

# 声道形状と声帯音源特性を用いたグロウル系 歌唱音声の印象付与の評価

溝淵翔平<sup>1,a)</sup> 西村竜一<sup>1,b)</sup> 入野俊夫<sup>1,c)</sup> 河原英紀<sup>1,d)</sup>

**概要：**本研究では提案法を用いて通常歌唱音声にグロウル系歌唱の特徴を付与した際の印象を評価した。これまでの研究よりグロウル系歌唱音声特有の物理的特徴として「1k~4kHzの帯域強調」、「基本周波数の振動」及び、「スペクトル形状の高速な時間変動」が確認された。従来法である「スペクトル形状の高速な変動」を付与したモデルは、観察された現象を表面的に模擬するために4個のガウス関数を組み合わせたものであり、声質の表現や発声の機構を考慮したものでは無かった。本研究では「スペクトル形状の高速な時間変動」を披裂喉頭蓋の形状変化と声帯音源波形の時間変化の相互作用としてモデル化することで、グロウル系歌唱音声の特徴を付与する手法をこれまでに提案した。本稿では、従来法と提案法を変換後の歌唱音声の対比較実験により評価した。結果をサーストンの対比較法により分析した結果、提案法がグロウル系歌唱音声の印象を付与するのに効果的であることが示唆された。

**キーワード：**グロウル, 声道音源特性, 声道形状, 声質変換, TANDEM-STRAIGHT

## Evaluation of singing voice conversion to growl-like singing based on vocal tract shape and glottal source characteristics

**Abstract:** We have investigated a system to convert usual singing voice to growl-like performance. Relatively high-speed periodic variations in spectral shapes, fundamental frequency trajectories and spectral level enhancement in the 1kHz to 4kHz frequency band were introduced as dominant features of growl-like singing in our previous works. In the previous method, the high-speed periodic variations in spectral shapes designed using a set of Gaussian functions did not take account of a vocalization mechanism. This study has proposed a method that spectral shape variations can be closely replicated by introducing vocal tract shape variations around supra-laryngeal structure and shape variations in glottal source waveform using the LF-model. In this paper, we have evaluated the proposed method with the previous method by using a pair comparison method. The analysis of the experimental results suggested that the production of singing having the growl-like impression by the proposed method have better than that of the previous method.

**Keywords:** Growl, source characteristics, vocal tract shape, voice conversion, TANDEM-STRAIGHT

### 1. はじめに

歌唱では、情緒を豊かに表現する為に、様々な声質が用いられる。本研究では特に声帯振動に加えて、声帯上部にある披裂喉頭蓋ひだの振動との相互作用によって独特な印象を付与された歌唱音声に注目する [1]。これらの特徴

をもつ歌唱音声をグロウル系歌唱音声と呼び、伝統的な歌唱や民族歌唱からロックやメタルなど幅広い音楽のジャンルで使用されている。グロウル系歌唱音声にみられる独特な印象を自由に操作出来る技術の開発は、歌唱の応用に大きく貢献出来る可能性を有している。先行研究として、実際のグロウル系歌唱音声から取り出した特徴を、分析合成音声に転写することでグロウル系歌唱音声を実現する方法が提案されている [2]。また、歌声合成エンジンである VOCALOIDV4[3] では、新機能として「グロウル」のエフェクトが搭載されており、合成歌唱音声に「唸り声」の

<sup>1</sup> Wakayama University, Wakayama 640-8510, Japan

a) s155059@wakayama-u.ac.jp

b) nisimura@sys.wakayama-u.ac.jp

c) irino@sys.wakayama-u.ac.jp

d) kawahara@sys.wakayama-u.ac.jp

ような印象を与えることが出来る。

一方で、我々のこれまでの研究では、フィルタリングと変調からなるシンプルな処理により通常歌唱音声にグロウル系歌唱音声の物理的特徴を付与する方法を提案した [4]。本手法におけるグロウル系歌唱音声特有の物理的特徴とは「1~4kHz 帯域の強調」、「基本周波数の高速な変動」、「スペクトル形状の高速な変動」を示す [5]。また音声の聴取実験より、これらの物理的特徴を付与された音声はグロウル系歌唱音声の印象を持つことを確認した [6]。しかし、提案法において「スペクトル形状の高速な変動」を付与するために用いたモデルは、観察された現象を表面的に模擬するために4個のガウス関数を組み合わせたものであった。これは、声質の表現や発声の機構を考慮したものではなかった。

この課題に対し、[7]では声帯音源特性と声道形状の周期的変化を組み合わせることで通常歌唱のスペクトル形状にグロウル系歌唱音声に認められる周期的変化を付与する方法を提案した。本稿では従来法と、[7]で提案した声帯音源特性と声道形状の周期的変化の組み合わせによる特徴付与処理について、変換音声のグロウル系歌唱音声に対する印象を主観評価実験より調査する。

2節では提案法による声道形状と声帯音源特性に着目したスペクトル形状の高速な変動の付与について概要を述べる。3節では提案法と従来法を比較し、印象評価実験を行った結果から得た特徴付与による歌唱音声の印象の変化について述べる。最後に4節で全体のまとめと今後の課題を述べる。

## 2. 提案法によるスペクトル形状の高速な変動の付与

グロウル系歌唱音声の生成系においては、披裂喉頭蓋ひだが声帯振動の約1/3程度の周波数で振動するという特徴がある。この披裂喉頭蓋ひだが振動することにより、次の2つの変化が考えられる。

- 披裂喉頭蓋ひだによる局所的な声道形状の変化による共振特性の変化
  - 声道形状の変化に伴って、駆動点から見た声道のインピーダンスが変化することによる声帯音源特性の変化
- これまでに、声道形状の変化と声帯音源特性の変化を分析した。まず声道断面積関数の時間変化を分析することで、グロウル系歌唱音声特有の声道形状の時間変化を調査した [8]。結果、披裂喉頭蓋ひだの影響と考えられる13~15cmあたりの周期的な声道断面積関数の変動が確認され、声道形状の変化が示唆された。また、1オクターブ幅で平滑化した歌唱音声の長時間平均対数スペクトルとフレーム毎の対数スペクトルの差分について帯域ごとのRMSを求め、各歌唱におけるスペクトルの変動を可視化した [9]。結果は、グロウル系歌唱音声の長時間平均スペクトルとフレーム毎の差分は通常歌唱よりも全帯域において変動

量が大きく、また、LF modelのシミュレーションから、高域におけるスペクトル傾斜の変化が示唆された。

以上の知見から声道断面積関数と声帯音源特性を用いて、スペクトル形状の高速な変動を付与する手法を提案した [7]。まず、提案手法による声道断面積関数を用いた声道形状の変化を反映する特徴付与処理について述べ、次にLF modelを用いた声帯音源特性の時間変化を付与する処理について述べる。

### 2.1 声道断面積関数を用いたスペクトル形状の高速な変動の付与

#### 2.1.1 声道断面積関数の導出

次節では、まず歌唱音声についてTANDEM-STRAIGHTより求められたスペクトルに対して、F0に適応した声帯音源と放射特性の補償 [10]を用いて声道特性以外の特性を抑圧した。これは、LF model [11]のシミュレーションより得られた平均的な声帯音源特性を声帯音源と放射特性を補償に用いる方法で、適応時には入力音声の基本周波数に応じてスペクトルの概形をスケールリングしてある。シミュレーションは基本周波数を50Hzに固定して時間パラメタをランダムに生成し、「地声」「フライ」「気息性」の3つの声質についてそれぞれ3000回のシミュレーションを行った。時間パラメタの決定にはChildersらによる平均と標準偏差 [12]を用いた。

本研究で扱う声道断面積関数はKellyの声道モデル [13]に基づき、図1に示すような声道を断面積一定の微小音響管の縦列接続で表現したものである。図1の $l$ は微小音響管の長さ、 $A_n$ は区間 $n$ の断面積、及び $\kappa_n$ は区間 $n$ と区間 $n+1$ 間の反射係数を示す。声道特性以外の特性を抑圧したパワースペクトルに関して6000Hzまでの帯域を取り出し、自己相関関数を求め、レビンソンダービンの方法よりPARCOR係数を求めた。この時、LPC次数は12とした。ここで、PARCOR係数 $k$ は反射係数 $\kappa$ と対応しており、反射係数 $\kappa_n$ と断面積 $A_n$ の間には以下の関係が成り立ち、声道断面積関数を求めることが出来る。

$$A_n = \frac{1 + \kappa_n}{1 - \kappa_n} A_{n+1} \quad (n = p, p-1, \dots, 1) \quad (1)$$

$$A_{p+1} = 1 \quad (2)$$

また声道断面積関数は体積一定の制約のもと計算している。

#### 2.1.2 声道断面積関数の動的変化の付与

図1において声道断面積関数の区間の長さにあたる $l$  cmは音速 $c$  m/sとサンプリング周波数 $F_s$  Hzを用いることで次式のように表現出来る。

$$l = \frac{100c}{2F_s} \quad (3)$$

披裂喉頭蓋ひだは口唇から約13~15 cm付近にある為、今回は時間変動を加える区間として、9次の区間を中心にモ

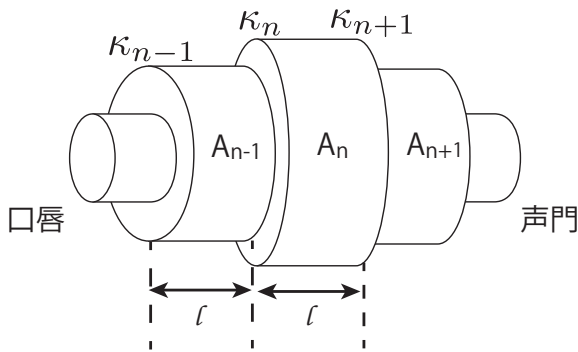


図 1 Kelly の声道モデルに基づく声道断面積関数

ジュレーションを付与することとした。また、式 (1) から分かるように変換後の声道断面積関数から式 (4) を用いることで変換後の反射係数を求めることが出来る。

$$k_n = \frac{A_n - A_{n+1}}{A_n + A_{n+1}} \quad (4)$$

変換された反射係数から変換後の LPC スペクトルを求め、変換後の LPC スペクトルと元の音声のスペクトルの差分を声道形状の変化を与えるフィルタの周波数特性とした。

音声変換の例として、グロウル系歌唱音声の LPC スペクトログラムと通常歌唱音声の LPC スペクトログラム、及び通常歌唱音声に声道形状変化の変換後の LPC スペクトログラムを図 2 に示す。図 2 はグロウル系歌唱音声が特にグロウルの特徴を持つ  $e/\epsilon$  の発話区間であり、通常歌唱音声はグロウル系歌唱音声と同一の発話区間である。変換された通常歌唱音声は音声の  $1/3$  の周波数で振動する  $\sin$  波状に 9 区間目の声道断面積を変動させた例である。図から分かるように、通常歌唱音声のフォルマントは時間経過に対して一定であるのに対し、グロウル系歌唱音声の  $3k \sim 4k$  Hz における第 3 フォルマント周辺は時間経過に応じて波状に変化していることが分かる。変換後の音声に注目すると、第 3 フォルマント周辺に波状の変化が付与されていることが確認され、グロウル系歌唱音声特有のスペクトル形状の高速な変動を付与出来ることが示唆された。

## 2.2 LF model を用いたスペクトル形状の高速な変動の付与

声門上部の構造の振動により、声門から声道側を見た駆動点インピーダンスは、周期的に変動する。その結果、声帯振動の速度が変化し、声門閉止の鋭さが変化する。ここでは、鋭い閉止の状況を modal 発声でモデル化し、緩い閉止の状況を breathy 発声でモデル化した。この仮定をもとに前述した声道形状の変化に合わせて、LF model による modal 時の声帯音源特性と breathy 時の声帯音源特性を発生させた。

### 2.2.1 LF モデルを用いた声帯音源特性

LF model[11] は、声門での体積流の波形に唇からの放射特性の影響を加えた波形を  $E(t)$  とし、次式により定義さ

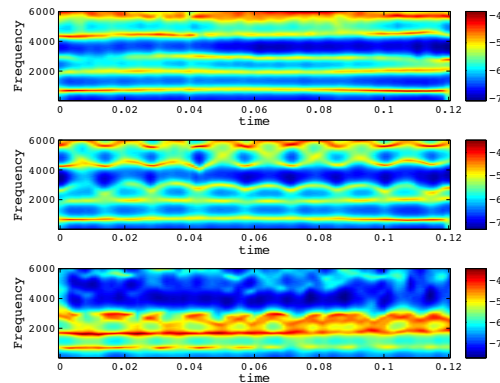


図 2 声道断面積関数の操作による LPC スペクトログラムの変化の比較 (上段：通常歌唱音声、中段：特徴付与された通常歌唱音声、下段：グロウル系歌唱音声)

れる。

$$E(t) = \begin{cases} E_0 e^{at} \sin \omega_g t & 0 \leq t \leq t_e \\ -\frac{E_e}{\epsilon t_a} [e^{-\epsilon(t-t_e)} - e^{-\epsilon(t_c-t_e)}] & t_e \leq t \leq t_c < T_0 \end{cases} \quad (5)$$

この時、 $t_p$  は声門体積流が最大となる時点、 $t_e$  は声門の閉止時点、 $t_a$  は声門閉止の時定数、 $t_c$  は声門が完全に閉止する時点を示す。また、 $T_0$  は声帯振動の一周期を指し、負の最大値を  $E_e$  としている。ここで、放射特性が周波数 0 に零点を有することから、以下の拘束条件があることがわかる。

$$\int_0^{T_0} E(t) dt = 0 \quad (6)$$

式 5 に含まれる  $\omega_g$  や  $\epsilon$  は、波形のタイミングを表す  $t_p, t_e, t_a, t_c$  から一意に定まる。また、 $E_e$  は、 $t_e$  における連続性から求められる。

図 3 に LF model を利用したシミュレーションの結果の平均として、求められた modal 時と breathy 時の声帯音源波形を示す。また、基本周波数は 50Hz としている。breathy は modal に比べ、 $t_p$  と  $t_a$  の値が大きいことが分かる。これは breathy の場合、声門開口と声門閉止の力が弱く、時間がかかる為である。この影響は声帯音源特性にも現れる。図 4 に modal と breathy の平均的な声帯音源特性を示す。実際の基本周波数を考慮して移動させると、100Hz から 300Hz の間では、breathy は、modal よりもスペクトルの傾斜が急であり、高域のレベルが低くなっていることが分かる。また、基本波と第二次調波のレベルを比較すると、breathy の方が基本波のレベルが高くなっていることが分かる。

### 2.2.2 声帯音源特性の時間変化の付与

前述のように声帯音源特性の時間変化の付与は声道断面積関数の操作に依存する。ここでは仮に、声道形状を広げる方向に変化させる部分は modal の声帯音源波形に近づき、声道形状を狭める方向に変化させる部分は breathy の声帯音源特性に対応させるように設定した。よって、本研

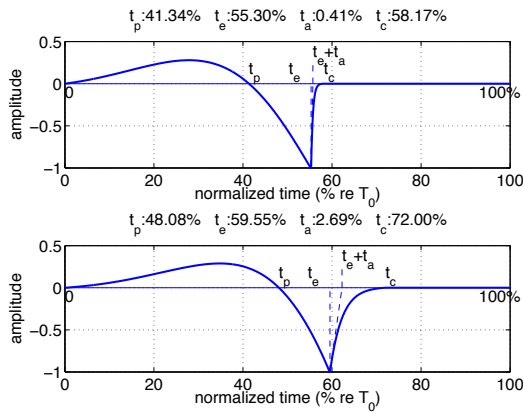


図3 LF modelによる微分声門体積流 (上段:modal, 下段:breathy)

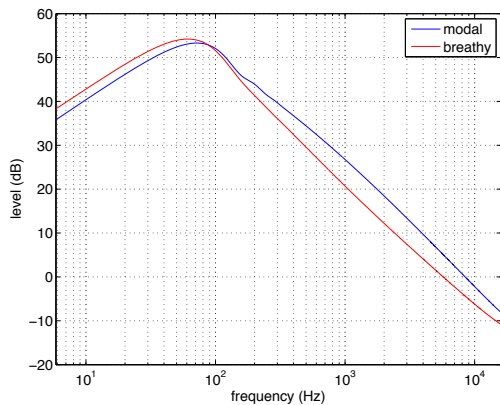


図4 LF modelによる声帯音源特性

究では声道形状を操作するモジュレーション信号  $m(t)$  が  $0 \leq m(t)$  のとき modal に近づき、 $m(t)$  が  $0 > m(t)$  のとき breathy に近づくよう LF model のパラメタが変化するように設計する。例えば、 $t_p(t)$  の場合であれば

$$t_p(t) = m(t)t_p^m + (1 - m(t))t_p^b \quad (7)$$

とする。ここで、 $t_p^m$  は modal 時の  $t_p$  の平均値であり、 $t_p^b$  は breathy 時の  $t_p$  の平均値を示す。ここで、それぞれの時間パラメタの平均値は Childers らによる平均 [12] を用いた。 $t_e, t_a, t_c$  についても同じようにパラメタを求めた。こうして設定した時間パラメタを用いて得られたフレーム毎の声帯音源特性と平均的な modal 時の声帯音源特性との差分を特徴付与フィルタとして用いた。

### 3. 主観評価実験

提案法によってグロウル系歌唱音声の特徴を付与された音声に対し、その印象を従来法による変換歌唱音声と一対比較することで評価した。

#### 3.1 音声試料

音声試料には著者が発話した歌唱音声を録音した通常音声と Reference 歌唱を録音した。この際、著者が想定するグロウルの歌唱方法を Reference となる歌唱音声として作

成した。歌唱音声は3回録音している。これは提示する刺激を1つに限定しない為であり、収録時の歌い方を意識的に変化させてはしていない。

### 3.2 提示する音声刺激

#### 3.2.1 従来法によるスペクトル形状の高速な変動の付与

提案法に対して、対数スペクトルで表現した特徴を、時間的に変動する 3000 Hz 付近の  $-10 \sim -15$  dB 程度の鋭い谷とやや緩いピークからなるペア  $R_1(f)$  と、その変動と逆の位相で変動する 2000 Hz 付近の緩いピークと谷のペア  $R_2(f)$  を用いて、次式のようにモデル化した。

$$R_1(f) = a_{p1} \exp\left(-\frac{(f - f_{p1})^2}{\sigma_{p1}^2}\right) - a_{d1} \exp\left(-\frac{(f - f_{d1})^2}{\sigma_{d1}^2}\right) \quad (8)$$

$$R_2(f) = a_{p2} \exp\left(-\frac{(f - f_{p2})^2}{\sigma_{p2}^2}\right) - a_{d2} \exp\left(-\frac{(f - f_{d2})^2}{\sigma_{d2}^2}\right) \quad (9)$$

ここで  $f_{p1}, f_{p2}$  は、ピークの周波数、 $\sigma_{p1}, \sigma_{p2}$  は、ピークの広がり、 $a_{p1}, a_{p2}$  は、高さを表す。同様に、 $f_{d1}, f_{d2}$  は、谷の周波数、 $\sigma_{d1}, \sigma_{d2}$  は、谷の広がり、 $a_{d1}, a_{d2}$  は谷の深さを表す。この2つの特性を、値域が  $[0, 1]$  になるように変換した変調に持ちいる信号を用いて補間し、時間的に変化する特性  $R_m(f, t)$  を作成した。今回の実験では、 $f_{p1} = 2000, f_{p2} = 3000, f_{d1} = 3000, f_{d2} = 2000, a_{p1} = 2, a_{p2} = 5, a_{d1} = 13, a_{d2} = 4, \sigma_{p1} = \sigma_{p2} = \sigma_{d1} = \sigma_{d2} = 400$  とした。

#### 3.2.2 音声刺激の作成

音声刺激はすべて TANDEM-STRAIGHT を用いて分析変形合成した。TANDEM-STRAIGHT のフレームレートの規定値は 1ms とした。STRAIGHT スペクトルに以下の8条件で特徴付与した。

- **Plane** : 平板に歌った音声の単純な分析合成音
  - **old** : 従来法による特徴付与と処理を施した合成音
  - **VTF** : 提案法による声道断面積関数の操作による特徴付与のみを反映した合成音
  - **VFC** : 提案法による声帯音源特性の変化をモデル化した特徴付与のみを反映した合成音
  - **NoShift** : 提案法による声帯音源特性と声道断面積関数の操作の両方の特徴付与を反映した合成音 (声帯音源特性の操作に用いる信号と声道断面積関数の操作に用いる信号の位相差なし)
  - **$\pi$ Shift** : 提案法による声帯音源特性と声道断面積関数の操作の両方の特徴付与を反映した合成音 (声帯音源特性の操作に用いる信号と声道断面積関数の操作に用いる信号の位相差が  $\pi$ )
  - **$\pi/2$ Shift** : 提案法による声帯音源特性と声道断面積関数の操作の両方の特徴付与を反映した合成音 (声帯音源特性の操作に用いる信号と声道断面積関数の操作に用いる信号の位相差が  $\pi/2$ )
  - **Growl** : Reference 歌唱の単純な分析合成音
- また、それぞれの特徴付与に用いる変調信号は全て同じ信



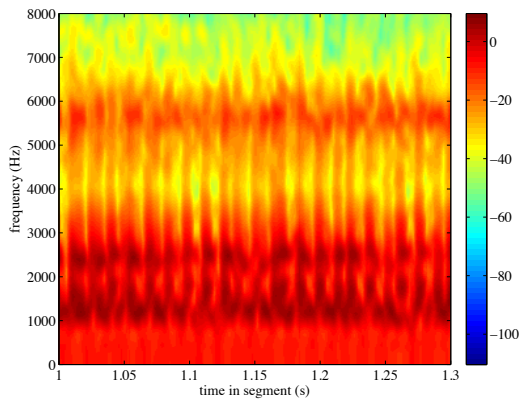


図 5 提案法による特徴付与を施した歌唱音声

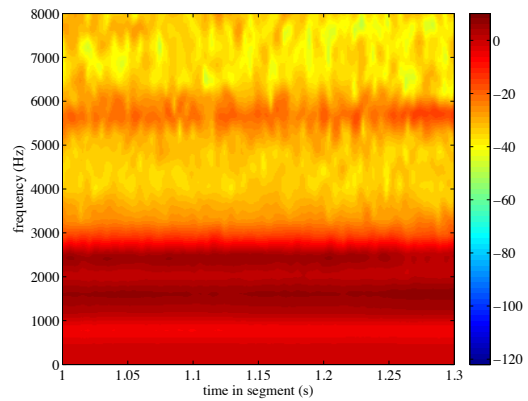


図 8 特徴付与と処理を行っていない通常歌唱音声

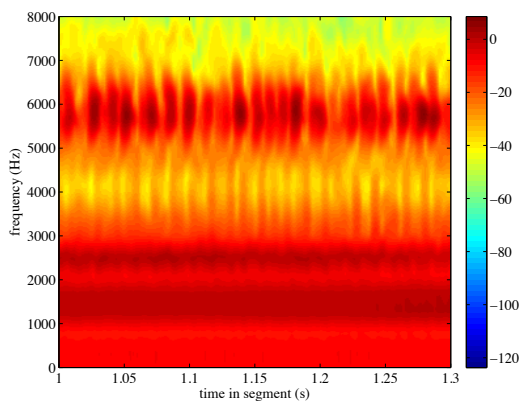


図 6 従来法による特徴付与を施した歌唱音声

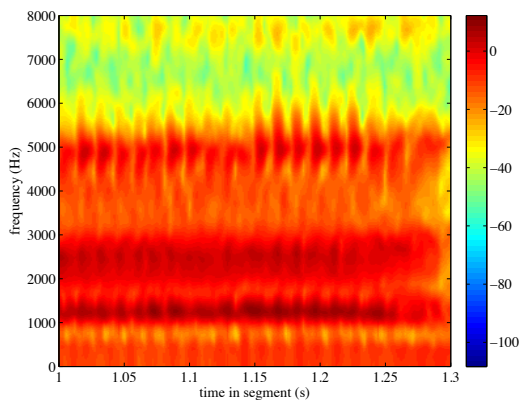


図 7 Reference の歌唱音声

号を用いており、基本周波数の 1/3 の周波数に位相同期した sin 波を用いた。図 5~8 にそれぞれの処理によって変換された音声の STRAIGHT スペクトログラムを示す。図 5 と図 6 はそれぞれ提案法と従来法による特徴付与と処理をした STRAIGHT スペクトルを表す。図 7 と図 8 はそれぞれ Reference 歌唱と特徴付与前の平板に歌った音声の STRAIGHT スペクトルを表す。

### 3.3 被験者

実験に協力を得た被験者は 21~24 歳の男性 2 名、女性 4 名の計 6 名である。被験者はいずれも日本語を母国語と

しており、聴力、発話に関する病歴のない健常者である。各被験者には文章にて実験参加の同意を得た後、実験を行った。

### 3.4 実験方法

実験では、防音室において、MacBook Pro から D/A converter(EDIROL UA-101)、ヘッドフォンアンプ(Audio-Technica AT-HA60)及び、ヘッドフォン(SENNHEISER HD-580)を介した両耳での受聴を行った。再生システムの音圧レベルは、LAeq で音圧が 70dB になるように設定した。

受聴時は、信号音に続いて、最初に印象評価の基準となる Reference 歌唱が提示され、続いて比較対象である合成音 A、合成音 B が順に提示される。被験者は A と B のうち、どちらがターゲット音声の印象に近いかを一対比較で行う。被験者にはグロウル系歌唱音声を選択する実験であることは明示せず、Reference 歌唱の印象により近い音声を選択するように依頼した。また、Reference 歌唱は前述のように 3 種類用意しており、Reference 歌唱の提示は毎回 3 種類の中からランダムに決定した。つまり実験は一人あたり、音声刺激 8 条件に対し、同じ条件の組み合わせを除く 56 条件を 3 個のオリジナル音声に対して実施するため、計 168 回の一対比較を行うこととなる。

### 3.5 実験結果

図 9 にサーストンの一対比較法による実験の分析結果を示す。図の横軸は、Reference 歌唱に対する印象の近さを示す間隔尺度を示す。スコアが大きいほど、Reference 歌唱の印象に近いことを表している。全ての条件において提案法は従来法に比べて Reference 歌唱の印象に近づいているということが確認出来た。

また、声道形状のみを操作した変換音声 (VTF) は、声帯音源特性のみを操作した変換音声 (VFC) 及び、声道形状と声帯音源特性を組み合わせて特徴付与した条件の 2 つ ( $\pi/2$ Shift、 $\pi$ Shift) よりも Reference 歌唱の印象に

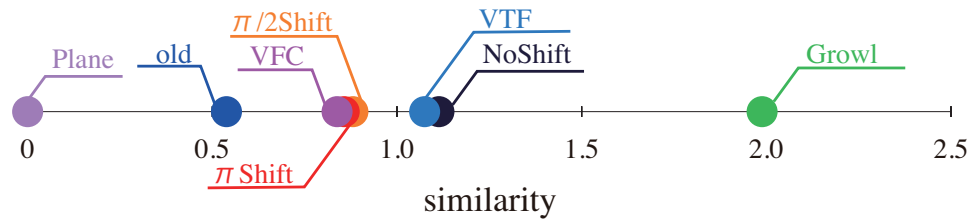


図 9 聴取実験結果より得られた刺激音声の間隔尺度

近いことを示す結果を得た。この結果から、声道形状のみを操作する処理 (VTF) は声帯音源特性を操作する処理  $\pi/2\text{Shift}$ 、 $\pi\text{Shift}$  よりもグロウル系歌唱音声の印象を付与することにおいて有効であることが示唆された。実験後、インタビューにより意見を聴取した際、声帯音源特性を操作した変換音声 (VFC、 $\pi/2\text{Shift}$ 、 $\pi\text{Shift}$ 、NoShift) は誇張が強すぎ、不自然であるという意見があった。

#### 4. おわりに

今回、従来法よりも声道形状の周期的変化と声帯音源波形の周期的変化を組み合わせた提案法が、グロウル系歌唱音声の印象の付与に有効であることが示唆される結果が示された。現段階では被験者の人数が 6 人で留まっているが、被験者を増やし、実験を継続する予定である。

今回の方法は、[4] で提案された「スペクトル形状の高速な変動」の特徴付与を、音声の機構を考慮したより合理的な特性に置き換えるためのものである。そのため、実時間でのグロウル系歌唱音声への変換の場合には、ここで得られた知見を、実時間処理に適した形で実装する必要がある。分析合成を行わずに、フィルタ処理と、時間軸の伸縮のみで実現することにより、品質劣化の回避することが出来る。LF model によるスペクトル形状の変化は、スペクトル傾斜の変化により近似することができ、簡単なフィルタで実現することができる。声道形状の変化によるスペクトルの変化は、LPC に基づく処理が軽いものであることから、そのまま実時間システムに組み込むことも可能である。これらを組み合わせて実際に動作するシステムとして実装することが、今後の課題である。

**謝辞** 本研究を進めるにあたり、実験計画の方針に関してご助言頂きました松井淑恵助教に厚く御礼申し上げます。また、実験結果の分析にあたりご助言を頂きました吉田駿氏に深く感謝致します。本研究は、科研費挑戦的萌芽研究 26540087 による支援を受けた。

#### 参考文献

- [1] K. Sakakibara, et al.: Growl voice in ethnic and pop styles, *Proc. Int. Symp. on Musical Acoustics* (2004).
- [2] B. Jordi, B. Merlijn, 才野, 久湊: スペクトルモーフィングによるグロウル系統の歌唱音声合成, 情報処理学会研究報告, MUS-100-24, pp.1-6 (2013).
- [3] <http://www.vocaloid.com/vocaloid4/>
- [4] 河原英紀, 溝渕翔平, 森勢将雅, 榊原健一, 西村竜一, 入野俊夫: 非線形振動子による変調と近似時変フィルタに基づくグロウル系統の歌唱への実時間変換の定式化について, 情報処理学会研究報告, MUS-102-14, pp.1-6 (2013).
- [5] H. Kawahara and M. Morise and K. Sakakibara: Interference-free observation of temporal and spectral features in “shout” singing voices and their perceptual roles, *Proc. SMAC-SMC*, pp.256-263 (2013).
- [6] 溝渕翔平, 西村竜一, 入野俊夫, 河原英紀: グロウル系統の歌唱音声にみられるスペクトルの時間変動に注目した分析と再現の検討, 日本音響学会 2014 年春季研究発表会講演論文集, pp.499-500 (2014).
- [7] 溝渕翔平, 西村竜一, 入野俊夫, 河原英紀: グロウル系歌唱から求められる声道断面積関数の特徴について, 電子情報通信技術研究報告, SP2014-80, pp.29-30 (2014).
- [8] 溝渕翔平, 西村竜一, 入野俊夫, 河原英紀: 声道形状と声帯音源特性を利用したグロウル系歌唱音声への変換について, 日本音響学会 2015 年春季研究発表会講演論文集 (2015).
- [9] 溝渕翔平, 西村竜一, 入野俊夫, 河原英紀: 声道形状と声帯音源特性を用いたグロウル歌唱音声への実時間変換の提案, 第 106 回音楽情報科学・第 35 回エンタテインメントコンピューティング合同研究発表会, 情報処理学会研究報告, Vol.2015-MUS-106, No.12, (2014).
- [10] 伊佐衣代, 西村竜一, 入野俊夫, 河原英紀: 声道断面積関数推定における声帯音源特性の補償について, 日本音響学会 2015 年春季研究発表会講演論文集, (2015).
- [11] G. Fant.: The LF-model revisited. Transformations and frequency domain analysis., *Speech Trans. Lab. Q. Rep., Royal Inst. of Tech.*, vol. 2-3, pp. 121-156 (1995).
- [12] D G Childers and C K Lee.: Vocal quality factors: analysis, synthesis, and perception, *The Journal of the Acoustical Society of America*, vol. 90, pp. 2394-2410 (1991).
- [13] J. L. Kelly and C. C. Lochbausondm: Speech synthesis, *Proc. 4th International Congress on Acoustics*, G42, pp. 1-4 (1962).