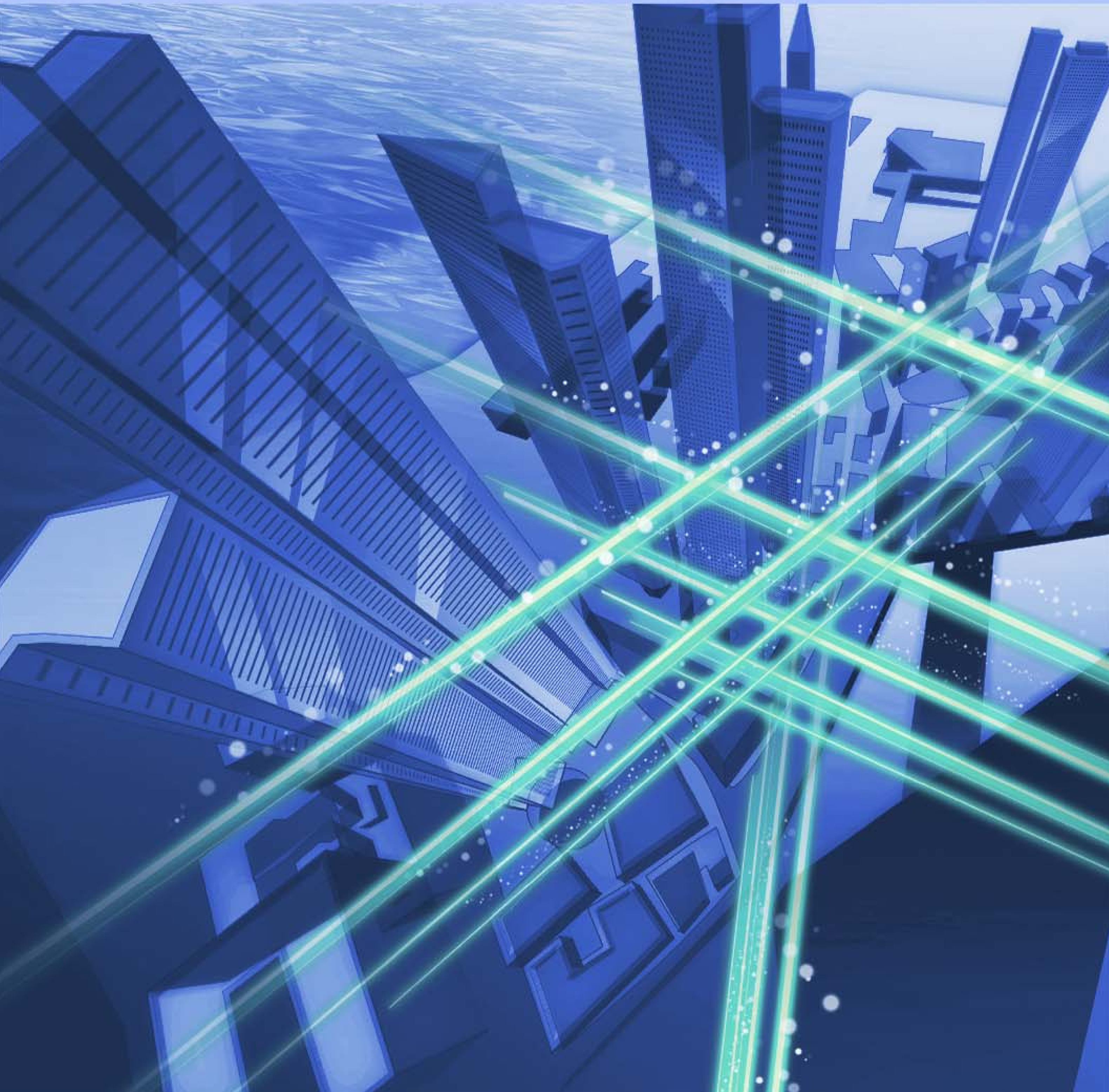
NTT Technical Review

6
2024



June 2024 Vol. 22 No. 6

NTT Technical Review

June 2024 Vol. 22 No. 6

Creating a Future Achieving Both Individual Well-being and Social Well-being That Is Friendly to the Planet by All Employees Becoming Marketers

Sachiko Oonishi
Executive Vice President, Head of Research and Development Market Strategy, NTT Corporation

Abstract

In 2023, the Research and Development Market Strategy Division was established at NTT Corporation. Its mission is to create new value by combining research and development on the basis of the conventional product-out approach with marketing while being committed to creating an exciting future. We asked Sachiko Oonishi, NTT executive vice president, head of Research and Development Market Strategy, about the Division and her mindset as a top executive.

Keywords: product-out perspective, market-in perspective, well-being

Establishment of the Research and Development Market Strategy Division—NTT's first of its kind

—Would you tell us about the Research and Development Market Strategy Division?

The Research and Development Market Strategy Division was established in June 2023 by merging two functionally independent departments, i.e., the Research and Development Planning Department, which has been managing NTT's research laboratories, and the New Business Promotion Department, which had been developing new businesses for the NTT Group. Mission of the Division is to create new value by combining research and development (R&D) on the basis of the conventional product-out approach with marketing.

In 1890, 134 years ago, the telephone first appeared in Japan, marking the birth of technologies that connect people. Following this first telephone development, the Electrical Communication Laboratory under the Ministry of Communications of Japan—the predecessor of NTT laboratories—was founded in 1948, and NTT began research on optical fibers in 1966. Technology that connects people has evolved into technologies that connect people and information, people and objects, and the real and virtual worlds, so it becomes possible to convey everything from sound to images and data as well as skills, experience, and space. Fifty years of research on optical

technology have led to envisioning the concept of the Innovative Optical and Wireless Network (IOWN). This evolution of technology was unimaginable at the time the Electrical Communication Laboratory was established. Nevertheless, the external environment of our business is changing dramatically, as the global environment, social structures, geopolitics, etc. change, and people's values are diversifying and shifting from material satisfaction to spiritual fulfillment.

In the midst of these changes, it has become increasingly difficult to satisfy market expectations with conventional approaches alone. Therefore, the Research and Development Market Strategy Division was established with a focus on how to advance R&D while incorporating market needs and insights to create markets.

—What is the structure of the Research and Development Market Strategy Division?

The Division consists of three departments: Research and Development Planning Department, Market Planning and Analysis Department, and Alliance Department. In addition to the existing Research and Development Planning Department, the Market Planning and Analysis Department was created. I believe that by incorporating a market-in perspective during the practical application phase of new technologies created in the conventional product-out manner, those technologies can be further developed leading to social implementation and market creation. Specifically, we will align our R&D technologies to the market by refining them from the viewpoint of to whom, what kind of value, and what features we appeal to the market. The mission of the Market Planning and Analysis Department also includes identifying promising R&D areas by deriving insights from market and technological trends. To accomplish this mission, we are also building a data-driven marketing infrastructure that will enable us to identify promising areas by using data to visualize NTT's capabilities, such as products and services, customer base, and human resources, as well as markets and external environments.

In a rapidly changing market, there is a limit on creating markets and innovation by NTT alone. Therefore, the Alliances Department is responsible for strategizing collaboration and co-creation with various partners to expand our R&D results "from 1 to 10 and 100."

Breaking through unconscious bias and creating new value

—I understand that the best and brightest of the NTT Group have been brought together in the Market Planning and Analysis Department. Would you tell us about the strengths of the Department?

The Market Planning and Analysis Department needs a variety of perspectives to enable us to constantly create services and value that people are excited about and want to continue to use with peace of mind by responding to their diverse values. Members of the Department include researchers, people who have worked on alliance businesses, people who have developed products, and people who have built their careers at other companies outside the NTT Group. They also vary in age and life stage. Members who have been in business with various roles and different perspectives on life bring diverse perspectives to the table. This diversity is a huge advantage.

However, we have not yet developed a framework, in terms of organization and structure, that translates these diverse perspectives into specific marketing functions and links them to activities and results. Firstly, I believe it is important to align the understanding of the market-in approach. Then, by combining the conventional product-out perspective and market-in perspective, it is possible to accelerate innovation and market creation.

I consider that the product-out perspective focuses on the "function" of a product and the market-in perspective focuses on the "value" of a product. Let me explain these two perspectives by taking a bicycle as an example. From the viewpoint of the function, "the tires move on all roads without becoming flat" or "the bicycle speeds up when the rider pedals lightly" can be considered. If we try to refine the bicycle from the viewpoint of the function only, we might pursue the durability of tires by making them wider and thicker. From the viewpoint of value, however, the value of a bicycle required by people changes according to the user and situation, as saying "My train station is far away, and I want to reduce my travel time to the station, but I cannot use a car because car parking is not available," "I have to carry heavy luggage," or "I need to drop off my child at daycare." To deliver such value, providing means of transportation other than bicycles can also be considered. In that case, instead of improving the durability of tires and making them wider and thicker, it may be necessary to improve other elements, thereby converting the bicycle to technology and products that people want to use and required by markets. In other words, a change in perspective may lead to different points of refinement and pursuit. I believe that by taking advantage of the diverse backgrounds and perspectives of our members, we can look at the products we are currently refining from the viewpoint of value and consider whether a different approach to refinement is needed or create new value by using a different refinement approach.

When incorporating these new perspectives into our R&D and commercialization processes, however, we tend to become stuck in our conventional ways of thinking and doing, and we may have an unconscious bias based on our experience and knowledge; thus, there are hurdles in incorporating new perspectives. It will take time, but we will work diligently to overcome these hurdles and incorporate new values and perspectives into our R&D and commercialization processes.

—NTT has announced new products one after another, and among them, IOWN and NTT's large language model (LLM) "tsuzumi" are attracting much attention both in Japan and abroad. Would you introduce them from a marketing perspective?

It has been five years since the concept of IOWN was announced. Now in the phase of shifting from conception to implementation, IOWN is becoming ever more popular as a group of technologies as a game-changing social infrastructure. In March 2023, NTT EAST and WEST launched the IOWN All-Photonics Network (APN)1.0 service, and in March 2024, NTT Communications began offering the "APN Leased Line Plan powered by IOWN," an inter-prefectural communication service that meets customer needs for a more-advanced communication

infrastructure.

Grown out of the exploration of technologies that connect people and R&D results based on the product-out approach, IOWN supports a data-driven information society in an energy-saving and sustainable manner through technological innovation by achieving a 100-fold increase in energy efficiency and 125-fold increase in transmission capacity. This view of IOWN is taken from a product-out perspective. Looking at IOWN from the market-in perspective raises two questions: what social issues can it help solve and what value can it create in people's lives? All information will be digitized, and artificial intelligence (AI) will be able to function with low power consumption, high speed, and high efficiency through IOWN. Therefore, new value will be created in ways such as eliminating food loss and clothing loss by enabling visualization, optimization, and personalization. I believe that such value creation in everyone's social lives from food, clothing, and shelter to healthcare and entertainment will enable us to build a future in which both social well-being that is friendly to the planet and individual well-being are achieved while bringing back humanity and the five human senses.

Regarding tsuzumi, which is an NTT's LLM and smaller and lighter than current generative AI chatbots such as ChatGPT but achieves the world's top-level Japanese processing performance, NTT DATA and NTT Communications launched commercial services in March 2024. Since the announcement of the development of tsuzumi in November 2023, we have received more than 500 inquiries from corporate customers and local governments. Two-thirds of those inquiries indicated that they wanted to use a customized LLM by training it on their internal data. We have received inquiries from a wide range of industries, many of which handle highly confidential data in, for example, manufacturing, local government, and finance. They have high expectations for using tsuzumi in the following scenarios: improving customer experience (CX) at customer touchpoints such as call centers; improving employee experience (EX) by streamlining internal operations such as through automatic creation and summarization of meeting minutes and creation of questions & answers from operational manuals; and automating information technology (IT) operations and software development. In preparation of the commercialization of tsuzumi in March, we addressed the above-described customer reactions, expectations, and needs, which were reflected in our solution items. It is another new initiative of the Research and Development Market Strategy Division that incorporates the market-in perspective I mentioned earlier.

The tsuzumi LLM has received strong interest from overseas as well. We plan to develop industry- and business-specific LLMs in multiple languages and incorporate them into NTT assets to be used in global markets. The tsuzumi LLM is simply one AI tool.

Using tsuzumi as a trigger, we will take on the challenge of creating a sustainable world for people and the planet by supporting various improvements in DX and CX.

Flat communication and an open mindset create an upward spiral

—In consideration of the businesses you have been involved in, what is important to you as a top executive?

At my previous job at an NTT operating company, I mainly engaged in creating new businesses and developing new services. For example, my mission was to create a service that would become the next main source of revenue after telephone services, and I was involved in the launch of FLET'S fiber-optic Internet services (FLET'S ADSL, B FLET'S, etc.) when Internet services were shifting from dial-up to flat-rate plans. At the Stadium Wi-Fi Promotion Office, which was established at NTT Broadband Platform after Tokyo was selected to host the Olympics, I focused on promoting the introduction of IT infrastructures in stadiums, such as for soccer, to make watching sports more exciting and emotional experience than ever before. I also visited the United States, which is leading the way in deploying such infrastructures.

What I have always valued is the perspective of users. In the process of creating new businesses or developing new services, we tend to focus solely on the provider's perspective and be biased toward the product-out approach. I have always tried to think from the perspective of users in terms of how to make them want to use our product, rather than thinking about how to sell it. In that sense, I tried to base my decisions on my own experience as a user; that is, I considered whether I would want to use a product or service myself and whether it would excite me.

As the head of the Research and Development Market Strategy Division, when incorporating marketing into the business process, it is essential to take the perspectives of users and consumers into account. To do so, I want NTT employees including myself to maintain a "work in life" attitude. Conventionally, society has viewed work and private life as separate; however, it is a fact that that work can create positive synergy in one's private life, and vice versa; namely, one's private life can create positive synergy for one's work. It is easy to forget that when we, the employees of NTT, provide value in the form of technology, human resources, and social infrastructure, we are also consumers. In that sense, we can bring our perspective as consumers and it is important to integrate such realistic perspective into business. For example, the world is changing in such a way that people are encouraged to take maternity and childcare leave, and when they actually do, they gain the perspective of being a consumer and not just a company employee. I believe that such life events will contribute to promoting marketing activities on the basis of the real perspectives of consumers. Having conducted market research for marketing for many years, I have seen firsthand the discrepancy between actual experience and survey results, and I have come to believe that survey has its limitations. It is therefore important that all 340,000 diverse employees of the NTT Group become marketers so that their perspectives and sensibilities as consumers are reflected in business.

For example, we use smartphones daily and enjoy their convenience, but when we try to sell smartphones as a business, for some reason, we switch our perspective from "how to use" to "how to sell." I want to change that mindset. To promote such change, I want the members of the Research and Development Market Strategy Division to enjoy their work every day.

—What do you keep in mind as a top executive?

I keep in mind flat communication and an open mindset. Information is shared among members regardless of one's position including peripheral, such as background and objectives. With remote work environments in place, it has become easier to create time for all members to connect via on-line tools such as Teams. "Town-hall meetings" are held to share primary information directly with members. By building a new relationship that differs from the hierarchical relationship at work and creating an environment in which people feel comfortable talking to each other, flat communication will gradually emerge in the workplace. Thus, it will be easier to discuss ideas not only from the provider's perspective but also from the consumer's perspective.

NTT President Akira Shimada's message that "Improving EX also improves CX" is also spreading. I hope that NTT researchers will take pride in the fact that they have accumulated an impressive track record and history of creating technologies that have become social infrastructure since the inception of NTT laboratories. The tsuzumi LLM is the result of their research. On top of that, I want them to add a market-in perspective to their research activities, look at the market and society, and pursue exciting research and technological development toward new value creation.

I invite our customers and partners to join us in addressing social issues in a variety of fields—from food to healthcare—and create a future achieving both individual well-being and social well-being that is friendly to the planet.

Interviewee profile

■ Career highlights

Sachiko Oonishi joined NTT in 1989 and became executive manager in charge of regional revitalization at the New Business Promotion Office in 2016. In her career at NTT Communications, she became a senior vice president, member of the board, head of the Third Business Solutions at the Business Solution Division in 2020 and senior vice president, head of the Third Business Solutions at the Business Solution Division in 2021. She has been in her current position since June 2023.

Meta-learning Achieves High Accuracy with a Small Amount of Training Data

Tomoharu Iwata
Senior Distinguished Researcher, NTT Communication Science Laboratories

Abstract

Large language models and generative artificial intelligence for images are increasing in popularity. Their accuracy is improved by training them with a vast amount of documents and images. If vast amounts of data are unavailable, however, it is difficult to achieve high accuracy. Meta-learning for attaining high accuracy through learning to learn with a small amount of training data has been attracting attention as an approach to solving this problem. We interviewed Tomoharu Iwata, a senior distinguished researcher at NTT Communication Science Laboratories, about research trends and future directions concerning meta-learning and his mindset of staying interested in and enjoying research.

Keywords: artificial intelligence, meta-learning, machine learning

Meta-learning—learning how to learn—expands application areas of machine learning

—Would you tell us about the research you are currently conducting?

In the field of machine learning, I'm researching meta-learning for achieving high-performance artificial intelligence (AI) using a small amount of training data. Since I joined NTT in 2003, I have been researching machine learning, and since around 2018, I have been focusing on meta-learning. The concept of meta-learning has been around for some time; however, around 2018, when deep-learning research was progressing apace, computer performance was improving, and various deep-learning libraries became available, meta-learning research gradually gained more attention.

Deep learning is widely used in natural-language processing and image processing, and achieves high performance by learning with large amounts of data. In contrast to deep learning, meta-learning learns how to learn from other related data, so its performance in regard to new tasks can be improved even when only a small amount of training data are available.

I will explain the basic framework of meta-learning by using an image-classification task as an example. For meta-learning, data for various tasks are prepared. An example task is classifying images of cats and dogs or cars and bicycles. For each of these tasks, when task-specific model parameters are learned

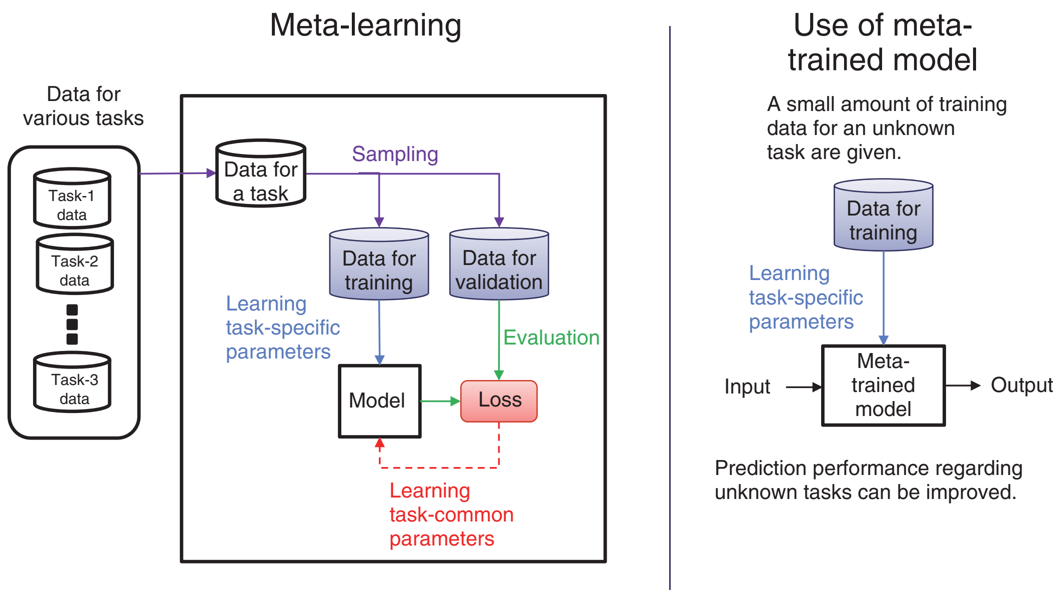


Fig. 1. Framework of meta-learning.

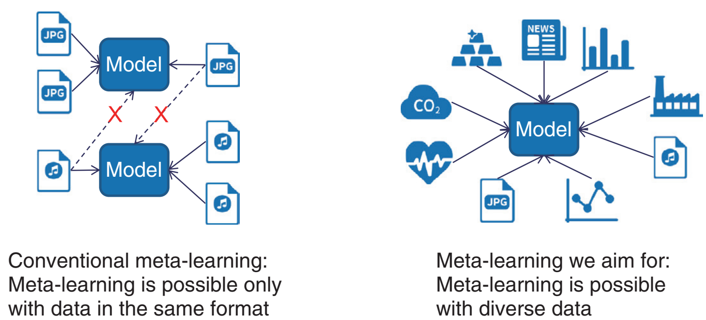


Fig. 2. Meta-learning from tasks in different feature spaces.

using a small amount of training data, the model parameters common to the tasks are updated to improve classification performance with the validation data. This process is referred to as "learning how to learn." Even if the task is a new one that has not been used during meta-learning (for example, classification of images of apples and oranges), it is possible to achieve high classification performance by simply learning task-specific model parameters with a small amount of training data (**Fig. 1**). In addition to the above example, meta-learning is also used for speaker recognition, which involves identifying who is speaking from audio data, and speakers can be classified even when a new speaker suddenly starts to speak; analyzing a language for which little data exist; and automatically classifying books by using only a small amount of data with a new book-classification system that differs from the previous systems.

In the course of my research on meta-learning, my colleague and I proposed a deep-learning model that enables meta-learning from tasks in different feature spaces (**Fig. 2**)—in contrast to existing meta-learning that assumes meta-learning from tasks in the same feature space—and presented the model at the Annual Conference on Neural Information Processing Systems (NeurIPS), a top machine-learning conference, in 2020 [1]. We also proposed a deep-learning model that combines large language models with meta-learning in a manner that enables the use of human-accumulated knowledge in machine learning and presented that model at NeurIPS 2022 [2]. I've also been conducting research to improve machine-learning performance by applying meta-learning to a variety of tasks such as clustering, uncertainty estimation, spatial analysis, causal inference, anomaly detection, and feature selection.

—Meta-learning can overcome the weaknesses of deep learning. What direction will its research and applications take in the future?

Machine learning and AI have made remarkable progress, and generative AI, such as ChatGPT, has emerged, but their success is based on having a large amount of data for training them. However, there are many cases in which sufficient data cannot be obtained due to the high cost of requiring expert knowledge, such as in medical images, and due to privacy protection. There is also a case in which there are almost no data because it is a new product or service in recommendation systems. Meta-learning makes it possible to use machine learning in application areas where only such a small amount of data are available.

Our meta-learning can currently handle a variety of data that can be expressed in tabular form but cannot handle more complex data such as data obtained on the factory floor or biometric data. Therefore, we are investigating meta-learning to expand the range of data that it can handle, which should allow us to expand the application area of meta-learning. In regard to the above-mentioned research on meta-learning using large language models, our research on meta-learning also aims to create AI that can address new challenges by using a variety of experiences and knowledge in the same manner as humans do.

Enjoying research while staying interested in it

—What do you keep in mind when choosing a research theme?

I choose research themes on the basis of my own curiosity. The meta-learning that I'm currently focusing on is also relevant to the evolution of life, which I was interested in and studied when I was a student. I hope that research on meta-learning, by which AI learns how to learn from various tasks, will also lead to understanding how humans have evolved to be able to learn from various things.

I used to read papers in areas of my interest and found problems to turn them into new research themes. I also combined things I was good at with things my colleagues were good at to create new themes. I sometimes found problems when I looked at new data and turned them into themes. When I was previously researching recommendation systems, for example, such systems were designed to recommend people to buy more products or use more services. However, in recommendation systems for subscription-type services or flat-rate services, a different perspective emerged, that is, the goal of recommendation systems is not to get people to buy more but increase their satisfaction.

When I read other researcher's papers, I naturally think about how the findings reported in those papers could be expanded or whether I could combine them with meta-learning, my specialty, or with the probabilistic models and topic models that I had investigated in the past. By looking at other researcher's papers and research on the basis of my specialty, I come up with new perspectives and new research themes. About ten years ago, I was posted to the UK for a year as a visiting researcher. Since I thought it was a great opportunity to collaborate with the people there, I asked students and post-doctoral researchers what they were researching, and I wondered if I could combine their research with my own specialties. I was therefore able to learn new things, write several papers on new themes, and enjoy life in the UK through researching with a variety of people.

—What do you keep in mind as a researcher?

I'm always willing to try different things. I like inventing and implementing new methods, so I come up with a lot of ideas, experiment with them, and repeat trial and error. Naturally, I often fail, but I believe that by trying new things, I can learn a lot and acquire unexpected knowledge that will lead to future research.

When researching machine learning, I want to solve problems as clearly and as simply as possible. I first clarify a goal in regard to the problem I want to solve, determine the loss that can be quantitatively evaluated, and minimize the loss. Solving problems neatly on the basis of theories, such as probability theory, will improve the performance of a machine-learning model. Its evolvability and extensibility will also increase in a way that leads to subsequent research and contributes to the research of others. Since that approach is easy to understand, I try to create a model by identifying only the important points so that the model can deliver dependable performance with the minimum amount of work required.

When I joined NTT, I was assigned to the machine-learning group, which was a completely different research field from the one I studied at my graduate school. At first, I didn't know anything about machine learning; however, my mentor gave me step-by-step guidance, and I was able to give a presentation at a prestigious international conference in my second year. During this process, I was able to learn about the format of machine-learning research, which starts with problem setting, continues through development and implementation of a model, experimentation, and ends with writing a paper and presenting it at conferences. After that, I came up with my own ideas many times and presented them to my mentor. At first, my ideas were constantly rejected, but as I kept repeating this process, I was able to acquire the ability to research proactively. My encounters with my mentor had a great effect on my life as a researcher, and I hope to have encounters in which I can have a similar effect on others. I would be happy if younger researchers grow as they create their own research styles by referring to other researchers, come up with their own ideas, and have straightforward discussions.

—Do you have a message for younger researchers?

My message is enjoy your research by staying interested in it. Because I enjoy doing research that I'm interested in, I can concentrate when I'm researching and find it rewarding, even if I'm exhausted. I enjoy learning new things that interest me, and I think it is exceptional that researchers have the opportunity to learn things that are yet unknown by investigating problems that have not yet been solved. It has been over 20 years since I joined NTT, and I think I have been able to continue my research by enjoying learning new things along the way. If you choose a research theme that interests you and enjoy learning about things you do not know in the course of your research, you will keep finding new interests and enjoyment.

References

[1] T. Iwata and A. Kumagai, "Meta-learning from Tasks with Heterogeneous Attribute Spaces," Proc. of NeurIPS 2020, Virtual, Dec. 2020.
[2] T. Iwata and A. Kumagai, "Sharing Knowledge for Meta-learning with Feature Descriptions," Proc. of NeurIPS 2022, New Orleans, LA, USA, Nov./Dec. 2022.

■ Interviewee profile

Tomoharu Iwata received a B.S. in environmental information from Keio University, Kanagawa, in 2001, M.S. in arts and sciences from the University of Tokyo in 2003, and Ph.D. in informatics from Kyoto University in 2008. In 2003, he joined NTT Communication Science Laboratories. From 2012 to 2013, he was a visiting researcher at University of Cambridge, UK. His research interests include data mining and machine learning.

High-frequency-band Distributed Antenna System Technology Exploring the Possibilities of Higher Frequency Bands for Mobile Communications

Daisei Uchida
Distinguished Researcher, NTT Access Network Service Systems Laboratories

Abstract

Today, as data capacity continues to increase, the demand for communications capable of higher transmission speeds and capacity is increasing. For sixth-generation mobile communication system (6G) wireless, which is scheduled to be put into practical use around 2030, extreme-high-speed and high-capacity communications with a wireless transmission speed of 100 Gbit/s or more per user is being discussed. To achieve this, new breakthrough wireless technologies are required. In this article, we spoke to NTT Distinguished Researcher Daisei Uchida about wireless technology that enables high-frequency bands to be comfortably applied to mobile communications. He talked about a high-frequency-band distributed antenna system for 6G.

Keywords: high-frequency-band distributed antenna system, WiGig, 6G

Technology for new higher frequency bands for a more comfortable communication society

—Mr. Uchida, what exactly is "high-frequency-band distributed antenna system technology?"

The demand for high-speed and high-capacity wireless communications is growing. "High-frequency-band distributed antenna system" is a promising technology to meet this demand. It can achieve stable high-speed and high-capacity wireless transmission by distributed multiple antennas from one base station when using high-frequency bands such as millimeter waves and sub-terahertz.

To increase data capacity in the future, extreme-high-speed and high-capacity wireless communications of 100 Gbit/s or greater per user is being discussed to support an extensive range of use cases in the sixth-generation mobile communication system (6G) scheduled for launch around 2030. These use cases include high-definition, uncompressed video, three-dimensional information, extended reality

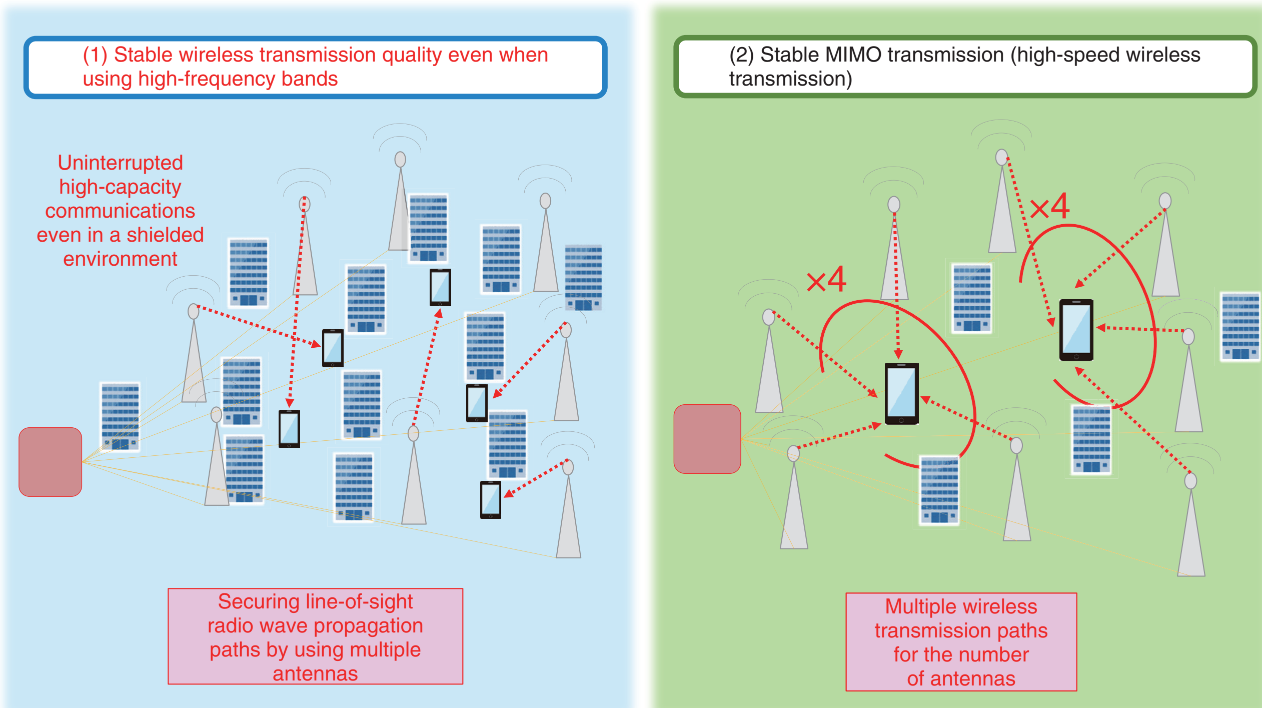


Fig. 1. Advantages of high-frequency-band distributed antenna system technology.

(XR) data, and data related to the five human senses. One means of achieving these high speeds is to expand the wireless signal bandwidth from the conventional range of several 100 MHz to 1–10 GHz and to enable millimeter-wave and sub-terahertz high-frequency bands that can secure such bandwidth to be applied to mobile communications. In this way, we aim to realize extreme-high-speed and high-capacity wireless transmission, which has been difficult to achieve until now (**Fig. 1**).

—What are you specifically undertaking in high-frequency-band distributed antenna system technology?

The first target problem that we had to deal with in achieving this technology is the "ease of interruption" of wireless communications in high-frequency bands. Compared with low-frequency bands, the advantage of high-frequency-band radio is that it can provide stable high-speed and high-capacity wireless transmission when radio wave propagation is unobstructed. Its disadvantage, however, is that radio waves are easily blocked by shielding objects, making wireless communications vulnerable to interruptions and difficult to use. Amid this advantage-and-disadvantage situation, I continued to have strong feelings about high-frequency-band radio such as "Certainly it would be useful to society if radio waves could be made to propagate freely, that is, if they could be made to fill the target area" and "I want radio waves to be useful to society." As a result, a search has been going on for some kind of method that can resolve this dilemma.

To this end, we are conducting research on distributed antenna technology that distributes the antennas radiating radio waves so that those signals can be transmitted from various directions. This technology has been studied since the 2010s in low-frequency bands. If radio waves are emitted from a base-station antenna at one location, the ease of interruption—the disadvantage of high-frequency bands—can easily occur. Consequently, in an environment in which reflected waves from all directions cannot be expected such as an outdoor or large-scale indoor environment, it is difficult to achieve stable high-speed and high-capacity communications in a mobile communications environment. On the other hand, if a distributed antenna technology could be achieved with a mechanism that distributes multiple antennas from a single base station so that a mobile terminal can be instantly switched to a more appropriate antenna at any time, it is thought that this problem of the ease of interruption of high-frequency bands could, in theory, be solved. Specifically, we are aiming to reduce the radio-wave blind areas through "passage" based on distributed antennas in a shielding environment. In this way, we aim to stably provide the potential for extreme-high-speed wireless transmission in high-frequency bands.

One of the key points of this technology is that it communicates with one base station no matter which distributed antenna transmits and receives, rather than switching base stations as in the past. Ordinarily, when switching between base stations, it is necessary to perform a control process called handover in upper layer protocol. The antenna switching of the proposed technology, however, adopts a method that switches on Layer 1, the physical layer. In short, it adopts a mechanism that maintains the connection with a single base station regardless of which antenna is transmitting or receiving signals. Thus, if nearby people or a passing car, for example, should cause instantaneous shielding to occur, this mechanism based on physical control makes it possible to switch in an uninterrupted manner to an optimal antenna for which radio waves are not shielded. Moreover, given that this antenna-switching control process is confined to the base station, it can be performed seamlessly without placing a burden on the network side or terminal side.

At the present stage of research, we are already conducting demonstration experiments in collaboration with NTT DOCOMO. Our final target is to enable the use of high-frequency bands unaffected by machines, people, or other obstacles in expansive areas like factories or event venues having many terminals in which large volumes of data are being transmitted back and forth.

—What other types of technologies are you working on?

In parallel with research on this technology, we are also researching 60-GHz-band wireless local area network (LAN) called Wireless Gigabit, or WiGig. With this technology, our aim is to provide stable high-speed and large-capacity communications even for mobile use cases. More specifically, we aim to contribute to wireless LAN that enables anyone to install base stations without a radio license, and in particular, to contribute to WiGig having an ultra-broadband signal of 1.76 GHz as a means of offloading mobile communications within an area like an event venue or factory having a super high concentration of terminals. These are scenarios that are difficult to support by only a mobile communications system. On first being exposed to WiGig, I was shocked and thought, "Is there really a wireless system with such high transmission speed?" This feeling was the trigger for me to start this research. WiGig allows us to download two hours worth of movie data in several seconds, and it has many advantages the same as conventional wireless LAN such as a compact configuration and useable by anyone with no need for a radio license. For these reasons, it was my wish to spread this technology in society.

However, as in the case of using high-frequency-band distributed antennas, WiGig radio waves are difficult to use due to interruptions caused by shielding, and it was generally recognized at first that "WiGig could be used only for fixed communications in radio relays but would be difficult to use in mobile communications." I therefore began research and development (R&D) to overcome this problem so that WiGig could also be used for mobile use cases.

The clue to solving this problem is the same as that of high-frequency-band distributed antennas. Throughout an area, if a mechanism can be achieved that can deliver signals to a certain location in an environment having nothing that obstructs those radio waves (or that results in just one reflection), I considered that it could be used in a stable manner in mobile use cases too. However, a WiGig antenna operates in the 60 GHz high-frequency band, so it takes on a structure that is included on a substrate to suppress cable loss, which makes it difficult to use a distributed-antenna mechanism. It was therefore necessary to create a mechanism within a wireless LAN framework that could provide a countermeasure to obstructions by using many base stations and enable uninterrupted transmission even for movement between those base stations. This problem could be solved by two technologies: "using a mechanism that switches between base stations with appropriate timing when a terminal moves among multiple base stations" (**Fig. 2**) and "equipping the terminal side with multiple wireless terminals and controlling the connection of those terminals to different base stations" (**Fig. 3**). Our target is to enable instantaneous switching to the appropriate radio link in both a shielded environment and a mobile environment so that WiGig could be provided to mobile use cases.

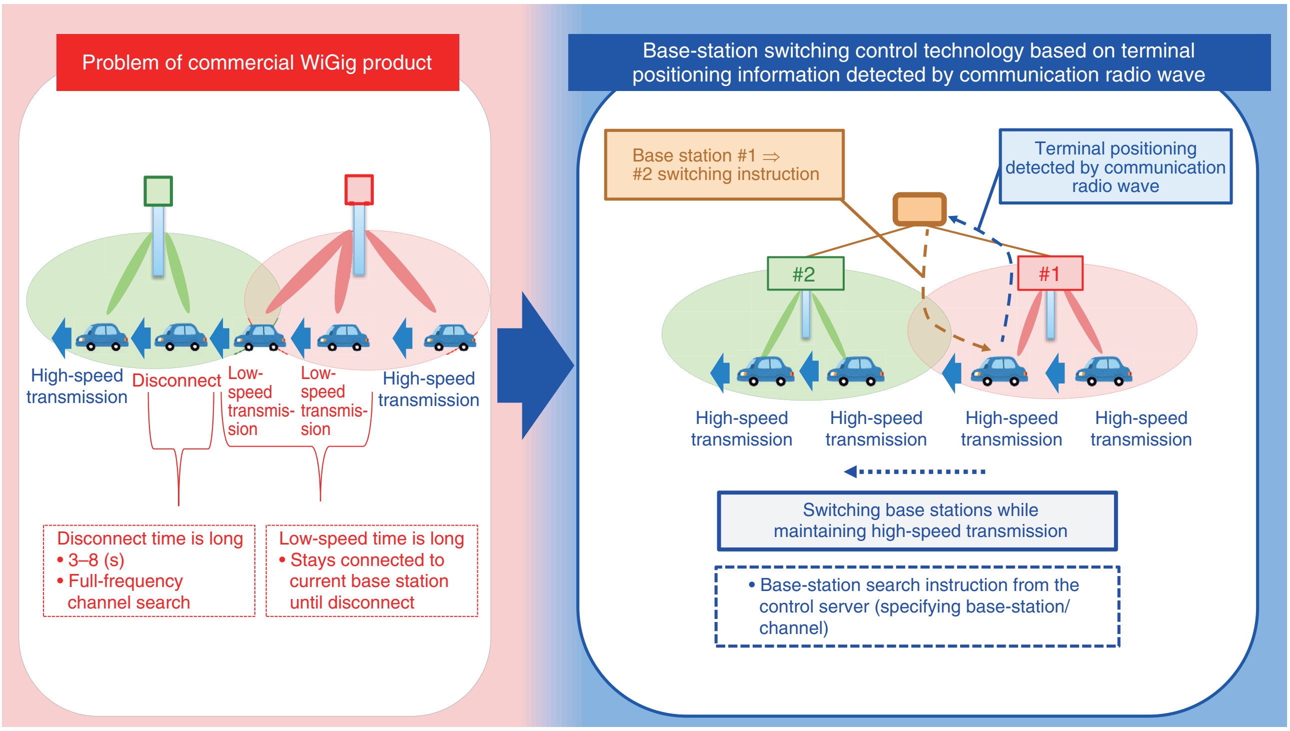


Fig. 2. WiGig mobile-support technology (1): base-station switching control.

—What kind of problems did you encounter during the course of your research?

The greatest difficulty that I encountered in advancing my research was "making it known that we need WiGig." WiGig was introduced with great expectations around 2016, but it was subsequently toned down becoming a "shadow entity" that was labeled as "unusable for other than radio relays." I myself continued to be told by many people "Why is this radio system needed when we already have 4G/5G and wireless LAN?" Nevertheless, I sensed that there was something amazing about WiGig, so I thought that one of my research missions was to make it known to many people and get them to use it. My first step was to put my efforts in getting people interested in WiGig itself and to get them to share my feelings about the need for it and the contributions that it could make. In particular, in the period of 2016–2019, more than half of my R&D activities were spent in getting people to learn about WiGig by making the rounds and holding briefings for many people.

There were a number of times when I encountered difficulties in continuing my research, but what made a great change to this situation and marked a new starting point was the racing-circuit experiment held in 2020. Connecting to a car moving at high speeds by WiGig was my first attempt and a highly risky experiment, but I thought it would have to succeed theoretically speaking, so I took up the challenge. As a result of being blessed with team members and vendors who worked together on this experiment, we achieved success to some extent. This success opened up a path toward the mobile use of WiGig in technical terms, and issuing a press release on the results of this experiment served to inform many people about the existence of WiGig. After this, we continued with our research, and in addition to racing-circuit experiments, we also conducted demonstration experiments in various types of mobile-terminal usage scenarios such as with drones and robots. In 2023, I received a commendation for my work from the Association of Radio Industries and Businesses, and as a result of this and other awards, I came to feel that my research was truly making progress.

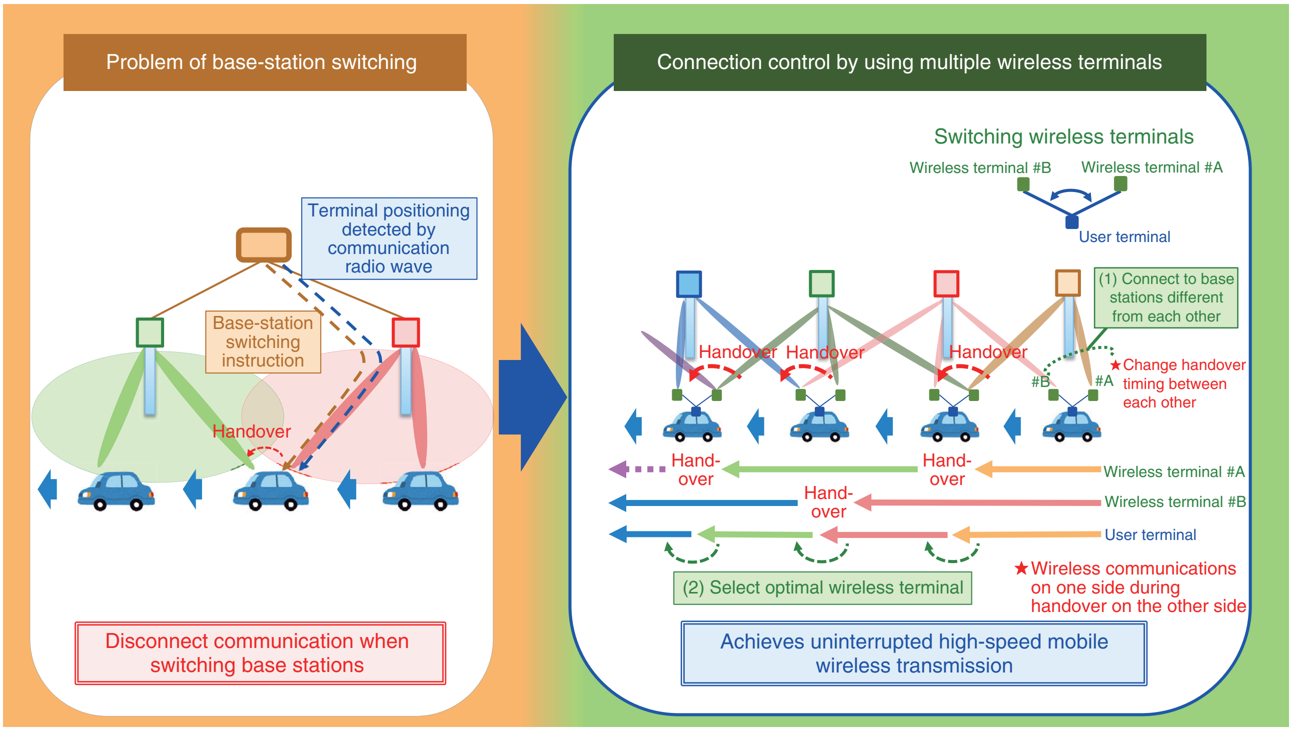


Fig. 3. WiGig mobile-support technology (2): high-speed mobile wireless transmission.

Opening up new frontiers in wireless access and contributing to the implementation of IOWN

—What is the outlook for your research going forward?

One problem that has not yet been solved is that WiGig radio waves can be easily disconnected by obstructions, so "WiGig is greatly dependent on the environment." Going forward, we plan to work on solving this problem by conducting experiments in a variety of environments both indoors and outdoors. Our target is to put WiGig to practical use in one to two years from now for "terminals that are expected to introduce advanced wireless communications in the future" such as robots, drones, XR terminals, cars, and trains.

In addition, I believe that the technical knowledge obtained by such experiments and the experience of conducting wireless transmission in actual environments can be applied to our high-frequency-band distributed antenna system technology using the same high-frequency-band radio waves.

Furthermore, to support the extreme-high-speed and high-capacity wireless transmission of 6G, research of high-frequency bands continues, and if radio waves having signal bandwidths of the 1–10 GHz class called sub-terahertz can be used, it will be possible to provide speeds of 100 bit/s or higher anywhere within a 100 m × 100 m area by applying the distributed-antenna mechanism of my research. This, I believe, will open up new frontiers in wireless access.

Additionally, in terms of a relationship between this technology and the Innovative Optical and Wireless Network (IOWN) proposed by NTT, we are studying whether multiple antennas can be economically deployed by connecting a base station to distributed antennas using the high-speed, high-capacity, low-latency, and low-power features of IOWN. At the same time, using a technology called analog radio-over-fiber (RoF) would make it possible to deliver radio waves as-is from a base station to a remote location over an optical circuit. In this way, the technology that we are researching would become a use case of IOWN natural transmission and a wireless base station format that embodies the worldview of photonics-electronics convergence in IOWN. Moreover, as a radio system with a good high-speed, high-capacity foundation requiring no license, WiGig can assist the high-speed transmission and low-latency features of IOWN. In this manner, WiGig can contribute greatly to mobile-terminal and mobile use cases connected to IOWN.

—Mr. Uchida, could you leave us with a message for researchers, students, and business partners?

NTT Access Network Service Systems Laboratories that I belong to is an organization that researches and develops access circuits that connect the NTT network with customers. This work involves a broad range of fields including circuit-line technology, optical technology, wireless technology, operations technology, and access communications technology. Additionally, in terms of the overall picture of NTT laboratories, another attractive point of NTT Access Network Service Systems Laboratories is its R&D in all sorts of fields in communications with diverse research phases from academic basic research to applied research close to actual implementation. In such an environment, there are many people within the organization to consult with whenever there is something you don't understand or something you'd like to work on in collaboration with another.

In this environment, I myself try to keep in mind the need to hold personal convictions with respect to motivation and starting points in R&D. It is important that all research opportunities start with thinking, "Would I like to use this?" It is through such motivation that I can passionately engulf many people while taking responsibility for my own research. And even if our proposal should receive opposing opinions from people around me or simply be something outlandish, I think it will still be an endeavor worth pursuing as long as I hold onto my convictions. Of course, a journey pursuing something different from what many people are working on can be very painful and filled with anxiety, but there's also the view that the greatest risk is not trying at all. If I don't take up the challenge, I'll never know the answer to "Would it be better to do it or not?" and I'd be filled with regrets while leaving no feedback to following research. Whenever I feel that I'm losing my way, I ask myself, "Is it OK that if I don't do it now I may never be able to do it in my life?"

At the same time, I have to keep in mind the need for judging with an altruistic heart. In actuality, it is extremely difficult to objectively judge my research from my own perspective. For this reason, it is my rule of thumb to stop during the course of my research and ask myself, "Will this research really contribute to the happiness of people as I would like it to?" Doing this often results in research that proceeds well. It is also important that either of these seemingly contradictory requirements of "passion" and "calm judgment" not be lacking. In particular, when promoting a project that involves many people, there is a need for passion to stir up the emotions of the people who will be advancing the research together, but at the same time, the research must be moved forward through calm analysis and judgment from a technical point of view.

At present, NTT laboratories are involved in many R&D projects that have the potential of bringing about a paradigm shift in society under the IOWN vision. Many people and facilities are being organized to make that a reality and many paths are being prepared. Anyone who would like to take up the challenge of developing new technologies, uncovering the possibilities of R&D, and contributing to society in the years to come, I look forward to being able to work with you.

■ **Interviewee profile**

Daisei Uchida received his B.E. degree from Tokyo Institute of Technology in 1994. He received his M.E. degree (applied physics) also from Tokyo Institute of Technology in 1997 and entered NTT in the same year. He has been an NTT Distinguished Researcher since 2022. He is engaged in the research of distributed MIMO technology using high-frequency bands and high-speed mobile support technology for private-radio communications. He received the IEICE Technical Committee on Ubiquitous and Sensor Networks Research Award in 2012, the 34th Radio Achievement Award from the Association of Radio Industries and Businesses (ARIB), Chairman's Commendation in 2023, and the 69th Maejima Hisoka Award in 2024 among other awards.

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

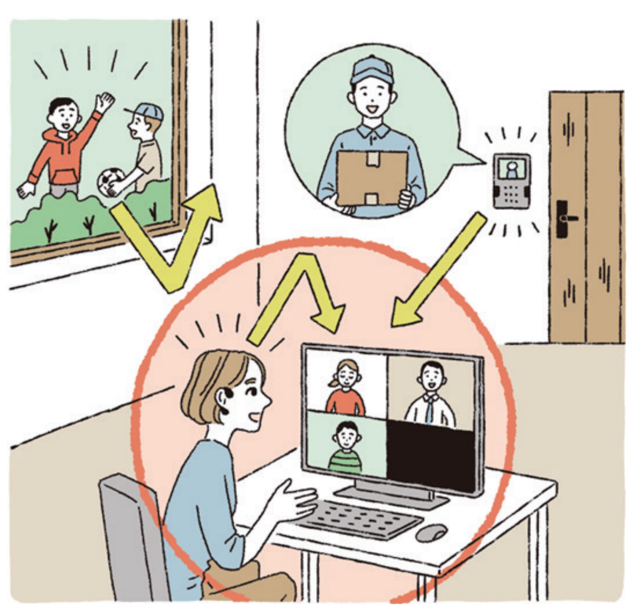


Fig. 1.   Example of an ideal sound environment.

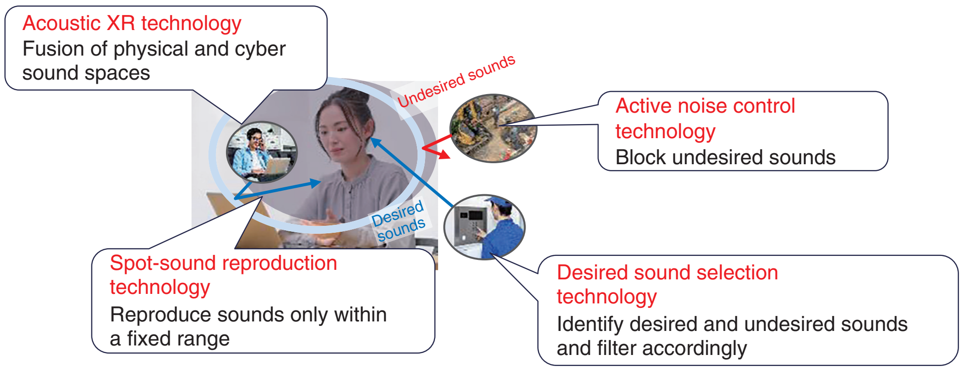


Fig. 2. Elemental technologies for achieving a PSZ.

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

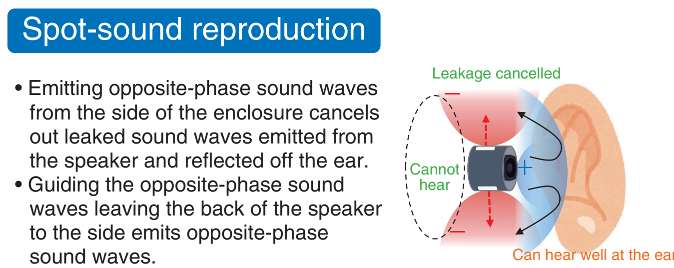


Fig. 3. Application to an earphone using spot-sound reproduction technology.

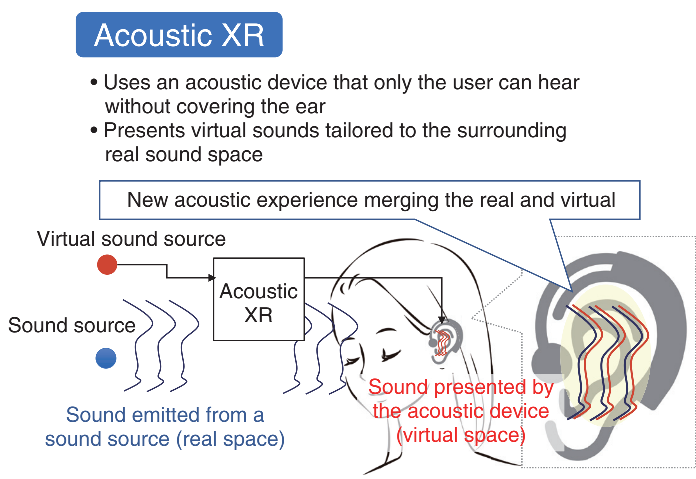


Fig. 4. Acoustic XR technology.

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 conventional 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

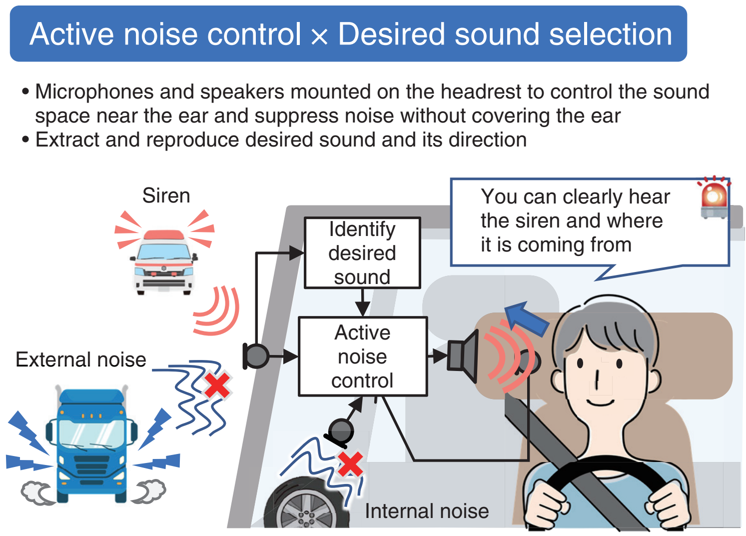


Fig. 5. Example of applying active noise control technology and desired sound selection technology to in-vehicle use.

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.

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 open-ear 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 positive-phase 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 positive- and 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

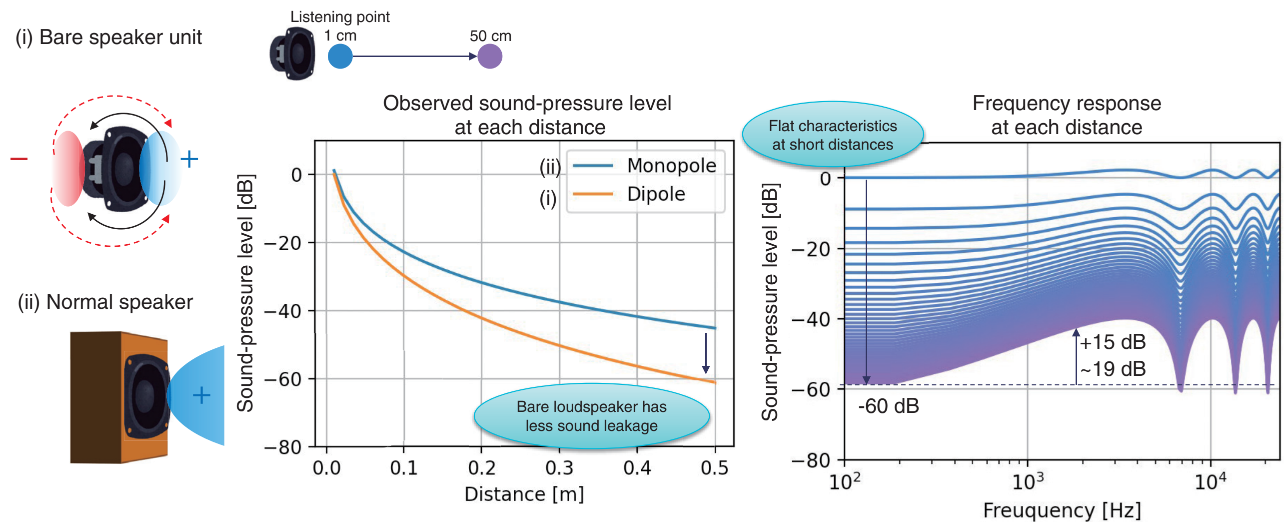


Fig. 1. Comparison of characteristics between bare and normal speakers at different distances.

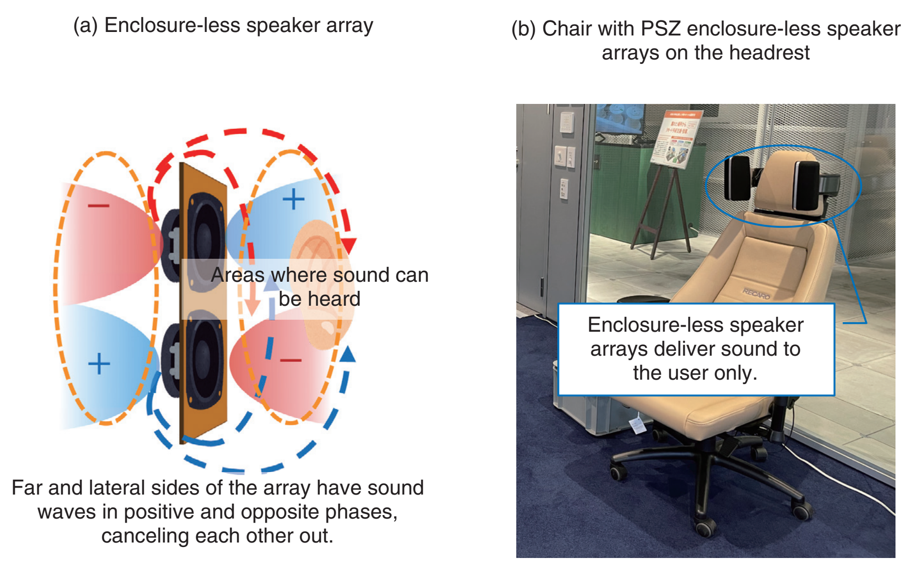


Fig. 2. Enclosure-less speaker array and chair equipped with the arrays on the headrest.

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.

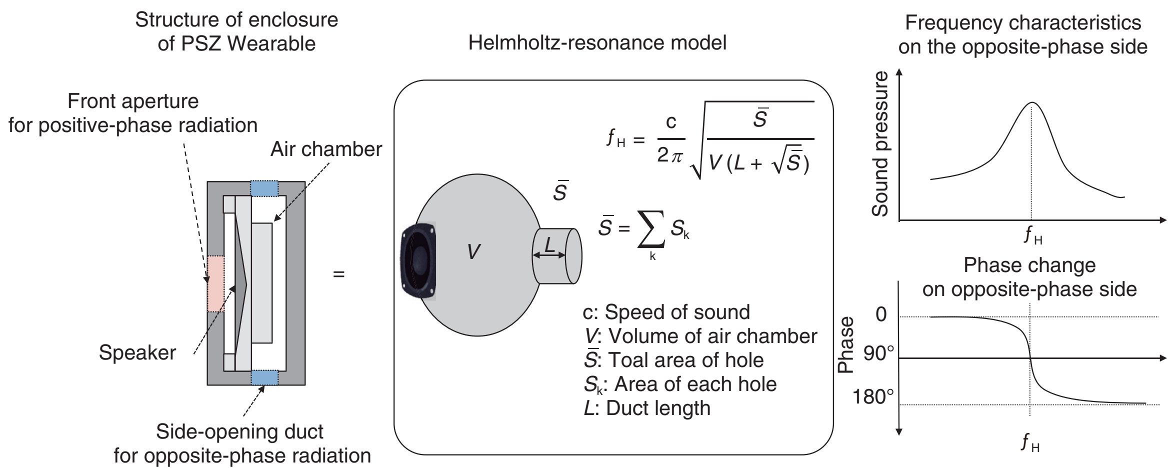


Fig. 3. Structure of enclosure of PSZ Wearable and model of Helmholtz resonance.

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 opposite-phase 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 positive-phase sound wave. By appropriately designing the volume and opening area of the enclosure, Helmholtz resonance can be generated at low frequency. A low-frequency 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

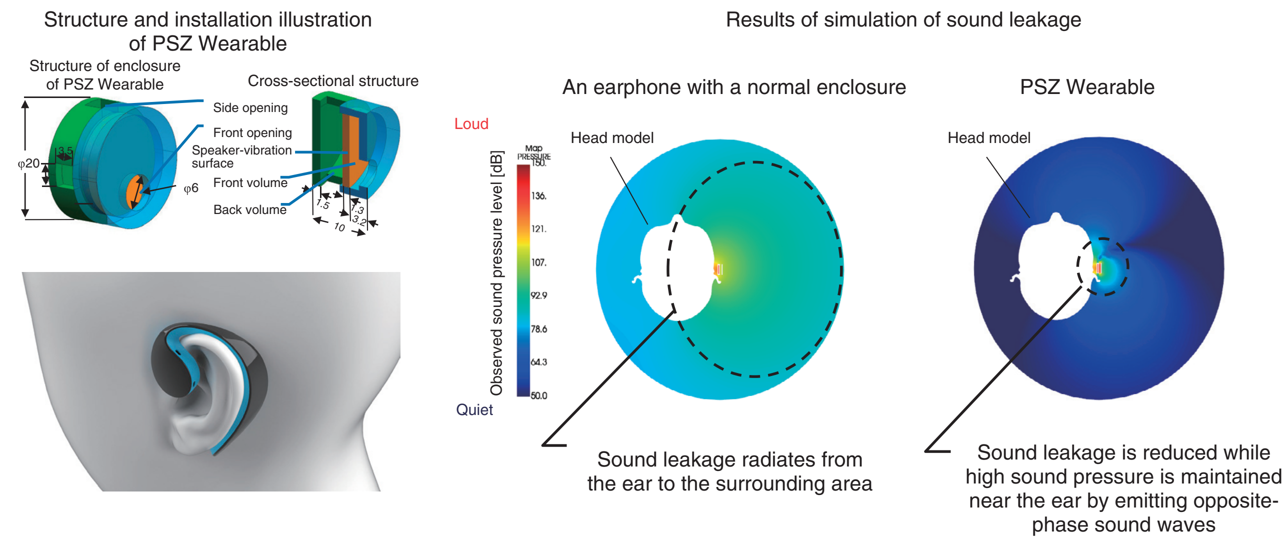


Fig. 4. Structure of PSZ Wearable and results of acoustic simulation of sound leakage.

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

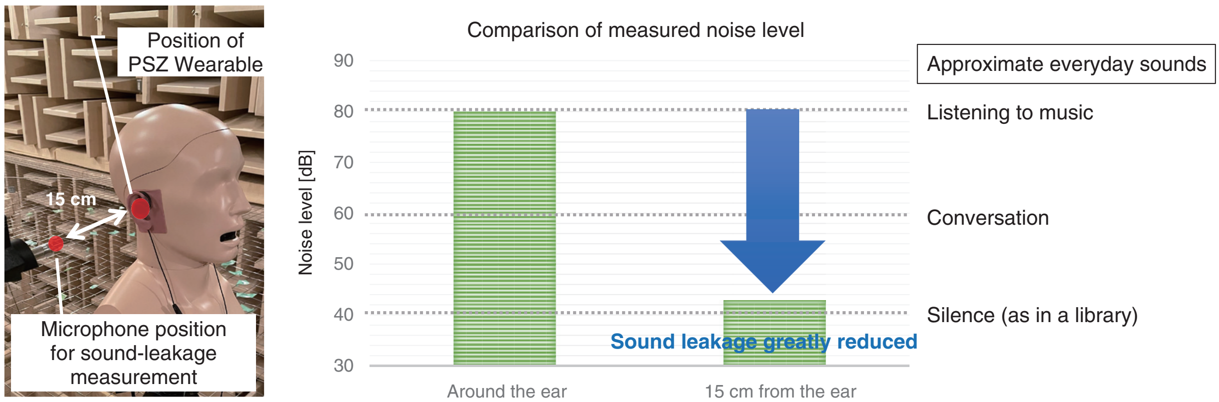


Fig. 5. PSZ Wearable sound-leakage measurement result.

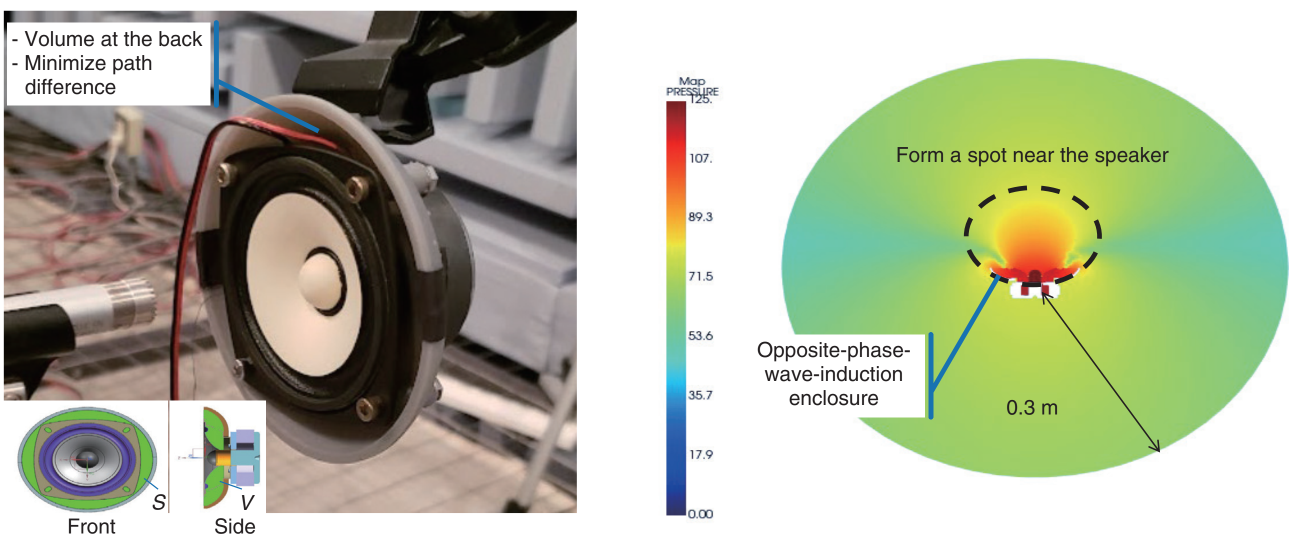


Fig. 6. Prototype of opposite-phase-wave-induction enclosure and results of acoustic simulation.

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, headrest-type speakers are placed further away from the ear, so they require higher output power. In accordance with the design guidelines for PSZ Wearable, the opposite-phase-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 positive-phase 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

[1] 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 recipient 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 acoustic-signal 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 DOCOMO 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, computer-supported 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.

Acoustic XR Technology Merging Real and Virtual Sounds

Kenichi Noguchi, Hironobu Chiba, Tatsuya Kako, Shihori Kozuka, Yoshiaki Kurokawa, Yuki Watanabe, and Akira Nakayama

Abstract

With the spread of open-ear earphones that do not cover the ear, new listening experiences are being proposed that combines real ambient sounds with virtual sounds heard from earphones. At NTT, we call this merging of the real and virtual sounds through open-ear earphones "acoustic XR (extended reality) technology," which we are now developing with a view to actual services. In this article, we describe this technology with a focus on actual trials and touch upon future developments.

Keywords: XR, PSZ, open-ear earphones

1. Merging real and virtual sounds by open-ear earphones

NTT is researching and developing the Personalized Sound Zone (PSZ) as the ultimate private sound space and has developed design technology for an earphone that enables the hearing of sounds by only the user without the ear needing to be covered [1]. A variety of open-ear earphones that do not cover the ear have been appearing on the market and been spreading rapidly. At NTT, we have been focusing on a key feature of open-ear earphones, namely, the ability to naturally hear sounds from one's surroundings, and proposed and begun research and development on acoustic extended reality (XR) technology that merges virtual sounds heard from earphones and real sounds heard directly by the ear.

Acoustic XR technology extends the sounds that can be heard by adding sounds from earphones while listening to ambient sounds. For example, when attending a stage performance or concert, while the sounds generated by venue speakers are ordinarily heard at venue seats, there are still problems in reproducing sounds occurring near the audience, controlling the sense of sound direction and distance, and representing spatial sounds. In contrast, acoustic XR technology that merges sounds from loudspeakers and sounds from earphones will make it possible to present acoustics optimized for a variety of acoustic representations and for individual audience members, which has thus far been difficult to achieve. When attending a sports event at a stadium, acoustic XR technology will enable a user to enjoy commentary from earphones while experiencing the surrounding cheers of the crowd. Similarly, at a multilingual international conference, it will be possible to listen to the translated speech of other participants from earphones while simultaneously sensing nuances in their spoken speech.

As shown in **Fig. 1**, there are two key technical issues in acoustic XR technology: spatial sound for open-ear earphones and virtual-sound spatial rendering. Sounds generated with earphones generally create sound images localized inside the head. However, spatial sound for open-ear earphones can reproduce sounds generated from any position in space outside the head as if they were originating, for example, from the position of a real object. This makes all types of acoustic representations possible. This accommodates shifts in the wearing position of open-ear earphones or differences in individual ear shape and presents spatial sound through such earphones.

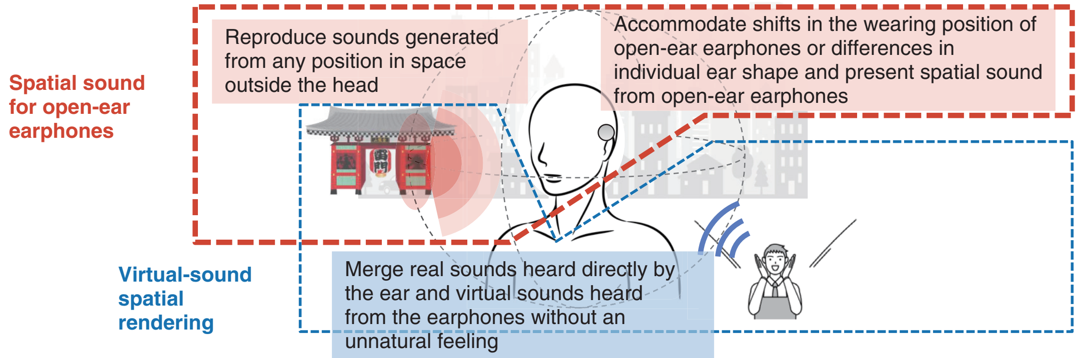


Fig. 1. Technical issues in acoustic XR technology.

Virtual-sound spatial rendering controls earphone-generated sound tailored to the user's cognitive characteristics and merges real sounds heard directly by the ear and virtual sounds heard from the earphones without an unnatural feeling.

NTT conducted several trials using acoustic XR technology in 2023, as described below.

2. Trial 1: Cho-Kabuki (Niconico Chokaigi 2023)

Cho-Kabuki is a performance that combines Kabuki, a traditional Japanese stage performance, with NTT advanced technology. "Cho-Kabuki Powered by NTT—Otogizōshi Koi No Sugatae—" was performed at the Niconico Chokaigi 2023 festival held at Makuhari Messe in Chiba Prefecture April 29–30, 2023. This program provided a spatial-sound performance in which real sounds in the venue and sounds flowing in the ear cross over each other. An audience member generally hears real sounds within the venue such as the actors' voices, music, and various sound effects emitted from venue speakers as well as shouts of admiration or encouragement from surrounding audience members (called *omuko*) directly at the ear. In this trial, we distributed open-ear earphones connected to a radio receiver to about 180 seats in front of the stage to provide those audience members with the experience of listening to sounds from earphones in addition to real venue sounds. These earphones played back sound effects synchronized with the program such as the hoofbeats of running horses, swishing sound of arrows flying through the air, and sound of wind. Combining these sounds with those from venue speakers can reproduce the sound of a horse running from left to right near the listener or the sound of an arrow flying above the listener, enabling a spatial sound production with a high sense of presence.

One problem with spatial sound using open-ear earphones is that the sound image may seem to be located upwards due to the acoustic characteristics of the enclosure, transmission characteristics of the sound propagating from the mounting position of the earphone to entrance of the external auditory canal, shape of the ear's surface, etc. To rectify this problem, we developed a compensating filter that cancels out this upward-sound-localization effect and applied it to sound effects that played back by the earphones (**Fig. 2**). This has enabled the generation of spatial sounds as intended by the content creator.

3. Trial 2: Audio guide for a satellite-transmission performance

A satellite-transmission performance was held to deliver high-quality sound of the 180th NTT EAST NHK Symphony Orchestra Concert held on November 2, 2023 at Tokyo Opera City Concert Hall (Shinjuku Ward, Tokyo) to Hokusai Hall in Obuse Town, Nagano Prefecture as a satellite venue. Hokusai Hall was equipped with a large screen for video playback and 5.1-channel sound equipment to play back high-quality sound. The performance also provided multi-angle delivery that enabled audience members to view the performance from any angle that they liked using a hand-held tablet or smartphone. Using open-ear earphones featuring minimal sound leakage based on NTT technology, we conducted a trial on delivering commentary on the music being performed. Since there was little sound leakage, it was possible to simultaneously hear the orchestra's performance and

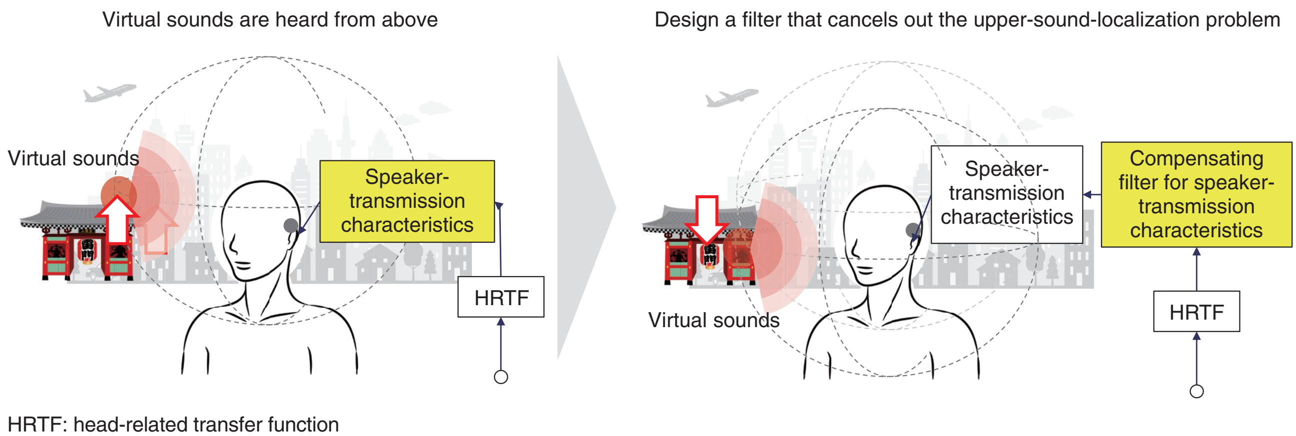


Fig. 2. Concept of applying a compensating filter to the localization position when wearing open-ear earphones.

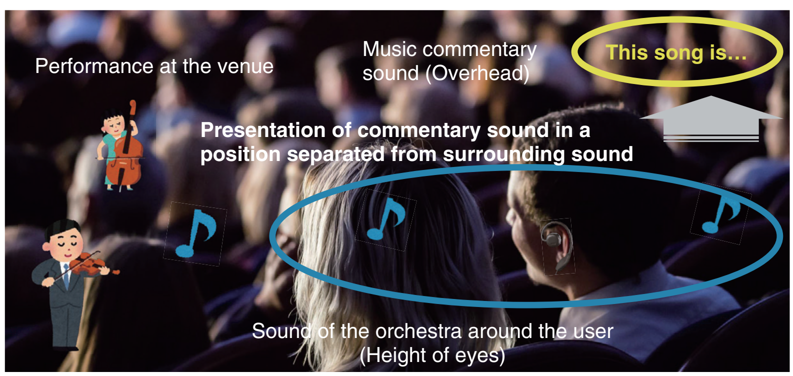


Fig. 3. Control the position of the sound image from the earphones.

commentary audio without bothering neighboring people and without needing to have the ear covered.

When using open-ear earphones, sounds are generally difficult to hear since real ambient sounds and sounds from the earphones overlap. To solve this problem, we controlled the position of the sound image from the earphones so that it appears to be coming from a position not covered by the real ambient sounds. This made it possible to present information by audio means even under conditions in which ambient sounds were being heard by the ear (**Fig. 3**). In this satellite-transmission performance, we used this technology to control the position from which commentary audio from the earphones could be heard somewhat above the user. This had the effect of separating the commentary audio from the sounds of the orchestra emitted from speakers, making it easier to hear and understand the commentary.

4. Trials 3: Audio guide for NTT History Center of Technologies

We conducted a trial of an audio guide using open-ear earphones at the NTT History Center of Technologies inside NTT Musashino R&D Center during the “NTT R&D FORUM 2023 — IOWN ACCELERATION” held November 14–17, 2023. We distributed to visitors a system that estimates user position by a smartphone sensor and plays back an audio guide automatically activated when approaching an exhibit from open-ear earphones connected to the smartphone. Although visitor guide services using earphones are provided at art galleries, museums, and other facilities, this trial stood out because of the following features.

(1) Appreciation of exhibits with open-ear earphones harmonized with surroundings

By not having the ear covered, ambient sounds can be heard naturally, enabling the user to understand surrounding conditions such as visitor congestion. Since earphone sound leakage is small, the user does not have to worry about bothering other visitors. The user can also naturally hear any sound emitted from building speakers or from the actual exhibits. Depending on the environment, it is also possible, for example, to enjoy exhibits while walking and chatting with a friend.

(2) Spatial sound heard as if coming from an exhibit

We enabled spatial sound that appears to be coming from an exhibit by making use of stereo sound playback and presenting signals that simulate the transmission characteristics of sounds propagating from the exhibit to the user's ears. We introduced technology that could compensate for the open-ear-earphone characteristic of localizing the sound image upwards. We also achieved spatial sound that appears to be coming from a stationary exhibit even while the user is moving by estimating user position in real time. To give a concrete example, a user may walk around the stationary exhibit of an old telegraph machine. The system would present sound effects mimicking the machine through the earphones by executing spatial sound processing in accordance with the user's position. This achieves acoustic direction that makes it appear as if sound is coming from the telegraph machine.

(3) Multilingual guide based on cross-lingual speech-synthesis technology

There is a growing demand for multilingual audio guides as the needs of inbound visitors to Japan increase. The cross-lingual speech-synthesis technology developed by NTT enables speech synthesis from speech data only in Japanese to a different language such as English or Chinese while maintaining the same voice quality. We prepared a Japanese/English audio guide using this cross-lingual speech-synthesis technology on the basis of Japanese speech data created by voice actors and enabled audio-guide switching on an app.

5. Future developments

We introduced acoustic XR technology for merging virtual sounds heard from earphones and real sounds heard directly by the ear and described recent trials. Open-ear earphones feature the ability to hear ambient sounds naturally, which suggests a variety of applications in scenarios other than tourism and entertainment such as business and everyday life. For example, they could be used to provide audio guidance to visually impaired persons. We actually tested the experience of walking around our office with a visually impaired person wearing open-ear earphones while listening to the audio guide. In interviews conducted after such experiences, we received similar comments to the following: "I was initially resistant to putting on earphones outside, but they hardly changed the way in which I heard outside sounds." Given that open-ear earphones do not cover the ear, they do not easily cause fatigue even after prolonged use. This should enable a variety of personalized acoustic XR services to be enjoyed depending on the user's current situation while wearing open-ear earphones all day. Going forward, we will work on solving whatever technical problems may arise in expected usage scenarios toward the provision of actual services.

Reference

[1] NTT press release, "Developed earphone design technology that only the user can hear without blocking the ear—NTT developed a single speaker with sound wave control that delivers sound to the user while counteracting sound leakage to the surroundings," Nov. 9, 2022. https://group.ntt/en/newsrelease/2022/11/09/221109a.html

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 the Institute of Electronics, Information and Communication Engineers (IEICE) and the Acoustical Society of Japan (ASJ).

Hironobu Chiba

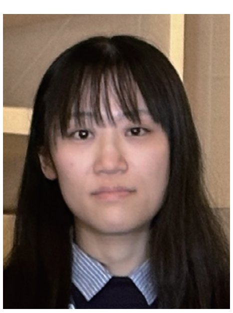
Research Engineer, Ultra-Reality Computing Group, NTT Computer and Data Science Laboratories.

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 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, IEICE, and the Institute of Electrical and Electronics Engineers (IEEE). He was the recipient of the Awaya Prize Young Researcher Award and the Technical Development Award by ASJ in 2023.



Shihori Kozuka

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

She received a B.E. and M.E. in mathematical engineering and information physics from the University of Tokyo in 2020 and 2022 and joined NTT in 2022. Her current research interests include acoustic-signal processing and mathematical engineering. She is a member of ASJ.

Yoshiaki Kurokawa

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

He received a Ph.D. in applied chemistry from Waseda University, Tokyo, in 1997. He joined NTT the same year and has been engaged in research and development on magnetic and holographic recording. Since 2008 he has been engaged in research and development of communication systems and acoustic engineering. He is a member of IEEE and IEICE.

Yuki Watanabe

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

She received a B.E. in electrical, information and physics engineering and M.E. in information science from Tohoku University, Miyagi, in 2021 and 2023 and joined NTT in 2023. Her current research interests include acoustic-signal processing and auditory psychology. She is a member of ASJ.

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, computer-supported 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.

PSZ Active Noise Control and Desired Sound Selection Technologies for Creating a Comfortable and Safe Sound Environment in Vehicle Cabins

Noriyoshi Kamado, Tomoko Kawase, Masahiro Yasuda, Shoichiro Saito, Shihori Kozuka, Hiroaki Ito, and Akira Nakayama

Abstract

Technologies for muting unwanted sounds and for letting wanted (must-hear) sounds pass through, which are elements of a Personalized Sound Zone (PSZ), are described in this article. The elemental PSZ technology for confining sound, called spot-sound-reproduction technology, is first described. The combination of this technology with the new active noise control technology and sound event localization and detection technology—and examples of their applications—are then introduced. Applying these technologies to in-vehicle sound control makes it possible to provide drivers of advanced safety vehicles with "superior ears" that can see the blind spots not shown in mirrors, radar, and cameras. These PSZ technologies act as an "ear" that not only enhances driver and passenger comfort but also greatly enhances safety and reliability while providing a sophisticated connection among the driver, vehicle, vehicle exterior environment, and automotive society.

Keywords: active noise control (ANC), sound event localization and detection (SELD), advanced safety vehicle (ASV)

1. Two technologies necessary for controlling a sound space

Thanks to technological advances, the sound spaces surrounding us are evolving to become more convenient and comfortable. For a sound space, it is necessary to be able to block out the sounds one does not want to hear and transmit only the sounds one wants to hear, and manufacturers are introducing a variety of wearable products, such as earphones, to meet this need. However, these products have not been able to solve the following two major problems.

The first problem is the heavy strain on the ears when earphones are worn for long periods. Technically, blocking the ear is the most-effective (and cheapest) way to block out unwanted sounds, and most earphones with such a function are "in-ear" earphones, which—as the name suggests—are inserted into the ear canal. Long-term use of in-ear earphones can increase stress on the user due to pressure and increase the risk of ear-canal problems [1], and such issues raise concerns about the health of the ears, which are fundamental to human social activities. It can therefore be said that it is necessary to devise technology for controlling the sound space in a manner that blocks sounds the one does not want to hear without blocking one's ears.

The second problem is that in addition to the sounds one *wants* to hear, the sounds one *must* hear are not adequately considered. People can react and respond

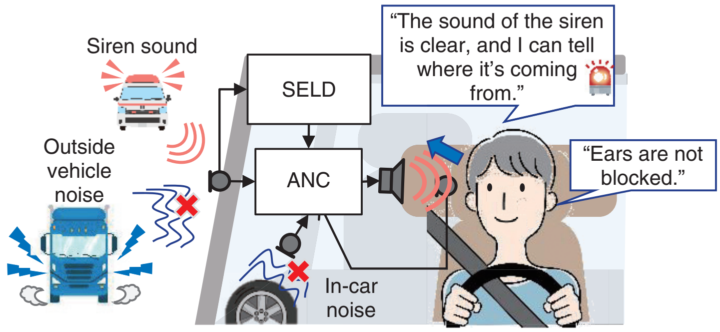


Fig. 1. Example applications of PSZ technology in vehicles.

to things that are happening in places invisible to the eye by hearing sounds from those places. For example, a person can avoid a bicycle approaching from behind by listening for the sound of the bicycle's ringing bell. These devices block out noise that does not need to be heard and must-hear sounds, and the person cannot always hear the sounds necessary for detecting danger. In other words, the ear loses its essential safety function of hearing sounds that must be heard. It can thus be said that the control of the sound space is required to ensure not only one hears the sounds one wants to hear but also one hears the sounds one must hear.

2. Two technologies for controlling the sound space in a vehicle: active noise control and sound event localization and detection

For vehicles, the above-described technical requirements are more pronounced than in other cases. Wearing a wearable device that blocks the ears is not only a possible violation of the Japan's Road Traffic Act but also a violation of certain other regulations. To create a comfortable sound space around a person traveling in a vehicle while ensuring their safety, it is thus necessary to be able to block the sounds one does not want to hear without blocking one's ears. The blocking of external sound with the vehicle's body and helmet, the increase in external noise including road noise while moving in a vehicle, and the fact that a vehicle is moving faster than a human make it difficult to hear sounds necessary to avoid danger, and that difficulty can be a cause of accidents. Consequently, the necessity to be able to hear the sounds that need to be heard is even greater, especially for advanced-safety vehicles (ASVs), which are essential for enhancing safety.

To meet the above-mentioned technological demands, we have been researching and developing elemental technologies for a Personalized Sound Zone (PSZ) that are highly integrated with spot-sound-reproduction technology and acoustic extended-reality (XR) technology. In this article, active noise control (ANC) technology, which suppresses noise without blocking the ears, and sound event localization and detection (SELD) technology are introduced (**Fig. 1**). SELD technology makes it easier to hear must-hear sounds to avoid danger even in environments where it is difficult to hear surrounding sounds.

3. ANC technology suppresses noise without blocking the ears

ANC technology for blocking sounds one does not want to hear without blocking the ears is explained. As mentioned above, popular earphones generally cover the ears to block sounds outside the ear. How commonly used in-ear earphones block sounds is shown in **Fig. 2(a)**. In-ear earphones are worn by inserting them into the ear canals. They act like earplugs by blocking the ear canals in a manner that makes it difficult to hear sounds outside the ear.

Sounds outside the ear, however, cannot be completely muted. Accordingly, as shown in **Fig. 2(b)**, two microphones (a reference microphone and error microphone) are fitted inside the earphone. On the basis of the sounds detected from moment to moment by these microphones, ANC reproduces sounds from the cancellation loudspeaker that eliminates the noise entering the ear. Therefore, sound-insulation performance improves. The reference microphone detects the ambient noise that is to be blocked, and the error microphone detects any sound missed by ANC

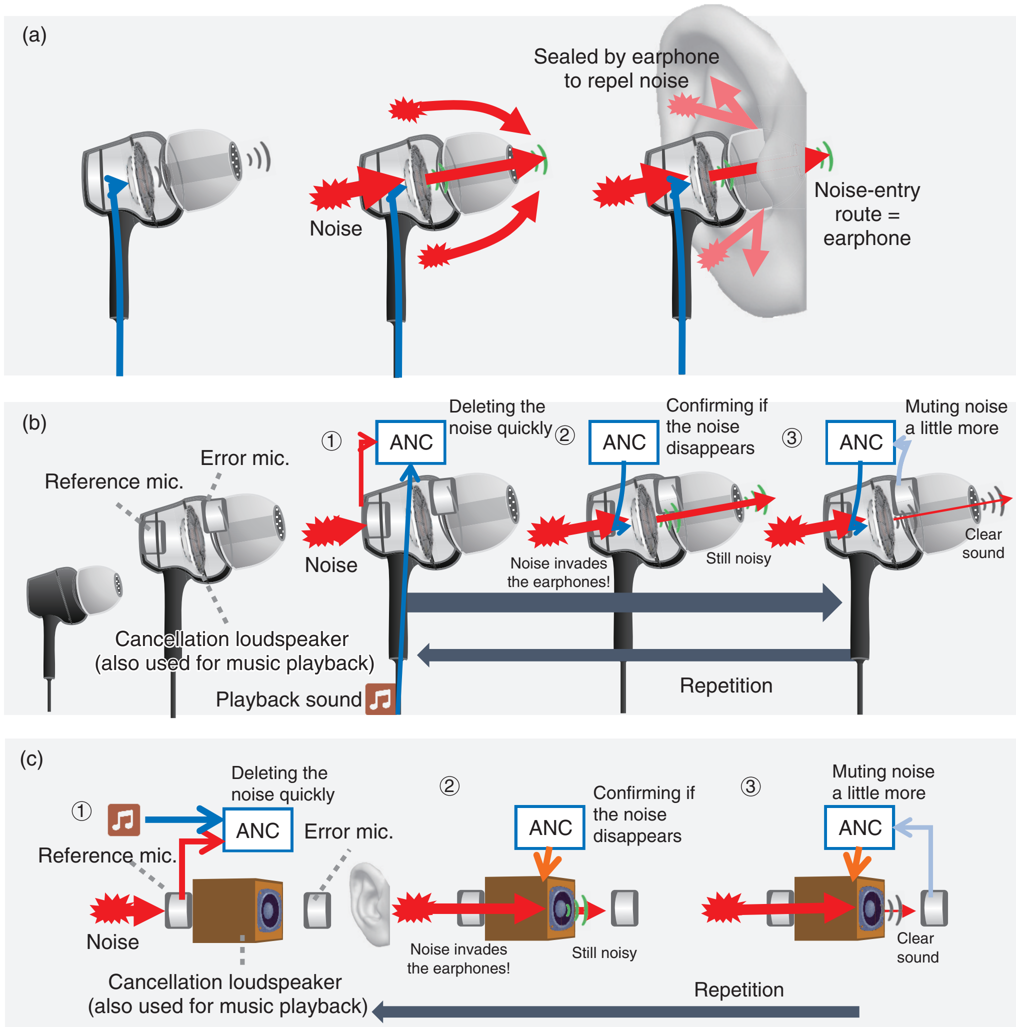


Fig. 2. (a) How in-ear earphones reduce noise. (b) How ANC works in in-ear earphones. (c) ANC that does not block the ear.

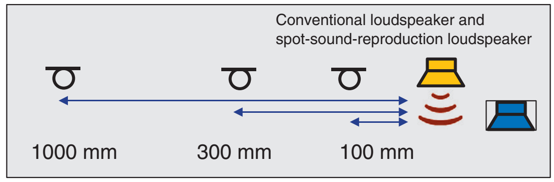
processing in the ear.

How to mute noise without covering the ears in the manner described above is considered. To avoid blocking the ear, the earphone loudspeaker must be positioned apart from the ear. However, a small loudspeaker, such as the one in an earphone, does not produce sufficient sound output, so a larger loudspeaker is required. The above-mentioned reference and error microphones should also be placed away from the ear because placing them near the ear, in the manner of in-ear earphones, would block the ear. Such a system is illustrated in **Fig. 2(c)**.

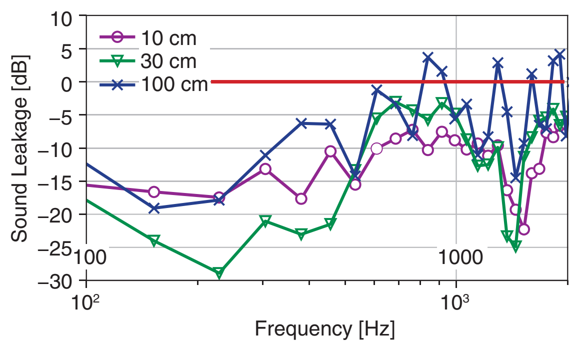
In consideration of the above circumstances, to suppress ambient noise without blocking the ear, it is necessary to cancel noise waves at the ear by using a loudspeaker and microphone placed apart from the ear. This configuration, however, faces the following three major problems:

(1) Control stability. The sound emitted by the loudspeaker (apart from the ear) to cancel out the noise is recorded by the reference microphone (which should record ambient noise only); as a result, so-called “howling” (also called “feedback”) occurs.

(2) As shown in Fig. 2(b), in-ear earphone ANC can correctly detect noise heard in the ear because the microphones are located near the ear entrance. On the contrary, as shown in Fig. 2(c), when the microphones are apart from the ear, they cannot correctly detect the noise



(a) System for measuring sound leakage from loudspeakers



(b) Sound-leakage comparison between conventional and spot-sound-reproduction loudspeakers

Fig. 3. Effectiveness of sound-leakage-suppression loudspeaker.

heard in the ear, and ANC outputs incorrect sound.

(3) Commercially available loudspeakers and computers take time to record and play back sound; thus, while the cancellation sound is being generated, the noise reaches the ears before being canceled. It is also undesirable to use large amounts of electricity in a vehicle's interior, so it is necessary to save power. Digital signal processors (DSPs) have conventionally been used for such applications. For ANC that does not block the ears, however, the problem of exceeding the computational power of the DSP must be solved. To solve problem (2), it is generally necessary to configure a large number of loudspeakers and microphones and implement signal processing for various compensations. This configuration requires a large amount of computation, and the cancellation sound cannot be generated in time.

To solve problem (1), it is necessary to reduce the sound leakage from the loudspeaker to the reference microphone. To meet that need, we developed a loudspeaker that applies the principle of the above-mentioned spot-sound-reproduction technology. The effect of reducing sound leakage is graphically shown in **Fig. 3**. This loudspeaker not only reduces the sound leakage to the entire surrounding of the loudspeaker but also creates an area where the sound leakage is very small, especially in the plane parallel to the diaphragm of the loudspeaker. The red line indicates the sound leakage of a conventional loudspeaker, and the other colored lines represent the amount of sound leakage from the loudspeaker using spot-sound-reproduction technology normalized by the sound pressure in front of the loudspeaker. Compared with the conventional loudspeaker, the spot-sound-reproduction loudspeaker suppresses sound leakage by several decibels to 30 decibels at 100 to 300 mm from the loudspeaker.

To solve problem (2), it is necessary to move the reference and error microphones closer to the ear. The spot-sound-reproduction loudspeaker reduces sound leakage, so by embedding it in the headrest of a vehicle, the reference microphone can be moved closer to the ear. The reference microphone is then able to detect sounds similar to noise heard with the ear.

Unlike the reference microphone, the error microphone does not suffer from the howling problem, so

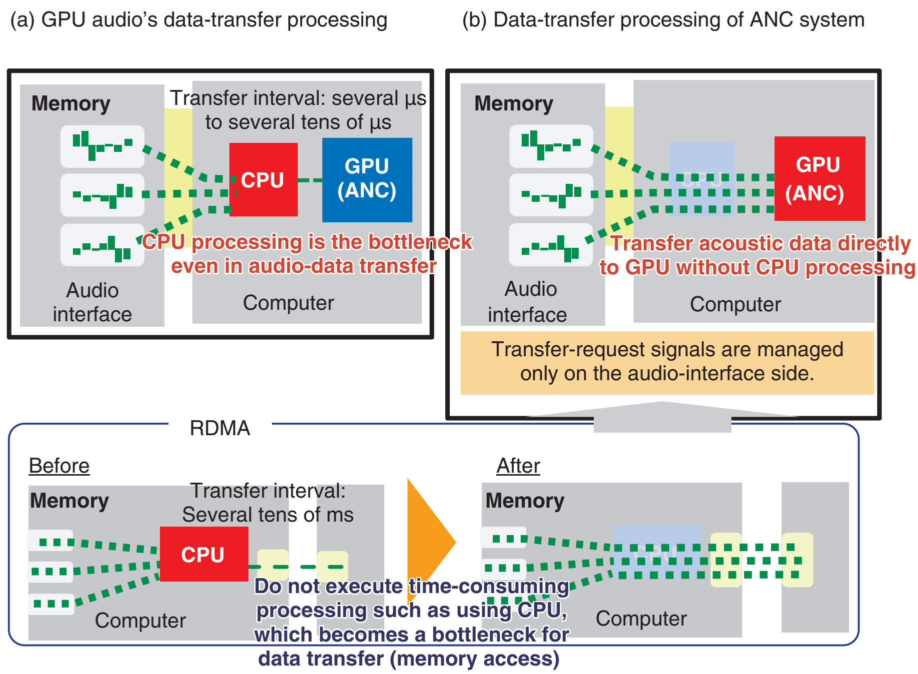


Fig. 4. GPU direct audio method for transferring data at high speed and low latency in short intervals with low power consumption.

it can be placed closer to the ear than the reference microphone. However, it is impossible to place the error microphone near the ear without blocking the ear. Therefore, to estimate the noise at the ear with the microphone slightly away from the ear, signal processing is required. This processing typically requires the placement of multiple error microphones, and it must be completed within a very short period (a few hundred microseconds) after the noise is detected by the reference microphone and before it reaches the ear.

For the above reasons, power-saving hardware that executes signal processing at high speed and with ultra-low latency is essential for ANC that does not block the ears. We solved this problem by applying general-purpose computing on graphics processing units (GPGPUs), which are widely used in the network and video-processing fields, to acoustic processing and optimizing it.

The principle of using GPGPUs for ANC is illustrated in **Fig. 4**. GPGPUs are known for their high speed, low latency, and low power consumption per unit of processing. However, as shown in Fig. 4(a), data cannot be directly input to and output from audio devices, and a central processing unit (CPU) must act as an intermediary between them, increasing processing delay. Therefore, by using remote direct memory access (RDMA) technology, as shown in Fig. 4(b), the audio signals of the microphone and loudspeaker are connected directly to the GPGPU without going through a CPU. The method was optimized to achieve processing of large amounts of small data-transmission packets (frame bursts) not handled in other signals and a distinct real-time nature unique to acoustic processing, which cannot tolerate even a 1-µs delay.

The result of the above-described RDMA application is hardware that can transfer acoustic data in about 1/50th the time required by conventional hardware and process large amounts of acoustic signals in real time and with low power consumption. This hardware not only enables ANC that does not block the ears, which has been difficult to achieve with DSP, but also enables the introduction of deep-learning technology, which is said to be difficult to implement in ANC due to its large computational load and significant processing delays.

An example of a test vehicle equipped with these technologies is shown in **Fig. 5**. In the test vehicle, loudspeakers with the spot-sound reproduction function are installed on both sides of the headrests of all seats, in positions that do not obstruct the driver's/passenger's line of sight, and reference and error

Fig. 5. Example of implementing an ANC system that does not block the ears.

microphones are placed around the loudspeakers. The shape of the headrest is designed to prevent the spot-sound-reproduction capability from deteriorating due to the spot-sound-reproduction loudspeakers being mounted inside the seat's headrests, thereby enhancing the accuracy of noise suppression near the ear. This configuration improves the comfort of the people inside the vehicle interior.

4. SELD technology enables one to hear only the sounds one must hear

As mentioned above, ANC technology can suppress unpleasant sounds when the vehicle is running; however, all sounds detected with the error microphones are subjected to suppression, which can lead to accidents by making it difficult to hear the sounds that must be heard to avoid danger. To make it possible to hear the sounds one needs to hear, it is therefore necessary to develop new sound-transmission technology. To meet this need, we focused on SELD technology.

SELD is overviewed in **Fig. 6**. SELD technology estimates when, where, and what happened from sound signals observed with microphones. By using this technology, it is possible to detect the sound (including its direction of arrival (DOA)) that the driver truly needs from the various sounds input into the microphones. SELD technology is now generally based on end-to-end deep-learning technology, which internally estimates the DOA corresponding to "where" and sound event detection (SED) corresponding to "what."

Since SELD technology uses deep learning, it requires a large amount of data related to necessary sounds that must be heard for its training. For example, an application of SELD technology—in which a siren is sounding in the driver's blind spot—is shown in Fig. 6. In this situation, the siren is quickly detected, and the driver is notified of the direction of the siren so that they can take the appropriate evasive action. Ideally, sirens from all directions and in all blind spots should be recorded and used for training the deep-learning model while considering all locations and situations (surrounding vehicles, buildings, weather conditions, etc.). However, it is difficult to record such a large volume of sound data comprehensively.

People can infer, to some extent, the direction of sound arrival, even in the presence of environmental differences and changes in sound due to self-motion [4–6]. In other words, a person can select the information necessary to estimate the direction of sound arrival from the information contained in the sound. Given this fact, we considered enabling SELD technology to imitate this human ability [7–9].

We developed echo-aware feature-refinement (EAR) - comprehensive anechoic data and sparse multi-environment data (CASM) technology [2] and motion-aware feature-refinement (MAR) technology [3] to enable SELD technology to imitate such human abilities. These technologies enable SELD technology

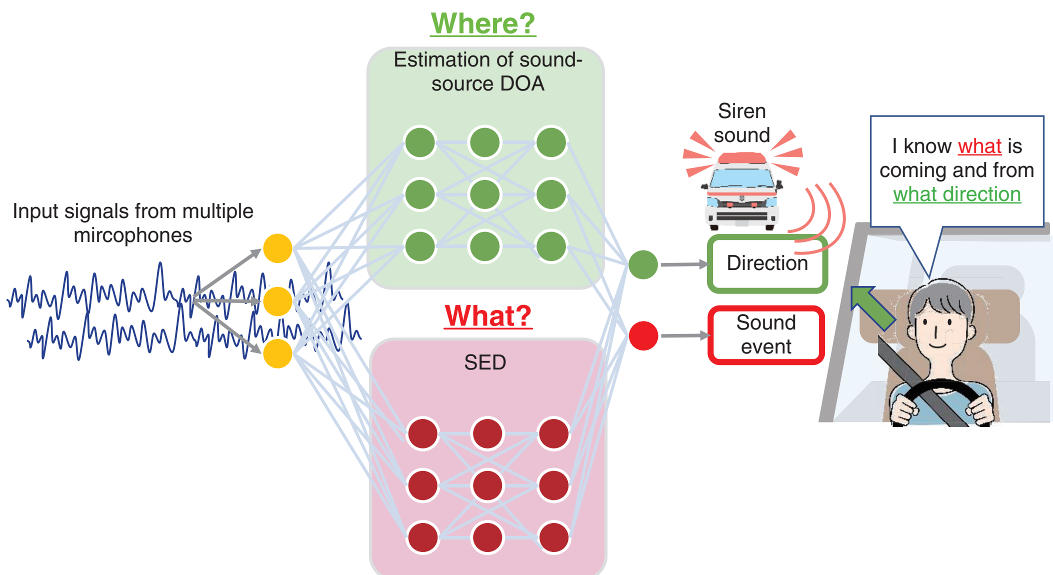


Fig. 6. Overview of SELD technology.

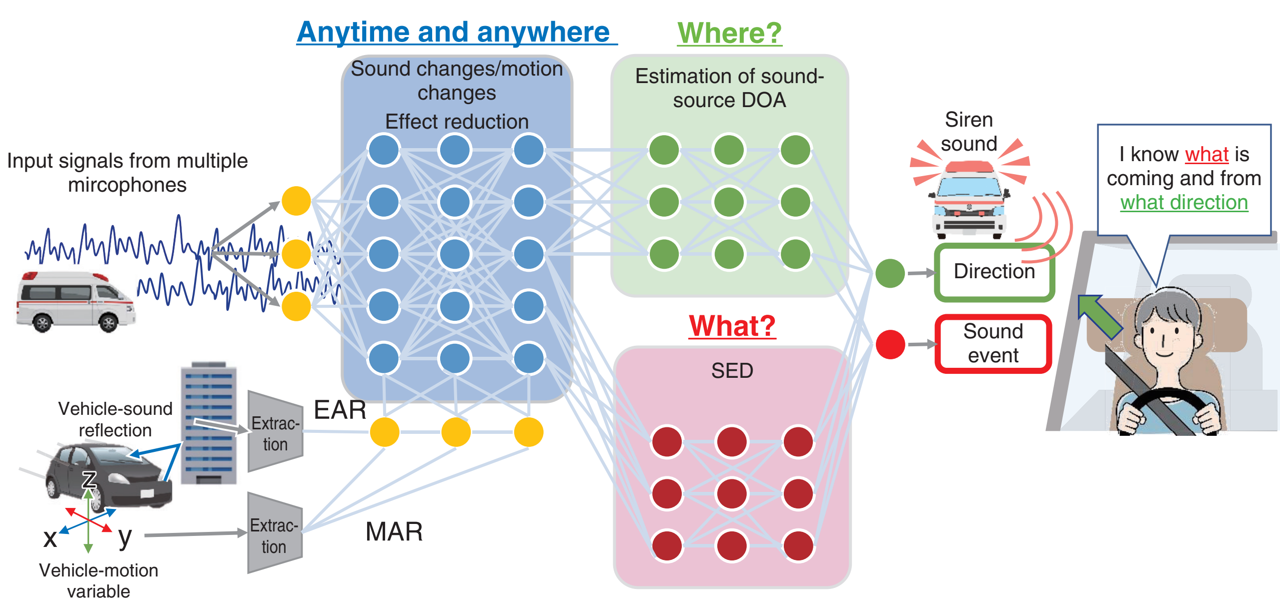


Fig. 7. Overview of automotive SELD technology applying EAR-CASM and MAR.

to operate robustly even when the environment changes or the user moves. It has thus become possible to apply SELD technology to mobile environments, such as cars, that were previously unrealistic at reasonable cost.

An overview of SELD technology for automotive applications applying EAR-CASM and MAR technologies is given in **Fig. 7**. The EAR-CASM technology provides the neural network with echoes of sounds emitted from the SELD-equipped vehicle (e.g., the car's running sound and sonar sound from its sensors) that are reflected to its surrounding as a cue for learning sounds not included in SELD technology's training data. The "improved" SELD technology works like the human ear, i.e., it suppresses the effects of unknown environmental sounds from these echoes by using the sounds it has learned up to that time. Moreover, MAR reduces the effect of changes in sound related to the vehicle's own motion by providing the neural network with various inputs from sensors (such as acceleration sensors fitted in the vehicle) and information cameras as information about the vehicle's motion.

Regarding EAR-CASM technology, exhaustive recordings of sounds from various environments are no longer necessary, and SELD technology can be implemented at a realistic cost. By applying EAR-CASM and MAR technologies in conjunction with ANC technology to ASVs, we have been creating a sound space for safe and comfortable transportation in which sounds one does not want to hear can be blocked—without blocking one's ears, and sounds one wants to hear—as well as sounds one must hear—can be heard.

References

[1] C. Mukhopadhyay, S. Basak, S. Gupta, K. Chawla, and I. Bairy, "A Comparative Analysis of Bacterial Growth with Earphone Use," OJHAS, Vol. 7, No. 2, April 2008.

[2] 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

[3] M. Yasuda, S. Saito, A. Nakayama, and N. Harada, "6DoF SELD: Sound Event Localization and Detection Using Microphones and Motion Tracking Sensors on Self-motioning Human," Proc. of the 49th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2024), Seoul, Korea, Apr. 2024.

[4] R. Gao, C. Chen, Z. Al-Halah, C. Schissler, and K. Grauman, "VisualEchoes: Spatial Image Representation Learning through Echolocation," Proc. of the 16th European Conference on Computer Vision (ECCV 2020), Aug. 2020. https://doi.org/10.1007/978-3-030-58545-7_38

[5] F. Antonacci, J. Filos, M. R. P. Thomas, E. A. P. Habets, A. Sarti, P. A. Naylor, and S. Tubaro, "Inference of Room Geometry from Acoustic Impulse Responses," IEEE Trans. on Audio, Speech, and Lang. Process., Vol. 20, No. 10, pp. 2683–2695, 2012. https://doi.org/10.1109/TASL.2012.2210877

[6] L. D. Rosenblum, M. S. Gordon, and L. Jarquin, "Echolocating Distance by Moving and Stationary Listeners," Ecological Psychology, Vol. 12, No. 3, pp. 181–206, 2000. https://doi.org/10.1207/S15326969ECO1203_1

[7] D. R. Begault, E. M. Wenzel, and M. R. Anderson, "Direct Comparison of the Impact of Head Tracking, Reverberation, and Individualized Head-related Transfer Functions on the Spatial Perception of a Virtual Speech Source," J. Audio Eng. Soc., Vol. 49, No. 10, pp. 904–916, 2001.

[8] Y. Iwaya, Y. Suzuki, and D. Kimura, "Effects of Head Movement on Front-back Error in Sound Localization," Acoust. Sci. & Tech., Vol. 24, No. 5, pp. 322–324, 2003. https://doi.org/10.1250/ast.24.322

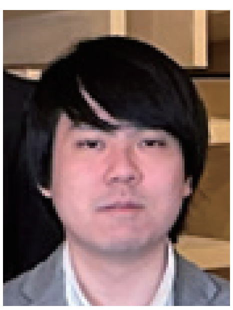
[9] B. C. J. Moore, "An Introduction to the Psychology of Hearing (3rd ed.)," Academic Press, 1989.

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 DOCOMO in 2015, where he has been working on speech-signal processing technologies for a speech-recognition system. He is a member of the Acoustical Society of Japan (ASJ), the Institute of Electrical and Electronics Engineers (IEEE), and the Audio Engineering Society (AES).

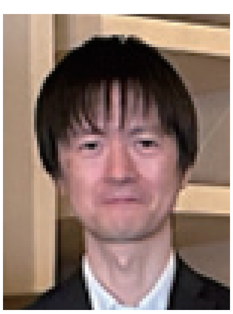
Tomoko Kawase
Research Engineer, Ultra-Reality Computing Group, NTT Computer and Data Science Laboratories.

She received a B.E. and M.E. in system design engineering from Keio University, Kanagawa, in 2011 and 2013 and Ph.D. in information engineering from Tsukuba University in 2018. Since joining NTT in 2013, her research interests have included microphone array signal processing and speech enhancement. She is a senior member of IEEE and a member of ASJ.



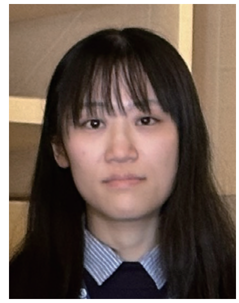
Masahiro Yasuda
Research Engineer, Ultra-Reality Computing Group, NTT Computer and Data Science Laboratories.

He received a B.E. and M.E. in physics from Tokyo Institute of Technology in 2017 and 2019. He joined NTT in 2019, and his current research interests include acoustic-signal processing and machine learning. He is a member of IEEE and ASJ.



Shoichiro Saito
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 the University of Tokyo in 2005 and 2007. Since joining NTT in 2007, he has been involved with the research and development of acoustic signal processing systems, including acoustic echo cancellers, hands-free telecommunication, and ADS. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) and ASJ.



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

She received a B.E. and M.E. in mathematical engineering and information physics from the University of Tokyo in 2020 and 2022 and joined NTT in 2022. Her current research interests include acoustic-signal processing and mathematical engineering. She is a member of ASJ.

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 acoustic-signal processing and sound-field control. He is a member of ASJ, IEICE, the Institute of Image Information and Television Engineers, 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, computer-supported 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.

NTT sonority's Pursuit of Innovation—New Businesses That Leverage PSZ and MAGIC FOCUS VOICE Technologies

Kaori Sasaki

Abstract

NTT sonority aims to create new lifestyle standards by leveraging two core sound technologies: Personalized Sound Zone (PSZ) technology, which confines sound to the ear without blocking the ears, and NTT's patented MAGIC FOCUS VOICE technology that cuts out ambient sounds to deliver only the user's voice. This article describes the foundation of our company, introduces our brand "nwm (pronounced noom)," showcases our flagship on-ear-speaker lineup, explores the dynamic landscape of the open-ear earphone market, and unveils our endeavors to develop the next-generation voice digital transformation service business, which began in 2024.

Keywords: PSZ technology, nwm, on-ear speaker

1. PSZ technology—from concept to a new business

NTT sonority was established in September 2021 to expand the application of Personalized Sound Zone (PSZ) technology, which confines sound to the user's ear without blocking the ear. PSZ technology originated from innovative thinking at NTT Computer and Data Science Laboratories and leverages inverted sound waves to minimize noise spillage outside the ear. PSZ technology uses noise-cancellation principles to create an inverted sound wave (reverse phase) layered over a sound wave to neutralize the sound. Driven to meet the needs of automotive and aircraft manufacturers, ongoing research has been dedicated to building individual acoustic environments where sound is heard only by the individual, and seamless communication with the individual's surroundings is made possible without the need for devices such as earphones or headphones.

The demonstration of aircraft seats equipped with PSZ technology at the NTT R&D Forum in November 2020 became the catalyst for the commercialization of this technology. The business plan initially targeted corporate clients in the automotive and aviation industries. However, NTT Chairman Jun Sawada (then President) emphasized the importance of bringing PSZ technology to the consumer sooner. This push led to the establishment of a consumer-focused venture, and the company was founded within less than a year of the demonstration. To establish the company as a full-service audio manufacturer from planning and development to manufacturing, quality assurance, distribution, and sales, we actively recruited talent from outside the NTT Group. In addition to experts in the fields of acoustic and mechanical engineering, we also hired professionals with backgrounds in advertising, media, and retail industries. The new hires constitute 80% of the company's workforce (**Fig. 1**).

2. Applying PSZ technology to earphones

In 2021, as the pandemic continued, new lifestyles,

Fig. 1. Embracing talent from diverse fields.

such as remote work, became increasingly prevalent. With the rise of digital content-mediated communication, headphones emerged as essential accessories. However, new challenges also emerged, such as discomfort from prolonged headphone use and the inability to hear ambient sounds when ears are blocked. Recognizing these needs, we began developing headphones equipped with PSZ technology.

Our first challenge was to minimize sound leakage while maintaining sound pressure in the ear and preserving audio quality. Countering inverted sound waves from the rear posed a significant challenge in product development as it was an unprecedented concept. We conducted countless simulations to determine the balance of various elements affecting sound quality, such as the placement size of the ports for emitting the sound waves, as well as the cavity space and volume within the enclosure. We also iterated hundreds of hardware improvements to ensure a design that looks great regardless of race or gender. The resulting product is a testament to the craftsmanship of NTT sonority engineers, who combined NTT's research technology with advanced expertise in acoustic design (**Fig. 2**).

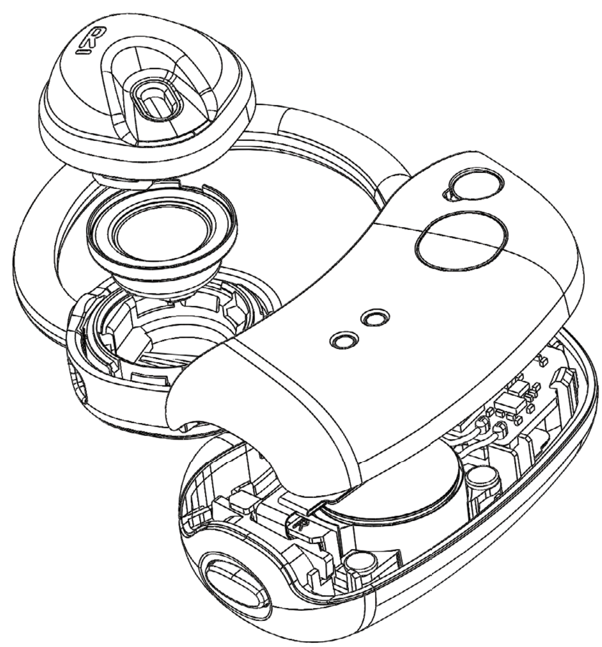


Fig. 2. A product refined through continuous trial and error.

2.1 Audio brand "nwm"

To promote the practicality of PSZ technology to consumers, we have been concurrently considering the development of our in-house audio brand alongside product development. As mentioned above, amidst the COVID-19 pandemic, we recognized emerging communication challenges from the prevalence of

Fig. 3. Brand tagline & statement.

digital content-mediated communication. Considering that PSZ technology is aimed at constructing individual acoustic spaces for uninterrupted communication, we conceptualized a brand that seamlessly connects an individual's world with their surroundings, embodying the concept of "coexistence."

In November 2022, we announced the launch of our audio brand "nwm (pronounced noom)," which stands for the "New Wave Maker" and is NTT Group's first audio brand. With a brand tagline "Seamless sound for a seamless world," we aim to connect people to their world and co-create new ways of working and enjoying life by listening to the consumer's real voices (**Fig. 3**).

The statement is as follows.

"Mom is in a remote meeting, Dad is on the phone, and the kids are watching videos. What if the whole family could be connected seamlessly, without being interrupted from their activities? With nwm, we want to work together with you to create the future of living and working and realize this with audio technology.

What we aim for is not immersion, but to connect and bring you closer to your world. Let's realize the future of sound, together with nwm."

2.2 new nwm on-ear speakers

As of January 2024, the lineup consists of three products: the wireless model "nwm MBE001," wired model "nwm MWE001," and wireless neckband model "nwm MBN001" with two color variations for two of the models. Although they belong to open-ear earphones, they stand out from competitors by minimizing sound leakage with PSZ technology, addressing the common issue with open-ear earphones (**Fig. 4**).

Taking into account the nwm concept, the design also visually communicates the idea that "unblocked ears = liberates communication," enabling others to know that the wearer can engage in communication. The on-ear speakers are worn by hooking them along the contours of the ears.

Thanks to PSZ technology, which minimizes sound leakage, users can enjoy their audio content comfortably whether in an office setting or quiet public spaces. By incorporating a 12-mm-diameter driver, the on-ear speakers provide a balanced and natural listening experience as if listening to background music. Equipped with a built-in microphone for calls, the user can speak and hear their voice clearly with the ears fully open. The lightweight, and comfortable fit with no added pressure on the ears allows for a stress-free listening experience even during long hours of remote work or online meetings. Whether enjoying music or watching videos, the user can hear the sounds of doorbells ringing as well as family members and pets in the next room and enjoy natural conversation while wearing the on-ear speakers. Our products are designed to facilitate seamless coexistence in all aspects of daily life.

The wireless-earphone market is on an expanding trajectory, with forecasts predicting a market size of

Fig. 4. nwm on-ear speaker lineup.

approximately 4 trillion yen globally by 2025. A survey conducted by NTT sonority in June 2023, targeting 2165 men and women nationwide, regarding prolonged usage of earphones and headphones, revealed that over 30% of respondents reported an increase in earphone usage. This underscores the transition of earphones into essential lifestyle products. Further internal research indicates that approximately 10% of users in the wireless-earphone market experience discomfort with earphones that block the ears, suggesting a latent demand for open-ear earphones.

2.3 nwm vision

We aim to provide a novel auditory experience through our on-ear speakers, and are engaging in exciting collaborations in entertainment, art, and culture. In September 2023, the nwm MWE001 (wired model) was used as part of an advanced physical performance using cutting-edge technology by a music/DJ artist with amyotrophic lateral sclerosis. The MWE001 was used to detect the DJ's brainwaves stimulated by the sounds of the music and ambient sounds of the venue, which was essential in the futuristic performance, enabling the DJ to be one with the audience.

In entertainment, the nwm audio technology was also used in events leveraging high-speed, high-capacity communication via extremely high frequency waves to offer a participatory augmented reality gaming experience. This enabled participants to enjoy both the sounds of the game content while strategizing with team members, significantly enhancing the sense of unity among players. In December 2023, "Hanakurabe Senbonzakura," a production that fused traditional Kabuki with cutting-edge technology, was held. This integration enabled audiences to enjoy the immersive sounds of the stage while listening to the simultaneous guided commentary to enhance the overall experience of the performance.

On the basis of current use cases, we are confident that the on-ear speakers will evolve into indispensable devices for new forms of communication, merging real and virtual spaces.

We are strengthening partnerships with various industries to meet the needs of individual acoustic spaces, especially with automobile and aircraft manufacturers, which are the foundation of PSZ-technology development, to drive business development.

Products equipped with MAGIC FOCUS VOICE technology are also being marketed under the brand. In October 2023, we launched the Beam Mic Speaker "LinkShell," a high-quality speaker specialized for meetings. By cutting out ambient noise and delivering only the necessary voices to the participants, it provides stress-free communication in an online environment, whether for remote work in a living room, in a noisy office, or in an online meeting during a workcation (**Fig. 5**).

Fig. 5. MAGIC FOCUS VOICE integrated Beam Mic Speaker “LinkShell.”

Fig. 6. BONX intro knot 3.5M.

3. Next steps: industry solution business

As the next step, in 2024, we launched a next-generation voice digital transformation (DX) service for businesses leveraging PSZ technology and MAGIC FOCUS VOICE, marking our entry into the industry solution market. As NTT sonority’s first service business, we are collaborating with a domestic startup, BONX Inc., to provide a comprehensive solution for on-site communication called “BONX WORK,” targeting industries with “deskless workers” who are engaged in activities other than desk work.

In Japan, the number of deskless workers is estimated to exceed 40 million in industries, such as retail, hospitality, caregiving, and construction, that rely mainly on vocal communication. While transceivers are essential tools, they face challenges such as ear fatigue from prolonged headphone use, the burden of carrying multiple devices, and communication breakdown in noisy environments. We believe that PSZ technology and MAGIC FOCUS VOICE are the optimal solutions for addressing these needs.

BONX was born from the founder’s passion to enhance the sense of togetherness and fun among friends who snowboard together, and their vision of valuing communication in every setting aligns well with NTT sonority’s goals to advance human communication.

In April 2024, we launched NTT’s OnSite DX: Zero to Transformation, a next-generation transceiver service that combines NTT sonority’s audio technology, including on-ear speakers, with the BONX WORK app. The BONX WORK app incorporates voice-activity-detection technology that instantly distinguishes between human voices and other sounds, enabling clear communication with simplicity and flexibility, even among a small number of users. Transcription, text, and photo-sharing features are also available on the app. Together with BONX, we also developed the BONX intro knot 3.5M, a single-ear model of wired ear speakers catering to specific use cases in various industries (**Fig. 6**).

The integration of apps, earphones, and networks alone is expected to bring about a dramatic improvement in communication within work environments. In the fall of 2024, we plan to release a push-to-talk device featuring MAGIC FOCUS VOICE, which enables clear communication even in high-noise environments (**Fig. 7**).

By incorporating these advancements into the next-generation transceiver service, we anticipate further expansion of its use cases (**Fig. 8**).

An entry campaign is also planned with NTT Group companies to facilitate service expansion into deskless-work industries. Building upon the advancements in artificial intelligence, we also aim to enhance existing features, such as voice-data analysis, and develop new functionalities such as simultaneous translation and guidance. We are also considering offering on-premises service options and providing industry-specific solutions that are also linked with third-party services.

Fig. 7. MAGIC FOCUS VOICE integrated push-to-talk device.

Fig. 8. The next-generation transceiver service in use.

4. Final thoughts

The nwm brand was developed to foster communication between people under the concept of coexistence. Our voice DX services deliver solutions that value the individual's identity by effectively conveying the subtlest nuances in the voice. Both businesses resonate with NTT's vision for IOWN (Innovative Optical and Wireless Network). As we mark the third year since NTT sonority's founding, we are committed to advancing the evolution of interpersonal communication through the innovative application of sound technology.

Kaori Sasaki
Press Staff, Marketing & Communications Group, NTT sonority, Inc.

Millimeter-wave-based Drone Automatic Landing-guidance System for Advanced Maritime Operations

Tatsuya Iizuka

Abstract

Drones have gained increased interests from a variety of fields such as logistics and environmental measurements. In the maritime domain, their applications are wide spreading, including cargo transportation, weather observation, fishery exploration, red-tide monitoring, port management, and detection of illegal ships. However, safely navigating drones to ships is challenging due to the swaying caused by waves and the occurrence of poor visibility conditions such as fog and rain. To address this challenge, our team has been investigating a novel automatic landing-guidance system using high-resolution millimeter-wave radar, which is tolerant to weather conditions. With this system, radio frequency identification (RFID)-based wide-range navigation and passive-landing port-based surface-inclination estimation are implemented to consistently guide drones located at a distance to the landing point. Specifically, the system consists of a wide-range flight-guidance method and a surface inclination-estimation method, both of which require a millimeter-wave radar mounted on a drone. The flight-guidance method, called MilliSign, uses corner reflector array-based RFID tags and enables a drone to be guided to a landing point by using the embedded information and position information of the tag. The inclination-estimation method uses a special arrangement pattern of corner reflectors and enables precise inclination of the landing port on a swaying ship for determining the appropriate start time of landing descent. This article investigates the overall system, which consistently guides a drone at a distant point to a landing point for all-weather operation.

Keywords: drone, millimeter-wave radar, maritime operation, autonomous landing

1. Introduction

Drones, with their advanced three-dimensional (3D) mobility capabilities, are gaining attention in various fields such as logistics, disaster response, and weather observation [1, 2]. They are expected to serve as an infrastructure for advanced logistics and environmental measurements by applying mobility, sensing, and information-processing technologies in unexplored areas such as the sky, marine areas, and mountainous regions that have been hard to reach.

Drones in the maritime domain provide a wide range of use cases, including, but not limited to, cargo transportation, weather observation, fishery exploration, red-tide monitoring, port management, and detection of illegal ships. They play a crucial role in understanding the maritime conditions of Japan, which is surrounded by the sea. To expand the area where drones can operate over a wide range of ocean, it is essential to use the ships as base stations for automatic takeoff and landing.

Due to the limited space, human labor, and equipment on a ship, advanced drone flight-guidance technology is necessary to achieve efficient navigation management. However, because the landing surface on the ship moves and sways due to waves, flight control systems based on not only absolute position obtained by the Global Positioning System (GPS) but also relative position are required. Navigation methods using image recognition have been proposed by

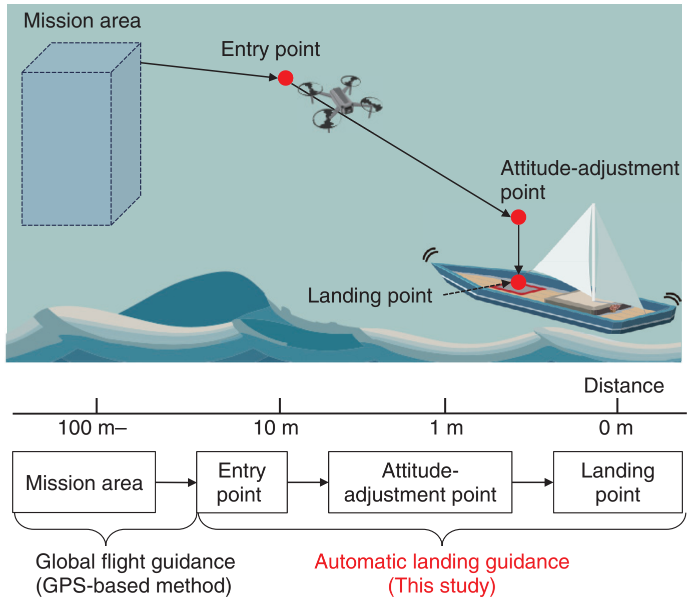


Fig. 1. Automatic landing-guidance system using millimeter-wave radar.

placing a QR (Quick Response) code as a marker on the landing platform to achieve accurate landing by recognizing the relative positional relationship with the ship [3]. However, the use of image recognition technology becomes difficult and the opportunity to use drones is limited in situations where visibility is poor, such as fog, rain, and nighttime.

To overcome this issue, our team has been investigating an automatic landing-guidance system that involves mounting a millimeter-wave radar on a drone that is less affected by weather and reading the radio frequency identification (RFID) tag installed on the ship. By using the RFID tag in the millimeter-wave band for the flight-guidance method, called MilliSign, that we previously developed [4] as a marker for the landing point, it is possible to transmit the embedded information and position to the drone from a position more than 10 m away and guide it accurately. We have also developed a method for estimating the inclination of the landing surface when the positional relationship between the drone and landing point is known using a landing port with multiple reflectors installed on the landing surface [5]. By combining MilliSign and the inclination-estimation method, it is possible to convey the appropriate flight path and landing timing to the rocking surface on which the position and inclination changes over time, enabling safe flight guidance to the ship. However, the working 3D ranges of the tag and inclination-estimation method are different, and to guide a drone located far away to the landing point consistently, it is necessary to appropriately arrange the tags and ports so that the flight path to guide is always within the range of either method.

An overview of this system is shown in **Fig. 1**. We aimed to guide a drone located about 10 m away to the landing point and divide it into three guidance points and two guidance routes. The point where the tag installed near the landing point is first detected with the millimeter-wave radar is set as the entry point. On the basis of the information obtained from the tag, we then guide the drone to the attitude-adjustment point and measure the inclination of the landing surface swaying with the waves there. On the basis of the measured inclination information, we start the landing descent at the appropriate time and finally reach the landing point.

In Section 2, we explain MilliSign using millimeter-wave RFID tags for guiding the drone from the entry point to the attitude-adjustment point and confirm its performance. In Section 3, we explain the method for estimating the inclination at the attitude-adjustment point related to the landing port and confirm its performance. We then discuss the overall automatic landing-guidance system in Section 4, and summarize this article in Section 5.

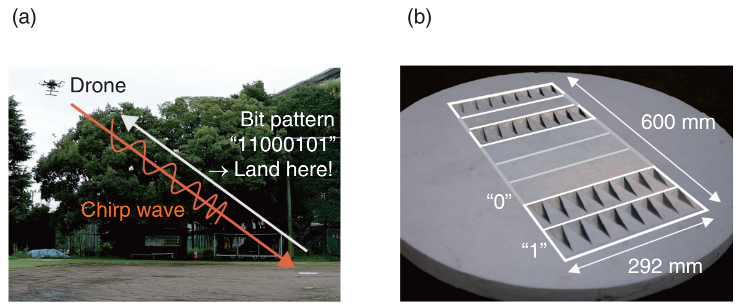


Fig. 2. MilliSign: (a) Readout of a bit sequence from the drone using slant-range radar. (b) Developed chipless RFID encoding "11000101."

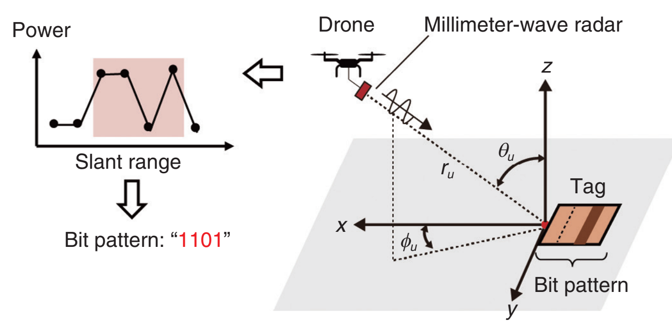


Fig. 3. Principle of bit-sequence reading by slant-range radar. The bits "1" or "0" can be decoded form the range-power analysis at the radar position.

2. Wide-range flight-guidance method, MilliSign, using RFID tags

In this section, we describe MilliSign [4] for guiding a drone from the entry point to the attitude-adjustment point, as shown in Fig. 1, and confirm its performance. An overview of MilliSign is shown in **Fig. 2**.

2.1 Readout of bits using slant-range radar

A slant-range radar is a radar that emits radio waves diagonally downward from a drone located in the air toward tags, measuring the scattering distribution on the ground. As shown in Fig. 2(b), it can read the arrangement order of reflectors that reflect millimeter waves as a bit string from the airborne radar and measure the position of the entire tag where reflectors are embedded. Therefore, it can read the position and information of tags installed on the ground. A spatially modulated chipless RFID tag [6] is then placed on the ground. The principle of bit string reading by slant-range radar is illustrated in **Fig. 3**.

We developed a tag-design method using corner reflectors (CR) to widen the reading range and a signal processing method [6] using the multiple signal classification (MUSIC) algorithm Root-MUSIC for eigenvalue analysis to accurately read the bit string stored in the tag.

2.2 CR-based chipless RFID tag for increasing 3D read range

We used a CR structure as an element constituting the developed spatially modulated chipless RFID tag. The CR is composed of three reflector plates connected perpendicularly, and due to the principle of retroreflective reflection, it has a high radar cross-section at a wide range of angles [7, 8]. The Van Atta array structure, which has 2D retroreflective reflection, has often been proposed as a component of chipless RFID tags [9, 10]. However, when considering reading from a reader with 3D positional freedom such as a drone, the CR, which has a wide angular characteristic in both azimuth angle ϕ_u and elevation angle θ_u, is suitable as a component of chipless RFID tags.

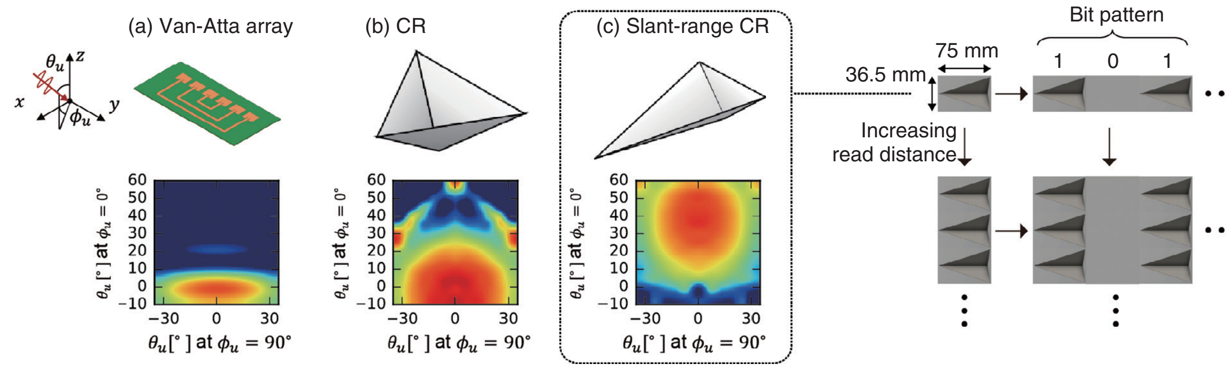


Fig. 4. Design method of developed CR-based chipless RFID tag.

In response to the problem that the maximum θ_u with retroreflective reflection in conventional CRs is limited to about 20°, as shown in **Fig. 4**(b), we have achieved an increase in θ_u (shallower angle) by changing the shape. By arranging this reshaped CR in two dimensions vertically and horizontally, the bit sequence to be stored and the maximum reading distance can also be determined. Figure 4 shows the design method of the developed CR-based chipless RFID tag.

2.3 Accurate readout of bits based on eigenvalue analysis

When CR tags are densely embedded in the distance direction, the frequency difference that constitutes the received signal changes as the path difference of the scatterers changes. Therefore, the bit-detection accuracy decreases with the fast Fourier transform (FFT), which fixes the frequency basis. Therefore, by using Root-MUSIC [11] for eigenvalue analysis, we can determine the frequency basis that constitutes the received signal and robustly detect bits even for frequency bases that change with the angle.

The procedure and principles of Root-MUSIC is as follows. After conducting eigen space analysis of the autocorrelation matrix R of the received signal, we first determine the frequency components that constitute the received signal. Once the size M of R and the maximum number of bases assumed (i.e., the maximum number of scatterers assumed) P are determined, a polynomial $Q(z)$ of order $(M - P)$ can be obtained by eigenvalue decomposition of R as follows:

$$Q(z) = \sum_{k=P+1}^{M} p^H (z^{-1})\, v_k v_k^H\, p(z) = 0.$$

Here, v_k is the k-th eigenvector of R, and $p(z)$ is the mode vector, represented as $p(z) = [1, z, z^2, \ldots, z^{M-1}]^t$. The roots z_m $(1 \le m \le M - P)$ are obtained by solving $Q(z) = 0$, where z_m is a complex number and represented as

$$z_m = e^{h_m \Delta t + j2\pi f_m \Delta t}.$$

Here, z_m is the rate of change of amplitude, and f_m is the frequency. The complex amplitude corresponding to each frequency component can be determined using least squares with $s(t)$.

The obtained frequency and complex amplitude can then be converted to distance as $r_m = f_m \cdot \frac{cT}{2B}$ and to signal power as $P_m = |c_m|^2$, respectively. When scatterers that strongly reflect in the radar direction are placed on the tag, roots corresponding to each scatterer's position can be determined and can be read as bit information. Since the accuracy of the obtained roots is less affected by slight changes in frequency components, it is possible to achieve bit detection that is resistant to changes in slant-range angles.

2.4 Experimental results

We evaluated the reading performance of the developed CR-type chipless RFID tag and the eigenvalue-analysis method using Root-MUSIC through experiments. In the coordinate system shown in Fig. 3, we present the measurement signal from the millimeter-wave radar when the drone's position is set at (r_u, θ_u, ϕ_u) = (12.6 m, 45°, 0°) in **Fig. 5**. This figure includes the results of both the eigenvalue-analysis method using Root-MUSIC and the conventional method using FFT. Upon examining the enlarged distribution of distance intensity around 13.0 m, where the tag is positioned (Fig. 5(b)), the eigenvalue-analysis method using Root-MUSIC accurately captured the bit

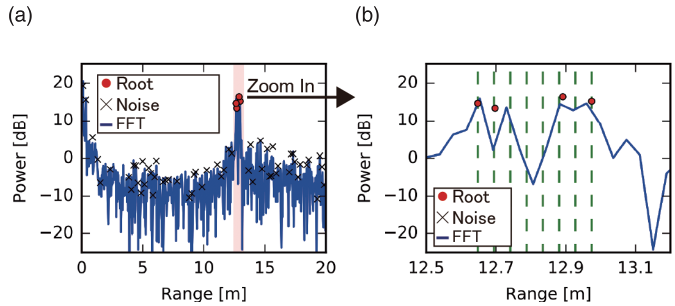


Fig. 5. Range-power profile of the received signal in front of the tag, in which eigenvalue-analysis- and FFT-based methods were used.

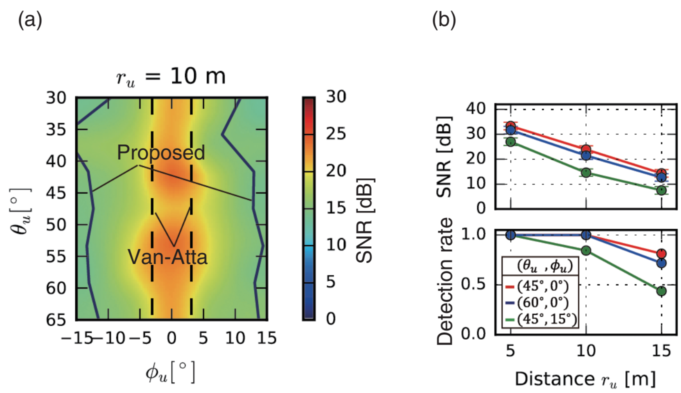


Fig. 6. (a) Distribution of SNRs and detection rates obtained from measurements when the ϕ_u and θ_u of the drone position were varied. (b) SNRs and detection rates obtained from measurements when the distance and angle of the drone position were varied.

sequence "11000101," demonstrating its ability to precisely estimate the tag's position and information. It is also evident that the tag's reflection from the backscattering from the ground can be read with a high signal-to-noise ratio (SNR) of over 15 dB.

We next evaluated the range within which the tag could be read in the coordinate system shown in Fig. 3. The reading performance was assessed using the detection-success rate and SNR when changing the distance r_u, ϕ_u, and θ_u. The detection-success rate was calculated by counting the successful reads among the 32 chirps used in a single-radar measurement and expressing the proportion of successful reads as the detection-success rate.

The variations in the SNR and reading rate at different positions are shown in **Fig. 6**. Even at a distance of 10 m from the tag, a reading-success rate of over 90% was achieved within a range of θ_u greater than 30° and ϕ_u greater than 20°.

3. Inclination estimation of landing surfaces on ships

In this section, we explain a prior study [5] on a method for estimating the inclination of the landing surface using the known arrangement of reflectors for the guidance method from the attitude-adjustment point to the landing point shown in Fig. 1, and confirm its performance.

This method uses a millimeter-wave radar chip, which is small and lightweight for mounting on drones. While a typical millimeter-wave radar chip achieves a high distance resolution of several centimeters by increasing the chirp bandwidth, the angular resolution is low, at about 10°, due to the limited number of array antenna elements. Therefore, this method is used for embedding multiple reflectors in the landing surface and estimating the inclination from the measured distance between the drone-mounted

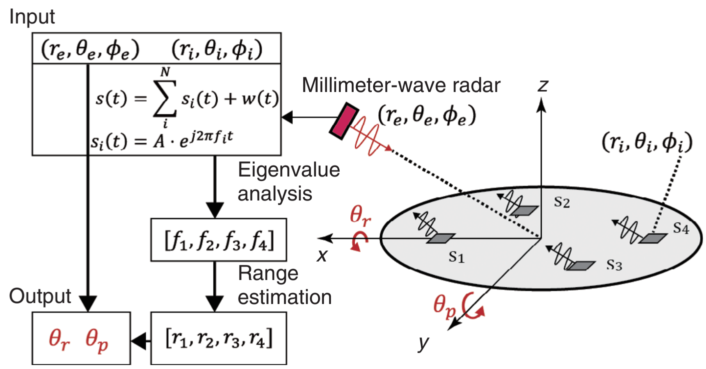


Fig. 7. Overview of inclination-estimation method using millimeter-wave radar.

radar and reflectors.

The overview of this method is shown in **Fig. 7**. In this study, to evaluate the inclination-estimation accuracy using distance-measurement information, one fixed radar and four reflectors were used, and the coordinates of the radar and each reflector were known and expressed.

We assumed the following situation: the drone is equipped with a millimeter-wave frequency-modulated continuous wave radar. There are four CRs s_i $\{1 \le i \le 4\}$ at the landing port, and the relative position of the reflectors to the center of the port are known as $\boldsymbol{P_i}$. The drone can obtain its position from the radar measurement as $\boldsymbol{P_e}$. If the rotation angle of the port around the x-axis as seen from the fixed coordinates on the ground is θ_r and the rotation angle around the y-axis is θ_p, the distance $D_i(\theta_r, \theta_p)$ from the drone to the i-th reflector can be calculated as

$$D_i(\theta_r, \theta_p) = |\boldsymbol{P_e} - \boldsymbol{R_x}(\theta_r) \cdot \boldsymbol{R_y}(\theta_p) \cdot \boldsymbol{P_i}| \ (1 \le i \le 4), \qquad (1)$$

where $\boldsymbol{R_x}$ and $\boldsymbol{R_y}$ are the rotation matrices around the x-axis and y-axis, respectively. By solving four simultaneous equations of Eq. (1), we can obtain rotation angles (θ_r, θ_p) and estimate the inclination of the surface.

In the experiments, the positions were expressed in a polar coordinate system and the coordinates of the drone were considered the same as the radar coordinates. The radar position $\boldsymbol{P_e}$ was set as (r_e, θ_e, ϕ_e) = (1 m, 45°, 0°). When we denote the position of i-th reflector as (r_i, θ_i, ϕ_i), then (r_i, θ_i) = (0.4 m, 90°) $\{1 \le i \le 4\}$ and (ϕ_1, ϕ_2, ϕ_3, ϕ_4) = (0°, 300°, 120°, 180°), which leads to the arrangement pattern shown in **Fig. 8**. We used a Texas Instruments' millimeter-wave radar [12] with a carrier frequency of 79 GHz and chirp bandwidth of 4.0 GHz, which provides a range resolution of 37.5 mm. We used the CRs mentioned in Section 2 as the embedded reflectors in the landing port. To obtain accurate distance measurements, the Root-MUSIC-based ranging method was used for processing the raw radar signal [11]. The direction of normal vector of the surface was used as the inclination-estimation-error metric. The measurements were conducted by varying roll and pitch angles as $-3 \le \theta_r, \theta_p \le 3$, respectively. The overview of experimental setup is shown in **Fig. 9(a)**.

Figure 9(b) shows measurement results of the range-power profile based on Root-MUSIC and range-FFT, which is a standard frequency analysis in radar-signal processing. The graph shows that the obtained roots plotted as red dots provided more accurate distance information than range-FFT. We then rotated the surface and observed the distance deviation caused by the inclination in accordance with Eq. (1). **Figure 10(a)** shows the distance deviation to each reflector D_i $\{1 \le i \le N\}$ when measured distances with the flat surface ($\theta_r = \theta_p = 0°$) were set as standards. The results indicate that radar-ranging measurements could capture the small distance deviation of several centimeters, considering that global navigation satellite system sensors have a root-mean-square error of 12.25 cm for vertical-position estimation. **Figure 10(b)** shows the inclination-estimation error for each combination of (θ_r, θ_p). Since the estimation error is less than 3° and relatively small when $\theta_r = 0°$, the results indicate that a larger θ_r deteriorates inclination-estimation accuracy. The reason may be that the close placement of s_2 and s_3 reduced the inclination-estimation accuracy of θ_r since distance-estimation errors of D_2 and D_3 increase the estimation error of θ_r. Therefore, optimizing reflector allocation

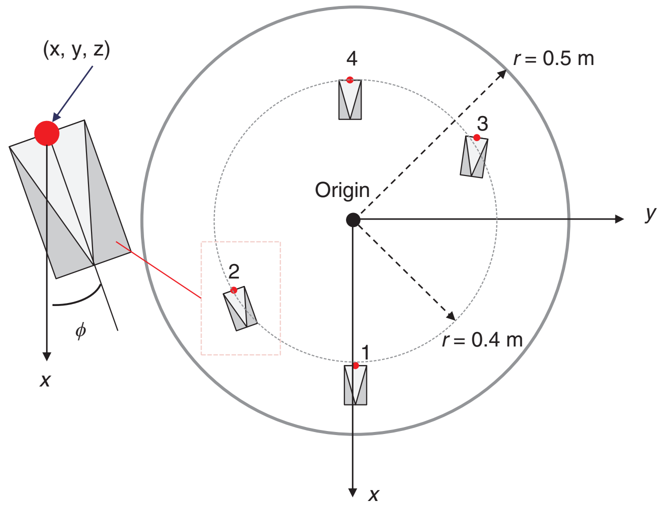


Fig. 8. Reflector-alignment pattern.

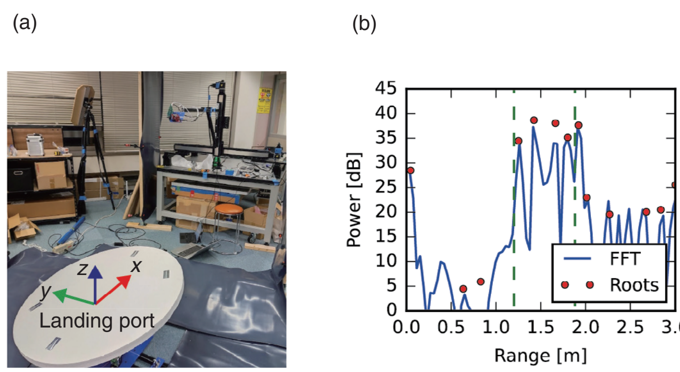


Fig. 9. (a) Experimental setup. (b) Range-power profile obtained from millimeter-wave-radar measurements.

will improve inclination-estimation accuracy.

4. Automatic landing-guidance system from detecting landing point to precision landing

4.1 Arrangement pattern

In this section, we discuss the automatic landing-guidance system that can consistently guide a drone from the approach point through the attitude-adjustment point to the landing point.

MilliSign and the inclination-estimation method described in Sections 2 and 3 can be used for information and position transmission and inclination estimation, but the application ranges of these two methods are different. For example, with MilliSign, a reading-success rate of over 90% is achieved for a drone 10 m away within a range of θ_u over 30° and ϕ_u over 20°. To apply the inclination-estimation method, however, the drone needs to be located near the point $(r_u, \theta_u, \phi_u) = (1.0\text{ m}, 45°, 0°)$, and the relative position between the drone and landing point must be known. Due to the different application ranges and usable conditions of the two methods, it is necessary to consider an appropriate arrangement method.

Figure 11 shows a schematic of the automatic landing-guidance system with the landing port and CR-type chipless RFID tags arranged in parallel.

Figure 11(b) shows the view of Fig. 11(a) on the x-z plane. The angles in Fig. 11(b) are set at $\theta_e = 45°$ and $\theta_u = \arctan\left(\frac{1.5}{1.0}\right) = 56.3°$. Since the developed CR-type chipless RFID tag functions up to an θ_u of 60°, both the landing port and CR-type chipless RFID tag

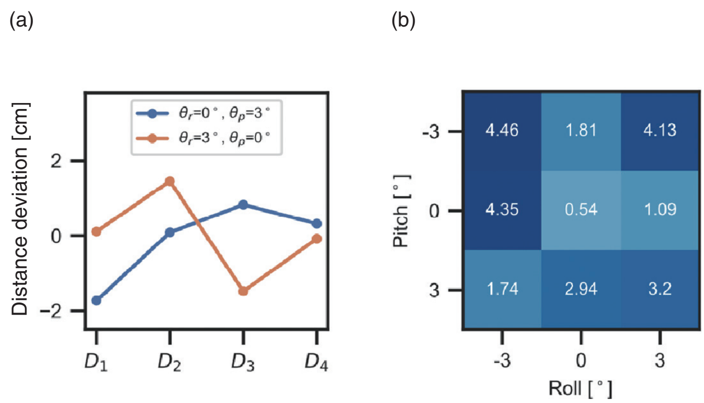


Fig. 10. (a) Deviations in distance to the *i*-th reflector caused by inclination. (b) Inclination-estimation errors for varying angles.

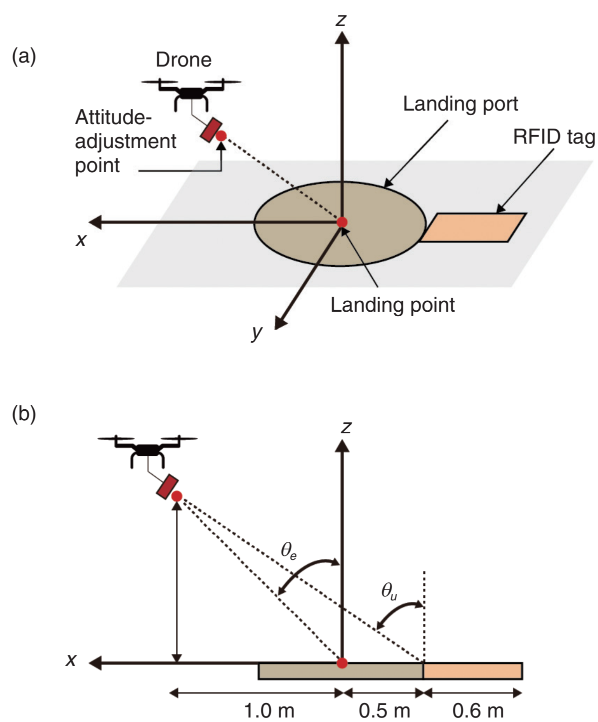


Fig. 11. Schematic of the automatic landing-guidance system using developed millimeter-wave based RFID tags and the port developed for inclination estimation.

can be read from the attitude-adjustment point. Using the relative position of the tag and the drone that can be read from the chipless-RFID tag measurement, the relative position of the landing port and the drone can be calculated, which can be used as an input for inclination estimation. By estimating the inclination of the landing port at the attitude-adjustment point and starting the descent flight at the appropriate time to be horizontal at landing, the drone can be guided to the landing point.

4.2 Applicability

We discuss the automatic landing-guidance system from the perspective of practicality. In terms of the system's area, the space on board a ship is generally limited, so a space-saving design is required. The system has side lengths of 1.6 and 1.0 m. In general industrial drones, including the propellers, the drone has a side length of about 1.5 m, so this system does not require more area than the drone occupies.

In terms of maintainability, this system does not use any batteries or electronic devices, so there is no need for battery replacement or electrical testing. The system is considered robust and highly maintainable. However, if debris gets into the grooves of the CR, the radio-wave-reflection performance will decrease, so measures are needed to make it less susceptible to dirt. Such measures can be taken, for example, by placing a radome that transmits millimeter waves on top of the tag.

5. Conclusion

We presented an automatic landing-guidance system to guide a drone located at a distance to a landing point on a ship. We introduced MilliSign for enabling position and information transmission using a CR-type chipless RFID tag and our inclination-estimation method for the landing port. We confirmed that it is possible to consistently guide a drone from the approach point through the attitude-adjustment point to the landing point with this system. For future work we will verify the system's performance through experiments and demonstrate the feasibility of landing on a ship.

References

[1] B. D. Song, K. Park, and J. Kim, "Persistent UAV Delivery Logistics: MILP Formulation and Efficient Heuristic," Computers & Industrial Engineering, Vol. 120, pp. 418–428, 2018. https://doi.org/10.1016/j.cie.2018.05.013

[2] J. Inoue and K. Sato, "Toward Sustainable Meteorological Profiling in Polar Regions: Case Studies Using an Inexpensive UAS on Measuring Lower Boundary Layers with Quality of Radiosondes," Environmental Research, Vol. 205, p. 112468, 2022. https://doi.org/10.1016/j.envres.2021.112468

[3] Z. Lyu, W. Ding, X. Sun, H. Sang, Y. Zhou, P. Yu, and L. Zheng, "Dynamic Landing Control of a Quadrotor on the Wave Glider," Journal of Marine Science and Engineering, Vol. 9, No. 10, 1119, 2021. https://doi.org/10.3390/jmse9101119

[4] T. Iizuka, T. Sasatani, T. Nakamura, N. Kosaka, M. Hisada, and Y. Kawahara, "MilliSign: mmWave-based Passive Signs for Guiding UAVs in Poor Visibility Conditions," Proc. of the 29th Annual International Conference on Mobile Computing and Networking (MobiCom '23), Madrid, Spain, Oct. 2023, Article no. 50, pp. 1–15. https://doi.org/10.1145/3570361.3613264

[5] T. Iizuka, T. Sasatani, T. Nakamura, N. Kosaka, M. Hisada, and Y. Kawahara, "An Inclination Estimation Method for UAV Landing Surfaces Using Millimeterwave Radar," Proc. of 2023 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2023), Pasadena, CA, USA, 2023, pp. 2049–2052. https://doi.org/10.1109/IGARSS52108.2023.10282437

[6] D. H. Nguyen, M. Zomorrodi, and N. C. Karmakar, "Spatial-based Chipless RFID System," IEEE Journal of Radio Frequency Identification, Vol. 3, No. 1, pp. 46–55, 2019. https://doi.org/10.1109/JRFID.2018.2887162

[7] A. Lazaro, J. Lorenzo, R. Villarino, and D. Girbau, "Modulated Corner Reflector Using Frequency Selective Surfaces for FMCW Radar Applications," Proc. of the 45th European Microwave Conference (EuMC 2015), Paris, France, Sept. 2015, pp. 111–114. https://doi.org/10.1109/EuMC.2015.7345712

[8] K. R. Brinker and R. Zoughi, "Corner Reflector Based Misalignment-tolerant Chipless RFID Tag Design Methodology," IEEE Journal of Radio Frequency Identification, Vol. 5, No. 1, pp. 94–105, 2020. https://doi.org/10.1109/JRFID.2020.3034483

[9] J. G. Hester and M. M. Tentzeris, "Inkjet-printed Flexible mm-Wave Van-Atta Reflectarrays: A Solution for Ultralong-range Dense Multitag and Multisensing Chipless RFID Implementations for IoT Smart Skins," IEEE Transactions on Microwave Theory and Techniques, Vol. 64, No. 12, pp. 4763–4773, 2016. https://doi.org/10.1109/TMTT.2016.2623790

[10] E. Soltanaghaei, A. Prabhakara, A. Balanuta, M. Anderson, J. M. Rabaey, S. Kumar, and A. Rowe, "Millimetro: mmWave Retro-reflective Tags for Accurate, Long Range Localization," Proc. of the 27th Annual International Conference on Mobile Computing and Networking (MobiCom '21), New Orleans, LA, USA, Oct. 2021, pp. 69–82. https://doi.org/10.1145/3447993.3448627

[11] T. Iizuka, Y. Toriumi, F. Ishiyama, and J. Kato, "Root-MUSIC Based Power Estimation Method with Super-resolution FMCW Radar," Proc. of 2020 IEEE/MTT-S International Microwave Symposium (IMS), Los Angeles, CA, USA, Aug. 2020, pp. 1027–1030. https://doi.org/10.1109/IMS30576.2020.9224082

[12] Texas Instruments, "IWR1443 Single-chip 76-GHz to 81-GHz mmWave Sensor Evaluation Module," Accessed: Mar. 16, 2023. https://www.ti.com/tool/IWR1443BOOST

Tatsuya Iizuka

Researcher, NTT Space Environment and Energy Laboratories.

He received an M.E. in information science and technology and Ph.D. in electrical engineering and information systems from the University of Tokyo in 2018 and 2023. He joined NTT in 2018 and is currently engaged in research on global environmental future-forecasting technology at NTT Space Environment and Energy Laboratories. His research interests include radar, wireless sensor networks, and earth observation systems. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).

Recent Trends in GlobalPlatform: Digital Trust – Evaluation & Certification, Trusted Execution Environment, and Digital Identity –

Eikazu Niwano, Akira Nagai, and Fumiaki Kudoh

Abstract

In addition to Internet-of-Things security, such as zero trust and supply chain issues, the environment surrounding secure components is rapidly changing due to advances in digital trust technologies such as confidential computing and digital identity. This article introduces the standardization trends in GlobalPlatform in response to these latest changes in the security environment.

Keywords: Secure Element, Trusted Execution Environment, trust

1. What is GlobalPlatform?

A smart card is a type of computer that has strong resistance to external attacks, which is called tamper-resistance. In the late 1990s, in response to the growing need for multipurpose use of smart cards, particularly in the public sector, technology was developed to remotely operate and manage multiple applications after issuance of smart cards. Standardization has since progressed in line with the era of mobile and Internet of Things (IoT).

GlobalPlatform (GP) [1] is a leading standards organization in the operation and management of multi-application smart cards since its establishment in 1999 [2]. GP specifies the following four technical standards (**Figs. 1** and **2**).

Firstly, the most representative technology of GP is Secure Element (SE), which is implemented in smart cards. Embedded subscriber identity modules (eSIMs), which can change the operator after it is issued, and integrated SIMs (iSIMs), which are integrated into system on chips (SoCs), low in cost, and used for secure boot, are emerging and diversifying.

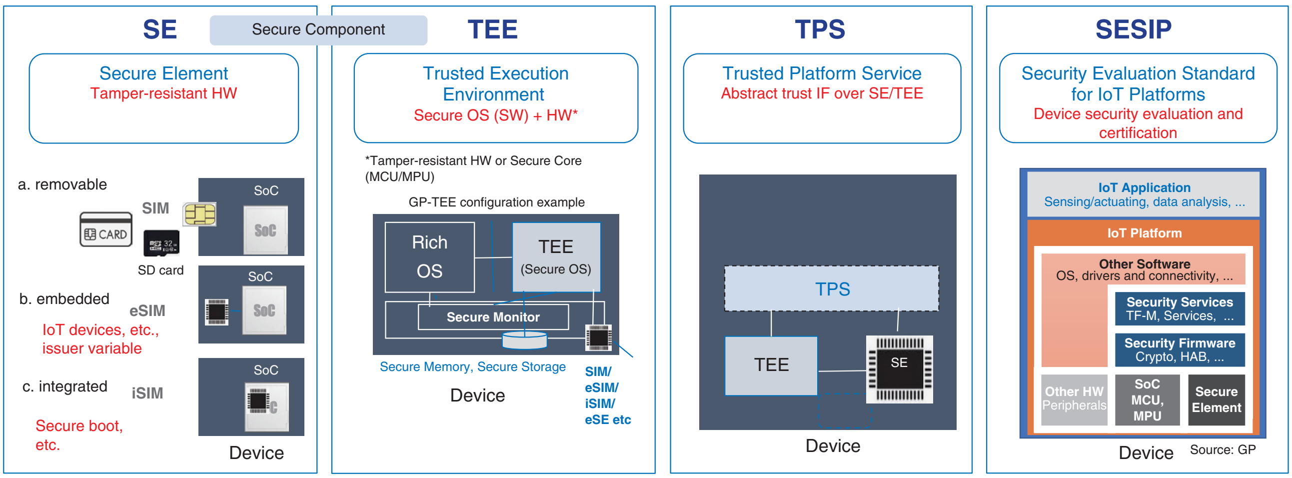
Secondly, GP standardizes a secure operating system (OS) technology called Trusted Execution Environment (TEE), which is installed in a device along with a normal device OS such as Android. This standard is widely used mainly in consumer devices such as mobile phones, Internet protocol televisions, wearable devices, and automotive onboard equipment software. GP has begun studying RISC-V*1 support and a scheme for installing multiple TEEs on a hypervisor.

Thirdly, GP has started to standardize the Trusted Platform Service (TPS), which is provided with trust-related services such as a remote attestation*2, as a higher-level interface to the two traditional secure components SE and TEE [3].

Finally, the Security Evaluation Standard for IoT Platforms (SESIP), an evaluation and certification

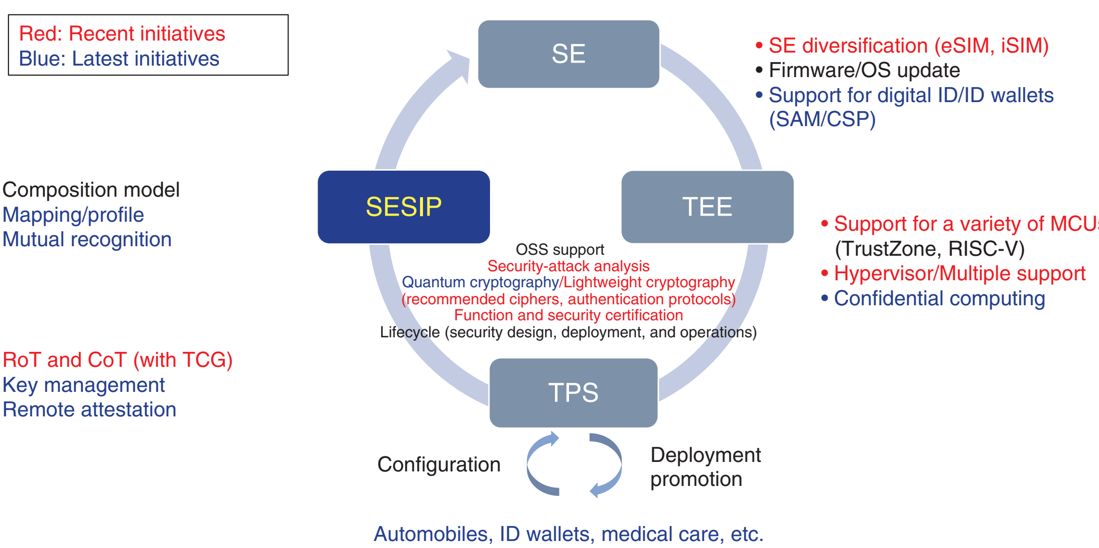
*1 RISC-V: An open-source instruction set architecture based on the reduced instruction set computer (RISC) concept.

*2 Remote attestation: A mechanism for remotely proving and verifying the authenticity of the configuration of device resources. A number of organizations, including the IETF, have made progress in studying the issue.



HAB: high availability boot
HW: hardware
HSM: hardware security module
MCU: microcontroller unit
MPU: memory protection unit
SD card: secure digital card
SW: software
TF-M: TrustedFirmware-M

Fig. 1. GP technologies.



SAM: Secured Application For Mobile
CoT: chain of trust

Fig. 2. Advances in GP technologies.

scheme for device security, is currently being promoted.

As a common approach, open source software (OSS) support and recommended cipher lists have been established, and post-quantum cryptography (PQC) has been studied (Fig. 2). The technology to be standardized by GP continues to expand from the secure component to the device level, and GP is currently developing into a standardization organization for secure devices and services.

2. Changes in GP circumstances and GP technology

The environment surrounding secure components and secure devices has changed greatly along with the expansion of these technologies. Given the increasing number of cybersecurity attacks, especially in the IoT field, there is an urgent need to address device-security issues such as zero trust and supply chain. As data consolidation advances, technologies related to data protection and privacy protection are advancing, and new data-security technologies, such as confidential computing and privacy-enhancing technologies (PETs)*3, are emerging.

In the user-security field, the need for mutual use of digital identities (IDs) in the public sector is increasing, mainly in Europe, and the examination and standardization of ID wallets that enable the unified management and use of various digital IDs in a wide area are rapidly advancing. Therefore, GP is making progress as discussed in the following sections.

2.1 Zero trust—Root of trust

In response to the complexity of the IoT environment, which is increasingly becoming an ecosystem, zero trust is an approach and architecture that assures security on a per-asset basis, such as devices, software, and IDs, rather than the traditional borderline security that protects against attacks at the border. A key issue for ensuring the trust of a device is how to configure the root of trust (RoT)*4 [3]. Secure components are becoming an increasingly important technology for this purpose.

To promote RoT study, GP is strengthening cooperation with the Trusted Computing Group (TCG), which promotes the Trusted Platform Module (TPM), an SE that is installed in personal computers for the purpose of secure boot, and is developing an RoT framework. As a mechanism to guarantee the authenticity of devices by using RoT, GP is also urgently standardizing key management and stipulating the aforementioned remote attestation service [3].

In response to this situation, the IoT/machine-to-machine (M2M) standards organization OneM2M specified the application of this secure component in the standards TS0003/TS0016, which were recently officially ratified by the International Telecommunication Union - Telecommunication Standardization Sector (ITU-T) Study Group 20 (Smart Cities). The application of secure components will progress in various IoT fields.

2.2 Supply-chain issues—Evaluation and certification of device security

As the arrival and distribution of IoT products across national borders is expanding, countermeasures against supply-chain issues to avoid the introduction of vulnerabilities are being rapidly promoted, particularly in Europe and the United States.

In Europe, since the enactment of the NIS Directive on cybersecurity for networks and information systems in 2016, the Cybersecurity Act as a cybersecurity-certification scheme, the Cyber Resilience Act requiring submission of the software bill of materials (SBOM)*5 for digital products and conformity assessment, and the European Union Cybersecurity Certification Scheme on Common Criteria (EUCC) were drafted. The European Telecommunications Standards Institute (ETSI) has defined TS 103 645, and the National Institute of Standards and Technology (NIST) has defined IR 8259 A in relation to the evaluation and certification for IoT/consumer devices. ETSI also established the TS 103 732 security requirement for mobile devices, and the Global System for Mobile Communications Association (GSMA) has been following suit.

In these circumstances, GP has established SESIP, an evaluation and authentication scheme for device security, and is working with the above organizations to map these guidelines. SESIP has been fully adopted by the European Committee for Standardization (CEN)/European Committee for Electrotechnical Standardization (CENELEC), the official European standardization committees, and was published as the European standard EN17927 in November 2023.

In Japan, ECSEC Laboratory*6 [4] has been certified as an SESIP evaluation body, and products of Renesas obtained SESIP certifications. GP has concluded a memorandum of understanding (basic agreement) with the Connected Consumer Device Security Council (CCDS)*7 in relation to secure

*3 Confidential computing and PETs: Both involve data protection. The Confidential Computing Consortium is considering the former on the basis of TEE. The latter uses TEE as a method to enable secure computation.

*4 RoT: Defined by NIST and other organizations to ensure device trust.

*5 SBOM: A mechanism to guarantee the configuration of software components. SBOM has been attracting attention since the issuance of the US Executive Order 14028.

*6 ECSEC Laboratory: A national evaluation organization and accredited as an evaluation body for SESIP in 2023.

*7 CCDS: A review organization of the security of domestic consumer devices and is active in more than 200 companies. It provides IoT security guidelines and IoT device authentication programs.

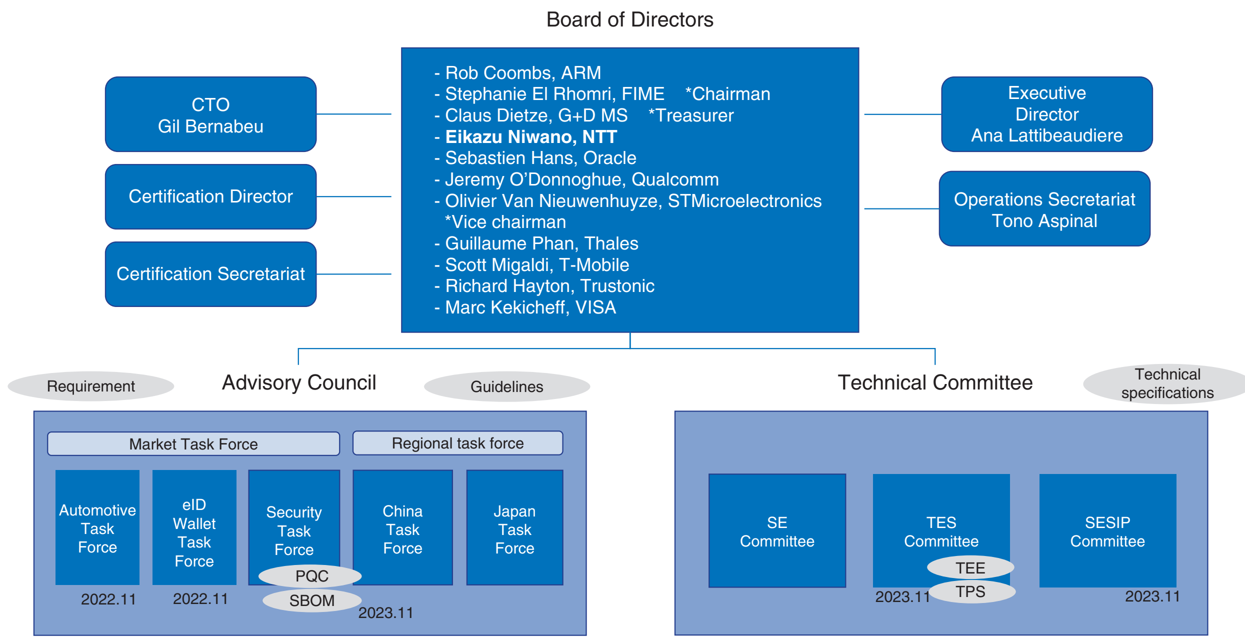


Fig. 3. GP organizations.

components and certification program. Other activities have started in the area of automotive.

To accelerate this standardization work and development, GP established a technical committee dedicated to promoting SESIP study in November 2023 (**Fig. 3**). To promote SBOM, which is important in addition to evaluation and certification, a subordinate organization related to SBOM was established and has moved under the Security Task Force (Fig. 3).

2.3 Confidential computing/PETs—TEE

Confidential computing, which is defined and discussed by the Confidential Computing Consortium, is a technology for data encryption and protection while using TEE as a base technology. There are many discussions on the use of TEE as one of the ways to achieve secure computing (security technology that enables computation while keeping data secret), which is one of the methods of PETs.

To respond to such needs, the TEE Committee and TPS Committee, which were previously independent committees within GP, were integrated and the Trusted Environment and Services (TES) Committee was established in November 2023. We expect to see progress in standardization related to confidential computing.

2.4 ID wallet—Digital ID

The European Digital Identity Wallet (EUDIW) is being aggressively promoted in Europe as a measure to enhance the utilization and sophistication of eIDAS (Electronic Identification, Authentication and Trust Services), a European public ID measure that has been implemented in Europe, and to enhance interoperability across Europe. This enables the presentation and use of various public IDs, credentials, and certificates owned by individuals in Europe.

In response, GP is working to standardize how to configure ID wallets in SE and the cryptographic service provider (CSP) technology that can be used not only for cryptographic services but also for certification of digital ID applications. The Secure Identity Alliance (SIA), an organization that promotes standardization related to the wide-area operation of digital IDs, has also started cooperation related to the technology Open Standard Identity APIs (OSIA), and it is expected that cooperation and standardization with organizations related to ID management will progress as a new area of GP. GP established the eID Wallet Task Force in November 2022 to promote this activity (Fig. 3).

3. Acceleration of market expansion—Automotive

In response to these changes in the security environment, GP is working to identify promising industries and markets and promote the diffusion of its technologies into these markets. The most promising market is the automotive market in which cybersecurity guidelines are currently being standardized by leading standards bodies. There are UN-R155/UN-R156 stipulated by the United Nations Economic Commission for Europe (UNECE), ISO/SAE 21434 stipulated by International Organization for Standardization (ISO) and the Society of Automotive Engineers (SAE) in the United States, and SAE J3101 (Hardware Protected Security for Ground Vehicles), a guideline related to secure components. GP established the Automotive Task Force in November 2022 (Fig. 3), started cooperation with the Automotive Open System Architecture (AUTOSAR), SAE, and Automotive Information Sharing and Analysis Center (AUTO-ISAC), which are global automotive organizations, and started joint work on mapping SESIP with SAE J3101.

4. Future issues and prospects

The standardization of secure components, which started with smart cards, is rapidly expanding. This is in response to recent changes in various information technology environments. It will become more important to pursue mutual use and operation with related standards to promote global interoperability. Therefore, it is necessary to cooperate with TCG for the integrated use of various SEs, standardize inter-TEE connections in a distributed environment, develop a global mutual recognition scheme for evaluation and certification technology, support further digital ID technologies such as decentralized identity (DID)*8 and self-sovereign identity (SSI)*8, verifiable credential (VC)*8 that can be verified online, and blockchain technology, and cooperate with related standardization organizations such as the Internet Engineering Task Force (IETF).

NTT has participated in GP activities since the establishment of GP and has proposed a system for managing SEs based on public key infrastructure, which is now being developed in various fields including eSIM/IoT. In addition to these technical contributions, significant organizational contributions have been made, including serving as a long-time board member, establishing a regional organization scheme in GP, and leading domestic activities as a representative of the Japan Task Force.

In 2023, in recognition of such long-standing efforts, Eikazu Niwano (the first author) became the second person in the world to receive the Kekicheff Award (the first in the world after Marc Kekicheff, from whom the award is derived). This is awarded to a person who creates outstanding long-term and continuous achievements in GP.

NTT is currently promoting the Innovative Optical and Wireless Network (IOWN) initiative, which is based on photonics-electronics convergence technology, and the IOWN Global Forum, an organization that promotes this initiative. On the basis of the results of these activities, NTT hopes to continue to contribute to the development of this field and industry by disseminating and contributing technologies from Japan, promoting cooperation between GP and domestic organizations, and contributing to the creation of a global environment for mutual use through domestic feedback on GP technologies.

References

[1] Website of GlobalPlatform, https://globalplatform.org/
[2] E. Niwano and H. Goromaru, “Standardization Trends at GlobalPlatform,” NTT Technical Review, Vol. 4, No. 11, pp. 48–52, 2006. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr200611048.pdf
[3] E. Niwano, “New Standardization Trends at GlobalPlatform—Secure Components for the IoT Era,” NTT Technical Review, Vol. 17, No. 2, pp. 63–69, 2019. https://doi.org/10.53829/ntr201902gls
[4] ECSEC Laboratory, IoT Evaluation (in Japanese), https://www.ecsec.jp/publics/index/42/

*8 DID/SSI/VC: These are core technologies of digital ID technology, and discussion has been progressing with IETF.

Eikazu Niwano

NTT Research Professor, NTT Social Informatics Laboratories.

He received a B.S. and M.S. in mathematics from Waseda University, Tokyo, in 1987 and 1989. He joined NTT in 1989 and has been engaged in research on distributed system architecture, social information platform, and applied security including secure component. From 2002 to 2005, he worked at an NTT office in Paris, where he was involved in European and International standardization activities. He received the Minister of Internal Affairs and Communications Prize of the Information and Communication Technology Award from The Telecommunication Committee in 2018, the Chairman's Prize of the SCAT Award from the Support Center for Advanced Telecommunications Technology Research in 2021 and the Kekicheff Award from GlobalPlatform in 2023. He has been a board of directors of GlobalPlatform since 2005 and chair of the Japan Task Force of GlobalPlatform. He was a member of the smart card and ICT city-related study group of Ministry of Internal Affairs and Communications of Japan. He is also a member of the Institute of Electronics, Information and Communication Engineers of Japan.

Akira Nagai

Senior Research Engineer, NTT Social Informatics Laboratories.

He received a B.S., M.S., and Ph.D. from Tokyo University of Science in 2006, 2008, and 2017. He is currently involved in applied research on post-quantum cryptography.

Fumiaki Kudoh

Senior Research Engineer, Social Information Sharing Research Project, NTT Social Informatics Laboratories.

He received a B.E. and M.S. in engineering from Waseda University, Tokyo, in 2009 and 2011. Since joining NTT in 2011, he has been studying the field of ID federation and authentication. He is currently involved in applied research on post-quantum cryptography and confidential computing. He was with NTT DOCOMO, INC. from 2017 to 2020, where he was engaged in security enhancement of authentication infrastructure and planning, development, and sales of an electronic Know Your Customer (eKYC) service.

Case Studies of Telecommunication Problems Caused by Conducted Disturbance in AC Power Lines

Technical Assistance and Support Center, NTT EAST

Abstract

This article describes the mechanism of telecommunication problems caused by conducted disturbance in AC (alternating current) power lines and case studies of troubleshooting investigated by the Technical Assistance and Support Center, NTT EAST. This is the eighty-second article in a series on telecommunication technologies.

Keywords: conducted disturbance, noise filter, fast Fourier transform

1. Introduction

A common telecommunication problem that occurs in the customer environment or telecommunication center (telecom center) is malfunctions caused by conducted disturbance generated from alternating current (AC)-power-line-connected equipment. The generation mechanism of conducted disturbance in AC power lines includes deterioration and malfunctions of equipment such as lighting and home appliances. Such conducted disturbance can be transmitted to telecommunication equipment connected to the same AC power line directly or induced into telecommunication equipment through telecommunication lines.

In telecom centers, conducted disturbances in AC power lines can infiltrate telecommunication equipment and cause audible noise on analog telephone lines, telecommunication failure, and equipment alarms. This article describes the telecommunication problems caused by conducted disturbance in AC power lines and case studies of troubleshooting investigated by the Technical Assistance and Support Center (TASC), NTT EAST.

2. Mechanism and countermeasures of telecommunication problems caused by conducted disturbance in AC power lines

Conducted disturbance in AC power lines can infiltrate telecommunication equipment through AC power lines and telecommunication lines. As shown in **Fig. 1**, conducted disturbance in AC power lines can infiltrate telecommunication equipment through two paths: (1) direct transmission through the power line of the telecommunication equipment and (2) induction from another (nearby) power line into the telecommunication line.

There are two major effects of conducted disturbance in AC power lines on telecommunication equipment. One is that conducted disturbance interferes with telecommunication signals, which causes audible noise, bit errors, link downs, etc. The other is that conducted disturbance interferes with the current flow of internal circuit wiring of the telecommunication equipment, which causes malfunction such as emergency stop or power off operation.

To prevent these problems, it is crucial to identify and eliminate the source of conducted disturbance. If the identification or elimination is difficult, it is

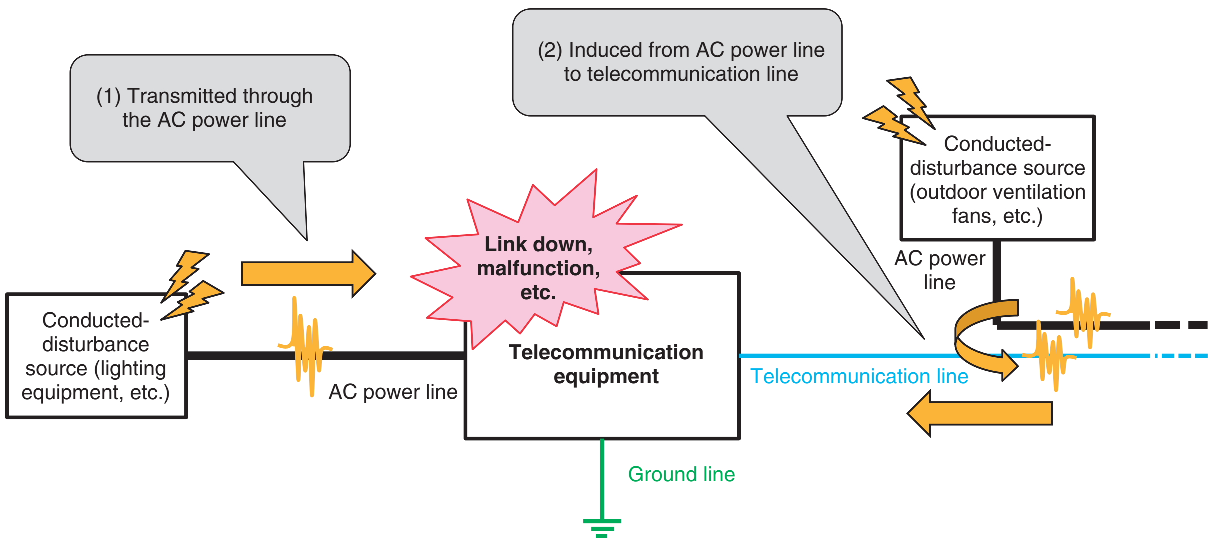


Fig. 1. Infiltration paths of conducted disturbance in AC power lines into telecommunication equipment.

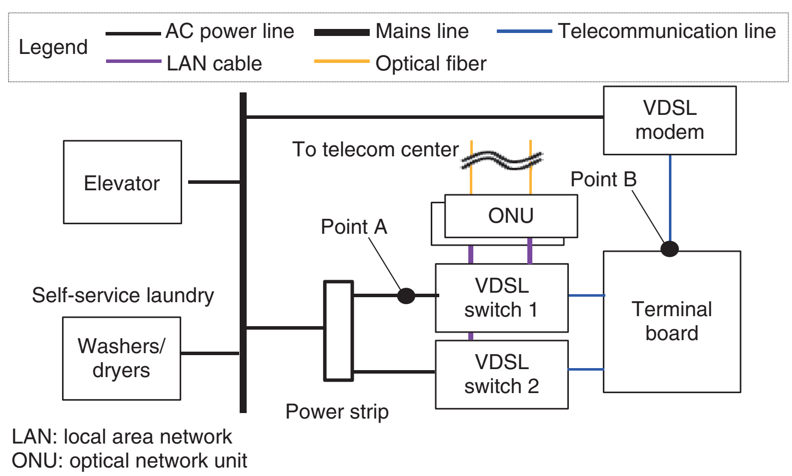


Fig. 2. Configuration of equipment.

necessary to reduce the conducted-disturbance level by installing noise filters or other countermeasures in the interference path.

3. Case study 1: Link down of VDSL switch due to conducted disturbance emitted by a washer/dryer

3.1 Overview of problem

Some customers who use very high-bit-rate digital subscriber lines (VDSLs) in their apartment building reported that they cannot use the Internet during the daytime. Local maintenance staff found that multiple ports of their VDSL switch were frequently linked down (particularly in the morning and evening). Since the cause of the problem was unclear, TASC was asked to investigate the cause and suggest countermeasures.

3.2 Equipment configuration and details and results of investigation

The configuration of equipment in the apartment building is shown in **Fig. 2**. Two VDSL switches are connected to the same AC power line, and

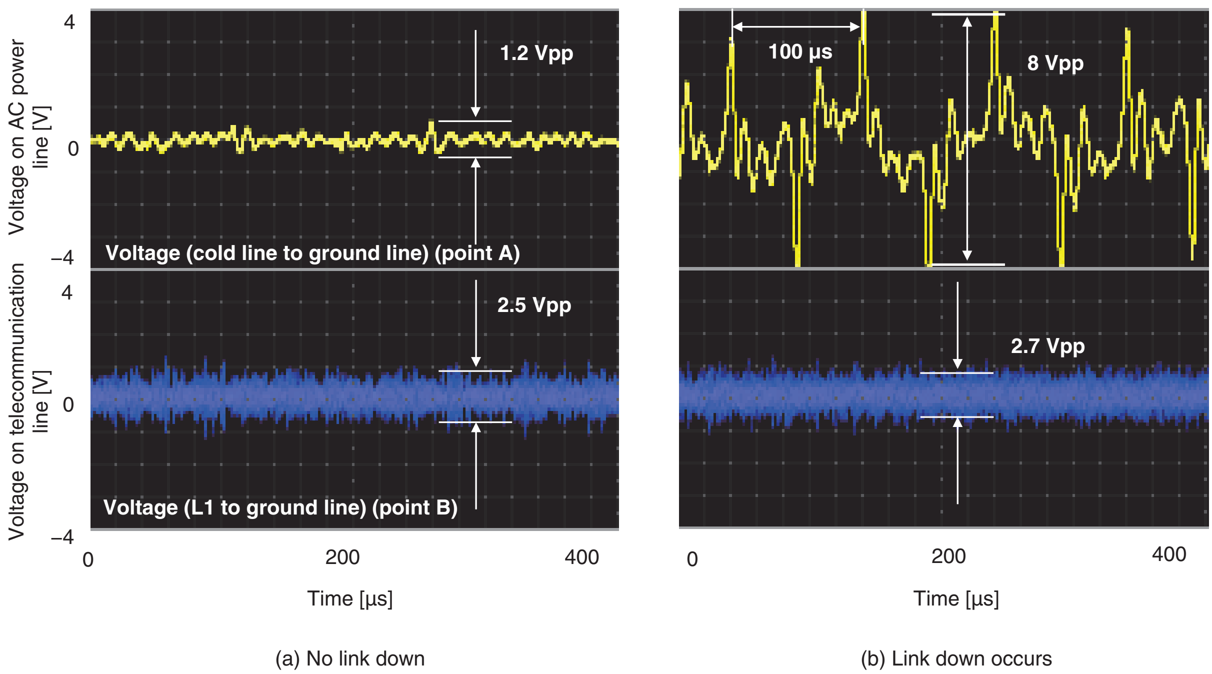


Fig. 3. Measured waveforms of conducted disturbance.

telecommunication lines are wired from the VDSL switches through the terminal board to the VDSL modem. The elevator and washers/dryers located in a self-service laundry annexed to the first floor of the building shared the same AC power-distribution system with the VDSL switches.

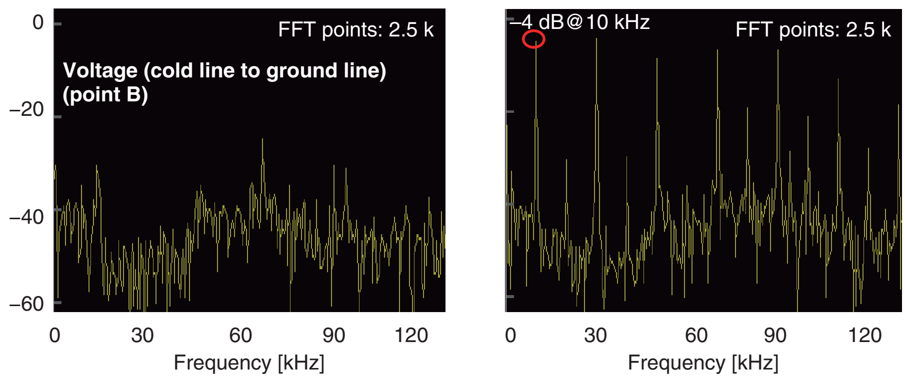
To confirm the conducted disturbance interfering with the VDSL switches, we used an oscilloscope (DLM4058 from Yokogawa Electric) to measure the voltage between the cold line of the AC power line and ground line of VDSL switch 1 (Point A in Fig. 2) and the voltage between L1 on the telecommunication line and the ground line (Point B in Fig. 2). We found conducted disturbance up to 8 Vpp on the AC power line when the link down occurred (**Fig. 3**). However, no obvious conducted disturbance on the telecommunication line was found. When the measured conducted disturbance was subjected to fast Fourier transform (FFT) processing, a fundamental frequency of 10 kHz (red circle in **Fig. 4**) and its harmonics were observed. The fundamental frequency and third harmonic had the highest conducted-disturbance level, i.e., −4 dBV.

Since we confirmed that conducted disturbance was interfering with the VDSL switches, we investigated the source of the disturbance by focusing on the elevator and washers/dryers installed in the laundry, which share the same AC power line with the VDSL switches. However, we could not conduct measurements in the power-supply control room, so we operated the elevator and washers/dryers and investigated the generation of conducted disturbance at point B in Fig. 2.

We first operated the elevator up and down several times but detected no conducted disturbance. We then operated multiple washers/dryers in the laundry and detected conducted disturbance. By turning each washer/dryer individually, we confirmed conducted disturbance when one large washer/dryer was operating (when the motor was running) and all ports of the VDSL switches exhibited a link down.

3.3 Cause and countermeasures

The results of the investigation suggested that the link down was caused by conducted disturbance of 10 kHz generated from one large washer/dryer, which was interfering with the VDSL switches through the AC power lines. As a countermeasure, it is necessary to inspect and repair the washer/dryer, which is the source of conducted disturbance. To suppress the impact of conducted disturbance during the investigation period, however, we implemented a countermeasure using noise filters for VDSL switch 1 and investigated its effectiveness. We tested two different



(a) No link down

(b) Link down occurs

Fig. 4. Measured waveforms of conducted disturbance (FFT-processing results).

noise filters for AC power lines (NCT-i3 from DENKENSEIKI Research Institute and CNF-5A from SANRITZ ELECTRONICS), each was installed on the AC power line (at point A in Fig. 2) in separate tests when the washer/dryer was operated for 20 minutes. We confirmed that the link down did not occur in VDSL switch 1 regardless of which filter was installed.

4. Case study 2: Clock-system failure of switching equipment caused by conducted disturbance emitted from a ventilation fan

4.1 Overview of problem

A clock-system failure (a malfunction detected in a clock signal) occurred in switching equipment installed in telecom center A. The local maintenance staff found that when they turned the ventilation fan on or off, the same malfunction occurred. Since the cause of the malfunction was unclear, TASC was asked to investigate the cause and suggest countermeasures.

4.2 Configuration of equipment and details and results of investigation

The configuration of the equipment in telecom center A is shown in **Fig. 5**. Telecom center A consists of multiple units, and the switching equipment was supplied with a clock signal via a lightning-protection isolation filter (IF-C) from a clock-supply module installed in a different unit of telecom center A.

The ventilation fan was installed in the same unit as the switching equipment and supplied with AC power from another unit. The clock line from the IF-C to the switching equipment was wired closely to the AC power line of the ventilation fan for about 10 m.

To investigate the correlation between clock-system failure and conducted disturbance when the power of the ventilation fan is turned on or off, we used an oscilloscope (DLM4058 from Yokogawa Electric) to measure the current in three locations: (1) the AC power line connected to the ventilation fan, (2) the clock line from the IF-C to the switching equipment, and (3) the direct current (DC) power line of the switching equipment.

The measurement results are shown in **Fig. 6**. When the ventilation fan was turned on, conducted disturbance of 1740 mApp was generated on the AC power line (1). In synchronization with that conducted disturbance on the AC power line, conducted disturbances of 70 mApp on the clock line (2) and 50 mApp on the DC power line (3) were also measured.

The analysis results of applying FFT processing to the measurement (2) are shown in **Fig. 7**. These results indicate that the conducted-disturbance level increased in the 0.7–4.5-MHz range.

4.3 Cause and countermeasures

From the investigation results, it was found that the clock-system failure was caused by conducted disturbance from the ventilation fan, which was propagated through the AC power line and induced into the clock line.

As a countermeasure, it is necessary to eliminate the effect of the induction of conducted disturbance by re-routing the clock line. However, it was not possible to reroute the clock line during the investigation period; therefore, we installed a noise filter on AC

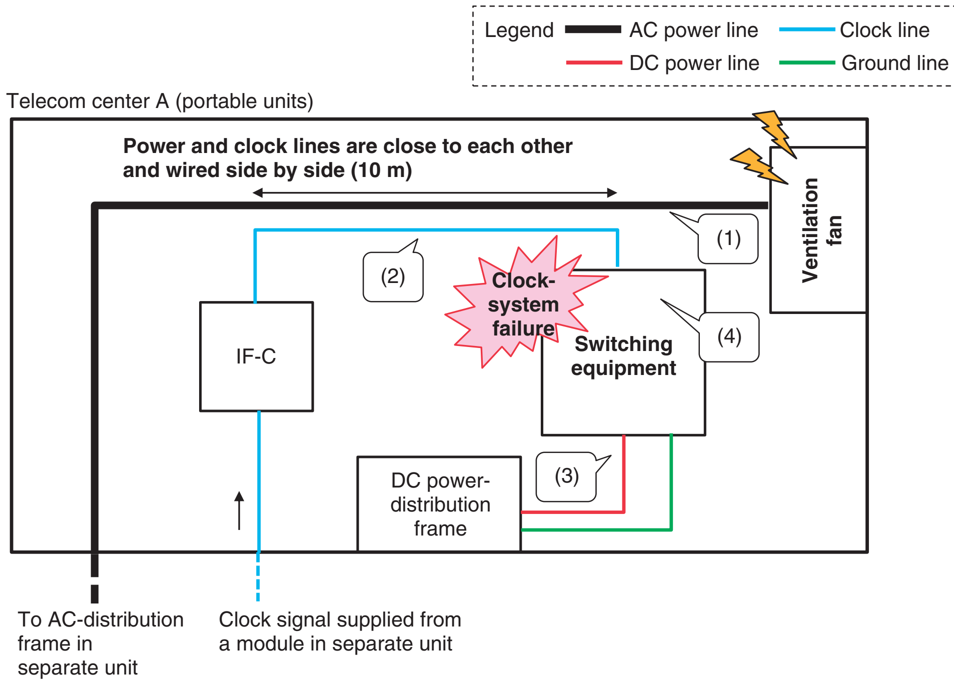


Fig. 5. Configuration of equipment.

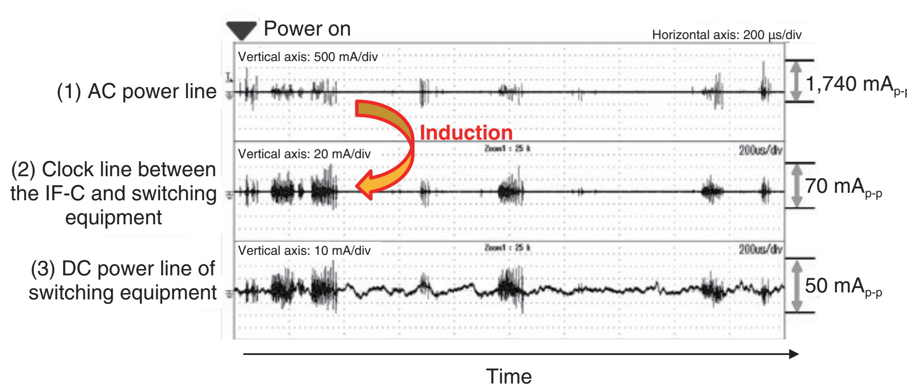


Fig. 6. Measured waveforms of conducted disturbance.

power lines (GT-215J from KEMET) between the ventilation fan and AC power line and a ferrite core (E04SRM from SEIWA ELECTRIC) on the clock line. We confirmed that the clock-system failure no longer occurred when the ventilation fan was turned on or off.

5. Conclusion

The mechanism of telecommunication problems caused by conducted disturbance in AC power lines and two case studies of troubleshooting investigated by TASC were described in this article.

Equipment connected to AC power lines, such as lighting equipment and home appliances, become sources of conducted disturbance as they deteriorate over time or break down, and may adversely affect

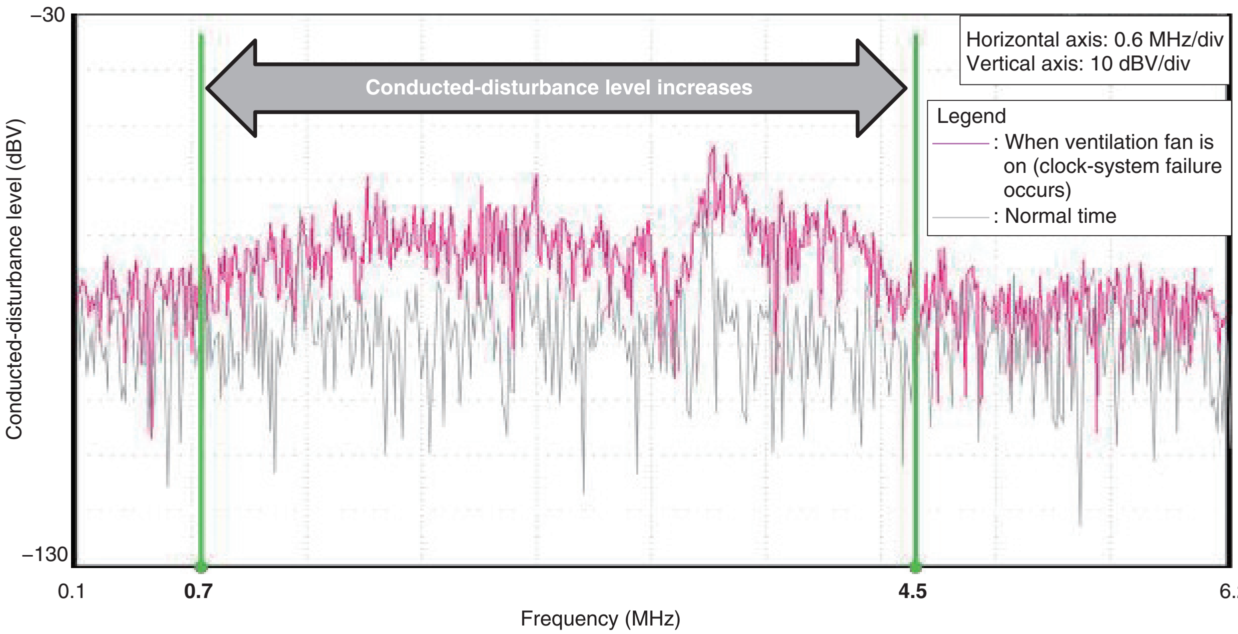


Fig. 7. Measured waveforms of conducted disturbance (FFT-processing results).

telecommunication equipment. It is therefore important to inspect and repair such equipment or take countermeasures with noise filters. Even if metal-cable sections will be converted to optical cables in the future, interference of conducted disturbance from AC power lines still may affect optical network units, home gateways, etc. At TASC, we will continue to reduce maintenance operation in the field by providing the knowledge we have accumulated.

To reduce telecommunication problems caused by electromagnetic disturbance, radio waves, induction, lightning, and other factors and improve the reliability of telecommunication services, the EMC Engineering Group in TASC will continue to engage in technical cooperation, development, and dissemination of technology through technical seminars and other activities.

External Awards

Network Systems Research Award

Winners: Hiroyuki Kitada, NTT Network Service Systems Laboratories; Daiki Fukudome, NHK Science & Technology Research Laboratories; Kento Noguchi, NTT Network Service Systems Laboratories; Takafumi Okuyama, NTT Network Service Systems Laboratories; Satoshi Nishimura, NHK Science & Technology Research Laboratories; Hisayuki Ohmata, NHK Science & Technology Research Laboratories
Date: February 29, 2024
Organization: The Institute of Electronics, Information and Communication Engineers (IEICE) Technical Committee on Network Systems

For "Implementation and Evaluation of Delay Adjuster for In-network Video Processing."
Published as: H. Kitada, D. Fukudome, K. Noguchi, T. Okuyama, S. Nishimura, and H. Ohmata, "Implementation and Evaluation of Delay Adjuster for In-network Video Processing," IEICE Technical Report, Vol. 123, No. 198, NS2023-106, pp. 172–176, 2023.

Fellow

Winner: Jun Terada, NTT Device Technology Laboratories
Date: March 5, 2024
Organization: IEICE

For research and development for practical application and advancement of high-speed optical access network systems.

JSAP Young Scientist Award

Winner: Megumi Kurosu, NTT Basic Research Laboratories
Date: March 22, 2024
Organization: The Japan Society of Applied Physics (JSAP)

For "Buckling-induced Quadratic Nonlinearity in Silicon Phonon Waveguide Structures."
Published as: M. Kurosu, D. Hatanaka, H. Okamoto, and H. Yamaguchi, "Buckling-induced Quadratic Nonlinearity in Silicon Phonon Waveguide Structures," Jpn. J. Appl. Phys., Vol. 61, SD1025, 2022.

Tingye Li Innovation Prize

Winner: Josuke Ozaki, NTT Innovative Devices Corporation
Date: March 26, 2024
Organization: Optical Fiber Communications Conference and Exhibition (OFC)

For "Net-1.8 Tbps/λ Transmission Enabled by C+L-band InP-based Coherent Driver Modulator."
Published as: J. Ozaki, "Net-1.8 Tbps/λ Transmission Enabled by C+L-band InP-based Coherent Driver Modulator," OFC 2024, San Diego, CA, USA, Mar. 2024.

Journal of Information Processing Outstanding Paper Award

Winners: Kazuki Nomoto, Waseda University; Takuya Watanabe, NTT Social Informatics Laboratories; Eitaro Shioji, NTT Social Informatics Laboratories; Mitsuaki Akiyama, NTT Social Informatics Laboratories; Tatsuya Mori, Waseda University
Date: March 29, 2024
Organization: Information Processing Society of Japan (IPSJ)

For "Understanding the Inconsistencies in the Permissions Mechanism of Web Browsers."
Published as: K. Nomoto, T. Watanabe, E. Shioji, M. Akiyama, and T. Mori, "Understanding the Inconsistencies in the Permissions Mechanism of Web Browsers," Journal of Information Processing, Vol. 31, 2023 (online).

Prime Minister's Award at Japan Industrial Technology Awards

Winners: Institute of Physical and Chemical Research, National Institute of Advanced Industrial Science and Technology, National Institute of Information and Communications Technology, Osaka University, Fujitsu Limited, and NTT Corporation
Date: April 3, 2024
Organization: The Nikkan Kogyo Shimbun

For the development of an ultra-high-performance computing platform using a 64-qubit superconducting quantum computer and provision of the platform on the cloud for corporate use.

Maejima Hisoka Encouragement Award

Winner: Kohki Shibahara, NTT Network Innovation Laboratories
Date: April 11, 2024
Organization: Tsushinbunka Association

For research on long-haul optical repeatered transmission systems using mode division multiplexing.

Papers Published in Technical Journals and Conference Proceedings

Unconditional Verification of Quantum Computation with Classical Light

Y. Takeuchi and A. Mizutani

arXiv:2403.14142, March 2024.

Verification of quantum computation is a task to efficiently check whether an output given from a quantum computer is correct. Existing verification protocols conducted between a quantum computer to be verified and a verifier necessitate quantum communication to unconditionally detect any malicious behavior of the quantum computer solving any promise problem in **BQP**. In this paper, we remove the necessity of the communication of qubits by proposing a "physically-classical" verification protocol in which the verifier just sends coherent light to the quantum computer.