

楽器の内部モデルに基づくフィードフォワード制御による テルミン演奏ロボットの開発

水本 武志[†] 辻野広司[‡] 高橋 徹[†] 尾形 哲也[†] 奥乃 博[†]

[†]京都大学大学院 情報学研究科 [‡](株)ホンダ・リサーチ・インスティテュート・ジャパン

1. はじめに

近年、ヒューマノイドロボットのエンターテイメント分野への応用が行われている。特に、音楽に合わせて動作するロボット(音楽ロボット)に関しては、人との音楽を通したインタラクション [1-3] や人の演奏スキルの解明 [4]、などを目指した研究がある。我々は、音楽ロボットの音楽理解能力が音楽認識能力と音楽表現能力から構成されるとし、汎用のアーキテクチャを設計した [5]。

本稿では、テルミン演奏で音楽表現を行うロボットのための、テルミンの音高制御のモデル化と演奏システムについて報告する。テルミンは物理的接触なしで音高が変化するので、ロボットのハードウェアに依存しない演奏システムを設計できる。この結果、あるロボットで得た演奏技術の別ロボットへの移転が期待できる。従来のテルミン演奏ロボットは音高を目標値としてフィードバック制御されていたが、[6]、目標値への収束に時間がかかるので、時間の制約が厳しいメロディ演奏には不十分である。我々は、ロボットの身体と音高の関係のモデル化に基づくフィードフォワード制御を行い、学習時に音を聞き、演奏時は音を聞かない目標音高演奏を実現した。

2. 従来の音楽ロボット

音楽ロボットの音楽表現は 3 種類に分けられる。

1. 身体動作: 足踏み [1, 2], 社交ダンス [7]
2. 音声: 歌唱 [2], ビートカウント [5],
3. 楽器音: フルート [4], マリンバ [3], テルミン [6].

1. は、最も一般的に用いられる表現だが、評価が主観的になる。定量的評価の為の身体動作の特徴量が提案されてはいるが [8]、客観的な評価は依然困難である。2. は身体の拘束がないので最も自由に表現できるが、ロボットの身体性を十分生かせない。一方 3. はロボットの身体動作と表現が密接に関連しており、表現である演奏音はビートやピッチで評価できる。

従って、我々は演奏音の評価を通した身体動作の評価を期待し、3. 楽器演奏ロボットを採用した。

従来の楽器演奏ロボットには次の 2 点の問題がある。

問題点 1 ハードウェアへの依存性の高さ

従来はロボットの演奏スキルとハードウェアが不可分であるので、あるロボットで得た演奏スキルを別のロボットへ移転するのは困難であった。

問題点 2 音高の制御が中心的な課題

従来の楽器演奏の対象である管楽器や打楽器などはそもそも演奏装置の作成自体が困難であり、その研究が中心であった。そのため、ロボットによる高度な演奏表現に関する研究は少ない。

3. テルミン

3.1 発音の仕組みと特長

まず、テルミンの仕組みについて述べる。標準的なテルミンには 2 本のアンテナがあり、片方が音高制御、もう一方が音量制御に用いられる。本稿では音高だけを扱うので、音量制御の説明は割愛する。

テルミンは、発振周波数がわずかに異なる 2 つの発振回路から構成され、その差が音高になる。周波数差の出力は、両回路の出力信号を乗算し、生じたうなりをローパスフィルタで取り出すことで実現される。片方の発信回路のキャパシタがアンテナの形で外部にあり、演奏者はアンテナと手との距離を変えて静電容量を変化させることで、周波数差を制御する。

音高は静電容量で決まるので、演奏ではハードウェア(ロボットの腕や手)の位置制御が主要な機能となる。つまり、あるロボットで得た位置制御の見解は比較的容易に別ロボットに移転できる。また、音生成は他の楽器に比べて非常に単純であるので、従来より容易に演奏技法や表情付けなどの研究へ取り組むことが可能な楽器である。

3.2 テルミン演奏における問題

テルミンの演奏は鍵盤楽器のように運指を記憶するだけでは演奏できない。なぜなら、手とアンテナの距離と音高の関係(距離-音高特性)は非線形であり、その特性は周囲の環境に応じて敏感に変化するからである。このため、テルミンの演奏は難しく、音楽の経験がない場合は目標音への手の移動すら難しい [6]。ロボットの場合も作り込みの動作では演奏できず、実際の演奏音を聞く必要がある。音階の無い楽器には、他にトロンボーン、バイオリンなどがある。

この問題に対して、本稿では、テルミンの距離-音高特性をモデル化し、モデルに基づくフィードフォワード制御によって、音を聞かずに目標音周辺を出力し、出力音を元にモデルパラメータを適応する手法を開発した。

4. ロボットによるテルミン演奏

4.1 音高制御のモデル化

テルミンの音高は、手とアンテナの距離が離れている間はほとんど変化しないが、距離が小さくなると音高の増加量が増加する。すなわち、距離-音高特性は音高を f 、距離を r として微分方程式 $f = df/dr$ の解と同様になると考えられる。そこで、テルミンの距離-音高特性 f を指数関数でモデル化する。

$$f(x) = p_1 \exp(-p_2 x) + p_3 \quad (1)$$

ここで、 p_1, p_2, p_3 はモデルパラメータであり、 x はロボットとアンテナの距離である。ただし、 x は距離に比例する値であればよく、本稿では腕の関節角度を用いる。

本モデル化により、以下 2 つの効果が期待できる: (1) 音を聞かずに目標音高へ制御するので、メロディ演奏のような時間の制約が厳しい場合も演奏可能。(2) 本モデルはテルミンの全距離-音高特性を記述するのでパラメータの更新で環境変化に追従可能。

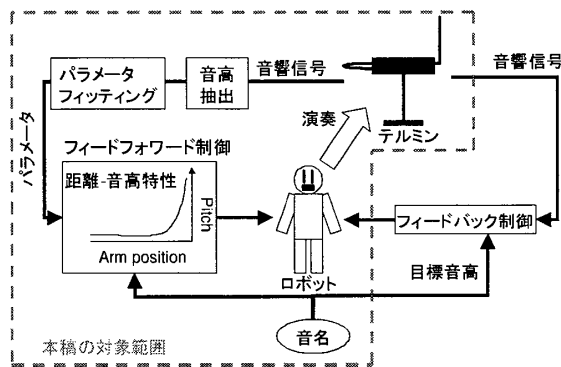


図 1: システム図: 破線で囲んだ範囲が本稿の対象

4.2 演奏システム

テルミン演奏システムの概要を図 1 に示す。ただし、本稿で開発したのは距離-音高特性のモデル化とフィードフォワード制御なので、その部分だけを述べる。システムは以下の 2 フェーズより成る。

学習フェーズ 身体形状を変化させながらテルミンの音を収録し、式 (1) のモデルパラメータを推定する。

演奏フェーズ 与えられた楽譜 (音符列) を平均律[‡] に従って目標周波数に変換しモデルを用いて目標身体形状に変換する。このとき、実際の演奏音を元にパラメータをオンラインで再推定し、環境変化に追従する。

5. 実験

本手法を (1) モデルの汎化性能と (2) 環境の変化に対するモデルの追従能力で評価した。(1) では、ロボットとアンテナの距離を 80 分割して各点でテルミンの音高を測定し、そのうち 5 点を用いてフィッティングした。そして、フィッティング曲線と測定点との誤差を評価した。(2) はテルミンのアンテナ付近に金属製の箱を 3 段階で近づけ、静電場の変化による距離-音高特性を評価した。箱なしを条件 0、箱ありを条件 1~3 (3 が最も近い) とした。

5.1 実験条件

ロボットは HRP2、テルミンは Moog 社の Etherwave Theremin を使用した。ロボットとテルミンとの距離は 1[m]、スピーカとの距離は 2[m]、どちらもロボットの正面に設置した。使用した軸は右肩 Roll 軸、可動域は $\theta = -90 \sim 0$ [deg] とした (0 [deg] が最も近い)。

テルミンの音高推定には、自己相関に基づく基本周波数推定手法を窓長 4096 [pts] で用いた [9]。腕の角度ごとに演奏音を 2 [sec] ずつ収録し、その点の音高は各フレームごとに推定した基本周波数の中央値とした。また、腕を近づけると基本周波数は単調増加するので、制約に合わないデータは除去した。式 (1) の x は関節角度 θ とした。また、フィッティングには最小二乗法を用いた。

5.2 実験結果

(1) の結果を図 2 に示す。測定点を菱形で、使用した点を円で、フィッティング曲線 $f(\theta) = 945.8 \exp(0.07\theta) + 161.8$ を実線で表す。平均絶対誤差は 9.9 [Hz] なので、等間隔であれば 5 点程度でほぼ適切なパラメータを推定できた。

(2) の結果を図 3 に示す。条件 0 から 3 の順で最低周波数が増加するが、最高周波数は変化しないので、金属の

[‡]平均律では、音名 n ($A = 0, A\sharp = 1, \dots, G\flat = 11$)、オクターブ o ($o = 0, 1, \dots$) の音の基本周波数は、 $440 \cdot 2^{o-4} \cdot \sqrt[12]{2^n}$ [Hz] で計算できる。ただし、音名 A4 の基本周波数を 440 [Hz] とした。

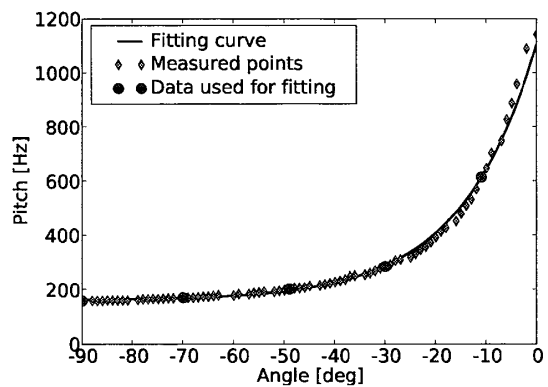


図 2: フィッティング結果: 5 点使用。平均絶対誤差 9.9 [Hz]

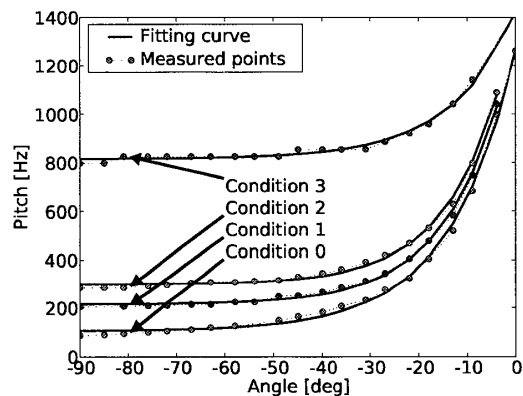


図 3: 環境変化に伴うモデルの変化: 条件 0 は金属箱なし、条件 1-3 の順に箱が接近

接近で音域は狭まった。また、全条件で平均絶対誤差は 8-14 [Hz] に収まったので、特性変化は本モデルで表せたといえる。

6. おわりに

本稿では、テルミン演奏ロボットのためにテルミンの距離-音高特性をモデル化し、有効性を検証した。実験より、本モデルは少数の点でモデルを推定でき、環境による特性変化も表せることを示した。

本モデルの誤差は、フィードバック制御で補正する必要がある。また、本稿と同様の方法で音量制御はモデル化できる。これらした後、音高と音量によるメロディの表情付け、別のロボットによる汎用性の評価を行う。

謝辞 本研究の一部は、科研費、GCOE の支援を受けた。

参考文献

- [1] K. Yoshii 他. A biped robot that keeps steps in time with musical beats while listening to music with its own ears. *IROS-07*, 1743-1750.
- [2] K. Murata 他. A robot uses its own microphone to synchronize its steps to musical beats while scattting and singing. *IROS-08*, 2459-2464.
- [3] G. Weinberg 他. The interactive robotic percussionist - new developments in from, mechanics, perception and interaction design. *HRI-07*, 456-461.
- [4] J. Solis 他. Development of waseda flutist robot WF-4RIV: Implementation of auditory feedback system. *ICRA-08*, 3654-3659, 2008.
- [5] T. Mizumoto 他. A robot listens to music and counts its beats aloud by separating music from counting voice. *IROS-08*, 1538-1543.
- [6] A. Alford 他. A music playing robot. *FSR-99*, 29-31.
- [7] K. Kosuge 他. Dance partner root - ms dancer. *IROS-03*, 3459-3464.
- [8] 中田ら. ロボットの身体動作表現と生成される印象とのラバン特徴量を介した定量的相関分析. *RSJ*, Vol. 19, No. 2, 1-8, 2001.
- [9] A. Camacho. *SWIPE: A sawtooth waveform inspired pitch estimator for speech and music*. PhD thesis, University of Florida, 2007.