

タイトル	音響放射パワーに基づく3次元声道モデルの伝達特性 評価法
著者	元木，邦俊；松崎，博季
引用	北海学園大学工学部研究報告，35：131-141
発行日	2008-02-00

音響放射パワーに基づく 3 次元声道モデルの伝達特性評価法

元 木 邦 俊・松 崎 博 季

An Evaluation Method of the Transfer Characteristics of 3-dimensional Vocal-Tract Models Based on the Acoustic Radiation Power

Kunitoshi MOTOKI and Hiroki MATSUZAKI

要 旨

3次元声道モデルでは、非平面波的な音波伝搬や鼻孔放射の干渉により口唇近傍の音場が複雑なものとなり、声道の共鳴特性を評価する出力点を明確に定めることが容易ではない。本稿では、放射される音響パワーに着目して、特定の放射位置を仮定することなく声道の伝達特性を評価する方法について述べる。声道伝達特性の評価で重要な極や零点の周波数について、従来の1次元モデルで利用される体積速度伝達関数とほぼ同等に評価することができることを示す。

1 はじめに

音声は、口唇、鼻孔、皮膚から放射され、比較的遠方で観測される時系列信号なので、声帯音源から観測音声への伝達関数を求めることが音声生成過程の特徴抽出では重要な課題である。筆者らは、MRIなどにより得られた声道の3次元形状データに基づいて幾何学的な3次元声道形状モデルを構成し、その音響解析を行っている¹⁻⁴⁾。従来から利用されている1次元音響管モデルやその電氣的等価回路表現では、信号の入力と出力を定義する位置が明確なため、声門部を声道入力端、口唇部を出力端として、その間の伝達関数（主に体積速度伝達関数）を計算することにより声道モデルの共振特性を評価している。一方、3次元声道モデルでは、非平面波的な音波伝搬や鼻孔放射の干渉により口唇近傍の音場が複雑なものとなり、声道の共鳴特性を評価する出力点を口唇近傍に定めることが容易ではない。数値解析の対象とする空間を比較的遠方の観測位置まで考えると、その点での音圧を出力と考えることができるが、実際の数値計算では膨大な要素分割が必要となるため現実的ではない。このため、声道の出力とし

* 北海学園大学工学部電子情報工学科

* Department of Electronics and Information Engineering, Hokkai-Gakuen University

て、口唇付近の特定の位置における音圧や粒子速度、あるいは空間的に広がる口唇と鼻孔近傍の放射領域の音圧や粒子速度の平均を出力信号と考えることが多い。しかしながら、このような信号を出力信号として考えた場合には平面波以外の高次モードによる音響放射による寄与を評価することができない恐れがある。筆者らは、モード展開に基づいてパラメトリックに3次元音場を表現する声道モデルを提案している⁷⁻¹⁰⁾。このモデルでは声道から外部へ放射される音響パワーを平面波と高次モードに分けて求めることができる。外部へ放射されるパワーの総量を容易に求めることができる一方、口唇端に相当する点においては、出力として考えることができるモード音圧（あるいはモード粒子速度）が複数存在し、音源体積速度と1対多の対応関係となってしまう。このため、モデルの伝達特性は位相特性の含まれない放射パワーに基づいた評価を行っている。複雑な3次元形状を対象に音場解析を行う有限要素法においても、音響インテンシティの空間分布を求めることができるので、放射パワーの総量を求めることができる。本稿では、音響放射パワーを用いて、音源体積速度と遠方点音圧との関係を伝達インピーダンスの形で近似し、これにより、声道の伝達特性として最も重要な極や零点の評価を行うことが可能であることを示す。放射される音響パワーに着目することで、幾何学的に複雑な3次元モデルであっても特定の信号出力場所を仮定することなく声道の伝達特性を求めることができることを示す。

2 出力信号と伝達特性

2.1 1次元モデルの伝達関数

1次元音響管モデルの伝達関数 H_1 は、図1のように声門体積速度源 U_G と口唇端での体積速度 U_L を用いて $H_1 = U_L/U_G$ として定義される。声道部を表現する縦続行列の要素を A_1 , B_1 , C_1 , D_1 , 声門部と口唇部での音圧をそれぞれ P_G , P_L , 放射インピーダンスを Z_L とすると、次の関係が成立する。

$$\begin{bmatrix} P_G \\ U_G \end{bmatrix} = \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix} \begin{bmatrix} P_L \\ U_L \end{bmatrix} \quad (1)$$

$$P_L = U_L Z_L \quad (2)$$

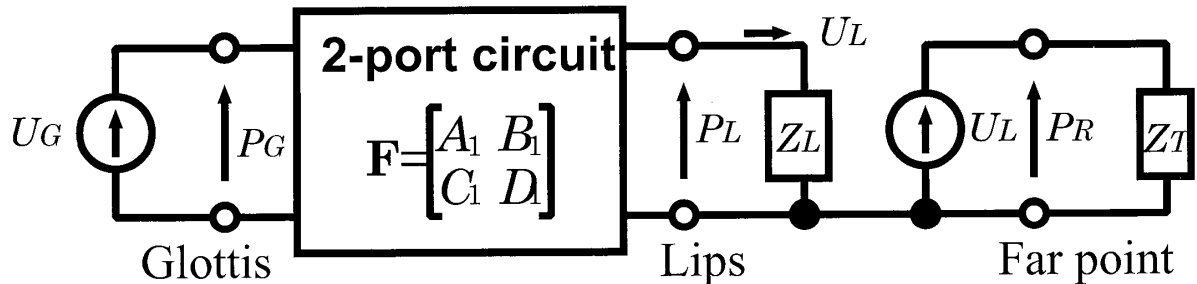


図1：1次元音声生成モデルの電氣的等価回路表現。

これより, H_1 は,

$$H_1 = \frac{1}{C_1 Z_L + D_1} \quad (3)$$

として定まる. また, 遠方点での音圧 P_R は U_L を用いて

$$P_R = U_L Z_T = U_G H_1 Z_T \quad (4)$$

として考えることができ, ピストン音源による放射過程からの類推により Z_T として1階微分特性(6 dB/oct.)を有するコイルを仮定することが多い.

2.2 3次元モデルの放射音場

図2 (a)は, MRIデータに基づいて作成された鼻腔が結合した有限要素モデルの例であ

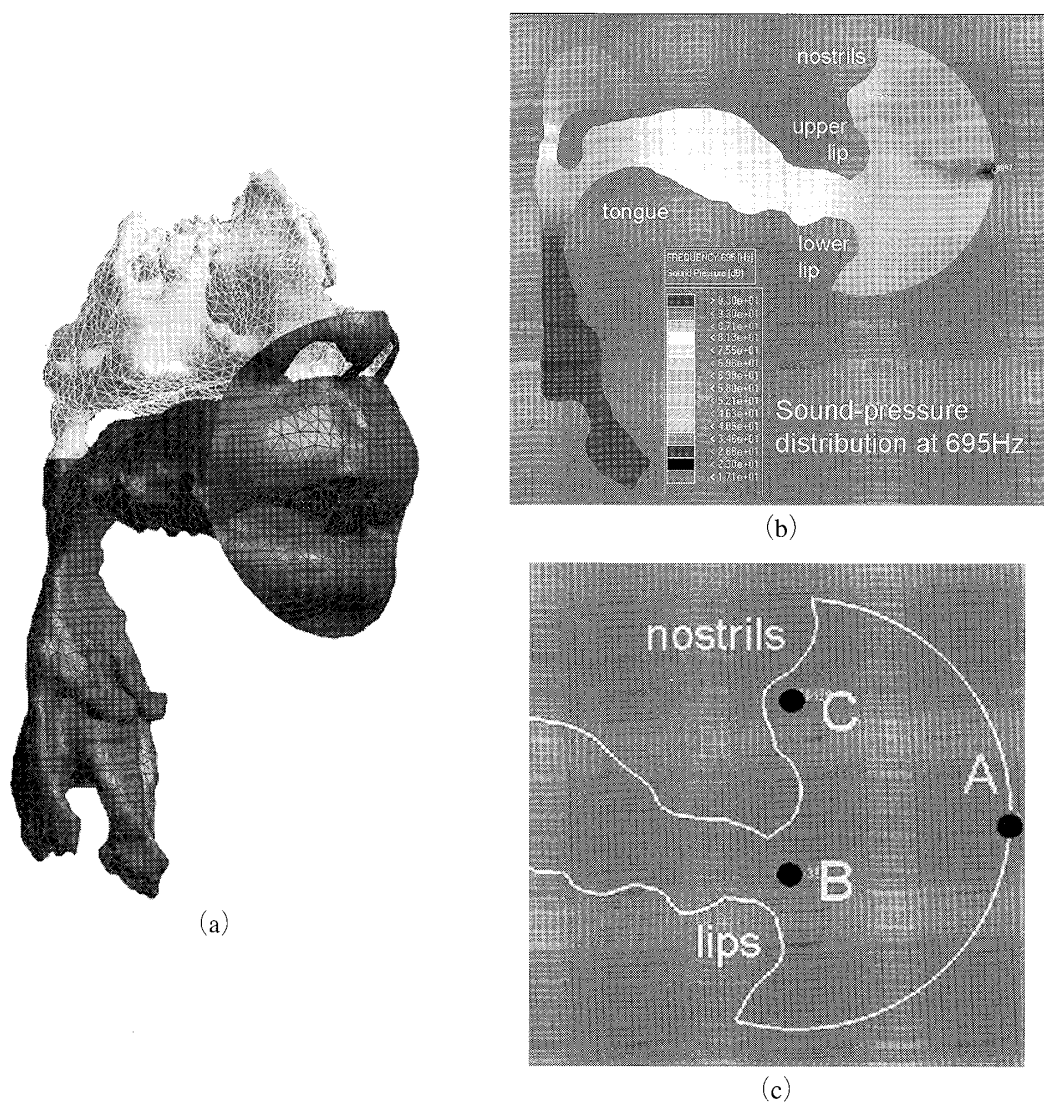


図2 : (a) 鼻腔を付加した3次元声道モデル (b) 正中矢状面上での音圧分布 (695Hz) (c) 点A, B, Cは図7の周波数特性の計算位置を示す.

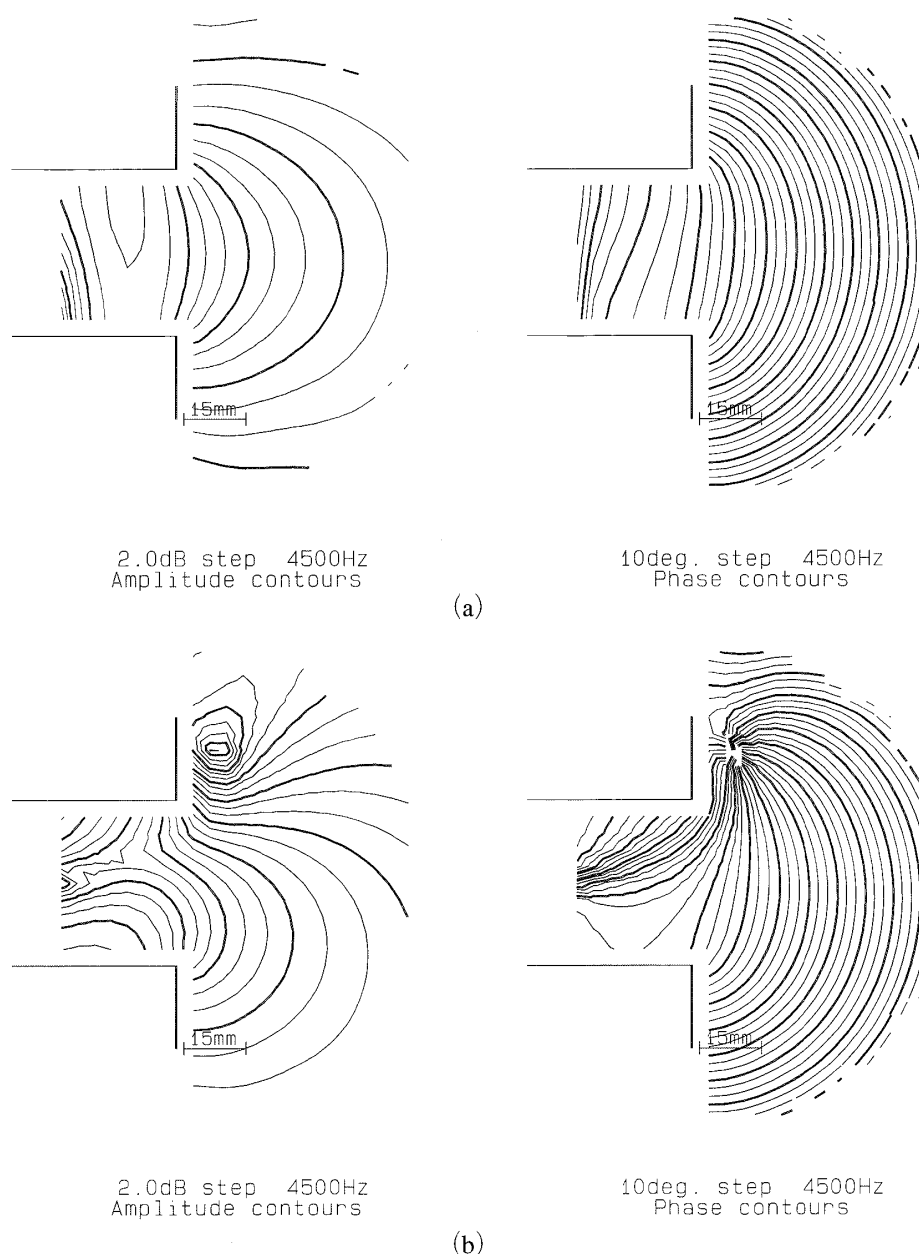


図3：バッフルにセットされた矩形音響管からの放射音場（4.5kHz）．左図は振幅分布を右図は位相分布を示す．(a) 平面波のみが放射される場合 (b) 平面波と高次モードが放射される場合．

る²⁾．3次元声道モデルは、放射域が連続的に広がるためにモデルの物理的形狀から終端位置を明確に定めることは困難である．また、鼻腔結合を考慮する場合には、放射空間では口唇と鼻孔からの放射音が重畳する．このため、特定の位置、周波数において音圧の極めて低い領域が放射空間に発生することがある（図2（b））．さらに、平面波伝搬の仮定が成立しない高い周波数域では、高次モードの放射が生じることもあり得る．図3は、平面バッフル面上の矩形音響管からの放射音場の測定例である⁵⁾．周波数が同じであっても駆動音源位置を変えることで矩形音響管内部の高次モードの励振状態が変化する．高次モードが伝搬モードとなると、平

面波成分と高次モード成分の両者が放射され、開口形状が同一であっても放射域に生じる音場は複雑なものとなる。開口部付近での音圧、または粒子速度の空間的な平均値には平面波以外の成分が反映されないため、高域での伝達特性を評価する場合には適当ではない。

2.3 放射パワー

図4のように3次元声道モデル（放射空間を含む）の位置 $\mathbf{r}$ における音圧 $p(\mathbf{r})$ ，粒子速度 $\mathbf{v}(\mathbf{r})$ とする。声道内部の断面 S ，開口部（口唇と鼻孔）を覆う曲面 C とするとき， S と C を通過する音響パワー W_S ， W_C は，アクティブ音響インテンシティ $\mathbf{I}(\mathbf{r}) = \text{Re}\{p(\mathbf{r})\mathbf{v}^*(\mathbf{r})/2\}$ （ $\text{Re}\{\cdot\}$ は実部を， $*$ は複素共役を表す）を用いて，

$$W_S = \iint_S \mathbf{I}(\mathbf{r}) \cdot d\mathbf{s} \geq \iint_C \mathbf{I}(\mathbf{r}) \cdot d\mathbf{s} = W_C \quad (5)$$

となる。ここで等号は声道内部を無損失とした場合に成立する。 W_C は，放射される全音響パワーを表す。 W_S は，音源が声門にのみ存在する場合には，声道内損失に応じて音源側で評価するほど大きくなるが，無損失の場合には声道内部のどの面で評価しても同じであり，音源が供給するパワーと放射されるパワーが等しいことを表す。

2.4 伝達インピーダンスによる評価

遠方点 R では，遠方点の音圧 P_R を用いて， $P_R = p(R) \approx \rho c v(R)$ （ ρc は空気特性インピーダンス， v は $\mathbf{v}$ の成分）となるので，自由空間における無損失を仮定すると，

$$W_C = \iint_C \mathbf{I}(\mathbf{r}) \cdot d\mathbf{s} \approx \frac{1}{2\rho c} \iint_C |P_R|^2 d\mathbf{s} \quad (6)$$

となる。遠方点での速度ポテンシャルを ϕ_R ，口唇正面前方での値を $\phi_{R,0}$ とすると，放射の指向特性 D は

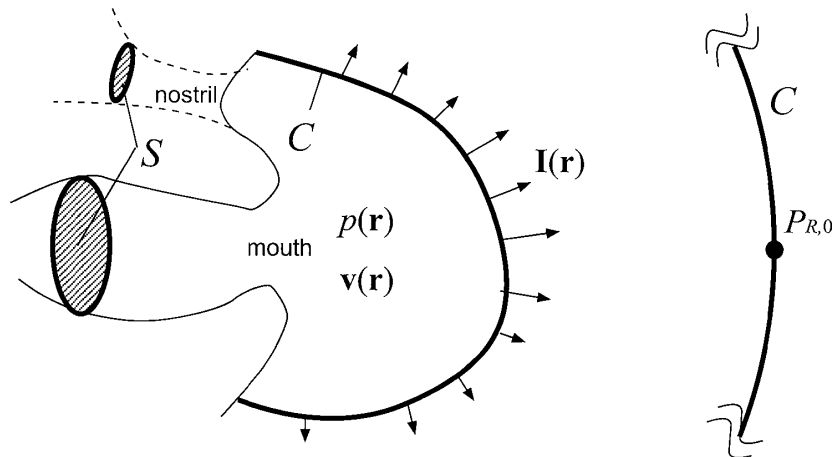


図4：声道内部の任意の断面 S と開口部を覆う自由空間側の任意の閉じた面 C 。

$$D = \left| \frac{\phi_R}{\phi_{R,0}} \right| \quad (7)$$

と書ける．口唇正面前方での音圧を $P_{R,0}$ とすると， $|P_R| = |P_{R,0}D|$ なので，

$$|P_{R,0}| = \sqrt{\frac{2\rho c}{\iint_C D^2 ds}} W_C \quad (8)$$

となる．バツフル面上の振動板の指向性は周波数の増加とともに鋭くなることから $\iint_C D^2 ds$ は周波数の増加に伴い減少すると考えられる．この積分項の特性に急峻な変化は現れないと仮定すると， K を周波数に関して緩やかに変化する関数として次のように書ける．

$$|P_{R,0}| = K\sqrt{W_C} \quad (9)$$

ここで，

$$K = \sqrt{\frac{2\rho c}{\iint_C D^2 ds}} \quad (10)$$

とおいた．一般に，振動板の指向性は周波数の増加とともに鋭くなることから， K は高域強調特性となると思われる． W_C は放射部近傍のインテンシティ分布から，あるいは，無損失の場合には音源が供給するパワーとして求めることができるから，音源体積速度 U_G から $P_{R,0}$ への伝達インピーダンスを Z_P とすると，その大きさは次のようになる．

$$|Z_P| = \left| \frac{P_{R,0}}{U_G} \right| = K \frac{\sqrt{W_C}}{|U_G|} \quad (11)$$

$|Z_P|$ により，放射端を特定することなく3次元モデルの伝達特性を評価することが可能である．なお，図1の1次元モデルでは，

$$W_C = \frac{\text{Re}\{P_L U_L^*\}}{2} = \frac{|U_L|^2 \text{Re}\{Z_L\}}{2} \quad (12)$$

となるので，1次元モデルの伝達関数 $|H_1|$ と $|Z_P|$ は，放射インピーダンスの実部を R_L として，次の関係となる．

$$|Z_P| = \sqrt{\frac{R_L}{2}} K |H_1| \quad (13)$$

平面バツフル面の任意の開口面 S に対する Z_L は，開口面上の微小面積要素 dS の振動による音圧をRayleigh積分により求め，その全体をさらに開口面上で積分することにより次式のように表現できる．

$$Z_L = \frac{jk\rho c}{2\pi S^2} \iint_{S'} \iint_S \frac{e^{-jkr}}{r} dS' dS \quad (14)$$

面積分は開口面($S = S'$)上で行い， r は dS と dS' の間の距離を表す．平面バツフル面上での半径 a の円形開口を仮定した場合には， Z_L は円形ピストン振動板の放射インピーダンスとして

よく知られており，その実部は，

$$R_L = \frac{\rho c}{\pi a^2} \left(1 - \frac{J_1(2ka)}{ka} \right) \quad (15)$$

となる．ここで， k は波数， $J_1(\cdot)$ は第一種Bessel関数である．円形開口に対する R_L の周波数特性は，低域では周波数の2乗にほぼ比例して増加し，高域では一定値となる．図5に開口面積をパラメータとした R_L の周波数特性の例を示す．矩形開口やバツフル面が球形となった場合にも R_L の周波数特性は同様な傾向を示す^{11,12)}ことから，伝達特性の評価として重要なピークと零点の位置は， $|H_1|$ と $|Z_P|$ でほぼ同じになる．式(14)は，積分操作により音圧分布の平均を求めているため，平面波に対するインピーダンスを意味する．高次モードの伝搬，放射を含むような高い周波数域における放射パワーの計算と計測法については，文献⁶⁾にて報告している．

3 伝達特性の評価

3.1 3次元有限要素法による結果

図6は，図2(a)の3次元声道モデルの有限要素解析（声道内部を無損失として計算）に

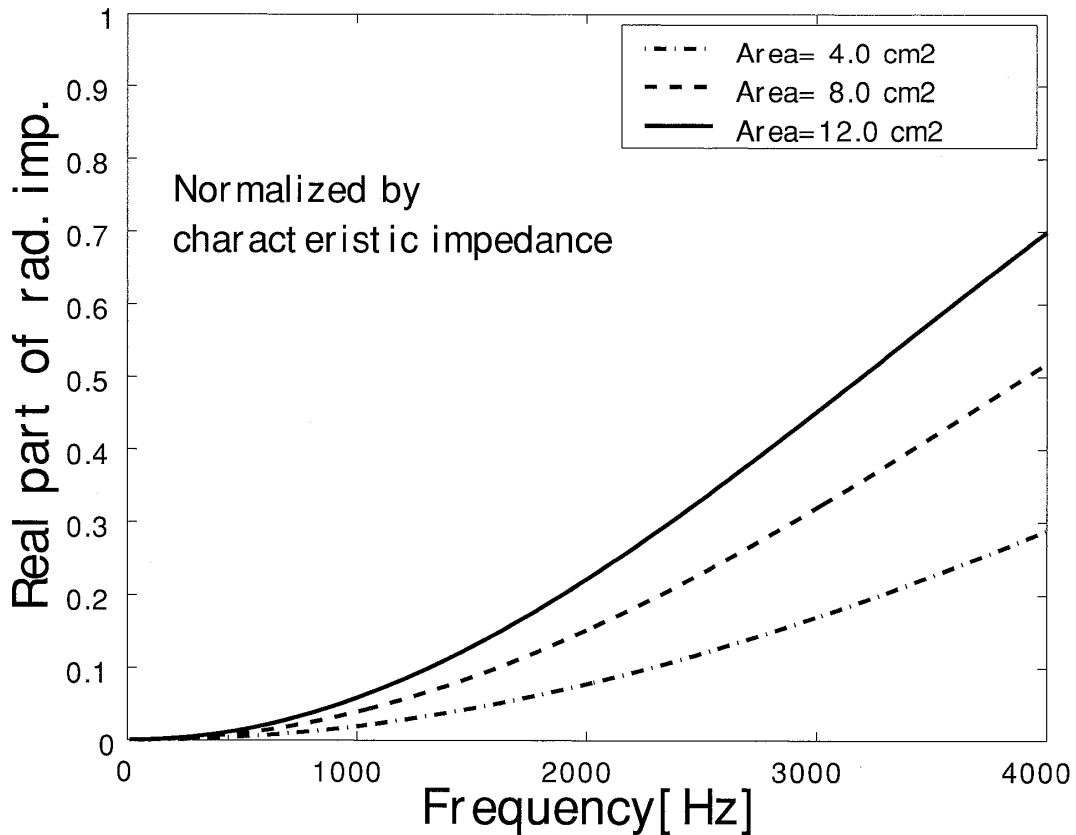


図5：円形ピストン振動板の放射インピーダンスの実部の例．縦軸の値は特性インピーダンス $\rho c / (\pi a^2)$ で規格化している．

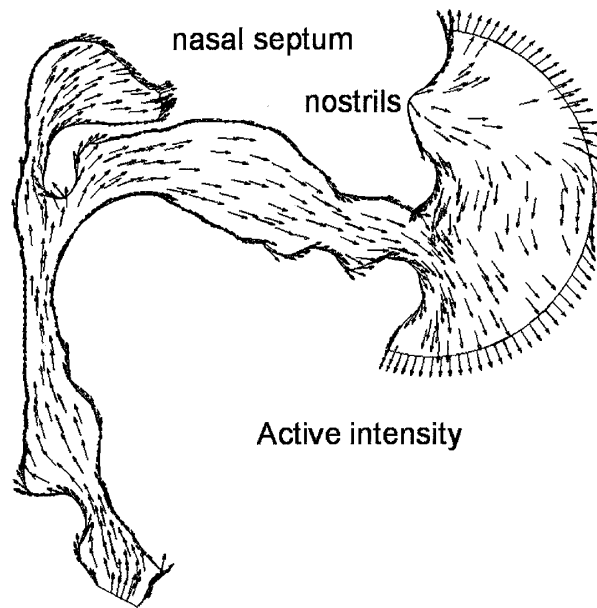


図6：アクティブ音響インテンシティの分布(695Hz)．矢印はインテンシティベクトルの向きを表す．

より得られた695Hzにおけるアクティブ音響インテンシティ $\mathbf{I}(\mathbf{r})$ の正中矢状面での分布図である．矢印の向きは $\mathbf{I}(\mathbf{r})$ の向きを表す． $\mathbf{I}(\mathbf{r})$ の大きさは位置により異なるが，音響パワーフローの様子が分かり易いように矢印は同じ長さで描いてある．口蓋上方は，鼻中隔（nasal septum）の部分なので矢印が描かれていない．音響パワーは口腔と鼻腔に分かれて伝搬し，鼻腔内ではさらに左右の鼻道に分かれる．口唇および鼻孔（nostrils）から放射される様子が分かる．放射空間の境界面（図2（c）の点Aを通る半球）を開口部を覆う曲面Cと考えて，この面上での $\mathbf{I}(\mathbf{r})$ の総和を取ることで全放射パワー W_c を求めることができる．図7の太線は，式（11）により伝達特性を評価したものである．ただし，ここで K は周波数によらない定数とした．また，細線は，図2（c）のA，B，Cの各位置で音圧 $p(\mathbf{r})$ を求め，音源 U_c から各点への伝達インピーダンス $|p(\mathbf{r})/U_c|$ の周波数特性を示したものである．点Aでは，700Hz付近で急峻な低下がみられる．図2（b）の音圧分布から分かるように，695Hzでは口唇と鼻孔放射の干渉により音圧が極めて低くなるので，点Aの伝達特性上では零点のように見える．このように，出力信号として口唇付近の特定の位置における音圧 P_s を考えた場合には，鼻孔と口唇からの放射の干渉により，特定の周波数において出力が極めて小さくなる場合がある．このため，伝達インピーダンス $Z_s = P_s/U_c$ により伝達特性を評価すると，アーチファクトとしての零点が出現する可能性がある．実効的な放射パワーによる式（11）に基づく評価では，この周波数では急峻な零点とはならない．外部への音響パワーの放射が小さくなる複数の零点が3～4.5kHzに生じることが分る．

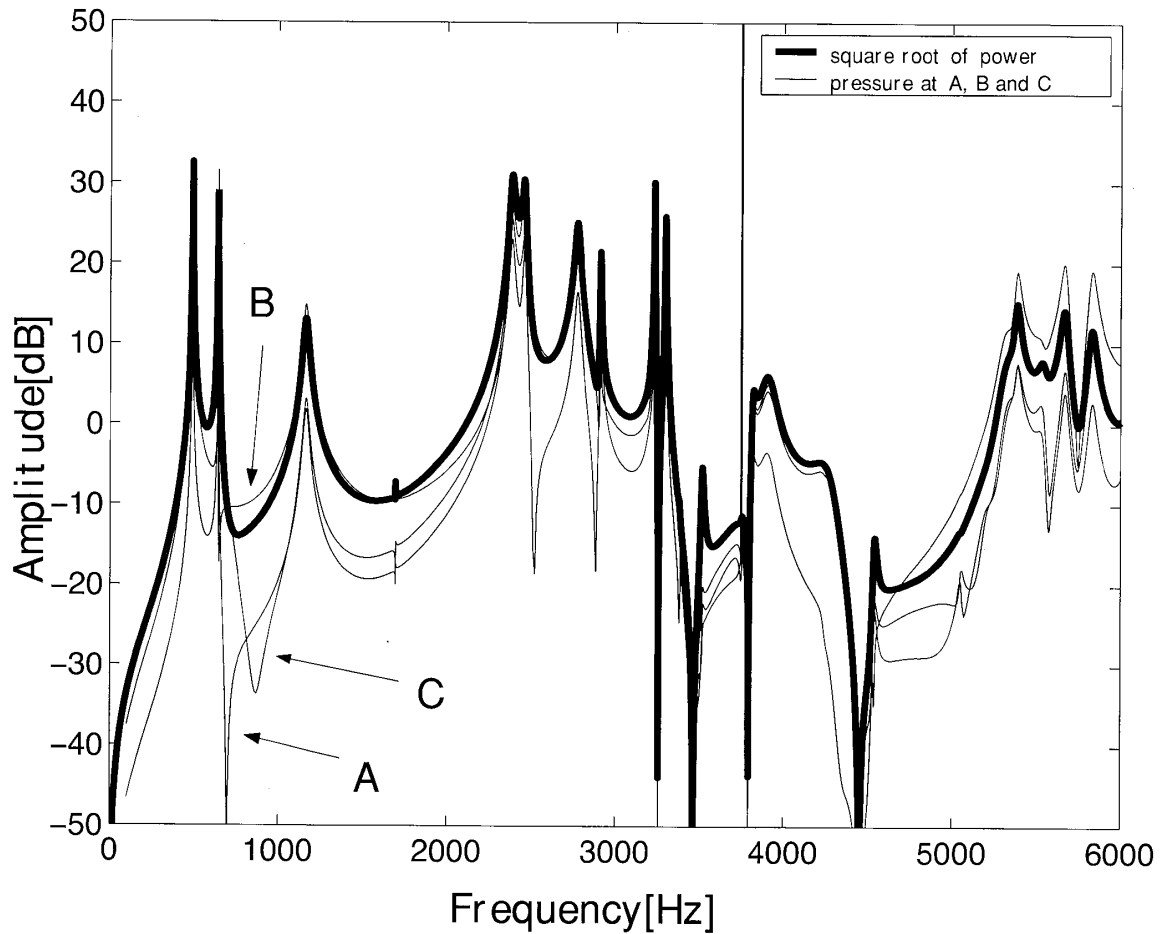


図7：太線：放射パワーに基づく伝達特性，細線：図2（c）の点A,B,C各点における音圧に基づく伝達特性．

3.2 伝達特性の近似表現

式（11）では，声道の伝達特性に関する位相は失われるが，通常は位相に関する情報は人には知覚されないため，声道モデルはその振幅特性で評価される．周波数域での解析は，駆動音源の時間因子を $\exp(j\omega t)$ として周波数域で行われるが，得られた伝達特性を s 領域（ s は Laplace 変換において用いられる複素変数）で近似的に表現できれば，周波数域での解析や時間域への変換に際して有用なことが多い．

複数のピークを持つ式（11）の伝達特性を各ピークに対応する2次システムが並列に接続されたものとして近似すると， i 番目のピークに対応する2次システムの分母多項式 $D_i(s)$ は，共振角周波数を ω_i ，共振回路の Q を Q_i とし次のように表現できる．

$$D_i(s) = (s - s_i)(s - s_i^*) = s^2 + \frac{\omega_i}{Q_i}s + \omega_i^2 \quad (16)$$

ここで， s_i は極を表す．式（11）から得られる伝達特性からピーク周波数 f_i と帯域幅 B_i を求めると， $Q_i = f_i/B_i$ なので，

表1：ピーク周波数 f_i [Hz]，ピーク値 A_i (A_i はピークでの振幅絶対値)，帯域幅 B_i [Hz]，Quality factor Q_i ，極 $s_i = \alpha_i$ [Np/s] + $j\omega_i$ [rad/s].

i	f_i	A_i	B_i	Q_i	$s_i = \alpha_i + j\omega_i$
1	488	42.4	2.1	238	$-6.5 + j 3066$
2	639	28.2	2.7	233	$-8.6 + j 4015$
3	1163	4.5	33.7	34	$-106.0 + j 7307$
4	2386	35.9	33.6	71	$-105.4 + j 14992$
5	2455	33.4	19.8	124	$-62.1 + j 15425$
6	2773	17.9	31.6	88	$-99.3 + j 17423$
7	2912	11.8	5.5	531	$-17.2 + j 18297$
8	3231	32.0	2.3	1419	$-7.2 + j 20301$
9	3297	19.7	3.7	885	$-11.7 + j 20716$
10	5387	5.7	47.0	115	$-147.8 + j 33848$
11	5665	5.2	44.4	127	$-139.6 + j 35594$
12	5830	3.9	48.7	120	$-153.0 + j 36631$

$$s_i = \alpha_i + j\omega_i = -\pi B_i + j2\pi f_i \quad (17)$$

となり，近似的に極の位置を求めることができる．図7で，+10dB以上の振幅値となっているピークを対象に近似的に得られた極の位置を表1に示す．この結果は声道内部を無損失としたシミュレーションモデルによるものであるため，各ピークの帯域幅は実音声やアクリル素材による物理モデルによる合成音声¹⁴⁾に比べても非常に狭くなっている．声道の壁インピーダンスを考慮したシミュレーションでは，帯域幅は大幅に広がり，近接した2つのピークが1つに融合する現象もみられることが報告されている⁴⁾．

音声合成などの時間域での処理を考える場合には，伝達特性のインパルス応答と音源波形を畳み込むことになるが，零位相の伝達関数のインパルス応答では時間の原点に波形が集中する．式(17)で近似した極の位置を用いてピーク周辺の位相特性を求め，これを図7の振幅特性に対する位相特性として付加すると，時間的にやや広がったインパルス応答を得ることができる．強い周波数成分に対して2次システムで近似した場合の位相を与えたものと解釈でき，信号の飽和を防ぐ点で効果があると考えられる．

4 おわりに

本稿では，放射パワーに基づいて3次元声道モデルの伝達特性を評価する方法を述べた．極や零点の位置については，従来から用いられている声門と口唇での体積速度の比として定義される伝達関数とほぼ同等に評価することができる．放射パワーは，放射空間でのインテンシティの和として求めることができるので，口唇付近が複雑な形状であっても特定の出力点を設定することなく声道モデルの共振特性を評価できることを示した．

謝辞 本研究の一部は、北海学園大学ハイテクリサーチセンタープロジェクト、及び、科学研究費補助金（18300069）の助成により行なわれたものである。

参考文献

- 1) 松崎博季, 元木邦俊, 三木信弘, “母音/u/の口唇部突出しと声道分岐の影響の有限要素法による解析”, 北海学園大学工学部研究報告, **32**, 99–105 (2005).
- 2) 松崎博季, 元木邦俊, “母音/a/発声時における鼻腔を伴う声道の音響特性”, 北海学園大学工学部研究報告, **33**, 171–181 (2006).
- 3) 松崎博季, 元木邦俊, “日本語母音/a/発声時における鼻腔を伴う声道の音響特性—声道壁インピーダンスの影響—”, 北海学園大学工学部研究報告, **34**, 73–81 (2007).
- 4) H. Matsuzaki and K. Motoki, “Study of acoustic characteristics of vocal tract with nasal cavity during phonation of Japanese /a/”, *Acoustical Science and Technology*, **28**, 2, 124–127 (2007).
- 5) K. Motoki and N. Miki, “Effects of higher-order modes on sound radiation characteristics at mouth opening”, *Proc. 3rd Joint Meeting ASA and ASJ*, 2pSC42, 895–898 (1996).
- 6) 元木邦俊, “放射過程におけるモード結合を考慮した音響放射パワーの計測法”, 北海学園大学工学部研究報告, **25**, 355–370 (1998).
- 7) 元木邦俊, “高次モードを考慮した声道音響モデルの構成法”, 日本音響学会誌, **54**, 12, 850–856 (1998).
- 8) 元木邦俊, 松崎博季, “インピーダンス変換に基づく3次元声道モデルの音響特性の計算手法”, 北海学園大学工学部研究報告, **27**, 139–157 (2000).
- 9) 元木邦俊, 松崎博季, “矩形音響管の非対称接続による声道音響特性の計算”, 北海学園大学工学部研究報告, **29**, 293–303, (2002).
- 10) K. Motoki and H. Matsuzaki, “Computation of the acoustic characteristics of vocal tract models with geometrical perturbation by using higher-order modes”, *Bulletin of the Faculty of Engineering Hokkai-Gakuen University*, **32**, 75–85 (2005).
- 11) P. M. Morse and K. U. Ingard, *Theoretical Acoustics*, McGraw-Hill (1968).
- 12) 早坂寿雄, 吉川昭吉郎, 音響振動論, 丸善 (1974).
- 13) A. D. Pierce, *Acoustics*, *Acoust.Soc.Am* (1989).
- 14) E. Maeda, N. Usuki, T. Arai, N. Saika and Y. Murahara, “Comparing the characteristics of the plate and cylinder type vocal tract models”, *Acoustical Science and Technology*, **25**, 1, 64–65 (2004).