

奄美大島民謡風歌声合成のためのコブシに着目した歌声の特徴分析[†]村主大輔[†] 森勢将雅^{†,††} 北原鉄朗^{†,††} 片寄晴弘^{†,††}[†] 関西学院大学理工学部 ^{††} 立命館大学情報理工学部 ^{†††} JST CREST CrestMuse

1 はじめに

日本各地には様々な民謡があり、それぞれに独特の歌い回しが存在する。その中でも、奄美諸島の民謡（島唄と呼ばれる）は、近年の日本のポピュラー音楽にもそれに近い歌い方をする歌手が現れるなど、その注目度は高まりつつある。しかし、その歌い方は伝承によって伝えられることが多く、特徴を定量的に分析した研究はほとんど存在しない。

そこで、本研究では、島唄の顕著な特徴の1つである「ゲイン」（コブシの一種）に着目し、その特徴をスペクトル分析によって分析する。また、通常の歌唱音声にゲインを付加する音声加工システムを開発し、analysis-by-synthesisによる分析も行う。

2 ゲインのスペクトル分析

「ゲイン」とは、島唄特有の歌唱技法でコブシの一種である。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 [1]。そのため、ゲインは基本周波数（F0）が変動するだけでなく、それに合わせて音色も変動する。ただし、実際の F0 やスペクトルを分析した事例はほとんど見られない。

ゲインの特徴のスペクトル分析には、河原らの STRAIGHT [2] を用いる。分析対象として奄美大島出身の歌手である中孝介の『花』の「言い訳さえも」というフレーズを選んだ。このフレーズから求めた F0 とスペクトログラムを図 1 に示す。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。図 1 から次の 2 つの事柄が確認できる。

- 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）。
- 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が非常に小さくなっている。これは、文献 [1] での指摘の通り、裏声を用いられたからであると考えられる。

3 一般歌唱に対するゲインの付加によるゲインの分析

前章の分析により、(1)F0 の隆起と沈降、(2) 裏声による高次倍音の欠落がゲインの特徴として観察された。本章では、これらの特徴を一般歌唱に付加するとゲインらしく聞こえるかどうかを検証することで、この 2 つの特徴がゲイン特有の特徴であることを明らかにする。

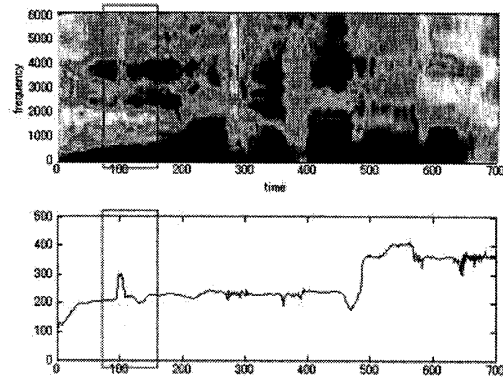


図 1: ゲインにおけるスペクトログラム (上図) と F0 (下図) の特徴的な変動

3.1 一般歌唱に対するゲイン特徴の付加方法

本手法では、ゲインを付加したい元の音声信号（対象音声と呼ぶ）の他に、同一歌唱者による裏声の音声信号（ゲインを付加したい箇所と同じ音素・音高）およびゲインが用いられている音声信号（参照音声と呼ぶ）が必要である。今回は、参照音声は対象音声と同一楽曲とした。この 3 つの音声信号を STRAIGHT によって F0 成分とスペクトル成分とに分解した後、F0 については、ゲイン付加箇所の対象音声の F0 を参照音声の F0 で置き換えることで、スペクトルについては、ゲイン付加箇所のスペクトルを音声モーフィング [3] によって裏声のスペクトルに近づけることでゲインの特徴を付加する。処理の概略を図 2 に示す。

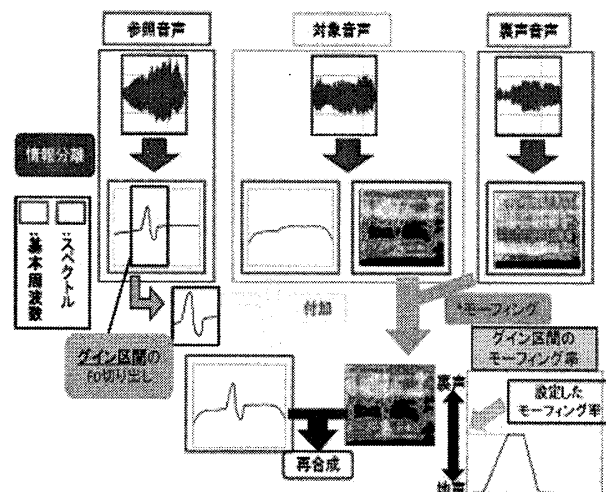


図 2: アルゴリズムの概略

[†] Analysis of Japanese traditional singing technique 'kobushi' for synthesis of singing voices with the style of Amami folk songs by Daisuke Suguru[†], Masanori Morise^{†,††}, Tetsuro Kitahara^{†,††}, and Haruhiro Katayose^{†,††} ([†]Kwansei Gakuin Univ., ^{††}Ritumeikan Univ., ^{†††}CrestMuse, CREST, JST)

3.2 ギン特徴の付加によるギンの表現例

上記での方法でギン付加を行った。一般歌唱者に中孝介の『花』の「言い訳さえも」の部分歌ってもらい、中孝介の原曲を参照音声として用いた。

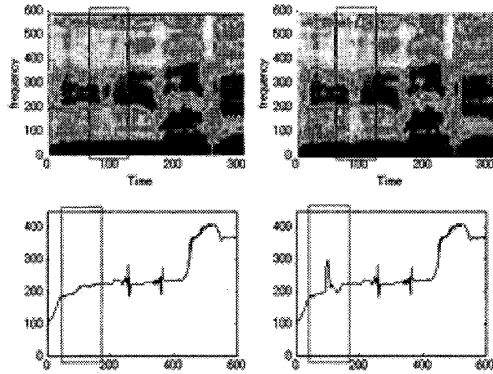


図 3: 一般歌唱者の音声 (左側) とギン特徴を付加した音声 (右側) のスペクトルと F0 軌跡

生成された音声の F0 とスペクトログラムを図 3 に示す。図 3 の F0 変動の付加前後の四角で囲われた部分と、図 1 の F0 軌跡部分の四角で囲われた部分を比較すると、中孝介の F0 変動を付加出来ているのがわかる。また同様に、図 3 のスペクトル変動の付加をした部分のスペクトルは、図 1 のギン部分のスペクトルのようにパワーが減少している。

3.3 聴取実験

F0 の参照音声とのモーフィング率 R_{F0} およびスペクトルの裏声音声とのモーフィング率 R_{Sp} をそれぞれ 0%, 50%, 75%, 100% とした 16 個の音声を作成し、モーフィング率の異なる全 240 組の対に対してどちらが中孝介の歌い回しに近いかを 2 名の被験者に回答してもらった。なお、中孝介の音声は各々の対を聴く直前に毎回聴いてもらった。

実験結果を図 4 に示す。被験者 A に着目すると、 $(R_{F0}, R_{Sp}) = (100\%, 0\%)$ や $(R_{F0}, R_{Sp}) = (0\%, 100\%)$ のときよりも、 $(R_{F0}, R_{Sp}) = (75\%, 75\%)$ や $(R_{F0}, R_{Sp}) = (100\%, 100\%)$ のときの方が、より中孝介に近いと判定された。これは、ギンにおいて F0 の変動とスペクトルの変動の両方の存在が重要であることを示している。また、スペクトル変動を導入しない場合よりも裏声とモーフィングさせた方が中孝介に近いと判定されている。このことは、文献 [1] で指摘されているように、裏声がギンならではの特徴であることを示唆している。

一方、被験者 B では上述の結果が当てはまっていない。この原因として、被験者 A に比べて被験者 B は歌唱の特徴の細かな違いにさほど敏感ではない可能性がある。実際、被験者 A が楽器経験者であるのに対して、被験者 B は楽器未経験者である。このことを確かめるため、次の実験を行った。2 章で用いた中孝介の音声信号から STRAIGHT を用いて意図的にギンを取り除いた音声信号を用意する。この音声と元音声とを聞き比べ、元音声を回答する実験を行った。その結果、被験者 A は正答率が 100% だったのに対し、被験者 B は、ギンを F0 のみ取り除いたときに正答率が 65% だった (図 5)。このことから、被験者 A に比べて被験

者 B は、ギンの細かな特徴の違いを聞き分けられなかったと考えられる。

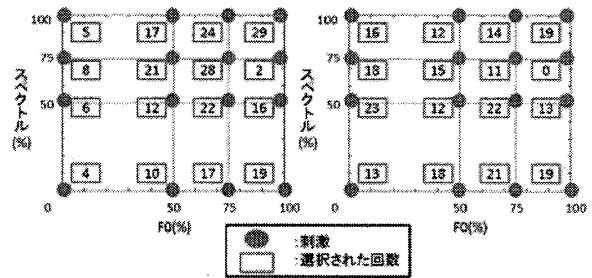


図 4: 中孝介の歌い回しに近いと選ばれた回数 (左図:A, 右図:B)

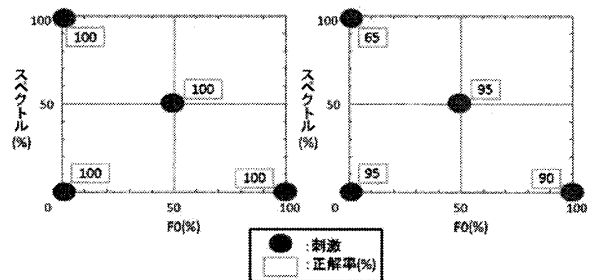


図 5: 原曲以外の比較音声ごとの正解率 (左図:A, 右図:B)

4 おわりに

本稿では、島唄特有のコブシであるギンをスペクトル分析したところ、(1)F0 の隆起と沈降、(2)裏声による倍音成分の欠落が観測された。この観察結果を元に、一般歌唱の音声信号に F0 とスペクトルの変動を付加したところ、ギンに近く聞こえるという結果を得た。

この分析を進める過程で、ギンらしく聞こえるには、F0 変動とスペクトル変動の同期性が重要であることが分かってきた。今後は、本稿で報告した実験を被験者数を増やして実施するだけでなく、F0 変動とスペクトル変動の同期性についても検討を行っていく。

参考文献

- [1] <http://blog.uta-net.com/kizuki/>
- [2] Kawahara, H., Masuda-Katsuse, I. and de Cheveign e, A.: Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequencybased F0 extraction, *Speech Communication*, Vol.27, No.3-4, pp.187-207 (1999).
- [3] 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).