

Development of a Personalized Sound Zone and Future Outlook

Sumitaka Sakauchi

Abstract

Personalized Sound Zone (PSZ) is the ultimate sound space that enables a world in which one hears only the sounds one wants to hear and others hear only the sounds that one wants them to hear. It will enable new lifestyles in which people can enjoy work and entertainment experiences regardless of location, provide a new acoustic experience by merging real space and virtual sound space, enable self-driving cars in which people seated apart from each other can comfortably have conversations in a space as quiet as a living room, and improve the quality of life by enhancing hearing ability. These Feature Articles introduce the challenges involved in achieving PSZ.

Keywords: Personalized Sound Zone, ultimate sound space, new acoustic experience

1. Acoustic environments toward new lifestyles

The conventional work style of going to an office is being re-evaluated due to work style reforms and the effects of the COVID-19 pandemic; thus, a flexible work style unbounded by place or time is attracting attention. Remote support and remote theatergoing that began as countermeasures to COVID-19 are now taking root as a new culture. Therefore, it is important to establish an environment that can provide work and entertainment experiences in a comfortable manner wherever the user may be. I believe that the “sound” environment (sound space), in particular, is an important factor in this regard.

I describe an ideal sound space taking remote work (working from home) as an example. The scene in **Fig. 1** shows a user participating in a web conference. The voices of the other participants do not leak to the outside so that only the user can hear them, noisy chatter of children playing outside is blocked so that neither the user nor fellow participants can hear it, and sounds that the user would want to hear such as doorbell chimes pass through. Achieving such a sound space in which one hears only the sounds one wants to hear and deliver only the sounds that one wants others to hear should make comfortable remote work a reality.

I believe that this kind of ideal sound space will not simply make everyday life more convenient but also be of help to people suffering from all types of problems. It is said that more than 430 million people

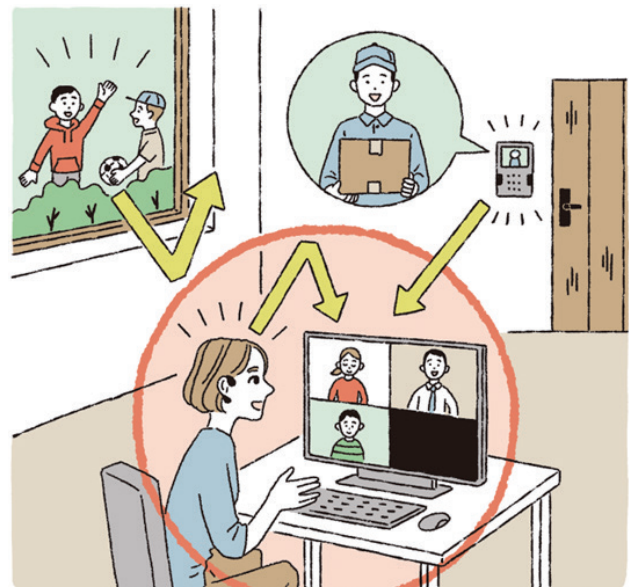


Fig. 1. Example of an ideal sound environment.

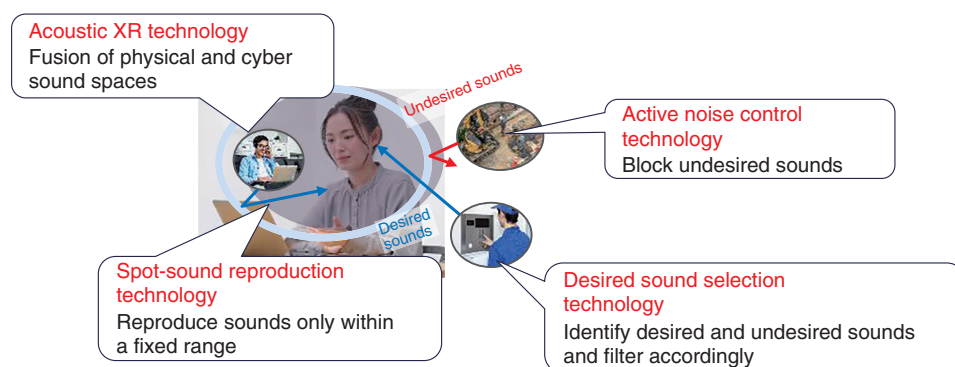


Fig. 2. Elemental technologies for achieving a PSZ.

throughout the world (5% of the world population) suffer from hearing impairments. For these people, it will become possible to adjust the voice of the person one is talking to, if difficult to make out, to a voice that can be heard at an appropriate volume, suppress irritable sounds that one is sensitive to, and detect or notify one of dangerous sounds that would otherwise be missed.

2. What is Personalized Sound Zone?

We at NTT Computer and Data Science Laboratories proposed such an ultimate sound space for each and every person as a Personalized Sound Zone (PSZ) [1]. The concept of a PSZ is to create what is truly one's sound space or sound zone that blocks sounds one does not want to hear from ambient sounds, enables one to hear only those sounds one wants to hear, and prevents "leakage" of one's sounds to other people. The aim is to control this sound space on a person-by-person basis to create a world in which each person can enjoy a comfortable living space as desired.

Achieving PSZ requires a combination of technologies for obtaining information on ambient sounds, understanding that information, and controlling those sounds appropriately. It is necessary to research a variety of technical areas including psychoacoustics that studies the way that people perceive sounds, wave equations that describe the propagation of sound waves, the structure of hardware such as acoustic devices that input/output sounds (microphones and speakers), signal processing, and an understanding of acoustic scenarios. To this end, we are engaged in developing four elemental technologies—spot-sound reproduction technology, acoustic

extended reality (XR) technology, active noise control technology, and desired sound selection technology—by taking a hardware/software fusion approach (Fig. 2).

2.1 Spot-sound reproduction technology: Letting sounds be heard by only the person who wants to hear them

Listening to sounds while preventing those sounds from being heard in the surrounding area had been accomplished by wearing earphones or headphones. This approach, however, has a number of problems such as the inconvenience of wearing such devices, fatigue or even hearing loss due to prolonged use, and difficulty in perceiving surrounding conditions or danger. These problems could be solved and listening made more convenient if spot reproduction could be achieved enabling only the target listener to hear desired sounds without using earphones or headphones. For this reason, we have undertaken the development of spot-sound reproduction technology that eliminates sound leakage by devising speaker enclosures or hardware configurations to emit opposite-phase sound waves, which have the effect of confining sound to an area near the ear [2, 3] (Fig. 3).

2.2 Acoustic XR technology: Creating new acoustic space to customize sound to one's liking

It had been commonplace to enjoy sounds presented from earphones or headphones or enjoy sounds that can be heard from speakers. We proposed acoustic XR technology that enables a new acoustic experience to be enjoyed by using a structure that naturally takes in peripheral sounds—a feature of acoustic devices that prevents sound leakage without covering the ear—and combining and blending ambient

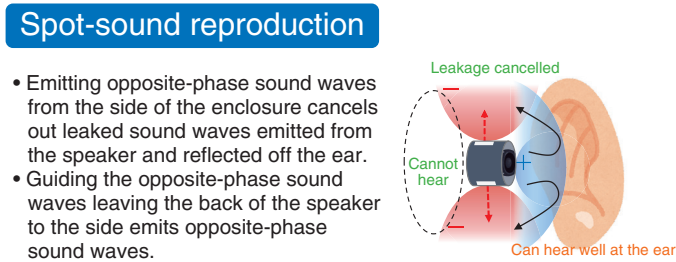


Fig. 3. Application to an earphone using spot-sound reproduction technology.

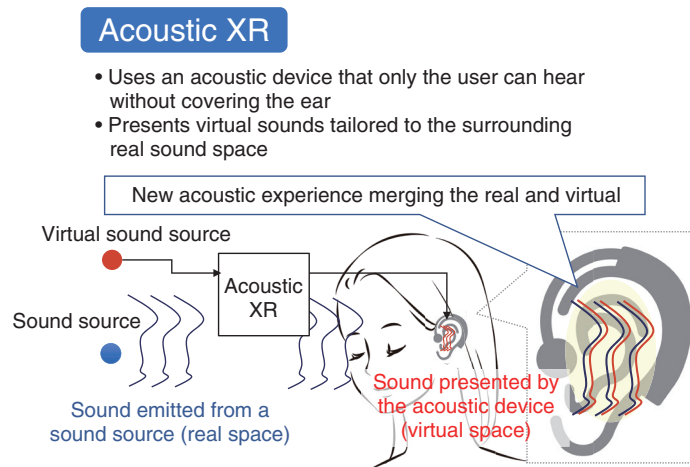


Fig. 4. Acoustic XR technology.

real-world sounds and device sounds for listening. We are applying this technology to stage-based acoustic performances that blend speaker sounds and earphone sounds and the superpositioning of audio guides taking into account real-space sounds [2, 4] (Fig. 4).

2.3 Active noise control technology: Preventing unnecessary sounds from being heard

Noise canceling that is widely used today in earphones and other devices is easy to achieve since the space targeted for canceling sound is small and the path that noise takes to reach the ear is simple. However, wearing earphones for a prolonged period raises the risk of outer-ear inflammation and compromises comfort. If technology that can cancel out unnecessary sounds with a device that does not need to be worn can be achieved, we can expect a greater variety of scenarios that use noise canceling in a more convenient way. This problem can be solved using conven-

tional technology in the form of many microphones and speakers, but we are developing technology for canceling noise without covering the ear with a minimum number of microphones and speakers by optimizing high-speed, low-latency processing and the arrangement of speaker enclosures and microphones [5, 6].

2.4 Desired sound selection technology: Enabling one to hear only desired sounds

This is a “pass only necessary sounds” technology, an element of PSZ. While in the cabin of an automobile, for example, one would want external noise to be blocked but would also want the sound of an ambulance’s siren to reach inside the car’s cabin so it could be noticed as soon as possible. We have achieved this by developing technology that takes acoustic signals observed by microphones and infers the direction that those signals are coming from and the types of sounds that are being generated using a

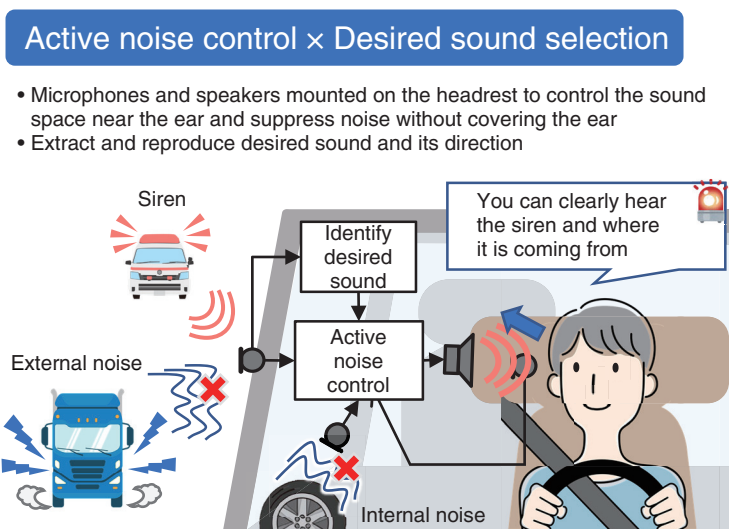


Fig. 5. Example of applying active noise control technology and desired sound selection technology to in-vehicle use.

deep neural network (DNN), and we are applying this technology to cars for extracting desired sounds and their direction and reproducing those sounds (Fig. 5). In such an application, it is necessary to instantly analyze desired sounds arriving from afar, but since they are susceptible to a considerable amount of noise, echoes, etc., identifying their direction and type is difficult. In the face of this problem, we have made it possible to make inferences robust to environmental changes by adapting to the peripheral environment on the basis of the analysis of echoes and background noise [7] and have enabled real-time operation while maintaining accuracy by using the physical features of sound waves based on microphone-array signal processing [6, 8].

Many companies and research institutions are researching each of these technologies. However, the focus is often on improving the performance of a single technology, which means there are various hurdles toward practical application such as the use of acoustic devices that cover the ear or the need for many hardware components. We are focusing on open-ear and no-user-load technologies with the aim of creating new lifestyles and new entertainment experiences by merging actual and cyber sound spaces.

3. Establishment of NTT sonority

NTT sonority, Inc. was established on September 1, 2021 to conduct acoustic-related business using PSZ

elemental technologies developed by NTT Computer and Data Science Laboratories. The company provides PSZ elemental technologies to businesses to incorporate them into aircraft seats, automobile seats, and office chairs, develops, manufactures, and sells portable speakers and wearable devices (earphones and neck speakers) for consumers, and provides next-generation voice digital transformation (DX) services for businesses. In July 2022, NTT sonority began selling earphones using this technology developed by NTT Computer and Data Science Laboratories for confining sound in front of the ear. These products can be used for acoustic production at various types of events and are making it possible to deliver PSZ elemental technologies at an even faster pace to customers [9].

4. Transformation of people's lifestyles tied to sound by PSZ

A variety of technical problems still remain in achieving a PSZ, so we will continue our research and development efforts on the basis of a hardware/software fusion approach. We will collaborate with NTT sonority and other partners inside and outside NTT and work to improve the feasibility of this technology by conducting actual field tests including acoustic productions. Going forward, we would like to make a major impact with this technology by creating new lifestyles in which people wear open-ear acoustic devices as part of daily life much like

eyeglasses so that they can continuously receive acoustic services. In short, we would like to promote transformation of people's lifestyles tied to sound in a wide range of usage scenarios such as remote work, office work, entertainment, and mobility. Finally, we would like to help create a world that is even more enjoyable for people who wear acoustic devices.

References

- [1] M. Fukui, S. Saito, and K. Kobayashi, "Media-processing Technologies for Ultimate Private Sound Space," NTT Technical Review, Vol. 18, No. 12, pp. 43–47, 2020. <https://doi.org/10.53829/ntr202012fa6>
- [2] T. Kako, "Development of Open-ear Earphones that Minimize Sound Leakage by Opposite-phase Sound Waves and Realization of Acoustic XR Service," IEICE Technical Report, Vol. 123, No. 170, EA2023-27, pp. 53–60, 2023 (in Japanese).
- [3] H. Chiba, T. Kako, H. Ito, K. Noguchi, N. Kamado, and A. Nakayama, "PSZ Spot-sound-reproduction Technology: New Sound-confinement Method Using Opposite-phase Sound Waves," NTT Technical Review, Vol. 22, No. 6, pp. 23–29, June 2024. <https://ntt-review.jp/archive/ntttechnical.php?contents=ntr202406fa2.html>
- [4] K. Noguchi, H. Chiba, T. Kako, S. Kozuka, Y. Kurokawa, Y. Watanabe, and A. Nakayama, "Acoustic XR Technology Merging Real and Virtual Sounds," NTT Technical Review, Vol. 22, No. 6, pp. 30–34, June 2024. <https://ntt-review.jp/archive/ntttechnical.php?contents=ntr202406fa3.html>
- [5] H. Ito, S. Kozuka, T. Kawase, and N. Kamado, "Toward the Implementation of an Open-ear Noise Control System," The Journal of the Acoustical Society of Japan, Vol. 80, No. 5, 2024 (in Japanese).
- [6] N. Kamado, T. Kawase, M. Yasuda, S. Saito, S. Kozuka, H. Ito, and A. Nakayama, "PSZ Active Noise Control and Desired Sound Selection Technologies for Creating a Comfortable and Safe Sound Environment in Vehicle Cabins," NTT Technical Review, Vol. 22, No. 6, pp. 35–43, June 2024. <https://ntt-review.jp/archive/ntttechnical.php?contents=ntr202406fa4.html>
- [7] M. Yasuda, Y. Ohishi, and S. Saito, "Echo-aware Adaptation of Sound Event Localization and Detection in Unknown Environments," Proc. of the 47th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2022), pp. 226–230, Singapore, May 2022. <https://doi.org/10.1109/ICASSP43922.2022.9747603>
- [8] M. Yasuda, Y. Koizumi, S. Saito, H. Uematsu, and K. Imoto, "Sound Event Localization Based on Sound Intensity Vector Refined by DNN-based Denoising and Source Separation," Proc. of the 45th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2020), Virtual, pp. 651–655, May 2020. <https://doi.org/10.1109/ICASSP40776.2020.9054462>
- [9] K. Sasaki, "NTT sonority's Pursuit of Innovation—New Businesses That Leverage PSZ and MAGIC FOCUS VOICE Technologies," NTT Technical Review, Vol. 22, No. 6, pp. 44–49, June 2024. <https://ntt-review.jp/archive/ntttechnical.php?contents=ntr202406fa5.html>



Sumitaka Sakauchi

Vice President, Head of NTT Computer and Data Science Laboratories

He received a B.S. in physics from Yamagata University in 1993, M.S. in physics from Tohoku University in 1995, and Ph.D. in systems and information engineering from Tsukuba University in 2005. He joined NTT in 1995 and conducted research on acoustics and speech and signal processing. He is currently engaged in managing the research and development of computer science and data science. He received the Paper Award from the Institute of Electronics, Information and Communication Engineers (IEICE) in 2001, and Awaya Kiyoshi Science Promotion Award from the Acoustic Society of Japan (ASJ) in 2003. He is a member of the IEICE and the ASJ.