ACOUSTICS 2012
HONG KONG

Comparison of transfer functions between the actual pinna and the simple pinna model which is composed of a rectangular plate and three rectangular cavities

Yohji Ishii[1], *Hironori Takemoto*[2], *and Kazuhiro Iida*[1]

[1] *Chiba Institute of Technology,*

[2] *National Institute of Information and Communications Technology*

Abstract: The authors have been studying a method to provide the appropriate HRTFs to the listener using simple pinna models, whose size and proportion are obtained by the picture of listener's pinnae. In this study, following two issues are examined to achieve the estimation process of the subject’s lowest two spectral notches (N1 and N2) of the HRTFs, which are the spectral cues for vertical localization, from the anthropometric data of subject’s pinnae using the three-step models, which are composed of successive three rectangular cavities. At first, the fundamental relation between the size and proportion of the three-step model and the frequencies of N1 and N2 was obtained in the median plane, by expanding its width, length, and depth. The results show that; 1) the frequencies of N1 and N2 become higher as the model becomes smaller, 2) the width of the model has a dominant influence on the frequencies of N1 and N2, 3) for 120°, which is the direction parallel to the long axis of the model, the length of the model affects the frequencies of N1 and N2 a lot.Then, in order to examine the accuracy of N1 and N2 obtained by the three-step model, the transfer functions of the models, whose cavities were fitted to the actual pinnae, were measured in the median plane. The results show that the frequencies of N1 and N2 of the model were close to those of HRTF in some models.
Keywords: Head-Related Transfer Function, Individualization of HRTF, Spectral cue, three-step model, 3D sound image control

1. Introduction

It is well known that 3D sound image localization is accomplished when the listener’s own head-related transfer functions (HRTFs) are reproduced at the listener’s eardrums [1]. However, the HRTFs of other listeners often cause degradation of localization accuracy. Therefore, the HRTFs which conform to the listener are necessary for the accurate 3D sound image localization.

Blauert showed a scenario on the method to provide the appropriate HRTFs to the listener as follows: *“A Person who enters a multimedia shop is scanned by a camera and some instants later his/her individual HRTF set is ready to be sold for the use in advanced 3D applications”* [2]. However, this scenario has not been realized since it is hard to estimate the complete information of HRTFs of the listener from the simple anthropometric data of the listener’s pinnae captured by a camera. Then, the authors try to estimate the cues for localization in the listener’s HRTFs.

It is well known that interaural differences are the cue for the lateral localization. On the other hand, the cues for vertical and front-back localization are two spectral notches above 4 kHz. Iida *et al*. [3] proposed a parametric HRTF model to clarify the contribution of each spectral peak and notch as a spectral cue for vertical localization. The parametric HRTF is recomposed using all or some of the

[1] kazuhiro.iida@it-chiba.ac.jp

spectral peaks and notches above 4 kHz extracted from the measured HRTF, which are expressed parametrically by their frequency, level, and sharpness (Fig.1). They showed that sound localization on the median plane can be archived with the first second notches (N1 and N2) and the first peak (P1) (Fig.2), and that the frequencies of N1 and N2 depend strongly on the source elevation (Fig.3).

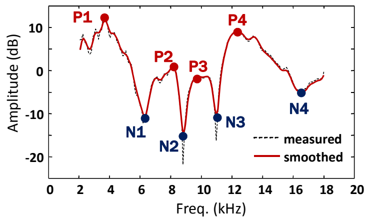

Figure1 – An example of extracted spectral peaks and notches from a measured HRTF

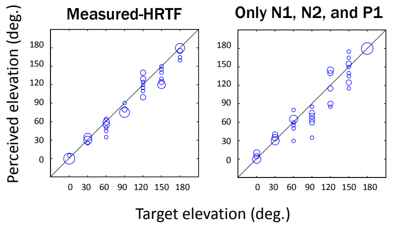

Figure 2 – Localization in the median plane by measured HRTF and parametric HRTF (N1, N2, P1) [3]

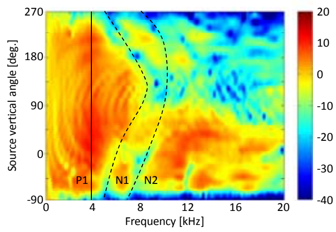

Figure 3 – Distribution of frequencies of N1, N2, and P1 in the median plane

Then, how do we obtain the listener's information on the interaural differences, N1, and N2? The interaural differences can be estimated easily by the distance between left and right ears. However, N1 and N2 information is hard to estimate because the relation between these spectral notches and the anthropometric data of the listener's pinnae is unknown. One of the approaches to clarify the relation is to obtain the relation between the frequencies of N1 and N2 and the dimension of a simple model of pinnae (Fig.4).

On the pinna model, Genuit proposed a simple geometrical model of pinna consisting of three pure cylinders with different diameters and length [2]. Takemoto *et al.* [4] proposed simple rectangular plate model with successive three rectangular cavities (three-step model). They reported that the frequencies of N1 and N2 of three step model showed the similar incident angle dependency to those of HRTFs of the actual ear (Fig.5) [4]. However, it is not clear whether the three-step model can express the listener's N1 and N2 frequencies or not.

In other words, the appropriate HRTFs can be provided by the individualization process shown in Fig.4, if the relation between N1 and N2 frequencies and the dimension of the three-step model is clarified

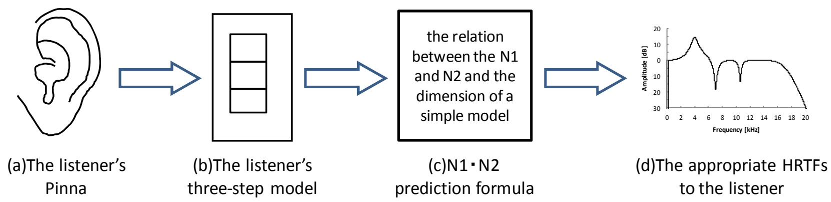

Figure 4 – Proposed individualization process of HRTF using simple pinna model

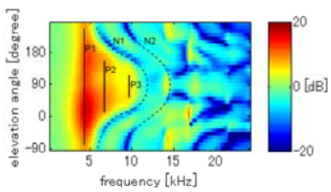

Figure 5 – Transfer functions of three-step model

In this study, following two issues are examined to achieve the individualization process;

(1) obtain the fundamental relation between the size and proportion of the three-step model and the frequencies of N1 and N2,

(2) confirm the accuracy of frequencies of N1 and N2 of the subject estimated using the subject's three-step model.

2. Influence of expansion of the pinna model on the N1 and N2 frequencies

In order to clarify the relation between the size and the proportion of the three-step model and frequencies of N1 and N2, influence of expansion of the model on the N1 and N2 frequencies is examined.

2.1 Pinna model

The actual pinna and its three-step model are shown in Fig.6. The sizes of x1 to x12 were measured for the both ears of 111 subjects. The mean values of x1 to x12 were used for the reference model (Table 1). The width, length, and depth of the reference model were expanded in 4 patterns shown in Table 2. In pattern 1, the width, length, and depth were expanded all together. In patterns 2 to 4, one of the width, length, and depth was fixed and others were expanded concurrently. The expansion ratio was 0.70 to 1.45 in 0.15 steps. Consequently, 24 three-step models (6 expansion ratio x 4 patterns) were created. The models were made of silicon.

The lengths of 3 cavities (x6, x7, and x8) were 1/3 of x9. x10 is the depth of the bottom cavity. The depth of the middle and upper cavities were 2/3 and 1/3 of x10, respectively. The location of a microphone, which corresponds to the entrance of the ear canal, was 4mm and 7mm from the left and bottom walls of the bottom cavity.

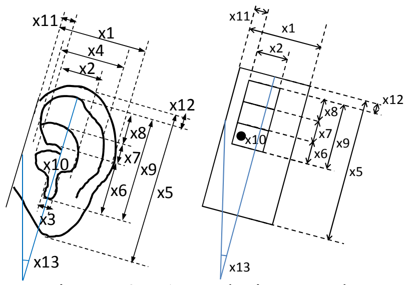

Figure 6 – Actual pinna and three-step model

Table 1 – The sizes of reference model

parts	sizes [mm]
x1	34
x2	19
x5	66
x6	15
x7	15
x8	15
x9	45
x10	13
x11	6
x12	5

Table 2 – The expansion pattern of the pinna model (○: expanded part)

	width (x1,x2)	length (x5,x9)	depth (x10)
pattern 1	○	○	○
pattern 2	○	○	
pattern 3		○	○
pattern 4	○		○

2.2 Measurement method of transfer functions of the pinna models

The transfer functions of the pinna models were measured in an anechoic chamber. The sound source was a swept-sine signal. The sampling frequency was 48 kHz. The directions of the sound source were seven elevations from 0° (front) to 180° (rear) in 30° steps in the upper sagittal plane. The distance from the sound sources to the entrance of ear canal was 1.2 m. The pinna model was tilted at 5° to the sagittal plane [4], and 25° to the transverse plane.

2.3 Results of measurements

Figure 7 shows the transfer functions of the front direction of pattern 1, as an example of measurement results. The frequencies of peaks and notches become higher as the expansion ratio becomes smaller.

Figure 8 shows the relation between the frequencies of N1 and N2 and the sound source elevation. The frequencies of N1 and N2 become higher as the size of the model becomes smaller for pattern1 (Fig.8 (a), (b)). However, N1 and N2 were not appeared at the above directions of the expansion ratios of 0.85 and 0.70. For pattern 2 (Fig. 8 (c), (d)), the frequencies of N1 and N2 become higher as the size of the model becomes smaller, as well as pattern 1. For pattern 3 (Fig. 8 (e), (f)), the frequency changes of N1 and N2 are smaller than those for patterns 1 and 2. For pattern 4(Fig.8 (g), (h)), the frequencies of N1 and N2 changes very little at the elevations of around 120°. The elevation of 120° lies in the direction of the long axis of the pinna model, as the model was tilted at 25° to the transverse plane. Since the width and the depth of the model don't affect the frequencies of N1 and N2, it can be considered that N1 and N2 frequencies depend on the length of the model for the elevation of 120°. However, for the other elevations, N1 and N2 frequencies become higher as the size of the model becomes smaller, as well as pattern1.

The comparison of the frequency of N1 and N2 among patterns leads following discussions. It can be considered that the depth of the model affects the frequencies of N1 and N2 very little because the frequencies of N1 and N2 of pattern 1 are similar to those of pattern 2. Pattern 3 has a similar tendency to that of pattern1, however, the change of frequencies of pattern 3 is smaller than that of pattern 1. This indicates that the width of the pinna model affects the frequencies of N1 and N2 a lot. Pattern 4 also has a similar tendency to that of pattern1 except for 120°. For 120°, the frequencies of N1 and N2 are not affected by expansion ratio. As mentioned above, the depth of the pinna model has very little effects on the frequencies of N1 and N2. Therefore, the width of the pinna model has a dominant influence on the frequencies of N1 and N2 except for 120°, while the length of the pinna model affects the frequencies of N1 and N2 a lot for 120°.

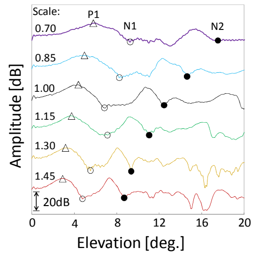

Figure 7 – Transfer functions of front direction for pattern 1.

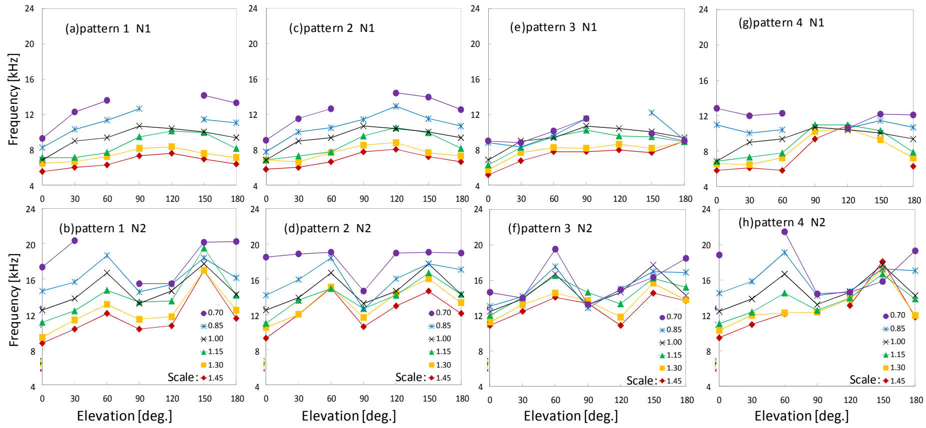

Figure 8 – Frequencies of N1 and N2 for patterns 1 to 4.

3. Comparisons of transfer functions between pinna models and actual pinnae

In order to examine the accuracy of N1 and N2 obtained by the pinna model, the transfer functions of the three-step models, which were created using the anthropometric data of the listener's actual pinnae, were measured.

3. 1 Conditions of pinna models

Following four kinds of cavity shape (Fig.9) were used to model the actual pinna (subject A), considering the importance of the width of the pinna model as mentioned above.

Nineteen models (A-1 to D-4) were created by silicon as parameters of the length of each cavity (x6 to x8), the depth of the cavities, and the location of the entrance of the ear canal (Table 3). x9 (the length of the cavities), x1 (the width of the outer shape of the pinna), and x5 (the length of the outer shape of the pinna) were fitted to those of the actual pinna.

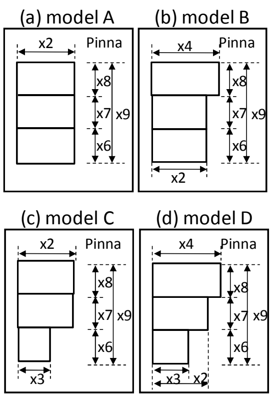

Figure 9 – Shape of cavities

Table 3 – Conditions of pinna models

model	length of cavities	depth of cavities	location of the entrance of the ear canal
A-1	actual pinna	actual pinna	actual pinna
A-2	equal	actual pinna	actual pinna
A-3	equal	actual pinna	4, 7 [mm]
A-4	equal	7, 11, 15 [mm]	4, 7 [mm]
A-5	equal	7, 11, 15 [mm]	actual pinna
A-6	actual pinna	actual pinna	4, 7 [mm]
B-1	actual pinna	actual pinna	actual pinna
B-2	equal	actual pinna	actual pinna
B-3	equal	7, 11, 15 [mm]	actual pinna
B-4	actual pinna	actual pinna	4, 7 [mm]
C-1	actual pinna	actual pinna	actual pinna
C-2	equal	actual pinna	actual pinna
C-3	equal	7, 11, 15 [mm]	actual pinna
C-4	actual pinna	actual pinna	4, 7 [mm]
C-5	equal	actual pinna	4, 7 [mm]
D-1	actual pinna	actual pinna	actual pinna
D-2	equal	actual pinna	actual pinna
D-3	actual pinna	7, 11, 15 [mm]	actual pinna
D-4	equal	7, 11, 15 [mm]	actual pinna

3.2 Results of measurements

As an example of the results, Fig. 10 shows the HRTF and the transfer functions of the pinna models for front direction. P1, N1, and N2 of the HRTF are shown around 4 kHz, 6 kHz, and 8.5 kHz.

For the pinna models, P1 appears around 4 kHz, exactly same as HRTF. The models whose N1 and N2 appear at the close frequency to the HRTF were A1 to A6, B-3, and B-4. However, N1 frequencies of the pinna models tend to be lower than HRTFs.

Next, the difference of the actual pinna and the pinna models in the frequencies of N1 and N2 was 6evaluated quantitatively by NFD (Notch Frequency distance) [5]. NFD for $HRTF_j$ and $HRTF_k$ is defined by equations (1) to (3),

$$NFD_1 = \log_2\left\{f_{N1}(HRTF_j) / f_{N1}(HRTF_k)\right\} \quad \text{[oct.]} \qquad (1)$$

$$NFD_2 = \log_2\left\{f_{N2}(HRTF_j) / f_{N2}(HRTF_k)\right\} \quad \text{[oct.]} \qquad (2)$$

$$NFD = |NFD_1| + |NFD_2| \quad \text{[oct.]} \qquad (3)$$

where, f_{N1} and f_{N2} are N1 and N2 frequencies, respectively.

Table 4 shows the elevations where NFDs between the HRTF and the pinna model are less than just noticeable difference (0.2 oct.) [5]. The pinna model which has the most number of elevations of NFD of less than 0.2 oct. was B-4 (6 of 7 elevations), and the next were A-3, A-4, and A-6 (5 of 7 elevations).

N1 and N2 frequencies of any model A are not close to those of HRTF for the elevation of 0 and 30°. On the other hand, N1 and N2 of the model C-1 and C3 appear near those of HRTF.

Elevation of 30° corresponds to the direction parallel to the short axis of the pinna model because the pinna model was tilted at 25° to the transverse plane. This indicates that the width of the lower cavity acts an important role for this incident angle because the difference of shape between model A and model C is the width of the lower cavity.

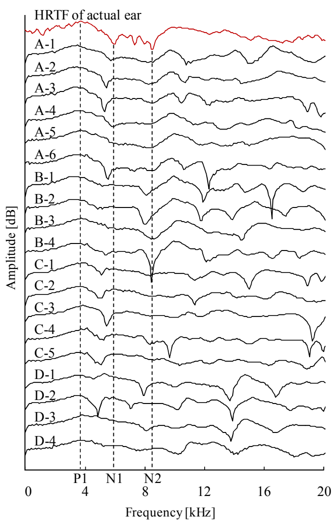

Figure 10 – HRTF and transfer functions of pinna models for front direction

Table 4 – Comparison between HRTF and transfer functions of pinna model (○ : NFD＜0.2oct.)

model	Sound source direction (deg.)						
	0	30	60	90	120	150	180
A-1	○						○
A-2	○			○			○
A-3	○			○	○	○	○
A-4	○			○	○	○	○
A-5	○				○		○
A-6	○			○	○	○	○
B-1		○		○	○	○	○
B-2		○					○
B-3	○						○
B-4	○	○		○	○	○	○
C-1		○	○			○	
C-2		○		○		○	
C-3		○	○		○	○	
C-4	○	○		○			
C-5					○	○	○
D-1		○					
D-2		○				○	
D-3							
D-4		○		○	○	○	

3. 4 Examination of the validity of the pinna model using several subjects

Besides subject A, another four subjects (B to E) were participated to examine the validity of the pinna models A-3, A-6, and B-4. However, the pinna models for subjects B, C, and D were essentially two-step models because the depth of the lower and the middle cavities were exactly same for these subjects.

As an example of the results, Fig. 11 shows the HRTF and the transfer functions of the pinna models A-6 for the front (0°) and rear (180°) direction.

For the front direction (Fig.11 (a)), the spectrum shapes of model A-6 are generally similar to that of HRTF for all the subjects. The frequencies of P1 of the model are close to that of HRTF. The frequencies of N1 are a little lower than that of HRTF. The results of chapter 2 infer that the size of the cavities of the model might be larger than actual cavities because circumscribed rectangular solids were adopted to model the cavities. The frequencies of N2 of the model are almost close to that of HRTF. However, they are a little lower (subject Band D) or a little higher (subject C).

For the rear direction (Fig.11 (b)), the spectrum shapes of model A-6 are generally similar to that of

HRTF for all the subjects. The frequencies of P1 of the model are close to that of HRTF for subject A, B, and C, however, they are higher than that of HRTF for subject D and E. The frequencies of N1 are close to that of HRTF for all the subjects except for subject C. The frequencies of N2 of the model are close to that of HRTF for subject A, C, and E. However, it is higher than that of HRTF for subject B, and lower for subject D.

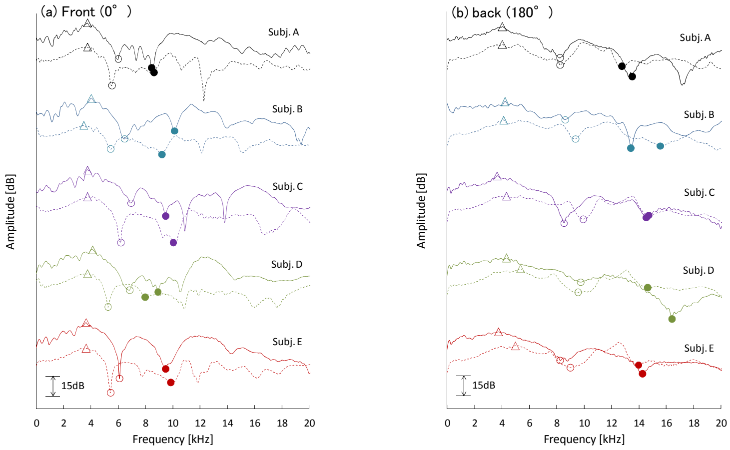

Figure 11 – HRTF (solid line) and transfer functions of model A-6 (broken line). △: P1 ○: N1 ●: N2

4. Conclusions

In this study, following issues are examined to achieve the estimation process of the subject’s N1 and N2 frequencies from the anthropometric data of subject’s pinnae;

1) The transfer functions were measured for various size and proportions of three-step models. The results show that:

a) the frequencies of N1 and N2 become higher as the pinna model becomes smaller,
b) the width of the pinna model has a dominant influence on the frequencies of N1 and N2,
c) for elevation of 120°, the length of the pinna model affects the frequencies of N1 and N2 a lot

2) The transfer functions of the three-step models, which were created using the anthropometric data of five listener’s actual pinnae, were measured. Comparisons of the N1 and N2 frequencies between pinna models and the actual pinna infer the possibilities of the accurate estimation of the subject’s own N1 and N2.

Acknowledgement

A part of this works is supported in part by Grant-in-Aid for Scientific Research (A) 22241040. The authors would like to thank Prof. M. Morimoto (Kobe University) for providing the ear molds. Thanks also to Mr. S. Nishioka, T. Okamatsu, S. Sakaguchi, and H. Tsuchiya for his cooperation in the measurements of transfer function.

References

[1] M. Morimoto and Y. Ando, “On the simulation of sound localization,” J. Acoust. Soc. Jpn(E) 1, pp. 167-174 (1980).
[2] R. Sottek and K. Genuit, “Physical modeling of individual head-related transfer functions,” Proc. DAGA, (1999)
[3] K. Iida, M. Itoh, A. Itagaki, and M. Morimoto, “Median plane localization using parametric model of the head-related transfer function based on spectral cues,” Applied Acoustics, 68, 835-850 (2007)
[4] H. Takemoto et al., “A simple pinna model for generating head-related transfer functions in the median plane,” Proceedings of 20th International Congress on Acoustics, ICA 2010 (2010)
[5] K. Iida and Y. Ishii, "3D sound image control by individualized parametric head-related transfer functions,” Proc. internoise 2011 (2011).