IWPASH
INTERNATIONAL WORKSHOP
ON THE PRINCIPLES AND APPLICATIONS OF SPATIAL HEARING
NOVEMBER 11-13, 2009, ZAO, MIYAGI, JAPAN

Motion-tracked 3D sound image localization system with individualization of head-related transfer functions

Yohji ISHII, Kazuhiro IIDA*
Chiba Institute of Technology

1 Introduction

It is well known that head-related transfer functions (HRTFs) have individual difference due to the difference in shape of the head and the pinna of the listeners. A method of searching appropriate HRTFs for each listener is required to accomplish an accurate 3D sound image control when the listener's own HRTFs are not available. This study introduces motion-tracked 3D sound image localization system with individualization of HRTFs.

2 Outlines of 3D sound image localization system

Figure 1 shows the total system configuration. This system is composed of a PC, headphones, ear-microphones, a motion sensor, and an audio inter face. A PC includes following functions;

1) individualization of HRTFs,
2) cancellation of the transfer functions between the headphones and the entrances of ear canals of a listener,
3) convolution of the sound sources and the HRTFs,
4) tracking of the head movement.

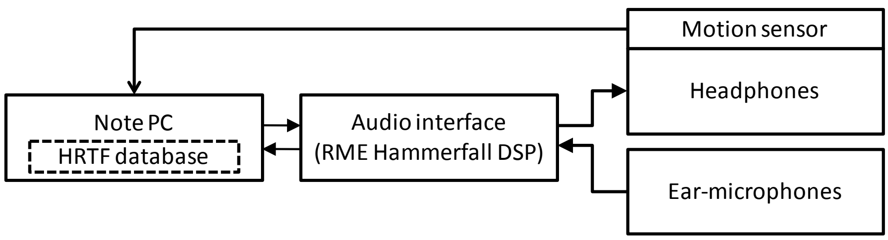

Fig.1 System configuration.

3 Individualization of HRTFs

This system extracts the most appropriate HRTFs for a listener from the HRTF database. In this study, the system doesn't use the measured HRTFs, but parametric HRTF for the database. Iida *et al*. [1] carried out localization tests in the median plane using a parametric HRTF, which is recomposed of all or some of the spectral peaks and notches extracted from the measured HRTF. Their results revealed that the parametric HRTF recomposed of the first and second notches (N1 and N2) and the first peak (P1), in order of frequency above 4 kHz, provides almost the same localization accuracy as the measured HRTF. Observations of the spectral peaks and notches indicate that the frequencies of N1 and N2 change remarkably with changes in the source elevation, whereas the frequency of P1 is independent of the source elevation. Thus, Iida *et al*. concluded that N1 and N2 could be regarded as spectral cues and that the human auditory system could use P1 as a reference for analyzing N1 and N2.

The system uses the parametric HRTF recomposed of N1, N2, and P1 (Fig.2). The individual difference of HRTFs is described as the difference in the frequency, level, and Q of two spectral notches and a peak. Among them, the frequencies of N1 and N2 are the most important characteristics. Thus, the system searches the appropriate parametric HRTFs from the database by changing the N1 and N2 frequencies systematically.

The HRTF database should be small in order to reduce the searching time and the burden of the listeners. The HRTF database, which is composed of required minimum parametric HRTFs, is created as following procedure.

1) Obtain the range of individual differences in N1 and N2 frequencies from many listeners (Fig. 3).
2) Divide the distribution range by the values of jnd of N1 and N2 frequencies for localization (0.1 - 0.2 oct.).
3) Extract the parametric HRTFs of which N1 and N2 frequencies are those of the grid point of Fig.4.

As for the front direction, at which front-back localization error occur frequently due to the individual difference, 38 parametric HRTFs were extracted. It could be considered that this is a required minimum amount of HRTFs for searching the appropriate HRTFs for each listener. It takes only 3 minutes to search them by the localization test.

* E-mail address: kazuhiro.iida@it-chiba.ac.jp

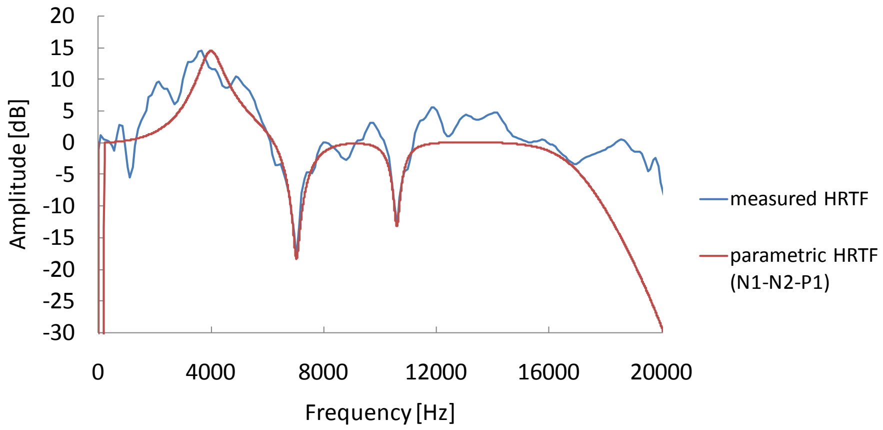

Fig.2. An example of parametric HRTF recomposed by N1, N2, and P1.

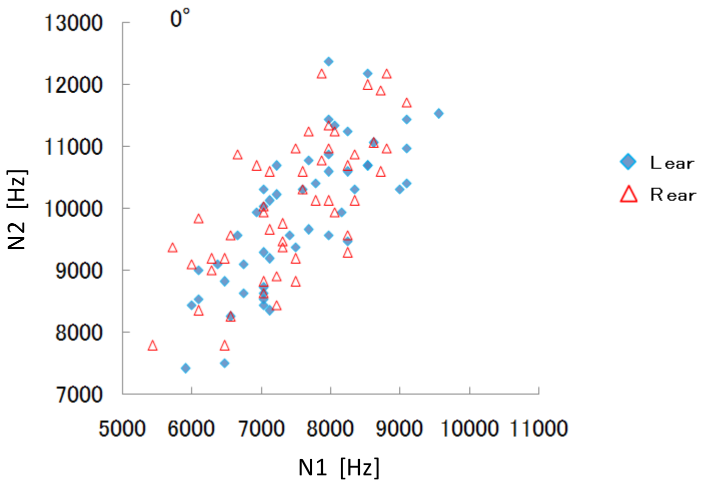

Fig. 3 Distribution of N1 and N2 frequency of 50 subjects' HRTF for front direction.

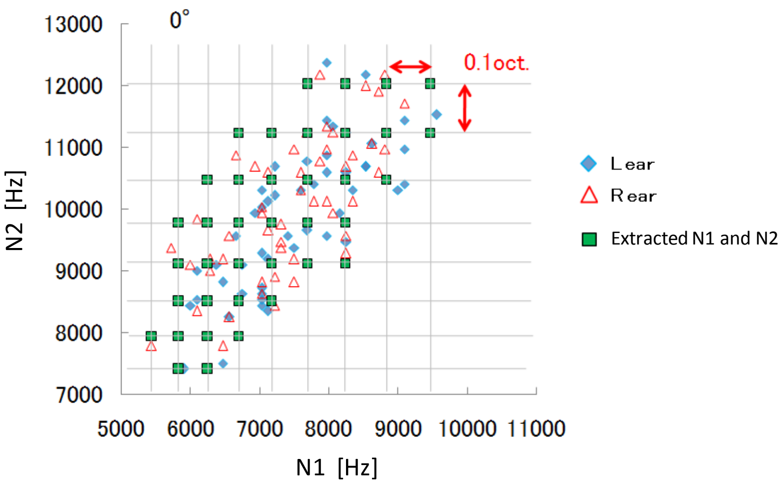

Fig. 4 Extracted N1 and N2 frequency by dividing the distribution range
by the jnd of N1 and N2 frequencies (0.1 - 0.2 oct.)

4 Cancellation of headphones transfer functions

Ear-microphones [1] are used in the cancellation of the transfer function between headphones and the entrances of the ear canals of a listener. The ear-microphones fabricated using the listener's ear molds are used for particular listeners. Miniature electret condenser microphones of 5 mm diameter (Panasonic WM64AT102) and silicon resin are put into the ear canals of the ear molds and consolidated (Fig.5). However, it is hard to get the ear molds for general listeners. Therefore, the conventional one was developed for general listeners who don't have his/her ear molds (Fig.6).

In order to cancel the headphones transfer functions, the ear-microphones are put into the ear canals of a listener. The diaphragms of the microphones are located at the entrances of the ear canals. Therefore, this is so called the "meatus-blocked condition" [2], in other words, the "blocked entrances condition" [3] Then, the listener wears the open-air headphones (SONY MDR-F1, altered), and the time-stretched pulses are emitted through them (Fig.7). The signals are received by the ear-microphones, and the transfer functions are obtained. Then, the ear-microphones are removed, and stimuli are delivered through the open-air headphones. Stimuli $P_{l,r}(\omega)$ are created by Eq. (1):

$$P_{l,r}(\omega) = S(\omega) \cdot H_{l,r}(\omega) / C_{l,r}(\omega), \quad (1)$$

where, $S(\omega)$ and $H_{l,r}(\omega)$ denote the source signal and HRTF, respectively. $C_{l,r}(\omega)$ is the transfer function between the open-air headphones and the ear-microphones.

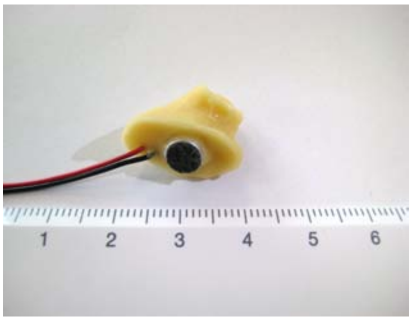

Fig. 5 An ear-microphone for particular listener

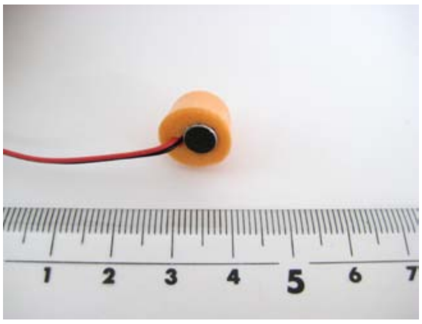

Fig. 6 An ear-microphone for general listeners

Fig. 7 A listener putting ear-microphones into his entrances of ear canals and wearing open-air headphones

5 Summary

Authors introduced motion-tracked 3D sound image localization system which includes functions of individualization of HRTFs and cancellation of the transfer function between the headphones and the entrances of the ear canals for general listeners.

References

[1] Iida K, Itoh M, Itagaki A, and Morimoto M "Median plane localization using parametric model of the head-related transfer function based on spectral cues," Applied Acoustics, 68: 835-850 (2007).

[2] Shaw EAG, Teranishi R "Sound pressure generated in an external-ear replica and real human ears by a nearby point source," J Acoust Soc Am. 44:240-249 (1968).

[3] Hammershoi D, Moller H "Sound transmission to and within the human ear canal," J Acoust Soc Am. 100:408‐427 (1996)