

## The role of low frequency components in median plane localization

Masayuki Morimoto<sup>1,\*</sup>, Motoki Yairi<sup>1</sup>, Kazuhiro Iida<sup>2</sup> and Motokuni Itoh<sup>1</sup>

<sup>1</sup>*Environmental Acoustics Laboratory, Faculty of Engineering, Kobe University, Rokko, Nada, Kobe, 657-8501 Japan*

<sup>2</sup>*Multimedia Solution Laboratories, Matsushita Communication Industrial Co., Ltd., 600 Saedo, Tsuzuki, Yokohama, 224-8539 Japan*

(Received 30 April 2002, Accepted for publication 4 September 2002)

**Abstract:** The high frequency components of an auditory stimulus are considered to be the primary contribution to median plane localization. However, a number of studies have demonstrated that the low frequency components of a stimulus are also important in median plane localization. Asano *et al.* concluded that important cues for front-back discrimination seem to exist in the frequency range below 2 kHz. In the present paper, localization tests were performed in order to examine the contribution of low frequency components to median plane localization. In these tests, the higher (above 4,800 Hz) and lower (below 4,800 Hz) frequency components, respectively, of a wide-band white noise were simultaneously presented from different directions so that the individual components provided different directional information. The results of these tests reveal that: (1) when the source is a wide-band signal, the higher frequency components are dominant in median plane localization, whereas the lower frequency components do not contribute significantly to the localization, and (2) important cues for front-back discrimination do not exist in the low frequency range.

**Keywords:** Median plane localization, Spectral cues, Low frequency components, Head-related transfer function

**PACS number:** 43.66.Qp, 43.66.Pn

### 1. INTRODUCTION

It is generally known that the amplitude spectrum of the head-related transfer function (HRTF) provides cues, called spectral cues, for median plane localization [1]. However, the mechanism by which spectral cues are extracted has not yet been revealed.

Generally speaking, two assumptions have been discussed concerning this mechanism. One assumption is that some spectral features such as peak and dip provide the cues. Blauert [2] demonstrated that the direction of a sound image for a 1/3 octave band noise was a function of center frequency only and did not depend on the source elevation angle. He referred to the frequency band by which the direction of sound image is determined as the directional band. Moreover, he inferred that the direction of sound image for a broadband stimulus is perceived according to the directional band, which corresponds to the frequency components boosted in the amplitude spectrum of the HRTF. Hebrank and Wright [3] compared the results of localization tests performed using various band-pass

stimuli to HRTF spectra, and speculated that some peaks and dips in the spectrum act as spectral cues. Similar results were indicated by Butler and Belendiuk [4] and Bloom [5]. However, these results are not consistent with the results of experimental tests conducted by Asahi and Matsuoka [6], who reported that although the perceived direction of sound image for a dip-filtered stimulus changed as the dip frequency changed, the direction of sound image did not coincide with the direction of sound source, in which HRTF included the same dip as the stimulus.

The second assumption is that the entire spectrum of the stimulus contributes to the localization of the sound image. Morimoto [7] demonstrated that the localization accuracy increased as the bandwidth of a stimulus having the center frequency of approximately 8 kHz increased. In other words, Morimoto suggested that neither specific peaks nor dips act as spectral cues, and that the amount of the cue is proportional to the bandwidth of the stimulus. Middlebrooks [8] proposed the vertical angle perception model, in which the sound image is localized in the direction in which the correlation coefficient between the

---

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

spectrum of the input signal to the ears and the spectrum of the subject's HRTF is a maximum.

Meanwhile, it is assumed that the high frequency components of a stimulus are the primary contribution to median plane localization. Mehrgardt and Mellert [9] showed that the HRTF spectrum changes systematically above 5 kHz as the elevation angle changes. Morimoto [7] observed the localization accuracy using several high- or low-passed stimuli and concluded that when the stimulus contains frequency components of 4,800–9,600 Hz, the localization error is small, and that when the stimulus does not contain frequency components above 4,800 Hz, the sound image always appears on the horizontal plane. Moreover, when the frequency components above 2,400 Hz are lost, front-back confusion occurs.

On the other hand, a number of studies have demonstrated that the low frequency components of a stimulus also contribute to median plane localization. Asano *et al.* [10] conducted virtual localization tests using the HRTF smoothed by applying the ARMA (Auto-Regressive Moving-Average) model. They concluded that important cues for front-back discrimination seem to exist in the frequency range below 2 kHz. Algazi *et al.* [11] showed that subjects could estimate the elevation angle even if a stimulus low-pass filtered to 3 kHz was used, when the source was located away from the median plane. The contribution of low frequency components can also be found in Morimoto's results [7] which reveal that front-back discrimination can be achieved when the stimulus contains frequency components of 2,400–4,800 Hz.

However, the contribution of low frequency components to median plane localization has not yet been clarified. The importance of low frequency components is not clear from the results by Morimoto [7] and Algazi *et al.* [11], and the results reported by Asano *et al.* [10] are not sufficient to discuss the importance of low frequency components, since the smoothed HRTF used in their experiments can be regarded as providing different information than that provided by the true HRTF.

In the present paper, localization tests were performed in order to examine the contribution of low frequency components to median plane localization. In these tests, the higher and lower frequency components of a wide-band white noise are simultaneously presented from different directions so that each components may provide different directional information.

## 2. EXPERIMENTAL METHOD

### 2.1. Subjects

Subjects were three male students (SM, TM, TU) with normal hearing sensitivity.

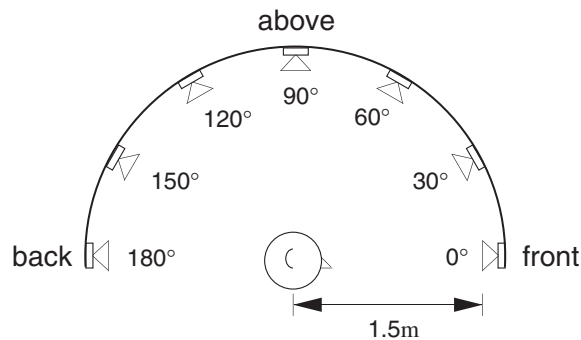


Fig. 1 Loudspeaker arrangement in the experiment.

### 2.2. Apparatus

The localization tests were conducted in an anechoic chamber. Seven cylindrical loudspeakers (diameter: 108 mm, length: 350 mm) were located at every 30° in the upper median plane, as described in Fig. 1. The loudspeaker radius was 1.5 m relative to the center of the subject's head. The frequency characteristics of the seven loudspeakers were flattened to within  $\pm 3$  dB in the frequency range of the stimulus by a frequency equalizer (Technics SH-8065).

### 2.3. Stimuli

The source signal was a wide-band white noise from 280 Hz to 11.2 kHz which was divided into two frequency components at the cutoff frequency of 4,800 Hz, as indicated in Fig. 2. In this paper, two localization tests were performed. In the first test, the two components were simultaneously presented from different loudspeakers. In the second test, the two components were presented from the same loudspeaker simultaneously. Hence the number of stimuli was 49 (seven directions for the high-pass components  $\times$  seven directions for the low-pass components). Note that seven out of the 49 stimuli had lower and higher frequency components that were presented from the same direction. The stimuli were delivered at  $50 \pm 0.2$  dBA for 1,682 ms, including 341 ms onset and offset ramps, followed by an interval of 5,000 ms.

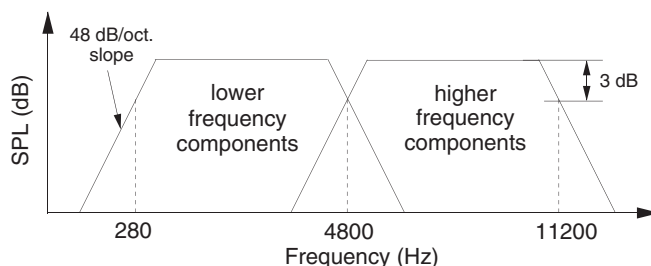


Fig. 2 Schematic diagram of the frequency characteristics of the stimuli.

## 2.4. Procedure

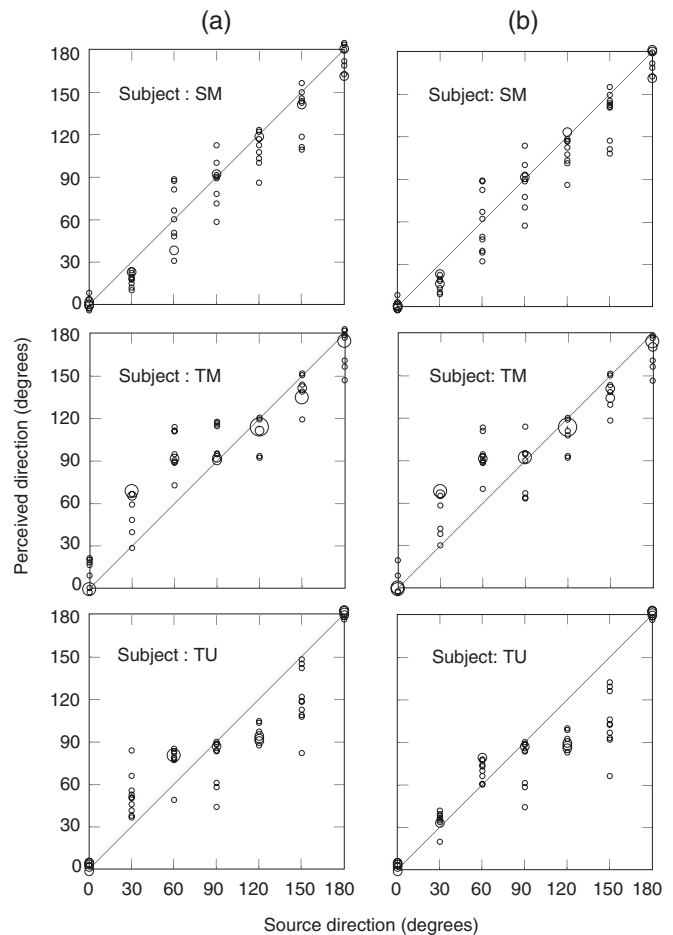
Each subject was tested individually while seated, with his head fixed in a stationary position, in a partially darkened anechoic chamber. Recording sheets with a circle intersected by perpendicular lines on it were supplied to the subjects. The circle indicates the median plane, and the four intersections of the circle by the cross lines indicate  $-90^\circ$ ,  $0^\circ$ ,  $90^\circ$ , and  $180^\circ$  relative to the front. The subject's task was to mark down the perceived directions of all sound images for each stimulus presentation. The 5,000 ms inter-stimulus interval allowed the subject to pick up 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 recording sheets. Each stimulus was repeated ten times in random order. In the first test in which the lower and higher frequency components were presented from different directions, 420 stimuli (42 kinds of stimuli  $\times$  10 times) were separated into five sessions. Each session was completed in approximately 9.5 minutes. In the second test in which both components were presented from the same direction, 70 stimuli were presented in one session. This test was completed in approximately 8 minutes. Each subject made a total of 490 judgments for the entire task.

## 3. EXPERIMENTAL RESULTS AND DISCUSSION

### 3.1. Distribution of Subject's Responses

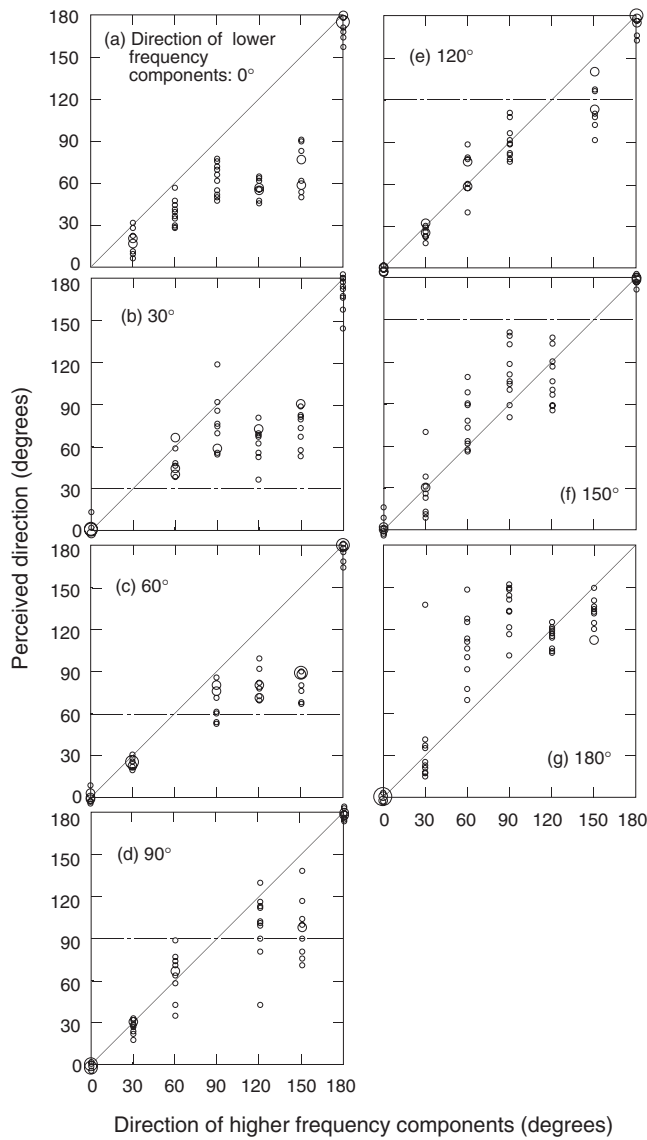
The direction marked by the subject was read with a protractor to an accuracy of one degree. Figure 3(a) shows the responses to the stimuli whose lower and higher frequency components were presented from the same direction. The diameter of each circle plotted is proportional to the number of responses. The ordinate of each panel is the perceived direction, and the abscissa is the source direction. Figure 3(b) shows the results obtained in the preparatory tests, in which normal wide-band noise stimuli (280–11,200 Hz) were presented from loudspeakers in the median plane. Comparing panels (a) and (b), no significant difference exists in the distributions of responses. This means that, even if a source signal is divided into lower and higher frequency components, all listeners can localize the sound images as accurately as the normal wide-band stimulus when both components are presented from the same direction.

Figures 4–6 show the responses for each subject to the stimuli for which the lower and higher frequency components were presented from different directions. Results are presented for the directions of the lower frequency components. The abscissa in each panel is the direction of the higher frequency components, and the dot-dashed line indicates the direction of the lower frequency components.



**Fig. 3** Distribution of responses. (a) shows the responses to the stimuli whose lower (280–4,800 Hz) and higher (4,800–11,200 Hz) frequency components were presented from the same direction in the median plane. (b) shows the responses to normal wide-band (280–11,200 Hz) noise stimuli in the median plane.

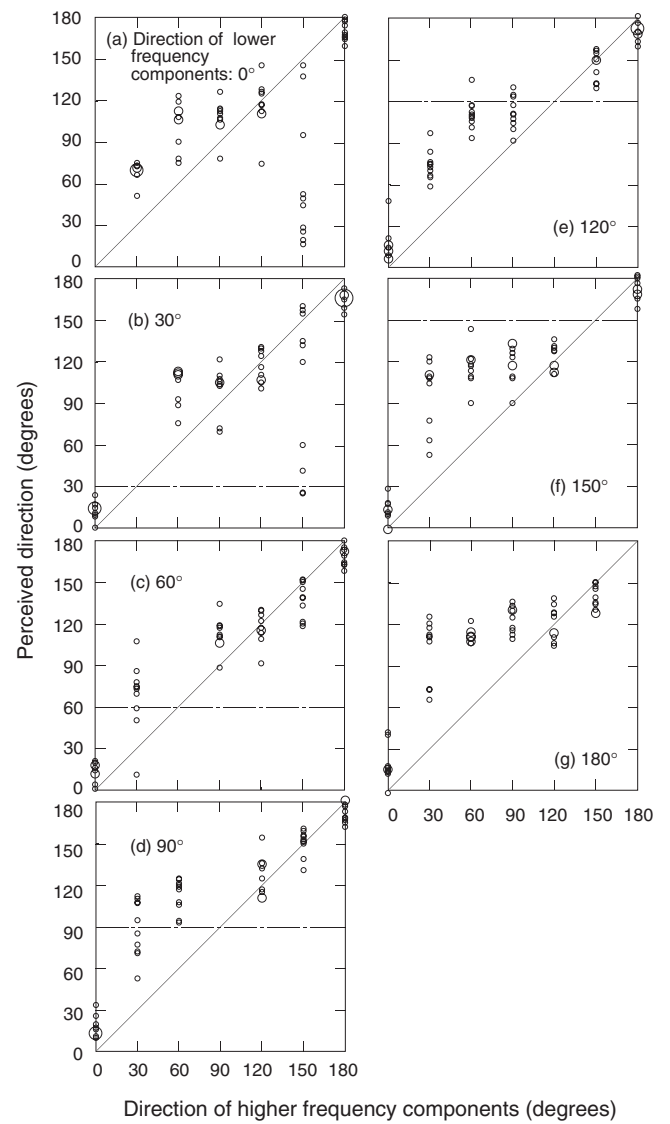
Subject SM (Fig. 4) perceived a fused sound image for each stimulus, even when the lower and higher frequency components of the stimulus were presented from different directions. When the lower frequency components were presented from  $0^\circ$  (panel (a)) and the higher frequency components were presented from  $90^\circ$ ,  $120^\circ$  and  $150^\circ$ , sound images were localized between the directions of the two components. When the lower frequency components were presented from  $30^\circ$ ,  $60^\circ$  and  $90^\circ$  (panels (b)–(d)) and the higher frequency components were presented from  $120^\circ$  and  $150^\circ$ , sound images were localized between the directions of the two components or near the direction of the lower frequency components. When the lower frequency components were presented from  $120^\circ$  (panel (e)) and the higher frequency components were presented from  $150^\circ$ , sound images were localized around the direction of the lower frequency components. When the lower frequency components were presented from  $180^\circ$  (panel (g)) and the higher frequency components were presented from  $60^\circ$  and  $90^\circ$ , sound images were localized between the



**Fig. 4** Responses to the stimuli whose higher and lower frequency components were presented from different directions in the median plane for subject SM.

directions of the two components. In all other cases, the perceived directions coincided approximately with the directions of the higher frequency components.

Subject TM (Fig. 5) also perceived a fused sound image for each stimulus. When the lower frequency components were presented from  $0^\circ$  and  $30^\circ$  (panels (a) and (b)) and the higher frequency components were presented from  $150^\circ$ , the perceived directions were distributed over all directions. When the lower frequency components were presented from  $60^\circ$  and  $90^\circ$  (panels (c) and (d)) and the higher frequency components were presented from  $30^\circ$ , sound images were localized around the direction of the lower frequency components. When the lower frequency components were presented from  $150^\circ$  and  $180^\circ$  (panels (f) and (g)) and the higher frequency components were presented from  $30^\circ$ ,  $60^\circ$  and  $90^\circ$ , sound



**Fig. 5** As Fig. 4 for subject TM.

images were localized between the directions of the two components. In all other cases, the perceived directions coincided approximately with the directions of the higher frequency components.

Subject TU (Fig. 6) sometimes marked down two directions on a recording sheet for the stimuli of which lower frequency components were presented from  $150^\circ$  and  $180^\circ$  (panels (f) and (g)) and higher frequency components were presented from  $0^\circ$ . According to his report, he perceived a sound image extended straight from the front to the rear, and so marked two directions. Nevertheless, the images were localized primarily in the front. Therefore, the responses near  $180^\circ$  (triangles in Fig. 6) are omitted from the examination. Most of the perceived directions coincided approximately with the directions of the higher frequency components.

Summarizing these results, although a few sound images were perceived as being shifted toward the

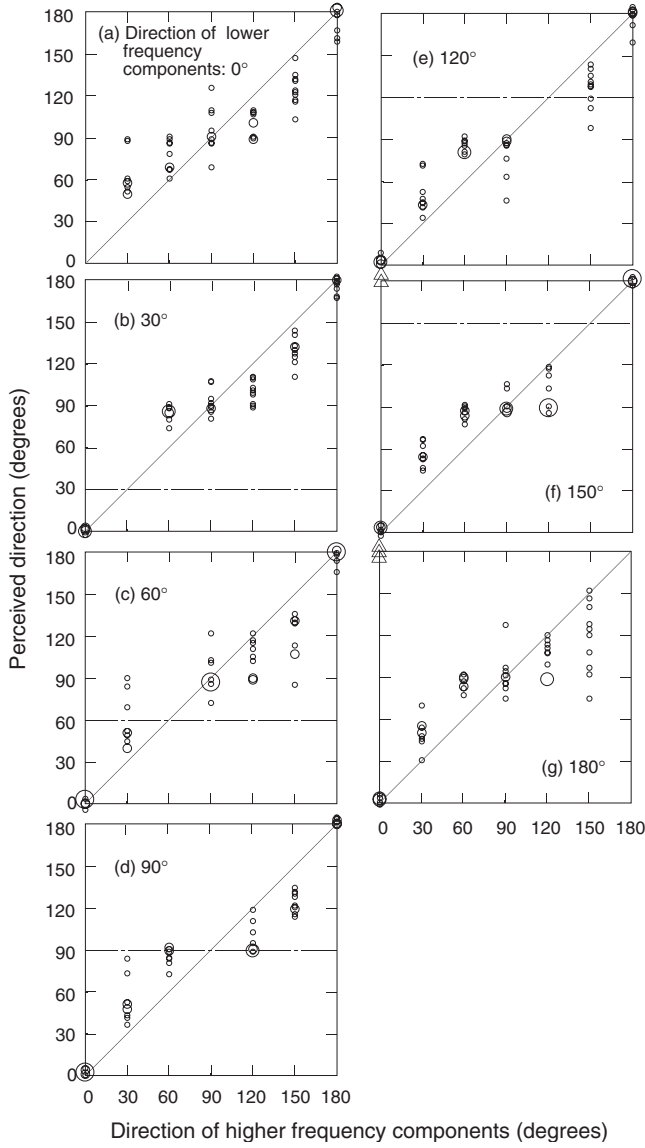


Fig. 6 As Fig. 4 for subject TU.

directions of the lower frequency components, most of the perceived directions agreed with the directions of the higher frequency components. In other words, when the source is a broadband signal, the higher frequency components are dominant in median plane localization and the contribution of the lower frequency components is slight.

### 3.2. Average error

Table 1 shows the localization error  $E$  [12] for each subject defined as:

Table 1 Averaged localization error  $E$  (degrees).

| Subject | $E_{\text{both}}$ | $E_{\text{high}}$ | $E_{\text{low}}$ |
|---------|-------------------|-------------------|------------------|
| SM      | 11                | 22                | 64               |
| TM      | 17                | 28                | 66               |
| TU      | 19                | 16                | 74               |

$$E = \frac{1}{J} \frac{1}{K} \sum_{j=1}^J \sum_{k=1}^K |S_{jk} - R_{jk}| \quad (1)$$

where  $S_{jk}$  indicates the presented direction,  $R_{jk}$  is the perceived direction, every  $j$  corresponds to a direction, and every  $k$  is the stimulus for each direction.  $E_{\text{both}}$  in Table 1 is the error when both components were presented from the same direction ( $J = 7, K = 10$ ), and  $E_{\text{high}}$  and  $E_{\text{low}}$  are the error to the direction of the higher and lower frequency components ( $J = 7, K = 60$ ), respectively, when the two components were presented from different directions. Clearly, for each subject,  $E_{\text{low}}$  is much larger than  $E_{\text{both}}$  or  $E_{\text{high}}$ , which are approximately equal. A factor analysis of variance indicates that both factors, type of error and subject, are significant ( $p < 0.01$ ). Moreover, the results of a t-test indicate significant differences ( $p < 0.01$ ) both between  $E_{\text{low}}$  and  $E_{\text{both}}$  and between  $E_{\text{low}}$  and  $E_{\text{high}}$ . This means that the higher frequency components are dominant in median plane localization, and that the contribution of the lower frequency components is slight.

In order to examine details of the experimental results, the localization error for each presented direction was calculated as follows:

$$E_j = \frac{1}{K} \sum_{k=1}^K |S_{jk} - R_{jk}| \quad (2)$$

where  $j$  is the presented direction. As described earlier,  $E_{\text{both},j}$  is the error when both components were presented from the same direction  $j$ , and  $E_{\text{high},j}$  and  $E_{\text{low},j}$  are the errors in the direction  $j$  of the higher and lower frequency components, respectively, when the two components were presented from different directions. T-tests were performed in order to determine whether the difference between  $E_{\text{both},j}$  and  $E_{\text{high},j}$  or that between  $E_{\text{both},j}$  and  $E_{\text{low},j}$  was significant ( $p < 0.01$ ). The results of these tests were classified into four cases.

Case A: The difference between  $E_{\text{both},j}$  and  $E_{\text{low},j}$  was significant and the difference between  $E_{\text{high},j}$  and  $E_{\text{both},j}$  was not significant. The sound images were assumed to be localized in the direction of the higher frequency components.

Case B: The difference between  $E_{\text{both},j}$  and  $E_{\text{high},j}$  was significant and the difference between  $E_{\text{low},j}$  and  $E_{\text{both},j}$  was not significant. The sound images were assumed to be localized in the direction of the lower frequency components.

Case C: Both of the differences were significant. Sound images were assumed to be localized in neither the direction of the higher frequency components nor in that of the lower frequency components.

Case D: Neither of the differences were significant. In this case, each components can be regarded as dominant in



**Table 2** Classification of significant difference between localization errors  $E_{\text{low,high}}$  and  $E_{\text{both}}$ .

| Subject | Direction of higher frequency components | Direction of lower frequency components |     |     |     |      |      |      |
|---------|--|---|-----|-----|-----|------|------|------|
|         |  | 0°                                      | 30° | 60° | 90° | 120° | 150° | 180° |
| SM      | 0°                                       |   | A   | A   | A   | A    | A    | A    |
|         | 30°                                      | A                                       |     | A   | A   | A    | A    | A    |
|         | 60°                                      | A                                       | D   |     | A   | A    | A    | C    |
|         | 90°                                      | C                                       | A   | D   |     | A    | A    | C    |
|         | 120°                                     | C                                       | C   | B   | D   |      | A    | A    |
|         | 150°                                     | C                                       | C   | B   | B   | D    |      | A    |
|         | 180°                                     | A                                       | A   | A   | A   | A    | D    |      |
| TM      | 0°                                       |   | D   | D   | A   | A    | A    | A    |
|         | 30°                                      | A                                       |     | D   | B   | A    | C    | C    |
|         | 60°                                      | A                                       | A   |     | B   | D    | C    | C    |
|         | 90°                                      | A                                       | A   | A   |     | D    | A    | C    |
|         | 120°                                     | A                                       | A   | A   | A   |      | A    | A    |
|         | 150°                                     | C                                       | A   | A   | A   | A    |      | A    |
|         | 180°                                     | A                                       | A   | A   | A   | A    | D    |      |
| TU      | 0°                                       |   | D   | A   | A   | A    | A    | A    |
|         | 30°                                      | A                                       |     | D   | A   | A    | A    | A    |
|         | 60°                                      | A                                       | A   |     | D   | A    | A    | A    |
|         | 90°                                      | A                                       | A   | D   |     | D    | A    | A    |
|         | 120°                                     | A                                       | A   | A   | D   |      | D    | A    |
|         | 150°                                     | A                                       | A   | A   | A   | A    |      | A    |
|         | 180°                                     | A                                       | A   | A   | A   | A    | D    |      |

sound localization; however, determination of the dominant components was not possible. This case often occurred when the directions of two loudspeakers were close to each other.

These results are summarized in Table 2. All of the results for subject TU are classified into either case A or case D, which means that this subject always localizes sound images in the direction of the higher frequency components. On the other hand, a few of the results for subject SM are classified into either case B or case C when the higher frequency components were presented from 60°, 90°, 120°, or 150°. A similar tendency is observed for subject TM when the higher frequency components were presented from 30°, 60°, 90° or 150°. For these directions, the responses scatter more widely than the other directions, even for normal wide-band stimuli, as shown in Fig. 3(b). Hence sound images might have easily shifted to the direction of the lower frequency components because spectral cues provided by the higher frequency components were not sufficiently effective. Nevertheless, a number of results are classified into case A, that is, the higher frequency components are thought of as having contributed to median plane localization more effectively in most cases.

### 3.3. Front-back Discrimination

Asano *et al.* [10] concluded that important cues for front-back discrimination exist in the frequency range below 2 kHz. Here, the contribution of the lower frequency

components to front-back discrimination is examined. Morimoto [7] and Kurosawa *et al.* [13] demonstrated that just noticeable differences in median plane localization are large when a white noise signal is presented from the above directions from 60° to 120°. Namely, this means that front-back confusion frequently occurs at these directions. Therefore, in the present paper, front-back discrimination is examined only when either lower or higher frequency components were presented from 0° or 30°, and the other components were presented from 150° or 180°. Note that a sound image perceived less than 90° is regarded to be in the front, and vice versa.

For subject SM, when the higher frequency components were presented from 150° and the lower frequency components were presented from either 0° or 30° (panels (a) and (b) in Fig. 4), most of the responses indicated the front. For subject TM, when the higher frequency components were presented from 30° and the lower frequency components were presented from 150° or 180° (panels (f) and (g) in Fig. 5), most of the responses indicated the rear. Moreover, when the higher frequency components were presented from 150° and the lower frequency components were presented from 0° or 30° (panels (a) and (b) in Fig. 5), half of all responses indicated the front. However, for both subjects, when the higher frequency components were presented from 0° or 180°, all responses coincided with the directions of the components. Moreover, for subject TU, the responses always coincided with the directions of the higher frequency components.

In summary, the lower frequency components contribute to front-back discrimination only slightly compared to the higher frequency components. Hence, important cues for front-back discrimination cannot be assumed to exist in the lower frequency components. This is not inconsistent with the results reported by Morimoto [7] and Algazi *et al.* [11], because these studies do not mention about the importance of low frequency components. However, this conclusion is obviously inconsistent with the conclusion of Asano *et al.*

Such inconsistency can be explained as follows: In the localization tests performed by Asano *et al.*, the subjects might have perceived the sound images inside their heads, because the stimuli were presented to the subjects through headphones. Laws [14] demonstrated that sound images were localized inside the head when stimuli were presented through headphones without accurate compensation for the transfer function from the headphones to the ears. It is possible that the smoothed HRTF used by Asano *et al.* corresponds to providing stimuli without such compensation. When the subject perceives sound images inside his or her head, the sound images are usually perceived in the back part of the head. The subjects in the tests performed by Asano *et al.* perceived most of the sound images in the rear when low frequency components were smoothed (see Figs. 5 and 8 in [10]).

#### 4. CONCLUSIONS

Localization tests were performed in order to examine the contribution of low frequency components to sound localization in the median plane. In these tests, the higher (above 4,800 Hz) and lower (below 4,800 Hz) frequency components of stimuli are presented simultaneously from different directions. The results indicate that: (1) when the source is a wide-band signal, the higher frequency components are dominant in median plane localization, whereas the lower frequency components do not contribute

significantly to the localization, and (2) important cues for front-back discrimination do not exist in the low frequency range.

#### REFERENCES

- [1] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT, Cambridge, MA, 1997).
- [2] J. Blauert, "Sound localization in the median plane," *Acustica*, **22**, 205–213 (1969/70).
- [3] J. Hebrank and D. Wright, "Spectral cues used in the localization of sound sources on the median plane," *J. Acoust. Soc. Am.*, **56**, 1829–1834 (1974).
- [4] R. A. Butler and K. Belendiuk, "Spectral cues utilized in the localization of sound in the median sagittal plane," *J. Acoust. Soc. Am.*, **61**, 1264–1269 (1977).
- [5] P. J. Bloom, "Creating source elevation illusions by spectral manipulation," *J. Audio Eng. Soc.*, **25**, 560–565 (1977).
- [6] N. Asahi and S. Matsuoka, "Effect of the sound anti resonance by pinna on median plane localization—Localization of sound signal passed dip filter—," *Tech. Rep. Hear. Acoust. Soc. Jpn.*, H-40-1 (1977).
- [7] M. Morimoto, "On sound localization," *Dissertation*, Tokyo University (1982).
- [8] J. C. Middlebrooks, "Narrow-band sound localization related to external ear acoustics," *J. Acoust. Soc. Am.*, **92**, 2607–2624 (1992).
- [9] S. Mehrgardt and V. Mellert, "Transformation characteristics of the external human ear," *J. Acoust. Soc. Am.*, **61**, 1567–1576 (1977).
- [10] F. Asano, Y. Suzuki and T. Sone, "Role of spectral cues in median plane localization," *J. Acoust. Soc. Am.*, **88**, 159–168 (1990).
- [11] 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).
- [12] M. Morimoto and Y. Ando, "On the simulation of sound localization," *J. Acoust. Soc. Jpn. (E)*, **1**, 167–174 (1980).
- [13] A. Kurosawa, T. Takagi and Z. Yamaguchi, "On transfer function of human ear and auditory localization," *J. Acoust. Soc. Jpn. (J)*, **38**, 145–151 (1982).
- [14] P. Laws, "On the problem of distance hearing and the localization of auditory events inside the head," *Dissertation*, Technische Hochschule, Aachen (1972).