タイトル	Computation of the Acoustic Characteristics of Vocal-tract Models with Geometrical Perturbation by Using Higher-order Modes
著者	MOTOKI, Kunitoshi; MATSUZAKI, Hiroki
引用	北海学園大学工学部研究報告, 32: 75-85
発行日	2005-02-21

Computation of the Acoustic Characteristics of Vocal-tract Models with Geometrical Perturbation by Using Higher-order Modes

Kunitoshi MOTOKI* and Hiroki MATSUZAKI*

Abstract

This paper presents computational results of sound pressure distributions and transfer characteristics for a large number of vocal-tract models which contain geometrical perturbations. A small change of the vocal-tract shape is regarded as a geometrical perturbation of the axis position of each vocal-tract section. Computation of the acoustic field in the models are performed using higher-order modes. The results indicate that acoustic characteristics in the higher frequencies are highly sensitive to the small change of the vocal-tract shape.

1 Introduction

With the progress of data acquisition techniques such as magnetic resonance imaging, analyses of the acoustic characteristics of three-dimensional vocal-tracts have been performed by many researchers. A finite element method (FEM) is a standard numerical simulation technique suitable for computing the acoustic characteristics of complex vocal-tract shapes at higher frequencies where the assumption of plane wave propagation does not hold. In order to extract important features such as individuality in higher frequency characteristics of vocal-tract resonance, it is presumably necessary to analyze on a large number of vocal-tract shapes. However, FEM requires a large amount of time not only for the computation itself, but also for the creation of finite element meshes. As a matter of fact, proper meshes for the complex vocal-tract shape are only obtained with the assistance of manual adjustment. The simplification of the vocal-tracts with respect to the cross-sectional shape has been examined for the improvement of the efficiency of model creation [1,2].

On the other hand, a parametric method to compute the acoustic characteristics of three-dimensional vocal-tract shape has been developed by using higher-order modes in cascaded rectangular tubes [4,5].

* Department of Electronics and Information Engineering, Hokkai-Gakuen University, Sapporo, Japan
E-mail: {motoki,matsu}@eli.hokkai-s-u.ac.jp

Compared with FEM, although less performance in representing the three-dimensional vocal-tract shape is achieved, much faster computation can be performed and the modification of the model shape is highly facilitated.

Using this method, this paper presents computational results of sound pressure distributions and transfer characteristics for a large number of vocal-tract models which contain geometrical perturbations. A small change of the vocal-tract shape is regarded as a geometrical perturbation of the axis position of each vocal-tract section. These results indicate that acoustic characteristics in the higher frequencies are highly sensitive to the small change of the vocal-tract shape.

2 Computational method with higher-order modes

A cascaded structure of rectangular acoustic tubes, connected asymmetrically with respect to their axes as illustrated in Fig.1, is introduced as an approximation of the vocal tract geometry. The area, ratio of width and height, and the position of the axis of each tube can be specified independently. The walls of the tubes are assumed to be rigid and no loss factors inside the tubes are considered. The three-dimensional acoustic field in each tube can be represented in infinite series of higher-order modes. The sound-pressure $p(x,y,z)$, z being the direction of the tube axis, and the z-direction particle velocity $v_z(x,y,z)$ in each tube are expressed as :

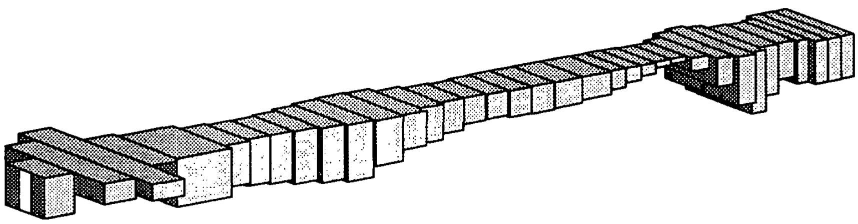


Fig.1 : Connected rectangular tubes as a model of vocal-tract.

$$\begin{aligned} p(x,y,z) &= \sum_{m,n=0}^{\infty} (a_{mn}e^{-\gamma_{mn}z} + b_{mn}e^{\gamma_{mn}z})\phi_{mn}(x,y) \\ &\approx \boldsymbol{\phi}^T(x,y)\{\mathbf{D}(-z)\mathbf{a}+\mathbf{D}(z)\mathbf{b}\} \\ v_z(x,y,z) &\approx \boldsymbol{\phi}^T(x,y)\mathbf{Z}_C^{-1}\{\mathbf{D}(-z)\mathbf{a}+\mathbf{D}(z)\mathbf{b}\} \end{aligned} \tag{1}$$

where m and n stand for the numbers of the higher-order modes in x and y directions, γ_{mn} and $\phi_{mn}(x,y)$ are the propagation constant and normal function(eigen function),respectively. γ_{mn}'s are imaginary numbers for propagative modes and real numbers for evanescent modes. In the matrix notation in (1), the infinite series are truncated to a certain value. $\mathbf{a}$,$\mathbf{b}$ and $\boldsymbol{\phi}(x,y)$ are column vectors composed of a_{mn}, b_{mn} and $\phi_{mn}(x,y)$, respectively. $\phi_{mn}(x,y)$'s are chosen to have the following orthogonal property :

$$\frac{1}{S}\int_S \phi(x,y)\phi(x,y)^T dS = \mathbf{I} \tag{2}$$

where S is the area of the tube, and $\mathbf{I}$ a unit matrix. $\mathbf{a}$ and $\mathbf{b}$ are determined by the boundary conditions at the both ends of each tube. $\mathbf{D}(z)$ and $\mathbf{Z}_C$ are defined as :

$$\mathbf{D}(z) = \mathrm{diag}[\exp(\gamma_{mn}z)],\ \mathbf{Z}_C = jk\rho c(\mathrm{diag}[\gamma_{mn}])^{-1} \tag{3}$$

where k,ρ and c are wave number, air density and sound speed, respectively. A modal sound-pressure vector $\mathbf{P}$ and a modal particle velocity vector $\mathbf{V}$ can be defined as :

$$\mathbf{P} = \mathbf{D}(-z)\mathbf{a} + \mathbf{D}(z)\mathbf{b} \tag{4}$$
$$\mathbf{V} = \mathbf{Z}_C^{-1}\{\mathbf{D}(-z)\mathbf{a} - \mathbf{D}(z)\mathbf{b}\}.$$

If each component of $\mathbf{P}$ and $\mathbf{V}$ is regarded as a voltage and a current at position z, each higher-order mode can be represented by an equivalent electrical transmission line. It should be noted that transmission lines with imaginary characteristic impedances correspond to evanescent higher-order modes.

A sound source can be specified as an arbitrary vibrating surface on the entrance of the first section. The last section corresponding to the mouth is open to a free space. The radiation of sound is taken into account.

With these assumptions, the three-dimensional acoustic field in this asymmetric tube can be calculated by the superposition of plane waves and several higher-order modes [5]. As one higher-order mode corresponds to one transmission line, the equivalent electrical circuit can be obtained as illustrated in Fig.2. The mode coupling at the junction between tubes is represented by an ideal transformer (I.T.,Φ). $\mathbf{Z}$'s are input impedance matrices looking toward the arrowed directions. The number

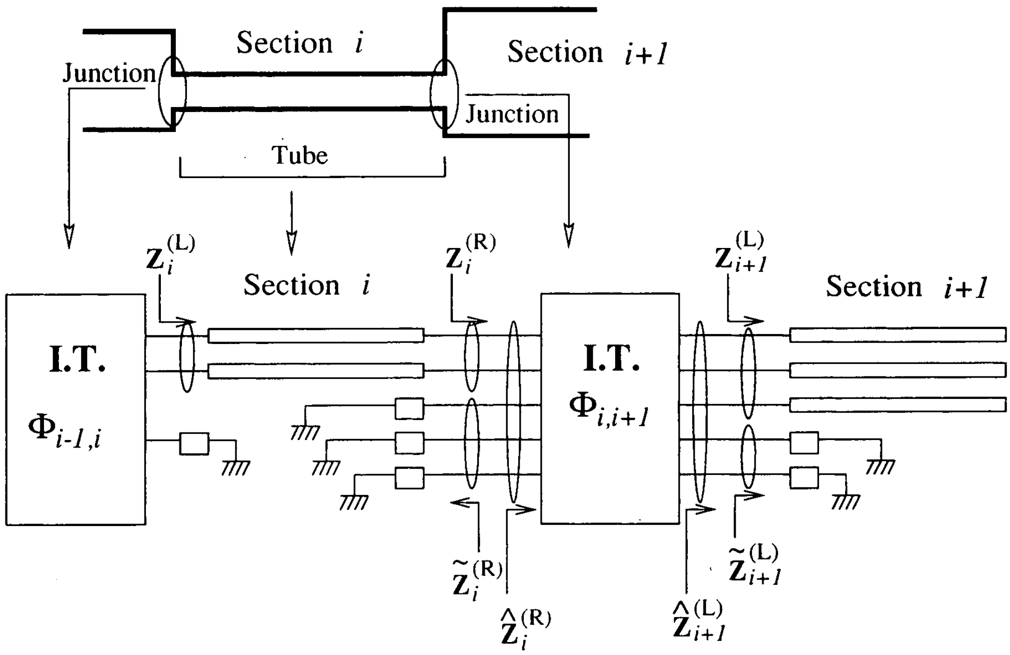


Fig.2 : Equivalent electrical circuit of 3-D vocal-tract model.

of higher-order modes including evanescent modes in each tube is an important parameter. The evanescent modes are acoustical modes that are mainly excited at the part of discontinuity in the shape, and are localized in the vicinity of the excitation part. If no higher-order mode is considered, i.e. plane waves only, this model is identical to the ordinary one-dimensional transmission line model of speech production.

3 Geometrical perturbation

A small change of the vocal-tract shape may have certain influences on the acoustic characteristics, especially in the higher frequencies where the assumption of the plane wave propagation is not always valid. As the geometric parameters can be easily modified in the method described above, the influences of a small change in the model shape onto the internal acoustic field and the transfer characteristics at higher frequencies can be evaluated.

3.1 Random change of axis position

To imitate the small change, the axis position (x_i, y_i) of the i-th section is modified while the area of that section is kept unchanged. Random perturbation is added to the initial position $(x_{i,0}, y_{i,0})$ both in the vertical and lateral directions as :

$$\begin{aligned} x_i &= x_{i,0} + \varepsilon L_{x,i} R[-1,1] \\ y_i &= y_{i,0} + \varepsilon L_{y,i} R[-1,1] \end{aligned} \tag{5}$$

where $L_{x,i}$ and $L_{y,i}$ are the dimensions of the i-th section, and $R[-1,1]$ a uniform random number between -1 and 1. ε is a parameter to specify the maximum perturbation of the axis position. Fig.3 shows an example of this modification.

3.2 36-section uniform tube

A uniform rectangular tube of the length 16.56cm with the cross-sectional dimension L_x=1.5 cm and L_y=3.9cm is divided into 36 sections of equal lengths. Fig.4 (a) shows the sound pressure distributions in the uniform tube at 2,4 and 8kHz. A sound source in the shape of a slit of 5mm in width is centered on the left end of the tube. Plane waves (mode(0,0)) and 5 higher-order modes of (1,0), (2,0),(0,1),(0,2),(1,1) are considered in the computation. It is clearly seen that wave fronts are spreading in an arch shape in the vicinity of the sound source. These patterns are represented by the superposition of plane waves and the higher-order modes that are all evanescent. Consequently the acoustic field in the central region of the tube is almost one-dimensional. A slight pressure increase can also

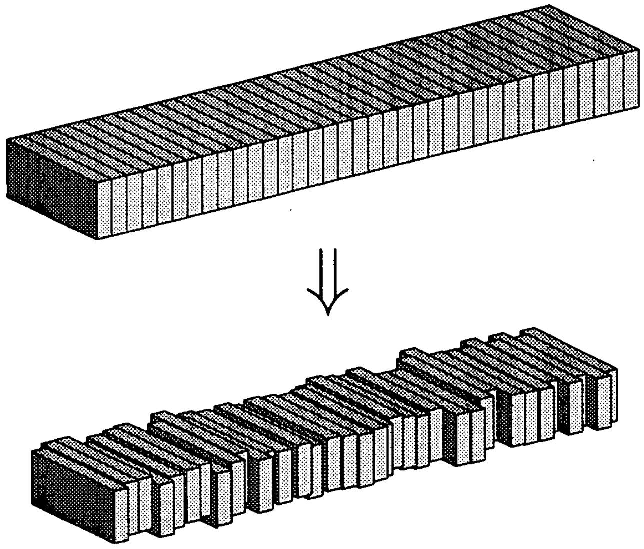


Fig.3 : Geometrical perturbation caused by a random change of axis positions.

be seen in the vicinity of the right end because of the sound radiation. Fig.4 (b) shows the sound pressure distributions in the model with the geometrical perturbation (ε=0.08).In this model higher-order modes are excited not only at the sound source section, but also at all the junctions between tubes. The cut-off frequency $f_{(m,n)_i}$ of mode (m,n) in the i -th section is calculated as follows :

$$f_{(m,n)_i} = \frac{c}{2\pi}\sqrt{\left(\frac{m\pi}{L_{x,i}}\right)^2 + \left(\frac{n\pi}{L_{y,i}}\right)^2} \tag{6}$$

where c is sound speed. This model has the cut-off frequencies of $f_{(0,1)_i}$=4.53kHz, $f_{(0,2)_i}$=9.06kHz for all sections. The curved patterns of sound-pressure contours at 2 and 4kHz are composed of plane waves and the higher-order modes that are again all evanescent. Comparing Fig.4 (b) with (a), it is observed that influences of the geometrical perturbation is not large in low frequency. However, above the first cut-off frequency, the first higher-order mode begins to propagate resulting in an asymmetric acoustic field. The spatial phase distributions corresponding to Fig.4 are shown in Fig.5. At 8kHz spatial phase rotation can be seen in the vicinity of the sound source. It is interesting to note that the sound-pressure becomes absolutely zero (Fig.4 at 8kHz) at the center of phase rotation since the phase becomes discontinuous at these points.

Fig.6 shows the transfer characteristics with increasing ε from 0 (no perturbation) to 0.08. For each ε except 0, 50 models are constructed. The transfer characteristics of the uniform tube show three sharp peaks above 9kHz where mode (0,2) becomes a propagating mode. In the uniform model, mode (0,1) is not excited since the model shape is completely symmetric. In other models the transfer characteristics fluctuate above 4.5kHz because of the influence of mode (0,1) that becomes a propagating mode. It is seen that the transfer characteristics tend to fluctuate more as ε increases. The geo-

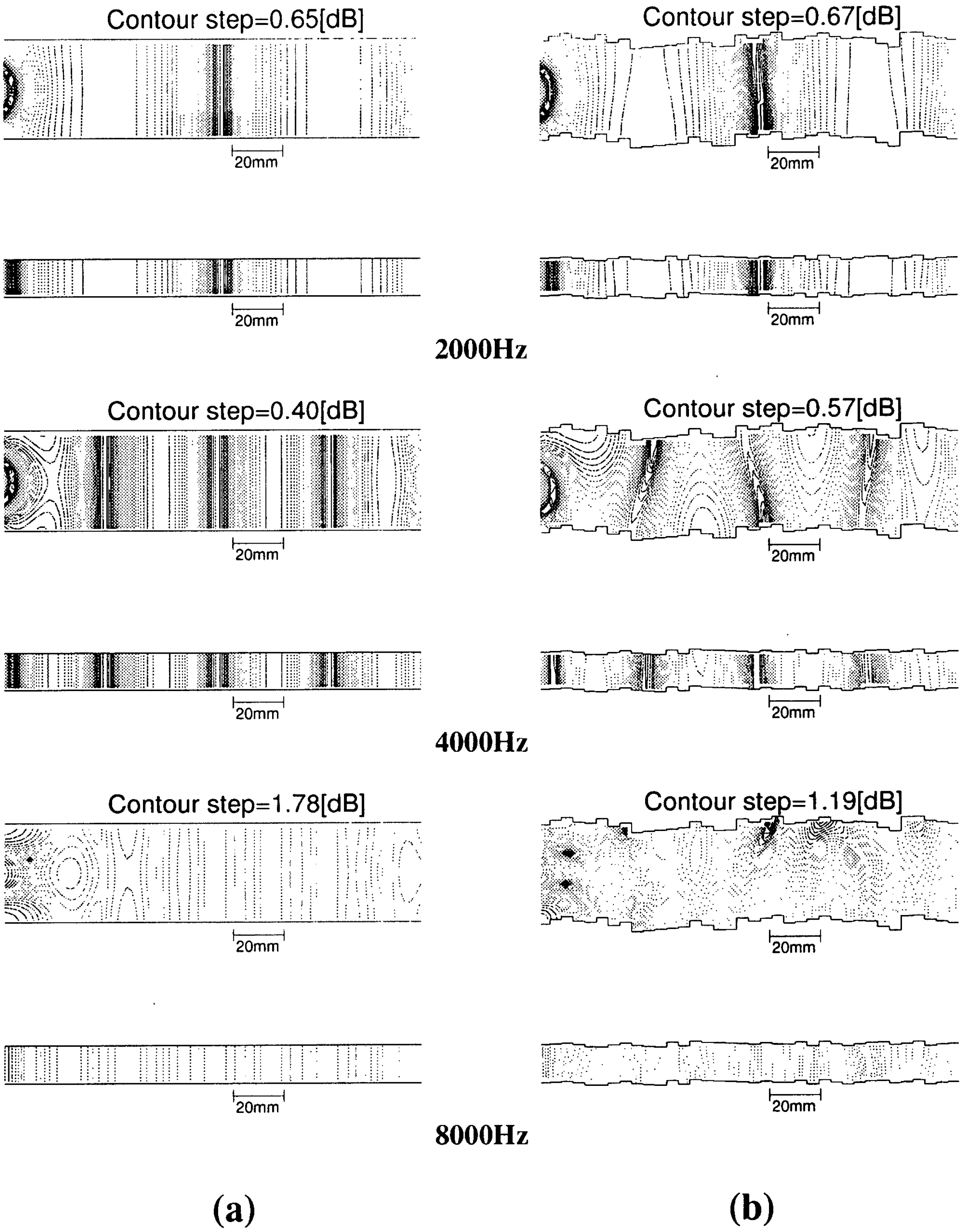


Fig.4 : Amplitude contours in (a) uniform rectangular tube and (b) tube with geometrical perturbation with $\varepsilon=0.08$. Sound source is located at the center of the left end of the tube. Upper and lower charts at each frequency correspond to lateral and vertical distribution of sound pressure.

metric perturbation has little influence on the transfer characteristics below 4.5kHz. However, small downward shifts of peak frequencies are observed with the increase of ε. This can be explained by the increase of the offset of axis positions that can induce the extension of the propagation path.

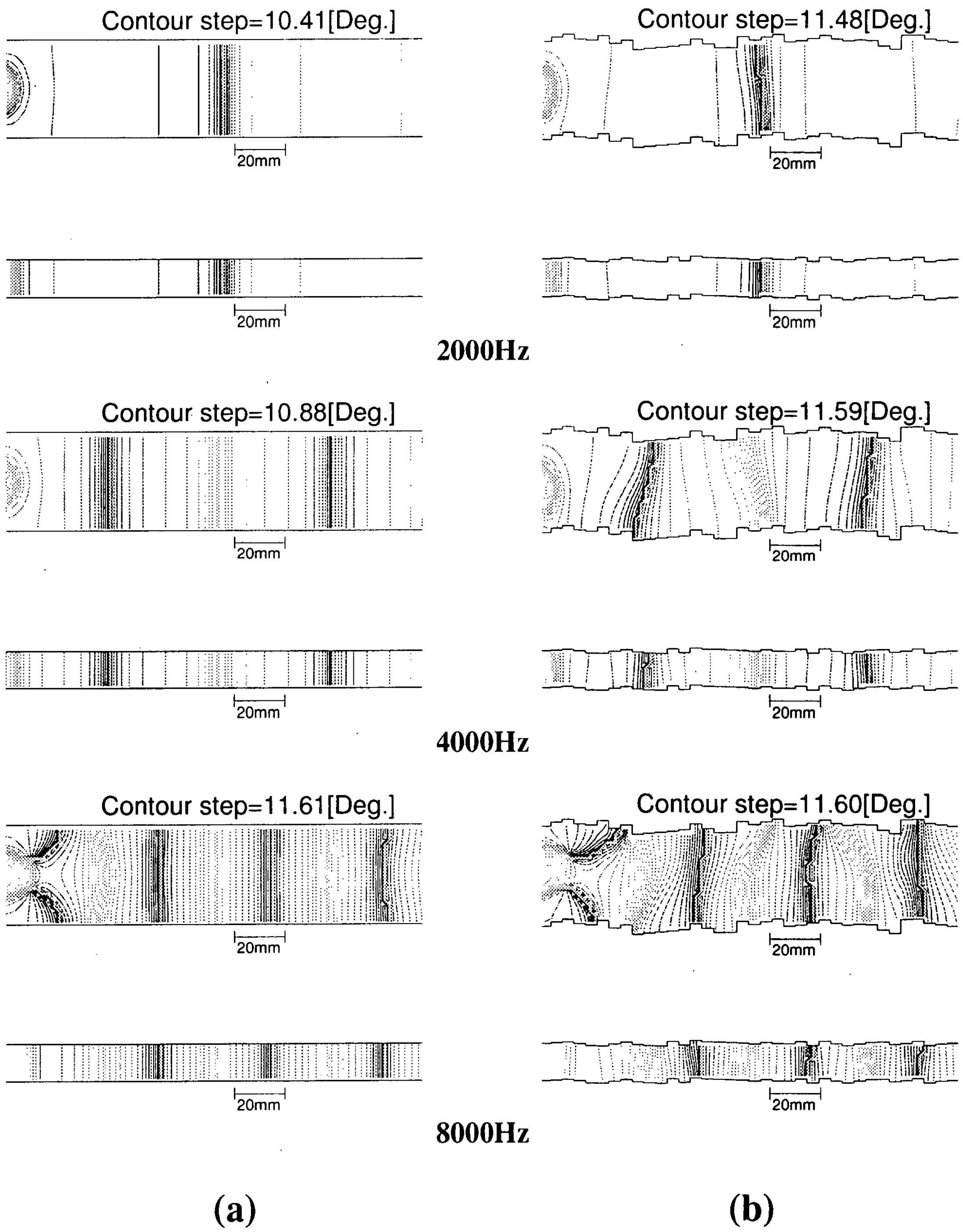
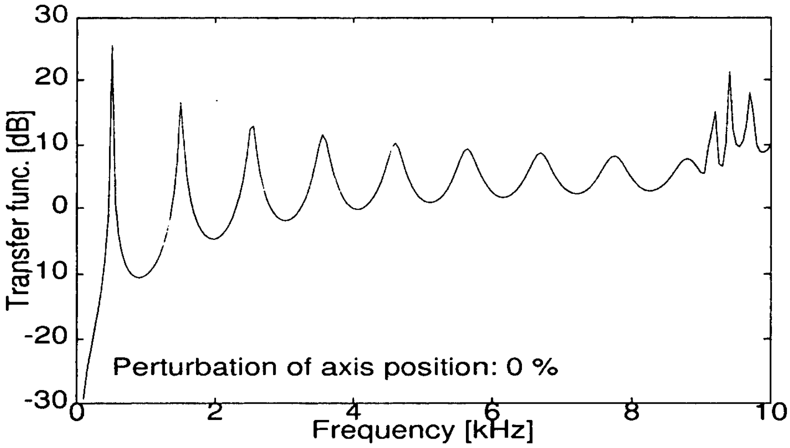


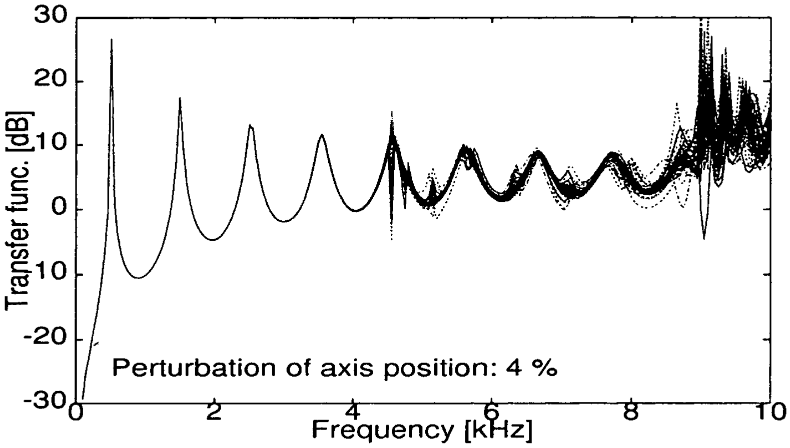
Fig.5 : Phase contours corresponding to Fig.4. (a) uniform rectangular tube and (b) tube with geometrical perturbation.

3.3 Model from MRI data

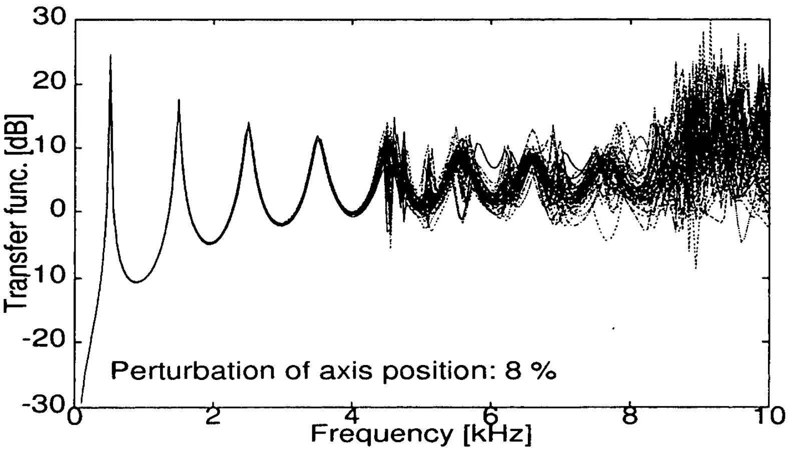
The three-dimensional shape of a / ʃ / measured by MRI has been converted into 36-section models. First each tube is aligned to a common horizontal plane, keeping the symmetry with respect



uniform tube ($\varepsilon = 0$)



$\varepsilon = 0.04$



$\varepsilon = 0.08$

Fig.6 : Transfer characteristics for different perturbation ε.50 models are constructed for each ε.

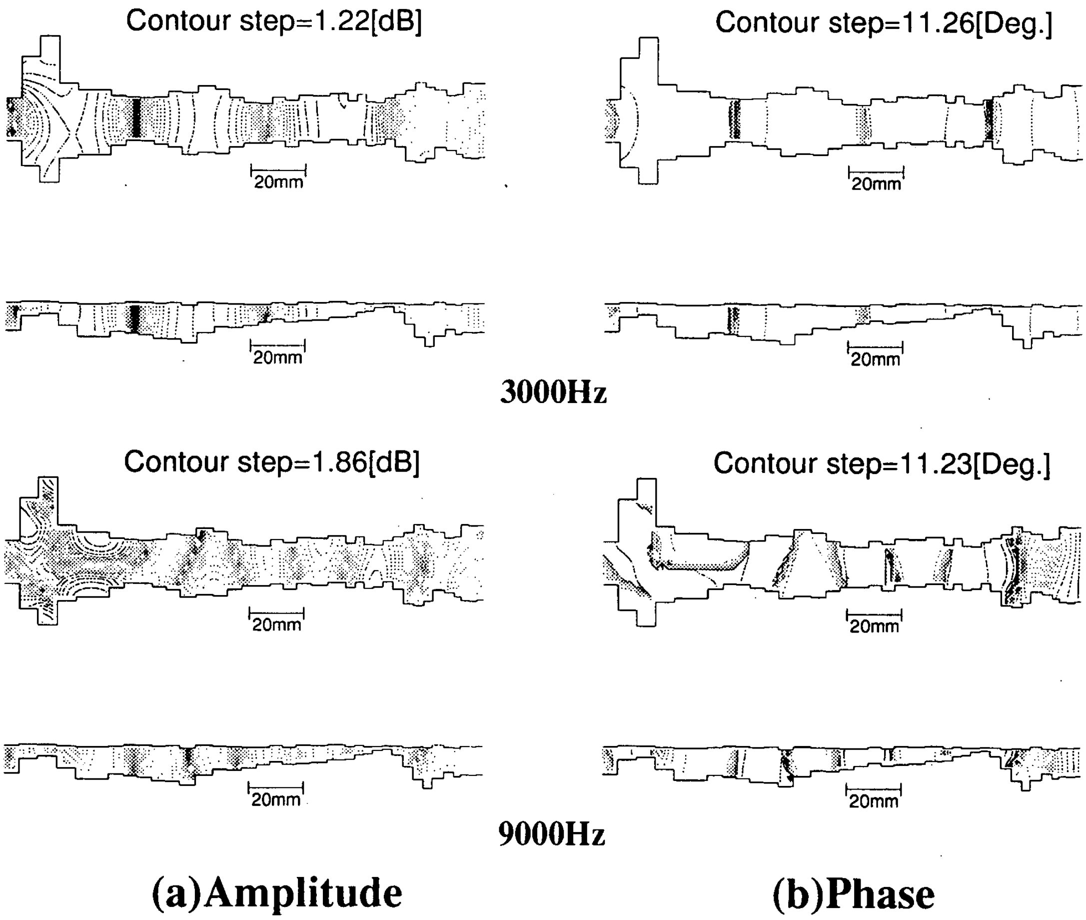


Fig.7 : (a) Amplitude and (b) Phase contours of asymmetrical 36-section model. Upper and lower charts at each frequency correspond to lateral and vertical distribution of sound pressure.

to the lateral direction. Then the axis position of each tube is perturbed randomly both in the vertical and lateral directions with ε=0.08. As the cut-off frequencies differ in each section, 5 higher-order modes are selected in order of frequency, and are considered in the computation together with plane waves. Fig.7 shows the outline of one of the models and internal sound pressure distributions at 3 and 9kHz. Both amplitude and phase distributions are presented. It is observed that even though the model is slightly asymmetric with respect to the lateral direction, the acoustic field becomes asymmetric as the frequency increases. Fig.8 shows the transfer characteristics for 100 models when ε=0.08 is specified. It is clearly seen that the transfer characteristics above 4kHz are largely influenced by the small change of the axis position. The appearance of zeros and their frequencies tend to be more sensitive to the small change of axis position. As there is no difference in area functions for these models, transfer characteristics in the lower frequencies are almost the same. It can be said that a small difference in the vocal-tract shape greatly affects the acoustic field, and consequently changes the resonant characteristics at higher frequencies.

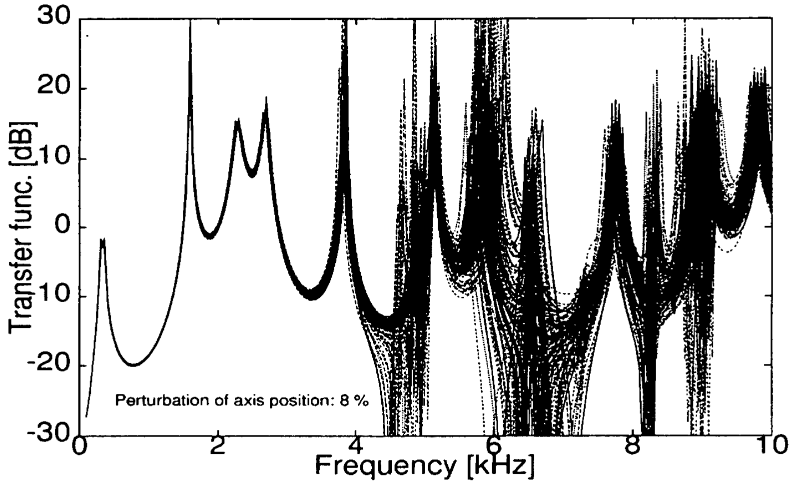


Fig.8 : Transfer characteristics of 100 models with 8% perturbation of axis position.

4 Conclusions

It has been shown that faster computation and evaluation for a large number of the vocal-tracts can be performed using the parametric method with the higher-order modes. The computational results indicate that acoustic characteristics in the higher frequencies are highly sensitive to the small change of the vocal-tract shape while the transfer characteristics in the lower frequencies are very stable. These computational results can be related to the results of psychoacoustic experiments which show that the speaker individualities exist in spectral envelopes in the higher frequencies [6]. One implication is that a part of the information of individuality is carried through the difference in spectrum in higher frequencies that may be expressed by a small difference in the vocal-tract shape.

Acknowledgment

We wish to thank Dr. P.Badin and Dr. X.Pelorson, ICP, France, and Prof. N.Miki, Future University-Hakodate, Japan. Part of this work has been supported by a research project of High-Tech Research Center, Hokkai-Gakuen University.

This paper has been presented at the International Symposium on Communications and Information Technologies 2004 (ISCIT 2004), Sapporo, Japan.

References

[1] H. Matsuzaki, and K. Motoki, "FEM analysis on acoustic characteristics of vocal tracts shape with different geometrical approximation,"Proc. ICSLP2000, Beijing, pp.897-900, 2000.

[2] H. Matsuzaki, K. Motoki, and N. Miki,"A study of the simplification of the three-dimensional vocal-tract model using finite element method," Proc. 18th ICA, Kyoto, Tu4.H.1, 2004.

[3] K. Motoki, and N. Miki, "Effects of higher-order modes on sound radiation characteristics at mouth opening," Proc.3rd Joint Meeting ASA and ASJ, Honolulu, 2pSC42, pp.895-898, 1996.

[4] K. Motoki, and H. Matsuzaki, "A model to represent propagation and radiation of higher-order modes for 3-D vocal-tract configuration," Proc. ICSLP98, Sydney, Fr1R14, pp.3123-3126, 1998.

[5] K. Motoki, X. Pelorson, P. Badin, and H. Matsuzaki, "Computation of 3-D vocal tract acoustics based on mode-matching technique," Proc. ICSLP2000, Beijing, pp.461-464, 2000.

[6] T. Kitamura, and M. Akagi, "Speaker individualities in speech spectral envelopes," J. Acoust. Soc. Jpn.(E), vol.16, No.5, pp.283-289, 1995.