FEM ANALYSIS OF 3-D VOCAL TRACT MODEL WITH ASYMMETRICAL SHAPE

Hiroki MATSUZAKI and Kunitoshi MOTOKI
Department of Electronics and Information Engineering, Hokkai-Gakuen University
S-26, W-11, Chuo-ku, Sapporo, 064-0926, JAPAN
Email: `{matsu,motoki}@eli.hokkai-s-u.ac.jp`

ABSTRACT

The acoustic features of speech characterized by 3-D vocal tract shapes are investigated with a simulation method of a 3-D finite element method (FEM). A simulation model has a cascaded structure of 36 cross sections based on magnetic resonance imaging data (MRI) of the vocal tract for the Japanese vowel /a/ of an adult male. Each cross section of the model is obtained by converting the MRI data into foursquare elements of the equal size. The cross-sectional shape of the simulation model corresponds well with that of the original MRI data. Vocal tract transfer functions are computed from the simulation results, and the effects of the asymmetrical and bent shapes of the vocal tract are discussed.

1 INTRODUCTION

The authors have examined the acoustic characteristics of the vocal tracts using a 3-D FEM which is a simulation method suitable for computing the acoustic field inside an arbitrary 3-D shape.

As a first step, we used a simple simulation model of the vocal tract constructed from cascading acoustic tubes with elliptical shapes[1, 2]. The geometries of each tube were determined from the MRI data in order to have the same perimeter and area. To simulate 3-D radiation from the lips, a 3-D radiation model[3], which is hemi-spherical in shape, was attached to the radiation end. Although this model can not represent the asymmetrical structure of the vocal tract, there are some computational advantages such as less memory and time, small amount of the data of results, and easieness to make finite elements. It is also easier to compare the results with those of an 1-D analysis. The visualization of the acoustic field showed the existence of some higher-order modes in the model below 5kHz. The appearance of zeros in the vocal tract transfer function (VTTF) were observed at particular frequencies because of the 3-D wave propagation. As the effects of the 3-D shapes of vocal tracts on the VTTF are found to be larger, the analyses for an asymmetrical shape of the vocal tract is needed for more precise evaluation.

In this paper, we examine the effects of the asymmetrical shape of the vocal tract using a simulation model which has a cascaded structure of 36 cross sections based on the MRI data of the vocal tract for the Japanese vowel /a/ of an adult male[4]. Each cross section of the model is obtained by converting the MRI data into foursquare elements of equal size. The cross-sectional shape of the simulation model corresponds well with that of the original MRI data. Each element of the model is in the shape of a cube and the same size. The utilization of the cube elements facilitates to construct the model. Furthermore a bent version of the above straight configuration is constructed in agreement with the vocal tract shape. In the case of the bent configuration, all the elements are not the same cubical shape. In these models, the 3-D radiation model with a radius of 4cm is attached to the radiation end. A rigid wall is assumed. The driving surface is driven by the sine waves. The results show that the first three formant frequencies of the VTTFs move to lower frequencies compared to those of the simulation model with the elliptical shape. Especially in the case of the bent configuration, the shift of the third formant frequency becomes larger than that of the straight configuration. This fact suggests that the asymmetrical shape of the vocal tract has an effect on the length of the acoustic vocal tract, which becomes longer in length than the real length. In the VTTFs some sharp peaks appeared at higher frequencies above 3kHz, which might be effects of the asymmetrical shape or the effects of the discontinuous shape of the boundary elements.

2 SIMULATION METHOD

It is well known that an acoustic wave equation in a steady state is represented using velocity potential ϕ as

$$\nabla^2\phi = -k^2\phi \tag{1}$$

where k is the wave length constant. A 3-D FEM is applied to Eq.(1)[5]. From the simulation results of velocity potential ϕ, the particle velocity v is computed by the 3-D FEM applied to the following equation.

$$v = -\nabla\phi \tag{2}$$

3 3-D VOCAL TRACT MODEL WITH ASYMMETRICAL SHAPE

A 3-D vocal tract model has a cascaded structure of 36 cross sections, from the glottis to the lips, and based on MRI data of the vocal tract for the Japanese vowel /a/ of an adult male[4]. Each cross section of the model is obtained by converting the MRI data into foursquare elements of equal size with a length of 2mm as shown in Fig.1. Special cares by sight are taken into account in order to obtain the cross-sectional area and the shapes of the simulation model as consistent as possible to those of the original MRI data. The solid line shows the cross sectional shape of the original MRI data, and the broken line shows the meshing of the foursquare finite element. Each

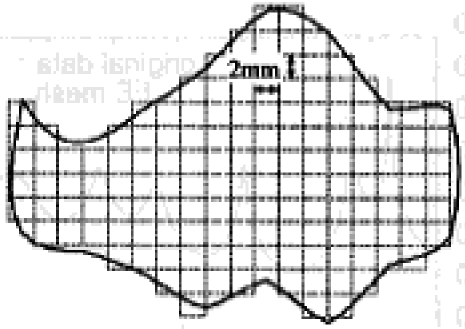


Figure 1: Example of meshing.

element of the model is in the shape of a cube and of the same size. This approach has some advantages such as less memory and time, small amount of the data of results, and it's easier to make finite elements.

Area functions and perimeter functions for the original MRI data and the finite element mesh(FE mesh) are shown in Fig.2 and Fig.3, respectively.

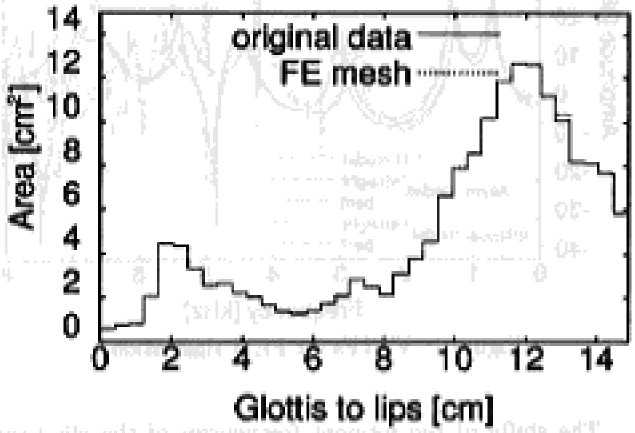


Figure 2: Area functions.

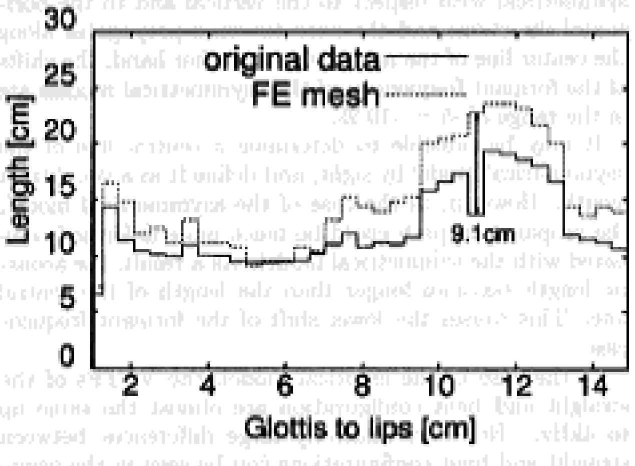


Figure 3: Perimeter functions.

The area function for the FE mesh is almost the same with that of the original data therefore the cross sections are divided as close as possible to that of the original data. The perimeter function for the FE mesh does not correspond well with that of the original data since the perimeter is not considered in the procedure of the mesh creation. The maximum difference is 9.1cm in length. Fig.4 shows the cross sections of this part. The difference in the perimeter becomes larger when the original sectional shape contains an oblique line with the angle ±45 degrees. Note that the perimeter of the simulation model is stepwise.

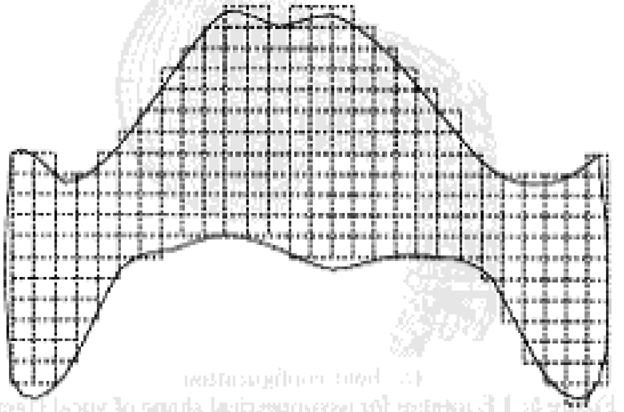


Figure 4: Cross section with the maximum difference in the perimeter functions.

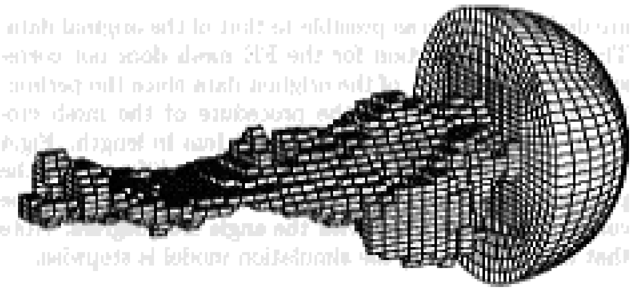
The simulation models are constructed with a cascaded structure of these approximated cross sections. FE meshes are shown in Fig.5(a) for a straight configuration and (b) for a bent configuration, respectively. The bent configuration is obtained by bending the straight configuration. The angle of the bend is approximately the same as that of the original MRI data. A rigid wall is assumed. To simulate the 3-D radiation in a free space, the 3-D radiational model[3] with a radius of 4cm, which is hemispherical in shape, is attached to the aperture surface (the lips side). The driving surface (the glottis side) is driven by particle velocity exp($j\omega t$).

4 SIMULATION RESULTS

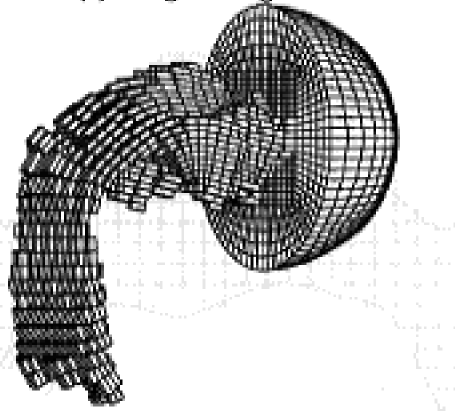
4.1 VTTFs for 1-D model

Before we describe the results of the FEM simulation, we address the differences between the original MRI data and the FE mesh using the VTTFs computed from the traditional 1-D equivalent circuit model[6] with the area functions shown in Fig.2 and the perimeter functions shown in Fig.3.

The VTTFs are shown in Fig.6. Both of results are almost in agreement without a slight difference above 6kHz. Although loss factors are the function of the perimeters, there is little influence due to the differences in the perimeters since the two area functions are almost the same, and the wall is assumed to be rigid



(a) straight configuration



(b) bent configuration

Figure 5: FE meshes for asymmetrical shape of vocal tract.

4.2 VTTFs for FEM simulation

From the results of the FEM simulation, the VTTF $H_v(\omega)$ is defined as

$$H_v(\omega) = 20 \log_{10} \left| \frac{\sum_{n_r} A_r v_r(\omega)}{\sum_{n_d} A_d v_d(\omega)} \right| \quad (3)$$

where $v_d(\omega)$ and $v_r(\omega)$ are the normal components of the particle velocity of the driving surface and aperture surface, respectively. A_d and A_r are the areas of the driving surface and aperture surface, respectively.

The VTTFs computed from the results of the FEM simulation are shown in Fig.7. The solid line shows the VTTF of the 1-D model, and is the same as in Fig.6. The VTTFs for the asymmetrical models are shown as a broken line for the straight configuration and a dotted line for the bent configuration, respectively. For a purpose of comparison, the VTTFs for our FEM models with an elliptical cross-section[2] shown in Fig.8 are also shown as a one-dot chain line for the straight configuration and a two-dot chain line for the bent configuration.

The first three formant frequencies and percentages of shifts of the formant frequencies for the FEM models from the frequencies of the 1-D model are shown in Table.1. The percentage of the shift is defined as

$$\frac{FFo/FEM - FFo/1D}{FFo/1D} \times 100 \quad (4)$$

where FFo/FEM and FFo/1D are the formant frequencies of the FEM model and of the 1-D model, respectively.

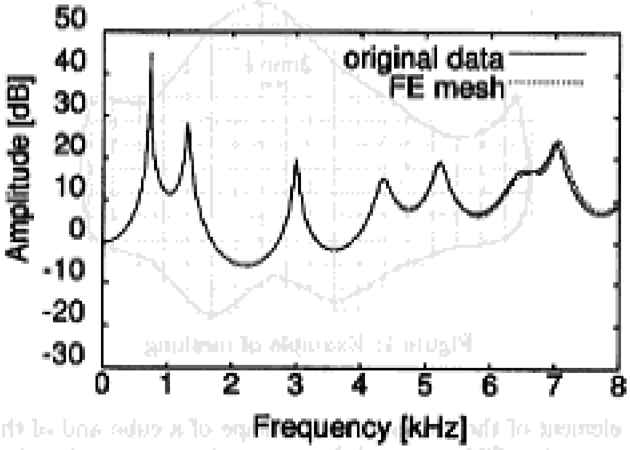


Figure 6: VTTFs computed from the 1-D equivalent circuit model.

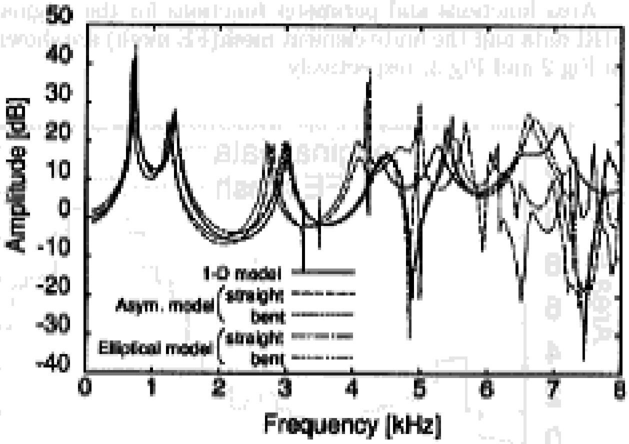
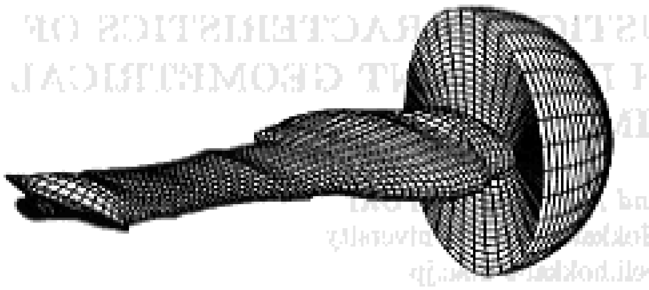


Figure 7: VTTFs for FEM simulation.

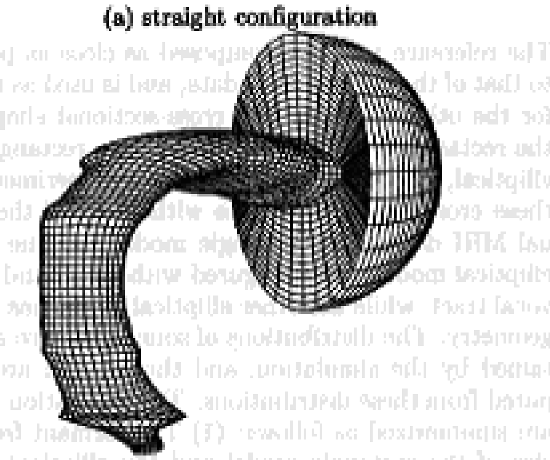
The shifts of the formant frequencies of the elliptical models from those of the 1-D model are only in the range of 0 – -2 %. It seems that the elliptical cross section is symmetrical with respect to the vertical and to the horizontal directions and the acoustic wave propagates along the center line of the model. On the other hand, the shifts of the formant frequencies of the asymmetrical models are in the range of -5 – -10 %.

It may be possible to determine a central line of the asymmetrical model by sight, and define it as a vocal tract length. However, in the case of the asymmetrical model, the propagation path could be much more deformed compared with the symmetrical model. As a result, the acoustic length becomes longer than the length of the central line. This causes the lower shift of the formant frequencies.

In the case of the elliptical model, the VTTFs of the straight and bent configuration are almost the same up to 4kHz. However, relatively large differences between straight and bent configurations can be seen in the asymmetrical model. In both models, lower shifts of the first three formant frequencies compared to the 1-D model are observed. As for the effect of the bent, Sondhi[7] has concluded based on the modal analysis for a uniform bent duct



(a) straight configuration



(b) bent configuration

Figure 8: FE meshes with elliptical cross section.

Table 1: Formant frequencies[Hz] and percentages of the shifts. The shifts are shown between parentheses.

	1-D	asym. model		elliptical model	
		straight	bent	straight	bent
1st	734	683 (-6.9)	676 (-7.9)	722 (-1.6)	726 (-1.1)
2nd	1320	1232 (-6.7)	1248 (-5.5)	1321 (0.1)	1335 (1.1)
3rd	3018	2795 (-7.4)	2718 (-9.9)	2956 (-2.1)	2979 (-1.3)

that a minor increase in upward shift percentage in the formant frequencies might be possible. However, for a bend with varying the area or the cross-sectional shape, it seems that the effects of the bend are smaller than those of the change in cross-sectional shapes or in areas when they are compared with the 1-D model.

In the region of the higher frequency, the VTTFs are completely different. It seems reasonable to suppose that the asymmetrical shape makes the propagation path of the acoustic wave more complex.

The fact that the 1st and 2nd formant frequencies for the asymmetrical models are different from those of the 1-D models and the elliptic models suggests that the influence of the asymmetrical shape can be an important factor of acoustic characteristics.

We see some sharp peaks in the VTTFs of the asymmetrical models above 3kHz. One of the possible reasons may be the effect of the stepwise discontinuous shapes of the asymmetrical models because we do not see any sharp peak in the elliptical models.

5 CONCLUSION

Using the 3-D FEM we have simulated the acoustic wave propagation in the 3-D vocal tract models with the asymmetrical shape.

In the VTTFs of the asymmetrical models, the formant frequencies were appreciably lower compared to those of the elliptical models. The formant frequencies for the bent configuration was also lower than those for the straight configuration. From these facts it can be said that the asymmetrical shape of the vocal tract has a large influence on the acoustic characteristics. The effect of the stepwise discontinuous shapes, which are regarded as artifacts in the vocal tract model, should be clarified. From this point of view we might go on to an even more detailed examination of the asymmetrical shape of the vocal tract using a precise approximated model with smooth connection. The effective way to convert the complex 3-D vocal tract into the simplified shape suitable for the simulation may be found from the results of the precise model.

ACKNOWLEDGMENTS

Part of this work has been supported by a research project of High-Tech Research Center (Hokkai-Gakuen University), and by Crest of JST (Japan Science and Technology).

REFERENCES

[1] H. Matsuzaki, N. Miki and Y. Ogawa, "3-D FEM analysis of vocal tract models using elliptic tubes with volume radiation," The XIIIth International Congress of Phonetic Science, Stockholm, Sweden, vol.4, 82.12, pp.440-443 (1995).

[2] H. Matsuzaki, N. Miki and Y. Ogawa, "3-D Finite element analysis of Japanese vowels in elliptic sound tube model," Jpn. IEICE Trans(A), vol.J81-A, no.6, pp.899-906 (1998)(in Japanese).

[3] H. Matsuzaki, N. Miki and Y. Ogawa, "FEM analysis of sound wave propagation in the vocal tract with 3-D radiation model," J. Acoust. Soc. Jpn. (E) 17, 3, pp.163-166 (1996).

[4] N. Kamiyama, N. Miki and N. Nagai, "Study of the vocal tract wall impedance using viscoelastic model of the wall," Jpn. IEICE Trans(A), vol.J75-A, no.11, pp.1649-1656 (1992)(in Japanese).

[5] H. Matsuzaki, T. Hirohku, N. Miki and N. Nagai, "Analysis of acoustic characteristics in the vocal tract with inhomogeneous wall impedance using three dimensioinal FEM model," Jpn. IEICE Trans(A), J77-A, No.3, pp.325-333 (1994) (in Japanese).

[6] J.L. Flanagan, "Speech analysis synthesis and perception," 2nd Ed. pp.25-35, Spriger-Verlag (1975).

[7] M. M. Sondhi, "Resonances of a bent vocal tract," J. Acoust. Soc. Am. b79, pp.1113-1116 (1986).