

声帯一仮声帯振動の喉頭音源モデルによる喉歌合成

今川 博*¹ 榊原 健一*² 小西 知子*¹ 村野 恵美*³ 新美 成二*¹

*¹ 東京大学大学院 医学系研究科
音声・言語医学教室
〒113-0033 東京都文京区本郷 7-3-1
03-5841-3575
imagawa@m.u-tokyo.ac.jp

*² NTTコミュニケーション科学基礎研究所
〒243-0198 神奈川県厚木市森の里若宮 3-1
046-240-3657
kis@brl.ntt.co.jp

*³ 東京大学大学院 医学系研究科
耳鼻咽喉科学教室
〒113-0033 東京都文京区本郷 7-3-1
03-5800-3641

あらまし 筆者等は先に、トウパのフーメイ、モンゴルのフーミーなどアルタイ山脈周辺部の民族に伝承される喉歌の歌唱時の喉頭調節をEGGと高速度デジタル撮影法で観測し、報告した。本稿では同音声資料について逆フィルタ法による喉歌の喉頭音源の推定を試みた。その結果と先の生理的観測結果から、喉詰発声では声帯と仮声帯の振動はほぼ逆位相になっていること、カルグラ一発声では声帯が推定音源体積流の1周期に2回振動し、そのうち1回は声門閉鎖が不完全であることが推測された。そこで、声帯の2質量モデルを拡張した声帯一仮声帯自励振動モデル(4質量モデル)を作成した。このモデルにより、仮声帯を狭めた時に声帯と仮声帯が、推測された振動パターンで振動する可能性が高いことを確認した。本稿では、さらに声帯と仮声帯の振動パターンを考慮した喉頭音源信号モデルを提案し、喉歌合成を試みた。その結果、従来の喉頭音源モデルを用いた場合に比べ、実際の喉歌音に近い聴覚印象が得られた。

キーワード 喉歌、声帯振動、仮声帯、逆フィルタ、喉頭音源モデル、自励振動モデル

Throat Singing Synthesis by a Laryngeal Voice Model Based on Vocal Fold and False Vocal Fold Vibrations

Hiroshi Imagawa*¹ Ken-Ichi Sakakibara*² Tomoko Konishi*¹ Emi Zuiki Murano*³ Seiji Niimi*¹

*¹ Department of Speech Physiology,
Graduate School of Medicine,
The University of Tokyo
7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan
+813-5841-3575
imagawa@m.u-tokyo.ac.jp

*² NTT Communication Science Laboratories
3-1, Morinosato Wakamiya, Atsugi-shi, Kanagawa,
243-0198, Japan
+8146-240-3657
kis@brl.ntt.co.jp

*³ Department of Otolaryngology,
Graduate School of Medicine,
The University of Tokyo
7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan
+813-5800-3641

Abstract In this paper, the authors estimated laryngeal voices of throat singing using an inverse-filtering method and proposed a laryngeal voice model based on vocal fold and false vocal fold vibrations. In the previous paper, the physiological observations on laryngeal movements of throat singing by means of the simultaneous recording of singing sounds, EGG, and high-speed digital images were reported. From the physiological observations and estimated laryngeal voices, we assumed (1) in pressed voice production, vocal folds and false vocal folds vibrated in almost opposite phase; (2) in karygraa voice production, two closed phases of vocal folds were appeared in one period of a glottal volume flow waveform, and at one of the two closed phases, vocal folds were incompletely closed. The authors also proposed a four-mass model as a self-oscillating model of vocal fold and false vocal fold vibrations which is developed from a two-mass model. The simulation of laryngeal movements using a four-mass model agreed with the above assumption for two laryngeal movement patterns of throat singing. Furthermore, the authors also proposed a laryngeal voice model for throat singing and synthesized throat singing sounds.

key words Throat singing, Vocal fold vibration, False vocal fold, Inverse filtering, Laryngeal voice model, Self-oscillating model

1. はじめに

喉歌とは、アルタイ山脈周辺に分布する民族にみられる、一人の歌手が同時に2つの音高（ピッチ）が知覚される歌声を発する歌唱法の総称である[1,4,5,8]。代表的なものとして、トゥバのフーメイ（ホーメイ）やモンゴルのフーミー（ホーミー）などがある。喉歌の音色の不思議さは多くの研究者の興味を集め、音響的、生理学的手法による研究がいくつか報告されてきた[1-8,17]。そして現在、喉歌の複数音発音の主要因は舌などの特殊な調音により、鋭いホルマントを形成することによるものであることが明らかになっている[6]。しかしながら喉歌歌唱法とその音響的な特徴を説明するためには、調音方法だけでなく、喉頭における独特な音源発声方法についても解明する必要がある。喉頭音源に関する研究はまだ数少なく[7,17]、喉歌音の合成に向けた研究報告は無い。

喉歌の喉頭音源としての発声法は、喉詰発声とカルグラ発声とに大別される。喉詰発声は喉を詰めて発声し、いわゆるダミ声のような印象を与える喉歌の基本的発声であり、カルグラ発声は地声の声区より低いピッチの知覚される発声である。

筆者等は、この2つの異なる発声法について、ファイバースコープを用いた高速度デジタル撮影[9]とEGG[10]の同時記録を行い、歌唱時の喉頭調節を生理学的手法を用いて観測し、前稿で報告した[11-14]。その要旨は以下のようなものであった。両発声法において歌唱時に声門上部を構成する器官（本稿では簡単にそれらの器官を総称し、以後「仮声帯」と呼ぶことにする）の狭めが全周にわたって起り、そしてそれが振動することが観察された。喉詰め発声における仮声帯の振動周期は音声波形の周期と一致するが、仮声帯振動の位相は声帯振動とほぼ逆位相と推測された。カルグラ発声では仮声帯の振動周期は声帯の振動周期の2倍であり、それは音声波形の周期と一致していた。

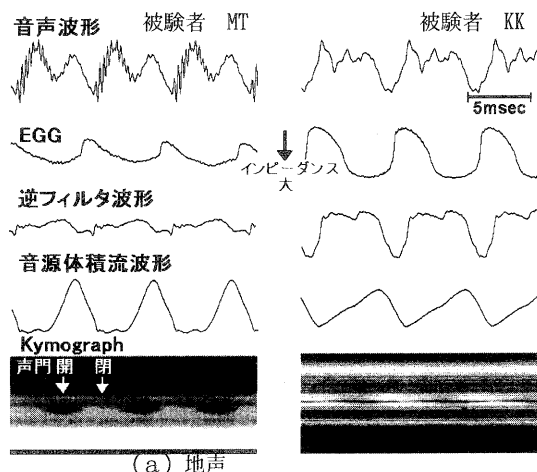
本稿では、この生理学的観測結果に加え、(1) 逆フィルタ法による音源推定、(2) 仮声帯振動も考慮した「声帯-仮声帯自励振動モデル」の作成とシミュレーションにより、喉歌歌唱時の喉頭音源についてより詳細な検討を行った。さらにその結果を基に、(3) 喉歌音合成のための喉頭音源モデルを考案し、喉歌の合成を試みた。

2. 逆フィルタ法による喉歌音源推定

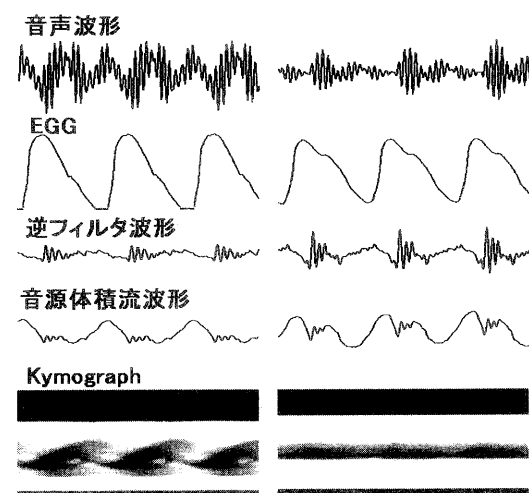
喉歌音も通常音声と同様に、その音声は音源・フィルタ理論に基づいて生成されるものと考え、逆フィルタ法による音源体積流波形の推定を行った[15]。

2.1 方法

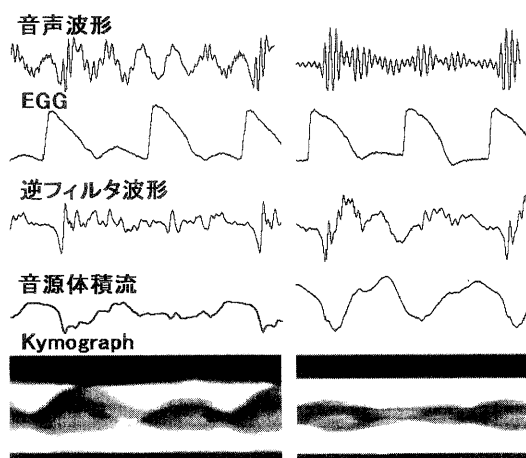
対象とする音声試料は高速度デジタル撮影と同期



(a) 地声



(b) 喉詰発声



(c) カルグラ発声

図1. 地声、喉詰発声、カルグラ発声の推定音源体積流波形（左MT、右KK）

して収録された音声であり、12 b/s, 18kHz sf で記録される。まず、原音声を LPC スペクトル分析し、第1ホルマントから第5ホルマントまでの周波数および帯域幅をそれぞれ求め、音源体積流推定のための探索初期値とする。次に、原音声中、カットオフ周波数 4300Hz のローパスフィルタを通し、続いて反共振周波数と帯域幅可変の従続接続反共振フィルタに通す。その出力を、放射特性を補正するため数値積分し、音源体積流波形とする。スペクトル分析によるホルマント推定は正確なものとは限らないので、各ホルマント周波数、帯域幅を僅かづつ変化させ、得られる音源体積流波形が、EGG 波形等によって示される声帯振動様式に矛盾しないように、目視により微調整を行う。

2.2 喉歌音の推定音源体積流波形

図1に逆フィルタ法により推定された喉歌音の音源体積流波形を示す。図に示した音声試料は前稿[11-14]と同じもので、日本人男性喉歌歌手2名、MT (35歳)、KK (31歳) による地声と喉歌である。図1(a)は地声、(b)は喉詰発声、(c)はカルグラ発声である。各図の最下段には高速度デジタル撮影画像から作成された Kymograph を示す[16]。Kymograph とは高速度デジタル撮影画像の冠状方向の断面として1本の走査線だけを選択し、その1走査線画像を時間軸に沿って並べたものである。画像の暗いところが声門等開口部を示す。本来ならば高速度デジタル撮影画像から声帯や仮声帯の開口面積を計測して表示すべきであるが、計測困難であったので参考データとして Kymograph を示した。音声収録時、喉頭からマイクロフォンまでの距離が約 25cm であったので、図中の音声、逆フィルタ波形、音源体積流波形は EGG 波形、Kymograph より約 1msec 位相が遅れている。

(b)の喉詰発声時においては、EGG 波形と Kymograph から、声帯と仮声帯のそれぞれの振動がほぼ逆位相になっていることがわかる。このため、仮声帯閉鎖時には、声帯の振動を見るのが不可能である。喉詰発声時の推定音源体積流は MT においては地声発声に近いが、KK では三角波の頂点右肩付近に細い切れ込みが見られる。EGG 波形と音声波形の励振時点を考慮すると、この細い切れ込み時点が声門閉鎖開始時点と考えられる。通常、声門閉鎖

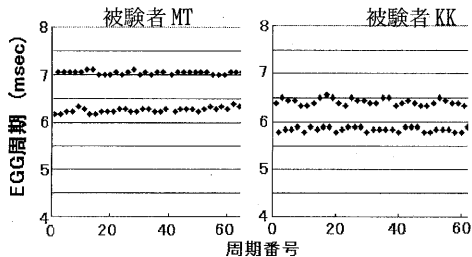


図2. カルグラ発声時のEGG周期系列(左MT, 右KK)

区間では声門体積流は零になるが、この例ではそうはなっていない。このことから KK においては声門の閉鎖が不完全であると推測される。また、声門体積流が最小になるのは仮声帯の開口面積最小時点と推測される。MT、KK ともに、声門閉鎖開始時点と思われる時点から 2400Hz 程度の周波数の減衰振動波形が認められる。

(c)のカルグラ発声においては、2名ともに推定音源体積流波形は1周期に2回、極大値をとる。声帯は、音源体積流波形の1周期に2回振動し、そのうち1回は声門閉鎖が不完全であると推測される。また、Kymograph から、仮声帯の開口面積が最小の時点付近に声門閉鎖が不完全になることがわかる。EGG の波形と周期が声帯振動の1周期おきに異なっていることが観察されたので、カルグラ発声時の EGG 波形周期系列を計測した。図2に、EGG 波形のゼロクロス(負から正)点の周期計測結果を示す。2例ともにその周期系列は長短の繰り返しパターンを持つことが確かめられた。

2.3 推定音源体積流のスペクトル

図1の各推定音源体積流波形のパワースペクトルを

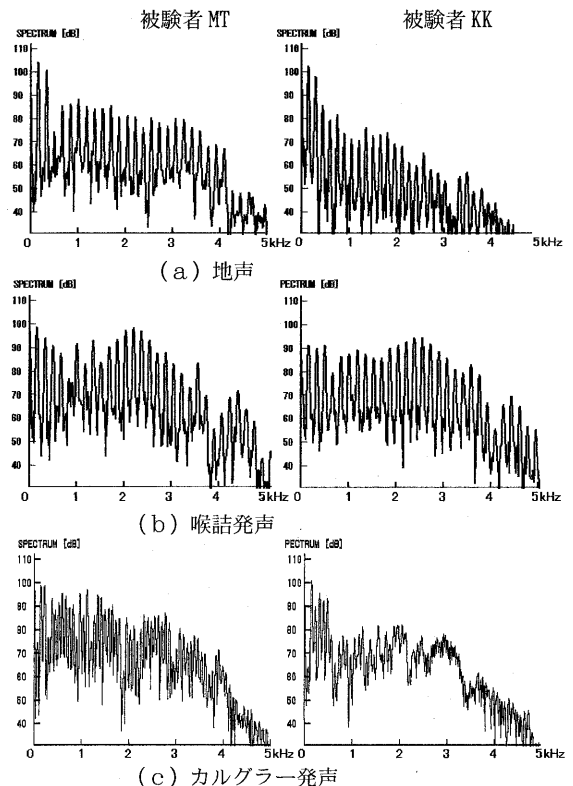


図3. 推定音源体積流波形のパワースペクトル(左MT, 右KK)

図3に示す。図3(a)は地声、(b)は喉詰発声、(c)はカルグラ発声である。2名の被験者とも、喉詰発声およびカルグラ発声における音源体積流波形は地声の場合に比べて、喉歌の上音に相当する第2ホルマントのピークの存在する帯域(1kHz—3kHz付近)のパワーが大きくなっている。

3. 声帯—仮声帯自励振動モデルによるシミュレーション

これまでの結果から、喉詰発声においては声帯と仮声帯は同じ周期で振動するがその位相はほぼ逆になっていること、またカルグラ発声では、仮声帯は声帯振動周期の2倍の周期で振動し、声帯は1周期おきに閉鎖が不完全になることが推測された。このような振動パターンでの発振の可能性を確かめるために、声帯—仮声帯自励振動モデルを作成し、シミュレーションを行った。

3.1 モデルの概要

声帯および仮声帯の振動する物理モデルを作成するために、まず声帯振動の2質量モデル[18,19]を、生理学的観測結果に基づき[12,17]、仮声帯も振動する4質量モデルに拡張し(図4)、計算機上に実装した。本モデルは2質量モデルを、気管から声道の間に2段階構えにしたものであり、声帯と仮声帯の間の機械的な振動の伝播は考慮していない。また、声帯と仮声帯の間にある喉頭室は断面積 2cm^2 、長さ 0.5cm の円筒とし、形状および体積は変化しないとされた。結合する声道は断面積が 5cm^2 で、長さは 16cm とした。本モデルに与える声帯の機械的パラメータは、2質量モデルの標準値[18,19]とし、仮声帯のパラメータは基本的に声帯と同じものとした。

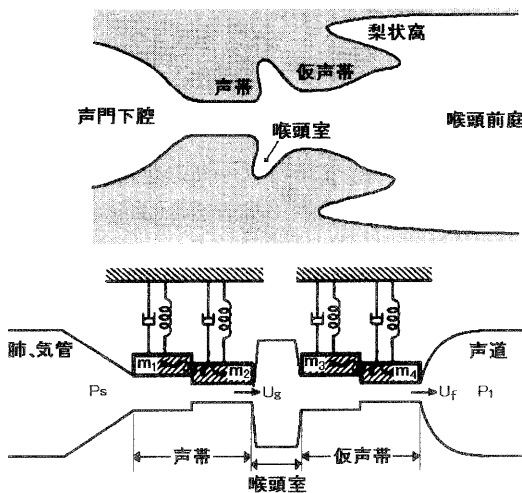


図4. 声帯—仮声帯自励振動モデル

3.2 喉詰発声のシミュレーション結果

図5に喉詰発声シミュレーション結果を示す。モデルのパラメータ値として与えた仮声帯の機械的パラメータは基本的には声帯と同じである(表1)。この結果、声帯と仮声帯は安定して振動した。そして、声帯と仮声帯がほぼ逆位相で振動することを確認することができた。本モデルの仮声帯休止開口面積パラメータの値を大きくした場合、声帯振動は仮声帯の影響を考慮しない2質量モデルとほぼ同じものとなる。図6に喉詰しない場合と喉詰した場合の両声門面積波形を示す。図6は、喉詰をした場合には、声門面積波形が対称三角波に近い形に変化している。Rothenberg[20]は声門体積流波形が声道特性の影響によって、声門閉鎖の速度が開大より速くなるため、閉鎖区間が相対的に短くなり、非対称三角波になると述べている。本モデルでは、喉詰をしたことにより声帯に与える声道特性の作用が減少したので対称三角波に近いものとも考えることもできる。

パラメータ	喉詰 発声	カルグラ 発声
呼気圧 [cmH ₂ O]	19	16
声帯 下唇休止開口面積 [cm ²]	0.005	0.04
声帯 上唇休止開口面積	0.005	0.04
仮声帯 下唇休止開口面積	0.03	0.05
仮声帯 上唇休止開口面積	0.03	0.05
声帯長 [cm]	1.4	1.4
仮声帯長	1.6	1.1
声帯 下唇質量 [g]	0.2	0.2
声帯 上唇質量	0.015	0.015
仮声帯 下唇質量	0.2	0.375
仮声帯 上唇質量	0.015	0.0375
声帯 下唇厚さ [cm]	0.2	0.2
声帯 上唇厚さ	0.04	0.04
仮声帯 下唇厚さ	0.2	0.316
仮声帯 上唇厚さ	0.04	0.0633
声帯 下唇スティフネス [dyne/cm]	120000	120000
声帯 上唇スティフネス	12000	12000
声帯間スティフネス	28000	28000
仮声帯 下唇スティフネス	120000	48000
仮声帯 上唇スティフネス	12000	4800
仮声帯間スティフネス	28000	15000
声帯 下唇損失係数	0.1	0.1
声帯 上唇損失係数	0.6	0.6
仮声帯 下唇損失係数	0.6	0.6
仮声帯 上唇損失係数	0.6	0.6

表1. 4質量モデルのシミュレーションにおける主要パラメータの値

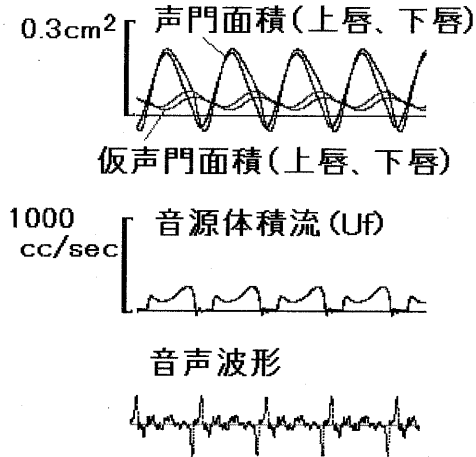


図5. 喉詰発声シミュレーション結果
(仮声門面積は仮声帯の開口面積を指す)

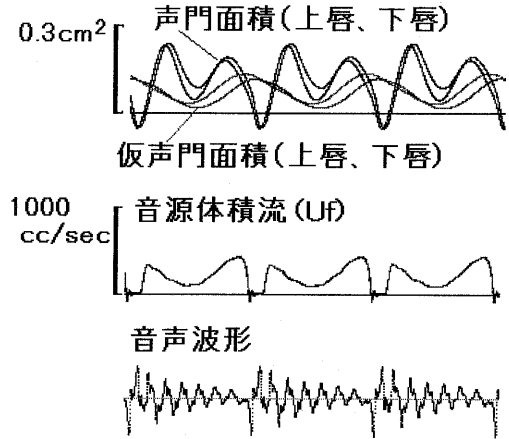


図7. カルグラール発声シミュレーション結果

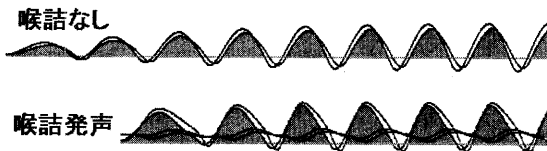


図6. 喉詰発声シミュレーションによる声門面積波形

3.3 カルグラール発声のシミュレーション結果

図7に、カルグラール発声シミュレーション結果例を示す。2質量モデルでは声帯の振動周波数を制御するために Q パラメータが導入された[18,19]。この Q パラメータの値を変化させると、声帯の振動周波数が Q パラメータに比例して変化する。カルグラール発声を出現させるために、本モデルの仮声帯の Q パラメータ(Q_f)と声帯の Q パラメータ(Q_g)の関係を、 $Q_f=Q_g/2.5$ に設定した。実際には、仮声帯のスティフネスを声帯のスティフネスの $1/2.5$ に、仮声帯の質量を声帯の質量の2.5倍に、仮声帯の厚さを声帯の厚さの $\sqrt{2.5}$ 倍に設定した(表1)。

図7の結果を見ると、仮声帯の振動周期は声帯の振動周期の2倍である。また、仮声門面積(仮声帯の開口面積)が最小になる時点近傍では声門は閉鎖しない。このため音源体積流波形は1周期に2回、極大値をとる。この結果は図1の観測結果と合致する。

3.4 音源体積流のスペクトル

図8に、本自励振動モデルで得られた喉詰発声とカルグラール発声の音源体積流波形のスペクトルを示す。

(a)は2質量モデル(仮声帯の影響無し)(b)は喉詰発声(c)はカルグラール発声である。モデルで得られた両発声法のスペクトルにおいて、2質量モデルのスペクトルに比べて約3kHzまでの区域のパワーが大きくなっている。このように本モデルは周波数スペクトルにおいても観測データと一致する。本モデルにおける、このパワーの増大は、声帯と仮声帯の間にある喉頭室の音響的コンプライアンスによるものと考えられる。

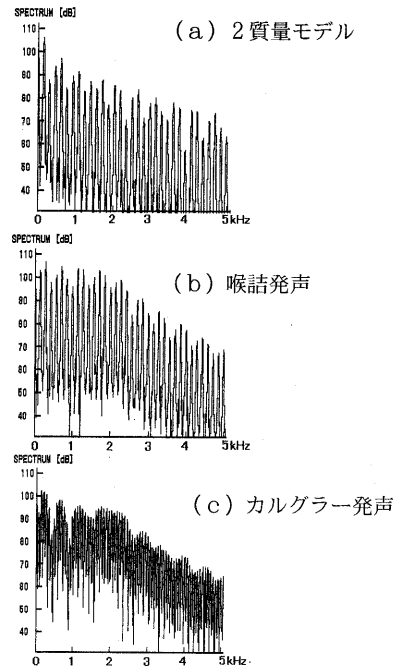


図8. 声帯-仮声帯自励振動モデルで計算された音源体積流のパワースペクトル

4. 喉歌の喉頭音源モデル

これまでに得られた結果を基に、喉歌合成のための喉詰発声、カルグラ発声の両方の喉頭音源を表現する信号モデルを提案する(図9)。

4.1 モデルの概要

本モデルは、声門体積流を、振動する仮声帯の開口面積(仮声門面積)により変調したものに、仮声帯の狭めに応じた乱流雑音を付加する。さらに、喉頭室の共鳴特性を考慮し、その音響的コンプライアンスによる単共振フィルタを通す。

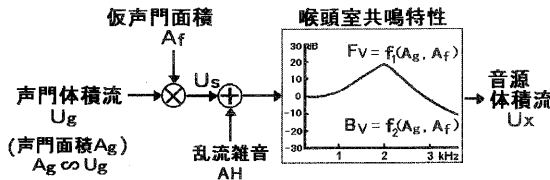


図9. 喉歌の喉頭音源モデル

4.2 音源波形生成計算式

[1] 声門体積流(U_g)は、図6の自動振動モデルの結果を参考にして、正の正弦波を基本とする(図10)。カルグラ発声を実現するために、 KQ (基本周期変動率)というパラメータを導入し、基本周期 T_0 の2倍周期単位で波形を計算する。

$$U_g(t) = (AV \cdot T_0) \{ \sin(t \cdot \pi / (OQ \cdot T_0)) \} \quad (0 \leq t < T_1)$$

$$U_g(t) = (AV \cdot T_0) \{ \sin(t \cdot \pi / (OQ \cdot T_0)) + \sin((t - T_1) \cdot \pi / (OQ \cdot T_0)) \} \quad (T_1 \leq t < OQ \cdot T_0)$$

$$U_g(t) = (AV \cdot T_0) \{ \sin((t - T_1) \cdot \pi / (OQ \cdot T_0)) \} \quad (OQ \cdot T_0 \leq t < T_1 + OQ \cdot T_0)$$

$$U_g(t) = 0 \quad (T_1 + OQ \cdot T_0 \leq t \leq 2 \cdot T_0)$$

ここで、 $T_1 = KQ \cdot T_0$ 、 AV は振幅、 T_0 は基本周期、 OQ は声門開放時間率パラメータである。 $T_1 < OQ \cdot T_0$ の場合は、1周期おきに声門閉鎖が不完全となる。

[2] 声門面積(A_g)は単純に U_g と同じ波形(但し最大面積を 0.2 cm^2 に正規化)とする。

$$A_g(t) = 0.2 \cdot U_g(t) / (AV \cdot T_0) \quad [\text{cm}^2]$$

[3] 仮声門面積($A_f [\text{cm}^2]$)は以下の様に計算する。

$$A_f(t) = (0.05 \cdot FQ) \{ -\cos(t \cdot \pi / (tp_i - tn_i)) + 0.05 \} \quad (tn_i \leq t < tp_i)$$

$$A_f(t) = (0.05 \cdot FQ) \{ \cos(t \cdot \pi / (tn_{i+1} - tp_i)) \} + 0.05 \quad (tp_i \leq t < tn_{i+1})$$

ここで、 FQ は仮声帯振動振幅率パラメータであり、 $0 \leq FQ < 1$ とする。仮声帯の振動周期は、カルグラ発声の場合は声帯振動の2倍であり、喉詰発声の場合は声帯振動周期と同一であった。また、観測された声帯振動と仮声帯振動の位相関係を考慮して tn_i と tp_i を以下の様に定めることにした。

$KQ < 0.98$ の場合はカルグラ発声と定め、

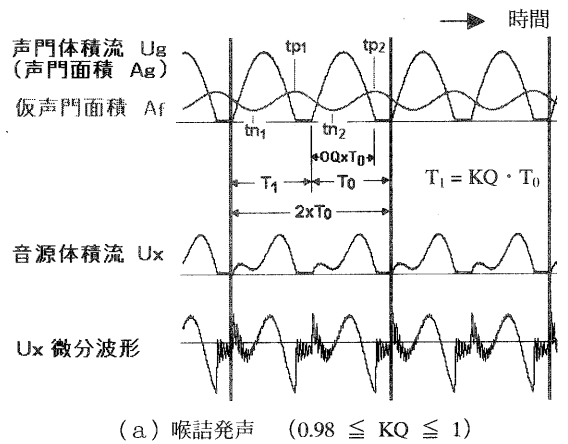
$$tp_1 = T_1 + OQ \cdot T_0 \quad tn_1 = (T_1 + OQ \cdot T_0) / 2$$

とした。 $0.98 \leq KQ \leq 1$ の場合は喉詰発声と定め、

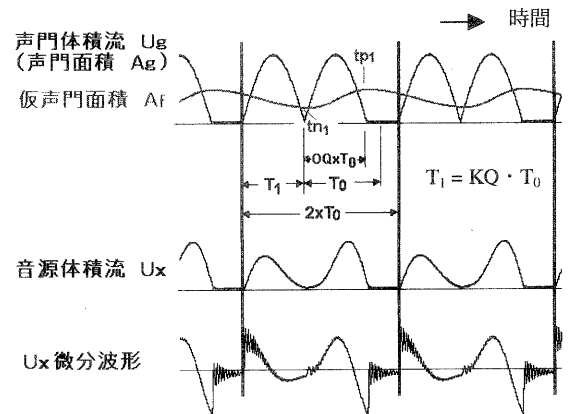
$$tp_1 = OQ \cdot T_0 \quad tn_1 = (OQ \cdot T_0) / 3$$

$$tp_2 = T_1 + OQ \cdot T_0 \quad tn_2 = T_1 + (OQ \cdot T_0) / 3$$

とした。



(a) 喉詰発声 ($0.98 \leq KQ \leq 1$)



(b) カルグラ発声 ($KQ < 0.98$)

図10. 喉歌喉頭音源モデルの各波形

[4] U_s は、 U_g を A_f で変調したものである。ここでは、 A_f を A_g のピーク値である 0.2 cm^2 で正規化したものを用いる。

$$U_s(t) = U_g(t) \cdot \{A_f(t)/0.2\}$$

[5] 喉頭乱流雑音 (AH) はレイノルズ数の2乗に比例する。ここでは、以下の様に計算し、 U_s に加算する。

$$AH(t) = AHA \cdot k \cdot \{U_s(t)/A_f(t)\}^2 \cdot [\text{一様乱数系列}]$$

$$k = 0.2^4 / (AV \times T_0)^2$$

ここで、AHA は乱流雑音振幅パラメータである。

[6] 喉頭室の音響的コンプライアンスによる単共振共鳴特性の共振周波数 F_v と帯域幅 B_v は、喉頭室断面積 2.5 cm^2 、喉頭室幅 0.5 cm と仮定して以下の式で計算する。

$$F_v(t) = 6434 \cdot \sqrt{A_g(t) + A_f(t)}$$

$$B_v(t) = 10.76 \cdot \{A_g(t) + A_f(t)\} / 2 + 175.1$$

4.3 音源波形例

図11に本モデルによる音源波形例と音源体積流微分波のスペクトルを示す。喉詰発声 (図左) の例の音源パラメータは $KQ=100\%$, $FQ=95\%$, $F_0=190\text{Hz}$, $OQ=90\%$ である。カルグラ一発声 (図右) の例の音源パラメータは $KQ=85\%$, $FQ=95\%$, $F_0=132\text{Hz}$, $OQ=99\%$ である。従来の喉頭音源モデルは、周波数が高くなるにつれパワーが単調かつ急速に減少するが、本モデルでは、喉詰発声、カルグラ一発声の両方において、喉歌の上音に相当する第2ホルマントのピークの存在する帯域 (1kHz-3kHz 付近) のパワーが大きくなっている。

4.4 喉歌の合成

図12に喉歌合成プログラムを示す。これは Klatt 合成器[21]の音源部を、本モデルによる喉頭音源に変更したも

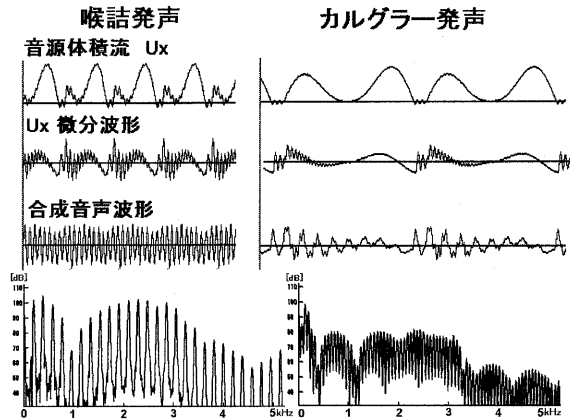


図 11. モデルによる喉歌喉頭音源波形とそのパワースペクトル

のである。実際の喉歌音声から推定されたホルマント (声道共鳴特性) 時系列により得られる合成音からは、従来の喉頭音源モデルを用いた場合に比べ、実際の喉歌の音色に近い聴覚的印象が得られた。

5. 考察

喉歌の上音の生成には、舌などを用いた調音による鋭いホルマントの形成が主な要因となっていることが知られている[6,8]。また、鋭いホルマントの形成には、舌による調音だけではなく、喉歌の特徴的な喉頭音源の関与も考慮されるべきであり、このことはしばしば指摘されてきた[12]。本研究での逆フィルタ法により推定された喉歌の喉頭音源の周波数特性は、喉歌の上音に相当する第2ホルマントのピークの存在する帯域 (1kHz-3kHz 付近) のパワーが地声発声よりも大きく、これは喉歌の上音の生成への喉頭音源の寄与を示していると言える。

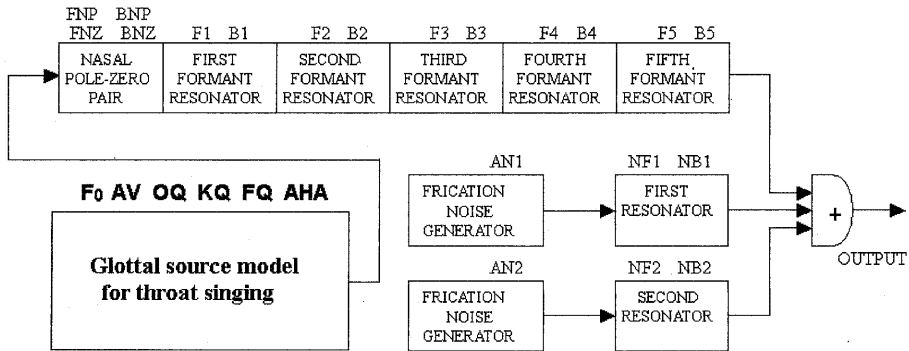


図 12. 喉歌合成プログラム・ブロック図

地声発声において、声帯の振動周期における、声帯の開大区間で、ホルマントエネルギーが急激に減少することが知られているが[22]、喉詰発声においては仮声帯と声帯がほぼ逆位相で振動しており、地声発声において0Q(声門開放時間率)が非常に小さい場合と同様の状況を作り、このことがホルマントエネルギーの減少を妨げ、上音生成に必要な帯域のパワーを支えていると言える。

一方、声帯-仮声帯自励振動モデルによるシミュレーションにおいても、喉歌の喉頭音源は地声発声よりも1kHz-3kHz付近のパワーが大きかった。これは、仮声帯の狭めと、振動により生じる喉頭室の共鳴特性とが、主な要因であると考えられる。

以上のことから、喉歌の歌唱においては、地声の発声では通常用いられない器官である仮声帯を用いて、狭めや、振動により、喉頭音源の1kHz-3kHzの帯域のパワーを増大させ、舌や口唇などによる調音と組み合わせて上音の生成を行なっていると言える。

声帯-仮声帯自励振動モデルにより、地声、喉詰に加え、カルグラの倍周期振動が模擬された。しかしながら、仮声帯に関しては解剖学的、生理学的に不明なことも多く、今後、生理計測によるモデルおよびモデルパラメータの妥当性の考察を通じ、より精密な喉頭の物理モデルを構築し、仮声帯および声帯の振動様式について検討していく必要がある。

6. むすび

喉歌歌唱時の喉頭の生理学的観測結果、および逆フィルタ法により推定された喉頭の体積流波形の考察に基づき、喉頭の物理モデルである4質量モデルを考案し、声帯-仮声帯の自励振動のシミュレーションを行なった。更に、シミュレーションの結果もふまえ、喉歌の喉頭音源モデルを信号モデルとして提案し、喉歌を合成した。

今後は、声帯-仮声帯自励振動モデルの精密化、喉頭音源の逆推定の精緻化、喉歌合成に基本周波数の揺らぎなど自然性を向上させるパラメータの追加、聴取実験によるモデルの有効性の評価、他の歌唱、音声へのモデルの適用、などを行なっていきたい。

謝辞 本研究に関してご討論頂いた、足立整治(愛知県立大学/ATR)、今泉敏(東大・医)、小坂直敏、白木善尚(NTT CS 基礎研)の諸氏に感謝します。

参考文献

[1] 等々力 政彦、榊原 健一、小坂 直敏、足立 整治、トゥバの喉歌フォーメイ(ホーメイ)の歌唱について、*音楽音響研究会資料*、Vol.MA99-18, pp.119-126, 1999.
 [2] A. N. Aksenov, Tuvin folk music, *Journal of the Society for Asian Music*, Vol 4, No.2, pp.7-18, 1973. Extract from

Tuvinskaia narodnaya muzyka(1964, Moscow: Musyka).
 [3] G. Bloothoof, E. Bringman, M. van Cappellen, J. B. van Luipen and K. P. Thomassen. Acoustics and perception of over tone singing. *J. Acoust. Soc. Am.*, Vol.92 pp.1827-1836, 1992.
 [4] 村岡 輝雄、武田 昌一、糸賀 昌士. モンゴル歌唱法「ホーミー」の音響的特徴の解析、*日本音響学会誌*、Vol.56, No.5, pp.308-317, 2000.
 [5] 近藤 和正. フーミーの調音音声学的分析試論. *日本モンゴル学会紀要*、Vol.28, pp.99-106, 1997.
 [6] S. Adachi and M. Yamada. An acoustical study of sound production in biphonic singing Xüömij. *J. Acoust. Soc. Am.*, Vol.105, No.5, pp.2920-2932, 1999.
 [7] T. C. Levin and M. E. Edgerton. The throat singers of Tuva. *Scientific America*, pp.80-87, Sep. 1999.
 [8] 山田 真司. 一人で歌う二重唱. *日本音響学会誌*、Vol.54, No.9, pp.680-685, 1998.
 [9] S. Kiritani. High-speed digital image recording for observing vocal fold motion. In R. D. Kent and M. J. Ball, editors, *Voice Quality Measurement*. Singular Publishing Group, 1999.
 [10] 石毛 美代子、新美 成二、森 浩一. Electroglottography (EGG). *音声言語医学*、Vol.37, pp.347-354, 1996.
 [11] 榊原 健一、足立 整治、近藤 和正、小西 知子、村野 恵美、熊田 政信、等々力 政彦、今川 博、新美 成二、トゥバ、モンゴルの喉歌における声帯振動の観測、*日本音響学会講演論文集I*、pp.171-172, 2000-9
 [12] 榊原 健一、足立 整治、小西 知子、近藤 和正、村野 恵美、熊田 政信、等々力 政彦、今川 博、新美 成二. 喉歌の発声における声帯振動の分析、*音楽音響研究会資料*、Vol.MA00-21, pp.33-40, 2000.
 [13] 近藤 和正、小西 知子、榊原 健一、村野 恵美、熊田 政信、等々力 政彦、今川 博、新美 成二. 喉歌の発声における喉頭調節. *日本音声学会全国大会予稿集*、pp.37-42, 2000-9.
 [14] 小西 知子、近藤 和正、榊原 健一、熊田 政信、村野 恵美、今川 博、新美 成二. 喉歌歌唱時における喉頭調節. *日本音声言語医学会総会・学術講演会予稿集*、pp.72, 2000-11.
 [15] B. Fritzell, Inverse Filtering, *J. Voice*, Vol. 6 No. 2, pp.111-114, 1992
 [16] J. G. Svec and H. K. Schutte. Videokymography: High-speed line scanning of vocal fold vibration. *Journal of Voice*, Vol.10, No.2, pp.201-205, 1996.
 [17] B. Chernov and V. Maslov. Larynx - Double-sound generator. *Proceedings XIth ICPhs*, (Tallin Estonia, U.S.S.R.), Sc 98.4.1, pp.40-43, 1987.
 [18] K. Ishizaka and J. L. Flanagan. Synthesis of voiced sounds from a two-mass model of the vocal cords. *The Bell System Technical Journal*, Vol.51, No.6, pp.1233-1268, 1972.
 [19] 石坂 謙三、J. L. Flanagan. 声帯音源の自励振動モデル. *日本音響学会誌*、Vol.34, No.3, pp.122-131, 1978.
 [20] M. Rothenberg and S. Zahorian. Nonlinear inverse filtering technique for estimating the glottal-area waveform. *J. Acoust. Soc. Am.*, Vol.61, No.4, pp.1063-1071, 1977.
 [21] D. H. Klatt. Software for a cascade/parallel formant synthesizer. *J. Acoust. Soc. Am.*, Vol.67, No.3, pp.971-995, 1980.
 [22] 大村 浩、田中 和世. 声帯振動の非線形効果を考慮した音声合成法とその評価. *日本音響学会誌*、Vol.53, No.6, pp.427-434, 1997.