

奄美大島民謡風歌声合成システム：グインレゾネータ

村主大輔^{†1} 馬場 隆^{†1}
森勢将雅^{†2} 片寄晴弘^{†1}

日本ではカラオケや DTM の普及によって音楽活動がますます一般化され、年間 200 組以上のアーティストがメジャーデビューしている。それに伴い、新たなジャンルや歌唱スタイルが生まれることは少なくない。その一つの例として、ポピュラーソングに沖縄や奄美大島などアーティスト出生地の特色を出した音楽表現のスタイルが近年注目されるようになってきている。そこで本研究は、歌唱スタイルが特徴的な奄美大島出身歌唱者の歌い回しに注目し、一般歌唱を奄美大島出身の歌唱者の歌い回しにするシステムの開発を目指す。具体的には、「グイン」と呼ばれる奄美大島出身歌手の歌唱音声を歌唱特徴の定量的な分析を実施し、その分析結果に基づいて、一般歌唱に「グイン」を付加するシステムの概要と、その動作結果について報告する。

GUIN-Resonator: A system synthesizing voice with the style of Amami folk songs

DAISUKE SUGURU,^{†1} TAKASHI BABA,^{†1}
MASANORI MORISE^{†2} and HARUHIRO KATAYOSE^{†1}

The recent spread of “Karaoke” and DTM has been promoting music production more generally, and more than 2 hundreds musicians make their debuts in Japan. This leads emergence of new singing styles. Among them, “Okinawa-style” or “Amami-style” is typical one that has been popular recently. We have been developing an assistance system for designing “Okinawa-style” or “Amami-style” vocal melodies. In this paper, we report acoustic analysis of “Amami-style,” especially singing style called “guin” and propose a “Amami-style” singing generator, called “Guin-Resonator.”

^{†1} 関西学院大学 理工学研究科

Graduate School of Science and Technology, Kwansai Gakuin University

^{†2} 立命館大学 情報理工学部

1. はじめに

近年、歌声合成技術の進歩が著しく、Vocaloid¹⁾ の登場を皮切りに、CGM(consumer-generated media)の世界で一般ユーザによる歌声合成作品が多数見られる。この Vocaloid は、人間の音声を録音し、収録音声を利用した連結的合成方式に分類される歌唱合成ソフトウェアであり、メロディーと歌詞を入力することによって歌声の生成が行われる。この技術によりそれまでの音声合成技術に比べ、リアルな人間の音声の合成を可能とした。この Vocaloid の歌声合成技術を用いた「初音ミク」²⁾ や「鏡音リン・レン」²⁾ は、同様の技術の歌声合成ソフトウェア³⁾⁴⁾ に比べ人気が高く、市販の歌声合成ソフトウェアでは異例の販売数を記録し、一般ユーザによりそれらのソフトウェアから作成された音声がニコニコ動画等で多数見受けられるほど歌声を自らの手でデザインできる技術が気軽に楽しめるようになってきている。一般ユーザが投稿しているの中には、かなり人間に近い表情豊かな Vocaloid の音声を作製している職人と呼ばれる人もいるが、Vocaloid のシステムではそのような高品質の音声を作るにはかなり煩雑で細かなパラメータの調整作業と労力を費やす必要がある。

そこで、産総研の Vocalistner⁵⁾ では煩雑なパラメータの調整作業を必要とせず、人間の音声に近い音声を合成できる歌唱デザイン技術を提供している。Vocalistner は、ユーザの歌声と歌詞を入力とし、入力音声の基本周波数と音量のエンベロープから独自の反復推定に基づく歌声合成パラメータ推定技術により、Vocaloid などのソフトウェア音源の音声を煩雑な操作をすることなく高品質な音声への加工を実現している。しかし、ユーザの歌声を入力とする Vocalistner は、歌声の情報を音響信号として入手できなければ利用できない。また、パラメータ推定がユーザの歌声情報に依存するため表現したい歌唱スタイルをユーザ自身が表現できるのかどうかという課題がある。

これらのシステムはパラメータ設定においても歌唱音声の利用においても、高度な歌唱表現を行うにはユーザの能力に頼らざるを得ない。高度な歌唱表現を実現するために低次のパラメータをフレームごとに指定などの煩雑な作業することなく、高次のパラメータ設定のみによる歌声合成はユーザの負担の軽減につながる。歌唱表現には様々なスタイルが存在し、それぞれによって音響的特徴は様々である。そのため、基本周波数などの低次のパラメータをフレームごとに指定するなどの煩雑な作業を必要とせず高次のパラメータを指定することによって自然な歌唱合成を実現するには、歌唱スタイルごとに歌唱スタイルに合わせた、

音響的特徴の分析・モデル化を行い、歌唱合成の手法を開発していく必要がある。

本研究では近年注目されつつある奄美大島出身歌唱者風の歌唱デザイン支援の一つの形態として、ユーザの歌唱音声を奄美大島出身歌唱者風に交換するシステムを提案する。奄美大島出身歌唱者の歌声は、2002年に日本のポピュラー音楽界にメジャーデビューした元ちとせの独特な歌唱スタイルによって広く知られるようになった。彼女の歌唱からは演歌のような節回しを耳にすることができるが、演歌界ではなく J-Pop での楽曲製作を行っている。また元ちとせ⁶⁾のデビュー後に同郷の中孝介がメジャーデビューするなど、奄美大島出身の歌唱者の楽曲は J-Pop の中でも音楽ジャンルの一つとして認識されつつある。しかし、注目度が増している奄美大島出身の歌唱者の歌声を生成するような歌唱デザイン技術は存在しない。本稿では、奄美大島出身の歌唱者の独特の歌い回しの特徴を明らかにし、明らかになった特徴を一般の歌唱に付加することで奄美大島出身者風の歌唱音声を生成するシステムについて報告する。

以下、2章では奄美大島出身者歌唱の特徴について述べる。3章では奄美大島出身者風歌声合成システムの概要と動作について述べる。4章では一般歌唱を奄美大島出身者風の歌唱音声を生成するための歌唱楽曲と音響信号に対する課題と対策について述べる。5章では今後の展望について述べる。

2. 奄美大島出身者歌唱の特徴とゲイン

この章では、奄美大島出身の歌唱者の特徴的な歌唱法について概況を述べ、その歌唱法の音響信号における特徴について紹介する。

2.1 奄美大島民謡（島唄）の特徴

奄美諸島出身の歌手は、奄美大島民謡（島唄）の歌い回しをしている場合がほとんどである。日本には民謡や長唄といった様々な伝統音楽があり、それぞれ歌唱様式が異なる。しかし、現在行われている日本民謡や長唄に関する歌唱の検討は、ピッチ推移を対象とした定性的な検討⁷⁾⁸⁾に留まっており、「島唄風」を対象とした歌唱デザインシステムは存在しない。

島唄の歌い方の特徴として、裏声の多用と「ゲイン」と呼ばれる島唄特有のコブシによる節回しがあげられる。裏声は近年のポピュラーソングの中でも非常によく使われるのに対し、「ゲイン」は他の民謡および奄美大島出身の歌唱者以外のポピュラーソングでは聴くことの出来ない技法である。ゲインの特徴的な性質として、歌唱時に裏声を使う点が島唄歌唱者によって指摘されている⁹⁾。そのためゲインは長唄や他の民謡で見られるコブシのように基本周波数（F0 と呼ぶ）が変動するだけでなく、それに合わせて音色も変化すると考えら

れるが、実際の F0 やスペクトルを分析した事例はほとんど実施されていない。次節では、ゲインの音響的な特徴について紹介する。

2.2 ゲインの音響特徴

典型的なゲインの分析対象事例として奄美大島出身の歌手である中孝介¹⁰⁾の歌唱音声（『花』の「いいわけさえも」というフレーズ）の F0 軌跡とスペクトログラムを図 1 に示す。なお、このスペクトル分析には、音声の分析・再合成において基本周波数・周期性指標・非周期性を独立に取り扱うことができる河原らの STRAIGHT¹¹⁾¹²⁾を用いた。ゲインは、図 1 において四角で囲まれた部分（2 つめの「い」）で用いられている v のこの図から次の 2 つの事項が推察される。

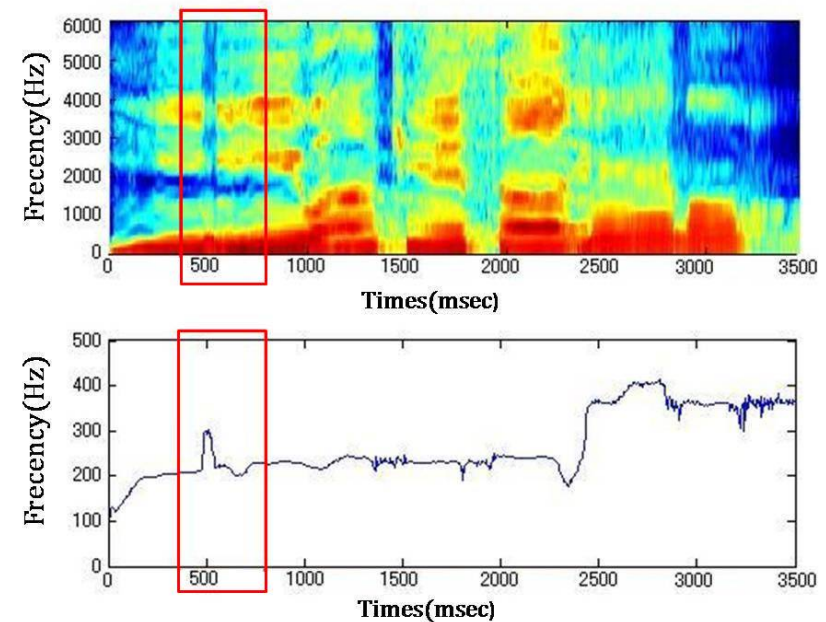


図 1 ゲインにおけるスペクトログラム (上図) と F0 (下図) の特徴的な変動

- F0 が急激に上がって元の F0 に戻った（本稿では「隆起」と呼ぶ）後、さらに F0 が下

がって元の F0 に戻っている（本稿では「沈降」と呼ぶ）。これは、長唄でも用いられている「あたり」⁷⁾¹³⁾ に非常に似た特徴である。ただし、「あたり」では「沈降」は見られない。そのため、この「沈降」がゲインの大きな特徴を表している可能性がある。

- 同一音素を発声中にも関わらずスペクトルが急激に変化している。特に、F0 の隆起時に高次倍音 (2kHz~4kHz 付近) が非常に小さくなっている。文献⁹⁾ での指摘の通り、裏声が用いられたからであると考えられる。

次に、図 1 で取り上げたフレーズの中孝介の歌い方を、奄美大島民謡の歌唱が可能なプロ歌手 S_T (奄美大島出身ではないスタジオミュージシャン) が模倣したものの分析結果を図 2 に示す。

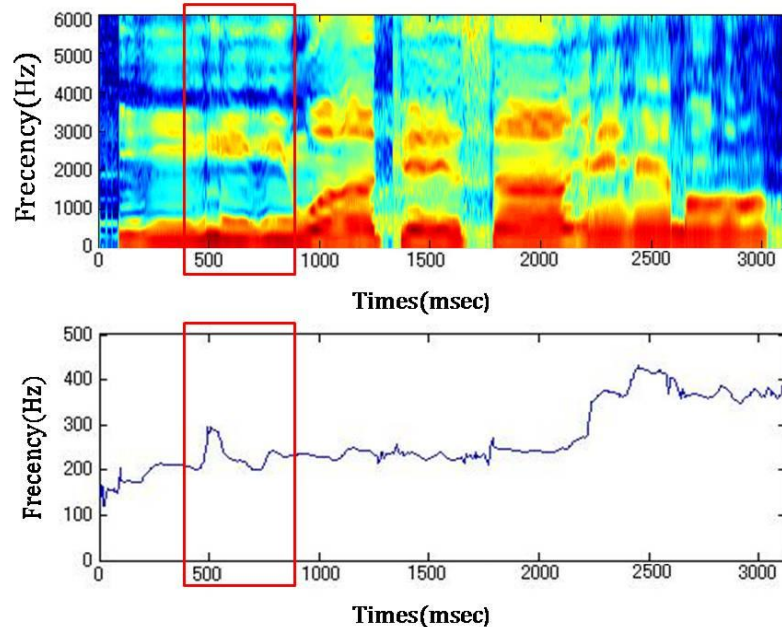


図 2 奄美大島民謡の歌唱が可能なプロ歌手によるゲインにおけるスペクトログラム (上図) と F0(下図) の特徴的な変動

奄美大島民謡の技法を使える S_T と中孝介のスペクトログラム・F0 では同様の特徴が図

2 では見られる。またこれらの特徴に加え、スペクトルでの高次倍音成分の減少が F0 の隆起区間中のみには発生することが確認できる。これらのことから、ゲインを表現するには

- (1) F0 における隆起と沈降
 - (2) 裏声によるスペクトルにおける倍音成分の減少
 - (3) F0 の隆起に合わせたスペクトル変動 (F0 変動とスペクトル変動の同期)
- が必要と言える。

3. 奄美大島民謡風歌声合成システム

本研究では、入力した「ゲイン」の含まれない一般歌唱に対して「ゲイン」を付加するシステム (ゲインレゾネータ) の開発を目指す。非専門家が簡易に使用できることを想定し、歌唱楽曲中で「ゲイン」の入る位置の候補を提示し、ユーザが指定した部分に対して「ゲイン」唱法へと変換する。歌唱デザインに利用のほか、技法や歌い回しの習得は困難であ

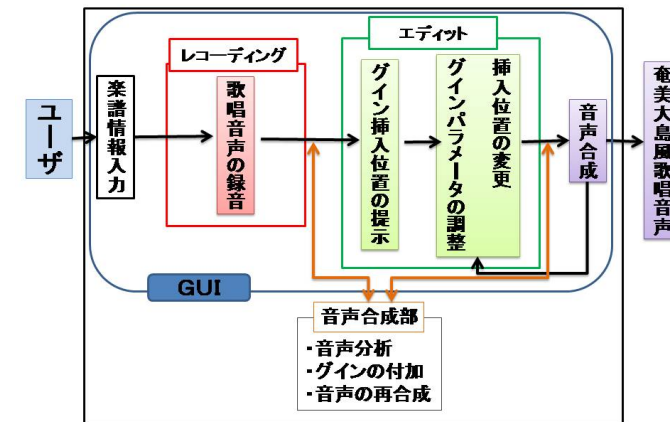


図 3 奄美大島民謡風歌声合成システムの概要

る¹⁴⁾ とされる民謡指導や学習における利用が想定される。本研究で提案する奄美大島民謡風歌声合成システム「ゲインレゾネータ」の処理概要を図 3 に示す。

このシステムは入力音声を GUI で行えるように設けたレコーディングモードと、「ゲイン」の入る候補の中からゲインを入れる箇所を選択するだけで入力音声を編集できるようなエディットモードを実装している。ゲインレゾネータへの入力データは一般歌唱者が歌う楽曲(対象楽曲と呼ぶ)の楽譜情報(ボーカルパートのメロディと歌詞情報)とレコーディングモードで収録される歌唱音声である。ボーカルパートのメロディは MIDI ファイル (STM), 歌詞はテキストファイル (.txt) として与えられる。

3.1 ユーザインタフェース

前節で記したゲインレゾネータの実用例と2つのモードについて以下で述べていく。ゲインレゾネータを起動し、楽譜情報をインプットすると図4のようにピアノロール形式で楽曲情報が表わされる。主な機能は、図4のメニューバーやツールバーに実装されている。

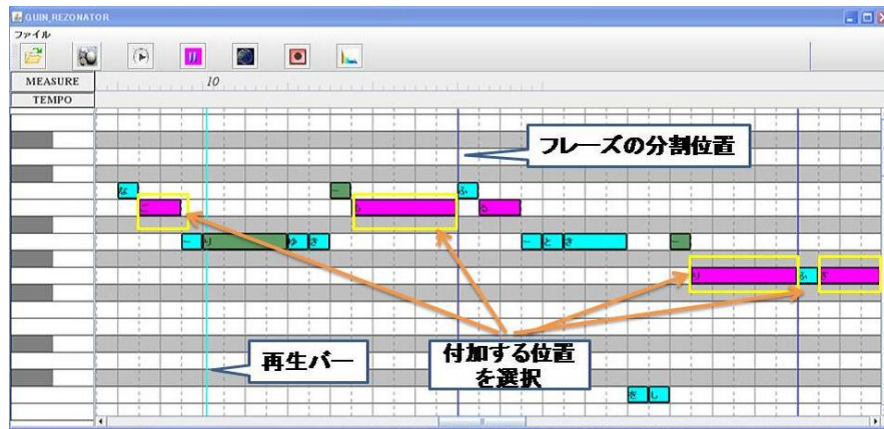


図4 ゲインレゾネータによるゲイン付加位置の提示

図4では、イルカのなごり雪¹⁵⁾を楽曲データとして与えている。ピアノロールのノートの色は付加位置の候補の優先度の違いを表現しており、水色はゲイン付加の候補がない音、赤はゲインが入りやすい音、緑はゲインは入るが赤ほど入りやすい音である。また、色の明るさによってユーザの選択した場所を分けている。図4の場合は、黄色の四角で囲った

赤色のノートでゲインの付加を行うことを選択している状態である。

3.1.1 レコーディングモード

レコーディングモードでは、ユーザが入力した MIDI 音源にあわせて歌唱を行い音声合成・編集を行うための音声を録音する。この際、カラオケのようにシステムから流れてくるガイドメロディを聞きながらインタフェース上の画面にある歌詞とノートの流れ歌唱するのだが、ユーザの音声だけを取得したいためヘッドフォンなどガイドメロディが同時に収録されないようにすることが必須となる。

3.1.2 エディットモード

エディットモードでは候補として挙げられたノートにゲインを付加するかどうかの選択ができることに加え、前章で明らかになった F_0 とスペクトルの変動の度合い(これらを音響信号を操作する際のゲインのパラメータとしてそれぞれ P_{f0} と P_{sp} とする)をそれぞれ独立に調整ができる。調整されたゲインの音響信号の変動の度合いはノートの彩度に表すようにする。これにより、単にゲインのオン/オフだけでなくユーザの意図するゲインの表現も可能な歌声合成を実現する。ただし、 F_0 とスペクトルの変動の調整幅は $0 \leq P_{f0} \leq 200$, $0 \leq P_{sp} \leq 100$ で単位はどちらも%とする。

一方で、このユーザによるゲインの位置の選択を行う機構とは別に、次章で述べる音響信号の処理を行う STRAIGHT を使った音声合成部という機構を用意している。音声合成部では主に、入力音声の分析、ゲインの付加、音声の再合成を行い、ユーザが指定した位置と度合いのゲインが表現された音声を生成する。この音声合成部で、図4の赤いノート位置(ユーザの選択位置)の状態を入力音声にゲインの付加が行われた音声の基本周波数とスペクトルはそれぞれ、以下の図5、図6のようになる。このようなゲインの付加位置を選択するインタフェースと音響信号の処理機構を使いユーザは奄美大島民謡の歌い回しをする歌唱音声を生成可能とする。

4. 「ゲイン」の変換処理

2章の分析により明らかになったゲインの音響信号における特徴を一般歌唱音声で表現するための処理と、3章で紹介したユーザインタフェースで提示されるゲインの付加する候補位置の算出について本章では述べていく。

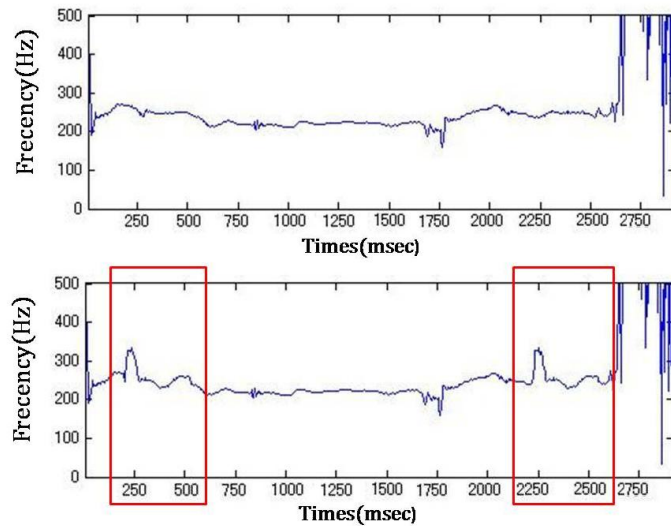


図 5 システムによる音響信号へのゲイン特徴の付加を行った基本周波数の遷移 (上図は入力音声, 下図は出力音声)

4.1 ゲインの付加位置の算出

奄美大島出身歌手 (以下, 奄美歌唱者と呼ぶ) の歌唱では, 楽曲中のある音の並び (メロディ) においては毎度もゲインは使われているが, 他のメロディでは全くゲインが使われないことがある. このことからゲインには楽曲中のあらゆる場所で使われるのではなく, 歌唱者によってゲインを表現しやすい場所が存在するといえる. これにより, ゲインの挿入においてはゲインの音響信号上での特徴表現とならびに適切な位置に付加されることが重要であると考えられる. しかし, 歌手毎にある程度の規則性を見いだすことができるものの必ずしも一意ではないことを奄美歌唱者の楽曲から確認している.

そこで本研究では, 奄美歌唱者が楽曲中でゲインが出現しやすい箇所の音の並びを調べ, 奄美歌唱者の楽曲でなくてもゲインが出現しやすいであろう音符とゲインが入ることのない音符を算出できるような機構を設ける. ゲインレゾネータでは, ゲインの挿入位置の候補を提示しユーザが選ぶ方式を採用することで, ユーザに選択の自由度を与えつつ, 音楽的にあり得ない位置でのゲインの挿入を抑制できる. ゲインの挿入位置算出に関しては, 歌手毎にゲイン挿入位置が異なるという可能性を考慮し, ゲインが出現する場所を歌手毎に決定す

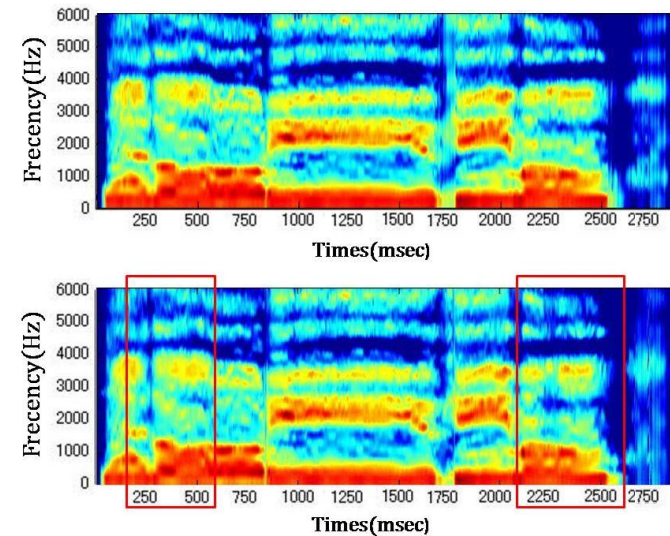


図 6 システムによる音響信号へのゲイン特徴の付加を行ったスペクトル (上図は入力音声, 下図は出力音声)

る学習アルゴリズムを用いることが有用と考えられる. 本稿では, 学習アルゴリズムを構成する前段階として, 奄美歌唱者の楽曲から筆者が設定したゲインが出現しやすい場所を算出する条件・制約をゲインの挿入位置を算出のルールとして以下で説明する.

- ゲインの挿入位置を算出するルール (挿入ルールと呼ぶ) の設定

奄美歌唱者の楽曲中でのゲインの入りやすさは, 音の並び (メロディ) が要因ということを上で述べた. 一方, 音の並びは音楽を階層的なまとまりとしてフレーズとして扱われることがあり, 作曲や演奏などで重要な楽譜上に明記されていない音楽の切れ目の役割をする. 歌においても歌詞や呼吸位置で楽曲のフレーズに分割され, 楽曲を把握するためなどに利用されるケースは少なくない. 特に呼吸位置は歌唱技法の付加やブレスコントロールなど, 歌唱表現を豊かにするための大切な楽曲のフレーズ分割の境界とされている. 本研究では楽曲をブレス位置でフレーズ分割し, ゲインの入りやすいメロディをそのフレーズ毎に各音符に対し前後の音符の音高情報, 音価, フレーズ内での相対的位置, 楽曲全体での相対的位置という情報をシステムに入力される楽譜情報から用いて算出する.

フレーズ毎のどの位置にゲインが入りやすいかを求める挿入ルールの例として以下が挙げられる。

- フレーズの頭の音には入らない
- 楽曲中において相対的に音価小さい音が、同音高・同音価の音列が続く場合は入らない
- フレーズのはじめから2音目, 3音目では、次の音符の音価が大きくなると、入りやすい
- フレーズ内で音価の最も大きい音符が、一つ前の音符から音高が下がる場合、ひとつ前の音符に入りやすい(選択ルール)
- フレーズの終わりで大きい音価の音符、一つ前の音符から音高が下がる場合、フレーズの終りの音符に入りやすい(選択ルール)
- 楽曲中において、前半部分では1フレーズに1か所、後半部分では1フレーズに2か所まで入れることができる
- 上記のルール以外で、フレーズ内で音価が比較的大きい音符

上記のルールにおいて、上に書かれているものから順にルールの優先度が高くなる。ただし、4つ目と5つ目の選択ルールについては、近年のポピュラーソング中でどちらも存在する節回し¹⁶⁾に近いので、優先度は同じとする。

● 挿入ルールを用いるための、楽曲の自動フレーズ分割

次に、上で設けた挿入ルールを適切に分割したフレーズで用いるために、楽曲中でのブレス位置を自動的に推定する方法について説明する。ユーザが入力するMIDI(楽譜)情報と歌詞情報に対し、以下の処理を行う。

- (a) MIDIから楽譜情報(各音符のノートナンバー、音価、発音開始時刻)、形態素解析により歌詞情報から切れてはいけないフレーズング位置の取得。形態素解析による情報は禁則として用いる。以下の処理はすべて、形態素によるフレーズング禁則出ない場合に行われるものである。
- (b) 休符情報により分割
- (c) フレーズ内での最も音価の大きい音符(音符(A))を算出
if フレーズ内での最も音価の大きい音符が一つの場合
音符(A)の後ろで分割
else フレーズ内での最も音価の大きい音符が複数の場合if 次の音が上への跳躍進行(音高があがる)の場合音符(A)の後ろで分割(Aは複数可) else 次の音が上の

跳躍進行でない場合分割すると短くなりすぎなければ、音符(A)の後ろで分割(d)(b)~(c)を各フレーズがある一定以上の長さ(秒)より、短くなるまで続ける



図7 フレーズ分割と挿入ルールにより算出されるゲインの候補位置を矢印で提示。(対象楽曲:「なごり雪」)矢印の色はゲインの入りやすさで異なる

ブレスの位置でのフレーズ分割と挿入ルールによって算出した楽曲(イルカの「なごり雪」)におけるゲインの付加位置の候補は図7のように表示される。ここでは、青色の線分はフレーズングする位置を示し、矢印はゲインの付加位置を表す。矢印の色は優先度を表現しており、赤色は非常にゲインが入りやすく、緑色はそれほど高くはないが入る可能性が0ではない場所を示している。

4.2 ゲイン付加に関する音響信号の処理

この節では、システムに楽譜情報を与えゲイン付加位置を算出した後、入力音声に対しゲイン特徴の付加をする音響信号上での処理内容を述べる。しかし、ゲインを挿入する処理では、2章で明らかになった基本周波数とスペクトルの特徴を付加するために次の2つの課題を考慮する必要がある。

- 音韻アライメントの問題
- 地声音声からの裏声への生成

これらの課題の詳しい内容と課題への対応を記し、その上で入力音声の基本周波数とスペクトルでのゲイン特徴の処理について説明する。

4.2.1 歌唱音声と楽譜情報から音韻アライメントの取得

適切なフレーズングとゲインを挿入する位置が楽譜情報において決定すると、次に音声情報のどの部分が楽譜上のどの位置に該当するかを決定する必要がある。

この課題への対策として、本システムはレコーディングモードでカラオケのようにガイドメロディーを流し歌うことで、ユーザの歌唱音声が入力された楽譜情報(MIDI)と同期が取

れているものとした。そして、MIDIによる楽譜の情報から各ノートに対する発音タイミングを割り出し、歌詞情報とあわせ入力音声の音韻アライメントの情報を保持する。

4.2.2 地声音声から裏声の自動推定

ゲインでは、裏声を歌唱中に一瞬入れることが報告されている。2章でのスペクトル解析により、F0の変動である隆起から沈降にかけて声質(スペクトル)でも裏声から地声と変化が生じているのが明らかになったため、ゲインを表現するには裏声の表現を欠かすことができないといえる。スペクトル解析の結果から、ゲインのF0の隆起中のスペクトルでは、同音素・同音高の他の箇所音声と比べて高次倍音成分が減少していたため、それをいかに表現するのが課題となる。ここでは、一般歌唱者3名の5母音の地声と裏声の音声を収録し、そのスペクトルを利用して、地声から裏声を生成するフィルタを構成した。以上の課題と対策を踏まえ、STRAIGHTにより音響信号の「基本周波数」「スペクトル」「非周期性成分」の分離を行った後、次の処理を施す。

4.2.3 基本周波数の操作

まず、2章で分析に用いた「中孝介」の「花」のフレーズ「言訳さえも」からゲイン部分のF0変動をテンプレートとして取得しておく。このテンプレートを歌唱音声の中のゲインを付加する位置で入力音声のF0と置き換えることによりゲインのF0変動を表現する。

ここで問題となるのは、歌唱者の音声のF0の遷移に合うようにゲインのF0の隆起と沈降を表現しなければならないことである。そこで楽曲中のゲインを付加するノートの音価(発音中の時間)に相当する入力音声の区間で、基準となるF0の値(目標周波数と呼ぶ)を入力音声から決定する。目標周波数の決定方法は、目標周波数を求める音声区間で区間の真ん中から区間の長さの1/4分までのF0の平均値をとる。区間の長さを音響信号のフレーム数Nとし、i番目のフレームのF0を F_i とすると目標周波数を F_t は

$$F_t = \frac{4}{N} \left(\sum_{i=N/2}^{3 \times N/4} F_i \right) \quad (1)$$

と表わされる。

次にテンプレートの基準となるF0を設定する。図1より、F0の隆起と沈降の間の上下動が少ない部分は沈降後のF0の安定した箇所とほぼ同じことが確認できるため、歌唱の際に目標とする音の高さに相当すると考えることができる。そこで、この隆起と沈降の間のF0の平均値をテンプレートの基準のF0(ゲイン周波数と呼ぶ)とする。ゲインは楽曲中で音高遷移が起こった直後に表現されるため、音韻のアライメント情報から発音開始時刻に合

わせて、ゲイン周波数と入力音声から決定される目標周波数が一致するように置き換える。ただし、ここでは歌詞情報からの音韻とスペクトルのフォルマントにより母音部分にゲインが入るように位置を後ろに移動させる処理も行う。

4.2.4 スペクトルの操作

スペクトルの処理では、ゲインの付加を行う場所の歌詞から、音韻の母音情報を取得しておく。F0における処理で決定した入力音声におけるゲインの挿入位置と合わせて、F0の隆起中のみ母音情報と一致する裏声フィルタを用いてスペクトルを変更する。その際に隆

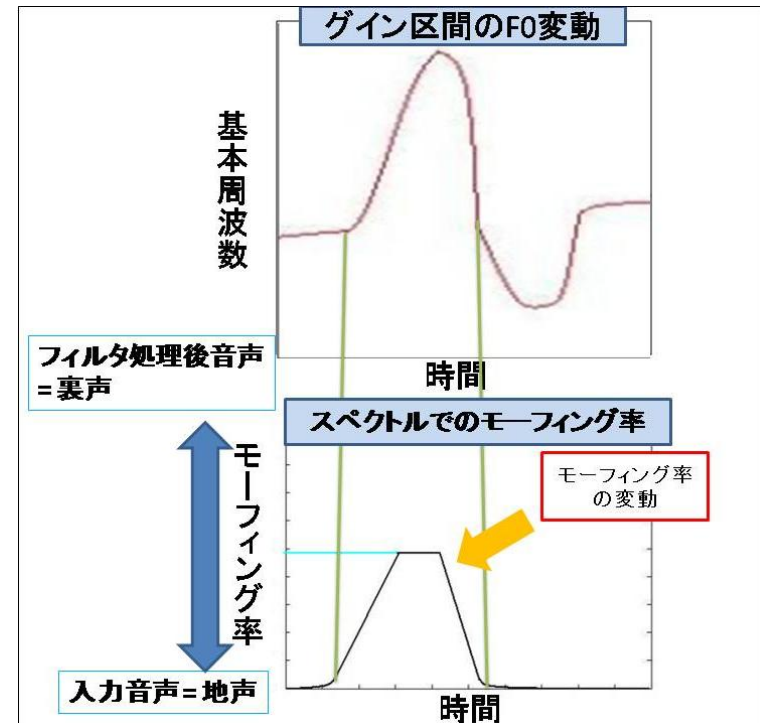


図8 ゲインの付加におけるF0の変動(上図)とそれにあわせたスペクトルのモーフィング率の変化(下図)

起中の音響信号が、隆起開始から隆起の上のピークまでは徐々に裏声フィルタで処理されたスペクトルに近づいてピーク後から隆起が終了までは地声に戻っていくようにスペクトル

の変換を行う。このスペクトルの変換では音声モーフィング¹⁷⁾を利用する。入力音声のスペクトルと裏声フィルタ処理後のスペクトルで図8のようにF0の遷移の仕方にあわせて、モーフィング率を変化させている。

本章の基本周波数の操作ではゲインのF0テンプレートを歌唱音声の目標周波数に置き換える方法を記したが、テンプレートが保持している基本周波数の変動の深さ(ゲイン周波数から隆起と沈降の各ピークまでの周波数の差)は3章で述べたパラメータ P_{f0} の値により調整する。同様に、スペクトルのパラメータ P_{sp} はスペクトルの操作におけるモーフィング率の最大値を決定し、入力音声とフィルリングによる擬似的な裏声の混ぜ合わせ具合を調整することができる。

5. 今後の展望

本研究では、これまで定量的分析の行われていなかった奄美大島出身の歌唱者の歌声の音響特徴の分析を実施し、奄美大島民謡の歌唱技法であるゲインの音響信号の特徴(1)F0の隆起と沈降、(2)裏声による倍音成分の欠落を明らかにした。その上で、一般歌唱にゲインの音響信号特徴(1)(2)を付加し、奄美大島民謡風の歌い回しにするゲインレゾネータの提案を行った。

提案したシステムでは、入力する楽譜情報の音高や歌詞情報からプレス位置での楽曲のフレーズ分割および各フレーズ内で奄美大島出身の歌唱者が楽曲中で付加するであろう位置の候補を推定し、ユーザにGUIのピアノロール上でその位置を提示する。ユーザはGUI上でその候補位置からゲインを付加したい箇所の選択を行い、その位置にゲインの音響信号特徴(1)(2)を付加することで簡単に奄美大島民謡風の歌声が生成できる。この方式によりユーザの自由度を保ちつつ、奄美大島歌唱者がゲインを行わない位置での付加を抑制するシステムとなっている。

本稿でのシステムでは、ゲインを付加する位置は筆者が設けたヒューリスティックなルール(4章)により入力される楽譜情報を使って算出される。現在はこのルールをより奄美大島出身の歌唱者に即したものとするために、実際の奄美大島出身の歌唱者の楽曲から歌唱者それぞれに合わせた、ゲインを付加しやすい位置を求めるための挿入ルールとゲインのパラメータの値を制御するルールが決定する学習機構を製作中である。機械学習などで用いられる混合ガウスモデル(GMM)からなる。現在構築中の学習機構では、観測変数を音符がもつゲインパラメータ、プレス位置でフレーズ分割した奄美大島歌唱者の楽曲から各音符に

対しての音価、前後の音符との音高差情報、フレーズ内での相対的位置を確率変数としてGMMをモデル化している。GMMは設定した確率変数はすべての種類を使い一つ作るのではなく、使う確率変数とその組み合わせ方を変えて複数作成する。そして、本システムでは確率変数が異なるGMMの中からもっとも性能評価が良いものをゲインの挿入位置を推定するためのGMMとする。この最も性能評価が良いGMMを構成する確率変数がゲインを付加しやすい位置を算出するルール(挿入ルール)となる。この学習機構を用いて、奄美大島出身の歌唱者の楽曲をそれぞれ別々に学習されることで歌唱者別のゲインの挿入場所のモデル化が可能となる。GMMを基盤とした学習機構の完成と学習機構によるゲイン挿入ルールの性能評価が今後の課題である。

参考文献

- 1) ヤマハ株式会社. Vocaloid, 2003年 <www.vocaloid.com/jp/index.html> (2010年2月現在)
- 2) クリプトン・フューチャー・メディア株式会社, キャラクターボーカルシリーズ. <<http://www.crypton.co.jp/mp/pages/prod/vocaloid/>> (2010年2月現在)
- 3) 林経正, 上村弥生, エンタテインメント業界への導入が進む歌声合成エンジン「ワンダーホルン」, NTT技術ジャーナル(2005)
- 4) 歌声合成ツール UTAU サポートページ. <<http://utau2008.web.fc2.com/index.html>> (2010年2月現在)
- 5) 中野 倫靖, 後藤 真孝, VocaListener: 事例を与えてインタラクティブに歌声合成できるメタ歌声合成インタフェース, 情報処理学会インタラクシオン 2009
- 6) Hajime Chitose Official Website. <<http://www.office-augusta.com/hajime/>> (2010年2月現在)
- 7) 松岡文吾, 長根における韻律的特徴のモデル化とそれに基づく五線譜からの長唄的韻律遷移の生成(東京大学修士論文(2003))
- 8) 桂博章, 秋田民謡についての覚書-「小節(こぶし)」について-, 秋田大学教育学部研究紀要(1998)
- 9) Uta-Net INTERVIEW. <<http://blog.uta-net.com/kizuki/>> (2010年2月現在)
- 10) 中孝介オフィシャルホームページ. <<http://www.atarikousuke.jp/>> (2010年2月現在)
- 11) Kawahara, H., Masuda-Katsuse, I. and de Cheveigné, A.: Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction, *Speech Communication*, Vol.27, No.3-4, pp.187-207 (1999).
- 12) 河原英紀, 生駒太一, 森勢将雅, 歌唱音声モーフィングに基づく声質と歌い回し転写の知覚的検討, 情報処理学会インタラクシオン 2007
- 13) 月溪恒子, 邦楽と洋楽の歌唱表現法 - 音響的特徴の比較 -, 大阪芸術大学芸術研究所 研究調査報告書(2001)
- 14) 虫明真砂子, 日本の伝統的な歌唱授業の試みに関する考察, 岡山大学教育実践総合センター紀要(2006)
- 15) イルカなごり雪 ZLS-7 (1988)
- 16) 中里南子, J・ポップにみられる装飾的旋律の歌い方-平井堅・桑田佳祐・ケミストリー・ドリカムの「コブシ」の分析を通して-
- 17) Kawahara, H. and Matsui, H.: Auditory morphing based on an elastic perceptual distance metric in an interference-free time-frequency representation, *Proc. ICASSP 2003*, Vol. I, Hong Kong, pp.256-259 (2003).