

Title	歌声におけるF0動的変動成分の抽出とF0制御モデル
Author(s)	齋藤, 毅
Citation	
Issue Date	2002-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1570
Rights	
Description	Supervisor: 赤木 正人, 情報科学研究科, 修士

歌声における F0 動的変動成分の抽出と F0 制御モデルに関する研究

齋藤 毅 (010046)

北陸先端科学技術大学院大学 情報科学研究科

2002 年 2 月 15 日

キーワード: 歌声知覚、歌声合成、基本周波数 (F0)、F0 動的変動、F0 制御モデル.

1 はじめに

歌声は話声に比べ、より複雑で動的な特性を持ち、その特性は歌声の音声波に含まれる特徴量、特に基本周波数 (以下 F0 と呼ぶ) に顕著に表れる。よって、高品質な歌声合成の実現を考えた場合、歌声の F0 における動的変動の制御が重要になり、その為には動的変動の持つ情報を明らかにすることが必要となる。

歌声の音響分析の研究では、これまで F0 における動的変動成分に着目したものが行われてきた。その中で、オーバーシュート・アンダーシュートや、ヴィブラート・微細変動といった動的変動が、歌声の F0 において重要な成分であると報告されている [1, 2, 3]。しかし、歌声の F0 に特化した動的変動成分として、上記の 4 つの成分だけで十分なのか議論されてないし、これらの歌声知覚に与える影響についても調べられていない。

一方 F0 制御モデルとして、話声に対応したものが数多く提案されているが [4, 5]、歌声の F0 変化を制御・記述できるモデルは提案されていない。

本研究では、歌声の F0 を分析し、心理物理実験によって F0 動的変動成分の歌声知覚に与える影響を調べる事で、主要となる F0 動的変動の抽出を行う。そして、それらを制御できる F0 制御モデルを提案し、歌声合成への応用を試みる。

2 歌声の F0 動的変動の分析

歌声特有の F0 動的変動成分を抽出する為に、本研究では音声分析合成システム STRAIGHT[6] の TEMPO で F0 推定を行い、分析を行った。使用したデータは、ボーカリスト 3 名が日本童謡 “七つの子” を日本語母音 /a/ で歌唱した歌声をサンプリング周波数 48 [kHz]、量子化ビット数 16 bit でデジタル録音し、16 小節からなる歌を 4 小節ごと

に切り出した計 12 個のデータを使用している。TEMPO2 で抽出した対数軸上での F0 変化を図 1 に示す。

2.1 歌声の F0 動的変動成分の抽出

先の研究では、話声には無く、歌声だけに特化した特徴として以下に示す 3 つが報告されている [1]。

1. F0 の大まかな変化はメロディに、その時の F0 の値は音階に対応する。
2. ある一つの音程区間における F0 変化は、下降せずに安定もしくは上昇する。
3. 歌声のみに観測される動的変動がある。

このうち、1、2 に関してはメロディに関する特性である。本研究では 3 番目の F0 動的変動成分に着目し、この成分が担う歌声としての情報を調べるために、図 1 の様な F0 波形から以下に示す 3 つの F0 動的変動に着目した。

オーバーシュート・アンダーシュート : 音程変化時に目的音の音高の値より大きく振れる現象。これらを以下 O-S、U-S と呼ぶ

ヴィブラート・微細変動 : 同一音高が持続した場合に観測される 4~7 [Hz] の周期的な振動、さらにはそれを取り除いた後に残る不規則的な細かい変動。

予備的变化 : 音程変化直前に観測される音程変化とは逆の方向に変化する瞬時的な変化。

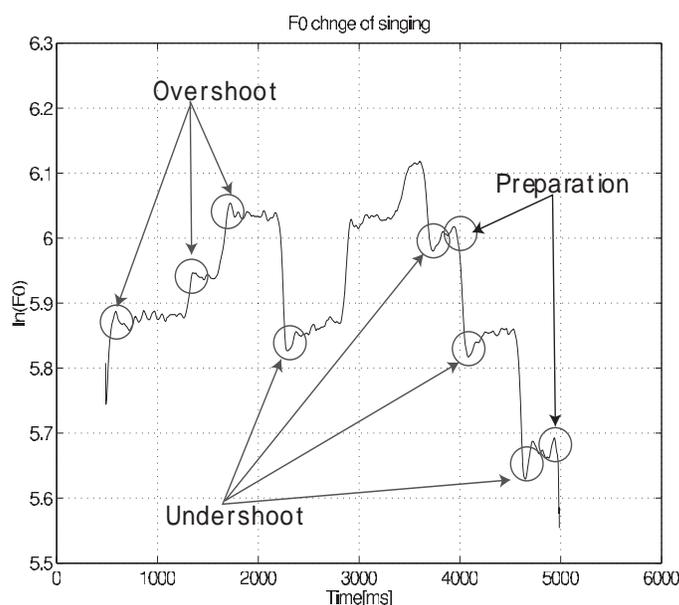


図 1: 歌声“かわいい七つの”歌唱時の F0 変化

最初の2つの成分は、過去の研究 [1, 2, 3] において歌声特有の特性であると報告されている。また、3つ目に関しては、今回新たに着目した変動成分であり、これら3つの成分が歌声知覚に与える影響を調べるために、次に示す心理物理実験を行った。

2.2 F0 動的変動成分の歌声知覚への影響

2.2.1 心理物理実験

先に示した3つのF0 動的変動成分の歌声知覚への影響を調べるために、以下の様な処理を施したF0を作成し、Klattのホルマント合成器によって合成音を作成した。

- O-S、U-Sを除去したもの(これを用いた合成音をNO-OUSと呼ぶ)
- ヴィブラート・細かい変動成分を除去したもの(NO-VIB)
- 予備的变化を除去したもの(NO-PRE)
- カットオフ周波数5Hzでスムージングし、動的変動を除去、及びメロディ構造を少し壊したもの(SMS)

この4つの刺激とF0に手を加えないで合成した歌声(NORMAL)を用いて実験を行った。実験方法はシェッフエの一対比較法(7段階評価法:-3~3)を採用し、被験者として正常な聴力を有する大学院生男女6名を対象に、防音室におけるヘッドフォンでの両耳受聴で行った。

2.2.2 実験結果と考察

実験の結果を図2に示す。刺激名の下値は、その刺激の歌声としての自然性を表し、NORMALから各F0 動的変動成分が除去されることで、歌声の自然性の劣化が確認できる。ここから、3つのF0 動的変動成分の歌声知覚に与える影響、中でもO-SとU-Sの影響が大きい事が分かる。以下、これらのF0 動的変動を制御できるモデルを構築する。

3 歌声のF0制御モデル

現在の話声に対応したF0制御モデル[4, 5]では、歌声のF0 動的変動を含むF0変化を制御・記述するのは困難である。そこで今回、歌声に対応したF0制御モデルを提案する。

3.1 F0制御モデルの概要

F0制御モデルの概要を図3に示す。このモデルは入力としてのメロディ成分に以下の4つの成分によってF0 動的変動成分を制御し、F0変化を記述できる2次系システムである。

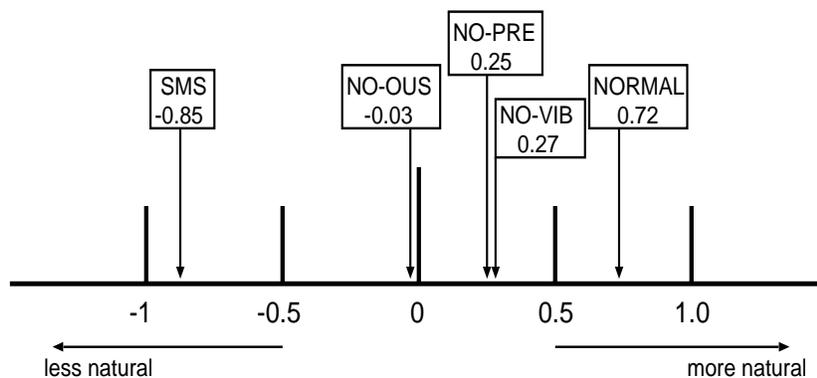


図 2: 聴取実験による歌声刺激の自然性の関係 1

歌声変動成分 1 : O-S、U-S を制御する

歌声変動成分 2 : F0 安定時にヴィブラートを制御する

歌声変動成分 3 : 音程変化直前の予備的变化 (preparation) を制御する

歌声変動成分 4 : F0 変化全体の微細変動 (fine fluctuation) を制御する

ここでシステムの入力であるメロディ成分とは曲の旋律概形を表し、F0 変化のベース成分で矩形パルスで記述される。次に歌声変動成分 1、2、3 は制動 2 次系の伝達関数

$$H(s) = \frac{\Omega}{s^2 + 2\zeta\Omega s + \Omega^2}$$

のインパルス応答で与えられ、その特性は、歌声変動成分 1、3 に関しては減衰振動、2 には定常振動で与えられる。そして歌声変動成分 4 は最大で 7 Hz 程度の振幅を持つ不規則的な細かい揺れ成分で表される。このモデルでは、歌声変動成分 1~3 それぞれに対して最適な制御パラメータ ζ と Ω の 6 つを与えてやることにより、図 4 に示す各動的変動を含んだ F0 変化を制御・記述できる。

3.2 歌声合成への応用とその評価

F0 制御モデルを STRAIGHT による歌声合成に適用し、聴取実験によって F0 制御モデル及び生成される F0 の評価を行った。実験で使用した刺激は以下の 6 つで、合成におけるスペクトル情報は一切手を加えていない。

NORMAL : 歌声データを STRAIGHT で分析合成したもの

SYN-All : モデルによってすべての変動成分を制御した F0 を用いた合成音

SYN-OUS : O-S、U-S のみ制御した F0 を用いた合成音

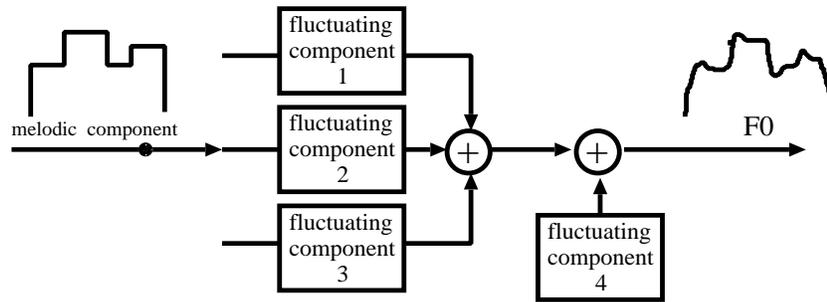


図 3: F0 制御モデルの概要

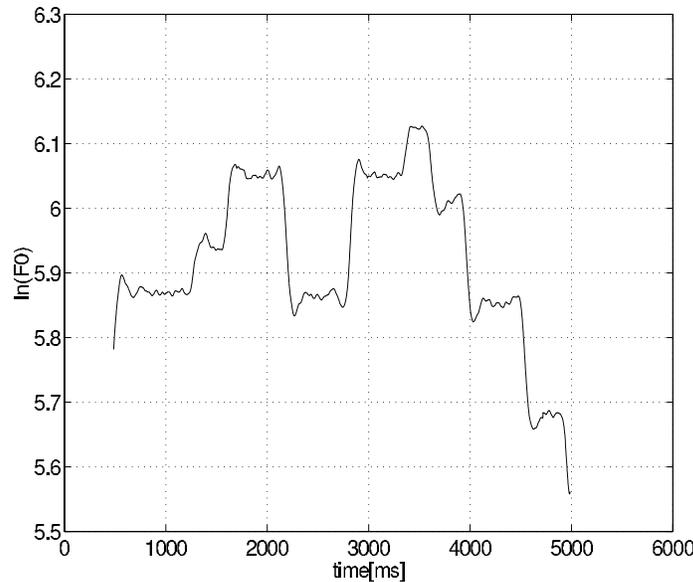


図 4: モデルによって生成された F0

SYN-PRE :予備的变化のみ制御した F0 を用いた合成音

SYN-VB :ヴィブラート・微細変動のみ制御した F0 を用いた合成音

SYN-BASE :メロディ成分のみ用いた合成音

これらの刺激を用いて、前述の実験と同じ方法・条件で心理物理実験を行った。その結果を図 5 に示す。F0 動的変動成分をすべて制御した F0 による合成音は実音声に近い自然性を有した歌声として知覚された。また、1つの動的変動をメロディ成分に加えるだけで、歌声の自然性が増す事が確認できた。この結果から、本 F0 制御モデルの有用性と、歌声合成への応用可能性を確認した。

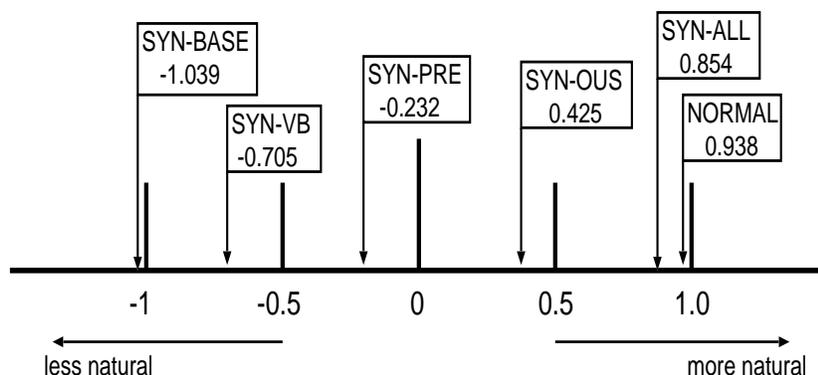


図 5: 聴取実験による歌声刺激の自然性の関係 2

4 まとめ

本研究では、歌声の F0 における F0 動的変動成分に着目し、心理物理実験を通じて歌声知覚に与える影響を F0 動的変動成分の抽出を行った。その結果、オーバーシュート・アンダーシュート、ヴィブラート、微細変動、予備的变化の 5 つの変動を、歌声知覚に大きな影響を与える F0 動的変動成分として抽出した。

次に歌声に対応した F0 制御モデルを提案し、歌声合成に応用した。その結果、入力としてのメロディ変化に F0 動的変動を付与できる 2 次系システムとして F0 制御モデルを構築し、少数のパラメータで F0 動的変動を正確に制御し、F0 変化を記述できる事を確認した。そして、F0 制御モデルによって生成された F0 を用いての歌声合成を行った結果、実音声に近い自然性を持った歌声を合成できる事が分かり、本 F0 制御モデルの高品質な歌声合成への応用可能性を確認した。

参考文献

- [1] 矢田部 学, 粕谷 英樹, “歌声の基本周波数の動特性” 日本音響学会秋季講演論文集, 3-8-6, 1998.
- [2] 小田切 わか菜, 粕谷 英樹, “歌声のヴィブラートの分析、合成、知覚に関する検討” 日本音響学会秋季講演論文集, 1-7-5, 1999.
- [3] 北風 裕教, “歌声の基本周波数の微細変動成分の知覚に関する研究” 北陸先端科学技術大学院大学修士論文, 2000.
- [4] H. Fujisaki, M. Tatsumi, “Analysis control in singing” “Vocal fold physiology”, UNIVERSITY OF TOKYO PRESS, pp.347-363, 1981.
- [5] T. Moriyama, H. Ogawa, S. Tenpaku, “A new control model based on rising and falling fundamental frequency” Proc. of ASA and ASJ Third Joint Meeting, pp.1171-1176, 1996.
- [6] H. Kawahara *et al.*, “Fixed point analysis of frequency to instantaneous frequency mapping for accurate estimation of F0 and periodicity” Proc. Eurospeech99, pp.2781-2784, 1999.