

科学研究費助成事業（学術研究助成基金助成金）研究成果報告書

平成25年 6月 5日現在

機関番号：13901

研究種目：挑戦的萌芽研究

研究期間：2011 ～ 2012

課題番号：23650088

研究課題名（和文）

相平面確率モデルを用いた合唱歌唱における相互作用の動的特性の分析

研究課題名（英文）Dynamical analysis of the F0 entrainment in chorus singing Through stochastic phase plane.

研究代表者

武田 一哉 (TAKEDA KAZUYA)

名古屋大学・情報科学研究科・教授

研究者番号：20273295

研究成果の概要（和文）：合唱では複数の歌唱者が同時に歌唱を行うため、他者の歌声から影響を受け、独唱とは異なった歌唱となる。本報告では、そのような合唱における歌唱者間の相互作用を歌声の‘引き込み’という観点から分析する。その第一段階として、歌声の音高(基本周波数, F0)に着目し、F0の動的な変動成分(ダイナミクス)をモデル化することを試みる。モデルパラメータは最急降下法に基づいて推定され、歌唱者の特徴はモデルパラメータによって特徴付けられる。評価実験では、歌唱力の異なる歌唱者ではF0ダイナミクスの特徴が異なると考えられるため、推定されたモデルパラメータを用いて歌唱者識別実験を行い、モデルの有効性を示した。

研究成果の概要（英文）：Unlike solo singing, singing behavior in a chorus is affected by other singers, therefore, acoustical characteristics may differ from that of solo singing. Such interaction between singers in a chorus is analyzed from the view point of ‘entrainment’ in singing behavior. As the first step, a model of F0 dynamics, i.e., the dynamic properties of fundamental frequency of singing voices in chorus is studied. The model parameters are estimated using a steepest decent method. The characteristics of the F0 dynamics of a voice singing in a chorus are parameterized. Preliminary experiments also show that experienced and novice singers can be correctly identified using the parameters of our model, because their entrainment behaviors are significantly different. It is confirmed that the our model can properly characterize individual singers by analyzing F0 entrainment during chorus singing.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
交付決定額	2,900,000	870,000	3,770,000

研究分野：総合領域

科研費の分科・細目：情報学・知覚情報処理

キーワード：①音楽情報処理 ②行動モデル ③相平面確率モデル

1. 研究開始当初の背景

合唱では、複数の歌唱者が同時に歌唱を行うため、他者の歌声から影響を受け、独唱とは異なった歌唱となる。例えば、独唱時と合唱時では、音色に関する音響特徴である歌唱ホルマント(singer's formant)のパワーが異なること、合唱では他者の音圧に合わせて歌唱

することが先行研究により明らかになっている。また、他者の歌声の音圧が非常に大きいと自身の音高を精確に制御することが困難となり、歌唱のパフォーマンスに影響を与えることも報告されている。このような歌唱の違いは、合唱歌唱を分析する上で重要な要素であり、歌声情報処理にとって重要な課題

である。

このような他者の行動に適応するような変化は、行動の「引き込み」と解釈することができ、音楽情報処理の分野においても様々な側面から研究されている。引き込みは、工学の分野において相互作用により振動体の位相や周波数が一致する現象として知られており、2つ以上の現象が同期に向かうことによる振る舞いの変化と定義することができる。様々な文献において、演奏者と聴取者の呼吸のリズムの相互作用や、ボーカルとギターによる共同演奏におけるリズム及び周波数の引き込みについて分析している。合唱に関する研究では、歌声の音高の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がある。このように、合唱や合奏のような共同演奏における演奏者間の関係について様々な研究が行われており、合奏時における演奏者のリズムの制御を数学的にモデル化した研究もある。しかしながら、合唱における歌唱の相互作用をモデル化する研究はほとんど行われていない。

2. 研究の目的

本研究では合唱における歌唱を特徴づける数学的モデルを構築することを目指す。その第一段階として、歌声の基本周波数(F0)に着目し、合唱におけるF0の動的な特徴(F0ダイナミクス)をモデル化することを目的とする。歌声のF0軌跡に含まれる動的な特徴には、ビブラートやオーバーシュートなどがある。これらのF0の動的な特徴は、音楽表現において主要な役割を果たす。合唱におけるF0ダイナミクスをモデル化することは、合唱歌唱の振る舞いを特徴付けるための重要な課題の一つである。

3. 研究の方法

(1) 合唱歌唱データの収録

合唱における歌声のF0の動的変動成分を歌声の引き込みという観点から議論するため、次のような条件の下、斉唱における6名の歌声を収録した。歌唱者は、大学の混成合唱団に所属する歌唱訓練歴3-7年の男性(経験者)3名と、歌唱訓練を積んでいない男性(素人)3名である。歌唱曲は「ふるさと」とし、その1番のバスパートを用い、低残響室において収録した。各歌唱者はそれぞれ、事前に収録した同じパートの歌声(随伴歌唱)と伴奏をヘッドホンで聴きながら3回ずつ歌唱した。伴奏音は、合唱パート練習用のChorus ONTA Vol.10のCDに収録された音源を用いた。随伴歌唱は、3名とは別の経験者が伴奏と同CDに収録された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である。また、伴奏をヘッドホンで聴きながら歌唱した独唱の歌声も収録した。つまり合計で、6名の

表1 F0推定の実験条件

サンプリング周波数	16 kHz
量子化ビット数	16 ビット
窓関数	ハニング窓
窓幅	64 ms
シフト長	10 ms
F0平滑化(移動平均)窓長	30ms

各歌唱者の独唱3回と随伴歌唱受聴時の歌唱(合唱)3回を収録した。

(2) F0軌跡の分析

収録した歌唱データを用いて独唱時と合唱時におけるF0軌跡の違いについて分析を行う。歌声のF0軌跡は、河原らの提案したTANDEM-STRAIGHTを利用して推定した。F0推定の条件を表1に示す。なお、本稿ではF0の周波数[Hz]は、式(1)を用いて対数スケールの周波数[cent]に変換する。

$$f[\text{cent}] = 1200 \log_2 \frac{f[\text{Hz}]}{440 \times 2^{\frac{3-5}{12}}} \quad (1)$$

このように対数スケールに変換することで、半音が100 centに対応する。

図1は、独唱時及び合唱時におけるF0軌跡と、その相平面表現である。図1には、楽譜の旋律情報及び随伴歌唱のF0軌跡も示してある。F0-ΔF0構成の相平面の図より、独唱時と比較し、合唱時の渦軌跡の中心が随伴歌唱の中心に近づいたことがわかる。これは、合唱歌唱においてF0の引き込みが生じた典型的な結果であると考えられる。そこで随伴歌唱を受聴することによるF0の変化を議論するために、独唱及び合唱における歌声の音高(F_0)と随伴歌唱音高(V)との差を計算した。

$$d(t) = V(t) - F_0(t) \quad (2)$$

素人と経験者のカテゴリ別に計算した $d(t)$ の平均と標準偏差を表2に示す。平均と分散についてt検定及びF検定を行った結果、独唱時と合唱時の間に1%の有意水準で有意差があることが確認された。この結果より、合唱歌唱では随伴歌唱を受聴することによる影響が生じることが明らかとなった。また、合唱における $d(t)$ の標準偏差が減少したことから、統計的に歌唱者の歌声の音高は随伴歌唱に近づくといえる。

(3) 合唱における歌声のF0ダイナミクスのモデル化

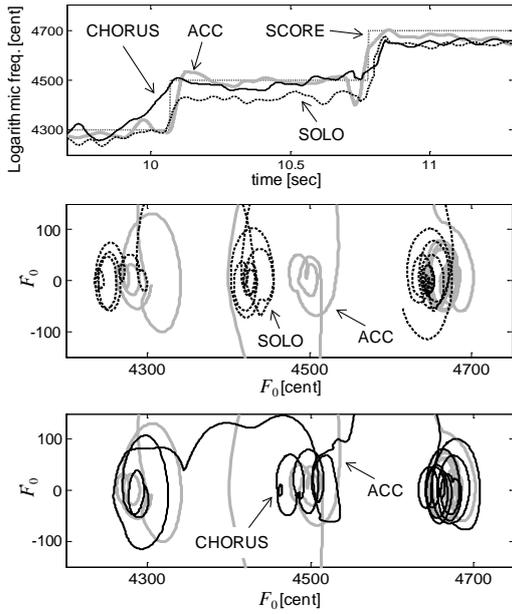


図1 独唱時及び合唱時における F0 軌跡, 上段: F0-ΔF0 構成の相平面上に描かれる独唱時の F0 軌跡. 点線: 楽譜の旋律情報(SCORE), 実線(グレー): 随伴歌唱音高(ACC), 破線: 独唱時の F0 (SOLO), 実線(黒): 合唱時の F0 (CHORUS)

① 合唱歌唱の OMTS モデル

前節において, 合唱歌唱時の音高が随伴歌唱音高に近づくことが統計的に確認された. そこで本節では, F0 の動的な特徴について更なる分析を行うため, 1つの質点と2つのばねから構成されるシステム (One-Mass Two-Spring (OMTS) モデル) を利用する. OMTS モデルは図2に示すような簡単なシステムである. このシステムの運動方程式は以下で与えられる.

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

ここで m は質点の質量, λ は動摩擦係数, 楽譜の旋律情報 U と随伴歌唱音高 V はそれぞれのばねの自然長の位置と定義される. k_U と k_V はそれぞれ楽譜の旋律情報と随伴歌唱音高に対応する振動子に接続するばねのばね定数である. OMTS モデルにおける2つのばねのばね定数の値は, 楽譜情報と随伴歌唱からの影響の度合いを表す. 例えば, k_V が大きければ合唱時の歌唱は随伴歌唱からの影響をより受けた歌唱となる. 逆に, k_U が大きければ随伴歌唱にあまり影響されない歌唱となる. よって, モデルパラメータ $\Theta = \{m, \lambda, k_U, k_V\}$ は, 合唱における歌声の

表2 dの平均と標準偏差

	平均 μ [cent]		標準偏差 σ [cent]	
	独唱	合唱	独唱	合唱
素人	25.18	22.81	38.13	37.27
経験者	12.72	7.93	34.93	33.08

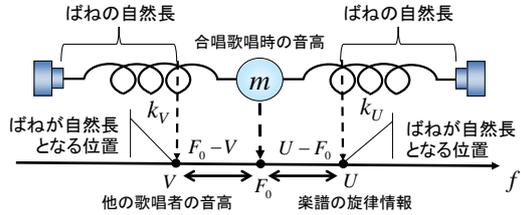


図2 合唱歌唱の OMTS モデル

引き込みを特徴づけるパラメータとなる.

② モデルパラメータ推定方法

本節では, 観測信号からモデルパラメータを推定する方法について述べる. 我々は, モデルパラメータを推定する方法として, 最急降下法を用いた. 最急降下法とは目的関数 $J(\Theta)$ を最小にする重みベクトル (モデルパラメータ Θ) を目的関数の勾配情報を利用して反復的に求める方法である. ここでは, 最小化する目的関数 $J(\Theta)$ として, 式(4)で与えられる二乗誤差を用いた.

$$J(\Theta) = \frac{1}{N} \sum_{n=1}^N (F_0(t_n) - y(t_n, \Theta))^2 \quad (4)$$

ここで, $F_0(t)$ は観測した歌声の F0, $y(t)$ は式(3)を解くことによって得られる関数である. 目的関数を式(4)とすると, モデルパラメータ Θ の更新式は以下ようになる.

$$\Theta^{(\tau+1)} = \Theta^{(\tau)} - \eta \frac{\partial J(\Theta^{(\tau)})}{\partial \Theta} \quad (7)$$

$$\frac{\partial J(\Theta^{(\tau)})}{\partial \Theta} = \frac{2}{N} \sum_{n=1}^N (F_0(t_n) - y(t_n, \Theta)) \frac{\partial y(t_n, \Theta)}{\partial \Theta} \quad (8)$$

ここで, η は各モデルパラメータに対する学習率パラメータである. 誤差関数 $J(\Theta)$ の Θ に関する導関数は, モデルパラメータセット Θ が与えられた下で式(3)を解くことによって計算することができる. 二階微分方程式である式(3)は, フーリエ級数展開とラプラス変換を用いて以下のように解くことが可能である.

4. 研究成果

(1) パラメータ推定による分析

表3 推定されたモデルパラメータと RMSE

	α	β_V	β_U	$\frac{k_V}{k_V + k_U}$	RMSE [cent]
素人	1.53	1.27	1.21	0.51	44.4
経験者	1.41	1.29	1.37	0.48	39.4

パラメータ推定を行い、素人と経験者のパラメータの違いについて分析する。実験に使用する歌声データ長は、収録データの歌い出しから 17 sec とした。素人と経験者について推定されたモデルパラメータの平均と、ばね定数 k_U, k_V の比 $k_V/(k_U+k_V)$ を表3に示す。なお、 α, β_U, β_V は、 λ, k_U, k_V を質量 m で割り、正規化した値である。この平均は、各歌唱者カテゴリに分類される3名の歌唱者の歌声データ9回分によって計算された値である。ばね定数比 $k_V/(k_U+k_V)$ は、楽譜の旋律情報と随伴歌唱音高のどちらからより影響を受けるかの指標となる。パラメータ推定の結果より、経験者より素人の方が随伴歌唱音高に対応する振動子に接続するばねのばね定数が大きいことがわかる。このことから経験者より素人の方が随伴歌唱に影響されやすいということが示唆される。また、経験者の α の方が素人より小さい値となっている。は OMTS モデルの動摩擦係数であり、変動の大きさの特徴付ける。よって、経験者の F0 ダイナミクスの方が変動が大きく、素人より表情豊かな歌唱であると推論できる。

(2) 歌唱者識別によるモデルの評価

合唱における引き込みの影響は各歌唱者によって異なると考えられるため、歌唱者毎に異なるモデルパラメータが推定されると仮定できる。そこで、観測した歌声の F0 系列と OMTS モデルから生成した F0 系列の二乗平均平方根誤差(RMSE)を評価尺度として歌唱者識別実験を行うことにより、歌唱の個人性を特徴付ける能力を評価する。本実験では、6名の歌声データを収録したため、6種類の識別実験となる。各歌唱者に対して3回分の歌声データがあるため、2回分を学習データとしてモデルパラメータ Θ を推定する。残り1回分を評価データとして RMSE の値を計算し、式(9)により歌唱者識別を行う。

$$\hat{s} = \arg \min E(F_0, y_s) \quad (9)$$

$$E(F_0, y_s) = \sqrt{\frac{1}{N} \sum_{n=1}^N (F_0(t_n) - y_s(t_n, \Theta_s))^2} \quad (10)$$

ここで s は歌唱者、 Θ_s は歌唱者 s のモデルパラメータ、 N は評価データ長、 F_0 は評価データ、 y_s はモデルパラメータ Θ から生成した信号を表す。

表4 RMSE の計算結果

	素-A	素-B	素-C	経-A	経-B	経-C
素-A	69.26	71.01	69.97	70.90	69.69	69.84
素-B	69.88	68.53	68.58	74.15	72.07	66.90
素-C	51.55	51.08	51.62	56.05	53.61	52.78
経-A	50.46	56.22	54.05	48.55	48.94	49.15
経-B	53.20	57.96	55.64	53.55	52.30	52.27
経-C	46.99	51.66	49.04	48.05	46.39	46.36

Note: “素”と“経”はそれぞれ素人と経験者を表す。横のラベルはモデルの名称、縦のラベルは評価データの各歌唱者の名称を表す。識別結果の正解は、対角の値が最も小さくなったときである。

結果を表4に示す。最小の RMSE の値を示すモデルをボールド体で強調した。横のラベルはモデルの名称、縦のラベルは評価データの各歌唱者の名称を表し、対角の値が最も小さくなったときが正解となる。表4より、6名中4名の歌唱者が自身のモデルにおいて最も小さい RMSE の値をとっていることが分かる。つまり、6名のうち4名の歌唱者を識別することができた。素-C と経-B は、歌唱者自身のモデルにおいて最も小さい値とならなかったが、自身のモデルにおいて比較的小さい RMSE の値となっている。また、RMSE の値から素人と経験者を正しく識別できていることが確認される。これらの結果から、合唱における F0 の引き込みの個人性を OMTS モデルにより特徴付けることができることが示唆された。

(3) まとめと今後の展望

合唱における歌声の F0 ダイナミクスを引き込みという観点から分析し、モデル化する手法を提案した。モデルの評価として歌唱者識別実験を行い、素人と経験者を正確に識別できることが確認された。このことから、F0 の引き込みをパラメトリックに表現できる手法として OMTS モデルの有効性が示唆された。本研究において、提案した合唱における F0 ダイナミクスを特徴付ける手法は国内外問わず新しい着想であり、査読付き国際会議(International conference on acoustics, speech, and signal processing)にも本研究の論文が採択されたことから、学術的貢献をしたと言えるだろう。

今後は、このモデルを応用しより人間ら

しい合唱音を生成するなど、アプリケーションとして適用することを試みる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

(学会発表)

- ① Motonari Kawagishi, Shota Kawabuchi, Chiyomi Miyajima, Norihide Kitaoka, Kazuya Takeda, "ANALYSIS AND MODELING OF ENTRAINMENT IN CHORUS SINGING," Proc. ICASSP 2013, (accepted), May 30, 2013, Vancouver, Canada, the Vancouver Convention & Exhibition Centre.
- ② 川岸 基成, 宮島 千代美, 北岡 教英, 武田 一哉, "ばね質量系を利用した合唱における歌声の FO ダイナミクスのモデル化," 情報処理学会研究報告, Vol. 2013-MUS-98, No. 12, Mar 16, 2013, 公立ほこだて未来大学.
- ③ 川岸 基成, 宮島 千代美, 北岡 教英, 武

田 一哉, "合唱における基本周波数軌跡のモデル化に関する研究," 日本音響学会 2012 秋季研究発表会講演論文集, 3-Q-33, pp. 433-436, Sept 21, 2012, 信州大学.

6. 研究組織

(1) 研究代表者

武田 一哉 (TAKEDA KAZUYA)
名古屋大学・大学院情報科学研究科・教授
研究者番号：20273295

(2) 研究分担者

なし ()

研究者番号：

(3) 連携研究者

なし ()

研究者番号：