

合唱における歌声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

歌声情報処理に関する多くの研究 [1] では, これまで歌唱者が単独で歌った独唱を研究対象とした. 一方で, 複数の歌唱者が同時に歌唱する斉唱や合唱を対象とする研究は少ない [2]. 本稿では合唱の基本周波数 (F_0) 軌跡に着目した分析を行う. 合唱の場合, 歌唱者は他者の歌声を聴きながら歌唱するため, 自身の歌声を他者の歌声に追従させる動作 (引き込み動作) が見られる. この引き込みに関する先行研究として, 引き込みにかかる時間に着目し, 歌唱技量を評価した研究 [3] があるものの, この引き込み動作を詳細に記述するモデルの構築には至っていない. そこで, 他者の歌声を聞きながら歌ったときに, 歌唱者にどのような引き込み動作が見られるか観察するための歌声収録実験を行い, その歌声の F_0 軌跡を分析する. そして, バネ質量モデルに基づいて, この引き込み動作を記述する合唱歌唱モデルを提案する. 最後に, 歌唱訓練を受けたことのある経験者と素人歌唱者で, 引き込み動作がどのように異なるか実験的に調査する.

2 合唱歌唱の F_0 軌跡

2.1 合唱歌唱の収録

引き込み動作を分析するための歌声を収録した. 歌唱力の異なる歌声を収録するために, 大学の混声合唱団に所属するバスパートの男性歌唱者 (歌唱訓練歴平均 4.5 年, 以後, 経験者と呼ぶ) 4 名と歌唱訓練を受けたことのない男性学生 (以後, 素人と呼ぶ) 3 名に協力を依頼した. 収録場所は低残響室とし, 歌唱者にはヘッドホンを装着させ, 右耳から伴奏音を左耳からは自身の歌声と他者の歌声が聴こえるように設定する. 歌唱曲は「ふるさと」とし, その 1 番歌詞のバスパートの斉唱を収録する. まず初めに, 経験者 1 名が伴奏とプロの音楽家の歌声をヘッドホンで聴きながら歌唱した歌声を収録する. この歌声を随伴歌唱と呼び, ヘッドホンの左耳から流れてくる他者の歌声として利用する. そして, 残りの 3 名の経験者と 3 名の素人がそれぞれ, 随伴歌唱と伴奏音をヘッドホンで聴きながら歌唱した歌声を 3 回収録した. 同時に, 伴奏音だけを聴きながらの歌声 (独唱) も 3 回収録した.

2.2 F_0 軌跡の分析

収録した歌唱データから F_0 を推定する. F_0 は, 後藤ら [4] の提案した F_0 推定手法を利用した. サンプリング周波数は 16 kHz, 窓関数はハニング窓, 窓長は 64 ms, 窓シフト幅は 10 ms, F_0 平滑化のための移動平均窓長は 50 ms とした. 抽出した F_0 は対数周波数 cent で表す. 受聴した随伴歌唱の影響による F_0 の引き込みを分析するため, 歌唱者の歌声, 随伴歌唱, 楽譜に記載される旋律の, 各々の F_0 の差分を

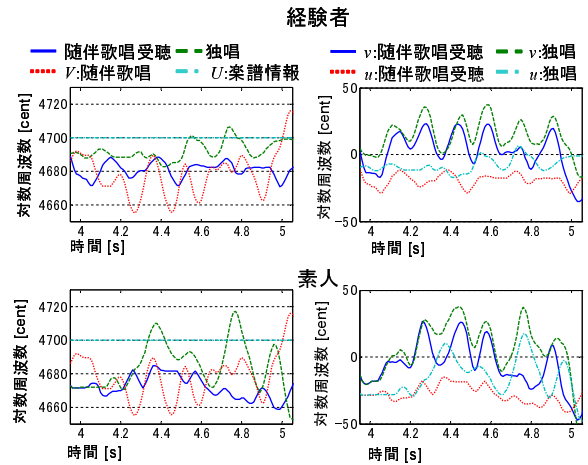


Fig. 1 経験者と素人の随伴歌唱受聴時と独唱の歌声の F_0 軌跡及び u, v 軌跡

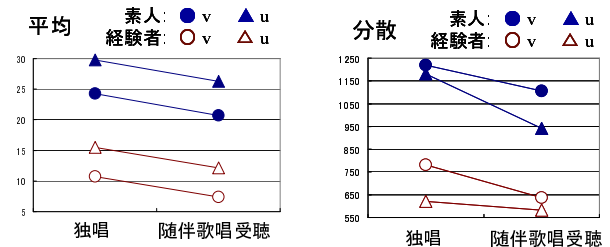


Fig. 2 独唱時と随伴歌唱受聴時の u と v の平均と分散の推移

以下のように求める.

$$v(t) = F_0(t) - V(t) \quad (1)$$

$$u(t) = F_0(t) - U(t) \quad (2)$$

V は随伴歌唱の F_0 , U は楽譜に記載される旋律の音高 (楽譜情報と呼ぶ) を表す. 随伴歌唱を受聴して収録した経験者と素人の歌声の F_0 , u , v の軌跡を図 1 に示す. 経験者, 素人ともに独唱時に比べて随伴歌唱を受聴して歌うことで, 随伴歌唱に近い F_0 軌跡が観測される. 随伴歌唱を受聴した斉唱と, 独唱の u と v の平均と分散の値を図 2 に示す. 経験者, 素人ともに随伴歌唱を受聴することで u と v の平均値が 0 に近づき, 分散の値も小さくなり随伴歌唱の音高に近づいた歌唱を行っていることが分かる. また随伴歌唱を受聴することで u が 0 に近づいたのは, 随伴歌唱がより楽譜情報に近い F_0 軌跡のためである.

3 合唱歌唱モデルを用いた分析

3.1 合唱歌唱モデル

合唱時に他方の歌唱する音高に F_0 が引き込まれる動作が確認できた. このことから, 合唱時の歌唱モデ

* Analysis of F_0 trajectory in Chorus by KAKO Tatsuya(Nagoya Univ.), OHISHI Yasunori, KAMEOKA Hirokazu, NAGANO Hidehisa, KASHINO Kunito(NTT), and TAKEDA Kazuya(Nagoya Univ.)

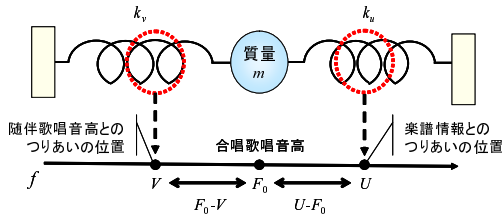


Fig. 3 合唱歌唱のパネモデル

ルには他の歌唱者の音高を無視することはできない。特に、歌唱経験が少なく、安定した目標音高をもたない素人歌唱者は、他者の音高の影響を受け、強く引きこまれる。一方、歌唱訓練を受けた経験者は、楽譜情報に相当する安定した目標音高を持つため、素人に比べて、他者の歌声の影響を受けにくいという仮説が立てられる。そこで歌唱者の F_0 は楽譜情報と他者の歌声の音高によって駆動するパネ質量モデル (図 3) に従うと仮定し、以下の合唱歌唱モデルを提案する。

$$m \frac{d^2 F_0}{dt^2} = \lambda \frac{dF_0}{dt} + k_u(U - F_0) - k_v(F_0 - V) \quad (3)$$

V と U が外力として加わり、パネ定数に相当する k_v と k_u はどちらの音高を重要視するかの割合となる。

3.2 モデルパラメータの推定

前節で立てた仮説を検証するために、式 3 を変形し、以下の係数 $\epsilon, a_1, a_2, a_3, a_4$ を観測可能な $F_0, \Delta F_0, \Delta\Delta F_0, V, U$ から重回帰分析によって推定し、経験者と素人で違いを調査する。

$$\frac{d^2 F_0}{dt^2} = \frac{\lambda}{m} \frac{dF_0}{dt} - \frac{(k_u + k_v)}{m} F_0 + \frac{k_v}{m} V + \frac{k_u}{m} U \quad (4)$$

$$\frac{d^2 F_0}{dt^2} = \epsilon + a_1 \frac{dF_0}{dt} + a_2 F_0 + a_3 V + a_4 U \quad (5)$$

ϵ は誤差項に対応する。 ΔF_0 は推定した F_0 から式 6 の回帰係数を用いて求めた。 $\Delta\Delta F_0$ も同様に ΔF_0 の微小区間の回帰係数を用いて求めた。

$$\Delta F_0(n) = \frac{\sum_{k=-2}^2 k \cdot F_0(n+k)}{\sum_{k=-2}^2 k^2} \quad (6)$$

重回帰分析を行う前に、各特徴量を各々の標準偏差で割り特徴量の正規化を行った。随伴歌唱受聴時の歌唱から重回帰分析によって求めたパラメータの推定結果を表 1 に示す。独唱歌唱の重回帰分析の結果を表 2 に示す。独唱に対しては、 V の項の係数 a_3 は推定しない。随伴歌唱受聴時の $a_4/(a_3 + a_4)$ の値が素人よりも経験者が高い値となった。このことから経験者は、たとえ随伴歌唱を聴きながら歌唱したとしても、自身のもつ目標音高 (楽譜情報) に重きをおいて歌唱すると言える。また独唱に比べて随伴歌唱を受聴することで a_4 の値が減少している。楽譜情報のみを意識して歌った独唱時と比べて、随伴歌唱を受聴することで楽譜情報を意識して歌唱する重みが減った。 ΔF_0 の係数である a_1 の絶対値は素人に比べ経験者が低い値であり、 a_1 は粘性項に相当するため、経験者の F_0 が減衰せずピブラートを維持する歌唱法であることが分かる。

Table 1 随伴歌唱受聴での歌唱の重回帰分析結果

	ϵ	a_1	a_2
素人 (随伴歌唱)	0.669	-0.140	-2.017
経験者 (随伴歌唱)	2.885	0.009	-3.960
	a_3	a_4	$a_4/(a_3 + a_4)$
素人 (随伴歌唱)	0.684	1.300	0.655
経験者 (随伴歌唱)	0.649	3.259	0.834

Table 2 独唱歌唱の重回帰分析結果

	ϵ	a_1	a_2	a_4
素人 (独唱)	-0.106	-0.149	-1.910	1.860
経験者 (独唱)	1.384	-0.067	-3.444	3.396

3.3 合唱歌唱モデルの評価

観測可能な $F_0, \Delta F_0, \Delta\Delta F_0, V, U$ の重回帰分析によって、経験者と素人による歌唱では、ヘッドフォンから受聴する楽譜情報と随伴歌唱への注視の比重が異なることを確認した。そこで本節では、相平面確率モデル [5] を用いて、 $F_0, \Delta F_0, \Delta\Delta F_0, V, U$ の同時確率分布に基づく経験者と素人の合唱歌唱自動識別器を作成する。具体的には $F_0, \Delta F_0, \Delta\Delta F_0, V, U$ を用いて、これらの特徴量の同時確率分布を学習し、経験者と素人、かつ随伴歌唱受聴有無の 4 分類の識別実験を行った。経験者と素人で各々 2 名を学習に用い、残りの 1 名で評価を行った。同時確率分布は混合ガウスモデル (GMM) を利用して学習される。共分散行列は、対角共分散行列とした。ただし、対角共分散行列では各特徴量間の相関情報を表現できないため、GMM の混合数を 4 から 128 まで変化させて実験を行った。識別の尺度として、GMM から計算される対数事後確率を用いた。その結果、混合数が 8 以上のときに、4 分類の識別が 100% 可能であることがわかった。混合数が 4 のとき識別誤りが得られた理由は本実験で用いる楽曲に含まれる音符数が混合数に比べて多いためである。この実験結果から、特徴量 $F_0, \Delta F_0, \Delta\Delta F_0, V, U$ の同時確率分布から、歌唱者および歌唱法 (独唱と合唱) を識別可能であることが示唆された。

4 まとめ

合唱の F_0 の値を分析し、楽譜情報と他者の歌声の音高によって駆動するパネ質量モデルに基づく合唱歌唱モデルを提案した。合唱歌唱モデルの各パラメータを推定し、経験者と素人で楽譜情報と随伴歌唱への注視の比重の異なりを確認した。また相平面確率モデルを用い歌唱法の違いが識別可能であることを示した。

参考文献

- [1] 後藤 他, 情処論, MUS-86, No.4, 2010.
- [2] 桑原 他, JAIST, 修士論文, 2010.
- [3] 野田 他, JAIST, 修士論文, 2008.
- [4] 後藤 他, 信学論, 83 (11), 2330-2340, 2000.
- [5] Kako *et al.*, ISMIR, 393-397, 2009.