PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Sound image localization by bone conduction sound

Kazuki TAKAI [1]; Takumi ASAKURA[1]

[1] Tokyo University of Science, Japan

ABSTRACT

An efficient voice-guidance technique by using bone conduction sound. By adopting binaural reproduction using the head-related transfer function, the playback sounds for guidance are reproduced to the users in order to indicate the traveling direction. However, the binaural reproduction technique by the bone conduction sound has not been established yet compared with the air-conducted sound. Therefore, in this study, a sound image localization experiment was carried out using a bone conduction sound reproduction method in which the intended direction of the reproduced sound images are quickly panned to left and right in order to enhance the perception of the sound direction. The results of the sound localization experiment with the panning angle of 10, 20, 30 degrees were parametrically discussed.

Keywords: Bone conduction sound, Head-related transfer function, Sound image localization

1. INTRODUCTION

Recently, the bone conduction sounds have been applied in various kinds of fields, such as the hearing aid technology [1] and human navigation [2, 3]. When a person utilizes bone conduction sound techniques, the excited vibration directly or indirectly transmits to the inner ears via complex path including the air-conducted and bone conduction transmission, and the person finally recognize the stimulation given by the actuator as sound. As for the bone conduction sound, many kinds of researches are performed [1-7, 10]. Tjellström et al. [1] investigated the effectiveness of an implanted bone conduction actuator (BAHA, the bone-anchored hearing aid). Such an implanted device contributes to efficient bone conduction transmission, however, it is hard to take the advantage of the bone conduction reproduction from the viewpoint of a person who has no difficulty in hearing. Walker et al. investigated navigation system for virtual environments by using bonephones [2]. From the viewpoint of more physical aspects, Purcell et al. [4] estimated the bone conduction transfer functions using the optoacoustic emissions. Chang et al. [5] developed a whole-head human finite-element models, and simulated the transmission of the bone conduction sound. McBride et al. [6] investigated the head sensitivity mapping for the bone conduction reception by a human head. Qin et al. [7] performed experimental study on the frequency characteristics of loudness of some kinds of bone conduction actuators.

On the other hand, from the viewpoint of subjective perception of sounds, various knowledge concerning three-dimensional sound field reproduction technology, especially in the field of air-conducted sound reproduction, has been constructed [8, 9]. In contrast, as for the bone conduction sound reproduction, such a three-dimensional reproduction technique has not yet been established whereas subjective evaluation for loudness of the bone conduction sounds are performed [10].

One of the advantages of using the bone conduction sound is that it can reproduce sound without blocking the ear canals, such as the reproduction using earphones or headphones. By taking such an advantage, it is possible to provide various kinds of voice information such as the route navigation for indoor/outdoor movement of a person without preventing smooth verbal communication with surrounding persons. Furthermore, it is also important not to hinder listening to environmental sounds such as the sounds of approaching vehicles by blocking the ear canal. Under such a situation with usage of bone conduction sound technique, reproduction of not only easy-to-hear clear sounds but also localized sounds which can indicate the direction of the navigation may enhance usefulness of the sound guidance technology.

In this research, for the purpose of constructing three-dimensional sound reproduction technique by bone conduction sound reproduction, a fundamental study on sound localization accuracy in bone conduction sound reproduction using the head related transfer functions (HRTF) measured by general air-conducted sound incidence into a head and torso simulator was performed.

[1] 7518533@ed.tus.ac.jp, t_asakura@rs.tus.ac.jp

2. DETAIL OF BONE CONDUCTION SOUND REPRODUCTION SYSTEM

2.1 Utilized device and contact condition between the actuator and the surface of the head

Figure 1 shows (a) adopted actuator utilized to the bone conduction reproduction, (b) the investigated contact positions between the actuator and the surface of each part of the head, and (c) the hairband-type fixing device of the actuator to the surface of the head. In this study, two actuators are attached to the left and right side of the face, respectively. The contact position between the actuator and the head surface were varied into four conditions including the auricular cartilage (P1), two points around the zygomatic arch (P2, P3), and the mastoid (P4) as shown in Fig. 1 (b-1) and (b-2), respectively. These points were chosen because they were often used as the excitation points of bone conduction reproduction [6]. As a fixing device of the actuators, four kinds of hairband-type fixing devices as shown in Fig. 1 (c) were adopted. These devices may have different contact pressures between the actuator and the surface of the face, because they have different widths of hairband □ as shown in Fig. 1(c), whereas they have similar dimensions and are made of similar resin-treated materials. The measured results of the contact pressures for each hairband-type device are described in the next section.

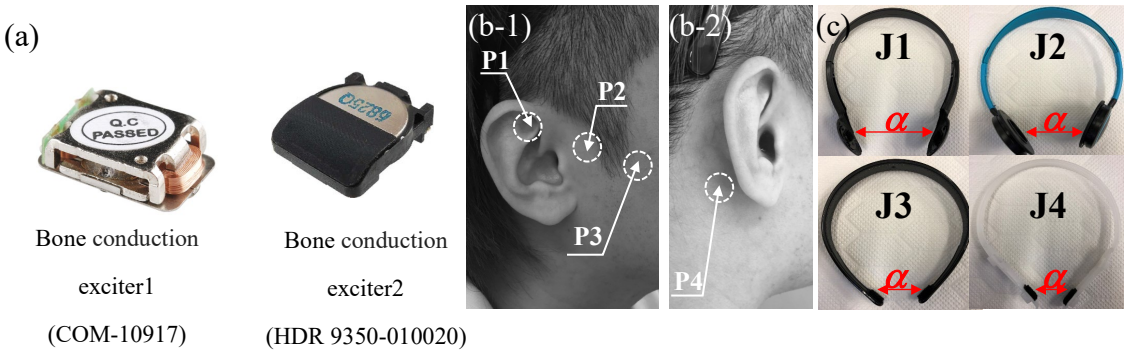

Fig. 1 Utilized device and contact condition.

2.2 Measurement of the excitation characteristics of the actuator

The contact pressure between the device of actuator and the surface of the subject's head may affect the frequency characteristics of the excitation force by the actuator. To investigate the transmission efficiency of vibration between the device and the contacting head part, the contact pressures caused by wearing the four kinds of fixing devices from J1 to J4 were measured by using a simple acrylic model imitating the head as shown in Fig. 2. The distance between the acrylic plates in this model was set to 15.7 cm of the average of young male and female face widths [11]. A thin gel sheet (thickness 1 mm, hardness 15 deg.) imitating human skin was sandwiched between the actuator and the acrylic plate. The contact pressure was measured using a dial tension gauge (TECLOCK, DTN-300). The measured results of the contact pressures for each hairband-type device are shown in Fig. 3. As shown in the figure, the contact pressures between the actuator and the surface of the acrylic plate were increased in accordance with the narrowness of the hairbands.

Next, the actuator was fixed to the surface of the abovementioned gel by the hairband-type devices of J1, J2, J3, and J4, and a white noise was generated with the same driving force acted to the actuator fixed by each hairband-type device. The vibration acceleration transmitted to the acrylic plate was measured by an acceleration pickup (Bruel & Kjaer, Type-4518) attached to the backside of the acrylic plate. The contact pressure and vibration characteristics measured in the condition from J1 to J4 are shown in Fig. 4. Although these frequency characteristics have peak and dip tendencies including the resonant mode frequencies of the contacting acrylic thin plates, the general tendencies of the difference between the excitation components by each bone conduction device can be discussed from the obtained results. In the conditions of J3 and J4 with relatively higher contact pressures, a higher vibration acceleration level was observed in relatively high frequency ranges compared to the conditions of J1 and J2 with relatively lower contact pressures.

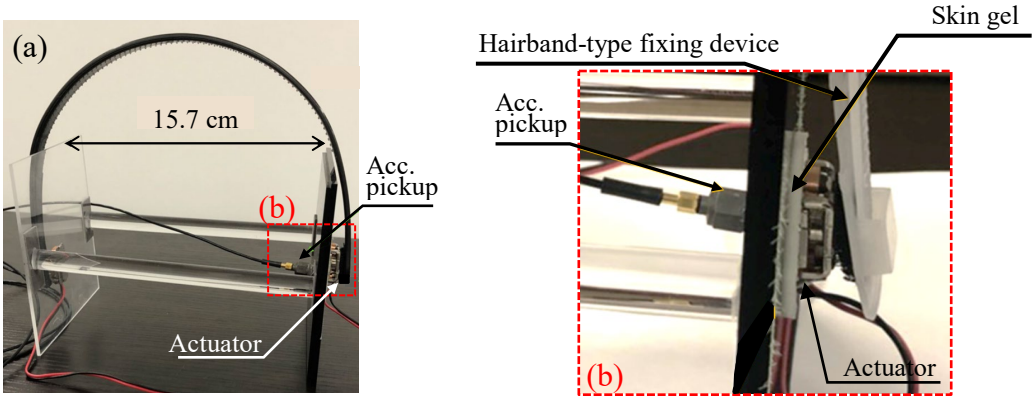

Fig. 2 (a)Simple acrylic model imitating the head, (b)Details.

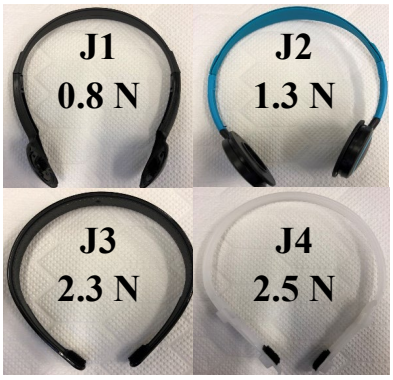

Fig. 3 Measured results of the contact pressures for each hairband-type device.

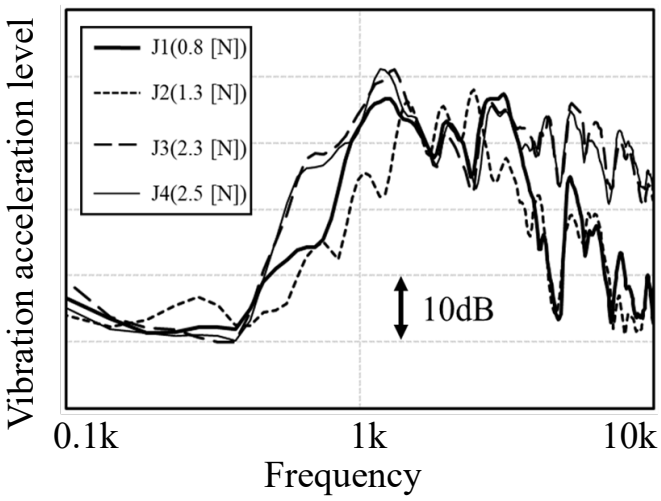

Fig. 4 Results between contact pressures and vibration characteristics.

3. EXAMINATION OF SUITABLE CONTACT POSITION OF ACTUATOR

To investigate the ease of hearing which is considered to be largely influenced by the contact position between the actuator and the face, a subjective experiment for evaluation of the ease of hearing of white noise, octave band noises, and female voices was performed by changing the contact positions from P1 to P4. It should be noted this experiment adopted the fixing device of J4 as shown in Fig. 1(c). As the participants in this experiment, seven males in their twenties were adopted. The experiment is composed of two parts. The first part investigates the subjective influence by the white noise and the octave band noises to the ease of hearing, whereas the second one investigates that by the female voice and the filtered sounds of that voice into each oct. band. The ease of hearing was

judged by using mono-polar five categories from “Considerably easy to hear” to “Considerably difficult to hear”.

Next, the scheme of the experiment is as follows. In this experiment, the subjects wore both the inner-ear type headphone and the bone conduction device in order to evaluate the relative extent of “ease of hearing” in comparison between the bone conduction and air-conducted sounds. It should be noted that the subjects keep evaluating all the conditions with both the inner-ear type headphone and the bone conduction device attached. This scheme was adopted in order to reduce the subjective influence of the detachment operation on the evaluated scores as much as possible. Firstly, the output volume of the bone conduction reproduction was determined so that the subjective loudness of the reproduced white noise including the frequency components from 500 Hz to 8 kHz bands by the bone conduction device was equivalent to that by the inner-ear type headphone. For this adjustment, following procedure was performed. The subject hears the reference air-borne sound of the abovementioned band-limited white noise from the inner-ear type headphone. Then, the subject can keep hearing the sound for enough time to judge the loudness of the sound. After that, the subject hears the same sound from the bone conduction device, and control the output volume by turning the volume controller of the audio interface. It should be noted that the subject can repeat the abovementioned adjustment for arbitrary times. The adjusted output volumes into the channels for bone conduction and air-borne actuators were fixed throughout the experiment including the first and second parts. Secondly, to familiarize the subjects with this experiment, three randomly chosen conditions were firstly evaluated by the subjects before the main experiment. It should be noted that the results of the trial evaluation were excluded from the analysis and discussion. Then, the ease of hearing of the abovementioned signal of the white noise, and octave band noises of 500 Hz, 1 kHz, 2 kHz, 4 kHz, and 8 kHz oct. bands, which were generated by filtering the white noise, were judged on the five categories. In each evaluation, the air-conducted sound was firstly reproduced, and after 5 second interval, the bone conduction sound was reproduced. Then, the subject answers the evaluated score for the bone conduction sound so that the score of the air-conducted sound is relatively evaluated as “3”. Lastly, the ease of hearing of a female voice spoken in Japanese with duration of 5 seconds, and the filtered signal of that voice into 500 Hz, 1 kHz, 2 kHz, 4 kHz, and 8 kHz oct. bands were judged on the five categories. In this part of the evaluation, the air-conducted sound was firstly reproduced, and after 5 second interval, the bone conduction sound was reproduced. Then, just the same as the procedure of the first part, the subject answers the obtained score for the bone conduction sound so that the score of the air-conducted sound is relatively evaluated as “3”. In these experiments, the air-conducted test sounds were reproduced with the inner-ear type headphone of ER-4S made by Etymotic research. It should be noted that the adopted inner-ear type headphone has relatively high sound insulation, and completely blocks the canal of the subject from the incoming sounds to their ears. The results of the first and second parts of the experiment are shown in Fig. 5 (a) and (b), respectively. It should be noted that the results of obtained by reproducing the white noise and the female voice including the frequency component from 500 Hz to 8 kHz band are indicated as “all” in the figure. In both results of Fig. 5 (a) and (b), it was found that presenting the bone conduction sound to the position of P2 is appropriate from the viewpoint of ease of hearing.

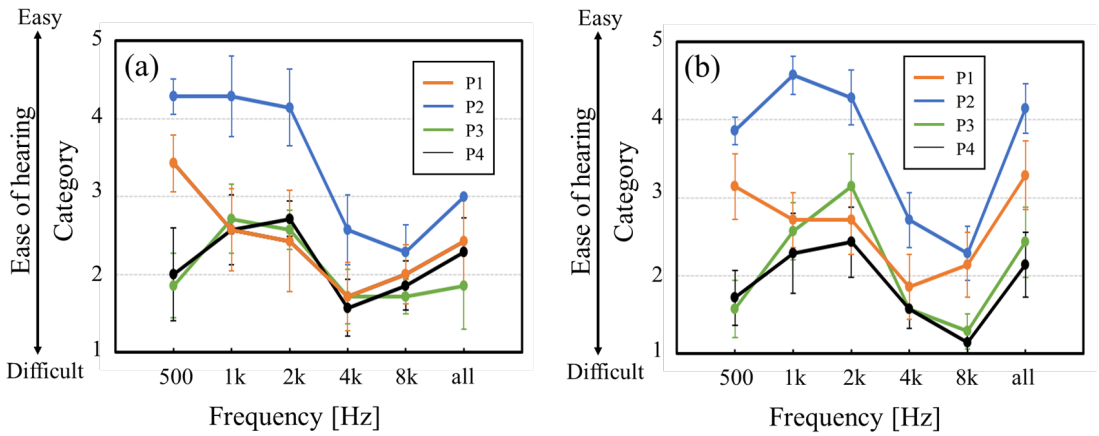

Fig. 5 Results of suitable contact position, (a)White noise, (b)Voice.

4. SOUND LOCALIZATION PERFORMANCE WITH BONE CONDUCTION REPRODUCTION

4.1 Basic performance of sound localization with binaural bone conduction reproduction

In order to confirm the performance of the sound field reproduction, a sound localization test was comparatively performed for both air-borne and bone conduction sounds. In this experiment, the sound field was simulated as follows. The subject was surrounded by 36 discrete virtual sound sources arranged in every 10 deg. As shown in Fig. 6, and the subject had to answer which source was radiating sound by the number of incident angle. Seven male subjects in their twenties were adopted.

Next, the processes of the binaural signal generation are described as follows. Firstly, white noise sound including the frequency components from 500 Hz to 8 kHz oct. bands having a duration of 5 seconds was adopted as the sound source. The HRTF measured using the head and torso simulator (ACO, Type 7828B) were used to synthesize the binaural signals.

The experiment is composed of two parts. The first part investigates the air-borne sound localization performance, whereas the second one investigates the borne-conducted sound localization performance. After the first part of the experiment, the subject detached the headphone for air-borne reproduction, and then put on the borne-conducted device. The sound incident from random directions was presented to the subject by the bone conduction device of J1 and J4, and the subject answered the incident direction of the sound, whereas the air-borne sound localization performance was measured by reproducing the same sound data as that of the bone conduction reproduction. In order to reproduce the air-borne sound, the over-ear headphone (Audio-Technica, ATH-W1000Z) was used instead of the inner-ear type headphone, because the subject had to detach the headphone only once between the first and second experiments.

It should be noted that the bone conduction sound was reproduced at the output volume so that the subjective loudness of the white noise including the frequency components from 500 Hz to 8 kHz bands reproduced by the bone conduction device is equivalent to that by the abovementioned over-ear headphone.

The result of the case in which the sounds were reproduced by the over-ear headphone is shown in Fig. 7. Some erroneous recognitions concerning misjudgment between the front and back directions were often seen as indicated by blue lines in the figure. This is caused by the use of a single HRTF obtained by the head and torso simulator for all subjects. Next, the results of the cases in which the sounds were reproduced by the bone conduction device of J1 and J4 is shown in Fig. 8 (a) and (b), respectively. The results of the sound reproduction by the device of J1 for the contact positions of P2 and P3 are shown in Fig. 8 (a) whereas those of J4 for the same contact positions are shown in Fig. 8 (b). Firstly, the sound localization performance of the case using the J4 device, especially in case of sound incidence of 90 and 270 deg., increased compared to that using J1 device. It is because the former device excites larger amount of bone conduction vibration especially in the higher frequency range, as shown in the vibration characteristics of Fig. 4. It should be noted that the erroneous judgement concerning the front and back direction indicated by the blue lines is also seen in these results with bone conduction sound reproduction. Secondly, there can be seen a slight difference in the sound localization performances of the cases of the contact position P2 and P3, and the former contact position indicates a little higher localization performance than the latter one, especially in case of sound incidence of 90 and 270 deg.

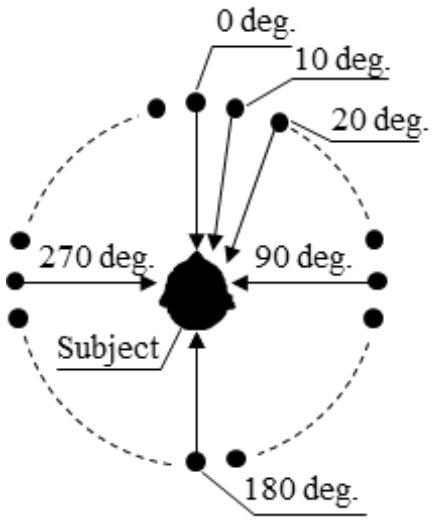

Fig. 6 Overview of experiment.

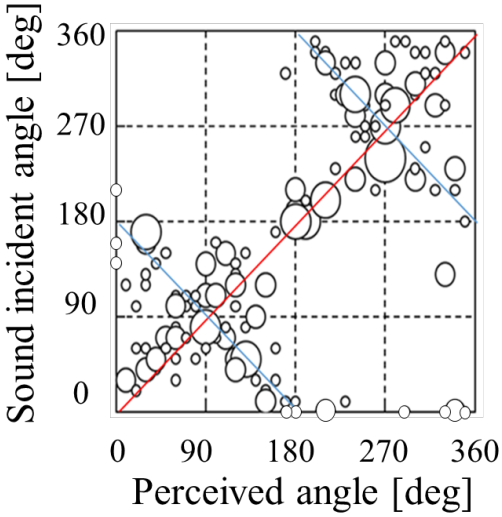

Fig. 7 Result of reproducing air-borne sound.

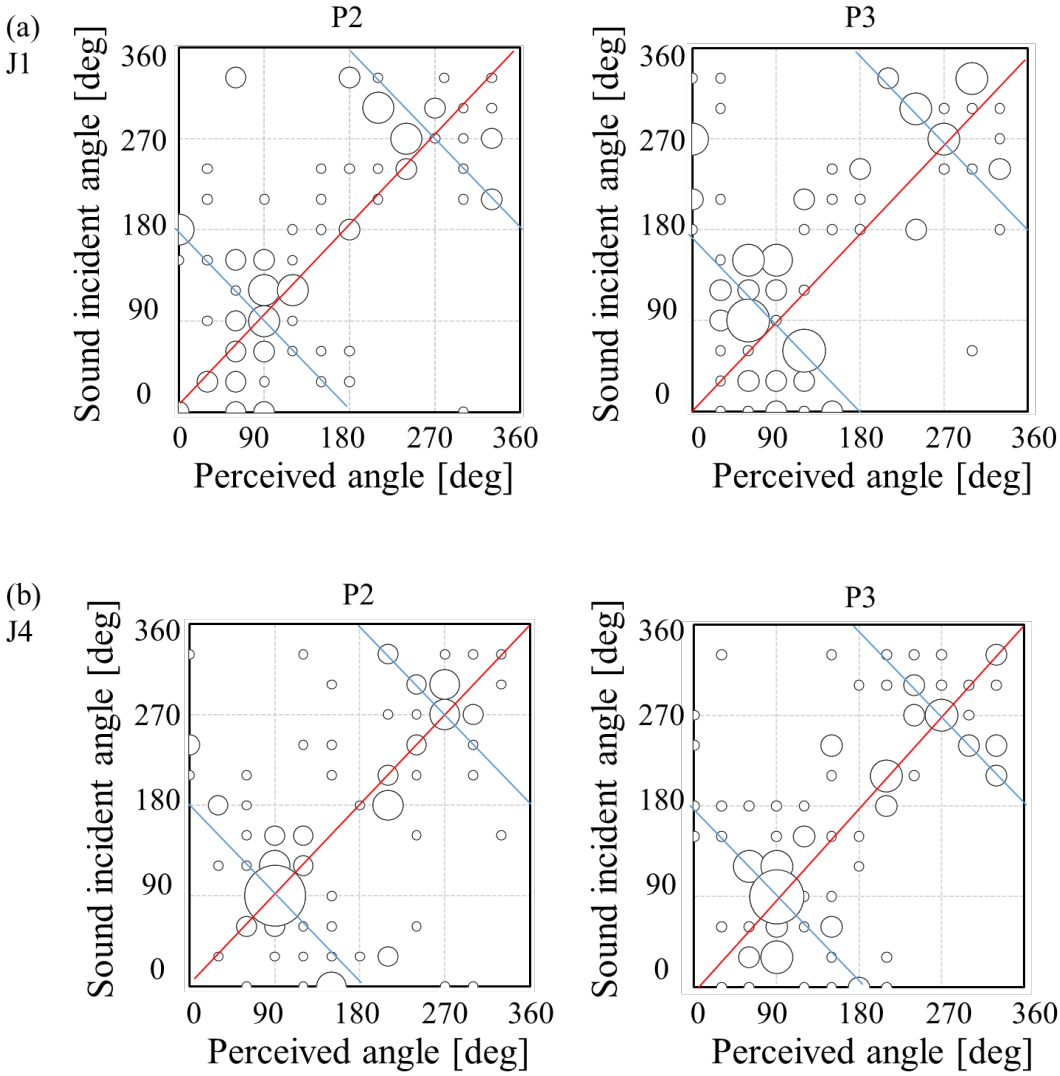

Fig. 8 Results of reproducing bone conduction sound, (a)Device J1, (b)Device J4.

4.2 Effect of panned binaural presentation with bone conduction reproduction

Further study on the sound image localization experiment using a bone conduction sound reproduction method, in which the intended direction of the reproduced sound images are quickly panned to left and right in order to enhance the perception of the sound direction, was carried out. The subject was surrounded by twelve discrete virtual sound sources arranged in every 30 degree. The overview of sound reproduction method is shown Fig. 9. The subject had to answer which source was radiating sound by the number of incident angle. As the participants of this experiment, five males in their twenties were adopted. In this study, angles of 10, 20 and 30 degrees were used. The reproduction sound was generated by convolution of the HRTF and the blink sound signal, which was adopted in this study aiming at future application for the human guidance. In this experiment, the HRTF measured with a KEMAR dummy head microphone provided by Massachusetts Institute of Technology [12] was used. The procedure of generating the panned sounds was performed as follows. Firstly, as indicated in Fig. 9(a), two virtual sound sources of Position A and B were additionally considered in addition to the original presentation point of Position O. The presentation of the sound source is quickly panned to Position A and B as shown in the time history of Fig. 9(b). Then, as can be seen in the figure, the sound incident from the Position O was made about 3 dB larger than that from positions A and B. The sound was created with a switching time of 0.5 s between each sound. The presentation conditions of the bone conduction element were P2 and J4 based on the examination results of the section 4.1. The results of sound localization experiments conducted under the above conditions are shown in Fig. 10. It should be noted that the red lines indicate correct answers, and the blue ones indicate the erroneous judgement concerning the front and back direction. The 20 and 30 degrees of panned angles tended to decrease erroneous judgement concerning the front and back than the 10 degree of panned angle. From this result, it turned out that it is appropriate to reproduce the panned angle at 20 degree. It also suggests that too large panned angle affects the sound image localization accuracy.

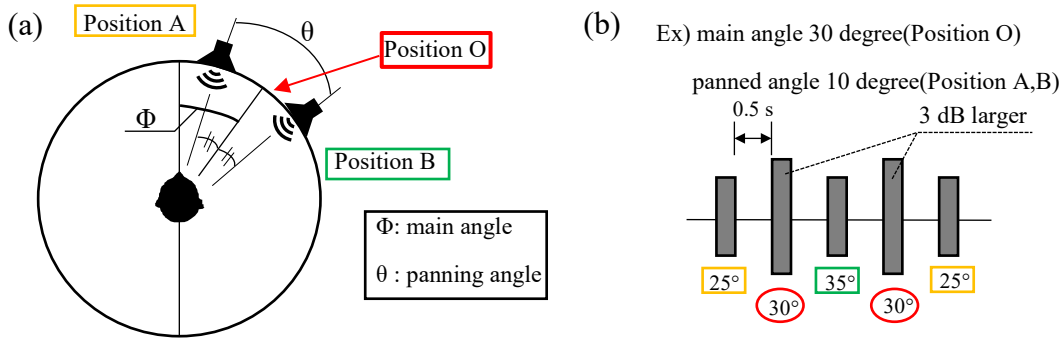

Fig. 9 Overview of sound reproduction method,
(a) Procedure of generating the panned sounds, (b)Example of reproduction sound.

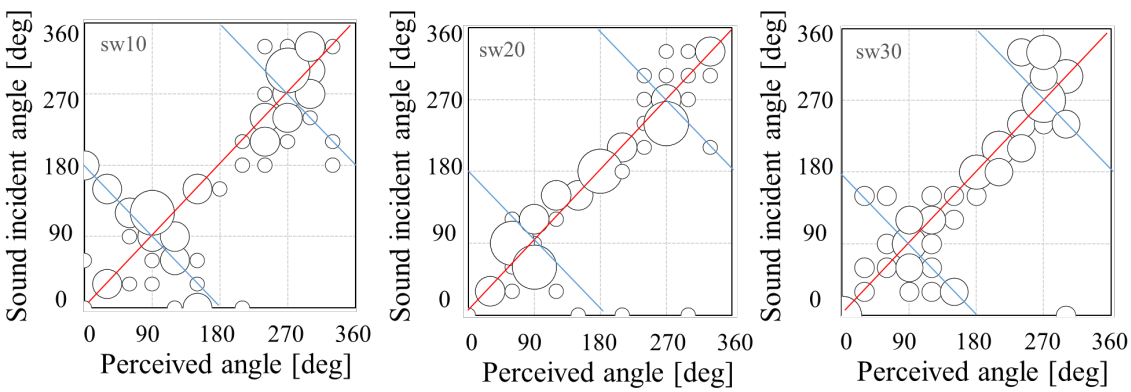

Fig. 10 Results of sound localization experiment.

5. CONCLUSION

The sound localization performance of binaural signals reproduced by the bone conduction device has been investigated. Detailed relationship between the sound localization performance and the contact conditions of the bone conduction reproduction devices has been discussed. It has been indicated that differences of the contact pressures between the head and the bone conduction device, and the contact positions of the bone conduction devices on the subject's head surface have large influence on the sound localization. Furthermore, as a results of a sound image localization experiment by using a bone conduction sound reproduction method, in which the intended direction of the reproduced sound images are quickly panned to left and right in order to enhance the perception of the sound direction, it was found that the condition of the twenty-degree panning was the best angle to clearly perceive the direction of incident sound.

REFERENCES

[1] A. Tjellström, B. Håkansson, G. Granström, "Bone-anchored hearing aids: current status in adults and children," Otolaryngol. Clin. North Am., 34, pp. 337–364 (2001).

[2] B. N. Walker and J. Lindsay, "Navigation performance in a virtual environment with bonephones," Proc. Int. Conf. Auditory Displays 2005 (2005).

[3] A. Väljamäe, A. Tajadura-Jiménez, P. Larsson, D. Västfjäll, and M. Kleiner , "Binaural bone-conducted sound in virtual environments: Evaluation of a portable, multimodal motion simulator prototype," Acout. Sci. & Tech., 29(2) pp. 149–155 (2008).

[4] D. W. Purcell, H. Kunov, and W. Cleghorn, "Estimating bone conduction transfer functions using otoacoustic emissions," J. Acost. Soc. Am., 114, pp. 907–918 (2003).

[5] Chang Y, Kim N, Stenfelt S., "The development of a whole-head human finite-element model for simulation of the transmission of bone-conducted sound," J. Acoust. Soc. Am., 140(3), pp.1635–1651 (2016).

[6] M. McBride, T. Letowski, and P. Tran, "Bone conduction reception: Head sensitivity mapping," Ergonomics, 51, 702–718 (2008).

[7] X. Qin and T. Usagawa, "Frequency characteristics of bone conduction actuators – Measurements of loudness and acceleration," Appl. Acoust., 126, pp. 19–25 (2017).

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

[9] K. Iida and M. Oota, "Median plane sound localization using early head-related impulse response," Appl. Acoust. 139, pp. 14–23 (2018).

[10] Y. Watanabe and H. Hamada, "A Study of Bone-Conduction and its Applications for Audio Signal," IEICE tech. rep., 105(136), pp. 19–24 (2005) (in Japanese).

[11] National Institute of Advanced Industrial Science and Technology (AIST), Digital Human Research Group, https://unit.aist.go.jp/hiri/dhrg/ja/dhdb/head/index.html (accessed 2018.9.27).

[12] HRTF Measurements of a KEMAR Dummy-Head Microphone, https://sound.media.mit.edu/resources/KEMAR.html (accessed 2018.5.24).