# A Study on individualization of Head-Related Transfer Function in the median plane

Hideki Hisatsune and Masato Akagi

School of Information Science, Japan Advanced Institute of Science and Technology
1-1 Asahidai, Nomi, Ishikawa 923–1292 Japan
Phone/FAX:+81-761-51-1699/+81-761-51-1149
Email: {h_hisatsune, akagi}@jaist.ac.jp

**Abstract**

Individualization of Head-Related Transfer Function (HRTF) is important. If an inappropriate HRTF is used, wrong sound localization and less sense of presence occur. In this paper, we studied the admissible ranges for P1, N1 and N2, which is spectral cue of the median plane for the sound localization. In order to verify the validness of the admissible ranges for the P1, N1 and N2, we investigated individualization differences in the P1, N1 and N2. As a result, we found less individualization difference on the P1, and big individualization difference in N1 and N2. And, we carried out the listening experiments on these findings. We revealed that P1 and N1 should be accurate; admissible range is narrow, and N2 is not strict, strict tuning of N2 is not required. Moreover, we observed the selected HRTFs in which variations of N1 and N2 according to the change of elevation. The result suggested that the control of P1 and N1 is especially important. This indicated necessity of adapting slope of the N1 and N2 when human perceive the change of elevation.

## 1. Introduction

Head-Related Transfer Function (HRTF) is the transfer function from a sound source to eardrums of a listener. When we convolve HRTF into a sound wave, three-dimensional presentation of the sound to listeners is possible [1]. However, there is a big problem that the HRTF is dependent on head, body, and auricle shapes of the listeners. If an inappropriate HRTF is used, wrong sound localization and less sense of presence occur. In order to present a highly accurate sound image in the three-dimensional space, HRTF should be as accurate as possible for each listener. However, measurement of HRTF requires large-scaled equipment and much time. Thus, it is important to provide individualizing method of HRTF for each listener.

HRTF required Interaural Time Difference (ITD) and Interaural Level Difference (ILD) for the cue of the horizontal plane. In order to judge the median plane, spectral cues are used [2]. The control of sound localization on the median plan is very difficult, due to no information of ITD and ILD rather than that on the median plan. Therefor in order to discuss about the indivisualization of HRTF, spectral cues are an important topic.

Iida et al. suggested that the sound localization in the median plane can be done by using only N1, N2 notches and P1 peak in the amplitude spectrum of HRTF [3]. This assumption was ensured by other reports that, when N1 and N2 are gone out by blocking the hollow of the ear of the listener, accuracy of elevation perception degrades [4]. Other reports showed that resonances in the ear involves the origin of peaks and notches [5].

However, there are still less discussion how large mismatches of individualization for P1, N1 and N2 affect on sound localization. If we can discuss an admissible range for P1, N1 and N2, this can provides new knowledge of individualization of the HRTF in median plane.

This study discusses admissible ranges for P1, N1 and N2. We carry out a listening experiments for this purpose. We select the n-best HRTFs from the database based on the listening experiments, and then analyze the P1, N1 and N2 and obtain the variance of the HRTFs. For this purpose, the section 2 describes investigation of individual difference of the P1, N1 and N2, to verify the validness of admissible ranges. The section 3 shows results of admissible ranges obtained by a listening experiment.

## 2. Individualization difference of P1, N1 and N2

In order to describe the validness of admissible range, this section describes the individualization difference of P1, N1 and N2. If the individualization difference of P1, N1 and N2 are small, there is no validness of admissible range. Using HRTF database we carried out investigation of the individualization difference of P1, N1 and N2.

### 2.1 Analysis method and conditions

In order to estimate P1 frequency, we apply cepstral smoothing to HRTF. N1 and N2 are sometimes not clear in some listener or elevation. For estimating N1 and N2 frequencies, we extract the initial impulse response (approximately 1-2 [ms] long) and apply FFT because of reducing effects of reflections in the ears [6]. 114 HRTF (of right ear) data obtained in the anechoic chamber in Tohoku university were used to individualize HRTFs for each listener.

### 2.2 Analyzed results

Analyzed results of P1, N1 and N2 are shown in Figure1. Additionally, average and standard deviation of P1, N1 and N2 on each elevation are shown in Figure2. The figures show as following tendency.

1. Individual difference about 1000 [Hz] in P1 on every elevation
2. Individual difference about 4000-6000 [Hz] in N1 and N2 on every elevation
3. Constant P1 even on elevation changes
4. Rising N1 on between 0-120 [degree] then decreasing to 180 [degree]
5. Rising N2 on between 0-120 [degree] then small changes to 180 [degree]

The tendencies above are similar to the previous study [7]. This ensures the analyzed results of N1 and N2.

The result show less individualization difference on the P1 and large individualization difference on N1and N2. The validness of the admissible ranges is verified for individualization of HRTF on the median plane.

## 3. Admissible ranges of P1, N1 and N2

This section discusses the admissible ranges. We carry out the listening experiments for this purpose. We select the n-best HRTFs from the database based on the listening experiments, and then analyze the P1, N1 and N2 and obtain the variance of the HRTFs.

### 3.1 Experimental method

The listening experiments choose the n-best HRTFs from the database, comparing 114 HRTF data. In order to choose n-best HRTFs accurately, we carried out the experiment in following three steps.

**Experiment I:** The n-better HRTFs are selected by comparing the HRTF database. The HRTF database is evaluated by checking sound localization. It is tried 2 times.

**Experiment II:** The 5-better HRTFs are selected by comparing the n-better HRTFs. The evaluation is the same as Experiment I. It is tried 2 times for the listener whose the n-better HRTFs is frequent, it is tried 3 times for the other.

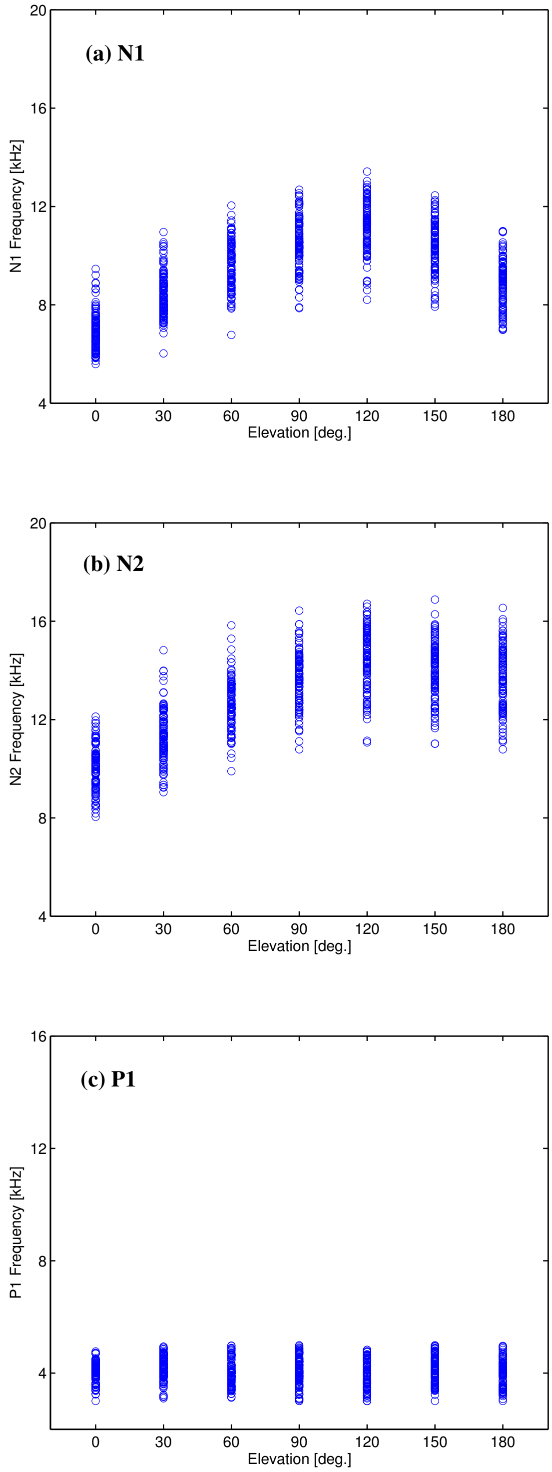


Figure 1: Frequency distribution of N1 and N2, P1 of upper-hemispheric median plane：(a) N1, (b) N2, (c) P1

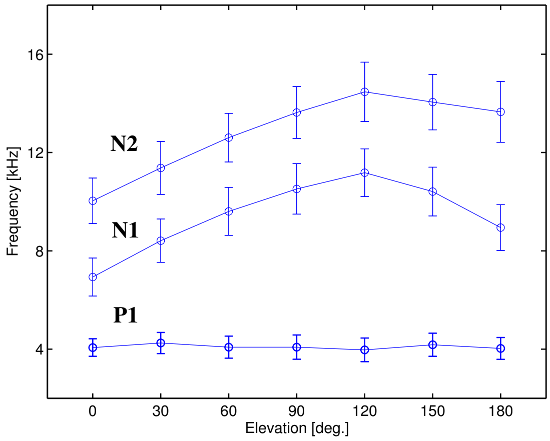


Figure 2: Average and standard deviation of P1, N1 and N2 of upper-hemispheric median plane

**Experiment III:** The best HRTF is selected by comparing the better 5-HRTFs. It is tried 2 times.

### 3.2 Experimental conditions

Figure3 illustrates the experiment system. The experiments are carried out in a sound-proof room by using Tucker-Davis Technologies (TDT) SystemIII. Stimuli are presented to each listener through headphone (STAX SR-404). Eight graduate students who had normal hearing are participated for the experiments.

We use white noise convoluted with the HRTFs in the database as stimuli. The sampling frequency of the stimuli is 48 [kHz]. Target elevations of the stimuli are from 0 degree (front) to 180 degree (back) via 90 degree (right above) in 30 degree steps in the median plane. In Experiments I and II, the sound for each elevation is presented in 3 [s]. In Experiment III, the sound for each elevation is presented in 1.5 [s]. In Experiments I and II, the listeners answer how sure they perceive the direction as indicated in each elevation. If the listeners show the right answer for all direction in one stimulus sound, we determine this sound is better HRTF for the listener. In Experiment III, the listeners hear paired stimuli and answer which is better.

### 3.3 Experimental results

Eight to thirteen HRTFs were selected for each listener in Experiment I. Experiments II and III are performed by five listeners, because three listeners could not perceive the virtual sound field. The better five HRTFs are determined in Experiment II. The best HRTF is determined in Experiment III.

The result obtain by the listener MA is shown in Figure4. The result for standard deviation and average values of the listener MA is shown in Figure5. About Figure4, thick solid line is the best HRTF and the dotted lines are the better HRTFs for the listener MA. The figure shows that the frequency band of the better HRTFs and the best HRTF is small for every elevations on P1 and N1. But it is large for N2. Moreover, the figure shows that variations of N1 and N2 according to the change of elevation is similar. Figure5 shows that the standard deviations of P1 and N1 are small, but N2 is large. The figure shows that the slopes of average of N1 and N2 are similar. This result is similar among the four out of five listeners.

We obtain that the admissible ranges for P1 and N1 are narrow in all elevations, and the range for N2 is not strict. Additionally, we observe that variations of N1 and N2 according to the change of elevation correlate. The result shows that the control of P1 and N1 is especially important, and this shows that adapting slope of the N1 and N2 when human perceive the change of elevation is required.

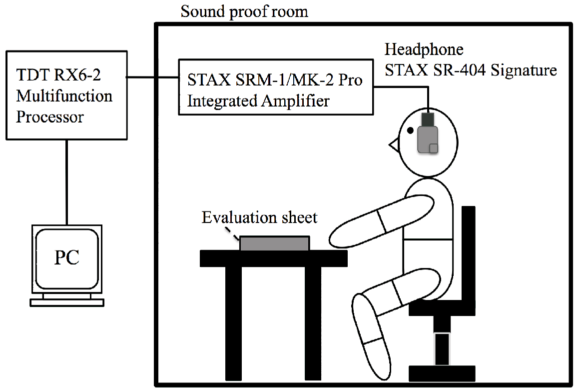


Figure 3: Experimental system

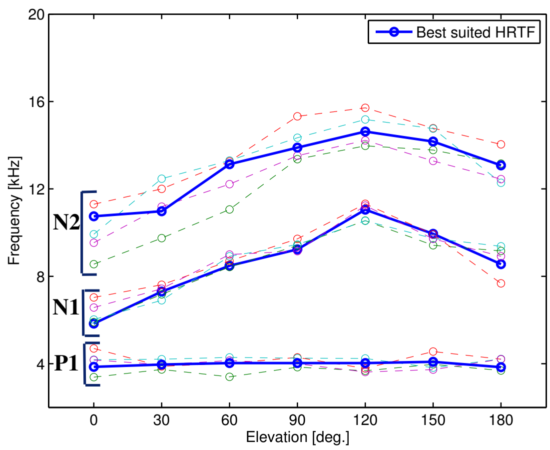


Figure 4: The selected better five HRTFs of P1, P2 and N2 (litener MA) : The solid line indicates the best HRTF for the subject

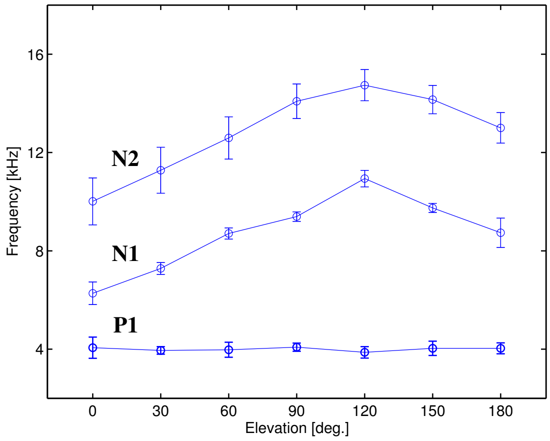


Figure 5: Standard deviation and average values of P1, N1 and N2 of selected HRTFs (litener MA)

## 4. Conclutions

In this paper, we studied the admissible ranges for P1, N1 and N2, which are spectral cues of the median plane for the sound localization. The admissible ranges were obtained by the listening experiments. The results showed that the control of P1 and N1 is especially important, and this showed necessity of adapting the slopes of the N1 and N2 when human perceives the change of elevation.

In our future work, if we could establish a control model of P1 and N1, we would be able to obtain easier individualization method of HRTF on the median plane.


## Acknowledgment

We thank Profs. Yukio Iwaya and Yôiti Suzuki for allowing us to use the HRTF database. This work was supported by JSPS A3 foresight program.


## References


[1] M. Morimoto, and Y. Ando, “On the simulation of sound localization,” *J. Acoust. Soc. Jpn.*, 1, pp.167–174, 1980.

[2] J. Blauert, “Spatial Hearing,” (MIT, Cambridge, MA), 1997.

[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.*, Vol. 68, No. 8, pp.835–850, 2007.

[4] K. Iida, M. Yairi, and M. Morimoto, “Role of pinna cavities in median plane localization, *”Proc.16th International Congress on Acoustics (Seattle)*, Vol.103, No.5, 845–846, 1998.

[5] H. Takemoto, P. Mokhtari, H. Kato, R. Nishimura, and K. Iida, “Basic investigation for effects of pinna shapes on head related transfer functions, *”Proc. Autum Meeting of the Acoustical Society of Japan,* 1445–1448, 2009.

[6] K. Iida, N. Gamoh, and Y. Ishii, “A consideration on a method to detect the 1st and 2nd spectral noches in the head-related transfer function, *”Proc. Fall Meeting of the Acoustical Society of Japan*, 473–476, 2010.

[7] Y. Ishii, S. Nishioka, and K. Iida, “Consideration on the individual difference of the pinna shape and the spectral notches on the median plane - Construction of the HRTF database with the quantitative individual difference information -,” *Proc. Fall Meeting of the Acoustical Society of Japan*, 463–466, 2012.