

音楽情報を用いたロボットのダンスモーションのリアルタイム制御システム

中原 直人[†] 宮崎 光二[†] 坂本 元[†]
中津 良平[‡] 藤澤 隆史[†] 長田 典子[†]

Real-Time Dance Motion Control of Robot using Music Information

NAOTO NAKAHARA[†] KOJI MIYAZAKI[†] HAJIME SAKAMOTO[†]
RYOHEI NAKATSU[‡] TAKASHI X. FUJISAWA[†] NORIKO NAGATA[†]

1. はじめに

現在までに、音楽音響からのビート抽出や、それらをダンス等の動作に関連付ける研究がなされてきた[1][2]。しかし、システムが情報を提示するだけではユーザーは、コミュニケーションにおいて重要な双方向のやりとりを行うことができない。そこでユーザー側からもリアルタイムにシステムに情報を与え、変化を与えられることがよりインタラクティブ性の高いシステムを実現する上で重要であると考えた。

本研究では、ユーザーが入力した、キーボードもしくは音楽ファイルの音楽音響信号からビートの情報を抽出し、さらにキーボードの音の強弱の情報を用いてロボットのダンスモーションのテンポと緩急をリアルタイムに変化させ、制御できるシステムを開発した。

2. システムの実現技術

2.1 音楽情報抽出

本研究では、Wav形式の音楽音響信号から、次のビートの時刻、テンポ等のビート情報をリアルタイムに予測し出力する後藤らの手法[1]をもとに、ビート抽出モデルの実装を行った。この手法の場合、入力曲はテンポがある程度一定である必要がある。本システムではユーザーからの入力も対象とするためテンポ変化により早く対応できる機能を実装した。

ユーザーによるキーボードの入力に対してはMIDIのノート番号の情報は用いず、ヘッドフォン出力端子をラインで入力端子と直接繋ぐか、出力された音をマイクから取り込んで、音響信号の解析を行う。

2.1.1 音の立ち上がり成分の抽出

音響信号にFFTを行ない、式(1)を満たす時刻 t における周波数 f のパワー $p(t, f)$ を音の立ち上がり成分とし、そのパワーの値 $d(t, f)$ を式(2)によって定めて条件を満たさない値は0と定める。

$$\min(p(t, f), p(t+1, f)) > prevP \quad (1)$$

$$d(t, f) = \max(p(t, f), p(t+1, f)) - prevP \quad (2)$$

$$prevP = \max(p(t-1, f), p(t-1, f \pm 1)) \quad (3)$$

$d(t, f)$ の合計値を時刻 t における全体の立ち上がり成分のパワー $D(t)$ とする。その合計値のピーク抽出をしたパワーの系列を発音時刻ベクトル $O(t)$ とする。

2.1.2 次のビート時刻の予測

発音時刻ベクトルの自己相関からビートの時間間隔 I を算出する。その際に入力曲の対象範囲である曲のビートの時間間隔である約350~1000msecに該当する自己相関の最大値を与えるずれ(τ)をビートの時間間隔であると定める。また、過去に算出された時間間隔の中で最も多く出現したものを自己相関結果として出力することによって一時的な検出の誤差を抑えることができる。

自己相関の結果から暫定的なビート系列を作り、 $0 \sim I$ の範囲で発音時刻ベクトルとの相互相関を計算する。その範囲で最大値を与える τ を現在時刻から次のビートの時刻までの時間間隔とし、次のビートの時刻を定める(詳細は(1)参照)。

2.1.3 テンポ変化の追跡

ユーザーがキーボードによって音を入力する場合、途中でテンポが変化することも想定される。しかし、前述したモデルは誤差によるテンポの変動を抑える為に過去に算出された自己相関結果の履歴を残すため、テンポが変化した時にそれが出力として反映されるま

[†] 関西学院大学

Kwansei Gakuin University

[‡] シンガポール国立大学

National University of Singapore

でに時間がかかってしまう場合がある。

そこで、自己相関で連続して 5 回誤差の少ない(約 50msec 以内)数値が算出され、その時間間隔 τ が前回出力された τ と約 70msec 以上違う場合は入力テンポが変化した判断し、過去の履歴を消去し、新たに履歴を取り直すようにする。以上の手続きによって、小さなテンポの変化に対しては感度を下げて誤差を抑え、大きなテンポの変化に対しては感度を上げ、より早く対応することが可能となった。

2.1.4 キー入力の強弱情報の保存

キーボードの入力と音の強弱の情報としては MIDI のベロシティの数値を用いる。入力されたキーのベロシティの平均値を 2 秒ごとに算出し、得られた 0~127 の値を 0~2 の 3 段階にスケールし直し強弱情報として保持する。

2.2 ロボットのダンスモーション制御

ロボットは株式会社ニルバーナテクノロジーの Tai-chi(太極)を用いた。

2.2.1 ダンスモーションの格納と再生

ダンスの一連の動作はキーフレームの 22 個の関節の角度をテキストファイル形式で格納しておく。キーポーズから次のキーポーズまでの移行時間として、ビート抽出で算出された時間間隔を次のビートの時刻のタイミングで与え、その間の関節の角度を補間して再生することによってビートとの同期を取る(図 1)

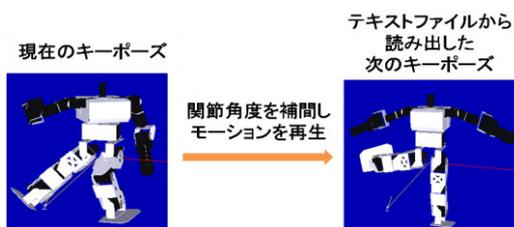


図1 モーションの再生

2.2.2 モーションの選択

今回の実装ではダンスの 4 つのキーポーズを 1 モーションとして、動きの大きさに応じて小, 中, 大に分類しそれぞれ 5 モーションずつデータベースとして用意した(4 ポーズ目はニュートラルな立ち状態する)。なお今回はモーターの角度変化などを基準に主観的に分類を行った。1 つのモーションを再生し終わったらシステムは強弱情報に基づいて次に呼び出すモーションの種類(小(0), 中(1), 大(2))を選択する。その種類の中でどのモーションを呼び出すかはランダムで選択される。こうすることによって毎回、強弱の情報を考慮した違うパターンのモーションが再生される。

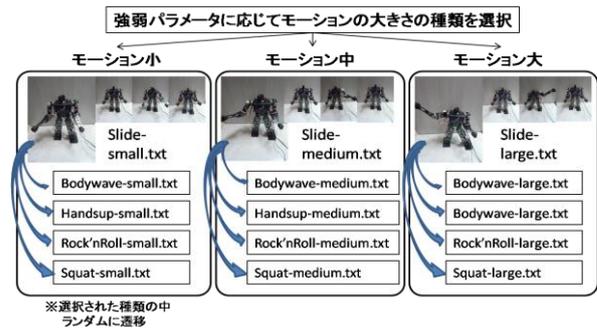


図2 モーションの選択と遷移

3. システムの 2 つのモード

● インタラクティブモード:

ユーザーによってキーボードから入力された音響信号のビート抽出を行い、その情報を用いてロボットのダンスモーションを制御し同期を取る。さらにユーザーの打鍵の強弱に応じて、3 段階にダンスの激しさが変化する。ユーザーはダンスの速度と激しさの 2 つの要素をコントロールできる。

● 鑑賞モード:

ユーザーが選択した音楽のビートに合わせてロボットがダンスを踊る。ユーザーは音楽を聞くだけでなく、それに合わせてダンスを踊るロボットを見ることによって、目と耳の両方で楽しむことができる。

4. おわりに

本稿では、音楽ファイルからビートを抽出し、リアルタイムにロボットのダンスモーションを制御するシステムを開発した。また、ユーザーによる入力音から抽出したビート情報と、キーボードの音の強弱の情報を用いてロボットのダンスモーションのテンポと緩急をリアルタイムに変化させ制御できるシステムを開発した。今後の検討課題としては、より複雑なモーションの結合方法、音楽音響信号からの盛り上がり情報のリアルタイム抽出などがある。

参考文献

- 1) Masataka Goto: An Audio-based Real-time Beat Tracking System for Music With or Without Drumsounds, Journal of New Music Research (2001).
- 2) 中澤篤志, 白鳥貴亮, 池内克史: 観察に基づく音楽およびモーションキャプチャデータからの舞踊動作生成手法, MIRU(2005).
- 3) Kazuyoshi Yoshii, Kazuhiro Nakadai, Toyotaka Torii: A Biped Robot that Keeps Steps in Time with Musical Beats while Listening to Music with Its Own Ears, RSJ International Conference on Intelligent Robots and Systems(2007).