3rd International Conference on
Spoken Language Processing (ICSLP 94)
Yokohama, Japan
September 18-22, 1994

Measurement of 3D shapes of Vocal Tract, Dental Crown and Nasal Cavity Using MRI : Vowels and Fricatives.

Masafumi MATSUMURA (1), Takuya NIIKAWA (1),
Koji SHIMIZU (2), Yasuji HASHIMOTO (2) and
Tatsuya MORITA (3)

(1) Department of Applied Electronics, Faculty of Engineering, Osaka Electro-Communication University, 18-8, HATSU-CHO, NEYAGAWA 572, JAPAN.
(2) Medical Systems Division, SHIMADZU Corporation, 1, Nishinokyo-Kuwabaracho, Nakagyo-ku, KYOTO 604, JAPAN.
(3) Department of Electrical Engineering, Faculty of Engineering, Osaka University, Yamada-Oka 2-1, SUITA 565, JAPAN.

ABSTRACT

This paper deals with the measurement of three dimensional shapes of vocal tract, dental crown and nasal cavity using magnetic resonance imaging (MRI). 3D MR images of the vocal tract and nasal cavity of 3 adult males were obtained in the steady state productions of Japanese vowels and fricatives. MR images of the dental crown that contain a small amount of water were obtained using a dental crown plate that is shaped so as to tightly attach to the subject's dental crown by thermoforming and that contain a contrast medium for MR imaging. The measurement time was 186 second for 28 transverse sections from the larynx to the nasal cavity at a 6 mm interval, and was 186 second for 32 coronal sections from the tip of the nose to the atlas at a 4 mm interval. A computer algorithm for air-tissue boundary tracing of arbitrary sections of the vocal tract and nasal cavity from the 3D MR image was proposed. The algorithm is based on the gray level interpolation and the threshold operation. Images of arbitrary sections of the vocal tract and nasal cavity from 3D MR image were reconstructed using the gray level interpolation based on the sampling theorem. Thresholds of the gray level were computed from gray levels at air-tissue border points. Shapes of arbitrary sections of the vocal tract and nasal cavity were obtained using the present algorithm. The boundary tracing error of the dental crown from the 3D MR image was less than 5.7%. The 3D shape and area function of vocal tract during the productions of fricatives /s/ were obtained. The 3D shapes of sinus frontails, sinus spenoidalis and sinus maxillaris were observed.

1. INTRODUCTION

Several techniques for investigating human speech productions have reported in speech literatures focusing on the vocal tract as the primary articulatory organ. Observations of three-dimensional shapes of the vocal tract and nasal cavity are urgently required for minutely understanding of the articulatory and acoustical processes in speech productions. It is difficult to observe 3D shapes of the vocal tract and nasal cavity using conventional techniques such as, X-ray technique [1], photoelectric device [2] and ultrasonic imaging [3].

Magnetic resonance (MR) imaging capable of noninvasely obtaining transverse, coronal and sagittal sections of the body is widely used for medical examinations. Rokkaku et al. applied the MR technique to measure the three-dimensional vocal tract shape [4]. Present the MR system has been performed high-quality imaging by using the superconductive magnet that develops a higher magnetic flux density, and high-speed imaging by advanced image processing [5]. Analysis and modeling of the vocal tract shape and dimensions using the MRI have been studied [6,7,8]. However, it is difficult to observe dental crown profiles that contain a small amount of water using MRI. For this disadvantage, there is not enough of the data on 3D shapes of the vocal tract during the productions of consonants that are produced by the teeth in steady state. Recently, the volume occupied by the teeth was modified using MR images of the dental impressions [9]. However, no data on 3D shapes of the vocal tract during the productions of consonants such as fricatives are obtained. And also, in traditional approach, there is no report that shapes of arbitrary sections of the vocal tract and nasal cavity were estimated from the 3D MR image. Few new data on 3D shapes of the vocal tract and nasal cavity have been added for analysis of acoustical coupling of the vocal tract and nasal cavity at the velum in recent years. Now, the MR technique for investigating human speech productions requires for imaging the dental crown shape and for estimating the shapes of arbitrary sections of the vocal tract and nasal cavity from the 3D MR image.

The purpose of this paper is to obtain 3D shapes of the vocal tract, dental crown and nasal cavity during the productions of Japanese vowels and fricatives /s/ using MRI. A method that obtain simultaneously MR images of the dental crown and vocal tract using a dental crown plate was developed. A computer algorithm for air-tissue boundary tracing of cross sections of the vocal tract and nasal cavity from the 3D MR image was proposed. The algorithm consist of a image reconstruction from the 3D MR image and a boundary tracing of cross sections of the vocal tract and nasal cavity. The validity of the present algorithm is confirmed by the comparison the profile of the dental crown obtained by the algorithm with dimensions of dental crown impressions.

In the following section, we will describe the measurement method of 3D MR images and profiles of the dental crown. The computer algorithm for boundary tracing of cross sections of the vocal tract and nasal cavity will be proposed. In section 3, the shapes of the vocal tract, dental crown and nasal cavity will be showed.

2. METHODS

2.1 Measurement of MR images

All MR images in this study were collected using a 1.0 Tesla superconductive MR system (MAGNEX100HP, Shimadzu Corp., Japan). MR images of mid-sagittal sections were measured by single-slice flip-back-spin-echo imaging methods that was performed a higher speed imaging. Each image was acquired with the repetition time TR=120ms, the echo time TE=25ms, using image matrix of 256×256 over a field of view of 25cm. A section thickness of each excited plane was 10 mm. Measurement time was 17.2 second for each image. Transverse and coronal MR images were measured by multi-slice T1-weighted spin echo imaging methods. Each image was acquired with TR=1300ms, TE=20ms, using image matrix of 256 ×256 over a field of view of 25cm.

3D MR images that consist of MR images of 28 transverse sections from the larynx to the nasal cavity at a 6 mm interval and 32 coronal sections from the tip of nose to the atlas

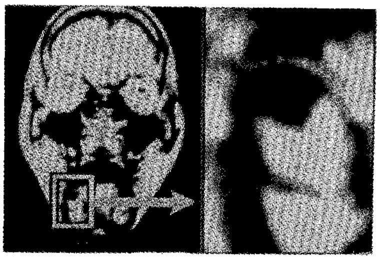

Fig.1. A general view (left) and expanded view (right) of MR Image.

at a 4mm interval were obtained. Measurement time was 186 second for each 3D MR image. The section thickness of each excited plane for the transverse and coronal MR images were 5 and 3.5 mm respectively. MR images and the vowel output uttered by a supine subject in the MR system were obtained simultaneously.

Data on dental crown shapes are required for analysis of fricative productions and for precisely estimating acoustical characteristics of the vocal tract. MR images of the dental crown that contain a small amount of water were obtained using a dental crown plate that is shaped so as to tightly attach to the subject's dental crown by thermoforming and that contain a contrast medium for MR imaging. A thickness of the dental crown plate is 0.6mm. By attaching the dental crown plate to subject's upper and lower teeth, profiles of the vocal tract and dental crown can be obtained simultaneously using MRI.

MR images of mid-sagittal, transverse and coronal sections of the vocal tract and nasal cavity were obtained during steady-state productions of Japanese vowels and fricatives. Experiments were performed on 3 Japanese adult male subjects who are no history of speech disorder. The subjects' voice were recorded using a high sensitivity condenser microphone placed inside the magnet 20 cm from the subjects' lips. Steady state vowel and fricative /s/ utterances were sustained for about several second and the intermittent utterances were sustained successively for about 3 minutes.

Figure 1 shows an example of coronal MR image. A profile of lower dental crown was clearly obtained using the dental crown plate.

2.2 Image reconstruction of cross sections from 3D image

A posterior-anterior axis is defined as x-axis and a superior-inferior axis as z-axis. A perpendicular axis to the x-z plane is defined as y-axis. The transverse section, coronal section and sagittal section are represented by x-y plane, y-z plane and x-z plane respectively. Images of arbitrary sections that are

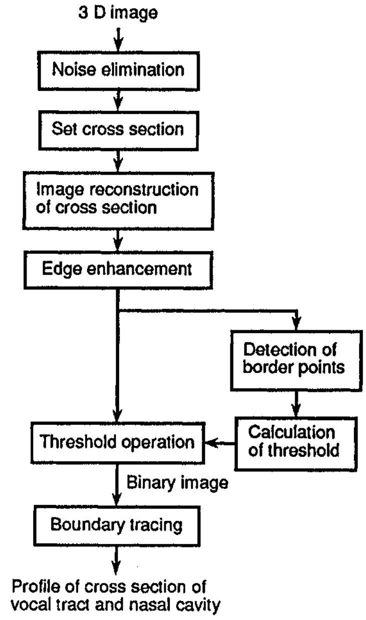

Fig.3. Processing flow.

perpendicular to the sagittal section (x-z plane) were estimated from the 3D MR image using the gray level interpolation which correct geometric distortions. A gray level f(x,z) at arbitrary point(x,z) in x-z plane was estimated from gray levels f(xk,zk) at lattice points (xk,zk) around the point(x,z) using the following equation.

$$f(x,z) = \sum_k \sum_l f(x_k, z_l) C(x_k - x) C(z_l - z) \qquad (1)$$

$$where\ C(x) = \frac{\sin(\pi x)}{\pi x}$$

The equation (1) is based on the sampling theorem in which original signals reconstruct from it's sampled signals. C(x) is the sampling function.

Figure 2(a) shows a MR image of midsagittal section estimated from the 3D MR image using the present algorithm. Figure 2(b) shows a MR image of a cross section at a slanting solid line in Figure 2(a). Profiles of the upper and lower teeth were observed in Figure 2(b).

2.3 Air-tissue boundary tracing

Figure 3 shows a computer algorithm for air-tissue boundary tracing of the vocal tract and nasal cavity. The algorithm is composed of three operations, a noise elimination, an edge enhancement and a threshold operations. The elimination

(a) *(b)*

Fig.2. Image of cross sections reconstructed from 3D MR Image.

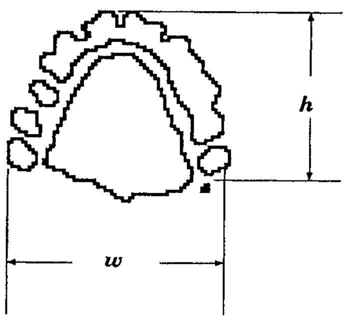

Fig.4. A dental crown shape of transverse section.

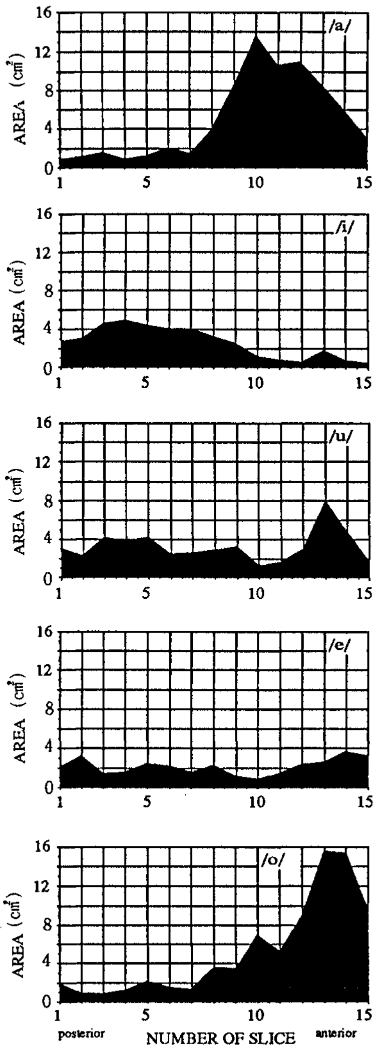

Fig.5. Vocal tract area functions.

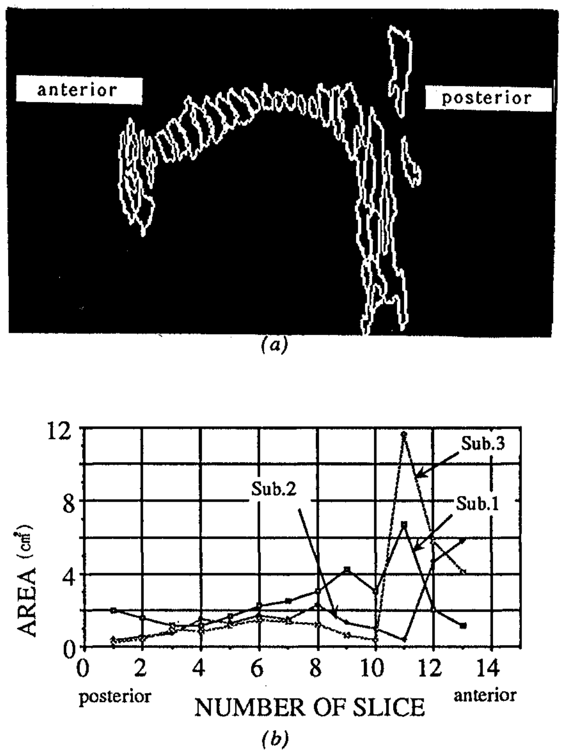

Fig.6. 3D shape (a) and area functions (b) of oral cavity during fricatives /s/ productions.

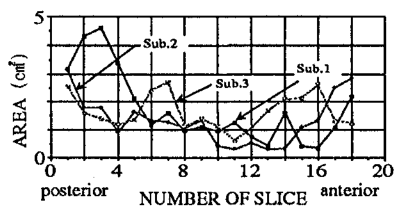

Fig.7. Area functions of nasal meatus.

of impulse noise was performed with the median filtering. The edge enhancement at air-tissue boundary was performed using the Prewitt's template matching which the direction of the edge is detected by the mask pattern. The air-tissue boundary was obtained using the threshold operation which a threshold is average value of the gray level at the air-tissue border points. The artifacts that tended to blur air-tissue boundaries were easily identified. Tracing errors generated by the algorithm were corrected by hand tracing.

3: RESULTS

3.1 Dental crown

Figure 4 shows a dental crown shape of transverse section obtained using the present algorithm. Distance w between both the first molar teeth and h between the first molar tooth and the incisor were measured from the dental crown shape. Measurement error of w and h obtained by the comparison with dimensions of subject's dental impressions were 4.7% and 5.7% respectively. Because the dental crown plate is thin in thickness and tightly attached to the upper and lower teeth, there was approximately no effect of the dental crown plate in place on the speech production.

Result shows that the measurement method of the upper and lower teeth profiles using the dental crown plate is useful and the validity of the boundary tracing algorithm is confirmed.

3.2 Vocal tract

Area functions were estimated from the 3D shapes of the vocal tract. The area functions shown in Figure 5 show approximately the expected patterns: narrow pharynx and wide oral cavity for /a/ and /o/, wide pharynx and narrow oral cavity for /i/, and wide pharynx and oral cavity with constrictions in the velar regions for /u/. A vocal tract length was 17.4 cm. The vocal tract shapes of cross sections that is perpendicular to the mid-sagittal plane were non-symmetric with respect to the mid-sagittal line. The formant frequencies of the vowels computed from the area functions without sinuses showed good agreement with the subject's original productions.

Figure 6 shows a 3D shape (a) and area functions (b) of the vocal tract during the productions of Japanese fricatives /s/ . The 3D shape shows a narrow oral cavity that are formed by the upper and lower teeth and the frontal tongue body for the fricative productions. The are functions show broadly the expected patterns : narrow oral cavity and wide radiation mouth

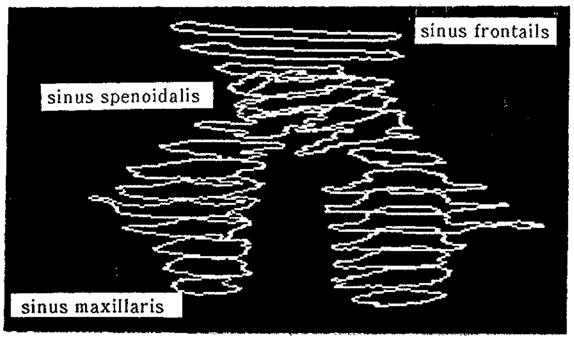

Fig.8. A 3D shape of paranasal sinus.

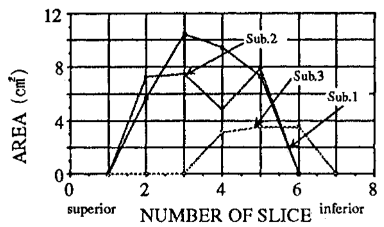

Fig.9. Sinus frontails.

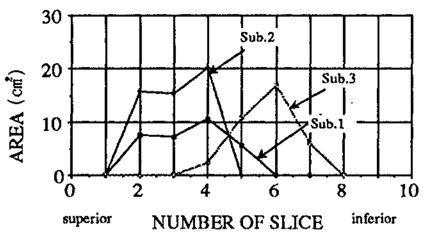

Fig.10. Sinus spenoidalis.

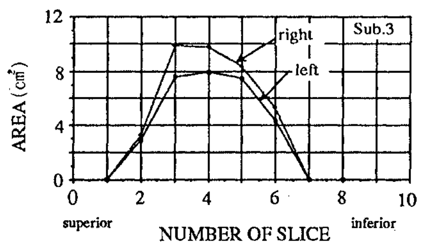

Fig.11. Sinus maxillaris.

with constrictions at the tongue tip (slice number 10 for sub.1 and sub.3, 11 for sub.2). Across subjects, the difference occurs for areas of the oral cavity and the radiation mouth, for which the sub.3 oral cavity is generally narrower and the sub.3 area at the radiation mouth is smaller than that for other subjects.

3.3 Nasal cavity

Nasal cavity shapes of coronal sections from the atlas to the nose tip were investigated. Profiles of the nasopharyngeal meatus, superior nasal meatus, middle nasal meatus, and inferior nasal meatus were obtained [10]. Results show that the nasal meatus shapes of coronal sections are non-symmetric with respect to the mid-sagittal line and vary by the conceivable reasons such as a cold.

Figure 7 shows area functions of the nasal meatus. The area functions for all subjects show many peaks and dips. Figure 8 shows a 3D shape of the paranasal sinus consist of the sinus-frontails, sinus-maxillaris, and sinus-spenoidalis. The sinus-ethmoidalis shape was not obtained.

Figure 9-11 show area functions of the paranasal sinus that is represented relations between positions on the superior-inferior axia and areas. Across sinuses, the area functions show the patterns : large area for the sinus spenoidalis, wide sinus maxillaris, and non-symmetric shapes of the sinus maxillaris with respect to the mid-sagittal line. Across subjects, the greatest difference occurs for the sinus frontails, for which a volume computed from the area function for sub.1 is 14.1cm3 and 6.04 cm3 for sub.3. The area function for sub.1 is narrower than that for other subjects. The validity of the results were confirmed by the comparison the profiles of the vocal tract and nasal cavity with the anatomical data and findings.

4. SUMMARY

This study was performed to measure 3D shapes of the vocal tract, dental crown and nasal cavity using MRI. A method that obtain simultaneously MR images of the dental crown and vocal tract using a dental crown plate was developed. A computer algorithm for air-tissue boundary tracing of cross sections of the vocal tract and nasal cavity from the 3D MR image was proposed. The validity of the present algorithm was confirmed by the comparison the profile of the dental crown obtained by the algorithm with dimensions of den ` crown impressions. The 3D shape and area function of vocal tract during the productions of fricatives /s/ were obtained. The 3D shapes of the paranasal sinus were investigated.

Acknowledgment. We give special acknowledgment to Takeshi Wada and Takashi Tachimura of Division for Oral-Facial Disorders, Faculty of Dentistry, Osaka University for the dental crown plate manufacturing.

REFERENCES

[1] S. Kiritani, K. Itoh, and O. Fujimura, "Tongue-pellet tracking by a computer-controlled x-ray microbeam system, "JASA, Vol.57, No.6, pp.1516-1520 (1975).
[2] C-K. Chuang and W.S-Y. Wang, "Use of optical distance sensing to track tongue motion, "J. Speech and Hearing, Vol.21, pp.484-496 (1978).
[3] F.D.Minifle, C.A.Kelsey and J.A.Zagzebski:"Ultrasonic scans of the dorsal surface of the tongue",JASA, Vol.49, No.6, pp.1857-1860 (1971).
[4] M. Rokkaku, K. Hashimoto, S. Imaizumi, S. Niimi, and S. Kiritani, "Measurements of the three-dimensional shape of the vocal tract based on the magnetic resonance imaging technique, "Ann. Bull. Res. Inst. Logopedics and Phoniatrics 20, pp.47-54 (1986).
[5] K. Shimizu et al. , "Visualization of moving fluid: quantitative analysis of blood flow velocity using MR imaging, "Radiology, Vol.159, No.1, pp.195-199 (1986).
[6] T. Baer, J.C. Gore, L.C. Gracco, and P.W. Nye, "Analysis of vocal tract shape and dimensions using magnetic resonance imaging: vowels, "JASA, Vol.90, No.2, pp.799-828 (1991).
[7] A.K. Foldvik, O. Husby, J. Karness, I.C, Nordli, and P.A. Rinck, "MRI (magnetic resonance imaging) film of articulatory movements. "ICSLP90,11.1, pp.421-422 (1990).
[8] M. Matsumura and A. Sugiura, "Modeling of three-dimensional vocal tract shapes obtained by magnetic resonance imaging for speech synthesis, "ICSLP90,11.2, pp.425-428 (1990).
[9] C.Yang and H.Kasuya:"Accurate measurement of vocal tract shapes from magnetic resonance imaging of child, female and male subjects", Technical Report of IEICE, SP94-6 (in Japanese) (1994).
[10] M.Matsumura:"Measurement of three-dimensional shapes of vocal tract and nasal cavity using magnetic resonance imaging technique",ICSLP92, Fr.fAM.1.3, pp.779-782 (1992).