

# 相平面確率モデルを利用した歌唱・楽器演奏 の基本周波数軌跡の分析\*

加古達也 (名大), 大石康智, 亀岡弘和, 永野秀尚, 柏野邦夫 (NTT), 武田一哉 (名大)

## 1 はじめに

歌声の音響的特徴として、歌唱フォルマントや基本周波数 ( $F_0$ ) の動的変動成分といわれるビブラート、オーバーシュート、プレバレーション、微細変動成分などの存在が明らかにされ、これらが歌声知覚や個人性知覚に影響を与えることが報告されている [1]。これまで、我々は歌声の  $F_0$  動的変動成分に着目し、これに現れる多様な歌唱様式をモデル化するために相平面確率モデルを提案した [2]。相平面確率モデルとは、 $F_0$  系列とその時間微分で構成される相平面上に  $F_0$  の軌跡を描くことでその動的変動成分を可視化し、軌跡の分布を確率分布として表現する手法である。この相平面確率モデルを利用すれば、歌唱者ごとの歌唱様式、技術の違いを識別可能であることを確認した。本稿では、この相平面確率モデルを歌声だけでなく楽器演奏音へ適用する。楽器音を分析する上で重要な特徴量として音色、高調波成分の豊富さ、周波数の変動の幅、ビブラートなどが挙げられており [3]、本稿では、各楽器の音高軌跡に着目する。そして、その動的変動成分を相平面確率モデルを利用して分析する。その結果、楽器のビブラートなどの動特性の違いを可視化し比較できることを確認した。

## 2 相平面を利用した $F_0$ 軌跡の確率表現

### 2.1 $F_0$ の相平面表現

歌声や楽器の  $F_0$  の性質を調べるために相平面を利用する。相平面とは微分方程式の大局的な解の性質を調べるために用いられる手法であり、 $F_0$  と  $F_0$  の 1 階微分や 2 階微分によって構成される平面である。 $F_0$  の 1 階微分は式 (1) のように微小区間の回帰係数  $\Delta F_0$  で近似し、同様に 2 階微分は  $\Delta F_0$  の回帰係数  $\Delta\Delta F_0$  で近似する。

$$\Delta F_0(n) = \frac{\sum_{k=-K}^K k \cdot F_0(n+k)}{\sum_{k=-K}^K k^2} \quad (1)$$

この  $F_0$ ,  $\Delta F_0$ ,  $\Delta\Delta F_0$  の 3 つの成分に着目し、相平面上に  $F_0$  軌跡を表現する。相平面上に描かれる軌跡の渦の中心は歌唱者の意図する目標音高であり、周期的に変動するビブラートであれば目標音高を中心に円を描く軌跡が観測できる。一方、オーバーシュートは螺旋を描きながら目標音高へ収束する動きとして観測できる。また、 $F_0$  と  $\Delta\Delta F_0$  で構成される相平面上の軌跡は負の方向に傾いた直線が描かれ、この傾きはビブラートの深さを表し、 $\Delta F_0$  と  $\Delta\Delta F_0$  で構成される相平面上にはビブラートの深さと周期から決まる楕円が観測される。

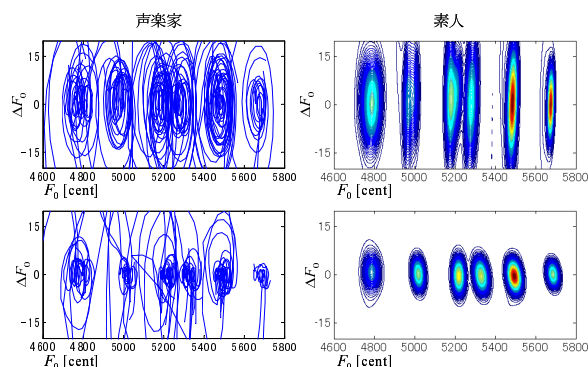


Fig. 1 相平面における声楽家と素人の歌声の  $F_0$  軌跡 (上段) と相平面確率モデル (下段)

### 2.2 相平面確率モデル

相平面上の渦軌跡の形状をパラメトリックにモデル化する。相平面上に描かれる  $F_0$  軌跡は、複数の目標音高を中心に分布している。この分布を確率密度関数で表現するために、混合ガウスモデル (GMM) を用いる。GMM は以下のように表される

$$\sum_{m=1}^M \lambda_m \mathcal{N}(\mathbf{f}_0(n); \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m), \quad (2)$$

ここで、 $\mathbf{f}_0(n) = [F_0(n), \Delta F_0(n), \Delta\Delta F_0(n)]^T$  であり、 $\mathcal{N}(\cdot)$  はガウス分布、 $\Theta = \{\lambda_m, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m\}$  はモデルパラメータである。各パラメータの  $\lambda_m$ ,  $\boldsymbol{\mu}_m$ ,  $\boldsymbol{\Sigma}_m$  はそれぞれガウス分布の重み、平均ベクトル、全共分散行列である。プロの声楽家女性と音楽経験のない素人女性の歌声の  $F_0$  の軌跡と  $(F_0, \Delta F_0)$  構成の相平面確率モデルを図 1 に示す。声楽家、音楽経験のない素人の相平面確率モデルを比較すると、声楽家では深いビブラートを使った歌唱法のためガウス分布の垂直方向に大きな分散が確認できる。一方、素人では音程が安定しない歌唱法のため水平方向に大きな分散が確認できる。文献 [2] では、このモデルを用いて、歌唱者ごとの歌唱様式、技術の違いを自動識別可能であることを確認した。

## 3 相平面確率モデルを利用した楽器演奏の $F_0$ 軌跡の分析

### 3.1 $F_0$ 推定法

相平面確率モデルを利用して、楽器演奏の  $F_0$  軌跡を分析する。楽器演奏の音源は RWC 研究用音楽データベースの楽器音データベースを利用した [4]。 $F_0$  は、後藤らの提案する  $F_0$  推定手法 [5] を利用した。 $F_0$  推定の分析条件として、サンプリング周波数は 16 kHz、フレーム長は 64 ms、分析窓関数はハニング窓、切り

\* Analysis of  $F_0$  Dynamics in Singing Voice and Instrument Sound Based on Stochastic Phase Plane Model by KAKO Tatsuya(Nagoya Univ.), OHISHI Yasunori, KAMEOKA Hirokazu, NAGANO Hidehisa, KASHINO Kunio(NTT), and TAKEDA Kazuya(Nagoya Univ.)

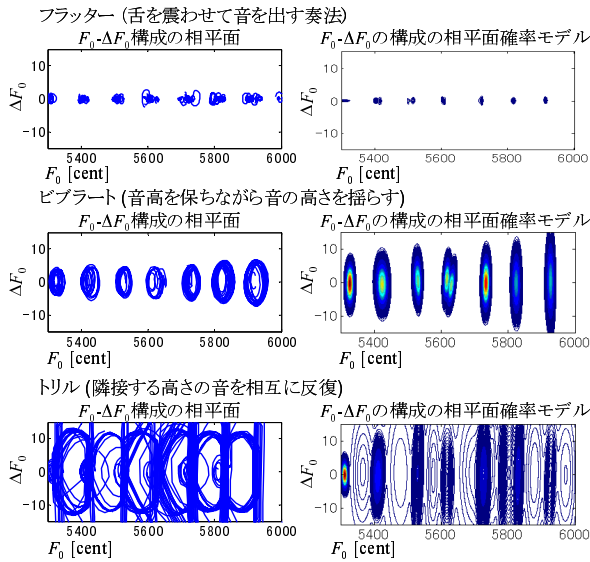


Fig. 2 フルートの各演奏技法での  $F_0$ - $\Delta F_0$  構成の相平面と相平面確率モデル

出し窓のシフト長は 10 ms とした．また  $F_0$  を平滑化するため窓長 50ms の移動平均フィルタを施した．また式 (1) の回帰係数は  $K = 2$  として  $\Delta F_0$ ,  $\Delta\Delta F_0$  を算出した．

### 3.2 奏法技法によって異なる $F_0$ 動特性の分析

データベース番号 No.33 フルーツは、4 種類の技法で演奏が収録されている．舌を震わせて音を出す技法のフラッターや、目標音高を保ちながら音の高さを揺らすビブラート、隣接する音高を相互に反復させるトリル、音を短くきって演奏するスタッカートである．そこで、フラッター、ビブラート、トリルからなる 3 つの技法の  $F_0$  軌跡から、 $F_0$ - $\Delta F_0$  構成の相平面確率モデルを作成した (図 2)．フラッター、ビブラート、トリルの順番で、相平面に描かれる渦軌跡の形状は大きくなった．フラッターは、目標音高の付近で小刻みに振動した．ビブラートは、目標音高を保ちながら音の高さを上下に揺らすため、目標音高付近で楕円状の軌跡が描かれた．トリルは、隣接する音高を交互に遷移するため、ビブラートに比べて、大きな楕円状の軌跡が観測された．さらに、GMM を用いた相平面確率モデルによって、これらの渦形状の違いをパラメトリックに表現できることを確認できた．ただし、ビブラートやトリルでは、渦の中心が空洞化しているため、GMM では表現できず、相平面の軌跡のモデル化方法について検討する必要がある．

### 3.3 楽器によって異なる $F_0$ 動特性の分析

データベース番号 No.45 歌声, No.15 ヴァイオリン (擦弦楽器), No.09 クラシックギター (撥弦楽器), No.21 トランペット (管楽器) のビブラートの演奏における  $F_0$  の動特性を分析する．図 3 より、相平面の渦形状から、ビブラートに関する発音機構の自由度の違いを観測できる．例えば、歌声のビブラートは、人間の複雑な発音機構に基づくため、最も制御の自由度が高く、相平面上に大きな渦軌跡が描かれる．歌声の相平面確率モデルでは、楽器とは異なる分散の大きな分布が確認できる．擦弦楽器であるヴァイオ

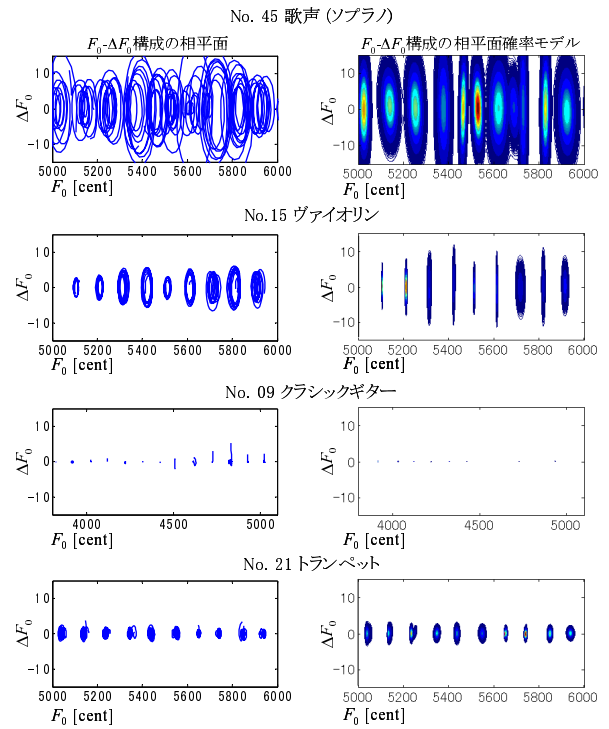


Fig. 3 各楽器の  $F_0$ - $\Delta F_0$  構成の相平面と相平面確率モデル

リンのビブラートは規則的な音高の振動があり、中心が空洞化した楕円の軌跡が観測できる．加えてビブラートをかける指の位置によって、制御の自由度が異なるため、相平面に現れる渦軌跡の広がり为目标音高によって異なるといえる．クラシックギターは、弦を弾いて発音するため、自由度が低く音高の変化が少ない．そのため相平面上では軌跡が一点に集中している．最後に、トランペットでは音高を上げる際に呼気の量を増やす必要があるため、音高が上がるにつれて音の変化を付けることが難しく、ビブラートが困難になる．そのため音高が上がるにつれ相平面確率モデルの分布の分散が小さくなると考えられる．このように楽器は発音機構によって、そのビブラートの制御の自由度は異なり、その動特性の違いを相平面確率モデルを利用することで可視化し比較することが可能となる．

## 4 まとめ

本稿では、歌声に利用していた相平面確率モデルを楽器演奏の分析に利用した．相平面確率モデルを用いることで発音機構の異なる楽器の動特性の違いを可視化し比較することができた．今後は、半音ごとの単音演奏ではなく、実際の楽曲演奏音源の  $F_0$  軌跡を確認する．また楽器の発音機構の違いから生じる動特性の違いを利用して、楽器の発音機構に基づく分類を行う予定である．

## 参考文献

- [1] 齋藤 他, 音講論, 1-R-21, 457-458, 2008 .
- [2] Kako *et al.*, ISMIR, 393-397, 2009 .
- [3] 北原 他, 音情研, 40 (2), 7-14, 2001 .
- [4] 後藤 他, 音講論 (春), 843-844, 2003 .
- [5] 後藤 他, 信学論, 83 (11), 2330-2340, 2000 .