PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Spatial sound segregation in the monaural listening condition

Daisuke MORIKAWA[1]; Daiki KOJIMA[1]; Tatsuya HIRAHARA[1]

[1] Toyama Prefectural University, Japan

ABSTRACT

A sound image revolves around the head with head rotation in the monaural listening condition. The sound image moves by a large angle around the listener's head when the sound source moves from the open-ear side to the blocked-ear side, and vice versa, based on the listener's head rotation. Sounds presented from a certain direction actually jump from rear to front or from the center line to the left/right side of a listener's head and vice versa, and a listener perceives segregated sound images at a certain head rotation angle. We measured the sound image trajectory of a fixed sound source with head rotation while listening under the monaural listening condition. Results show that sound image segregation tends to occur when the sound source crosses the boundary between the open-ear and blocked-ear sides as a result of head rotation. Two sound image locations calculated from the direct and diffracted waves along the head from the sound source can cause spatial sound image segregation under the monaural listening condition.

Keywords: Monaural, Sound image trajectory, Sound segregation

1. INTRODUCTION

The interaural time difference (ITD), interaural level difference (ILD), and spectral cue (SC) are well-known important cues for sound localization (1). However, ITD and ILD cannot be cues in the monaural listening condition because they are binaural cues. SCs are cues that can be used to localize sound in the monaural listening condition. Unilateral deaf persons (2) or normal-hearing persons with one ear blocked by an earplug (3) can actually localize sound sources on the horizontal plane of the available-ear side, suggesting that the SCs of the ear can be cues to localize sound when ITD and ILD are unavailable.

The temporal variations of the ITD, ILD, and SCs produced by head rotation are also well-known significant contributors to the improvement of binaural sound localization accuracy (4, 5, 6). Thus, the temporal variation of SCs produced by head rotation is also expected to improve the accuracy in monaural sound localization. However, monaural sound image localization experiments under the head rotation condition have shown that the sound image flies around the head with head rotation in the monaural listening condition. The sound image moves by a large angle around the head when the sound source moves from the open-ear side to the blocked-ear side, and vice versa, with head rotation (7, 8). All listeners perceive the sound image flying around the head with head rotation when the stimulus sound is presented from a certain direction. The sound image jumps from rear to front or from the center line to left/right side of a listener's head, and vice versa. Furthermore, some listeners perceive two segregated sound images at a certain head rotation angle.

This study describes this spatial sound segregation phenomenon in the monaural listening condition and discusses its cause.

2. METHODS

2.1 Experimental system

Figure 1 shows the experimental system. The sound reproduction system used herein comprised a Windows-based PC, two digital-to-analog converters (DACs; RME, Fireface UFX), 12 power amplifiers (ONKYO, CR-N755), and 12 loudspeakers (Vifa, MG10SD-09-08). The sampling frequency of the DACs was 48 kHz. The loudspeakers were placed around a chair centered on a horizontal circle of 1 m radius at 30° intervals. The loudspeakers were mounted at a height of 1.1 m.

[1] dmorikawa@pu-toyama.ac.jp

The answering system comprised a tablet-type device (apple, iPad), a stylus (Just Mobile, JTM-PD-000019), and a Mac-based PC. The tablet-type device has an answer interface function that sends answered coordinate values to the Mac-based PC via wireless UDP/IP communication. The communication protocol used was Open Sound Control (OSC). The answering interface was developed using an Mrmr OSC controller (9). The Mac-based PC received the coordinate values. Figure 2 shows the graphical user interface (GUI) of the answering interface on the tablet. The listener drew the sound image trajectory and the presence or absence of sound image segregation using the GUI. A discontinuous trajectory of the sound image was drawn when sound image segregation occurred.

A listener could push the "next" or "repeat" buttons after the sound image trajectory of the stimulus was drawn. When the listener pressed the "next" button, a trigger signal was transmitted from the Mac-based PC to the Windows-based PC via a serial port, and the next stimulus was presented. When the listener pressed the "repeat" button, another trigger signal was transmitted, and the same stimulus was presented again. The listener was allowed to listen to the stimulus as many times as he/she wanted until a satisfying trajectory was drawn.

2.2 Stimulus

Broadband noise (i.e., Gaussian-distributed random noise) was used as the stimulus in the experiment. The duration of each stimulus was 5 s, and a 30-ms linear taper window was applied at its beginning and end. The A-weighted sound pressure level of the stimulus L_A was set to 40 dB at the head center position. This low-sound-pressure-level stimulus was used to ensure the monaural listening condition. All normal-hearing listeners wearing earplugs (Moldex, Pure-Fit 6800s) in both ears could not hear the stimulus at this level. The earplug attenuated sound by 30 to 50 dB from 100 Hz to 20 kHz; however, some listeners wearing earplugs on both ears could still hear the stimulus presented at L_A = 60 dB or 50 dB, possibly because of incomplete/ill-fitting of the earplugs in the ear-canals.

The experiment was performed in an experimental room with the walls and the ceiling covered with sound-absorbing materials. The A-weighted background noise level of the room was 23 dB. Figure 3 shows the stimulus spectrum, the background noise spectrum of the room, the hearing thresholds (ISO 226-2003), and the hearing threshold with an earplug, considering the vertical axis as the sound pressure level, ref. 20 μPa.

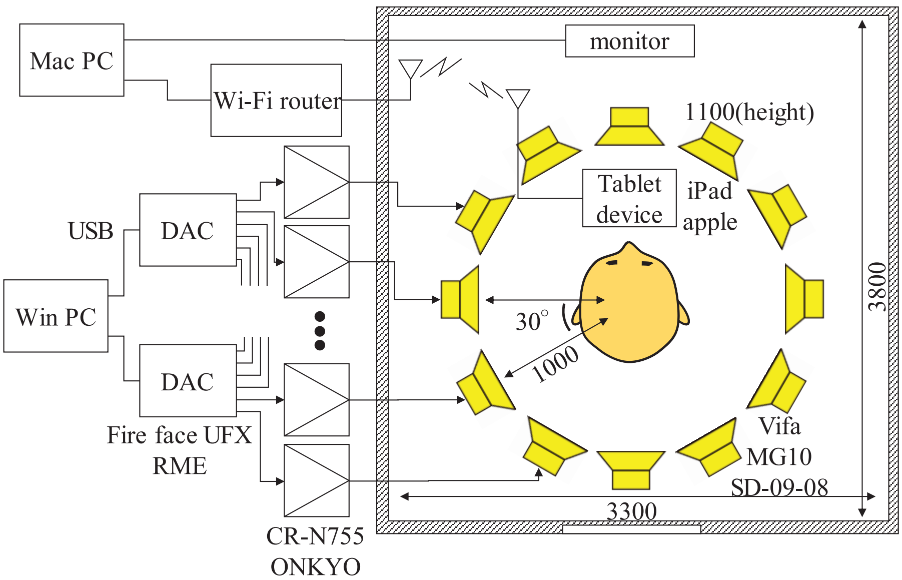

Figure 1 – Experimental system.

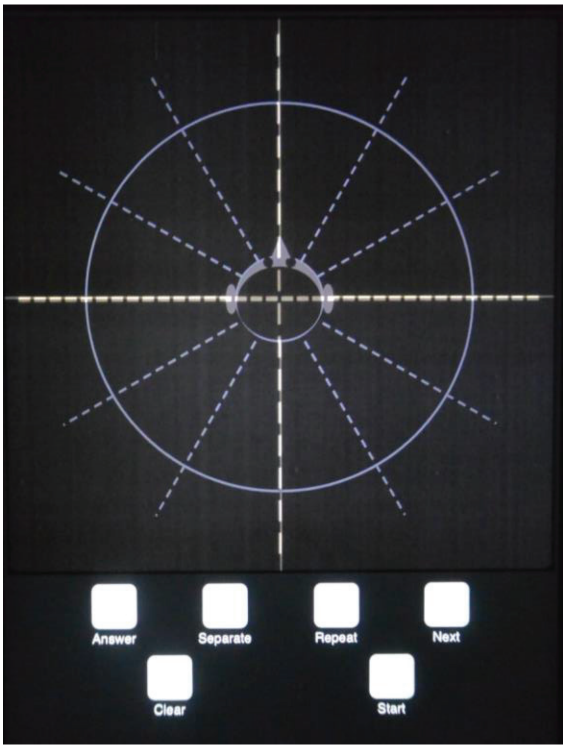

Figure 2 – Answering interface.

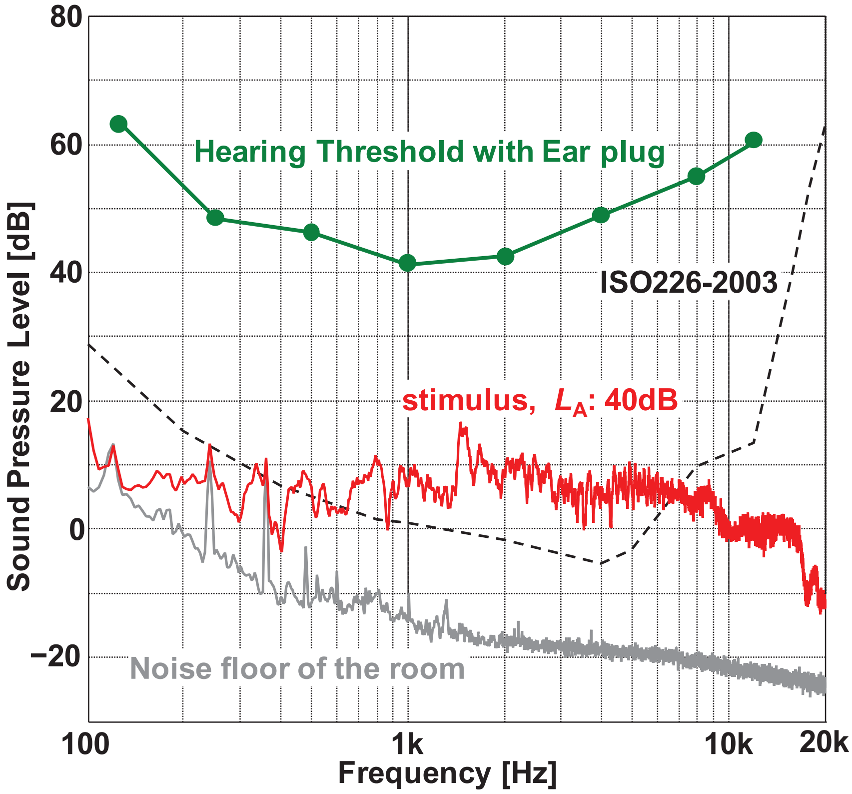

Figure 3 – Stimulus spectrum, background noise spectrum of the room, hearing thresholds (ISO 226-2003), and hearing threshold with an earplug.

2.3 Procedure

First, a listener put ear plugs in both ears and sat on the chair at the center of the circular speaker array. Second, the experimenter confirmed that the listener cannot hear the stimuli presented from the sides of either the left or the right ear (90° or 270°). The experimenter then removed one ear plug from the left or right ear, making that side the normal-hearing side. This procedure was performed to guarantee the monaural listening condition.

The listener under the monaural listening condition was instructed to rotate his/her head by approximately 60° to the left or right from the front when the stimulus was presented. Pitch and roll head rotations were not allowed. In order to ensure that the listener's head rotation was confined to a yaw angle within an acceptable range, the Euler angle of the listener's head was monitored using a motion sensor (Ascension, Flock of Birds or Logical Product, LP-WSD004-0A) mounted on the top of the listener's head. The listener's task was to draw a sound image trajectory on the tablet-type device, after listening to the stimulus while turning his/her head. The listener could listen to the stimulus as many times as needed to obtain a reliable sound image trajectory. Thus, the depicted sound image trajectory data were not linked to the head rotation angle data of the listener being monitored. In addition, the listeners were instructed to place the tablet-type device on their laps while the stimulus was being presented, to avoid producing a reflected wave.

Thirteen normal-hearing listeners participated in the experiments. All listeners participated in both right-ear blocked and left-ear blocked condition. They had experience in binaural sound localization experiments with static as well as dynamic head movement conditions.

3. RESULTS

Eleven listeners reported that sound image segregation occurred at some stimulus locations, whereas the remaining two listeners reported that no sound image segregation occurred at any of the stimulus locations. Figure 4 shows examples of the sound image trajectories of a typical listener who perceived the sound image segregation.

When a listener with a blocked right ear turned his head counterclockwise, he perceived that the sound image discontinuously moved around the head from the front to the rear with the stimulus presented from the frontal (θs = 0°) loudspeaker. He perceived two segregated sound images at approximately −30° and −120° when he turned his head leftward (Fig. 4(a)).

When a listener with a blocked right ear turned his head clockwise, he perceived a sound image that jumped from rear to front with the stimulus presented from the rear (θs = 180°) loudspeaker. He perceived two segregated sound images at approximately −15° and −165° when he turned his head rightward (Fig. 4(c)).

Furthermore, when a listener with a blocked left ear turned his head clockwise, he perceived a sound image that jumped from rear to front with the stimulus presented from the right-front (θs = 30°) loudspeaker. He perceived two segregated sound images at approximately 60° and 140° when he turned his head rightward (Fig. 4(b)).

When a listener with a blocked left ear turned his head counterclockwise, he perceived a sound image that jumped from rear to front with the stimulus presented from the left-front (θs = −60°) loudspeaker. He perceived two segregated sound images at approximately 75° and 25° when he turned his head leftward (Fig. 4(d)).

In this manner, 11 listeners perceived two segregated sound images when they rotated their heads to the angle where the sound image had just jumped. When sound segregation occurred, the sound image jumped from rear to front or from the center to the left/right side, and vice versa.

Figure 5 summarizes the number of responses in which the listeners perceived sound image segregation in each stimulus direction for each of the two head rotation directions and two ear occlusion conditions. The abscissa of each panel represents the azimuth of the stimulus presented. The azimuthal angle was 0° at the front of a listener and positive, clockwise.

When the listeners rotated their heads toward the open-ear side, most listeners perceived sound image segregation with the stimulus presented at 0° ± 30° (Figs. 4(a) and (b)). In contrast, when the listeners rotated their heads toward to the blocked-ear side, most listeners perceived sound image segregation with the stimulus presented at 60° to 120° with the blocked right ear (Fig. 4(c)) and 270° and 300° with blocked left ear (Fig. 4(d)).

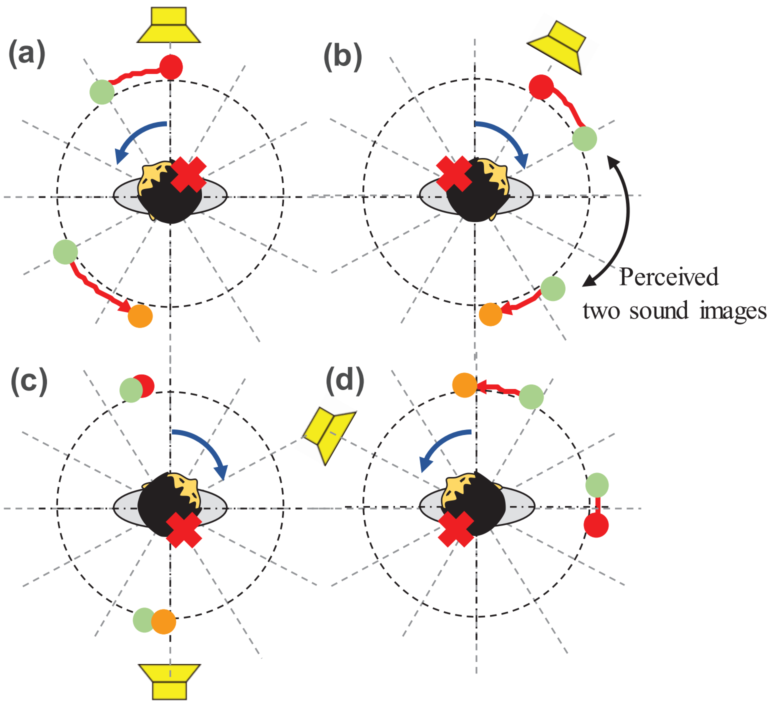

Figure 4 – Examples of the sound image trajectory. (Red dot: sound image position at start, Green dot: perceived two sound image, Orange dot: Sound image position at the end)

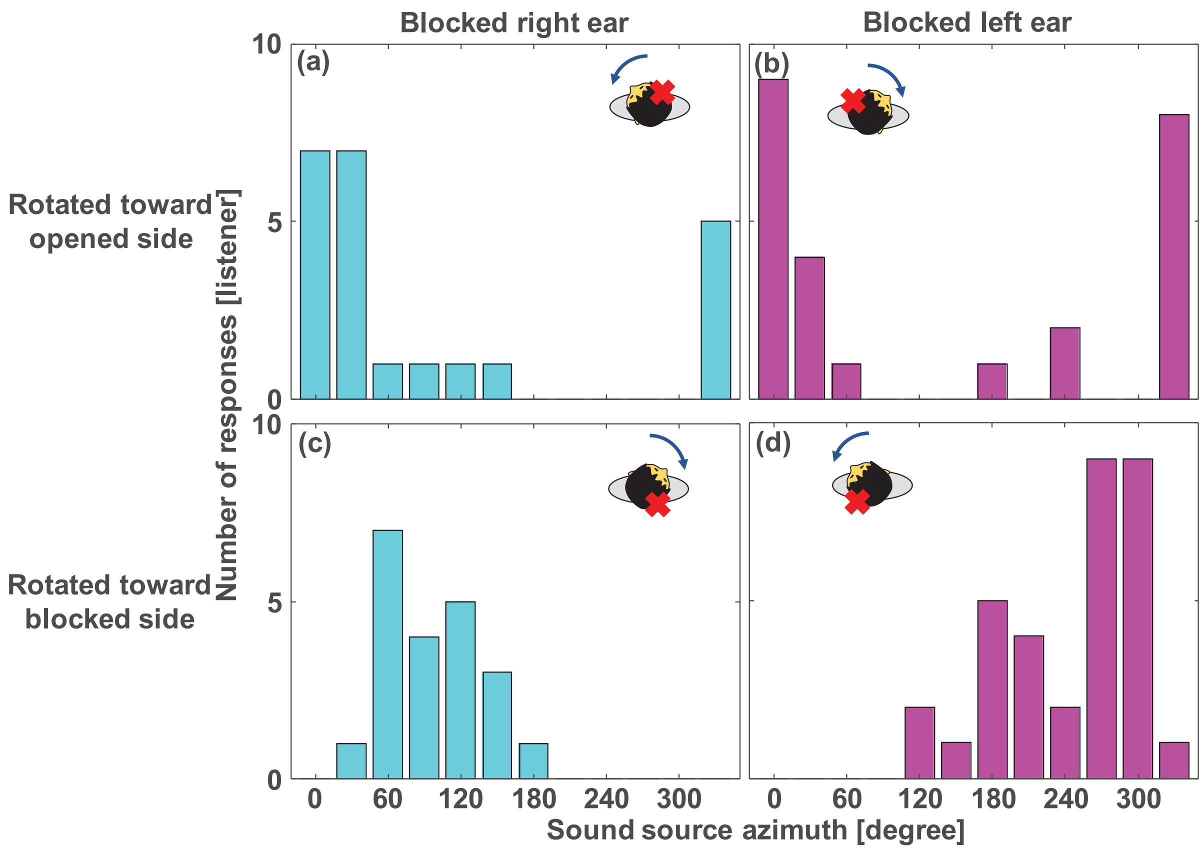

Figure 5 – Number of responses in which the listeners perceived sound image segregation. (a) Rotated to the open-ear side with a blocked right ear; (b) rotated to the open-ear side with a blocked left ear; (c) rotated to the blocked-ear side with a blocked right ear; and (d) rotated to the blocked-ear side with a blocked left ear.

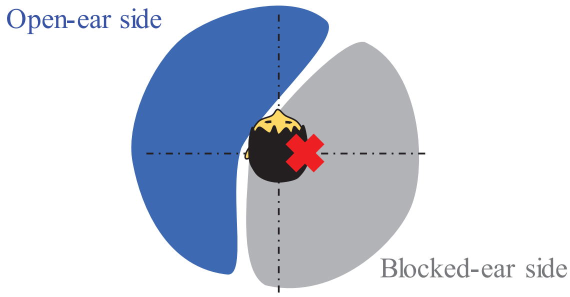

Figure 6 – Open- and blocked-ear sides

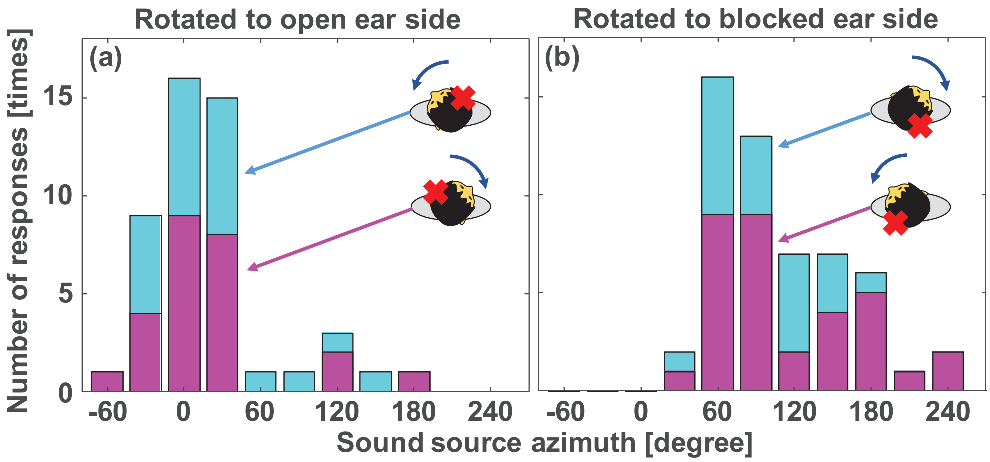

Figure 7 – Sum of the number of sound image segregation occurrences when the direction of the blocked-ear side was positive. (a) Rotated to the open-ear side and (b) rotated to the blocked-ear side.

4. DISCUSSION

The results showed that sound image segregation mostly occurred when the sound source direction crossed the boundary between the open- and blocked-ear sides (Figure 6) as a result of head rotation.

Assuming left-right symmetry in the shape of the head and pinnae, the stimulus direction observed from the open-ear would have been the same under the following two conditions: turning the head clockwise with an open left ear, and turning the head counterclockwise with an open right ear. Thus, each of the two pairs of Figures 5 (a) & (b) and (c) & (d), can be combined into one figure respectively by using the abscissa of the stimulus azimuth between −60° and 240°. Figures 7(a) and (b) show the sum of the number of sound image segregation occurrences for each stimulus azimuth when the listeners turned their heads to the open and blocked-ear sides.

As shown in Fig. 7, sound image segregation tended to occur with stimulus presented at 0° and ±30° when the head was turned toward the open-ear side and at ±60° and ±90° when the head was turned toward the blocked-ear side. The stimulus azimuths that provided a large number of image segregation occurrences differed by 60° or more between the two head-turning directions.

The sound presented at the front (0°, ±30°) transitioned from the open-ear side to the blocked-ear side when the head was rotated to the open-ear by approximately 30° or less. The sound presented at the blocked-ear side (60°, 90°) transitioned from the blocked-ear side to the open-ear side when the head was rotated to the blocked-ear side.

These facts suggest that front–back sound segregation occurs when sound is presented at the boundary between the open- and blocked-ear sides as a result of head rotation. Under the monaural listening condition, the sound image of a sound source in the open- and blocked-ear sides is localized in the open-ear side. Therefore, the perceived sound image jumps when a sound source location passes across the boundary between the open- and blocked-ear sides according to the head rotation, and two sound images are simultaneously perceived when the relative sound source location with respect to the head is on the boundary. This sound segregation cannot be considered as based on the cone of confusion, on which ITD and ILD adopt the same value, because binaural cues are unavailable under the monaural listening condition. This sound segregation can be considered to be based on the application of direct waves to the open ear and diffracted waves along the head to the open ear from the sound source. One sound image location is that computed from the sound source's direct wave convolved with the head-related transfer function (HRTF) of its incident angle to the open-ear. Another sound image location is that computed from the diffracted waves along the head convolved with the HRTFs of their incident angles to the open-ear.

Additionally, sound image segregation occurs more frequently with a stimulus presented between 120° and 180° when the head is rotated toward the blocked-ear side, rather than a stimulus presented between 60° and 120° when the head is rotated toward the open-ear side. The difference between the two head rotation directions appears to depend on whether the open-ear is in the front of or above the shoulder. The reflection from the shoulder might affect the perception of the sound image segregation.

The range of the open- or blocked-ear side must be vague because it depends on the sound pressure level of the stimulus and the hearing threshold at each frequency. The range and the boundaries should vary among listeners because the size and shapes of the head and the pinnae depend on the individual. Measurements of the HRTF and the absolute threshold of hearing of each listener are necessary to clarify the range of each listener. These remain as future work.

The head rotation angle at which the sound image segregation occurred was not measured herein because drawing the sound image trajectory on a tablet in real time was difficult while listening to the stimuli when the head is rotating. However, the measurement of head rotation angle along with the sound image trajectory is necessary to verify the relative angles of the head direction and the sound source location at which the sound image segregates. This will also be investigated in future work.

5. CONCLUSIONS

Herein, we measured the sound image trajectory of a fixed sound source with head rotation while listening under the monaural listening condition. The results showed that sound image segregation tended to occur when the sound source was located on the boundary between the open- and blocked-ear sides according to the head rotation. Two sound image locations calculated from the direct and diffracted waves along the head from the sound source can cause spatial sound image segregation under the monaural listening condition.

ACKNOWLEDGEMENTS

Part of this work was supported by KAKENHI (17K00244).

REFERENCES

1. Blauert J. Spatial Hearing. The MIT press, Cambridge, MA, 1997.
2. Takahashi K, Morikawa D. Horizontal Localization of Sound Images and Sound Sources for Monaural Congenital Deafness. Journal of Signal Processing 2017;21(4):167-170.
3. Wightman FL, Kistler DJ. Monaural sound localization revisited. J Acoust Soc Am. 1997;101(2):1050-1063.
4. Wallach H. The role of head movements and vestibular and visual cues in sound localization. J.Exp.Psychol. 1940;24:339-368.
5. Morikawa D, Hirahara T. Effect of head rotation on horizontal and median sound localization of band-limited noise. Acoust. Sci, & Tech. 2013;34(1):56-58.
6. Hirahara T, Sawada Y, Morikawa D. Sound Localization of Dynamic Binaural Signals Provided Using a Pinna-Less Dummy Head or a Stereo Microphone, Interdisciplinary Information Sciences. 2015;21(2):159-166.
7. Kojima D, Hirahara T. Monaural horizontal sound localization. IEICE Technical Report 2015;115(359):31-36. (in Japanese)
8. Hirahara T, Kojima D, Morikawa D. Monaural sound localization under conditions of active head rotation. IEICE Technical Report 2017;117(328):47.
9. 10base-t Interactive, Mrmr. http://ecumedesjours.com/Mrmr/ (accessed 22 May 2019).