



Sound-Image Icon with Aerial Haptic Feedback

Seunggoo Rim^(✉)^(iD), Shun Suzuki^(✉)^(iD), Yutaro Toide^(✉)^(iD),
Masahiro Fujiwara^(✉)^(iD), Yasutoshi Makino^(✉)^(iD), and Hiroyuki Shinoda^(✉)^(iD)

The University of Tokyo,
5-1-5 Kashiwanoha, Kashiwa-shi, Chiba-ken 277-8561, Japan
{rim,suzuki,toide}@hapis.k.u-tokyo.ac.jp,
Masahiro_Fujiwara@ipc.i.u-tokyo.ac.jp,
yasutoshi_makino@k.u-tokyo.ac.jp,
hiroyuki_shinoda@k.u-tokyo.ac.jp

Abstract. In this study, we attempt to define a novel invisible mid-air three-dimensional (3D) object, which informs users of its existence and location via sound and haptic feedback. The correlation between the senses of hearing and touch creates the feeling of touching a sound source, which the user recognizes as a virtual object: sound-image icon. The sound from the icon instantaneously notifies the user about its position without requiring vision. In addition, aerial tactile sensation enables users to freely interact with and manipulate these icons with no need to wear any devices. Therefore, this approach exhibits enormous potential in various situations, such as surgical operations, works in factories, driving cars, and button/switch operations in daily life. In this study, we prototyped the sound-image icon and experimentally examined their feasibility. We confirmed that users could estimate the location of the icons and measured the time required to access these icons. The results indicate that the sound-image icon is feasible as a novel 3D interface.

Keywords: Mid-air haptics · Haptic display · Sound-source localization

1 Introduction

In this study, we propose a sound-image icon, which is an invisible 3D object that integrates sound-source localization and mid-air tactile sensation. This sound and haptic feedback creates a virtual object with no visual appearance, without requiring the user to wear any devices.

A typical method of reproducing tactile sensation in mid-air [2,3,6,8] is installing an airborne ultrasound tactile display (AUTD) that presents the moving stimulus on a skin by remotely producing the radiation pressure of focused aerial ultrasound. Such mid-air haptic feedback has been integrated with visual floating images in previously conducted studies [9,11]. From the viewpoint of interface design, vision is the most efficient channel to transmit the spatial

© The Author(s) 2020

I. Nisky et al. (Eds.): EuroHaptics 2020, LNCS 12272, pp. 489–496, 2020.

https://doi.org/10.1007/978-3-030-58147-3_54

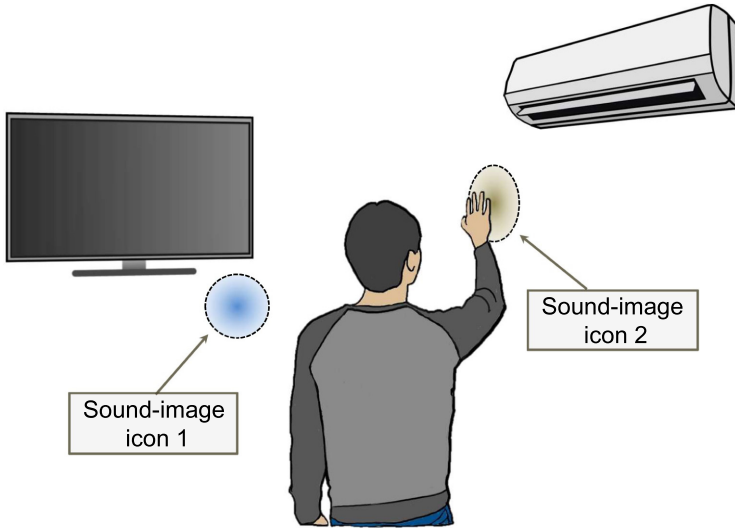


Fig. 1. Concept of a sound-image icon. The person selects the sound-image icon of the air conditioner to control the temperature. The sound from the icon informs the user about the position of itself. The icon provides tactile feedback without any visual image.

arrangement of an object to a user, while haptics facilitates the transmission of the will of a user to a computer system. Therefore, a 3D visual interface with mid-air haptics is a reasonable integration as an efficient interface. However, the eyes are sometimes occupied by a specific task such as in-car driving or a surgical operation. In addition, glassless 3D vision is still immature, where it is difficult to secure a wide view angle, while a head-mounted display sometimes causes fatigue and VR sickness. Instead of vision, the use of sound is another option to display the object position, as humans can instantaneously identify the direction of the sound when they are in an environment where sound can be clearly heard [1]. In addition, the sound can transmit words and tones that express various attributes.

Many studies have been performed to create virtual sound sources at specific locations under various conditions and environments. Recent studies have focused on virtual sound source positioning for acoustic navigation in unknown spaces [14] or vector base amplitude panning for creating 2D or 3D sound fields without considering the placements of any number of loudspeakers [13]. Although the aforementioned technologies form a wide research domain, we could not find studies that integrated virtual sound sources and mid-air haptics technologies. It would be intriguing to investigate whether auditory and haptic perceptions effectively complement each other.

The concept of a sound-image icon is depicted in Fig. 1. The user hears a binaural sound and identifies the direction of the sound-source. We refer to this

sound source as a “sound image.” The sound image can represent various functions, and users can recognize the role of the object by its sound. For example, an icon that represents an air-conditioner generates a sound to explain it in words, while an audio-volume icon produces a pleasant musical sound. The users utilize this sound as a clue to reach for the icons. Using tactile cues, they can recognize the exact positions of the icons and then perform fine tasks; for example, a user who accepts the objects of the air-conditioner can control the temperature by operating the sound-image icon. The haptic feedback is critical not only to improve the operability during the control but also to reliably guide and hold the user’s hand to the starting point of the operation.

In this study, we prototype sound-image icons and experimentally verify their feasibility. We aim to realize a system where a user can select the desired icon among multiple icons and operate it. In this study, for the first step, we examine whether users can find an icon and measure the time required to access it. The combination of auditory and tactile sensations enables users to accurately and effectively locate the icons.

2 Proposed Method

A sound-image icon is realized using a sound source with the sense of touch provided via acoustic radiation pressure. In this section, we describe the method used to create the sound-image icon.

2.1 Producing a Sound Image

The user specifies the direction of the sound source using a binaural sound. We plan to provide binaural sounds using ultrasound beams to reach the ears. However, in this experimental system, the binaural sound was provided to the users by an in-ear binaural headset (CS-10EM, Roland). The binaural sound was recorded using the microphone in the headset that was fixed to the ears of one of the authors, keeping the sound source at the icon position. By reproducing the recorded sound in both the ears, the listener perceived the same 3D sound image as the real sound, under the assumption that the head related transfer functions are common [10].

2.2 Aerial Haptic Feedback

The aerial tactile sensation is presented at the icon location by using AUTD. The users actively search for the ultrasound focus, where the acoustic radiation pressure produces a tactile sensation of the virtual icon.

AUTD is a phased array that generates an ultrasonic focal point at an arbitrary position in the air [5,6]. The acoustic radiation pressure is proportional to the sound energy density on the skin surface [3]. Though an AUTD can produce various pressure patterns by controlling the amplitude and phase of each

transducer at the frame-rate of 1 kHz [4], a single focus is created in this prototype. The users can perceive a certain stimulus around the focal point, where the tactile feel becomes vivid when the ultrasound amplitude is modulated in the amplitude or the focus position is laterally vibrated on the skin [15].

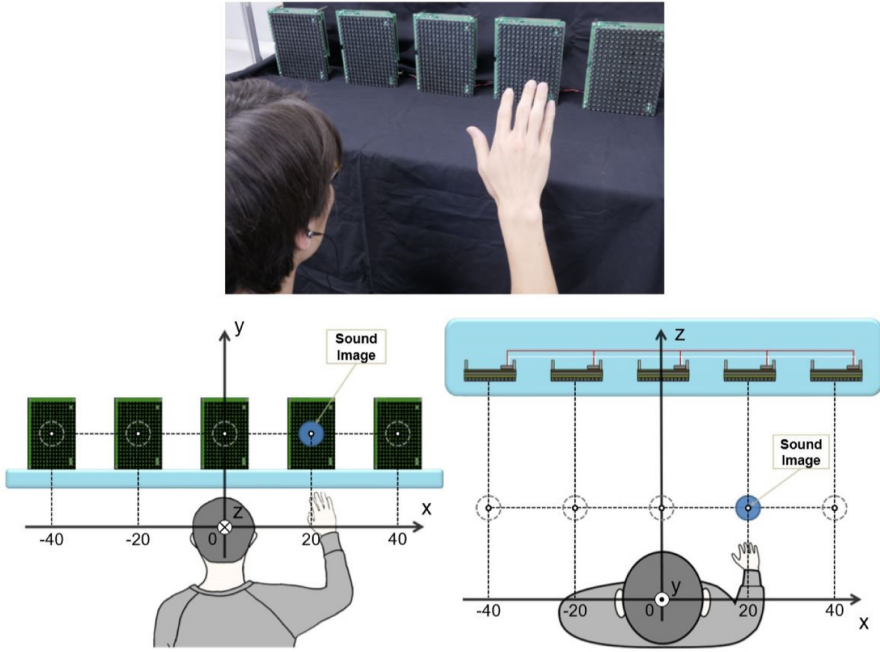


Fig. 2. Experimental setup. Five AUTDs were deployed 30 cm behind the five spots where the sound-image icon must be placed. The units of the numbers are in cm.

3 Experiment

We experimentally verified that humans could haptically identify the location of a particular icon by following the perception of sound. We measured the accuracy and time for haptic identification.

3.1 Procedure

The experimental setup is depicted in Fig. 2. In this experiment, we displayed five icons and examined whether the participants could identify them, and then we measured the time required for the identification. The icons were placed at $(-40, 20, 20)$, $(-20, 20, 20)$, $(0, 20, 20)$, $(20, 20, 20)$, and $(40, 20, 20)$ cm, where the origin was the center of the head, z-axis was parallel to the front direction,

y-axis was parallel to the vertical direction, and x-axis is set as forming a right-handed system. The five AUTDs were placed 30 cm behind the sound image. A single corresponding AUTD emitted a focused ultrasound with 200 Hz sinusoidal amplitude modulation. The unit of the AUTD is an ultrasound phased array (SSC-HCT1, Shinko Shoji Co., Ltd.) with 14×18 elements. The maximum force displayed by a single unit is 10 mN.

As sound sources, we recorded the solos of the following five kinds of musical instruments: drum, bass, acoustic guitar, piano, and flute. We defined the drum sound as the target sound, as it was the easiest one to locate the position. We instructed the participants that the drum was the sound of the target sound-image icon to locate. Several sound sources (five at maximum), which included the target sound, were randomly selected and played at random positions in the five locations. As tactile feedback, a single ultrasound focus of the target icon was created at the target position.

The finding-target experiment was conducted considering the following two conditions: auditory-only and auditory-with-tactile. In the auditory-only condition, only the binaural sound was presented, while in the auditory-with-tactile condition, both sound and tactile stimulation were simultaneously presented.

We asked the participants to estimate the location of the target sound-image icon as soon as possible after they recognized the start cue, i.e., the moment the audio was played. In addition, we instructed them to close their eyes during fumble to prevent the visual effect. After determining the position of the icon, the participants indicated the position number among the five options from 1 to 5 with the keyboard. We applied white noise to eliminate the effects of AUTD driving noise.

The participants in this experiment were twelve men in their twenties who had no problems with hearing or health.

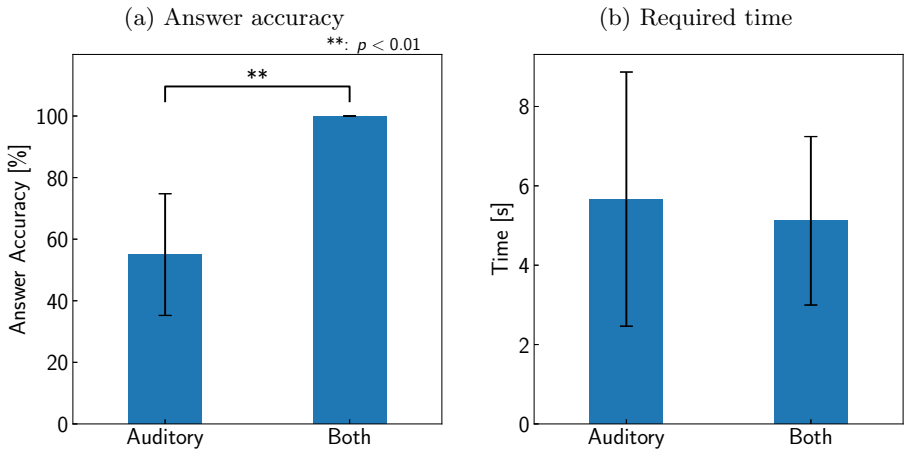


Fig. 3. Average of participants' (a) answer accuracy and (b) required time.

3.2 Results and Evaluations

The results are depicted in Figs. 3 (a) and (b). According to Fig. 3 (a), the answer accuracy was 55% when only the sound was informed of the location to the participants. Despite the high error rate of the “auditory-only condition,” the accuracy rate of the “auditory+tactile condition” was almost 100%. Using the *t*-test, we examined whether the correct answer rate could be significantly improved by adding a tactile sensation to the sound cue. As a result of the test, the *p*-value between “auditory” and “auditory+tactile” was smaller than 0.01. This result means that the participants could exactly pinpoint the location of the sound image when they were able to search using tactile sensations.

In addition, Fig. 3 (b) shows that the average required times of each case were almost the same. Accordingly, the average time for “auditory” was 5.67 s, and that for “auditory+tactile” was 5.12 s. The standard deviation for “auditory+tactile” was less than that for “auditory.” The standard deviation for “auditory+tactile” was approximately 2.12 s, and that for “auditory” was 3.20 s. This indicates that presenting both the stimuli reduced their standard deviations. For clarification, we used Levene’s test for the standard deviations and Welch’s *t*-test for the averages. As a result of Levene’s test, the *p*-value between “auditory” and “auditory+tactile” was smaller than 0.01. According to the test, a significant difference was observed between the variances of required time for the two conditions. Additionally, the significance between the average amounts was not noticed, as the *p*-value from the Welch’s *t*-test was 0.121.

4 Discussion

As depicted in Fig. 3 (a), in the case of sound alone, the exact position of the icon could not be estimated, and mistakes occurred. On the other hand, the correct answer rate became 100% by adding tactile feedback. This result indicates that sound localization was instantaneous, but it was inaccurate and unreliable. This drawback was compensated via haptic feedback, which offers reasonable cooperation between auditory and haptic perceptions. That is, the participants grasped the approximate position by hearing the sound and determined the exact position by touching the sound-image icon [7].

As additional information, it was possible to localize the position of the icon only by tactile sensation. To clarify this, we also conducted an additional experiment for the tactile-only condition. The participants and procedures were the same as those described in the Experiment section, and the start cue was an extra monaural audio. In this case, the time to perform localization was 5.30 ± 3.39 s. Although the average of the required time was comparable to that for the “auditory+tactile” condition, the variance was significantly longer.

Before this additional experiment, the time cost was expected to be the shortest in the case of “auditory+tactile.” However, this hypothesis was not observed in this experiment. Searching the entire space without prior information was not a time-consuming task, as it only took approximately 2 s for the participants to fumble around the area with their hands. Nevertheless, considering that the

standard deviation for the “auditory+tactile” condition is the smallest, it was confirmed that the combination of sound and haptic feedback facilitated the search of the icon.

To avoid confusion, we reconfirm the purpose of the combination of auditory and tactile sensations as follows. The role of the sound is to notify the user of the existence and attributes of the icon around the user. Tactile sensation is necessary to determine the exact location and operate the icon. Therefore, even if the localization time for the tactile-only condition is short, it does not mean that the auditory cue is unnecessary.

5 Conclusion and Future Works

In this study, we proposed a sound-image icon and examined the basic feasibility of icon localization. The sound-image icon represents the virtual existence of sound and haptics without visual presentation. Through the research, we confirmed that the participants could search and estimate the location of a single icon in an efficient manner using their tactile and auditory senses. For future work, we will investigate the possibility of efficiently displaying multiple icons.

We used a headset as a sound display device, as this was a feasibility study to examine the effectiveness of the auditory–haptic integration. However, it is also possible to produce binaural sounds in a non-contact manner using airborne ultrasound [12]. Performing a detailed operation using the sound-image icon was beyond the scope of this paper, and it would be the next important challenge of the sound-image icon.

References

1. Blauert, J.: *Spatial Hearing: The Psychophysics of Human Sound Localization*. MIT Press, Cambridge (1997)
2. Carter, T., Seah, S.A., Long, B., Drinkwater, B., Subramanian, S.: UltraHaptics: multi-point mid-air haptic feedback for touch surfaces. In: *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, pp. 505–514 (2013)
3. Hoshi, T., Takahashi, M., Iwamoto, T., Shinoda, H.: Noncontact tactile display based on radiation pressure of airborne ultrasound. *IEEE Trans. Haptics* **3**(3), 155–165 (2010)
4. Inoue, S., Makino, Y., Shinoda, H.: Scalable architecture for airborne ultrasound tactile display. In: Hasegawa, S., Konyo, M., Kyung, K.-U., Nojima, T., Kajimoto, H. (eds.) *AsiaHaptics 2016*. LNEE, vol. 432, pp. 99–103. Springer, Singapore (2018). https://doi.org/10.1007/978-981-10-4157-0_17
5. Iwamoto, T., Shinoda, H.: Two-dimensional scanning tactile display using ultrasound radiation pressure. In: *2006 14th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 57–61. IEEE (2005)
6. Iwamoto, T., Tatzono, M., Shinoda, H.: Non-contact method for producing tactile sensation using airborne ultrasound. In: Ferre, M. (ed.) *EuroHaptics 2008*. LNCS, vol. 5024, pp. 504–513. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-69057-3_64

7. Kaul, O.B., Rohs, M.: Haptiched: a spherical vibrotactile grid around the head for 3D guidance in virtual and augmented reality. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp. 3729–3740 (2017)
8. Korres, G., Eid, M.: Haptogram: ultrasonic point-cloud tactile stimulation. *IEEE Access* **4**, 7758–7769 (2016)
9. Makino, Y., Furuyama, Y., Inoue, S., Shinoda, H.: Haptoclone (haptic-optical clone) for mutual tele-environment by real-time 3D image transfer with midair force feedback. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, pp. 1980–1990 (2016)
10. Møller, H.: Fundamentals of binaural technology. *Appl. Acoust.* **36**(3–4), 171–218 (1992)
11. Monnai, Y., Hasegawa, K., Fujiwara, M., Yoshino, K., Inoue, S., Shinoda, H.: HaptoMime: mid-air haptic interaction with a floating virtual screen. In: Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology, pp. 663–667 (2014)
12. Ochiai, Y., Hoshi, T., Suzuki, I.: Holographic whisper: rendering audible sound spots in three-dimensional space by focusing ultrasonic waves. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp. 4314–4325 (2017)
13. Pulkki, V.: Virtual sound source positioning using vector base amplitude panning. *J. Audio Eng. Soc.* **45**(6), 456–466 (1997)
14. Storek, D., Rund, F., Suchan, R.: Virtual auditory space for the visually impaired-experimental background. In: 2011 International Conference on Applied Electronics, pp. 1–4. *IEEE* (2011)
15. Takahashi, R., Hasegawa, K., Shinoda, H.: Tactile stimulation by repetitive lateral movement of midair ultrasound focus. *IEEE Trans. Haptics* **13**, 334–342 (2019)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

