

複数自由度を用いて音高特性モデルに基づく音高制御を行うテルミン演奏ロボットの開発

水本 武志[†] 高橋 徹[†] 駒谷 和範[†] 尾形 哲也[†] 奥乃 博[†]

[†] 京都大学大学院 情報学研究科 知能情報学専攻

1. はじめに

我々は、人とロボットの合奏を実現するためにテルミン演奏ロボットを開発している [1]。テルミン (図 1 参照) は、音高アンテナ・音量アンテナと演奏者の手先とのそれぞれの距離を変えて、物理的接触なしで音高・音量を制御する電子楽器である。既開発のシステムでは、テルミン演奏における重要な演奏技法の一つであるビブラートを扱っておらず、楽譜に対応する目標位置の系列のみを考えていたので演奏が単調であった。本稿では、ビブラート演奏を腕の軌道を含めた問題として定式化し、腕の複数自由度を用いて振動を抑制しながらビブラート演奏を実現する。歌声に適切なビブラートを付与すると自然性が向上するという知見 [2] からビブラート演奏による演奏の質の向上見込まれる。ビブラート演奏に伴う腕からの振動は複数自由度制御により抑制されると期待される。また、ビブラートで演奏の音高誤差に対する聞き手の許容範囲が広がるという知見 [3] もあり、音高の制御誤差を抑える効果も期待できる。

2. 従来のテルミン演奏システム

問題設定 L 個の音符から成る楽譜に従ってテルミンを演奏する問題は次のように書ける:

入力: 長さ L の音名 (n_i, o_i) と持続時間 d_i の 3 変数の組の列 $(i = 0, \dots, L-1)$

出力: 入力された音名に対応する手先目標位置列 x_{pi}

ただし、 n_i は音名のインデックス ($C = 0, \dots, B\sharp = 11$), o_i はそのオクターブ (整数) とする。演奏上の課題は、手先位置と音高・音量の関係 (以下音高・音量特性と呼ぶ) が周囲の静電的環境から受ける影響である。そのため、楽譜と x_{pi} の対応は環境ごとに異なる。異なる静電的環境で演奏を容易にするために、少数の測定で 2 つの特性を学習することが要求される。

2.1 音高特性モデル

少数の測定でパラメータの推定が可能な音高特性モデル [1] で特性を学習する:

$$p = M_p(x_p; \theta) = \theta_2 / (\theta_0 - x_p)^{\theta_1} + \theta_3 \quad (1)$$

ただし、 $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ はモデルのパラメータである。パラメータは複数の x_p でのテルミンの音高 p を学習データとして非線形最適化手法で求める [1]。

2.2 フィードフォワード音高制御によるメロディ演奏

与えられた楽譜に対する目標点列 x_{pi} は、式 (1) で求め、それによって腕を制御する。まず、音名列 (n_i, o_i) を

Model-Based Pitch Control Method for a Theremin Player Robot using Multiple Degrees-of-Freedom: Takeshi Mizumoto, Toru Takahashi, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno (Kyoto Univ.)

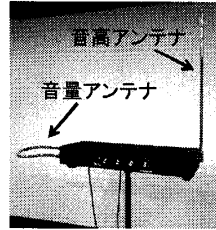


図 1: テルミンの写真



図 2: テルミンを演奏するロボット (HIRO)

平均律[‡] で音高列 p_i に変換する。次に、音高列 p_i に対応する目標位置列 x_{pi} を次の逆音高特性モデルで求める:

$$x_p = M_p^{-1}(p; \theta) = \theta_0 - (\theta_2 / (p - \theta_3))^{1/\theta_1} \quad (2)$$

そして、目標位置列 x_{pi} を次式から求める。

$$x_{pi} = M_p^{-1}(p_i; \theta) \quad \text{where } i = 0, \dots, L-1 \quad (3)$$

ただし、 θ は前節で求めたパラメータである。最後に、 i ごとに d_i だけ停止しながら手先を移動させて、フィードフォワード制御によるメロディ演奏を行う。なお、フィードバック制御は、単純に適用すると音高軌跡が不明瞭になる、すなわち楽譜上の音高に十分な時間滞在できずメロディを明示的に示せないのが、今回は使用していない。

3. 複数自由度を用いたビブラート演奏

3.1 ビブラート演奏の要求条件と設計方針

自然なビブラートは演奏の表情付けに重要だが、同時に、連続的に音高が変化するというテルミンの特性から、演奏音は不明瞭になりやすい。従ってビブラート演奏は、次の 2 つの要求を満たす必要がある: (1) 自然にビブラートを演奏する、(2) 演奏中の音高を十分演奏する。

要求 1 を満たすために、テルミンと同様に連続的な声高が変化する人の歌声に関する知見を用いる。具体的には、歌声のビブラートの振動数は時間と共に増加するという知見 [2] である。次に要求 2 に対しては、音高が変動している音高間の移動中と、目標音高への到達直後において、ビブラート演奏を止めることで聴衆に対して演奏する音高を明示する。

3.2 複数自由度への拡張

前節では x_p を関節角度としたが、複数自由度を用いて体の振動を抑えるため、手先座標 (x, y) を用いる。ここで、手先の範囲を $(x_0, y_0), (x_1, y_1)$ を両端とする線分

$$\begin{cases} x(s) = x_0 + s(x_1 - x_0) \\ y(s) = y_0 + s(y_1 - y_0) \end{cases} \quad (4)$$

[‡] 半音だけ異なる音程の周波数比を $1/\sqrt{2} : 1$ と定義する音律。

の上に拘束する。このとき、媒介変数 $s \in [0, 1]$ を手先位置 x_p とすれば、第 2 章の議論が適用できる。なお、実際の関節角度は逆運動学を解いて求める [4]。

3.3 目標位置関数の設計

x_p の移動速度 v を一定として目標位置系列 (式 (3)) を時間を変数とする関数 $x_p(t)$ に書き直す:

$$x_p(t) = \begin{cases} x_{pi} & \left(D_{i-1} \leq t < D_i - \frac{x_{pi} - x_{p(i-1)}}{v} \right) \\ x_{p(i+1)} - v(D_i - t)(x_{p(i+1)} - x_{pi}) & \left(D_i - \frac{x_{pi} - x_{p(i-1)}}{v} \leq t < D_i \right) \end{cases} \quad (5)$$

ただし、 $D_i = \sum_{n=0}^i d_n$ は i 番目の音高を演奏する時刻を表し、第 1 式が目標位置で停止中の状態を、第 2 式が次の目標位置へ移動中の状態を表す。

3.4 自然なビブラート演奏への拡張

3.1 節の要求 (1) を実現するために、ビブラート演奏を行う範囲を、「ある音程を出力し始めて時間 T_{th} 後から次の音高に移動を開始するまで」に限定する。次に要求 (2) を実現するために、ビブラートの軌跡には振動数 f 、振幅 a の正弦波を用い、 f, a を次の通り定める。振幅 a は対数軸で音高を考えたときに振幅が一定になるように、現在の音高に依存して次のように定める:

$$a_i = M_p^{-1}(\mu p_i; \theta) \quad (6)$$

ただし、 $\mu \in [0, 1]$ は振幅を定めるパラメータである。振動数 f は、人の歌声では時間と共に増加するので [5]、線形増加 $f_i = b(t - D_i) + c$; とする。これらの結果、式 (5) を次のように修正する:

$$x_p(t) = \begin{cases} x_{pi} & (D_{i-1} \leq t < T_{th}) \\ a_i \sin(f_i t) - a_i + x_{pi} & (T_{th} \leq t < D_i - \frac{x_{pi} - x_{p(i-1)}}{v}) \\ x_{p(i+1)} - v(D_i - t)(x_{p(i+1)} - x_{pi}) & (D_i - \frac{x_{pi} - x_{p(i-1)}}{v} \leq t < D_i) \end{cases} \quad (7)$$

4. 人の演奏との比較による評価実験

実験条件 本手法を川田工業のロボット HIRO (図 2) と HRP-2 に実装した。テルミンには Moog Music 社の Etherwave Theremin を、音声収録はロボット頭部のマイクを用いた。パラメータの学習サンプル数は 12, $\mu = 0.99$ (17.4 [cent] の振幅に相当), $T_{th} = 0.5$ [sec] とした。

実験 1 では、体の振動が強い HRP-2 と少ない HIRO について、C4 と D4 を交互に演奏する場合で比較する。実験 2 では、キーボード演奏経験のある被験者と HIRO が Schubert の Ave Maria を演奏し、音高軌跡を比較する。

実験 1: 複数自由度制御の有無による比較 図 3 に結果を示す。上段が HRP-2、中段が HIRO、下段が正解である。HRP-2 の演奏は、音高軌跡の振動が顕著に見られる。これは、HRP-2 の比較的重い腕を、1 つの関節角のみを用いて急激に動かしているからである。一方、HIRO の演奏には振動が見られず、正解とほぼ等しい演奏ができている。このことから、複数自由度を用いて振動を抑えれば意図しないビブラートを抑制できることが分かる。

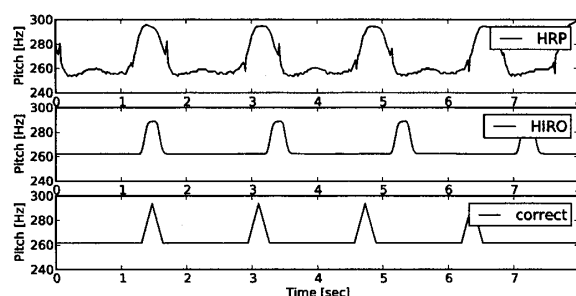


図 3: 複数自由度の効果

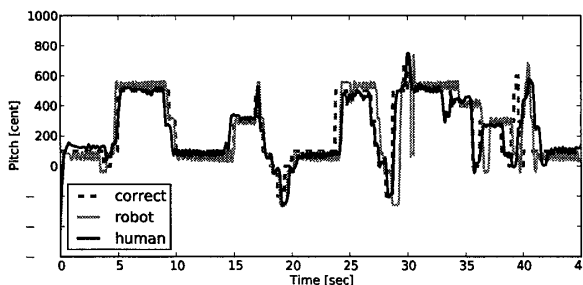


図 4: 人とロボットの音高軌跡の比較

実験 2: 人の演奏との比較 図 4 に音高軌跡を示す。横軸が時間、縦軸が 220 [Hz] を基準とした音高 ([cent]) である。赤線がロボットの、黒線が人の演奏軌跡を示す。また、青い破線は楽譜から計算した正解である。

HIRO の場合は手の移動が位置制御精度に影響しないので、ビブラート演奏の期間が長い。一方、人の場合は手を動かすと位置がずれるので振幅と周期に揺らぎが見られる。従ってロボットの方がビブラートが安定している。楽器演奏では安定したビブラートが好ましいとされるが、人の自然な歌声の特徴の一つに音高の微細振動があるという報告 [2] もあるので、ロボットの腕の位置に揺らぎを加えることで自然な演奏になる可能性がある。

5. おわりに

本稿では、テルミン演奏ロボットにビブラート演奏機能を実現するため、手先制御の目標を点列から軌跡に拡張した。実験では振動を抑えることで音高制御の精度が改善されることと、本手法で被験者よりも安定したビブラート演奏ができることを確認した。今後の課題は、フィードバック制御による音高制御の精度向上や、プロのテルミン奏者による演奏との比較、被験者実験等による見た目と演奏音の自然性の評価である。

謝辞 本研究の一部は、科研費、GCOE、CrestMUSE、HRI-JP の支援をうけた。

参考文献

- [1] T. Mizumoto 他. Thereminist robot: Development of a robot theremin player with feedforward and feedback arm control based on a theremin's pitch model. *IEEE/RSJ IROS09*, pp.2297-2302.
- [2] 斎藤 他. 自然性の高い歌声合成のためのヴィブラート変調周波数の制御法の検討. *信学技報* 2005-09, pp.13-18.
- [3] ヨハン・スンドベリ. 歌声の科学. 東京電機大学出版局, 2007.
- [4] 吉川恒夫. ロボット制御基礎論. コロナ社, 1988.
- [5] E. Prame. Measurements of the vibrato rate of ten singers. Technical report, KTH, 1992.