3-D FEM ANALYSIS OF VOCAL TRACT MODELS USING ELLIPTIC TUBES WITH VOLUME RADIATION

Hiroki Matsuzaki, Nobuhiro Miki and Yoshihiko Ogawa
Faculty of Engineering, Hokkaido University

ABSTRACT

Using a 3-D FEM, we compute the sound pressure and the particle velocity in our vocal tract model with a volume radiation in order to estimate the 3-D effect. From the simulation results, we were able to show the differences between 1-D and 3-D models for the formant frequency and band width in the following case: the cross section of the tube is flattened from circle to ellipse, and the shape of constriction is complex.

1 INTRODUCTION

The shape of the vocal tract and the aperture at the lips are important factors in characterizing speech sound . A vocal tract model with cascading elliptic tubes was used to represent the 1-D equivalent circuit model. In this model, however, a plane wave is assumed on the radiational part, and not only the 3-D effect of radiation, but also the 3-D effect of the constriction of the incisors is neglected. We showed that there was a 3-D effect in the elliptic tube by using a 3-D FEM[1]. In this paper, using the FEM, we try to evaluate the 3-D effect of the radiation and the constriction of the incisors. We compute the sound pressure and the particle velocity in our vocal tract model including the volume radiation. The volume radiation is hemisphere shaped. From the experimental results, we see that the 3-D effects are large on the Vocal Tract Transfer Function(VTTF). Finally, we propose a new vocal tract model with cascading non-uniform elliptic tubes. This model is based on MRI data of vocal tracts, and the shape of several cross sections is determined by the elongation factor and the area.

2 FORMULATION OF THE WAVE EQUATION

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

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

where k(ω(angular frequency)/c(sound velocity)) is the wave length constant, and that sound pressure p and particle velocity $\boldsymbol{v}$ are represented as

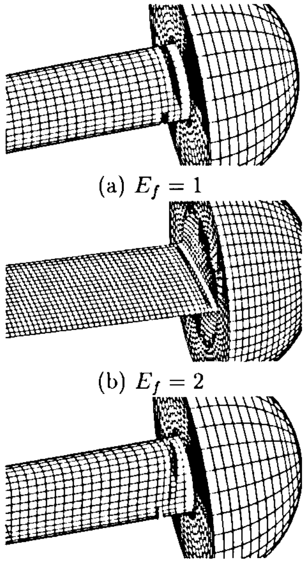
$$p = j\omega\rho\phi \tag{2}$$

$$\boldsymbol{v} = -\nabla\phi \tag{3}$$

where ρ is the atmospheric pressure density. Our FEM formulation was based on the above equations.

3 SIMULATIONS FOR CONSTRICTION

In order to evaluate the effect of the constriction of the incisors in vocal tracts, we computed acoustic characteristics in some straight sound tubes with the constriction. We use a simulation model of the sound tube with a cross sectional area of πcm^2, and a length of $15cm$. The two cross sectional shapes of the tube were determined by a parameter E_f(Elongation Factor)[2]; E_f=1 for a circle and E_f=2 for an ellipse. Its driving surface is driven by sound velocity $v_n = 1e^{j\omega t}$. A volume of radiation, which is hemisphereical in shape, is attached to the aperture surface[3]. The radius is $3cm$ for E_f=1 and $4cm$ for E_f=2[4]. As a boundary condition on the volume, the specific acoustic impedance of spherical radiation is assumed on the spherical surface, and a rigid wall baffle is assumed. The walls making constrictions are shaped by half of the cross section with a thickness of $0.17cm$ and are located in the upper and lower parts of the tube. The lower one is at a distance of $14.25cm$ from the driving surface and the upper one is at a distance of $14.85cm$. In Fig.1, we show the close-up figures of our finite element models with a circular cross section(E_f=1) in (a), and an elliptic cross section(E_f=2) in (b). Moreover, in the case of E_f=1, the constrictions are rounded as to approximate the shape of the incisors(See Fig.1(c)).



(a) $E_f = 1$

(b) $E_f = 2$

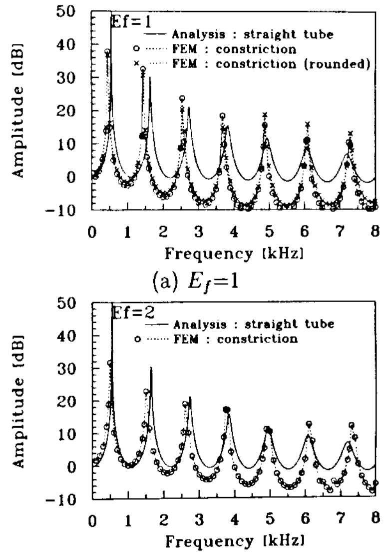
(c) $E_f = 1$: Rounded

Figure 1: *Finite element model*

In Fig.2, we show the VTTF computed from our FEM and the 1-D analytical model without incisors . From the particle velocity, simulated by the FEM analysis, the VTTF $H_v(\omega)$ is computed as

$$H_v(\omega) = 20\log_{10}\left|\frac{\sum v_l(\omega)/A_l}{\sum v_g(\omega)/A_g}\right| \tag{4}$$

where v_g (v_l) are the normal component of particle velocity at the driving (the aperture) surface, and A_g (A_l) is the area of the driving (the aperture) surface. In the results of our FEM



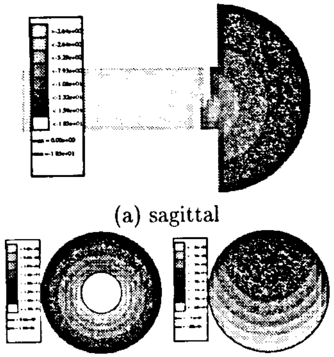
(a) E_f=1

(b) E_f=2

Figure 2: *Vocal tract transfer function*

model, up to the fourth (or fifth) formant, the formant frequencies shift to lower frequencies. This means that the acoustic length of the tube seems longer than the real length. The differences of the formant frequencies between the non-rounded and rounded constriction models are a few Hz. In the case of E_f=2, there are some peaks on the VTTF between $6k$ and $7kHz$;and these peaks are discussed in the following experiment.

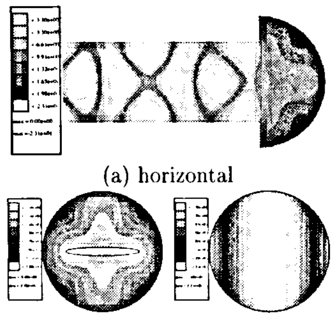
Fig.3 and 4 show the sound pressure distributions([dB]) for E_f=1(frequency is $1450Hz$) and E_f=2(frequency is $6800Hz$). Fig.3(a) is on the sagittal plane and Fig.4(a) is on the horizontal plane. (b) of both models is on the baffle of the volume, and (c) on the spherical surface. In Fig.3(a), we see that the sound wave propagates along



(a) sagittal

(b) baffle (c) surface

Figure 3: *Sound pressure distribution* ($Ef = 1, 1450Hz$)



(a) horizontal

(b) baffle (c) surface

Figure 4: *Sound pressure distribution* ($Ef = 2, 6800Hz$)

the acoustic pass of the constriction. If the analytical model is used with the equivalent length and area of the tube estimated from the FEM results, the VTTF may be in agreement with the FEM results. In Fig.3(b) and (c), we see that the sound wave propagates non-symmetrically with respect to the upper and lower direction because of the alternate constriction. In Fig.4(a), we see that a higher mode is formed. In Fig.4(b), the sound wave does not always propagate along the aperture. In Fig.4(c), the sound wave is distributed vertically, therefore we guess that the sound wave propagates in a vertically polarized wave through free space.

4 SIMULATION OF A NEW VOCAL TRACT MODEL

The traditional vocal tract model is modeled by cascading uniform circular tubes. In this model, the tubes are not connected smoothly, although the outline of the human vocal tract is smooth. The question then arises about the non-smooth connection. In this simulation model, we connect the elliptic tubes smoothly. Our model is based on MRI data of the vocal tract for the Japanese vowel /a/[5], and the shape of several cross sections is determined by E_f and the area. In Fig.5, we show a finite element model including a volume radiation (radius of $4cm$).

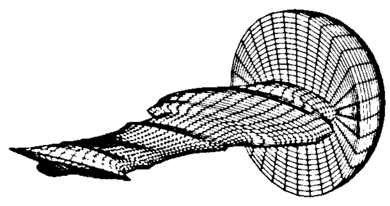


Figure 5: *Finite element model*

In Fig.6, we show the VTTF computed from our FEM and the traditional 1-D analytical model. The first and second formant frequencies of the FEM solution agree relatively with analytical solutions, but the third formant of FEM is about $100Hz$ lower than the analytical one. In the higher formant frequencies, there are difference in the formant frequency or band width. And there are two valleys(zeros) between $4.5k$ and $5kHz$ and between $6.8k$ and $7.6kHz$.

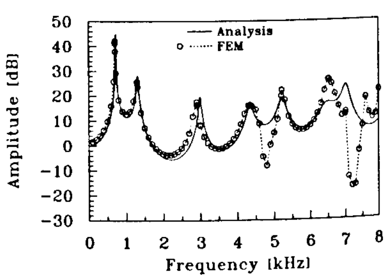
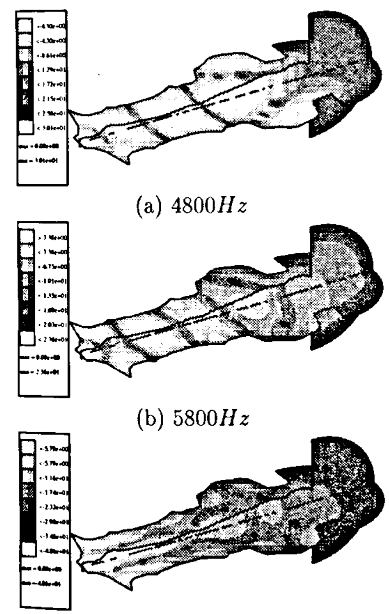


Figure 6: *Vocal tract transfer function*

In Fig.7, we show the sound pressure distributions($[dB]$) on the sagittal and the horizontal plane. The driving frequencies of $4.8k$ and $7.2kHz$ correspond to the two valleys(zeros), and $5.8kHz$ to the non-zero. In the large cavity corresponding to the oral cavity, the effects of the higher mode in Fig.7(a) and (c) are larger than in (b).



(a) $4800Hz$

(b) $5800Hz$

(c) $7200Hz$

Figure 7: *sound pressure distribution*

CONCLUSION

Using our 3-D FEM models, we simulated the sound wave propagation in a vocal tract model with the constriction of incisors and in our model for the Japanese vowel /a/. These models included the volume radiation. The VTTF was computed by simulated result. And the sound pressure distributions were shown.

The results of the experiment showed 3-D effects on the formant frequencies because of the 3-D shape of the vocal tract with constriction and the existence of higher modes. These facts never appear in the traditional 1-D analysis. We consider that these facts are useful for natural speech analysis.

ACKNOWLEDGEMENT

We are grateful to Dr. Naohisa Kamiyama for his offer of valuable data and Dr. Kunitoshi Motoki for valuable advice.

REFERENCES

[1] Matsuzaki,H., et al. (1994), "Analysis of acoustic characteristics in the vocal tract with inhomogeneous wall impedance using three dimensional FEM model", Electron.& Commun. Jpn.

[2] Kamiyama,N., et al. (1991), "A study on the effect of the viscoelastic vocal tract wall", Proc. Fall Meet. Acoust. Soc. Jpn., 1-6-12, pp.213-214(in Japanese).

[3] Matsuzaki,H., et al. (1994), "3D FEM analysis of vocal tract model of elliptic tube with inhomogeneous-wall impedance", ICSLP94, S12-17,Vol.2, pp635-638.

[4] Motoki,K. & Miki,N.(1994) "Distribution characteristics of acoustic impedance density around the radiation area", JASA, 1-8-8, pp.643-644(in Japanese).

[5] Kamiyama,N., et al. (1992),"Study of the vocal tract impedance using viscoelastic model of the wall", Jpn. IEICE Trans(A), **J75-A**, 11, pp,1649-1656(in Japanese).