

HMMに基づく音楽に同期した振り付けの生成

Music Driven Computer Animation Based On HMM

亀田 康介^{*1}
Kousuke Kameda

渡辺 郁美^{*2}
Ikumi Watanabe

山本 正信^{*3}
Masanobu Yamamoto

^{*1} 新潟大学大学院自然科学研究科
A Graduate School of Science & Technology, Niigata University

^{*2,3} 新潟大学工学部情報工学科
Faculty of Engineering, Niigata University

1. はじめに

動作とそれに同期した音をそれぞれ基本要素に分割し、分節化された基本動作を新たな音に対して再構成することによって動作の自動生成を行うことができる[1]。しかし、生成された動作の繋ぎ目では不自然な動きが見られる。その一部を図1に示す。



図1 動作の繋ぎ目

これは、動作の繋ぎを考慮せずに再構成を行ったことが原因と考えられる。本研究では音楽と動作の遷移を考慮した手法を提案する。

2. 学習

音楽に合わせて身振りをする動作をモーションキャプチャにより測定しておく。得られたモーションデータは姿勢の時系列である。動作とそれに同期した音楽との関係をHMM(隠れマルコフモデル)によりモデル化し、モデルの状態遷移確率分布と出力確率分布を学習する。

2-1. HMM(隠れマルコフモデル)

HMMとは、マルコフ過程に従って遷移する内部状態、および各状態における記号の出現確率分布から構成される確率モデルである。

動作生成のための隠れマルコフモデルは、5項組

$M = (Q, \Sigma, A, B, \pi)$ によって定義される。

(1) $Q = \{q_1, \dots, q_N\}$: 基本動作(状態)

(2) $\Sigma = \{o_1, \dots, o_M\}$: 音楽(出力)

(3) $A = \{a_{ij}\}$: 動作遷移確率分布

a_{ij} は基本動作 q_i から基本動作 q_j への遷移確率であり、 $\sum_j a_{ij} = 1$ を満たす。

(4) $B = \{b_i(o_t)\}$: 音楽出力確率分布

$b_i(o_t)$ は状態 q_i で記号 o_t を出力する確率であり、

$\sum_t b_i(o_t) = 1$ を満たす。

(5) $\pi = \{\pi_i\}$: 初期動作確率分布

π_i は q_i が初期状態である確率 $P(X_1 = q_i)$ である。

2-2. クラスタリング

全身を16の部位で表した多関節モデルで姿勢を表わす。時系列姿勢データから部位の角加速度を計算し極値を検出する。極値が極小、極大、極小と変化する区間を1つの動作として分節化する。動作の特徴を極小、極大、極小となる3フレームを姿勢データとし、クラスタリングにかけることで分節化された動作を基本動作に分類する。クラスタリングにはk-meansアルゴリズムを用いた。

音楽には正確な音の情報が得られるMIDIデータを用いる。人は主に打楽器音を聴きながら身振りをする傾向がある。そこで、MIDIデータに記載されている47種類の打楽器音を47通りの出力記号として分類した。

2-3. 動作と音楽の関連付け

同期した基本動作と音楽の記号列より、動作の遷移確率分布と音楽の出力確率分布を求めることで動作と音楽の関係を関連付ける。このとき、状態(動作)は見えているのでバウム・ウェルチアルゴリズムを使用する必要はない。

3. 生成

ここでは、学習結果を使ってビタビ・アルゴリズムを用いて音楽から動作の生成を行う。

まず、音楽をそのMIDIデータを使用して出力記号系列 $o_t^1 = o_1 \dots o_t$ に変換する。時刻 t で状態 q_i に到達する状態遷移系列は一般に複数あるが、このうち最大の確率値を与えるものだけを記憶していけば、最終的に最適な状態遷移系列を求めることができる。

4. まとめ

動作の遷移を考慮したアニメーションを自動的に作成するシステムを提案した。MIDIファイルを用いることで正確な音情報からのアニメーションを作成していく。

今後は、打楽器音だけでなくメロディーラインも考慮したアニメーションの作成を行う。また、元の動作の個性が生成動作に継承されていることを確認する。

参考文献

- [1] 青塚 寛之, 山本 正信: 「歌唱と身振りの教示に基づく音楽からの動作生成」, 電子情報通信学会論文誌D Vol. J90-D No. 11 pp. 3055-3064, 2007