Extraction Process of Spectral Cues from Input Signals to Two Ears in Median Plane Localization

K. Iida[a], M. Itoh[b], E. Rin[b], and M. Morimoto[b]

[a]*Multimedia Solution Labs, Matsushita Comm. Ind. Co., Ltd, 600 Saedo, Tsuzuki, 224-8539 Yokohama, Japan*
[b]*Environmental Acoustics Lab, Fac. of Eng., Kobe University, Nada, 657-8501 Kobe, Japan*

Most previous studies on median plane localization have treated spectral cues as a monaural phenomenon. Morimoto demonstrated that both ears contribute to the perception of the vertical angle even in the median plane. However, it is not yet clear whether spectral cues are extracted from the input signal to each ear independently, or they are extracted after some integration process of input signals to two ears. In this study, localization tests were carried out using the stimuli, which provided HRTF of different vertical angle in the median plane to the right and left ear, respectively. The results show that the subjects localized a sound image to either direction of HRTF, or sound images to both directions of HRTF. This suggests that a listener extracts spectral cues directly from the input signal to each ear, independently.

INTRODUCTION

Most previous studies on median plane localization have treated spectral cues as a monaural phenomenon. Morimoto [1] demonstrated that both ears contribute to the perception of the vertical angle even in the median plane. Moreover, the contribution of the near ear is more than that of the far ear, when a sound source shifts laterally from the median plane [2,3]. Though some models on the extraction process of the spectral cues have been proposed [2,3], it is not yet clear whether spectral cues are extracted from the input signal to each ear independently, or they are extracted after some integration process of input signals to two ears. In this study, two hypotheses were made and the validity of them was examined.

EXTRACTION PROCESS OF SPECTRAL CUES

According to the previous studies, the following two hypotheses on the extraction process of the spectral cues for the perception of the vertical angle could be built:

1) Hypothesis 1: Integration of the Spectra

The spectra of the input signals to two ears are integrated into one spectrum. The spectral cues are extracted from the integrated spectrum.

2) Hypothesis 2: Integration of the Cues

The spectral cues are extracted from the spectrum of the input signal to each ear independently. Listeners perceive the vertical angle of a sound image by integrating those spectral cues.

METHOD OF LOCALIZATION TESTS

The HRTF of the subjects in the upper median plane were measured at seven elevations of every 30° from the front to the rear. The source signal was the band-limited white noise (280Hz - 11.2kHz). Stimuli were prepared by convolving the noise with the measured subject's own median plane HRTF. The HRTF from the sound source at different vertical angles were provided to the left and right ears, respectively. Namely, directional information provided to the left and right ears are different. Stimuli were presented to the subjects by the sound field simulation system through near-ear loudspeakers. This system compensates the transfer function from the near-ear loudspeakers to the entrance of the ear canal of the subject by DSP, and the crosstalks between left and right ears are negligible small. Each stimulus was presented 10 times in random order at 60±0.4 dBA at the entrance of the ear canal of the subjects. The duration of the stimuli was 1s and the interval between two stimuli was 9s. The task of the subjects was to mark down the perceived elevation of the sound image on the recording sheet. Subjects were four males with normal hearing sensitivity.

RESULTS AND DISCUSSIONS

The subjects reported that they perceived all sound images outside the head. Figure 1 shows an example (Subject: IT) of the responses to the stimuli, in case that the elevations provided to two ears are the same. This shows that he localized sound images accurately with the simulation system.

Figure 2 shows an example (Subject: IT) of the responses to the stimuli, in case that the HRTF for

different elevations were provided to the left and right ears. Figure 2(a) shows the 10 responses in case of HRTF for 0° and 180°. The subject sometimes perceived two sound images simultaneously, and sometimes one sound image. The perceived elevations agree with the provided ones. This means that the listener perceives the elevation of a sound image from the spectrum of the input signal to each ear, independently. Figure 2(b) shows the responses in case of HRTF of 30° and 60°. The responses agree with either elevation of two HRTF. Figure 2(c) shows the responses in case of HRTF of 120° and 30°. The subject perceived a sound image around 30° in most trials, and two sound images once. These were common behavior to other subjects. Therefore, it seems reasonable to consider that the hypothesis 2 is credible.

The mean localization error, e was obtained by Eq. (1).

$$e = \overline{|R - S|} \qquad (1)$$

where R is the perceived angle, and S is the simulated angle. Table 1 shows the errors when the elevations provided to two ears are the same (Case 1), and when they are different (Case 2), are shown in Table 1. Since two elevations are provided in the latter case, the smaller error is regarded as the error. Table 1 indicates that the localization error in Case 2 is almost the same as that in Case 1.

CONCLUSION

These results of the localization tests infer that the spectral cues for the vertical angle perception were extracted from the spectrum of the input signal to each ear, independently. It seems reasonable to consider that the hypothesis 2 is credible.

REFERENCES

1. M. Morimoto, J. Acoust. Soc. Am. **109**, 1596-1603 (2001).

2. R. A. Humanski and R. A. Butler, J. Acoust. Soc. Am. **83**, 2300-2310 (1988)

3. J. C. Middlebrooks, J. Acoust. Soc. Am. **92**, 2607-2624 (1992)

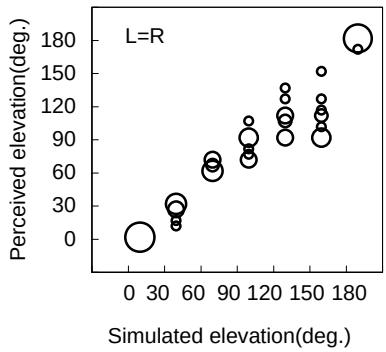

FIGURE 1. Responses to the stimuli. In case that the elevations provided to two ears are the same.

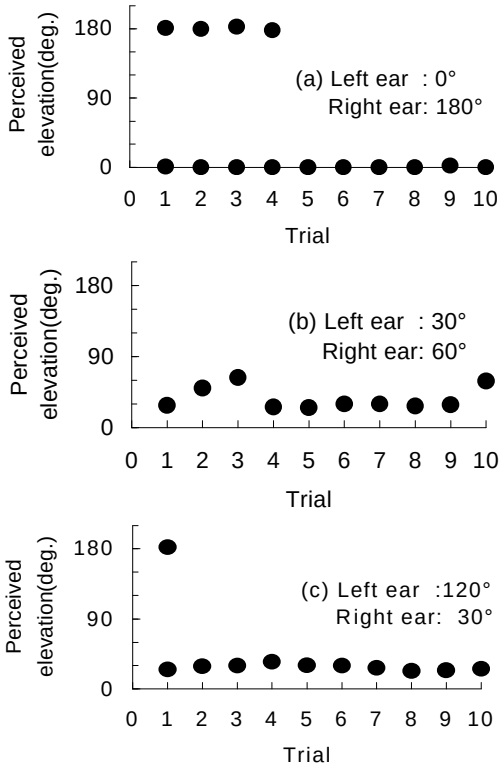

FIGURE 2. Responses to the stimuli. In case that the elevations provided to two ears are different.

Table 1. Mean localization error e

	Case 1 (the same HRTF)	Case 2 (the different HRTF)
e (deg.)	11.6	9.3