Acoust. Sci. & Tech. **24**, 5 (2003)

PAPER

Upper hemisphere sound localization using head-related transfer functions in the median plane and interaural differences

Masayuki Morimoto[1,*], Kazuhiro Iida[2,†] and Motokuni Itoh[1,†]

[1]*Environmental Acoustics Laboratory, Faculty of Engineering, Kobe University, Rokko, Nada, Kobe, 657–8501 Japan*

[2]*Multimedia Solution Laboratories, Matsushita Communication Industrial Co., Ltd, 600 Saedo, Tsuzuki, Yokohama, 224–8539 Japan*

(*Received 6 December 2002, Accepted for publication 29 May 2003*)

Abstract: Morimoto and Aokata [*J. Acoust. Soc. Jpn. (E)*, **5**, 165–173 (1984)] clarified that the same directional bands observed on the median plane by Blauert occur in any sagittal plane parallel to the median plane. Based upon this observation, they hypothesized that the spectral cues that help to determine the vertical angle of a sound image may function commonly in any sagittal plane. If this hypothesis is credible, sound localization in any direction might be simulated by using head-related transfer functions (HRTFs) measured on the median plane to determine the vertical angle, and by using frequency-independent interaural differences to determine the lateral angle. In this paper, a localization test was performed to evaluate the hypothesis, and to examine a simulation method based on the hypothesis. For this test, stimuli simulating HRTFs measured on the median sagittal plane combined with interaural differences measured on the frontal horizontal plane were presented to the subjects. The results supported the hypothesis and confirmed that the experimental simulation was not only possible, but also quite effective in controlling sound image location.

Keywords: Sound localization, Spectral cues, Head-related transfer function, Sagittal plane

PACS number: 43.66.Qp, 43.66.Pn, 43.38.Vk [DOI: 10.1250/ast.24.267]

1. INTRODUCTION

The present study has two aspects; one is a focus on sound localization cues, and the other is a test of a simulation method for localizing sound images.

It has been clarified that sound localization is accomplished by using two major cues, interaural difference cues and spectral cues [1]. Concerning the spectral cues, most former studies have concentrated on the cues in the median plane, and few have dealt with every point in three-dimensional space [2–5]. Morimoto and Aokata [2] introduced the interaural-polar-axis coordinate system shown in Fig. 1, and demonstrated that the lateral angle α and vertical angle β of a sound image are independently determined by human listeners based upon interaural difference cues and spectral cues, respectively. They also clarified that the same directional bands observed on the median plane by Blauert [6] occur in any sagittal plane. Middlebrooks [4] obtained a similar result. He showed that the horizontal component of a subject's response, which corresponds to the angle α in Fig. 1, is accurate when 1/6 octave-band noise was presented; and that the vertical and front/back component, which corresponds to the angle β, tended to cluster within restricted spatial ranges that were specific to each center frequency. Furthermore, Morimoto and Aokata [2] suggested that their subjects might readily use the spectral cues that are common across sagittal planes to determine the vertical angle β of their experimental sound stimuli.

Morimoto and Ando [7] had demonstrated earlier that the simulation of sound localization could be accomplished as long as head-related transfer functions (HRTFs) were accurately reproduced. Most recent studies on the simulation are based on this principle. However applications of this method face two difficult problems that must be solved. One problem is that a large number of HRTFs are required to simulate a sound image at any arbitrary direction. To solve this problem, data reduction has been performed on sets of HRTFs, for example, by applying principal components analysis [8–11] or by direct interpolation between HRTFs measured at only a few directions

*e-mail: mrmt@kobe-u.ac.jp

†Presently: Network Solution Development Center, Matsushita Electric Industrial Co., Ltd.

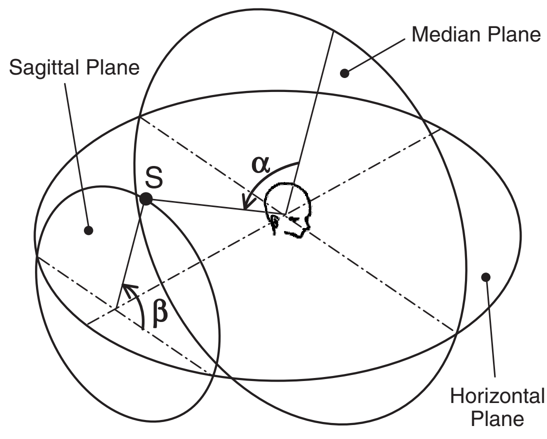

Fig. 1 Definition of the interaural-polar-axis coordinate system. α is the lateral angle and β is the vertical angle of a sound image S.

[12]. The other is the problem of individual differences between subjects' HRTFs [7,13]. Although some studies showed feasible solutions to the problem [14–16], a general-purpose simulation method for sound localization is not yet available. Needless to say, most of these approaches are based primarily upon mathematical methods that do not necessarily consider the localization cues, but rather operate upon all features of the HRTF data. Yet the information derived from the input signals to two ears, and used by the human auditory system in sound localization, may be based upon only part of the information present in the HRTF. A simulation method based upon specific sound localization cues might achieve a more effective and general-purpose result.

If the suggestion by Morimoto and Aokata mentioned above is credible, the vertical angle β of a sound image should be controllable using HRTFs for any sagittal plane, such as the median plane, regardless of the sagittal plane upon which the sound image is to be localized. Accordingly, sound localization cues for any direction can be simulated by using median-plane HRTFs to determine the vertical angle β, and interaural differences to determine the lateral angle α. Since the method requires HRTFs measured only in the median plane, the amount of required HRTF data dramatically decreases. Furthermore, the issue of individual differences in HRTFs could be addressed by capturing this small set of HRTFs on the median plane for each subject.

The purpose of the present paper is to evaluate the hypothesis regarding sound localization cues that was suggested by Morimoto and Aokata, and to examine the proposed simulation method for sound localization based on the hypothesis.

2. LOCALIZATION TEST

2.1. Method

2.1.1. Apparatus

For precise reproduction of HRTFs via headphones, accurate compensation for the transfer function from the headphone to the ear is required [1]. However, variation in this transfer function caused by uncertain headphone placement may not be negligible. Therefore, it is desirable to measure the transfer function and to calculate the compensation filter every time the subject puts on the headphones. Open-air headphones (AKG K1000) were employed for the tests done here, as they allow transfer functions to be measured while they are being worn (i.e., without removing the headphones). The difference between sound pressure levels measured at the entrances of the left and right ears exceeded 30 dB over the frequency range from 280 Hz to 11.2 kHz of the stimulus when an acoustic signal was presented from one of two transducers of the headphones. Thus it can be assumed that the interaural crosstalk from one ear's transducer to the opposite ear was negligible. A DSP board mounted on a PC was used for real-time convolution of the source signal with the sound-localization filter described below.

For the measurement of the subject's HRTFs, interaural differences, and the transfer functions from the headphone to the ear mentioned above, ear-microphones were developed individually for each subject. The ear-microphones were made using the following procedure. Molds of the ear canals of each subject were made. Then miniature electret condenser microphones (diameter: 5 mm) and silicon resin were put into the molds. For these measurements, the ear-microphones were placed within the ear canals of a subject to satisfy the condition of "blocked entrances" recommended by Hammershøi and Møller [17] for HRTF measurement.

2.1.2. Measurements of HRTF and interaural differences

The subjects' median-plane HRTFs in the upper hemisphere were measured in an anechoic chamber at seven vertical β angles ranging from frontal incidence to rearward incidence in 30 degree steps. The distance from the loudspeaker positions to the center of the subject's head was 1.5 m. First, a reference measurement of the electret condenser microphone used for the ear-microphone was made by placing it at the point corresponding to the center of the subject's head, but in a free field without the subject present. An M-sequence signal was reproduced by the loudspeaker, and 512-point impulse responses, $f_{l,r}(t)$, were measured at a 48 kHz sampling rate (with subscripts l and r indicating the left and right ears, respectively). The transfer functions from the loudspeaker to the microphones $F_{l,r}(\omega)$ were obtained by Fourier transformation of the $f_{l,r}(t)$. The $F_{l,r}(\omega)$ are expressed by

$$F_{l,r}(\omega) = \mathrm{SPK}(\omega) \cdot \mathrm{MIC}_{l,r}(\omega), \tag{1}$$

where $\mathrm{SPK}(\omega)$ is the transfer function of the loudspeaker, the $\mathrm{MIC}_{l,r}(\omega)$ are the transfer functions of the electret condenser microphone. Note that this measurement was taken before the ear-microphones were made. Next, the subject was seated with the ear-microphones inserted, and with head fixed. The impulse responses $e_{l,r}(t;\alpha,\beta)$ were measured, and the transfer functions from the loudspeaker to the ear-microphones $E_{l,r}(\omega;\alpha,\beta)$ were obtained by Fourier transforming $e_{l,r}(t;\alpha,\beta)$. The $E_{l,r}(\omega;\alpha,\beta)$ are expressed by

$$E_{l,r}(\omega;\alpha,\beta) = \mathrm{SPK}(\omega) \cdot \mathrm{HRTF}_{l,r}(\omega;\alpha,\beta) \cdot \mathrm{MIC}_{l,r}(\omega) : \quad \alpha = 0 \text{ in the median plane.} \tag{2}$$

Then the $\mathrm{HRTF}_{l,r}(\omega;\alpha,\beta)$ were obtained by

$$\mathrm{HRTF}_{l,r}(\omega;\alpha,\beta) = E_{l,r}(\omega;\alpha,\beta)/F_{l,r}(\omega). \tag{3}$$

In addition, interaural differences, consisting of a single ITD and ILD for each lateral angle α, were measured at four lateral angles ($\alpha = 0$, 30, 60, and 90 degrees) on the right side of the frontal horizontal plane ($\beta = 0$ degrees). The signals used for the measurements of ITD were the signals obtained by convolving a wide-band white noise source signal with the HRTFs measured at the four lateral angles. The ITD was operationally defined as the time lag at which the interaural cross-correlation of the signals reached a maximum. Also, ILD was directly measured using the ear-microphones response to the wide-band white noise presented from the loudspeakers at the same four lateral angles. Note that the frequency characteristics of the four loudspeakers were flattened to within ± 1.5 dB in the frequency range of the stimuli by a frequency equalizer (Technics SH-8065). These HRTFs and interaural differences were measured for each subject.

2.1.3. Stimuli

The source signal was a wide-band white noise ranging from 280 Hz to 11.2 kHz. The signal was shaped by a bandpass filter (NF 3625, -48 dB/Oct). The duration of the signal was one second with abrupt rise-fall time. The stimuli were delivered at 60 dB for the simulation of sound images in the median plane. The stimuli were presented as follows: At the beginning of each experimental session, the subject put both the headphones and the ear-microphones in place, and the transfer functions from the headphones to the ear-microphones $C_{l,r}(\omega)$ were measured. The $C_{l,r}(\omega)$ are expressed by

$$C_{l,r}(\omega) = \mathrm{HDP}_{l,r}(\omega) \cdot \mathrm{HM}_{l,r}(\omega) \cdot \mathrm{MIC}_{l,r}(\omega), \tag{4}$$

where the $\mathrm{HDP}_{l,r}(\omega)$ are the transfer functions of the headphones, and the $\mathrm{HM}_{l,r}(\omega)$ are the transfer functions from the positions of the transducers of headphones to the entrances of ear canals. The ear-microphones were removed after this measurement, while the headphones remained on the subject's head. The filters for simulating sound localization, $W_{l,r}(\omega;\alpha,\beta)$, were calculated as follows:

$$W_{l,r}(\omega;\alpha,\beta) = \mathrm{HRTF}'_{l,r}(\omega;\alpha,\beta)/C_{l,r}(\omega) = \frac{\mathrm{HRTF}'_{l,r}(\omega;\alpha,\beta)}{\mathrm{HDP}_{l,r}(\omega) \cdot \mathrm{HM}_{l,r}(\omega) \cdot \mathrm{MIC}_{l,r}(\omega)}, \tag{5}$$

where the $\mathrm{HRTF}'_{l,r}(\omega;\alpha,\beta)$ are the HRTFs that included both ITD and ILD. In practice, $\mathrm{HRTF}'_{l}(\omega;\alpha,\beta)$ were obtained by Fourier transformation of the impulse responses of the left ear measured for the median plane which were delayed by the time corresponding to the measured ITD and multiplied by the amplitude ratio corresponding to the measured ILD. $\mathrm{HRTF}'_{r}(\omega;\alpha,\beta)$ were $\mathrm{HRTF}'_{r}(\omega;0,\beta)$ measured for the median plane themselves. Stimuli were prepared by convolving the source signal $S(\omega)$ with the filters $W_{l,r}(\omega;\alpha,\beta)$ using the DSP board, and were presented through the headphones. The signals at the entrances of ear canals $P_{l,r}(\omega;\alpha,\beta)$ are expressed by

$$P_{l,r}(\omega;\alpha,\beta) = S(\omega) \cdot W_{l,r}(\omega;\alpha,\beta) \cdot \mathrm{HDP}_{l,r}(\omega) \cdot \mathrm{HM}_{l,r}(\omega) = S(\omega) \cdot \mathrm{HRTF}'_{l,r}(\omega;\alpha,\beta)/\mathrm{MIC}_{l,r}(\omega). \tag{6}$$

Here, the frequency characteristics of the $\mathrm{MIC}_{l,r}(\omega)$ were approximately flat within ± 2 dB in the frequency range of the stimulus. So the $\mathrm{MIC}_{l,r}(\omega)$ can be regarded as having unity gain, namely $P_{l,r}(\omega;\alpha,\beta) = S(\omega) \cdot \mathrm{HRTF}'_{l,r}(\omega;\alpha,\beta)$. Thus, the HRTFs measured on the median plane with the imposed interaural differences were accurately reproduced for the subjects.

For the localization test, 28 directions (seven measured HRTFs $\times$ four measured interaural differences) were simulated. Although the position at $\alpha = 90$ degrees is defined only for lateral angle α, and not for the vertical, the median-plane HRTFs for all seven vertical angles β were simulated. The idea here was to examine whether or not all of the responses would be concentrated around the target position at $\alpha = 90$ degrees, despite the variation in HRTFs associated with the seven β angles.

2.1.4. Procedure

The test was conducted in a partially darkened anechoic chamber. The subject was seated with chin fixed, and was instructed not to move his head. The task of the subjects was to mark the perceived azimuth and elevation of each sound image on a standard graphic response form. The response form displayed two circles intersected by perpendicular lines printed upon a sheet of paper. One circle was used to indicate the perceived azimuth angle, the other to indicate the perceived elevation angle (in reference to a spherical coordinate system containing a single vertical pole). The angles marked by subjects were read with a protractor to an accuracy of one degree, and were transformed into the angles α and β after the experiment. The duration of each stimulus was one second and the interstimulus interval was nine seconds (this interval giving the subject time to place the next recording sheet). The only light in the chamber was placed such that it provided just enough illumination for the subject to see and utilize the response recording sheets. Each stimulus set contained 28 different stimuli arranged in a random order. Twelve such sets were prepared for the test. The order of presentation of stimulus depended on the set. Twelve sets were divided into six sessions. Each session was completed in approximately ten minutes. Subjects were three males (IT, NS, YG), all with normal hearing sensitivity.

2.2. Results and Discussion

Responses given during the first session were regarded as practice and were excluded from the analysis of the results. The subjects reported that they perceived all sound images as well externalized (positioned well outside of their heads). Figures 2–4 show the responses of each subject. The circular arcs denote the lateral angle α, and the straight lines from the center denote the vertical angle β. The outermost arc denotes the median plane ($\alpha = 0$ degrees), and the center of the circle denotes the extreme side direction ($\alpha = 90$ degrees). The target α and β are shown in bold lines. The intersection of the two bold lines indicates the target direction. The diameter of the circular plotting symbols is proportional to the number of responses within each cell of a sampling grid with 5 degree resolution.

Broadly speaking, the responses are concentrated around the target directions. In order to distinguish the role of spectral cues and interaural difference cues, the lateral angle α and vertical angle β of the responses are discussed separately.

2.2.1. Distribution of perceived lateral angle

With subject IT (Fig. 2), for the target lateral angle $\alpha = 0$ degrees, that is, on the median plane (first column), the subject localized sound images in the median plane for all seven of the target vertical angles β. In the case of the target angle $\alpha = 30$ degrees (second column), the perceived angles α agreed with the target ones for the target vertical angles β of 0, 30, and 180 degrees. However, the responses were somewhat scattered, and shifted slightly towards the median plane for target angles β from 60 to 150 degrees. In the case of the target angle $\alpha = 60$ degrees (third column), the responses were scattered more than those for $\alpha = 30$ degrees, for all of the target β angles. Furthermore, shifts in the responses towards the median plane were observed for target β angles from 90 to 180 degrees. In the case of the target angle $\alpha = 90$ degrees (rightmost column), the perceived lateral angle α agreed closely with the target location for β target angles of 0 and 30 degrees. However, responses were scattered and shifted towards the median plane for target angles β from 60 to 180 degrees.

With subject NS (Fig. 3), for the target lateral angle $\alpha = 0$ degrees (first column), the subject localized sound images in the median plane for all seven of the target vertical angles β. In the case of the target angle $\alpha = 30$ degrees (second column), the perceived angles α agreed closely with the target ones for all of the target β angles. However, the responses were somewhat scattered. In the case of the target angle $\alpha = 60$ degrees (third column), the responses were scattered more than those for $\alpha = 30$ degrees, for target angles β from 0 to 60 degrees and 150 degrees. Furthermore, shifts in the responses towards the median plane were observed for target angles β from 60 to 180 degrees. In the case of the target angle $\alpha = 90$ degrees (rightmost column), the perceived angle α agreed with the target location for β target angles of 0, 30 and 60 degrees, except a few responses were shifted towards the median plane for the target angles β of 0 and 60 degrees. However, the responses were scattered and shifted towards the median plane for target angles β from 90 to 180 degrees.

With subject YG (Fig. 4), for the target lateral angle $\alpha = 0$ degrees (first column), the subject localized sound images around the median plane for all seven of the target vertical angles β. However, a tendency of the responses to appear outside the outermost arc, that is, on the left of the median plane was found on the whole. In particular, most of the responses were shifted slightly towards the left of the median plane for the target angle $\beta = 0$ degrees. In the case of the target angle $\alpha = 30$ degrees (second column), the perceived angle α agreed closely with the target location for the target angles β of 0 and 90 degrees, although the responses were somewhat scattered. However, the responses were scattered and slightly shifted towards the median plane for the other target angles β. In the cases of the target angles $\alpha = 60$ and 90 degrees (third and rightmost columns), the responses were scattered and shifted towards the median plane on the whole. Note that the responses were shifted towards the left for the target angle β of 0 degrees, except for the target angle α of 30 degrees, although they were expected to appear at the target angle α, since the interaural differences were simulated by using those measured at the angle β of 0 degrees. The shifts seem to be due to a kind of bias in the perception of interaural differences.

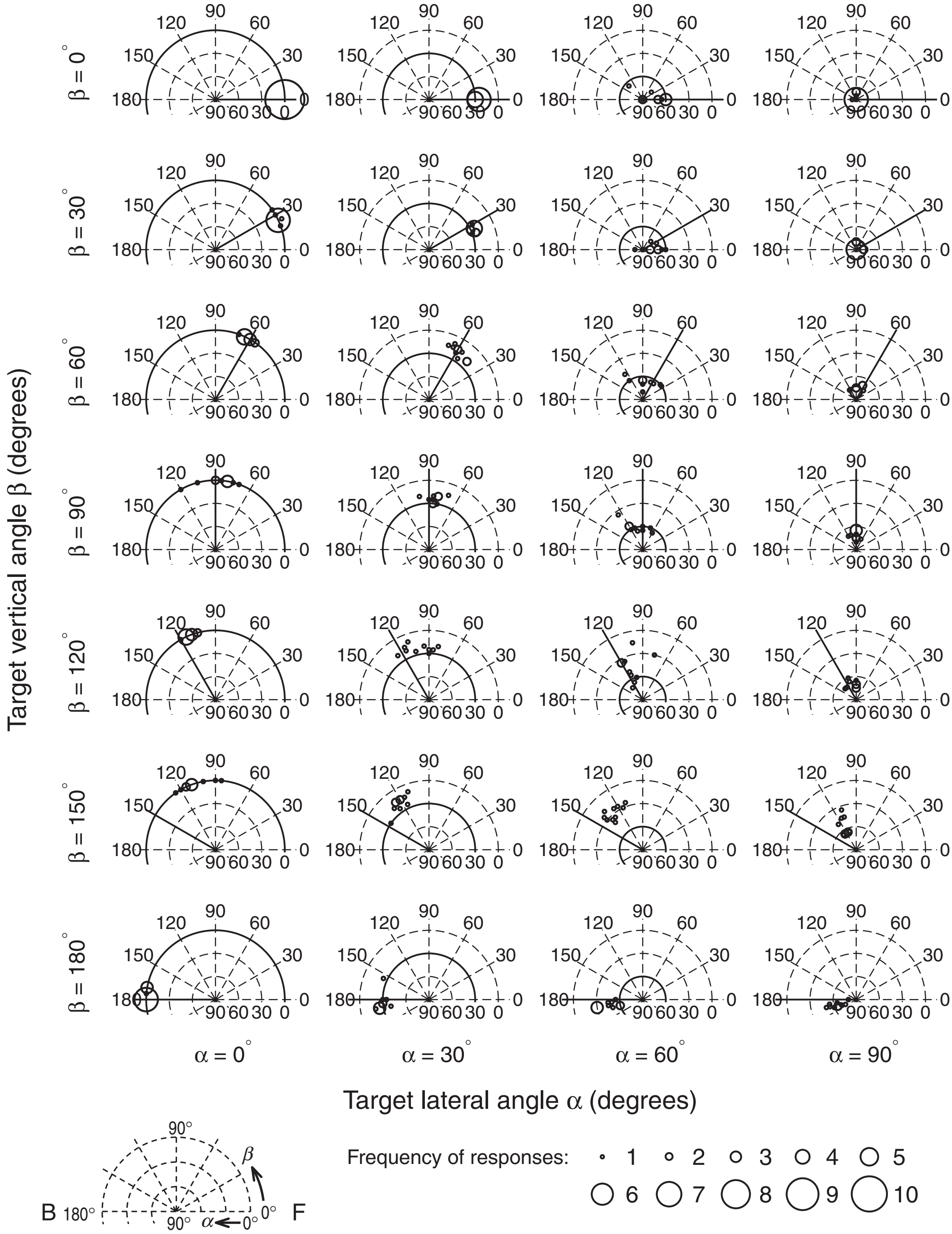

Fig. 2 Responses to the stimuli which simulated HRTF in the median plane and interaural differences for Subject IT. The circular arcs denote the lateral angle α, and the straight lines denote the vertical angle β. Bold lines show the target angles α and β.

Summarizing the results of three subjects, two kinds of error were observed commonly to three subjects. One is that the variance of responses increased as the target lateral angle α increased. This result is consistent with the just noticeable difference of horizontal plane localization [1]. The other is that the responses shifted toward the median plane for the target β angles from 60 to 180 degrees. According to ITD contours reported by Wightman and Kistler [18], both ITD and ILD for such vertical angles in a sagittal plane are larger than those for frontal directions. In this test, the interaural differences were simulated by using those measured at the target angle α only on the frontal horizontal plane, regardless of the target β angle. Accordingly, the simulated interaural differences for the target β angles from 60 to 180 degrees were smaller than those that could be measured. Thus it is inferred that the shift of the

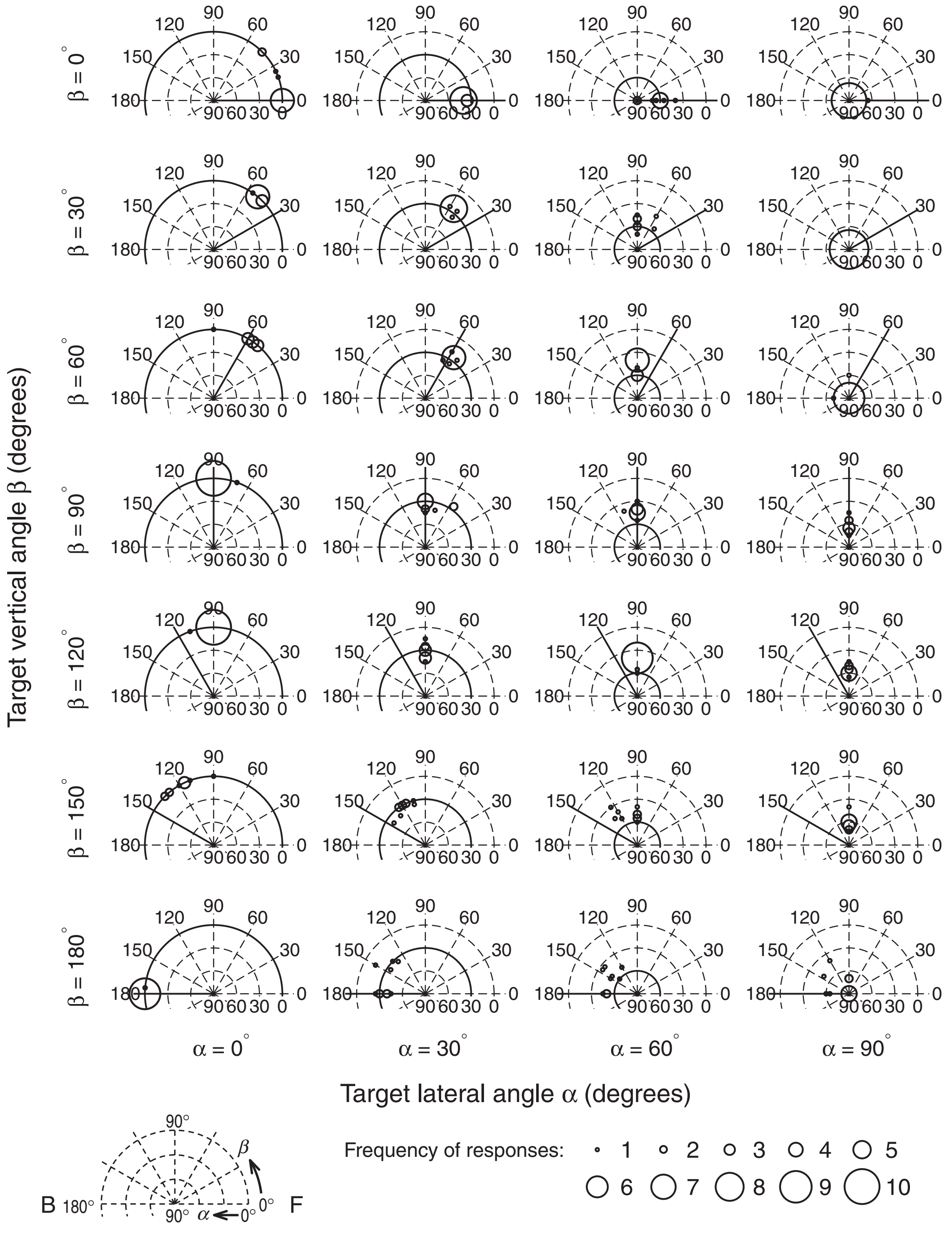

Fig. 3 As Fig. 2 for Subject NS.

responses towards the median plane was caused by the difference between real and simulated interaural differences.

2.2.2. Distribution of perceived vertical angle

With subject IT (Fig. 2), in the case of the target lateral angle $\alpha = 0$ degrees, that is, on the median plane (first column), the perceived β angles closely agreed with the target angles except for the target β angles of 90 and 150 degrees. The responses were somewhat scattered for these target angles, and were shifted towards $\beta = 120$ degrees for the target β angle of 150 degrees. This tendency that the responses for oblique directions in the median plane were sometimes shifted upwards coincides with responses observed for real sound sources [7]. Furthermore, this means that the simulation of sound localization was accomplished accurately without the effects of interaural crosstalk.

In the case of the target lateral angle $\alpha = 30$ degrees (second column), the responses showed a very similar tendency to that observed for the median plane, although

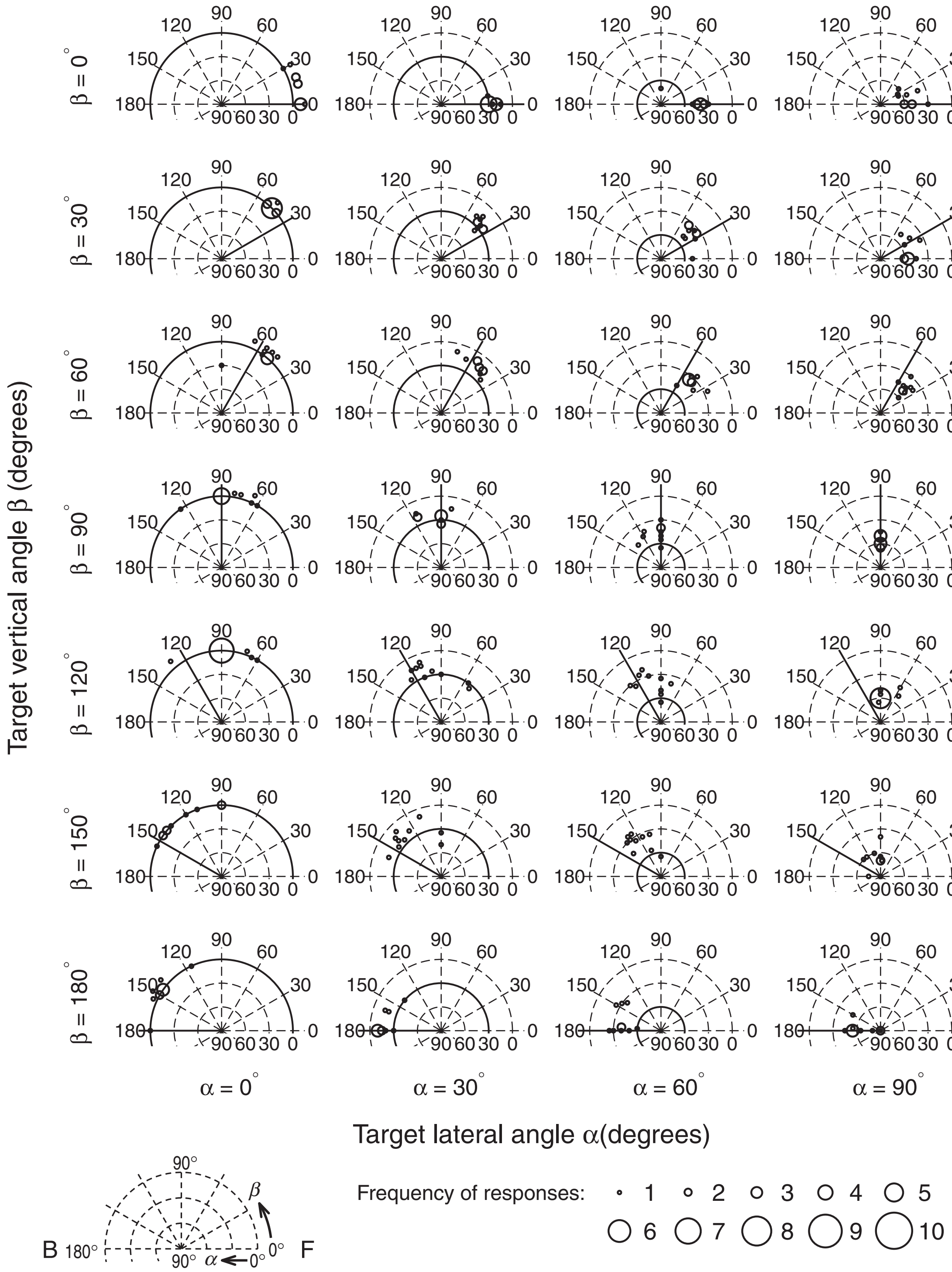

Fig. 4 As Fig. 2 for Subject YG.

the responses were scattered for the target β angle of 120 degrees. In the case of the target angle $\alpha = 60$ degrees (third column), a few front-back confusions occurred for the target β angles of 0 and 30 degrees. The responses shifted to $\beta = 0$ degrees for the target angle β of 30 degrees, and scattered for the target β angles of 60 and 120 degrees. Except for these few cases, the distributions of responses show the same tendency as those for the median plane. In the case of the target angle $\alpha = 90$ degrees (rightmost column), all responses were expected to appear at the position determined by $\alpha = 90$ degrees, regardless of the target angle β, since the position is defined only by the angle α. As a result, the responses appeared at the position of the angle $\alpha = 90$ degrees for target angles β of 0 and 30 degrees. Although the responses shifted towards the median plane for target β angles from 60 to 180 degrees, because of the mismatch in the simulation of interaural differences, the distribution of the perceived angle β are practically the same as those in the median plane.

With subject NS (Fig. 3), in the case of the target angle $\alpha = 0$ degrees (first column), the perceived β angles closely agreed with the target ones for the target β angles of 0, 60, 90 and 180 degrees, except for a few responses for the target β angle of 0 degrees. The responses were shifted upwards for the target β angles of 30, 120 and 150 degrees. However, such a tendency is sometimes observed for real sound sources as mentioned above. This means that the simulation of sound localization was accomplished as well as that for subject IT. In the case of the target angle $\alpha = 30$ degrees (second column), the responses showed a similar tendency to that observed for the median plane, except the perceived β angles agreed with the target ones for the target β angle of 0 degrees and a few responses were shifted towards $\beta = 150$ degrees for the target β angle of 180 degrees. In the case of the target angle $\alpha = 60$ degrees (third column), the responses were shifted to $\beta = 90$ degrees for the target angles β of 30 and 60 degrees, and a few responses were shifted upwards for the target β angles of 150 and 180 degrees. Except for these few cases, the distributions of responses show the same tendency as those for the median plane. In the case of the target angle $\alpha = 90$ degrees (rightmost column), the responses appeared at the position of the angle $\alpha = 90$ degrees for target angles β from 0 to 60 degrees as expected. Although the responses were shifted towards the median plane for target β angles from 90 to 180 degrees and shifted upwards for target β angles of 150 and 180 degrees, the distribution of the perceived angle β are practically the same as those in the median plane.

With subject YG (Fig. 4), in the case of the target angle $\alpha = 0$ degrees (first column), the perceived β angles closely agreed with the target ones for the target β angles of 0, 60, 90 and 150 degrees, although some responses were scattered and shifted upwards. The responses were shifted upwards on the whole for the other target β angles. Except for the target β angle of 180 degrees, such a tendency is sometimes observed for real sound sources as mentioned above. This means that the simulation of sound localization was accomplished as well as those for subjects IT and NS. In the case of the target angle $\alpha = 30$ degrees (second column), the responses showed a similar tendency to that observed for the median plane, except the perceived β angles closely agreed with the target ones for the target β angles of 0 and 180 degrees. In the case of the target angle $\alpha = 60$ degrees (third column), the responses show the same tendency as those for the target angle $\alpha = 30$ degrees. In the case of the target angle $\alpha = 90$ degrees (rightmost column), the distributions of the perceived angle β are practically the same as those in the median plane, although a few responses were shifted towards $\beta = 0$ degrees for the target β angle of 30 degrees and few shifts in the responses upwards were observed for the target β angle of 180 degrees.

Summarizing the results of three subjects, responses for any sagittal plane show a similar tendency to that observed for the median plane with minor exceptions. Accordingly, it can be concluded that the spectral cues for the perception of the vertical angle β provided by the median plane HRTFs played the same role in any other sagittal planes. This supports the hypothesis of Morimoto and Aokata [2] that the spectral cues observed on the median plane can be used to localize sound images on any sagittal plane.

2.2.3. Localization error

In contrast to the above reported directional biases in the distributions of judgments, an estimate of accuracy is possible using a measure of localization error obtained using Eq. (7):

$$e = \overline{|R - S|}, \tag{7}$$

where R is the reported perceived angle and S is the target one.

Table 1 shows the errors in the lateral and vertical angles for all subjects and for each target lateral angle. The localization error in the lateral angle α increases as the target angle α increases. This tendency agrees with the just noticeable difference in the perception of the lateral angle for naturally-heard sound sources [1]. Moreover, the average of these errors is practically equal to the localization error in the localization test by Morimoto and Ando [7] which reproduced the subject's own HRTFs accurately. This means that the angle α of a sound image can be accurately simulated on the average by using interaural differences measured only for the frontal horizontal plane.

The localization error in the vertical angle β is practically the same as that in natural median plane localization, as observed by Morimoto and Ando for any target lateral angle α. This result supports the hypothesis of Morimoto and Aokata [2] that the spectral cues to sound localization are common for all sagittal planes.

Consequently, these localization errors indicate that a sound image in any direction can be simulated by using only median-plane HRTFs and frontal-plane interaural differences, with much the same accuracy as real sound sources.

Table 1 Localization error in degrees when HRTFs in the median plane and interaural differences are simulated.

Error	Target angle α (deg.)			
	0	30	60	90
Perceived angle α	1	7	16	23
Perceived angle β	15	13	21	—

3. CONCLUSIONS

Stimuli simulating HRTFs measured on the median plane and interaural differences measured on the frontal horizontal plane were presented to three subjects for a localization test. The results showed the following: The vertical angle β of the sound images could be perceived with much the same accuracy as those of real sound sources, regardless of the lateral angle α. Similarly, the lateral angle α of the sound images could be also perceived with much the same accuracy as those of real sound sources, except for shifts toward the median plane for upper and rear sound images. These shifts could be explained by the difference between the simulated and the measured interaural differences for those angles. From these results, it can be concluded that the hypothesis suggested by Morimoto and Aokata [2] on sound localization cues is reasonable, and that spectral cues to sound localization are common in any sagittal plane. Moreover, these results indicate that it is basically possible to localize sound images in any direction via a simulation using median-plane HRTFs combined with frequency-independent interaural differences.

ACKNOWLEDGMENT

The authors would like to thank Prof. William L. Martens (University of Aizu) for his comments and copy-editing on the English version of this manuscript. Thanks also to Mr. E. Rin for his cooperation in the localization tests.

REFERENCES

[1] J. Blauert, *Spatial Hearing*, revised edition (MIT Press, Cambridge, Mass., 1997).

[2] M. Morimoto and H. Aokata, "Localization cues of sound sources in the upper hemisphere," *J. Acoust. Soc. Jpn.* (*E*), **5**, 165–173 (1984).

[3] S. R. Oldfield and S. P. A. Parker, "Acuity of sound localisation: a topography of auditory space. II. Pinna cues absent," *Perception*, **13**, 601–617 (1984).

[4] J. C. Middlebrooks, "Narrow-band sound localization related to external ear acoustics," *J. Acoust. Soc. Am.*, **92**, 2607–2624 (1992).

[5] V. R. Algazi, C. Avendano and R. O. Duda, "Elevation localization and head-related transfer function analysis at low frequencies," *J. Acoust. Soc. Am.*, **109**, 1110–1122 (2001).

[6] J. Blauert, "Sound localization in the median plane," *Acustica*, **22**, 205–213 (1969/70).

[7] M. Morimoto and Y. Ando, "On the simulation of sound localization," *J. Acoust. Soc. Jpn.* (*E*), **1**, 167–174 (1980).

[8] W. L. Martens, "Principal components analysis and resynthesis of spectral cues to perceived direction," *Proc. International Computer Music Conf.*, pp. 274–281 (1987).

[9] D. J. Kistler and F. L. Wightman, "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction," *J. Acoust. Soc. Am.*, **91**, 1637–1647 (1992).

[10] J. C. Middlebrooks and D. M. Green, "Observations on a principal components analysis of head-related transfer functions," *J. Acoust. Soc. Am.*, **92**, 597–599 (1992).

[11] S. Carlile, C. Jin and J. Leung, "Performance measures of the spatial fidelity of virtual auditory space: Effects of filter compression and spatial sampling," *Proc. 2002 International Conf. on Auditory Display*, pp. 375–380 (2002).

[12] T. Nishino, S. Kajita, K. Takeda and F. Itakura, "Interpolation of head related transfer functions of azimuth and elevation," *J. Acoust. Soc. Jpn.* (*J*), **57**, 685–692 (2001).

[13] E. M. Wenzel, M. Arruda, D. J. Kistler and F. L. Wightman, "Localization using nonindividualized head-related transfer functions," *J. Acoust. Soc. Am.*, **94**, 111–123 (1993).

[14] H. Møller, C. B. Jensen, D. Hammershøi and M. F. Sørensen, "Selection of a typical human subject for binaural recording," *Acustica*, **82**, S215 (1996).

[15] J. C. Middlebrooks, "Individual differences in external-ear transfer functions reduced by scaling in frequency," *J. Acoust. Soc. Am.*, **106**, 1480–1492 (1999).

[16] J. C. Middlebrooks, "Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency," *J. Acoust. Soc. Am.*, **106**, 1493–1510 (1999).

[17] D. Hammershøi and H. Møller, "Sound transmission to and within the human ear canal," *J. Acoust. Soc. Am.*, **100**, 408–427 (1996).

[18] F. L. Wightman and D. J. Kistler, "Resolution of front-back ambiguity in spatial hearing by listener and source movement," *J. Acoust. Soc. Am.*, **105**, 2841–2853 (1999).