Three-dimensional rectangular vocal-tract model with asymmetric wall impedances

Kunitoshi Motoki[1]

[1]Department of Electronics and Information Engineering, Hokkai-Gakuen University, Japan

motoki@eli.hokkai-s-u.ac.jp

Abstract

A method to compute the acoustic characteristics of a simplified three-dimensional vocal-tract model with asymmetric wall impedances is presented. The acoustic field is represented in terms of both plane waves and higher-order modes in tubes. This model is constructed using a connected structure of rectangular acoustic tubes, and can parametrically represent acoustic characteristics in higher frequencies where the assumption of plane wave propagation does not hold. The propagation constants of the plane waves and the higher-order modes are calculated taking account of the asymmetric distribution of the wall impedances which can be specified as different values on four sides of each rectangular tube. The frequency characteristics of the propagation constants and the transfer characteristics of multiple section models are presented.

Index Terms: wall impedance, vocal-tract model, higher-order modes, propagation constant, vocal-tract transfer function

1. Introduction

The acoustic analysis of three-dimensional vocal-tract models at higher frequencies has been performed by many researchers using the geometrical data obtained by magnetic resonance imaging (MRI). A finite-element method (FEM) is a standard numerical simulation technique for investigating the effects of the fine three-dimensional structure of the vocal tract, such as small dips or branches [1–3]. However, the FEM requires a large amount of time not only for the computation itself, but also for the creation of finite-element meshes. A finite-difference time-domain method (FDTD) is also used for the analysis of the three-dimensional vocal tract [4]. As the computation nodes can be specified at the same position as the voxels of the MRI data, the FDTD has an advantage for the creation of the entire configuration of the vocal tract. The FDTD is also suitable for parallel computing, but encounters difficulties in representing the impedance properties of the boundaries.

On the other hand, a parametric method to compute the acoustic characteristics at higher frequencies, where the assumption of plane wave propagation does not hold, has been developed by using higher order modes [5–8]. The vocal tract is approximated by a cascaded structure of rectangular acoustic tubes. The acoustic field in this model is represented in terms of both plane waves and higher order modes. Although the performance in representing the detailed structure of the vocal tract is limited, this model can represent the effects of the transverse dimension of the vocal tract at higher frequencies, and has the advantage of fast computation since this method can be regarded as the extension of a well-known one-dimensional vocal-tract model.

In the previous model [9, 10], a wall impedance of the vocal-tract was introduced for the computation of a propagation constant in the rectangular tube only when the same impedance is given on the whole wall of a unit section. The properties of the vocal-tract wall, however, are not symmetric. For example, relatively harder wall properties will be suitable on the upper wall, representing the hard palate, in the oral cavity and more yielding wall will be suitable for the lower wall corresponding to the tongue surface.

In this paper, a method to calculate the acoustic characteristics of a rectangular tube model having an asymmetric distribution of the wall impedance is presented. The propagation constants of the plane waves and the higher-order modes are calculated taking account of the wall impedances which can be specified as different values on four sides of each rectangular tube. As the computational results, the frequency characteristics of the propagation constants and the transfer characteristics of multiple section models are also presented.

2. Rectangular tube model

2.1. Overview of the model

The aim of this model is to present a new method of computing the acoustic characteristics of the vocal tract which has the advantages of less computing costs compared to the full three-dimensional modeling like FEM and FDTD, and more capability to represent a geometrical complexity compared to the conventional one-dimensional model.

A cascaded structure of rectangular acoustic tubes, connected asymmetrically with respect to their axes, as illustrated in Figure 1, is introduced as a simplified approximation of the vocal-tract geometry. The width, height, length, and the position of the axis of each tube can be specified independently. A sound source can be specified as an arbitrary vibrating surface at the entrance of the first section. The last section, corresponding to the mouth, is open to free space. The radiation of sound is taken into account. The wall of each rectangular tube has yielding properties that is represented in terms of a complex value of the wall impedance. The wall impedance can be specified as different values on four sides of each rectangular tube.

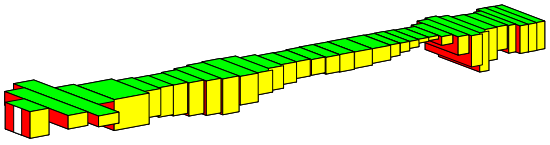


Figure 1: *Cascaded rectangular tube model.*

2.2. Calculation of propagation constant with asymmetric wall impedance

A sound pressure $p(x, y, z)$, z being the direction of the tube axis, and a particle velocity $\boldsymbol{v}(x, y, z)$ can be expressed as

$$p(x, y, z) = j\omega\rho\phi(x, y, z) \tag{1}$$
$$\boldsymbol{v}(x, y, z) = -\nabla\phi(x, y, z) \tag{2}$$

where ω and ρ are angular frequency and air density, respectively. $\phi(x, y, z)$ is a velocity potential satisfying the Helmholtz equation.

$$\nabla^2\phi(x, y, z) + k^2\phi(x, y, z) = 0 \tag{3}$$

$k = \omega/c$, c being sound speed, is the wave number. The three-dimensional acoustic field in each tube can be represented as an infinite series of higher order modes as

$$\phi(x, y, z) = \sum_{m,n=0}^{\infty} \{a_{mn}\exp(-\gamma_{z,mn}z) + b_{mn}\exp(\gamma_{z,mn}z)\}\psi_{mn}(x, y) \tag{4}$$

where m and n stand for the indices of the higher order modes in the x- and y-directions, respectively, and a_{mn} and b_{mn} are constants determined from the geometry of adjacent tubes. $\gamma_{z,mn}$ and $\psi_{mn}(x, y)$ are the propagation constant and normal function (eigenfunction), respectively. $\gamma_{z,mn}$ is required to satisfy the following equation,

$$\gamma_{x,m}^2 + \gamma_{y,n}^2 + \gamma_{z,mn}^2 + k^2 = 0 \tag{5}$$

where $\gamma_{x,m}$ and $\gamma_{y,n}$ are propagation constants in the x- and y-directions, respectively, and are determined by specifying the boundary condition on the wall.

When the same impedance Z_w is given on the both sides of the tube, a gradient of sound-pressure $\partial p/\partial x$, which is proportional to the particle velocity in the x-direction, is always 0 for even modes, and the sound-pressure p is always 0 for odd modes at the center of the tube as shown in Figure 2. In this case, the theory of lined duct [11, 12] is used to obtain the propagation constant $\gamma_{z,mn}$ in the z-direction.

If different wall impedances, Z_{x+} and Z_{x-}, are given on the walls at $x = \pm l_x/2$, respectively, the acoustic field in the tube becomes asymmetric, which induces a difficulty to establish a condition similar to the symmetric impedance case. To overcome this problem, a hypothetical position ξ_x is introduced as a position satisfying the following condition.

Even modes

$$\left.\frac{\partial p}{\partial x}\right|_{x=\xi_x} = 0,$$
$$Z_{x\pm} = \left.\frac{p}{\pm v_x}\right|_{x=\pm l_x/2} = -j\frac{\omega\rho}{\gamma_{x,m}}\coth\left(\gamma_{x,m}\left(\frac{l_x}{2} \mp \xi_x\right)\right) \tag{6}$$

Odd modes

$$p|_{x=\xi_x} = 0,$$
$$Z_{x\pm} = \left.\frac{p}{\pm v_x}\right|_{x=\pm l_x/2} = -j\frac{\omega\rho}{\gamma_{x,m}}\tanh\left(\gamma_{x,m}\left(\frac{l_x}{2} \mp \xi_x\right)\right) \tag{7}$$

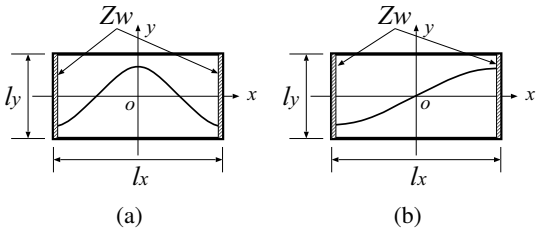


Figure 2: *Examples of modal sound-pressure distributions on the cross-section when the same impedance Z_w is given on the both walls.* (a) *even mode* (b) *odd mode.*

ξ_x satisfying the equations (6) and (7) is obtained algebraically as

$$\xi_x = \begin{cases} \dfrac{\bar{Y}_{x-} - \bar{Y}_{x+}}{\bar{Y}_{x-} + \bar{Y}_{x+}}\dfrac{l_x}{2} & m = 0 \\ \dfrac{jk}{2}\left(\dfrac{l_x}{m\pi}\right)^2\left(\bar{Y}_{x-} - \bar{Y}_{x+}\right) & m \geq 1 \end{cases} \tag{8}$$

where $\bar{Y}_{x\pm}$ is a normalized wall admittance defined as $\bar{Y}_{x\pm} = \rho c/Z_{x\pm}$. Then $\gamma_{x,m}$ is obtained as follows.

$$\gamma_{x,m} = \begin{cases} e^{j\frac{3}{4}\pi}\sqrt{\dfrac{2k\bar{Y}_{x+}}{l_x - 2\xi_x}} & m = 0 \\ \dfrac{j\pi}{l_x - 2\xi_x}\left(m + j\dfrac{2k(l_x - 2\xi_x)\bar{Y}_{x+}}{m\pi^2}\right) & m \geq 1 \end{cases} \tag{9}$$

The normalized wall admittance is assumed to be small satisfying $kl_x|\bar{Y}_{x\pm}| \ll 1$ and the following condition is also assumed for $m \geq 1$ in the above calculation.

$$|\xi_x|/l_x \ll 1 \tag{10}$$

As easily confirmed from equation (8), ξ_x is always 0 when the same wall impedance is given on the both sides of the tube. In the case of asymmetric wall impedance, the hypothetical position ξ_x becomes a complex value which is difficult to consider as an existent position. However, ξ_x may be interpreted as a parameter to modify the distributions of both amplitude and phase of sound-pressure so as to force $p = 0$ and $\partial p/\partial x = 0$.

Similar relations can be obtained in the y-direction by replacing a subscript x by y, and an integer index m by n in equations (8) and (9). Substituting $\gamma_{x,m}$ and $\gamma_{y,n}$ into the equation (5), the the propagation constant $\gamma_{z,mn}$ is obtained. Note that $\gamma_{z,mn}$ has an attenuation constant, (a real part of $\gamma_{z,mn}$), due to the given wall impedances on the four sides of the tube.

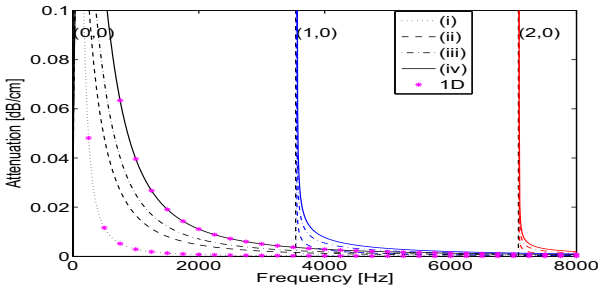
The rigid wall condition corresponds to $\bar{Y}_{x\pm} = \bar{Y}_{y\pm} = 0$, and $\gamma_{z,mn}$ is reduced to

$$\gamma_{z,mn} = \sqrt{\left(\frac{m\pi}{l_x}\right)^2 + \left(\frac{n\pi}{l_y}\right)^2 - k^2}. \tag{11}$$

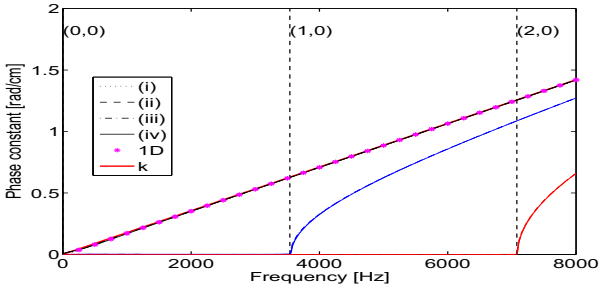
In this condition, the cut-off frequency $f_{c,mn}$ of mode (m, n) is calculated as a frequency where $\gamma_{z,mn} = 0$.

$$f_{c,mn} = \frac{c}{2\pi}\sqrt{\left(\frac{m\pi}{l_x}\right)^2 + \left(\frac{n\pi}{l_y}\right)^2} \tag{12}$$

The higher-order modes are propagative above $f_{c,mn}$, and are evanescent below $f_{c,mn}$.

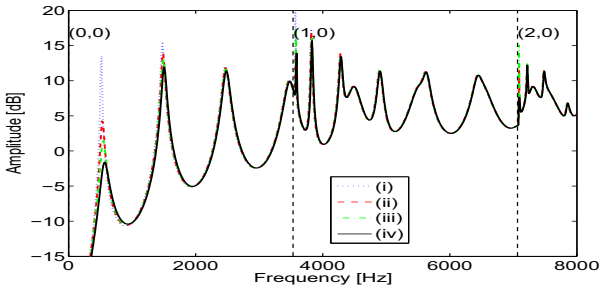


(a)

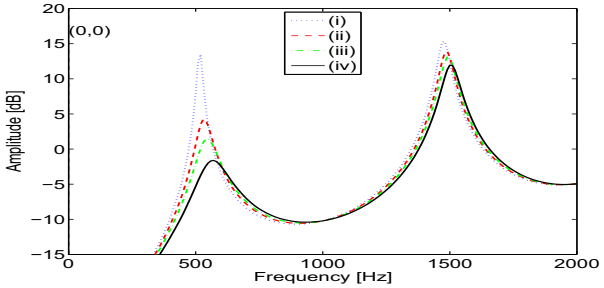


(b)

Figure 3: (a) *Attenuation of modes* (0,0),(1,0) *and* (2,0). (b) *Phase constant.* (i) $L = L_h$ *for all 4 sides,* (ii)$L = L_l$ *for bottom side only,* (iii)$L = L_h$ *for top side only,* (iv) $L = L_l$ *for all 4 sides. Tube size:* $l_x = 5.0$ cm, $l_y = 1.5$ cm. *Vertical dashed lines indicate cut-off frequencies for rigid wall condition. Asterisks indicate one-dimensional model with wall impedance of conditions* (i) *and* (iv). k in (b) is a wave number.



(a)



(b)

Figure 4: *Transfer characteristics of a uniform tube of* l_x =5.0 cm, $l_y = 1.5$ cm *and* $l_z = 17$ cm. *Wall condition is the same as that of Figure 3. Vertical dashed lines indicate cut-off frequencies for rigid wall condition.* (b) *is a magnification of* (a) *up to* 2000 Hz.

3. Computational results

As the wall impedance of the vocal tract is an important physical parameter, its value has been directly measured or indirectly estimated by various methodologies. The measured values showed a relatively wide range depending on the measurement methods and the position in the human body [13–18]. The wall impedance Z_w is composed of a resistive term R, an inertia term L and a stiffness term K, and is represented as $Z_w = R + j(\omega L - K/\omega)$. As the L has a greater influence than variations in R on both plane waves and higher order modes [10], the L is specified as either relatively higher value $L_h = 2.0$ g/cm^2 or lower value $L_l = 0.5$ g/cm^2. Note that L_h and L_l may be upper and lower limit of a proper value of L. The resistive term R is set constant as $R = 1500$ g/(cm$^2 \cdot$ s), and the stiffness term K was ignored in this computation.

Here, the attenuation constant $\alpha_{z,mn}$, a real part of $\gamma_{z,mn}$, and phase constant $\beta_{z,mn}$, an imaginary part of $\gamma_{z,mn}$, are computed for the following different conditions. The value of L on four sides of a rectangular tube is specified as follows.

(i) $L = L_h$ for all 4 sides,

(ii) $L = L_l$ for bottom side and $L = L_h$ for other 3 sides,

(iii) $L = L_h$ for top side and $L = L_l$ for other 3 sides,

(iv) $L = L_l$ for all 4 sides.

The sound speed $c = 3.533 \times 10^4$ cm/s and air density $\rho = 1.142 \times 10^{-3}$ g/cm^3 as the values at a temperature of 36 degrees Celsius are used in the following computation.

The value of $kl_x|\bar{Y}_{x\pm}|$, which is assumed to be small enough, depends on both the frequency and sectional size of the tube. For the above conditions, $kl_x|\bar{Y}_{x\pm}|$ is an order of 10^{-3} at 8 kHz, and $|\xi_x|/l_x$ is also an order of 10^{-3} for $m \geq 1$, which satisfies well the assumption of equation (10).

3.1. Propagation constant

Figure 3(a) shows frequency characteristics of the attenuation of modes (m, n) = (0,0),(1,0) and (2,0), per unit length, for the tube dimensions of $l_x = 5.0$ cm and $l_y = 1.5$ cm. Vertical dashed lines indicate cutoff frequencies $f_{c,mn}$ for the rigid wall condition. Asterisks indicate the attenuation of one-dimensional model with wall impedance of conditions (i) and (iv). For conditions (ii) and (iii), the wall has partially different wall properties. It is clearly seen that the attenuation increases when the inertia term L of the wall decreases. For the higher-order modes, the attenuation is especially large in the vicinity of the cutoff frequencies.

Figure 3(b) shows the phase constant $\beta_{z,mn}$. The difference in $\beta_{z,mn}$ depending on the condition is not eminent. However, closer observation indicates that $\beta_{z,mn}$'s of condition (i)-(iv) are always lower than that of the rigid wall case for all modes. For mode (0,0), the relative difference becomes large compared to the wave number k, i.e. the phase constant of the rigid wall case, in the frequency range below 2 kHz. This means that if a resonance frequency appears in this range, the resonance frequency will shift upward compared with the case of the rigid wall.

3.2. Transfer characteristics

Figure 4 shows transfer characteristics of a uniform rectangular tube with a length $l_z = 17$ cm. The sectional dimensions are the same as used in the previous section. Figure 4 (b) is a magnified characteristics up to 2 kHz. The lower peak frequencies and bandwidths are largely influenced by the wall condition. For the purpose of quantitative comparison, the frequencies (F_1

Table 1: *Peak frequencies* (F_n) *and bandwidths* (B_n). *Unit in Hz. Shift rates* (%) *of the peak frequencies and increase rates* (%) *of the bandwidths are relative to those of the rigid wall case.*

Cond.	F_1	%	F_2	%	F_3	%
Rigid	486	-	1465	-	2458	-
(i)	516	6.2	1476	0.8	2465	0.3
(ii)	534	6.9	1487	1.5	2472	0.6
(iii)	545	12.1	1494	2.0	2476	0.7
(iv)	568	16.9	1505	2.7	2483	1.0
	B_1	%	B_2	%	B_3	%
Rigid	6	-	45	-	106	-
(i)	19	217	47	4.4	108	1.9
(ii)	62	933	56	24.4	110	3.8
(iii)	89	1383	60	33.3	112	5.7
(iv)	135	2150	69	53.3	116	9.4

to F_3) and bandwidths (B_1 to B_3) of the first three peaks are listed in Table 1. The bandwidth of the lower peaks, especially those below 1 kHz, are strongly influenced depending on the condition of L. Compared with the case of the rigid wall condition, a small upward shift of peak frequencies is also seen. This result is consistent with the effects of wall impedance that is already suggested by the conventional one-dimensional model. It is also noted that some sharp peaks above 3.5 kHz are resulting from the resonances of the propagative higher order modes. As seen in the Figure 4 (a), these peaks are less sensitive to the condition of the wall impedance.

An example of 36-section vocal-tract model is illustrated in Figure 5(a), and the transfer characteristics corresponding to each wall condition (i)-(iv) are shown in (b). This model is constructed from the shape of /ʃ/ measured by MRI. Although the original MRI data was not for vowels, a sound source in the shape of a slit (5 mm in width) is assumed at the center of the entrance of the first section. The lower peaks are influenced by the variation of the wall condition. Although the model does not include any branch, the appearance of zeros in the vicinity of 4.6 kHz and 7 kHz are also noted. However, the zeros are also less sensitive to the condition of the wall impedance.

A possible reason for the appearance of zeros are the existence of the strong standing waves in the transverse direction. Figure 6 shows the distribution of the sound-pressure at 4601 and 6952 Hz corresponding to the frequencies of the zeros. A node of sound-pressure with a large change in amplitude in the transverse direction is observed at 4601 Hz in the left part of the model. Similarly two nodes can be seen at 6952 Hz. These are the standing waves that are well agree with the size of the transverse length.

4. Conclusions

A method to calculate the acoustic characteristics of a simplified three-dimensional vocal-tract model with asymmetric distribution of wall impedances was presented. The wall impedances can be specified independently as different values on four sides of each rectangular tube.

The computational results with changing the inertia term of the wall impedance showed that the lower inertia values have effects to shift upward the peak frequencies and increase the

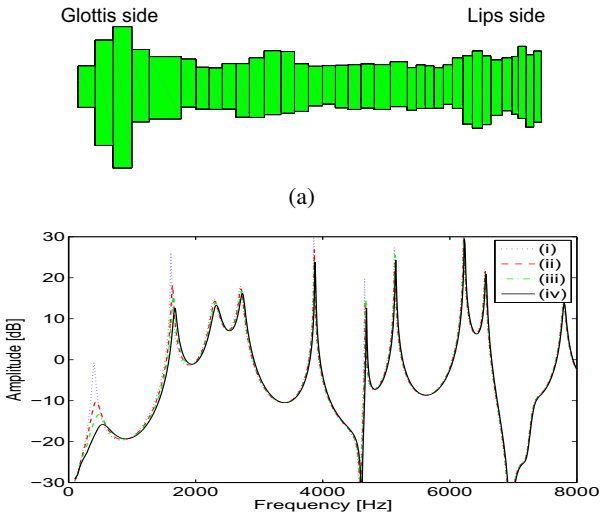


Figure 5: (a) *An example of 36-section vocal-tract model of asymmetric geometrical structure and* (b) *its transfer characteristics. Wall condition is the same as that of Figure 3 and is specified to all sections.*

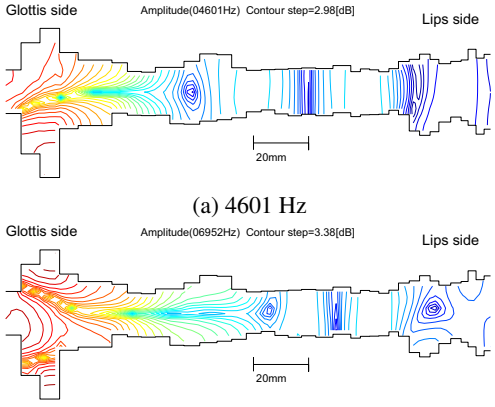


Figure 6: *Sound-pressure distributions at 4601* Hz (*first zero*) *and 6952* Hz (*second zero*). *Wall condition is* (iii).

bandwidths especially at the lower frequencies. Transfer characteristics at higher frequencies were less sensitive to the condition of the wall impedance both for peak and zero frequencies.

The transfer characteristics contains several sharp peaks at higher frequencies due to the resonances of the propagative higher-order modes which are dependent on the transverse dimension of the rectangular tubes. A proper method to convert a complicated shape of the real vocal tract into the simplified rectangular structure should be further studied.

5. Acknowledgements

The author would like to thank Dr. Hiroki Matsuzaki for valuable discussion on acoustic analysis of vocal tract. Part of this work has been supported by a research project of High-Tech Research Center, Hokkai-Gakuen University and JSPS KAKENHI Grant Number 23300071.

6. References

[1] Matsuzaki, H. and Motoki, K., "Study of acoustic characteristics of vocal tract with nasal cavity during phonation of Japanese /a/", Acoust. Sci. and Tech., 28:124–127, 2007.

[2] Matsuzaki, H., Serrurier, A., Badin, P. and Motoki, K., "3D and plane wave acoustic propagation comparison for a Japanese and a French vowel /a/ with nasal coupling," Proc. Autumn Meet. Acoust. Soc. Jpn., 471-474, 2007.

[3] Vampola, T., Horáček, J. and Švec, J.G., "FE modeling of human vocal tract acoustics. Part I: Production of Czech vowels", Acta Acust., 94:433–447, 2008.

[4] Takemoto, H., Mokhtari, P. and Kitamura, T., "Acoustic analysis of the vocal tract during vowel production by finite-difference time-domain method", J. Acoust. Soc. Am., 128:3724–3738, 2010.

[5] Motoki, K. "A method of constructing an acoustic model of vocal-tracts using higher-order modes", J. Acoust. Soc. Jpn., 54:850–856, 1998.

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

[7] Pelorson, X., Motoki, K. and Laboissière, R., "Contribution à l'analyse acoustique du conduit vocal", XXIIIèmes Journées d'Etude sur la Parole, O4-3, 131-134, 2000.

[8] Motoki, K. and Matsuzaki, H., "Computation of the acoustic characteristics of vocal-tract models with geometrical perturbation", Proc. Interspeech2004-ICSLP, TuB602p.16, 521-524, 2004.

[9] Motoki, K., "Effects of wall impedance on transmission and attenuation of higher-order modes in vocal-tract model", Proc. Interspeech2010, Tue-Ses2-S1.3, 1013-1016, 2010.

[10] Motoki, K., "A parametric method of computing acoustic characteristics of simplified three-dimensional vocal-tract model with wall impedance", Acoust. Sci. and Tech., 34(2):113–122, 2013.

[11] Morse, P.M. and Ingard, K.U., *Theoretical Acoustics*,Chap. 9, McGraw-Hill, 1968.

[12] Crocker, M.J. *ed.*, *Handbook of Acoustics*, Chap. 7, Wiley, 1998.

[13] Flanagan, J.L., *Speech Analysis Synthesis and Perception*, 2nd ed., p. 68, Springer-Verlag, 1972.

[14] Ishizaka, K., French J. and Flanagan, J.L., "Direct determination of vocal tract wall impedance", IEEE Trans. Acoust. Speech and Signal Process., ASSP-23:370–373, 1975.

[15] Suzuki, J., "Discussions on vocal tract wall impedance", J. Acoust. Soc. Jpn., 34:149–156, 1978.

[16] Lunde, P., "Acoustic transmission-line analysis of formants in hyperbaric helium speech", Proc. IEEE Int. Conf. on ICASSP'85, 1141-1144, 1985.

[17] Kamiyama, N., Miki, N. and Nagai, N., "Measurement of acoustic reflection characteristics of the human cheek", J. Acoust. Soc. Jpn. (E), 11:207–214, 1990.

[18] Dang, J., Nakai, T. and Suzuki, H., "Measurement of cheek impedance by sound pressure in oral cavity and acceleration of vibrating cheek", J. Acoust. Soc. Jpn., 48:621–628, 1992.