座標 3 次元と、21 のボーンの回転を表すパラメータ 63 次元の 66 次元で表されるが、消去の際は 2 つを別々に分けて考える。

提案手法では、線形補間によってダンス動作に違和感が発生するのを防ぐため、結合する前後の姿勢によって以下のような 3 つの場合に分けて処理を行う。

(1) 線形補間の結果に違和感がない場合

線形補間を行った結果に違和感がないようなケースでは、線形補間の結果をそのまま使用する。たとえば、図 1(a) のように、両足が地面についた状態から右足のみ上げる場合では、線形補間のみでも自然に動かすことができるため、線形補間の結果をそのまま使って分析区間どうしを結合する。

(2) 線形補間の結果に違和感がある場合

線形補間を行った結果に違和感がある場合は、動作を追加したり、位置座標を調整することで対応する。たとえば、図 1(b) のように両足を揃えた状態から左足を外側へ移動させる場合では、線形補間を行って結合すると、左足が平面上を滑るように移動するのに合

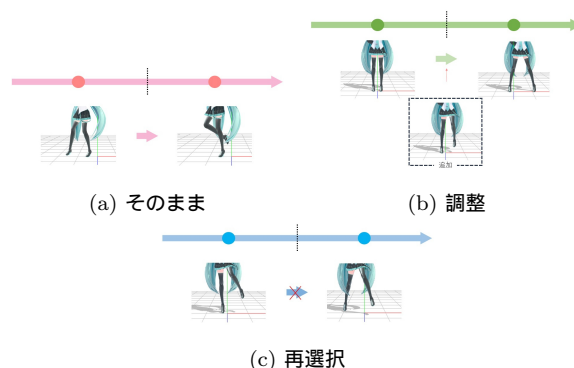


図 1: 動作の結合の場合分け

わせ身体も同じ方向へ動くことになる。そのため床の上をすべて平行移動しているように見え、違和感が生じる。この違和感を軽減するためには、2 つの動作の間に 1 度足を上げる動作を加え、結合を行う。そうすると、左足は元の状態から 1 度足を上げ、目的の位置まで足を下げつつ移動させるような、山形の動きになり、違和感が軽減される。他の動作でも、それぞれ適切な動作を追加した上で補間することで違和感を軽減する。

動作の追加や位置座標の調整でも違和感を取り除くことが難しい場合には、動作の再選択を行う。たとえば、図 1(c) のように片足が地面についている状態から、両足が図の左方向へ移動する場合を考える。2 つの姿勢を線形補間して結合しようとする、地面についている足が左方向へ動くのと同時に浮いている足が同じように左方向へ動くことになる。平面上を滑るような動作になるが、この場合は違和感を軽減させるような動作の追加が難しいため、後ろの分析区間の動作を再選択することで対応する。

3 計算機実験

計算機実験を行い、提案手法においてダンス動作と表情の生成が行えること、従来手法に比べて不自然な動作が軽減できていることを確認した。

参考文献

- [1] 深山 覚, 後藤 真孝: “MachineDancing: ダンス動作データの自動分析に基づく音楽に連動したダンス生成手法,” 情報処理研究報告, Vol.2014-MUS-104, No.14, pp.1-6, 2014.
- [2] 佐藤 大輝, 長名 優子: “隠れマルコフモデルを用いた音楽に連動したダンスと表情の自動生成,” 電子情報通信学会総合大会, 2019.