Acoust. Sci. & Tech. **36**, 1 (2015)

ACOUSTICAL LETTER

Experimental evaluation of binaural recording system using a miniature dummy head

Shunsuke Uchibori[1], Yuki Sarumaru[1], Kaoru Ashihara[2], Tetsuo Ohta[1] and Shizuko Hiryu[1,*]

[1] *Faculty of Life and Medical Sciences, Doshisha University, 1–3 Tatara Miyakodani, Kyotanabe, Kyoto, 610–0321 Japan*
[2] *National Institute of Advanced Industrial Science and Technology, 1–1–1 Higashi, Tsukuba, Ibaraki, 305–8566 Japan*

(*Received 16 September 2014, Accepted for publication 2 October 2014*)

Keywords: Ultrasound, Sound localization, Out of head externalization, Binaural technique, Biosonar
PACS number: 43.66.Pn, 43.66.Qp [doi:10.1250/ast.36.42]

1. Introduction

In unfamiliar places, some blind people use echolocation, i.e., the ability to sense obstacles using auditory information from sounds they make [1]. Psychoacoustic studies have compared human echolocation ability in blind and sighted people with normal hearing. For example, one study examined the ability of blind subjects to identify the presence of an obstacle by making click sounds and listening to the echoes [2,3]. Other studies examined target detection when blind subjects listened to echoes from a target using a loudspeaker that produced artificial sounds [4,5]. Furthermore, a mobility aid for blind people was developed that used digital signal processing to convert ultrasound echoes from a target into audible sounds in real time [6].

Bats and dolphins use echolocation to compensate for their having very poor eyesight. The use of biosonar in bats and dolphins has been demonstrated in many psychoacoustic and behavioral experiments [7,8]. Some biosonar researchers have examined human echolocation and conducted sonar discrimination experiments in humans to determine which discrimination cues were important for echolocation [9–11]. However, little research has been done in human echolocation. Biosonar researchers should be capable of studying echolocation in humans, and this might contribute to advances in biosonar research.

This research project aims to examine bat echolocation in an unconventional way, by comparing it with human echolocation ability. This study proposes a new system for ultrasonic binaural recording that humans can use to perform echolocation using ultrasound. The system consists of a miniature dummy head (MDH) and a device that converts ultrasonic echoes into audible sounds. The frequency of the sounds captured by the MDH shifts with the relative reduction in the size of the head, and listeners can perceive three-dimensional (3D) sounds through headphones.

This study evaluated the proposed MDH system for binaural recording. Sound localization experiments were conducted in humans with normal hearing to evaluate how the MDH system captures spatial information and conveys it via headphones.

*e-mail: shiryu@mail.doshisha.ac.jp

2. Materials and methods

2.1. Miniature dummy head

Figure 1 shows a 1/7 scale MDH with torso made from acrylonitrile-butadiene-styrene resin using a 3D printer (UP!3D Printer; Techno Solutions, Tokyo, Japan). The 3D shape data of a standard dummy head (4128C Head and Torso Simulator, HATS; Brüel & Kjær, Nærum, Denmark) were scanned using a 3D scanner (Model 2020i Desktop Laser Scanner; NextEngine, CA, USA). Because the HATS torso was too large for the 3D scanner, the shape of the torso was not scanned, but was formed so that it roughly imitated the shape of the HATS torso. The width, height, and depth of the MDH torso were 45 mm, 58 mm and 20 mm, respectively. 1/8-inch condenser microphones (FG-3329; Knowles, Itasca, IL, USA) were inserted at the entrance of the left and right ear canals of the MDH.

2.2. Measurement system

The loudspeaker (Pioneer Corp., PT-R7 III, Kanagawa, Japan) was positioned 1 m above the floor at the center of an anechoic chamber [3.4 m (H) $\times$ 4.4 m (L) $\times$ 3.4 m (W)]. The MDH was placed on a turntable 0.6 m from the center of the loudspeaker. Sounds were generated on a personal computer using MATLAB software. Output signals at a sampling rate of 192 kHz through an audio interface (OCTA-CAPTURE; Roland, Shizuoka, Japan) were amplified (Pioneer, A-D3) by 38 dB, high-pass filtered (cut-off frequency: 7 kHz), and sent to the loudspeaker. The sounds emitted from the loudspeaker were captured binaurally by the condenser microphones in the MDH, amplified by 46 dB using a custom-made amplifier circuit, and then digitized with a multi-channel PC recorder (PXIe-1036l NI, Tokyo, Japan, 500 kHz sampling rate). The total frequency response of the sounds emitted from the loudspeaker was corrected to a flat within $\pm$3 dB between 5 and 90 kHz. The sensitivity difference between the left and right microphones of the MDH was also corrected.

2.3. Sound localization experiments

White noise (5–90 kHz, 0.43 s duration including 0.04 s rise-fall time, sound pressure level of 100 dB at 1 m from the loudspeaker) was presented from the loudspeaker with a 30-ms repetition period. The MDH was rotated by 30° from 0 to 360° in the horizontal plane (around the cranio-caudal axis),

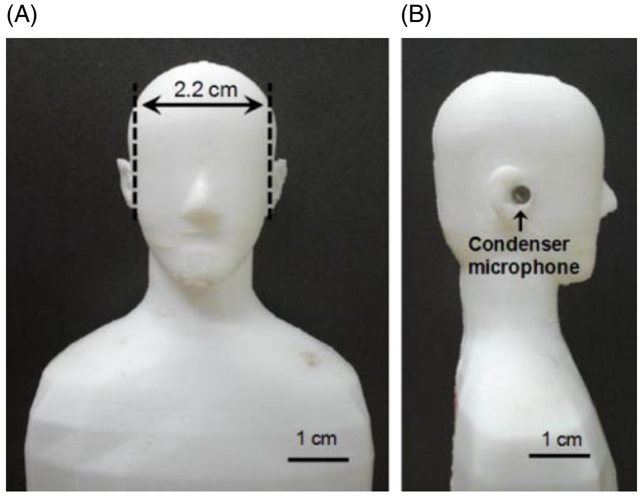

Fig. 1 Pictures of the miniature dummy head (MDH) showing the front (A) and side (B) views.

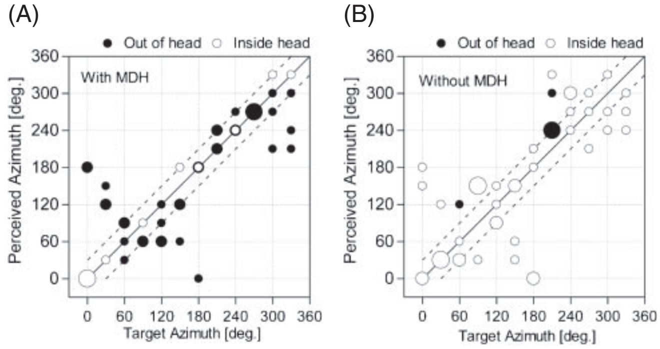

Fig. 2 Sound localization scores of subject B with (A) and without (B) the MDH under the rest condition. Open circles represent inside-head lateralization, and solid circles represent out-of-head localization. The sizes of the circles represent the number of the correct answers.

and the sound stimuli were captured in each position by the condenser microphones (12 recordings, referred to as rest condition). Additionally, recordings were conducted while the MDH was rotated manually by 90° around the cranio-caudal axis during sound emission. By changing the start angle of the rotation by 30° increments between 0 and 360° in the horizontal plane, 12 different recordings were made (rotate condition). Normal stereo recordings were also made using the same condenser microphones without the MDH under both the rest and rotate conditions (the distance between the two microphones was 22 mm).

Theoretically, the head-related transfer function (HRTF) of the MDH corresponds to the frequency-compressed version of that of the standard dummy head. Therefore, the frequency of the recorded high-frequency sounds was converted by a factor of 1/7 on the frequency domain by sample-rate conversion using Cool Edit 2000 (Syntrillium Software, AZ, USA). As a results, the recorded sounds were converted into the audible range (band-limited white noise from 0.71 to 12.9 kHz) at a sampling rate of 71.4 kHz (= 500 kHz/7), and the duration of a single burst noise was expanded to 3 s (= 0.43 s × 7). The frequency-converted sounds were presented to subjects through headphones (MDR-CD900ST, Sony, Tokyo, Japan).

Ten subjects (eight males and two females, 20–21 years old) with normal hearing (standard pure-tone audiometry, within 30 dB HL) participated. Before the experiment, the experimenter explained the purpose of this experiment. Training was conducted using the sounds under the rest and rotate conditions so that the subjects would understand how to report the perceived location of the presented sounds. The sounds used in the training were not used in the experiment.

Experiments were performed under rest and rotate conditions. Forty-eight stimuli [12 directions × 2 (rest and rotate) × 2 conditions (with and without MDH)] were presented in random order to each subject in an anechoic room [2.3 m (H) × 1.5 m (W) × 1.6 m (L)] using a custom-made program of EXPLAB (free software for computational experiments) on a personal computer. After each stimulus, the subject was asked to report the location (direction and inside/outside the head) of the sound source on an answer sheet that included a schematic top view of a human head with marks for every 30° of direction. After subject marked the position of the detected sound source on the answer sheet, they were instructed to push an enter key on the PC, and then the next stimulus was presented through the headphones after a 1-second delay.

After the experiment, the experimenter asked each subject about the externalization of the sound localization. The interview was recorded on an IC recorder (ICD-UX523; Sony, Tokyo, Japan). The experiment was approved by the ethics board of Doshisha University. Written informed consent was obtained from each subject before participation.

3. Results

Figure 2 shows representative results of the sound localization scores of one subject (subject B in Fig. 3) under the rest condition. The results for the binaural recording with the MDH and normal stereo recording without the MDH are shown in the left and right panels, respectively. The binaural sounds captured with the MDH system were localized outside the head more often than the normal stereo sounds were. Additionally, the sound localization score was improved with the MDH (percent correct for subject B: 50% with the MDH *vs.* 37% without the MDH). Although the front-back sound localization error was still observed, Fig. 2 shows that the front-back error was reduced slightly by MDH, as confirmed in all ten subjects.

Figure 3 compares the percent of correct sound localization scores for all 10 subjects (A–J) with and without the MDH in the rest condition. The means of correct answer were significantly different between the sound captured by the MDH and the normal stereo sounds (Student's t-test; $p < 0.001$). All subjects had higher sound localization scores with the MDH (mean ± SD: 41 ± 8%) than for the normal stereo sounds (25 ± 6%).

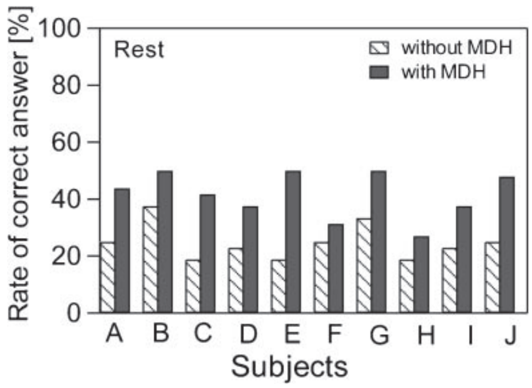

Fig. 3 Comparisons of the percent correct of the sound localization scores for all 10 subjects (A–J) with and without the MDH under the rest condition.

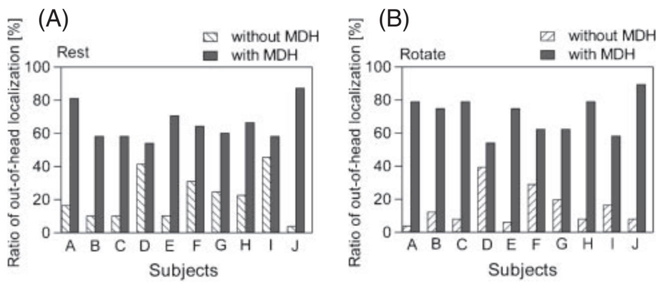

Fig. 4 Comparisons of the scores for out-of-head localization with and without the MDH under the rest (A) and rotate (B) conditions for all 10 subjects (A–J).

Figure 4 summarizes the scores for out-of-head localization in all subjects under the rest and rotate conditions. In the rest condition, all subjects experienced greater externalization for the sound captured by the MDH ($66 \pm 11\%$) than for normal stereo sounds ($22 \pm 14\%$) (Student's t-test; $p < 0.001$, Fig. 4A), and the difference was also significant in the rotate condition ($p < 0.001$, Fig. 4B). The out-of-head localization for the sounds captured by the MDH was improved for six of 10 subjects under the rotate condition compared with the rest condition. Although the difference was not significant in this study (Student's t-test; $p = 0.14$), on average, the rotate condition showed better out-of-head localization than the rest condition ($71 \pm 11\%$ under the rotate condition *vs.* $66 \pm 11\%$ under the rest condition). Additionally, in the interviews, all subjects answered that (1) the movement of the sound source was perceived more clearly with the MDH than without it, and (2) the externalization of the sound source was better under the rotation than under the rest condition, although the improvement in out-of-head localization varied among individuals.

4. Discussion

In this study, all subjects had better sound localization scores and experienced externalization more frequently for the sounds captured by the MDH than for those captured without the MDH. This suggests that ultrasound captured by a MDH potentially contains spatial information that can be used by a human listener after converting its frequency.

Sound localization studies using a standard dummy head (HATS) demonstrated that (1) the out-of-head localization for normal stereo sounds (without a dummy head) is difficult, (2) the front-back sound localization error is observed even for binaural sounds capture by a dummy head based on individual head shape of a subject, and (3) the sound localization error dramatically decreases under head movement condition especially when the head movement was controlled by the subject [12]. In this study, we found that it was difficult to externalize the sound source for normal stereo sounds (without the MDH), and the front-back error was observed not only for the stereo sounds but also for the binaural sounds captured by a miniature HATS (not "individual" miniature dummy head). These two results correspond to findings previously reported in sound localization studies using a standard dummy head [12]. Furthermore, all subjects answered that the externalization of the sound source was better under the rotation than under the rest condition, suggesting that the head movement is important for spatial localization of the sound source for binaural sounds captured by the MDH. In the present study, however, the rotation of the MDH was not related to either of the subject's self-motion or the subject's will. It can be speculated that if the rotation of the MDH was controlled by the subjects, the front-back error rate might have been lower.

Ultrasound has an important advantage over audible sounds in the context of echolocation; i.e., a high-frequency wave (short wavelength) does not get diffracted much, resulting in higher resolution of object sensing. Furthermore, competing background sounds are scarcer in the ultrasonic range than in the audible range in nature although ultrasound attenuation in the air is marked. Therefore, this suggests that binaural ultrasound has more potential and includes different acoustical information regarding the surrounding 3D space compared with audible sound. Therefore, the proposed MDH system should be a practical tool for human echolocation research. We plan to conduct a psychoacoustic experiment using the MDH system to investigate human echolocation ability.

Previously, a small-sized human dummy head was used to investigate the effect of the head size on sound localization using human listeners (the sounds were not frequency-converted by a scaling factor of the dummy head) [13]. In an architectural acoustics study, a miniature dummy head was also developed for the advanced prediction of the acoustic quality of a room [14]. Kumagai, *et al.* [15] suggested that the frequency-compressed HRTF of the MDH shows similar frequency response of the standard dummy head up to 4 kHz. Although further acoustical evaluation of difference in HRTFs between the standard dummy head and the MDH should be required, psychoacoustic experiment using human subjects was conducted first in this study so that we could verify the usefulness of the proposed MDH system for binaural recording. As a result, we demonstrated that frequency-converted binaural sounds using MDH system improved sound localization scores and more frequently localized outside the head, compared with frequency-converted stereo sounds. We will conduct acoustical measurements of the HRTFs for the proposed MDH system to reveal the physical accuracy in the context of binaural recording.

Recently, Morimoto, *et al.* [16] proposed a new auditory guide system using ultrasound in which only users with ultrasonic microphones could detect the guide signals that were converted to audible sounds. Our MDH system might also be applied to a new signal transduction system using binaural ultrasound, such as an auditory guide.

5. Conclusions

We proposed a binaural recording system that allows human listeners to localize the source of a sound. The system consists of a miniature dummy head (MDH) and a frequency converter. Sound localization experiments were conducted in subjects with normal hearing to investigate the efficiency of the MDH system. It was found that the sounds captured by the proposed MDH system were localized more correctly and outside the head more than normal stereo sounds were. Our finding suggests that the MDH is a novel tool not only for human echolocation research but also for providing new information for communication engineering, such as ultrasonic guidance signals.

Acknowledgements

We thank Dr. Hiroshi Kido for the 3D shape data of a standard dummy head. We also thank Masataka Kinoshita and Masataka Yoshida for assistance with the acoustical measurements.

References

[1] L. Thaler, "Echolocation may have real-life advantages for blind people: An analysis of survey data," *Front Physiol.*, **4**, 98 (2013).
[2] C. E. Rice, "Human echo perception," *Science*, **155**, 656–664 (1967).
[3] C. Arias and O. A. Ramos, "Psychoacoustic tests for the study of human echolocation ability," *Appl. Acoust.*, **51**, 399–419 (1997).
[4] B. N. Schenkman and M. E. Nilsson, "Human echolocation: Blind and sighted persons' ablity to detect sounds recorded in the presence of a reflecting object," *Perception*, **39**, 483–501 (2010).
[5] L. Thaler, S. T. Arnott and M. A. Goodale, "Neural correlates of natural human echolocation in early and late blind echolocation experts," *PloS One*, **6**, e20162 (2011).
[6] T. Ifukube, T. Sasaki and C. Peng, "A blind mobility aid modeled after echolocation of bats," *IEEE Trans. Biomed. Eng.*, **38**, 461–465 (1991).
[7] J. A. Simmons, "Perception of echo phase information in bat sonar," *Science*, **204**, 1336–1338 (1979).
[8] W. W. L. Au, *The Sonar of Dolphins* (Springer-Verlag, New York, 1993).
[9] C. M. Delong, W. W. L. Au, H. E. Harley and H. L. Roitblat, "Human listeners provide insights into echo features used by dolphins (*Tursiops truncatus*) to discriminate among objects," *J. Comp. Psychol.*, **121**, 306–319 (2007).
[10] W. W. L. Au and D. W. Martin, "Insights into dolphin sonar discrimination capabilities from human listening experiments," *J. Acoust. Soc. Am.*, **86**, 1662–1670 (1989).
[11] J. R. Fish, C. S. Johnson and D. K. Ljungblad, "Sonar target discrimination by instrumented human divers," *J. Acoust. Soc. Am.*, **68**, 1285–1293 (1976).
[12] T. Hirahara, M. Ohtani and I. Toshima, "Issues on HRTF measurements and binaural reproduction," *Fundam. Rev.*, **2**, 68–85 (2011).
[13] T. Hirahara, H. Sagara and W. Endo, "Sound localization with scaled dummy heads," *Proc. Spring Meet. Acoust. Soc. Jpn.*, pp. 433–434 (2006).
[14] N. Xiang and J. Blauert, "A miniture dummy head for binaural evaluation of tenth-scale acoustic models," *Appl. Acoust.*, **33**, 123–140 (1991).
[15] S. Kumagai, K. Ashihara and S. Kryu, "Simple measurement of head-related transfer function by using miniature dummy head," *Tech. Rep. IEICE*, (2013) (in Japanese).
[16] M. Morimoto, H. Sato, K. Suda and H. Sato, "A possibility of auditory guide signals using lateralization by bone-conduction headphones," *Proc. Spring Meet. Acoust. Soc. Jpn.*, pp. 589–590 (2014).