Feature Articles: Revolutionizing Living and Working Spaces with Personalized Sound Zone

PSZ Spot-sound-reproduction Technology: New Sound-confinement Method Using Opposite-phase Sound Waves

Hironobu Chiba, Tatsuya Kako, Hiroaki Ito, Kenichi Noguchi, Noriyoshi Kamado, and Akira Nakayama

Abstract

In response to the spread of telework and web conferencing as well as the growing need to value private time and space, we aim to create the ultimate private acoustic space, a Personalized Sound Zone (PSZ), that delivers only the sounds you want to hear and blocks the sounds you do not want to hear. As part of this effort, we have been researching and developing a new spot-sound-reproduction technology that confines unwanted sounds from speakers to a very small space. This technology will create an area where sound can only be heard near the speaker by appropriately controlling opposite-phase sound waves emitted from the back of the speaker. This technology will enable the NTT Group to provide a variety of unique PSZ audio equipment.

Keywords: spot-sound-reproduction technology, PSZ, enclosure

1. Spot-sound-reproduction technology that presents sound locally

Since individuals now have their own personal electronic devices, such as smartphones, tablets, and personal computers, and telework and web conferencing have rapidly become common, NTT Computer and Data Science Laboratories is conducting research and development aimed at constructing the ultimate private acoustic space for delivering only sounds you want to hear and blocking sounds you don't want to hear in response to the increasing need to value private time and space.

Acoustic devices, such as directional speakers and parametric speakers, had been used to reproduce sound in one part of an area only without the sound leaking to the surrounding area. These loudspeakers form an area where sound can be heard from a specific direction, but they require special devices and have not yet become widely used. At the research level, it has also been shown that spatially controlling a speaker array with multiple speakers can reproduce sound in a certain area only. However, it is problematic that many speakers are required, and the cost is extremely high.

Headphones and earphones are widely used to present sound to individuals only. Although they are inexpensive and readily available, wearing them for a long time puts pressure on the ears and ear canals, which causes fatigue and pain, and prolonged wear can lead to inflammation of the outer ear. Many openear earphones have been sold as a method of wearing earphones that do not block the ear canal, thus reducing the burden on the ear. They are configured with the loudspeaker placed close to the ear, from where it delivers the sound to the eardrum. Unlike earphones that are inserted into the ear canal, open-ear earphones leave the ear canal open, so they exert less pressure—thus less burden—on the ear. However, the distance between the ear and loudspeaker in such earphones is greater, so sound leakage becomes an issue.

We describe the spot-sound-reproduction technology we are researching and developing that uses opposite-phase sound waves emitted from the back of a loudspeaker to create an area where sound can only be heard in the area close to the loudspeaker. We first introduce our previously proposed enclosure-less speaker array, which uses general-purpose loudspeakers to achieve spot-sound reproduction at a reduced cost. We then introduce the Personalized Sound Zone (PSZ) Wearable earphone we developed for solving the sound-leakage problem with open-ear earphones. Since these two technologies can localize sound, they can expand the possibilities of new audio equipment from a conventional loudspeaker system, which delivers sound to a large number of people, to one that can produce sound to be heard by specific people only.

2. Spot-sound reproduction using opposite-phase sound waves from the back of the speaker

The principle on which speakers function is briefly explained. A speaker consists of a speaker unit and enclosure that houses the speaker unit. The speaker unit consists of a diaphragm and magnetic circuit that vibrates the diaphragm. When the diaphragm of the speaker unit is vibrated, it causes compression and expansion waves around the unit. People are able to perceive sound when the created sound waves reach their ears. The sound waves generated on the diaphragm are a physical phenomenon, so when the diaphragm vibrates, the sound waves are generated not only on the entire front surface of the diaphragm but on the back surface as well. The sound wave generated on the back surface-generally referred to as the opposite-phase sound wave—is exactly the inverse of the wave generated on the front surface (positive-phase sound wave). This opposite-phase sound wave can cancel out the front wave when timed appropriately with the positive-phase sound wave. Therefore, if the diaphragm vibrates without the enclosure, the opposite-phase sound wave diffracts around the speaker unit and cancels out the positivephase sound wave, so the sound disappears. No matter how much the diaphragm is vibrated, no sound

can be heard. Therefore, normal speakers are configured so that the speaker unit is housed in a box called an enclosure, which suppresses the radiation of the opposite-phase sound waves to the surroundings, thus enabling the speaker's audible sound to travel further. Regarding PSZ spot-sound reproduction, we conjectured that by appropriately controlling the opposite-phase sound waves generated from the back of the diaphragm of the speaker unit, it would be possible to suppress sound leakage and achieve spot-like sound reproduction. The results of sound-leakage analysis are shown in **Fig. 1**. The sound-pressure level of the bare (enclosure-less) speaker unit (i) is lower at a greater distance from the unit than that of the normal (enclosure-type) speaker (ii).

Regarding the frequency characteristics for each frequency (shown in the right graph), the distance from the speaker unit increases, the sound magnitude is attenuated more at lower frequencies than at higher frequencies. Closer to the speaker unit, however, the frequency characteristic was found to be flat, meaning a better sound quality. These results indicate that by using two enclosure-less speaker units, i.e., placing two speakers near the head, it is possible to make a low-cost speaker capable of spot-sound reproduction by which the generated sound can only be heard near the speaker. To improve the performance of enclosure-less speaker units, we proposed an enclosure-less speaker array in which two speaker units are arranged [1]. Our speaker array is shown in Fig. 2(a). It is a compact array with no enclosure mounted on the speaker unit; instead, the two speaker units are attached to a baffle plate. The positive-phase sound wave is radiated from the front of the diaphragm of each speaker unit, and the opposite-phase sound wave is radiated from the back of the diaphragm; without signal processing, a region of abruptly reduced sound pressure is formed at the side of the array. At the front and back of the array, the positiveand opposite-phase sound waves interfere with each other, canceling each other out, and the sound disappears. However, close to the front of the array, the sound waves do not cancel each other out because the opposite-phase sound waves from the back of the array do not diffract around the array in time. This phenomenon creates an area where sound is audible near the loudspeaker but inaudible away from it.

By using a speaker array structured with two speakers side by side and applying signal processing that emphasizes the sound near the speaker array and cuts the sound in other directions, it is possible to emphasize the area where the sound remains only in front of



Fig. 1. Comparison of characteristics between bare and normal speakers at different distances.



Fig. 2. Enclosure-less speaker array and chair equipped with the arrays on the headrest.

the speaker. The use of enclosure-less speakers instead of enclosure speakers also affects sound quality. In normal speakers, the speaker unit operates in a narrow enclosure, in which the air exerts a repulsive force that causes a lack of sound reproduction in the low-frequency range. However, this structure, which has no enclosure, is not affected by the repulsive force of air, so the speaker unit can reproduce sound down to its inherently reproducible lower frequencies. An armchair fitted with two of the proposed speaker arrays mounted on the headrest: one on the left and one on the right, is shown in **Fig. 2(b)**. Since the enclosure-less speaker arrays are mounted on the headrest, the listener does not have to wear anything and can hear the sound from the speaker arrays when sitting in the chair. Hardly any sound leaks to the surroundings, so the sound is comfortably delivered to the seated listener only.



Fig. 3. Structure of enclosure of PSZ Wearable and model of Helmholtz resonance.

3. PSZ Wearable for small wearable devices

With the enclosure-less speaker array (using the opposite-phase sound waves radiated from the back of the speaker array), the positive- and oppositephase sound waves cancel each other out at a certain distance from the front of the speaker array and mute the sound; however, in the vicinity of the array, the sound at the back and front of the speaker array is not muted, and sound leakage can be heard. This speaker array also requires signal processing, such as a beamformer using two speakers, which restricts the size and installation conditions.

Considering the above-described issues, we developed an earphone called PSZ Wearable by designing its structure to achieve nearby sound reproduction without signal processing by controlling the opposite-phase sound waves generated from the back of the speaker unit by means of an enclosure.

An enclosure structured with a hole in it—a vented box (or bass reflex)—has been in use since around 1971 [2]. Bass-reflex speaker units are used to lower the low-frequency range that can be reproduced and enhance bass reproduction. Volume, aperture area, and duct length of the enclosure can cause so-called Helmholtz resonance. When this resonance occurs, the phase of the sound is inverted (so-called phase inversion occurs), that is, the phase is shifted by 180 degrees so that the plus and minus portions of the sound wave are inverted, and the opposite-phase sound wave becomes in phase with the positivephase sound wave. By appropriately designing the volume and opening area of the enclosure, Helmholtz resonance can be generated at low frequency. A lowfrequency sound with inverted phase is then emitted from the opening of the enclosure, that is, sound with enhanced bass—but without sound cancellation—is emitted from the enclosure. Thus, a bass-reflex speaker unit can reproduce lower frequency sounds than a single speaker unit.

The structure of PSZ Wearable (Fig. 3) is based on this idea of a phase-inverted enclosure, which it uses for suppressing sound leakage rather than enhancing bass [3]. Since PSZ Wearable requires the use of the opposite-phase sound wave radiated from the back of the loudspeaker, its phase correlation with the positive-phase sound wave radiated from the front of the loudspeaker unit must be maintained up to high frequencies. Therefore, PSZ Wearable is designed to maintain the opposite-phase correlation by setting the Helmholtz resonance in a frequency band higher than that for which sound leakage must be suppressed. The controllable elements of Helmholtz resonance are volume of the enclosure, aperture area, and duct length; however, in the case of small wearable devices, duct length is very small, and aperture area cannot be freely designed due to the need to radiate the opposite-phase sound wave. PSZ Wearable is therefore designed to have the minimum enclosure volume to increase Helmholtz resonance and maintain the opposite-phase correlation.

The frequency band that can be suppressed by the structure of PSZ Wearable is determined by the path difference δ between the positive- and opposite-phase sound waves at the observation point and wavelength λ of the sound. The λ of sound is the distance between



Fig. 4. Structure of PSZ Wearable and results of acoustic simulation of sound leakage.

two points in phase on the wave (such as peaks), and the period is the time for one complete cycle of the wave. The λ thus represents the spatial extent of one complete cycle (period) of the wave. At higher frequencies, λ is shorter, and at lower frequencies, λ is longer. To suppress sound with PSZ Wearable, it is necessary to satisfy $\delta / \lambda \approx 0$. That is, PSZ Wearable works at sound frequencies at which δ is sufficiently smaller than λ . PSZ Wearable consists of a normal speaker unit and enclosure with an opening. Accordingly, there is a limit to which the δ that occurs between the radiation positions the positive- and opposite-phase sound waves can be reduced, and that limitation determines the frequency band that can be suppressed. Helmholtz resonance is also used to expand the bandwidth in which the sound leakage can be suppressed. By using phase inversion due to Helmholtz resonance and inducing inversion at high frequencies at which leakage cannot be suppressed, the path of the sound waves is apparently shortened by half a wavelength, so it becomes possible to suppress sound leakage at even higher frequencies.

The results of an evaluation through acoustic simulation [4] of the suppression of sound leakage with PSZ Wearable are shown in **Fig. 4**. When a sound with a frequency of 1 kHz is played from an earphone with a normal enclosure, the sound is spread around the head as well as around the ears; in other words, the sound leaks into the surroundings from a speaker with a normal enclosure. With PSZ Wearable, however, high sound pressure can be observed near the ear. However, as the distance from the ear increases, the sound pressure drops significantly; in other words, sound leakage is suppressed. The results of measuring sound leakage from PSZ Wearable in an anechoic chamber when a three-dimensional-printed housing was attached to the ear of the dummy head are shown in **Fig. 5**. When sound level of 80 dB is heard at the dummy head's ear, as would be normal when listening to music, the sound level is reduced to 42 dB at a distance of 15 cm from the ear. That sound level (42 dB) is generally considered to be about as quiet as a library, and the performance of PSZ Wearable is such that the sound is almost inaudible from a distance of only 15 cm.

4. Research and future development to expand the application scope of spot-sound reproduction

Spot-sound reproduction is now possible with enclosure-less speaker arrays and PSZ Wearable, which allow sound to be reproduced in the vicinity of the speaker by using one or two loudspeakers. Although spot-sound reproduction is suitable for open-ear earphones, where the speaker that can be located near the ear, and for speakers mounted on the headrest of a chair, our future research and development is aimed at usage scenarios that require a distance between the speaker and the ear and at more freely controlling the range of the sound spot. When speakers are mounted on the headrests of seating in aircraft and vehicles, to reduce the weight and cost of

Fig. 5. PSZ Wearable sound-leakage measurement result.

Fig. 6. Prototype of opposite-phase-wave-induction enclosure and results of acoustic simulation.

the sound equipment, it is necessary to structure the speaker with a single speaker unit at a lower manufacturing cost. We have therefore developed an opposite-phase-wave-induction enclosure that enables spot-sound reproduction with the enclosure structure of PSZ Wearable even with larger speakers. Speakers must be able to reproduce a wide bandwidth (from bass to treble) and be able to output high sound pressure; however, such a wide bandwidth and output sound pressure are affected by the size of the speaker's diaphragm. Compared with earphones, headresttype speakers are placed further away from the ear, so they require higher output power. In accordance with the design guidelines for PSZ Wearable, the oppositephase-wave-induction enclosure has a structure that reduces the enclosure volume and increases the frequency of Helmholtz resonance, while reducing the path difference between the opposite- and positivephase sound waves, to match the large loudspeaker aperture (**Fig. 6**). This structure enables cost-effective spot-sound reproduction with a single loudspeaker without the need for signal processing.

We intend to use the opposite-phase-wave-induction enclosure [5] to develop a smart speaker system that enables only those who need to hear the sound, such as able-bodied people or visually impaired people, by mounting it on car-seat headrests and public-announcement loudspeakers. We will promote research and development for using this smart speaker system in a wide range of places as a speaker that changes the conventional wisdom of sound.

References

- M. Fukui, K. Kobayashi, and N. Kamado, "A Seat Headrest Loudspeaker System with Personalized Sound Zone Capabilities," 2024 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, pp. 1–2, 2024. https://doi.org/10.1109/ ICCE59016.2024.10444201
- [2] A. N. Thiele, "Loudspeakers in Vented Boxes," JAES, Vol. 19, Nos. 5

and 6, pp. 382-392, 471-483, 1971.

- [3] H. Chiba, T. Kako, and K. Kobayashi, "Proposal of the Open-back Enclosure Design for Open-ear Hearable Devices to Reduce Sound Leakage," Proc. of the 2022 Autumn Meeting of the Acoustical Society of Japan, pp. 417–418, 2022 (in Japanese).
- [4] T. Kako, H. Chiba, and K. Kobayashi, "Simulation of Open-back Enclosure for Open Ear Hearable Device," Proc. of the 2022 Autumn Meeting of the Acoustical Society of Japan, pp. 415–416, 2022 (in Japanese).
- [5] T. Kako, H. Chiba, and K. Noguchi, "Proposal for an Opposite-phase Sound Wave Induction Enclosure for Near-field Reproduction with a Single Loudspeaker," Proc. of the 2023 Autumn Meeting of the Acoustical Society of Japan, pp. 351–352, 2023 (in Japanese).

Hironobu Chiba

Research Engineer, Ultra-Reality Computing Group, NTT Computer and Data Science Laboratories.

He received a B.E. and M.E. in computer science from the University of Tsukuba, Ibaraki, in 2013 and 2015. He started his career as an R&D engineer at Pioneer Corporation in 2015 and joined NTT in 2019. His current research interests include hearable devices and acoustic-signal processing. He is a member of the Acoustical Society of Japan (ASJ). He was the recipient of the Technical Development Award by ASJ in 2023

Tatsuya Kako

Senior Research Engineer, Ultra-Reality Computing Group, NTT Computer and Data Science Laboratories.

He received a B.E. and M.E. in information science from Nagoya University, Aichi, in 2009 and 2011 and joined NTT in 2011. He has been engaged in research on microphone-array signal processing. He is a member of ASJ, the Institute of Electronics, Information and Communication Engineers (IEICE), and the Institute of Electrical and Electronics Engineers (IEEE). He was the receipient of the Awaya Prize Young Researcher Award and the Technical Development Award by ASJ in 2023.

Hiroaki Ito

Senior Research Engineer, Ultra-Reality Computing Group, NTT Computer and Data Science Laboratories.

He received a B.E. in electronics and M.E. in information science from Nagoya University, Aichi, in 2007 and 2009 and joined NTT in 2009. His current research interests include acousticsignal processing and sound-field control. He is a member of ASJ, IEICE, the Institute of Image Information and Television Engineers, and the Audio Engineering Society (AES).

Kenichi Noguchi

Senior Research Engineer, Ultra-Reality Computing Group, NTT Computer and Data Science Laboratories.

He received a B.E. in electronic physics and M.E. in human system science from Tokyo Institute of Technology in 2001 and 2003 and joined NTT in 2003. His current research interests include audio-signal analysis and processing. He is a member of IEICE and ASJ.

Noriyoshi Kamado

Senior Research Engineer, Ultra-Reality Computing Group, NTT Computer and Data Science Laboratories.

He received an M.E. in electrical and electronic systems engineering from Nagaoka University of Technology, Niigata, in 2009 and Ph.D. from Nara Institute of Science and Technology in 2012. He joined NTT in 2012 and NTT DOCO-MO in 2015, where he has been working on speech-signal processing technologies for a speech-recognition system. He is a member of ASJ, IEEE, and AES.

Akira Nakayama

Senior Research Engineer, Supervisor, Group Leader of Ultra-Reality Computing Group, NTT Computer and Data Science Laboratories.

He received an M.E. and Ph.D. in computer science from Nara Institute of Science and Technology in 1999 and 2007. After joining NTT in 1999, he has been engaged in robotics, computersupported cooperative work, and recommendation and people-flow analysis. His current research interests are acoustics and signal processing. He is a member of the Information Processing Society of Japan, the Association for Computing Machinery, and the Robotics Society of Japan.