NTT Technical Review

11
2022

November 2022 Vol. 20 No. 11

NTT Technical Review

November 2022 Vol. 20 No. 11

Think Big, Start Small with High Aspirations and Passion

Yo Honma
President and Chief Executive Officer, NTT DATA

Abstract

It has been over half a century since the Data Communications Division was established within the Nippon Telegraph and Telephone Public Corporation (now NTT) in 1967, which was spun off as NTT DATA in 1988. NTT DATA has earned the trust of many of its clients by understanding their business operations and using its technology to create and operate new systems that support society. We interviewed Yo Honma, president and chief executive officer of NTT DATA, who is leading the company's transformation while maintaining long-term relationships with clients, about the business environment surrounding NTT DATA and his thoughts as a top executive.

Keywords: digital transformation, profitable global growth, sustainability

Pursuing profitable global growth through consistent belief and courage to change

—First, would you tell us about the business environment surrounding NTT DATA?

We have been striving to expand our business globally since FY2005. After executing the so-called Global 1st and 2nd Stages of this effort, we are currently implementing the Global 3rd Stage under which we aim to obtain the top-five status in net sales among the global information technology (IT) services industry by around FY2025 and become a global company trusted by clients worldwide. As the preliminary step of this stage, during the previous medium-term management plan from FY2019 to FY2021, we pursued profitable growth on a global scale with consistent belief and courage to change.

We have worked with many clients to construct IT systems that support their business and contributed to society. We will continue to maintain and nurture our long-term relationships of trust with our clients as our consistent belief. At the same time, we are thoroughly committed to making a solid profit and investing it in forward-looking projects. Outside Japan, we have carried out restructuring of our business, which resulted in expanding it through global offerings and acquisition of various digital businesses in each region. These were achieved through our courage to change.

We achieved three of our four business goals in FY2021; consolidated net sales of 2.5-trillion yen, a client base of more than 80 companies, and consolidated operating income margin of 8%. Although an EBITA (earnings before interest, taxes, depreciation, and amortization) margin of 6.5% did not quite meet the target goal of 7%, partly due to the impact of the COVID-19 pandemic, we achieved 7% in North America.

During the previous medium-term management plan, our business was greatly affected by the COVID-19 pandemic. We can say that our business-to-business

business, which is the core of our business, faced two "winds": a head wind and a following wind. First, clients in the transportation, travel, and service industries faced a difficult environment and curtailed their IT investments. That is the head wind we faced. Many clients wanted to reduce costs of system maintenance and operations. We have been responding to those client demands by transforming the way we work to further automate operations, increase productivity, and speed up development. Second, in accordance with the desire to build a new society, the trend to create new services, products, and business models by using IT and digital technologies has accelerated. We have been striving to harness this following wind and turn it into an opportunity for growth.

Supporting the growth of our clients' businesses and working together with them to build a sustainable future

—It's great to be able to achieve your goals even in the COVID-19 pandemic. What strategies will be deployed in the new medium-term management plan?

Although we have almost achieved the business goals set out in our previous medium-term management plan, we must aim even higher to accomplish the Global 3rd Stage. Our new medium-term management plan (for FY2022 to FY2025) announced in May 2022 further advances the principles set out in our previous medium-term management plan.

As I mentioned earlier, the purpose of IT investment is shifting from digitizing a company's existing operations to creating new services and business models through IT and digital technologies. Digital transformation to solve social issues and help preserve the global environment is also gaining pace. Various players are also expanding their service offerings in response to changes in society and technology, further intensifying competition.

Under such circumstances, we must further strengthen our competitiveness. Therefore, we will continue to pursue profitable global growth and further expand digital offerings. Companies are witnessing a shortage of IT personnel and engineers worldwide. It is no exaggeration to say that our business is all about human resources, so it is vital to enhance the capabilities of our employees and organization as a whole.

Under our new medium-term management plan, we also need to enhance the two strengths we have cultivated thus far to support growth of our clients' businesses and creating a sustainable future together with our clients and become a *trusted global innovator*. One strength is the ability to create systems while understanding our clients and using our advanced technical capabilities. The other strength is the ability to connect people, companies, and society by supporting various enterprise systems and social infrastructure.

The objective of our clients had been to replace existing business and social systems with IT systems, so how to create highly reliable systems had been our key theme. From now on, however, our clients will be required to create new services, products, and business models by using IT and digital technologies, and what to create is becoming the more important theme. To meet the needs of the times, we will strengthen our consulting capabilities on the basis of foresight that looks beyond the client's industry and collaborate with clients to create new services and business models. Of course, the architecture for creating such services and business models is also important. Our strength will be our ability to propose what is feasible by leveraging our advanced knowledge and technology.

Recent situation, such as climate change and the COVID-19 pandemic, have highlighted many social problems. The key to solving these problems is to use IT and digital technologies. We want to contribute to solving social problems through the digitization of society as a whole by enhancing not only our digital capabilities but also that of our partner companies and clients. Therefore, in our new medium-term management plan, we have positioned ourselves more clearly to contribute to solving social issues through our business activities under the slogan of "Realizing a Sustainable Future."

—So your new medium-term management plan focuses on supporting clients, partners, and the industry and solving social problems, right?

Since 1967, for more than 55 years, when NTT DATA's predecessor, the Data Communications Division, was established at the Nippon Telegraph and Telephone Public Corporation (now NTT), we have approached our business with an attitude focusing on benefiting society. The corporate mission of NTT DATA is "to use information technology to create new paradigms and values, which help contribute to a more affluent and harmonious society." With this mission, which has been passed down from generation to generation, in mind and by nurturing the long-term relationships that we value, we will work in harmony with our clients to build the society of the future.

In the current era of the Internet of Everything, all people and things are becoming connected through the advancement of digitalization. Some data are collected from the edge, while other data are collected from smartphones and other devices, which include consumer-contact data. Securely collecting, analyzing, and using data, which is often referred to as "the oil of the 21st century," beyond the boundaries of companies and industries will make it possible to address complex social issues.

To make such data distribution and utilization a reality, it is necessary to be able to handle such a massive amount of data with high reliability, and the NTT Group's Innovative Optical and Wireless Network (IOWN) is expected to meet this requirement. NTT DATA established the IOWN Promotion Office in January 2021 with the aim of understanding social trends and client needs, verifying the component technologies of IOWN, and promoting implementation of IOWN. We are focusing on a data-collaboration platform and Digital Twin Computing to envision and build a digital-twin platform that will be used across a wide variety of industries and business sectors for digital transformation of society.

The NTT Group is also working in unison to achieve its environmental and energy vision "NTT Green Innovation toward 2040." NTT DATA's green innovation initiatives are based on two pillars: "Green Innovation of IT" to reduce our greenhouse gas (GHG) emissions, and "Green Innovation by IT" to reduce GHG emissions of our clients and society as a whole through our business activities.

Regarding "Green Innovation of IT," we aim to achieve the science-based target of reducing GHG emissions by 60% by FY2030 compared with the FY2016 level. Approximately 70% of our electricity consumption is accounted for by the datacenters we run. By introducing innovative cooling and air-conditioning technologies and advanced IT technologies, we plan to replace all electricity used in our datacenters

with renewable energy by 2030. We also aim to become carbon neutral in terms of Scope 1 (direct emissions from resources owned or controlled by the company) and Scope 2 (indirect emissions from the purchased energy) by 2040 and net zero (which includes Scope 3 (indirect emissions in the supply chain)) by 2050.

Regarding "Green Innovation by IT," we will contribute to reducing emissions through the systems and services we provide. For example, Trade Waltz, a trade-information collaboration platform, digitizes trade administrative work such as contracts, settlements, and customs clearance, which used to be mostly paper based. This platform contributes to not only improved work efficiency and cost reduction but also conservation of forests and reduction of carbon dioxide (CO_2) emissions. We estimate that if all players involved in Japan's trade practices used Trade Waltz, up to 30,800 trees could be conserved and 431 tons of CO_2 could be reduced annually.

Focusing on the creation of strength, creation of value, and development of human resources

—What did you experience to reach your current position?

I joined the Nippon Telegraph and Telephone Public Corporation in 1980. Back then, Yasusada Kitahara, the vice-president of Nippon Telegraph and Telephone Public Corporation, launched a concept called the information network system (INS). Attracting attention from around the world, the INS aimed to digitize and integrate all communication networks and provide various services, such as telephone, data communication, and image communication, through an integrated network. After reading the book explaining the concept, I thought a job working on the INS sounded interesting, so I was inspired to join the company.

My 40 years at NTT has been a daily joy, although sometimes a struggle. Our system-development work is project based, so both joys and hardships are shared among the project team members. When the job is completed, everyone hugs and pats each other on the back. That joy is unmatched by anything else. The saying, "All for one and one for all" is a very important spirit to have. We also share hardships and joys with our clients. One of the most impressive projects I had ever been involved in was one to develop a travel-industry system. We were unfamiliar with the travel industry, and the project was becoming problematic. We first tried our best to keep things running smoothly, but over time, the project team members became exhausted. Talented people could be assigned to solve problems, but that approach can lead to the loss of new business opportunities.

With that in mind, we explained our situation to the client and asked for their cooperation in making the project a success. In fact, this simple act is extremely important. Things do not work well in a typical relationship between a contractor and contractee; it is important to recognize that we and our clients are in the same boat and that cooperating with each other leads to the success of the project and great satisfaction when it is completed.

Therefore, I have always valued the power of organization, where we work together as one team, on the same footing, to create something good together. A good organization cannot be established solely on the capabilities of individuals. NTT DATA attracts exceptional people, and combining their strengths with the vitality that pervades the organization, attractiveness of the work, attractiveness of the upper management and colleagues, and attraction that both clients and we feel for each other can lead to good results. Organizational strength can be represented as the multiplication of strength, vitality, and attractiveness.

Using other words to make a similar point, I can say "skill" and "will" are necessary to produce good results. Skill is very important, but it is not enough; will is also important. Combining skills and will with teamwork will produce outstanding results. Regarding the travel-industry project I mentioned, the project manager bought vellum paper and ribbons from a store next to the head office and used it to visualize the progress of the project and facilitate discussions, which created a positive mood in the team and increased the collective strength of the team.

—Listening to you, I get the impression that you have always had a comprehensive view of things. What do you think it is important as a top executive?

I have been working with clients while keeping the importance of our long-term relationships in mind. Long-term relationships with clients go through good times and bad times; all the same, it is extremely important to do one's best for the client even when times are tough. Clients look closely at our work and say, "You've gone this far, and you've thought that far?" That recognition leads to trust.

To maintain that trust, I place importance on thoroughness in all I do. I will continue to value the attitude of being thorough in what I decide to do, even in the most mundane matters.

I also place importance on creation of strength, creation of value, and development of human resources. In particular, the various laboratories of NTT have their own areas of expertise. By combining such expertise with our value, I'd like to provide our clients with better services.

Finally, I believe that top executives must approach their work with an attitude of think big and start small, high aspirations, and passion. However, even if you have a strong will and clear vision, you will face ups and downs along the way. You might even encounter an unimaginable event. I think it is vital that top executives have the ability to stay on course and move forward, even when their team faces difficult situations. In this day and age, we have no clear charts to navigate the ocean that is the business world. Still, top executives must have a compass to provide guidelines even in a changing environment.

■ Interviewee profile

Yo Honma joined Nippon Telegraph and Telephone Public Corporation in 1980. In his career at NTT DATA Corporation, he became executive vice president and general manager of the Third Enterprise Business Division in 2013, director and executive vice president, head of the Enterprise IT Segment in 2014, and representative director and senior executive vice president in 2016. He assumed his current position in June 2018.

Front-line Researchers

Researchers Are Like Mirrors that Reflect the Future

Junji Watanabe
Senior Distinguished Researcher, NTT Communication Science Laboratories, NTT Social Informatics Laboratories, and NTT Human Informatics Laboratories

Abstract

To develop information and communication technologies that enable heart-to-heart communication, researchers at Human Information Science Laboratory in NTT Communication Science Laboratories study the information-processing mechanisms underlying human perception and emotion on the basis of information science, psychology, and neuroscience. We interviewed Junji Watanabe, a senior distinguished researcher investigating the mechanism of tactile sensation and its application to transmission technology as well as supporting the improvement of well-being, about the progress of his research.

Keywords: haptics, well-being, Self-as-We

Pursuing methodologies to ensure the well-being of diverse people

—It has been three years since our last interview. Could you tell us about your current research activities?

My research focuses on haptic communication and its value to society. The sense of touch allows us to feel our state of being as well as contributes to building empathy and trust between people. In conjunction with my research on haptic communication, I'm investigating methodologies to improve well-being.

From 2020 to 2021, we held a series of workshops using haptic-communication tools with about 30 elementary-school students (11–12 years old) in Yokohama City [1]. This "Touch and Connect Sports Lab" was attended by members of NTT laboratories and the para-athlete Akihito Tanaka, a member of a special-purpose subsidiary of NTT called NTT Claruty, to create a hands-on learning environment toward an inclusive society. Since the workshops were held during the COVID-19 pandemic, we struggled to find a way to provide a genuine learning experience for the students while also taking measures to prevent the spread of infection.

This workshop consisted of four sessions called the "Four Steps of Sensory Experience" (**Fig. 1**). In Step 1 "Reflect on yourself," the students participated in a workshop called "Heartbeat Picnic" in which they could feel their own heartbeat through the sense of touch using their hand. When the students placed the stethoscope over their chest, the square box they were holding vibrated in sync with their heartbeat. Thus,

Fig. 1. "Four Steps of Sensory Experience" presented at the Touch and Connect Sports Lab workshop (an excerpt from NTT laboratories' haptics and well-being magazine "Furue," Vol. 35 [1]).

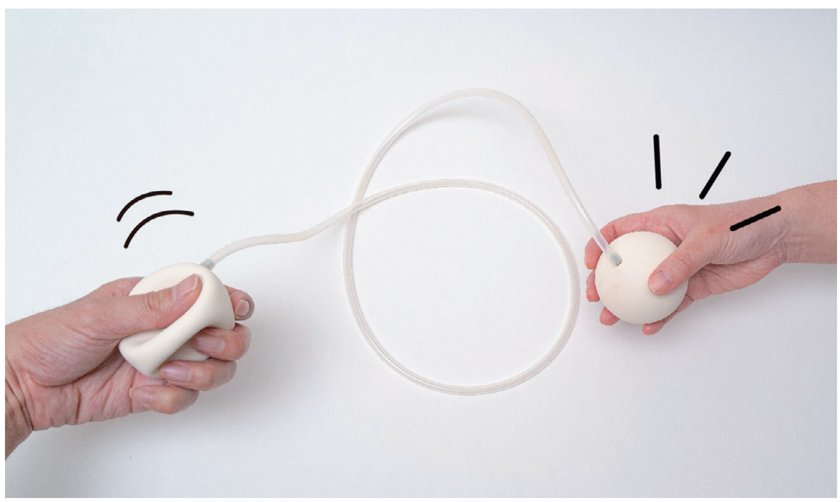


Fig. 2. Air-transmit haptic communication device.

they could feel their beating heart as a tactile sensation with their hand.

In Step 2 "Share sensations with others," the sensory experience was shared with surrounding people. As an extension of the work in Step 1, the box that vibrates to the heartbeat of the student was passed to another student so that the other student could feel the heartbeat of the former student on their hand and vice versa. By feeling the heartbeats of others, each student became keenly aware that the other students were alive in the same way as he or she was alive. The students also used an air-transmit haptic communication device (**Fig. 2**) to experience communicating emotions through tactile sensation alone. This device consists of two soft balls connected by a tube, and one of the balls (held by one student in one hand) inflates when the other (held by another student in one hand) is squeezed. The vivid experience of the ball swelling in one's response to the other person's feelings is an experience of connecting with that person that is not possible with words [2].

In Step 3 "Connect with special person," students experienced communicating with people having different senses and bodies that they would not usually come into contact with. During this step, para-athlete Akihito Tanaka joined in the workshop, and we used a tactile sharing table that vibrated when a student tapped in a manner that enabled the student to react to Akihito's story.

In Step 4 "Tell others about your experiences," students told lower-year students about their sensations they had experienced and the lessons they learned about inclusive societies.

I believe that through this workshop, students can

Fig. 3. Sports Social View (a translation of Judo).

discover their own senses, connect with others, meet other people having different senses and bodies, and learn with a sense of reality how we can live together in society. I believe that this is a fundamental experience for creating a society in which people live with respect of one another and ensure their well-being.

—Discovering one's senses sounds exciting. Could you tell us about another study?

I'll introduce a project called "Sports Social View" [3], which uses tactile sensation to enable normal-sighted people and the visually impaired to share the excitement of sports watching. The visually impaired typically attend sports events while listening to audio commentary. However, there are two problems with audio commentary. It is unable to effectively express the details of sports movements in words, and the visually impaired are left out of the excitement of spectators surrounding them. Considering these problems, we developed a method of spectating sports by which the action happening in a sports event is presented physically rather than verbally then shared with the visually impaired. In one example of using this method, the action of a judo match is conveyed to a visually impaired person. Two normal-sighted people each hold one end of a piece of cloth and, while watching the match, pretend to be one of the two judo athletes. As they spectate, they pull the cloth while imagining they are pulling the opponent's judo jacket in a manner that reproduces the power struggle between the athletes (**Fig. 3**). By grasping the cloth in the middle, the visually impaired person can feel the struggle of power and rhythm of the movement during the match, which are difficult to verbalize. We call this attempt to replace the essence of a sport with a particular physical movement "translation of sport." Later, we expanded the target of this project to include the general public and translated the essences of ten different sports into another physical movements together with athletic experts in an initiative called "Sports Guide without Sight." This initiative has been exhibited in a museum and published as a book [4].

Feeling each other through haptic technology and recognizing intrinsic value

—I feel that well-being is being re-examined in the midst of the COVID-19 pandemic.

Since emerging at the end of 2019, the COVID-19 pandemic has brought about changes in our way of life as well as an opportunity to rethink our well-being. The term "well-being" is used in a variety of ways, and I believe that well-being is achieved by actively recognizing a person's intrinsic value. Intrinsic value means that an action, thing, etc., has value in itself; in contrast, instrumental value means that something has value because it can be used to achieve a certain goal. Diversity & inclusion is a concept about respecting the intrinsic value of a person regardless of differences in race, culture, body, or sense, and extending that concept beyond people to nature and the global environment will lead to "sustainability." Well-being is about acknowledging the intrinsic value of each person and respecting different ways of living. The keys to achieving this goal are mutual understanding and trust. In my case, I'm attempting to enhance mutual understanding and trust among people through tactile sensation.

Well-being tends to be thought of as a state of being or a target to aim for; however, it may be easier to think of it as a way of being or an attitude when performing an action. For example, instead of "playing soccer while aiming for well-being," it would be "playing soccer in a manner of well-being." When we think of well-being as a state, we imagine that well-being exists externally and try to get closer to it; however, when we think of well-being as a way of being, we can think it exists within ourselves and our teammates. The key to achieving the "good way of being" for each person within the team is to understand what each person values and how it can be fulfilled. In the example of soccer, it means the style of play and the favorite moves of each player. However, even if people get together, there is a high barrier to suddenly start talking about what each person values without any prompting, so they need something to trigger the conversation. Therefore, I created "Cards for Collective Well-being" as a tool to help people become aware of and understand their well-being as well as that of other people around them (**Fig. 4**).

—How did the cards come about?

Based on a survey conducted involving approximately 1300 university students asking about what is important to their well-being, the cards describe factors that make people feel a sense of well-being. The factors are divided into four categories: "I," "We," "Society," and "Universe." Factors such as "Hyperfocus," for example, are in category "I," factors such as "Close relationships" are in category "We," factors such as "Contributing to society" are in category "Society," and factors such as "Peace" are in category "Universe." The 2021 edition consists of 27 cards, and the 2022 edition consists of 32 cards [5].

We held workshops using these cards at several elementary, junior-high, and high schools. In the workshop, in groups of about three to five students, each student first chooses three cards that are important to their well-being and explains the reasons behind their choice to the other students in the group. The team then chooses three cards they would like to value in their place of learning, such as schools. This process can be described as a process of co-creating our well-being in the place of learning while taking into account the personal well-being of each member of the group that was initially shared. The students commented that they have now "realized what affects personal well-being," "have a deeper understanding of what my friends consider happy," and "understand that some aspects of what I consider to be well-being are the same as those of my friends and others are different." We also devised materials and other resources that help students discuss well-being at home after experiencing this workshop.

In this manner, an approach shifting from "I" to "We" is important regarding learning as well as workstyle transformation, team building, and local community development. To deepen this approach, since 2019, we have been conducting joint research on one's idea of self with Professor Yasuo Deguchi, a philosopher at Kyoto University. Professor Deguchi advocates the "Self-as-We" concept, namely, a holistic idea of self that is based on East Asian traditions of thought. We developed a scale to examine how closely the personality traits of individuals follow this Self-as-We view [6] and showed that people with such personality traits were less prone to depression during the COVID-19 pandemic [7]. We are currently developing an evaluation index regarding the state of a team, namely, how much members of a given team feel the "Self-as-We," and examining the correlation between the state of a team and the collective well-being.

Translating research into practice

—Research on haptics and well-being will be of great academic and social value.

My current goal is to use haptic methodologies and technologies to support human connections and the sense of "We" [8], which are important factors contributing to well-being. This endeavor will lead to the establishment of a new field of research. In fact, I'm participating as a co-investigator in the project called "Foundation of Digital-Embodied Economics" [9], which was selected as a Grant-in-Aid for Transformative Research Areas (B) of the Ministry of Education, Culture, Sports, Science and Technology in 2021. I'm also participating in discussions at the International Telecommunication Union (ITU) Telecommunication Standardization Sector, a standardization agency for transmission methods of information to ensure the distribution of tactile information.

Fig. 4. A set of "Cards for Collective Well-being."

Another area of technology that I'm currently interested in is blockchain, which stores data in a decentralized manner. Blockchain is the foundation for enabling new forms of organizations that differ from the centralized organizations of the past. In such a decentralized autonomous organization, participants make decisions by voting with governance tokens issued by the organization on the blockchain, on which there is no specific administrator. By adopting this blockchain structure, I hope to create an organization in which like-minded people can be involved in a flexible and responsible manner.

—Finally, please tell us what it means to you to be a researcher.

Researchers are like mirrors that reflect the future. In information and communication technology, which is directly related to people's daily lives, we conduct research by imagining what society will be like a little further down the road as the technology changes rapidly. Although it is a little misleading to say, I feel that selecting which research area we should work on for an unpredictable future is similar to how venture capitalists decide which businesses to invest in.

I also believe that if one wants to make a difference in society, it is important to think of society as a river and choose where to place stones to change its flow. By working as a "We" with one's colleagues, one can respond with vigor to various changes in the situation. What "I" can do alone is small, and even if several "I"s gather together, things will end up being not quite right. I think it would be better if each person could be involved in society while keeping in mind a sense of "We."

Another important factor is timing. Since the value of things is determined by the receiver, it is necessary to communicate such value in accordance with the state and readiness of the receiver. For example, if you express your opinion to someone but they are not ready to hear it, they will not accept it. If you water your flowers at the wrong time, they will rot. In that sense, research is similar, that is, it is important for researchers to conduct research that society thinks it is necessary and have a good sense of timing.

References

[1] "Touch and Connect Sports Lab workshop," Furue, Vol. 35, July 2021 (in Japanese). http://furue.ilab.ntt.co.jp/book/202107/contents2.html

[2] J. Watanabe, K. Ai, T. Yoshida, K. Kuwano, K. Komazaki, and A. Hayashi, "Air-Transmit Haptic Communication for Sharing Excitement in Sports Viewing: 'WOW BALL,'" Transactions of the Virtual Reality Society of Japan, Vol. 25, No. 4, pp. 311–314, 2020 (in Japanese).

[3] A. Hayashi, A. Ito, and J. Watanabe, "Sports Social View—A Generative Sports Viewing Method, in Which Essences of Sports Are Physically Translated and Shared with Visually Impaired People," Transactions of the Virtual Reality Society of Japan, Vol. 25, No. 3, pp. 216–227, 2020 (in Japanese).

[4] A. Ito, J. Watanabe, and A. Hayashi, "Sports Guide without Sight," Shobunsha, Tokyo, Japan, 2020 (in Japanese).

[5] Website of "Wellbeing through Haptic Sense" (in Japanese), https://socialwellbeing.ilab.ntt.co.jp/tool_measure_wellbeingcard.html

[6] J. Watanabe, A. Murata, C. Takayama, M. Nakatani, and Y. Deguchi, "Evaluation of Self-as-We—Development of Self-as-We Measures," PROSPECTUS, Vol. 20, pp. 1–14, 2020 (in Japanese).

[7] A. Murata, J. Watanabe, and Y. Deguchi, "Relationship between Trends in Depression under the COVID-19 Pandemic and 'Self-as-We,'" PROSPECTUS, Vol. 20, pp. 15–33, 2020 (in Japanese).

[8] J. Watanabe and D. Chen (compilers and editors), H. Ando, K. Sakakura, and A. Murata (editors), "Towards Collective Wellbeing—Its Concept, Practice, and Technology," BNN Inc., 2020 (in Japanese).

[9] Website of Foundations of Digital Embodied Economics (in Japanese), https://embodiedecon.digital/

■ **Interviewee profile**

Junji Watanabe received a Ph.D. in information science and technology from the University of Tokyo in 2005 and joined NTT in 2013. His research is focused on haptic communication and collective well-being.

Software-based Optical Access Network Virtualization Technology: Facilitating New-service Prototype Provisioning

Takahiro Suzuki
Distinguished Researcher, NTT Access Network Service Systems Laboratories

Abstract

Optical access networks are the networks connecting users to telecom central offices, and current trends of increasing traffic, services for mobile devices, and low-latency or high-reliability edge computing are creating a range of new requirements that these networks must be able to accommodate. We have asked Distinguished Researcher Takahiro Suzuki to talk with us about software-based optical access network virtualization technology that will be able to handle such requirements.

Keywords: optical access network, virtualization, software

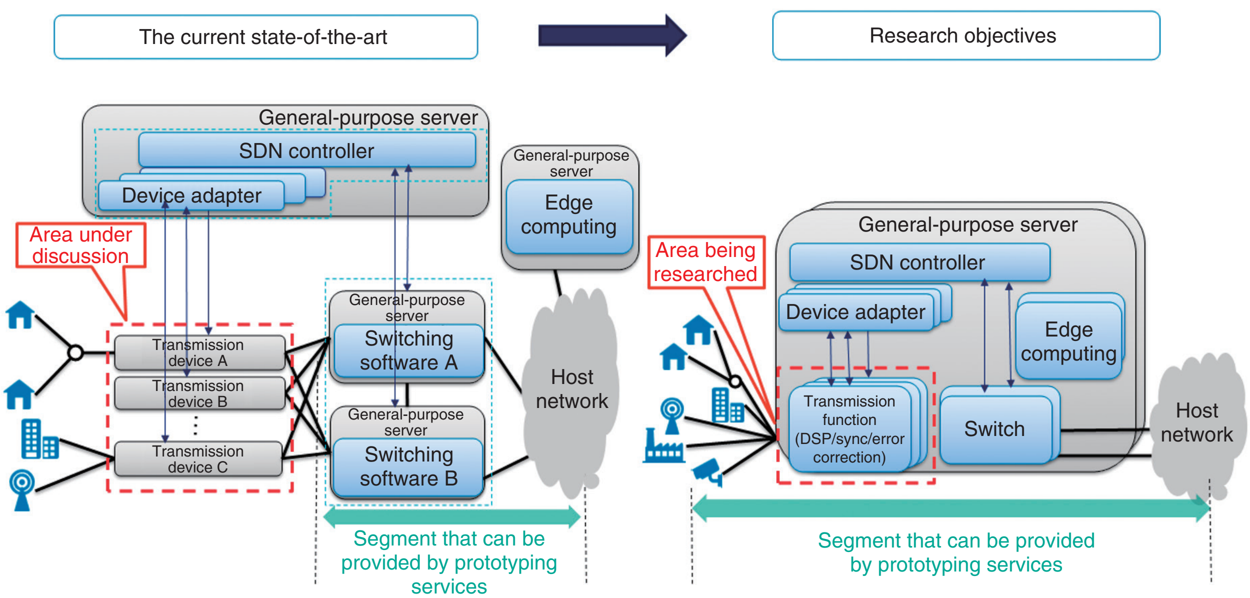
Software-based optical access network virtualization technology for implementing low-latency and high-reliability networks with flexibility

—Could you tell us about "software-based optical access network virtualization technology"?

Optical access network virtualization technology separates the control functions from the transmission functions in transmission equipment for optical access systems, and moving the control functions to a general-purpose server enables unified control and management of communications equipment. We are also researching use of software technology to implement functions of the transmission equipment itself, so that it can also be implemented on general-purpose servers.

Conventional optical access systems have incorporated network specifications, which have been different for each vendor, in dedicated hardware, so they could only be used for their original purpose. Such dedicated hardware also required a huge amount of time and expense to complete the design and development of each system, so that when introducing a new service, the transmission equipment and control systems had to be developed at large scale, and it was necessary to carefully consider in detail whether the service would be able to generate the revenue required when it was introduced. Software-based optical access network virtualization technology provides a solution for such difficulties.

By virtualizing optical access networks with software, functions in the optical access system transmission



SDN: software defined network

Fig. 1. Software-based optical access network virtualization technology overview.

equipment itself can be replaced easily, so that various requirements for access can be accommodated. With software, development cycles can also be shortened, so that even if there is a problem after the system is installed, it can be corrected repeatedly with updates. Thus, this research enables services to be prototyped and verified easily, so that new services can be created one after another.

—What are some methods used for software-based optical access network virtualization technology?

Some specific transmission functions required for software-based optical access network virtualization technology include frame synchronization, error-correction coding, a scrambler function, and digital signal processing (DSP), and we implement these functions on a general-purpose server. Software cannot achieve the performance possible with dedicated hardware, so we utilize general-purpose accelerator hardware to increase performance. We are also working on other ways to resolve this issue, such as devising technology to transmit the input signals to the processing devices at high speed, making radical changes to the transmission algorithms themselves to create low-computation algorithms that achieve the same performance, and creating high-speed processing technologies that coordinate between central processing units and accelerators efficiently.

Figure 1 shows our initiatives in software-based optical access network virtualization and what we ultimately aim to achieve. Currently, my research focuses on a network prototyping service, and with the latest results, we can implement the control functions with SDN Enabled Broadband Access (SEBA), which is open-source software from the Open Networking Foundation (ONF). In the future, by also implementing transmission system components in software on general-purpose servers, we aim to expand the domain over which "free prototyping" can be done, from only the host network as was the case earlier, to enable free network prototyping across both the user and host networks. This will allow testing of various types of services.

Future telecommunications services implemented using software-based optical access network virtualization technology

—Could you tell us about your future expectations for use of software-based optical access network virtualization technology?

Currently, research is advancing on DSP for future high-speed access systems and this is still done with a basic assumption that it will be implemented in

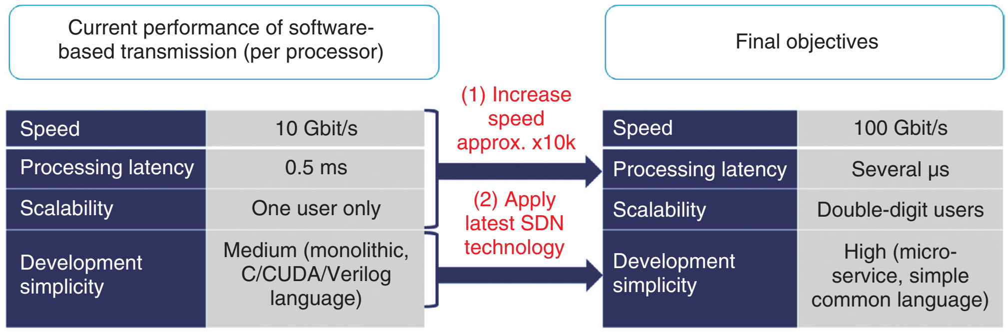
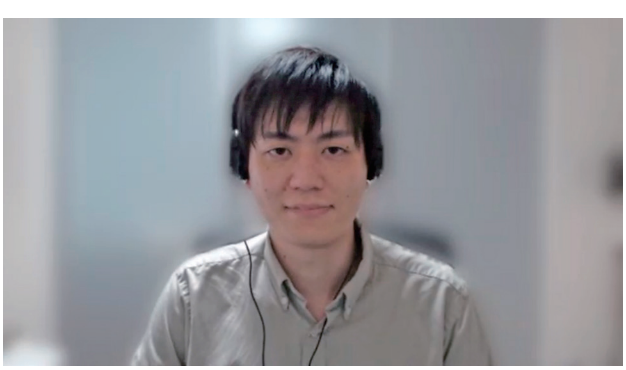


Fig. 2. Software-based optical access network virtualization technology initiatives and current final objectives.

dedicated hardware, but we hope to implement it in software. In fact, several of the functions comprising this DSP can be processed using a graphics processing unit, as is already being done in areas such as artificial intelligence, achieving results with throughput of 10 Gbit/s. However, when considering expandability and how many users can be connected, we are still limited to a single user, which is inadequate, and it is still not "developer friendly" enough to enable prototyping of various services.

Figure 2 shows the current state and final target for software-based optical access network virtualization technology. Our next target is to build a 100 Gbit/s-class system to handle the constantly increasing user traffic, which is an increase of approximately 10,000 times current speeds. Other targets are to reduce processing latency to a few microseconds, and to increase the number of users that can be accommodated into the double-digits. To implement network prototyping that is very easy, we will also need to create micro-services, reducing the size of functional components that can be developed, and use a common language to simplify development.



Then, when we have achieved this, it will be possible, for example, to implement most of the systems for fiber-to-the-home communication, fifth-generation and sixth-generation mobile communications systems using the same components, and provide these services at lower cost. Currently, users are being provided with the access segment for ordinary Internet connection services, but in the future, it will be possible to provide segments to access new networks. As a concrete example, by enabling simple access network prototyping, we could create a service that extends the access segment to provide industrial networks requiring low latency and high reliability, for controlling robots and other industrial equipment in a factory.

With general-purpose servers in the optical access system, edge computing that offloads most of the user processing onto the telecom central office will be possible, eliminating the need for the user to have processing equipment at all. To play a game, for example, a game console or personal computer would no longer be needed, and games could be played by just connecting a monitor and controller to the communications device. By having all processing equipment at the network or service provider in this way, users would no longer need to manage these various devices themselves.

—Do you have a message for other researchers and students?

I currently belong to NTT Access Network Service Systems Laboratories, but I am often in situations where I can see that NTT has international influence extending beyond Japan. My past experience includes being stationed for about one year at ONF in Silicon Valley in the USA, developing open software for optical access system virtualization with the ONF organization and several overseas carriers. There were many functions that could not be developed because the developers objected that they would add complexity to the software specifications, but when I wanted to add functionality as an NTT requirement, they had confidence in NTT and were willing to expand the scope of the software, based on the fact that NTT considered it important. This helped the research to advance smoothly. It is clear that NTT has name value that is valuable in advancing international research, and that NTT has a very good environment for telecommunications research in Japan.

For researchers and students that do not have much experience but want to work hard in research, in the future, it will be important to focus on work that produces published papers and other results. This is because it is very satisfying to gain knowledge through the process of studying a research theme, performing experiments, publishing a paper and being peer reviewed. To get frustrated part-way through without reaching the phase of publishing a paper is not much different than not producing anything at all. It was a very long road before a paper I had written was finally published, but having my paper published in a well-known journal was a joy like nothing else, and gave me confidence that I had achieved something as a researcher. So, I would encourage everyone to keep going till you have produced your results in a paper, and not to be satisfied with stopping at an earlier phase. It is through this process that entirely new technologies are produced and technologies that are useful around the world are completed.

■ Interviewee profile

Takahiro Suzuki completed a Ph.D. in engineering at Waseda University in 2017. He joined NTT in 2014, with NTT Access Network Service Systems Laboratories. He became a distinguished researcher for the Laboratories in 2022. He is engaged in research on software-based optical-access network virtualization. He received a Best Paper Award from the Transmission, Access, and Optical Systems Technical Committee of the GLOBECOM international conference.

Toward the Remote World in Which Life Is Many Times More Enjoyable—From "Remote by Necessity" to "Remote by Choice"

Yushi Aono

Abstract

The novel coronavirus (COVID-19) pandemic has forced people to change their lifestyles and become *remote*. To enable people to choose a remote lifestyle or combine remote and real-world lifestyles so as to enrich their lives, NTT Human Informatics Laboratories is aiming to enable the Remote World unique to the post-pandemic era in two ways: (i) identifying and analyzing issues specific to new lifestyles from a wide range of perspectives including not only technology but also social science and humanities and (ii) promoting research and development regarding such issues.

Keywords: Remote World, human augmentation, harmonic reproduction

1. A new lifestyle arrives: Remote World

The novel coronavirus (COVID-19) pandemic has drastically changed people's lives and ways of thinking. One such change is the adoption of a remote lifestyle as epitomized by remote work, remote education, and remote sightseeing. Entering the year 2022, Europe and the US quickly regained their pre-pandemic lifestyles as real, face-to-face, offline activities become the norm again, which raises the question: Is remote lifestyle no longer necessary?

Many people were forced to choose a remote lifestyle during the pandemic. However, after such experience, people concluded that what can be done remotely will be done remotely and enhance the real world and giving up is the only possibility if the only option is to participate in an activity in the real world; however, it is now possible to participate remotely. In other words, many people realized that by exploiting being remote, they can aim for a more fulfilling life, and by combining both real-world and remote lifestyles, they can increase opportunities for new experiences. We therefore predict that an era will come when people will actively and wisely choose with the mindset that some jobs or events are better in the real world, some are better in the remote world, and some are better in a combination of the real and remote worlds. When that era arrives, what will be required of the Remote World and the technologies that support them? We believe it is important to provide a remote lifestyle that offers unique value and user experiences that make people think that being remote has its unique advantages (**Fig. 1**).

When COVID-19 spread quickly throughout the world and had to be controlled as quickly as possible, we were suddenly forced to live a remote-centered life, and we were surprised and frustrated by the difference between the remote world and real world. The main requirement of technology was to resolve these issues and create a remote world with little difference from the real world and no sense of discomfort. At NTT Human Informatics Laboratories, we were among the first to tackle these issues, and in a very short period, we created a number of technologies and made achievements that have greatly contributed to the success of the international sports

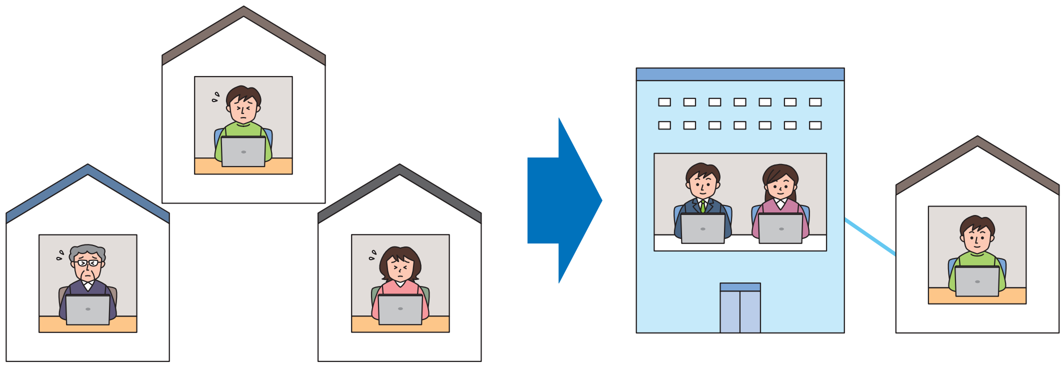


Fig. 1. Provide value and user experience unique to remote lifestyle.

Photo 1. Sailing spectating by using Kirari!.

event held in Tokyo in 2021 by overcoming the challenging situation of holding the event behind closed doors and supporting spectators remotely. These achievements include providing a highly realistic sailing experience by using Kirari! (ultra-realistic communication technology) (**Photo 1**) and support for spectators remotely watching a marathon by connecting Sapporo and Tokyo using ultra-low-latency communication technology (**Photo 2**).

As of August 2022, the status of global COVID-19 infections, including Japan, has been steadily declining; however, it is expected to take a considerable amount of time before the disease is completely controllable like influenza. However, without a doubt, humankind will overcome this hardship and regain its pre-pandemic life. In fact, people's desire to return to pre-pandemic life has become strong, especially in Europe and the United States. Under this circumstance, NTT Human Informatics Laboratories is anticipating a world in which this unprecedented situation comes to an end and people choose to be remote, as the occasion demands, to enrich their lives. We will research and develop technologies to mitigate dissatisfaction and discomfort in a society that is forced to be *remote by necessity* as well as technologies that enable people to be *remote by choice* i.e., choose a remote lifestyle or combine remote and real-world lifestyles.

Photo 2.   Remote support using ultra-low-latency communication technology.

2. Scenes in life where being remote is required

What situations can be envisioned where the remote lifestyle can be applied? We envision the following situations.

The first situation that comes to mind is work. The NTT Group is actively promoting remote work, and as an organization setting remote standards, we are further shifting to a remote-work-centered work style by which we basically use the home as the work place. Although it is relatively easy to shift to remote work for work using office automation equipment such as personal computers, it is not so easy to shift to remote work for work that involves both human and physical interactions. Equipment and facility maintenance and medical care are examples of tasks that are difficult to shift to remote work. If such tasks can be performed remotely, it would be possible to solve certain social issues such as (i) reducing energy consumed by cooling datacenters by locating them in cold regions (like the polar regions) and (ii) compensating for the shortage of nursing resources by allowing former registered nurses (who have left the nursing profession to raise children or other reasons) to perform nursing tasks remotely for short periods.

The next situation to consider is education. During the COVID-19 pandemic, many places of learning were operated remotely, and problems such as a lack of communication between teachers and students and among students became apparent. However, it also became apparent that it is possible to have lessons even if the teacher and students are separated by considerable distance. For example, it is possible to receive instruction from an admired teacher regardless of distance and national borders. In such a remote-instruction situation, it is problematic when it is necessary to teach how to move the body, such as when playing a musical instrument or playing sports, because it is difficult to provide direct guidance using the current means of video-audio communication. To solve these problems and proactively use remote education and reap its benefits, it is essential to communicate—even remotely—the use of the body and body language.

We also believe it is important to enhance leisure time, especially entertainment and sports. During the pandemic, every opportunity to experience the arts, such as concerts and plays, became remote; therefore, people have become dissatisfied over the loss of the *live* feeling and excitement compared with seeing the arts at venues and theaters. Such experiences, however, have shown that remote viewing is much easier to participate in because it is not restricted in terms of distance from the venue or seat availability. Some people like the remote viewing experience, which cannot be recreated in real life. Therefore, we believe

Fig. 2.   Scenes of life enhanced by remote lifestyle.

that if performers and many spectators in remote locations can feel a sense of unity and excitement, and if this feeling can be created in a personal environment, it will further enrich leisure and hobby time. The same can be said about sports. If we can work up a sweat with like-minded people even when we are far from each other, we will have more opportunities to be physically active, and we can make ourselves and society as a whole healthier (**Fig. 2**).

3. Technologies supporting the Remote World

The Feature Articles in this issue introduce four technologies from among these various remote application scenarios: (i) *human augmentation* for transmitting motor skills and embodied knowledge of experts directly to people at a distance [1], (ii) *telepresence technology* for supporting remote interpersonal and interactive tasks [2], and (iii) *harmonic reproduction technology* [3] and (iv) *emotional-perception-control technology* [4] for creating a sense of unity and excitement between the venue audience and remote audience, and among remote audience members.

As mentioned above, the world will eventually overcome this pandemic, even if it takes time. When that happens, we will not simply return to pre-pandemic life; instead, we will seek and obtain a new lifestyle—perhaps one that uses the knowledge we have gained over the past few years and one in which everyone can enjoy life many times more than before—by using real or remote styles or by combining the two. In collaboration with our partners, NTT Human Informatics Laboratories is aiming to achieve the Remote World unique to the post-pandemic era in two ways: (i) identifying and analyzing issues specific to new lifestyles from a wide range of perspectives including technology as well as social science and humanities and (ii) promoting research and development regarding such issues.

References

[1] T. Inoue and Y. Koike, “Efforts Concerning Enhancement of Human Ability to Achieve the Remote World,” NTT Technical Review, Vol. 20, No. 11, pp. 21–26, Nov. 2022.
https://ntt-review.jp/archive/ntttechnical.php?contents=ntr202211fa2.html

[2] S. Kondo, D. Sato, M. Goto, M. Takagi, and N. Matsumura, “Telepresence Technology to Facilitate Remote Working with Human Physicality,” NTT Technical Review, Vol. 20, No. 11, pp. 27–32, Nov. 2022.
https://ntt-review.jp/archive/ntttechnical.php?contents=ntr202211fa3.html

[3] T. Kurozumi, K. Hasegawa, E. Matsumoto, T. Eura, and S. Fukatsu, “Harmonic Reproduction Technology between a Real Venue and Remote Audiences,” NTT Technical Review, Vol. 20, No. 11, pp. 33–39, Nov. 2022.
https://ntt-review.jp/archive/ntttechnical.php?contents=ntr202211fa4.html

[4] R. Mochizuki, M. Makiguchi, M. Kohjima, M. Yokoyama, and R. Yamamoto, “Emotional-perception-control Technology for Estimating and Leading Human Emotions,” NTT Technical Review, Vol. 20, No. 11, pp. 40–45, Nov. 2022.
https://ntt-review.jp/archive/ntttechnical.php?contents=ntr202211fa5.html

Yushi Aono
Executive Research Engineer, Supervisor, Laboratory Manager, Cybernetics Laboratory, NTT Human Informatics Laboratories.

He received a B.E. in biotechnology and M.E. and Ph.D. in control engineering from Osaka University in 1994, 1996, and 1999. Since joining NTT in 1999, he has been engaged in research on speech synthesis, speech recognition, and spoken dialogue systems. He is a member of the Institute of Electronics, Information and Communication Engineers and the Information Processing Society of Japan.

Efforts Concerning Enhancement of Human Ability to Achieve the Remote World

Teruhisa Inoue and Yukio Koike

Abstract

To achieve the Remote World for providing a unique remote-user experience, NTT Human Informatics Laboratories is developing technologies to enhance human abilities. We are developing two technologies that target motor skills and embodied knowledge (which are difficult to verbalize). The first, *motor-skill-transfer technology*, collects objectively observable information (such as surface electromyograms and electroencephalograms) and shares this information with people by directly transmitting it through electrical stimulation of muscles. The second, *embodied-knowledge technology*, captures subjective sensations that occur within a person and transmits and shares these sensations so that others can experience similar sensations.

Keywords: human augmentation, motor-skill-transfer technology, embodied-knowledge technology

1. Efforts to enhance human abilities via the Remote World

To achieve the Remote World, we believe that it is not just a matter of connecting real spaces remotely, i.e., in cyberspace, but a world in which the real and cyber aspects are fused to create a new user experience. Our research targets education, especially education and training involving physical exercise such as playing musical instruments and sports. We aim to create new user experiences in the Remote World by enhancing human abilities by transmitting and sharing skills and physical techniques—regardless of time and place—that have been directly transmitted and shared between people in the real world.

We introduce two technologies to combine the remote and real worlds. The first is *motor-skill-transfer technology*, which collects objectively observable information, such as surface electromyography and electroencephalography on the skills and physical techniques (tips, etc.) of professional athletes and craftsmen, and transmits and shares this information directly to people using electrical muscle stimulation (EMS). The second is *embodied-knowledge technology*, which captures the subjective sensations that occur within a person and transmits and shares these sensations so that others can experience similar sensations.

2. Initiatives concerning motor-skill-transfer technology

We are developing motor-skill-transfer technology to create new forms of education and training in the Remote World. We are trying to create a world in which people from all walks of life can teach each other, regardless of time or physical distance, and gain unprecedented teaching and learning experiences, especially in learning about movement, which is a difficult task, even face-to-face. The concept of this technology is to reproduce and transfer the movement of a skilled person, that is, to support movement by sensing the human body's motor control and intervening in it [1] (**Fig. 1**).

The relationship between motor control and motor support in regard to the human body is explained as follows. People move their bodies when motor instructions from the brain are transmitted to the

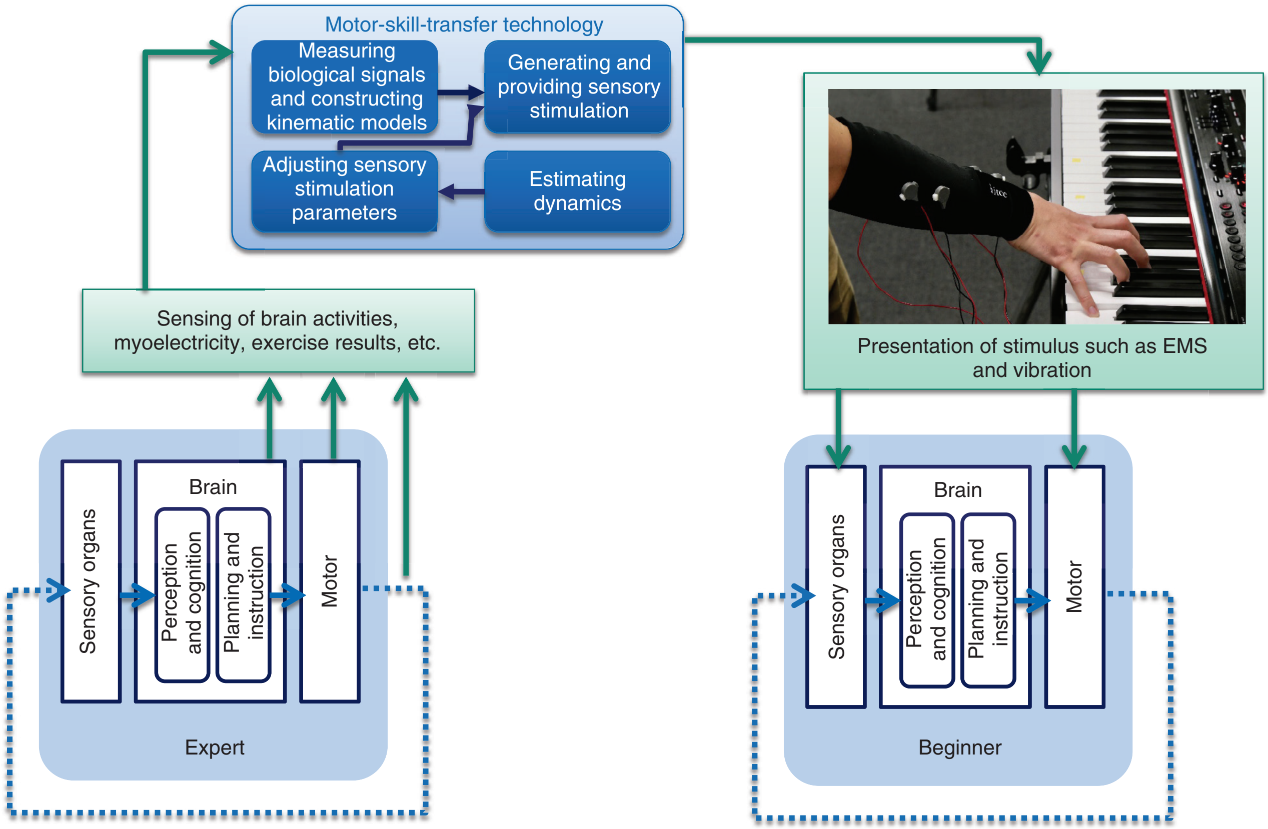


Fig. 1. Conceptual diagram of remote training by using motor-skill-transfer technology.

muscles and the muscles contract. The brain then perceives and recognizes the results of the movement as stimuli through sensory organs and repeats planning and giving new movement instructions. In contrast, the combination of sensing and feedback locations, as well as the specific sensing and feedback methods, changes the content of motion support. Motor-skill-transfer technology supports communication and control in the human body through analysis and electrical-stimulation feedback on the basis of the sensing of biological signals such as brain waves and muscle potentials. Regarding sensing and intervention regarding motor instructions from the brain to muscles and the resulting motor activity, we have begun investigating two topics: (i) measuring muscle activity (via electromyography), sensing of motor status, and motor support (intervention) using EMS based on that motor activity and (ii) sensing and intervention regarding perception in the brain in response to input from sensory organs. We have also started to focus on vision, somatosensation, and vestibular sensation, which are important in postural control; the basis of movement.

We now discuss new forms of education and training using motor-skill-transfer technology. Typically, in the education and training concerning exercise, the instructor or trainer and the trainee are present face-to-face in the same space, and the trainee is instructed through words and gestures. In contrast, the technology of transferring motor skills through biometric sensing and intervention has two key features: (i) enabling instruction regardless of time and physical distance and (ii) creating a more-effective space (Remote World) that goes beyond face-to-face education and training. As an example of training in the Remote World, we describe the remote instruction of a piano tremolo performance.

Beginners and experts use their arm muscles differently when playing a tremolo on the piano; that is, beginners focus on finger movement, while experts focus on wrist rotation. Focusing on this difference in muscle activity, we are developing a technology that uses EMS to directly *transfer* the way an expert uses their muscles to the muscles of a beginner. We

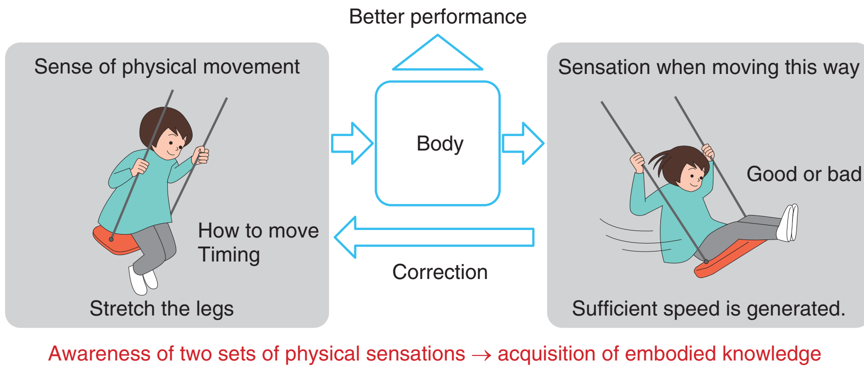


Fig. 2. Acquisition of embodied knowledge.

confirmed that the body can directly learn how skilled pianists efficiently move their bodies in a manner that enables them to play with less unnecessary effort in their forearms [2].

With this technology, we developed a conceptual system for training in the Remote World. A simple motion sensor is attached to the skilled pianist's forearm to measure the rotational motion of the wrist and transmit it to the system. An EMS device is connected to the arm of a beginner, and the EMS is applied to alternate rotations of the wrists as the skilled pianist moves. Experienced players can be taught through audio and video as well as directly experiencing—through the system—the playing techniques of expert pianists in a manner that creates a new teaching and learning experience for both instructor and student. We are focusing on the coordinated movement of multiple muscles during a performance and investigating a system that allows multiple muscles to move in a manner coordinated through EMS. This system will help beginners play smoothly and better.

By applying the various sensing and intervention technologies we are developing for our motor-skill-transfer technology, we will continue to create new forms of education and training in the Remote-World era for specific cases such as daily movements, sports, and musical-instrument performances.

3. Initiatives concerning embodied-knowledge technology

We are researching the extraction and sharing of *embodied knowledge* (i.e., skills) that cannot be acquired through verbal understanding to (i) clarify the mechanism of acquisition of embodied knowledge in sports and (ii) establish technology for supporting people in remote areas in acquiring embodied knowledge possessed by others.

The acquisition of embodied knowledge is considered the ability to (i) modify the way one moves one's body (muscles and bones) on the basis of the unique sensations that arise in one's body while performing a certain physical action and (ii) perform the action more skillfully and appropriately (**Fig. 2**). However, the unique sensation of the body is confined to the person who performed the physical action, and it is difficult to directly capture and communicate that subjective sensation to others. In this initiative, using windsurfing as a subject, our embodied-knowledge technology extracts and shares the unique experience of professional athletes during high-performance sports by capturing their physical activities and the behavior of the natural environment and tools (i.e., state information) and reproducing those factors to form a similar experience and acquire the embodied knowledge of others.

This technology extracts embodied knowledge in four steps: (1) sensing video and behavior data of actual athletes during competition; (2) conducting retrospective interviews with the obtained data about the intention and awareness of physical actions; (3) identifying situations in which unique physical sensations occur during high-performance activities; and (4) combining information on the state of the body at that time with media such as onomatopoeia and models that express the physical sensations.

We aim to improve extraction accuracy by evaluating whether the athletes actually experience and are aware of the extracted information during actual competitions and providing feedback (**Fig. 3**). We are improving the means of extracting more effective information to support the acquisition of embodied

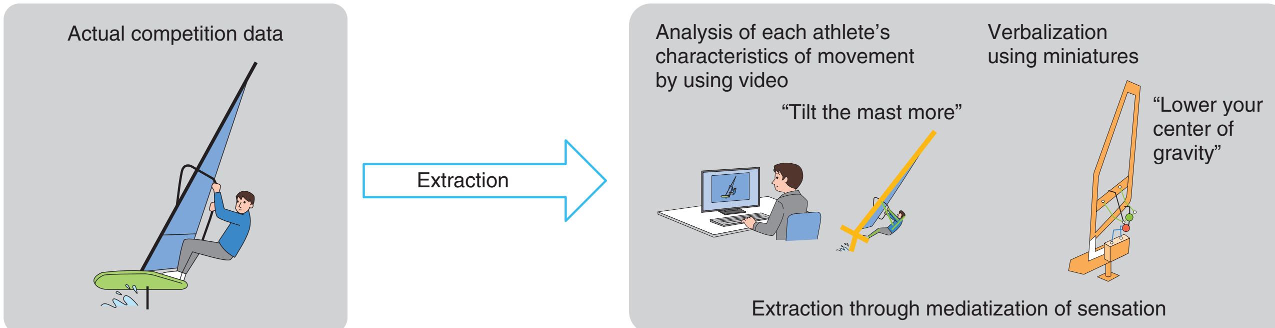


Fig. 3. Extraction of athlete-specific experiences.

knowledge by sharing the state information and media that form each unique experience among athletes and extracting common and different shared media by comparing them.

We have begun developing this technology by constructing a bodily-sensory-reproduction simulator to reproduce state information in a manner that reproduces the unique sensations that others have on the basis of the extracted information (**Fig. 4**). This simulator is used to form unique sensations that others have by replaying in tandem the use of windsurfing equipment by a professional windsurfer, wind (natural environment), and video the windsurfer watched. We believe that this simulator will make it possible to share subjective experiences, which are normally difficult to share and "closed" in the individual, as experiences for other people. Professional windsurfers are currently using this simulator to evaluate it and provide feedback. Through these efforts, we are striving on a daily basis to improve extraction and sharing technologies with the goal of acquiring embodied knowledge to achieve, for example, speeds of up to 60 km/h or more in future windsurfing competitions. We will continue to study this technology with an eye on expanding it to non-sports fields in which skills involving physical movements are used.

Fig. 4. Bodily-sensory-reproduction simulator for sharing the unique experience of an athlete.

4. Collaboration with NTT DOCOMO Human Augmentation Platform

We are promoting efforts to enhance human capabilities in the Remote World in collaboration with the Human Augmentation Platform®, which is a platform for enabling human augmentation that NTT DOCOMO has identified as one of the new values to be provided in the sixth-generation mobile communications system (6G) era as a means of extending human senses via networks [3, 4]. We believe that our goal of enhancing human capabilities through networks can be effectively reached by linking with NTT DOCOMO's network technology, which transmits and extends human senses through networks.

At docomo Open House'22 (January 17–19, 2022), for visualizing the state of collaboration with NTT DOCOMO's Human Augmentation Platform, we presented two exhibits. One was a demonstration of

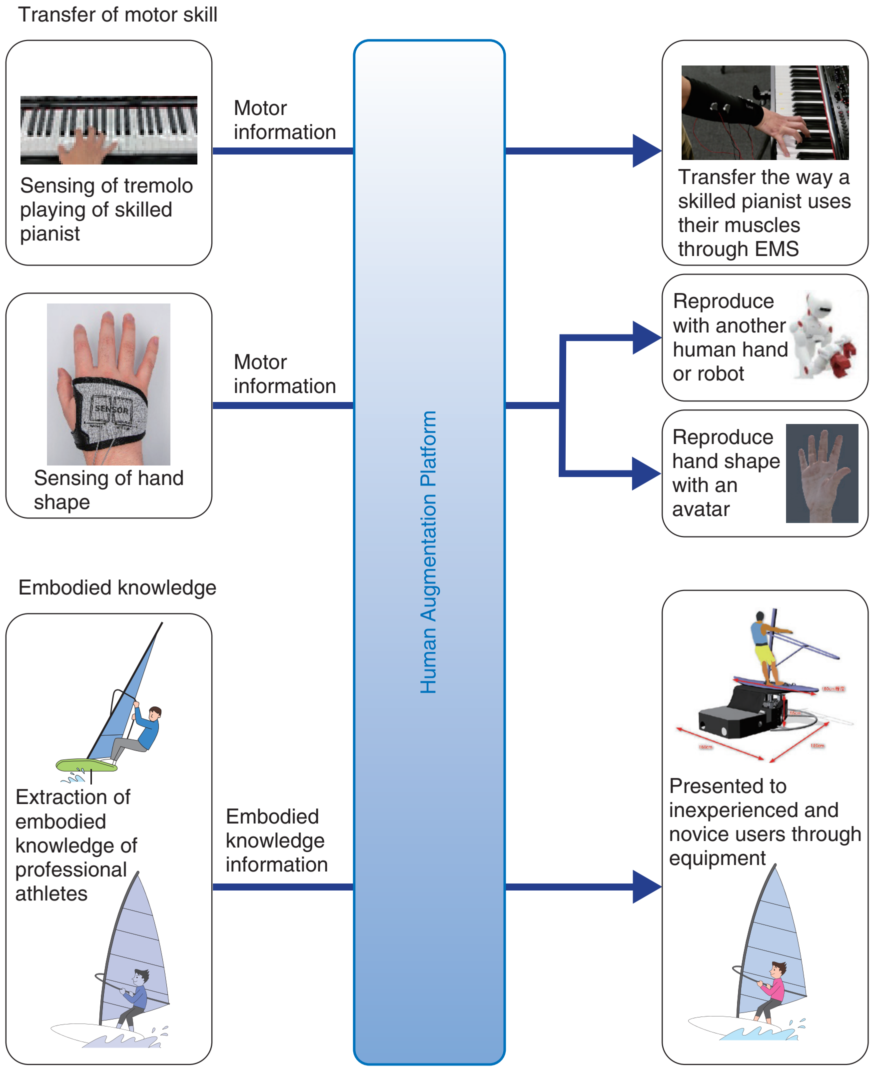


Fig. 5. Overview of the exhibition in collaboration with the Human Augmentation Platform (NTT DOCOMO).

our motor-skill-transfer technology, which involved active bioacoustic sensing [5] of hand and finger shapes to reproduce another human hand or robot, and a demonstration of a piano tremolo performance using the movements of an expert. The other exhibit on our embodied-knowledge technology showed our bodily-sensory reproduction simulator that allows beginners and amateurs to experience the skills of a professional windsurfer (**Fig. 5**). We will promote (i) collaboration with NTT DOCOMO to implement real-time augmentation of human ability by taking advantage of ultra-low latency and other features of 6G and (ii) investigation of more-valuable technologies for the Remote World in collaboration with more external partners.

5. Future developments

To achieve the Remote World, we will focus on the objective and subjective aspects of augmentation of human ability as a new user experience, especially in terms of transferring and sharing skills and physical techniques used during exercise regardless of time and place. We will combine these technologies and collaborate with external partners to provide even more valuable technologies.

References

[1] Y. Aono, H. Seshimo, N. Matsumura, Y. Koike, and M. Matsumura, "Cybernetics Technologies for Symbiosis between People and Machines," NTT Technical Review, Vol. 19, No. 12, pp. 36–41, Dec.

2021. https://doi.org/10.53829/ntr202112fa4

[2] A. Niijima, T. Takeda, K. Tanaka, R. Aoki, and Y. Koike, "Reducing Muscle Activity when Playing Tremolo by Using Electrical Muscle Stimulation to Learn Efficient Motor Skills," Proc. of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Vol. 5, No. 3, Article no. 123, Sept. 2021. https://doi.org/10.1145/3478110

[3] Press release issued by NTT DOCOMO on Jan. 17, 2022 (in Japanese). https://www.docomo.ne.jp/info/news_release/2022/01/17_00.html

[4] H. Ishikawa, "What Is the Human Augmentation Platform that Shares Human Movements and Senses," ITU Journal, Vol. 52, No. 7, pp. 16–18, July 2022 (in Japanese).

[5] Y. Kubo, "Fine-grained Hand-posture Recognition for Natural User-interface Technologies," NTT Technical Review, Vol. 19, No. 3, pp. 37–39, Mar. 2021.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr202103fa7.html

Teruhisa Inoue

Senior Research Engineer, Cybernetics Laboratory, NTT Human Informatics Laboratories.

He received a B.E. and M.E. from Yokohama National University, Kanagawa, in 2003 and 2009. He joined NTT EAST in 2009 and transferred to NTT in 2020. His research interests include human-computer interaction, human-robot interaction, and non-verbal behavior of communication.

Yukio Koike

Senior Research Engineer, Cybernetics Laboratory, NTT Human Informatics Laboratories.

He received an M.S. in physics from Nagoya University in 2004 and joined NTT the same year. His research interests include network technology for Internet-of-Things services and biological signal processing. He is a member of the Institute of Electronics, Information and Communication Engineers.

Telepresence Technology to Facilitate Remote Working with Human Physicality

Shigekuni Kondo, Daisuke Sato, Mitsuhiro Goto, Motohiro Takagi, and Narimune Matsumura

Abstract

Remote work allows people to free themselves from geographical and temporal constraints, and there are calls for it to be made available as an option for essential workers for whom it is not readily applicable. This article introduces our efforts to develop telepresence technology with which a person can use an artificial body or robot to sense the situation at a remote location and interact with other people and the environment in the same way as when working in the field. Thus, they can carry out their required tasks and engage in face-to-face communication without any discomfort or hindrance.

Keywords: remote working, telepresence, cybernetics

1. Introduction

In 2020, our values and common practices were changed by a new strain of coronavirus that rapidly spread throughout the world. One such change was the general acceptance of remote working, in which employees are no longer required to perform their jobs at a fixed location [1]. With the spread of remote working, it has become possible for people to carry out various tasks that would previously have been performed face-to-face without having to go to the actual location, thus releasing them from the geographical and temporal constraints associated with working in a set location. This provides people with various benefits, such as being able to freely choose where to live and perform other activities during the time they would have spent commuting to and from work. However, certain people, such as essential workers, can find it difficult to perform their jobs remotely, and even when remote working is possible, they can still experience problems such as communication failures and increased psychological burden when they have to accommodate a remote and decentralized work style.

NTT Human Informatics Laboratories is researching and developing telepresence technology to address the problems of remote working to provide the Remote World where people can have the option of working remotely when necessary in any line of work.

2. What is telepresence?

In some cases, remote work may not be an option because a person's job primarily involves on-site work that would be less efficient and productive when performed remotely or that requires face-to-face communication.

To facilitate remote working in jobs that require on-site work, it is necessary to have some way of actually performing this work on site by remote control (e.g., by machines or robots, i.e., artificial bodies). However, a problem with the current situation is that the content and timing of work operations performed at the remote site can differ from what was intended by the remote operator due to a lack of repeatability or delays in the response of the operation interface or artificial body. This can significantly affect the efficiency and productivity of the work. It should also be possible to accurately ascertain the local situation from a remote location, but this raises the issues of the lack of information for recognizing changes in the current environment or behavior of other people on site, and the tendency for people who are present on site to feel uncomfortable and anxious about working with someone who is in a remote location.

With the telepresence technology we are developing, we aim to make it possible for a person to use an artificial body to sense the situation at a remote location and interact with other people and the environment in the same way as when working on site. Therefore, they can carry out their required tasks and engage in face-to-face communication without any discomfort or hindrance. We are currently researching and developing technologies that will be needed to make this telepresence a reality: *zero-latency media technology*, which improves the operability at remote locations by predicting motions and reactions to compensate for latency and lack of operational information, *lifelike communication technology*, which enables information presentation methods that deliver the same quality of experience as being present on site, and *embodied-knowledge-understanding technology*, which clarifies and systematizes the connections between knowledge, experience, and physical behavior to predict changes in the remote operator or in the on-site environment and ascertain information from the on-site environment that is necessary for operating from the remote location. This article introduces our efforts in developing these technologies to enable telepresence.

3. Zero-latency media technology

Zero-latency media technology is used to create a system with which human operators can exert at least 100% of their work from a remote location by supporting the remote operation of robots. The main factors that hinder people's ability to work by remote control are latency, missing information, and differences between the human and robot body structures, each of which requires countermeasures. We are currently engaged in research and development focused on countermeasures against latency, which has a particularly large impact.

In remote control work, the operator controls a robot's position, force, and speed to appropriate values on the basis of feedback information such as camera images sent from a remote location. If there are any delays in this process, the feedback needed for remote control cannot be obtained at the required timing, reducing the accuracy with which the robot's position, force, and speed can be controlled remotely and resulting in issues such as operational errors and reduced work efficiency. For example, if an object's position shifts while trying to grasp it with a robot arm, the arm will fail to grasp it properly. If too much force is applied, the object could be destroyed, and if too little force is applied, it will not be possible to hold it well. When working on or painting an object, a shift in speed will change how the object is affected.

Zero-latency media technology addresses the latency issues of remote operations to improve work efficiency and reduce the burden on workers by implementing motion correction and cognition correction on the basis of an understanding of people and the environment (**Fig. 1**). Motion correction mitigates the effects of latency by compensating for the operational information sent by the operator before moving the robot. We are studying two approaches for implementing this compensation, predictive control and semi-autonomous control. In predictive control, we aim to reduce the latency of the entire system by predicting the operator's movements and controlling the robot on the basis of these predicted movements, making it operate in a preemptive manner. Semi-autonomous control is used to eliminate operational difficulties due to latency by estimating the operator's intentions and controlling the robot semi-autonomously on the basis of these estimations. Cognition correction mitigates the effects of latency by compensating the feedback information sent from the remote work environment before it is presented to the operator. We are studying two approaches for achieving this compensation, control-object compensation and work-object compensation. In control-object compensation, feedback information related to the robot (the controlled object) is compensated for with techniques such as augmented reality (AR), whereas in work-object compensation, compensation is applied to the feedback information about the object on which this remote work is performed.

To verify the effects of motion correction and cognition correction on mitigating latency effects, we remotely operated a simple pen-type robot and measured the time taken to accomplish the alignment task of touching points marked on a sheet of paper. The results of this experiment confirmed that the task took longer to accomplish as the delay time increased when no compensation was applied, but that compensation was effective at preventing delays from adversely affecting the task efficiency. We are

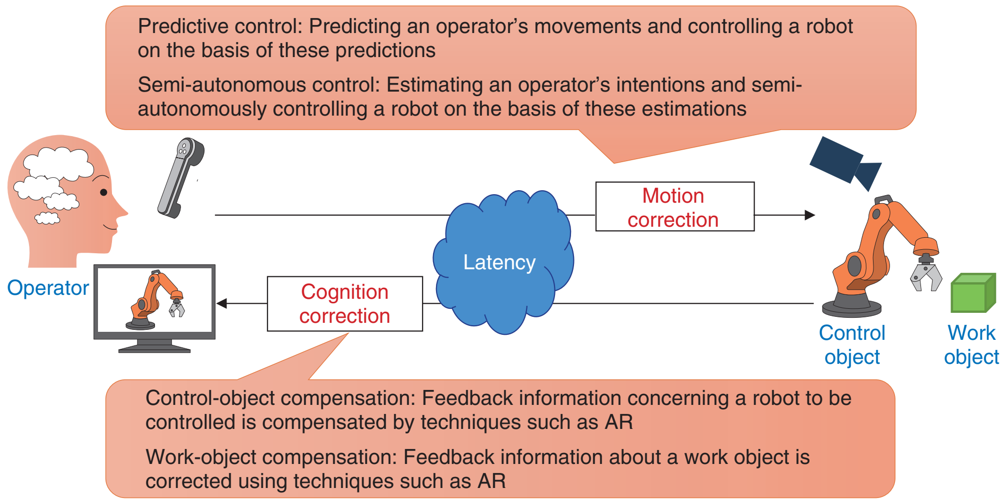


Fig. 1. Zero-latency media technology.

currently using remotely operated robot arms and humanoid robots to verify the delay-mitigation effects in more practical tasks and working to improve various correction techniques by deepening our understanding of people and the environment.

4. Lifelike communication technology

Lifelike communication technology focuses on how communication is experienced between people online. People have been presented with more opportunities to communicate online by exchanging video and audio with remote partners using tools such as web-conferencing applications. However, compared with direct face-to-face communication, this sort of online communication can tend to feel somewhat unsatisfactory. With lifelike communication, our aim is to reveal the factors that cause these dissatisfactions with online communication and provide the best possible online information that delivers the same quality of experience as face-to-face communication. In online communication, it is particularly difficult to sense the other person's presence and distance through the screen. The conventional approach to addressing these issues has been to provide high-quality and high-definition reproductions of sounds and images perceived in real space. With lifelike communication technology, instead of faithfully reproducing the audiovisual experience of communication in a real space, it is sufficient if the user ultimately feels satisfied and convinced even if the information reaching the eyes and ears is crude. Therefore, even the users of online communication can be provided with a face-to-face experience, as if the other person was close by (**Fig. 2**).

In 2021, based on this technical concept, we conducted a case study to enhance online technical exhibitions [2]. Currently, visitors to online exhibitions are only able to browse the content of exhibition pages they are interested in and do not get to feel the satisfaction of receiving an explanation from a helpful staff member, as is the case with face-to-face exhibitions. We, therefore, proposed an information-presentation method that uses stereophonic sound to present visitors with an audio commentary that corresponds to the part of the page they are viewing, thus providing them with the experience of receiving explanations directly from those people while allowing them to experience the distance and position of exhibition guides and attendants in the same way as at a face-to-face exhibition. We asked several visitors with experience in participating in online exhibitions and exhibition planning to evaluate the effectiveness of the proposed method. The results indicate satisfaction with and acceptance of the exhibition. We are considering applying this method to technical exhibitions as well as online customer service and art exhibitions.

5. Embodied-knowledge-understanding technology

As part of our effort to support telepresence, we introduce our research on the understanding of

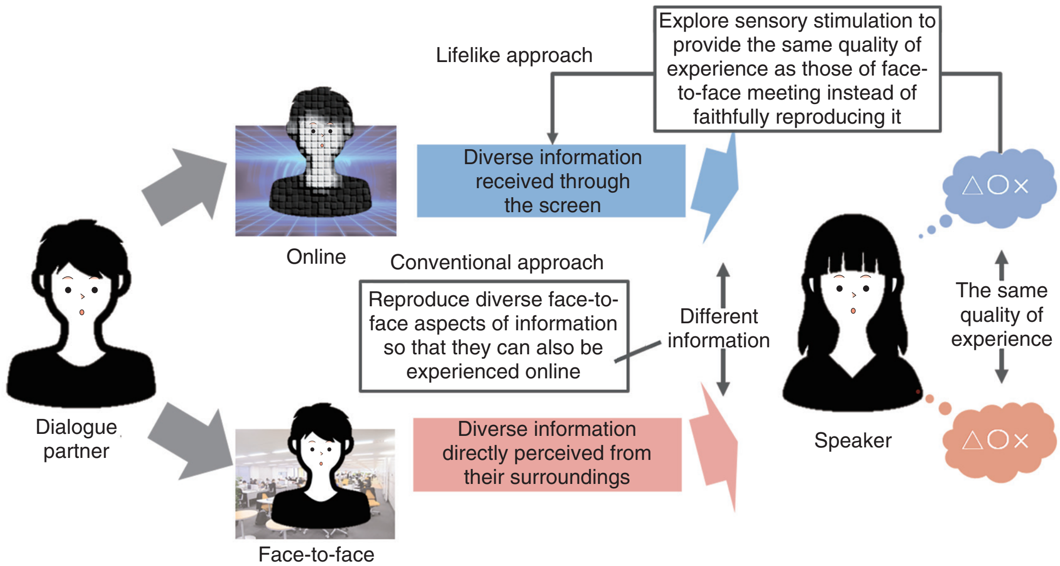


Fig. 2.   Lifelike communication technology.

embodied knowledge for systematization of physical behavior. For an essential worker to perform their tasks by means of telepresence, it is necessary to reproduce the essential worker's skills on an artificial body or robot. The use of robots has been considered for tasks that are relatively simple for humans, such as restocking the shelves in convenience stores and replacing local area network cables in datacenters. However, the complex jobs of professionals, such as nurses and caregivers, involve many skills with embodied-tacit knowledge that must be acquired in the field and are therefore difficult to perform with robots. These skills, which can be considered embodied knowledge, cannot be reproduced by artificial bodies or robots because there are no clear requirements for handling them digitally, making them difficult to perform remotely. To reproduce the skills required for tasks such as nursing and caregiving that are complicated even for humans, the challenge we face is how to express these skills digitally with artificial bodies and robots.

To digitally represent the skills demonstrated in high-level professional work, it is necessary to clarify and systematize the physical behaviors that are based on the specialist expertise that these skills require. For example, when dealing with patients in a hospital ward, the situational knowledge that the patients in this ward have asthma and the occupational knowledge that the breathing of asthma patients should always be checked with a stethoscope will result in the physical action of using a stethoscope to check their breathing. Thus, by systematically organizing the embodied knowledge that is needed in different environments and situations, and what physical behavior needs to be expressed based on this knowledge, it is possible to define how artificial bodies and robots should function in each environment and situation.

In our effort to systematize physical behavior, we are conducting research and development on how to systematically define the physical behavior information needed for the movement of artificial bodies and robots and how to acquire and express this information. Some of the physical behaviors that are targeted by physical-behavior information can be expressed as words in the form of instruction manuals, while others cannot and exist only as embodied-tacit knowledge within a particular field. There may also be discrepancies between the actions described in a manual and those that take place in the field. Our goal is to reproduce the skills of physical knowledge in artificial bodies and robots by grasping how these physical behaviors are performed in practice and combining them with expert knowledge so that they can be handled digitally. In 2021, we developed a technique for understanding the physical behavior of customers and employees in a convenience store business with finer granularity than could be achieved from manuals [3] and a dataset for recognizing the behavior of employees working in high places in the construction of telecommunications infrastructure from a first-person perspective and evaluating whether they were adhering to safety rules defined in manuals [4]. We will investigate physical-behavior-recognition technology for grasping the actual situation in the field as well as establishing methods for linking specialized knowledge and physical behavior.

6. Conclusion

We have introduced the zero-latency media technology, lifelike communication technology, and embodied-knowledge-understanding technology that we developing to achieve telepresence technology with which a person can use an artificial body to sense the situation at a remote location and interact with other people and the environment in the same way as when working on site. Therefore, they can carry out their required tasks and engage in face-to-face communication without any discomfort or hindrance. We will proceed with technical development and verification in specific fields such as nursing and equipment maintenance in which it is currently difficult to work remotely and expand the concept of telepresence technology and promote its research and development to liberate individuals from physical constraints and extend their capabilities.

References

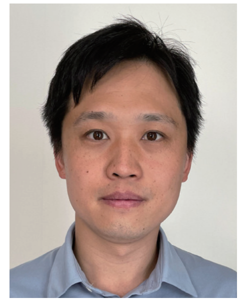
[1] The Ministry of Health, Labour and Welfare, "Whitepaper on the Labour Economy 2021," 2021 (in Japanese). https://www.mhlw.go.jp/wp/hakusyo/roudou/20/dl/20-1.pdf

[2] S. Uchida, M. Goto, and H. Seshimo, "A Study of Increasing Satisfaction during Online Communication by 3D Audio," IEICE Technical Report, Vol. 121, No. 423, MVE2021-88, pp. 255–260, 2022.

[3] "Human behavior understanding technology for improving CX/EX—Visualize human behavior and promote better CX/EX," Exhibits of NTT R&D Forum – Road to IOWN 2021, C27, 2021. https://www.rd.ntt/_assets/pdf/e/forum/2021/C27_e.pdf

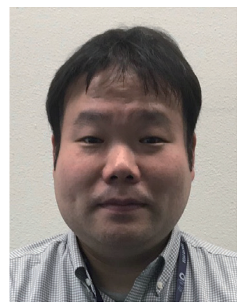
[4] R. Furuta, M. Takagi, Y. Sugano, and Y. Sato, "A Dataset for Work Instruction Deviation Detection from Field Work Videos," IEICE Technical Report, Vol. 121, No. 304, PRMU2021-30, pp. 37–42, 2021.

Shigekuni Kondo
Senior Research Engineer, Cybernetics Laboratory, NTT Human Informatics Laboratories.
He received a B.E. in applied physics and M.S. in integrated design engineering from Keio University, Kanagawa, in 2003 and 2005 and joined NTT in 2005. His research interests include software modularity, home networks, and user interfaces. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).



Motohiro Takagi
Researcher, Cybernetics Laboratory, NTT Human Informatics Laboratories.
He received a B.E., M.E., and Ph.D. from Keio University, Kanagawa, in 2009, 2011, and 2020. In 2011, he joined NTT. His research interests include human behavior understanding through machine learning, computer vision, and natural language processing.

Daisuke Sato
Senior Research Engineer, Cybernetics Laboratory, NTT Human Informatics Laboratories.
He received a B.E. and M.E. in mechanical engineering from Waseda University, Tokyo, in 2007 and 2009. He joined NTT in 2009 and has been engaged in the research and development of information retrieval, geographic information, and telerobotics services. He is a member of the Information Processing Society of Japan (IPSJ).



Narimune Matsumura
Senior Research Engineer, Cybernetics Laboratory, NTT Human Informatics Laboratories.
He received a B.E. and M.E. in information science from Kyushu Institute of Technology, Fukuoka, in 2002 and 2004 and joined NTT in 2004. His research interests include human-machine interaction, robotics, and digital transformation of nursing and care.

Mitsuhiro Goto
Senior Research Engineer, Cybernetics Laboratory, NTT Human Informatics Laboratories.
He received a B.E., M.E., and Ph.D. from the University of Electro-Communications, Tokyo, in 2008, 2010, and 2022. In 2010, he joined NTT. His research interests include human-computer interaction, human-robot interaction, and non-verbal behavior of communication.

Harmonic Reproduction Technology between a Real Venue and Remote Audiences

Takayuki Kurozumi, Keisuke Hasegawa, Eiichiro Matsumoto, Toshihiko Eura, and Shinji Fukatsu

Abstract

NTT Human Informatics Laboratories has been researching and developing technology to reproduce the appearance of an audience enjoying a live-streamed event online remotely from home (remote audience) at the venue where the event is being held (real venue) while harmonizing it with the situation at the real venue. At the 34th Mynavi TOKYO GIRLS COLLECTION 2022 SPRING/SUMMER held on March 21, 2022, a demonstration experiment was conducted to support the excitement of the event by using low-latency video communication and cross-modal sound search to reproduce pseudo cheers at the real venue for both real-venue and remote audience members who could not cheer due to the COVID-19 pandemic. This article introduces the activities of this demonstration experiment.

Keywords: two-way video communication, remote viewing, harmonic reproduction

1. The necessity of harmonic reproduction of video and sound

NTT has been engaged in research and development of interactive video communications that interconnect multiple remote-viewing environments and deliver high-definition video with low latency with the primary focus on delivering highly realistic video. In the Real-time Remote Cheering for Marathon Project [1], ultralow-latency communication technology with uncompressed transmission and low-latency media-processing technology were used to connect the marathon course in Sapporo, Hokkaido with the cheering venue in Tokyo in real time. This enabled spectators to send their support to the athletes from remote locations, creating a sense of presence similar to that of cheering along the course and a sense of unity between athletes and spectators, thus enabling a new way to watch the race.

To develop this initiative for home users, NTT began research on bidirectional video and audio communication between the venue of a live-streaming event (real venue) and home environment. However, there are inconveniences that arise when trying to achieve bidirectional, highly realistic video communication between the home environment and a real venue. For example, in situations such as web conferencing using video and sound between a remote-working home environment and the workplace, the background of the camera image from the home environment may show a room with a lived-in feel, or a family member's voice may be mixed in with the microphone audio, which can be awkward. In situations such as a live sporting or entertainment event, the audience members who participate remotely (remote audience) want to be present in the video to share the excitement with the audience at the real venue (venue audience) but would like to avoid information they do not want seen or heard from being distributed to the venue and other remote audience. Therefore, it is necessary to suppress unnecessary information and reproduce information that is desired to be reproduced at the real venue with a high sense of presence and harmony.

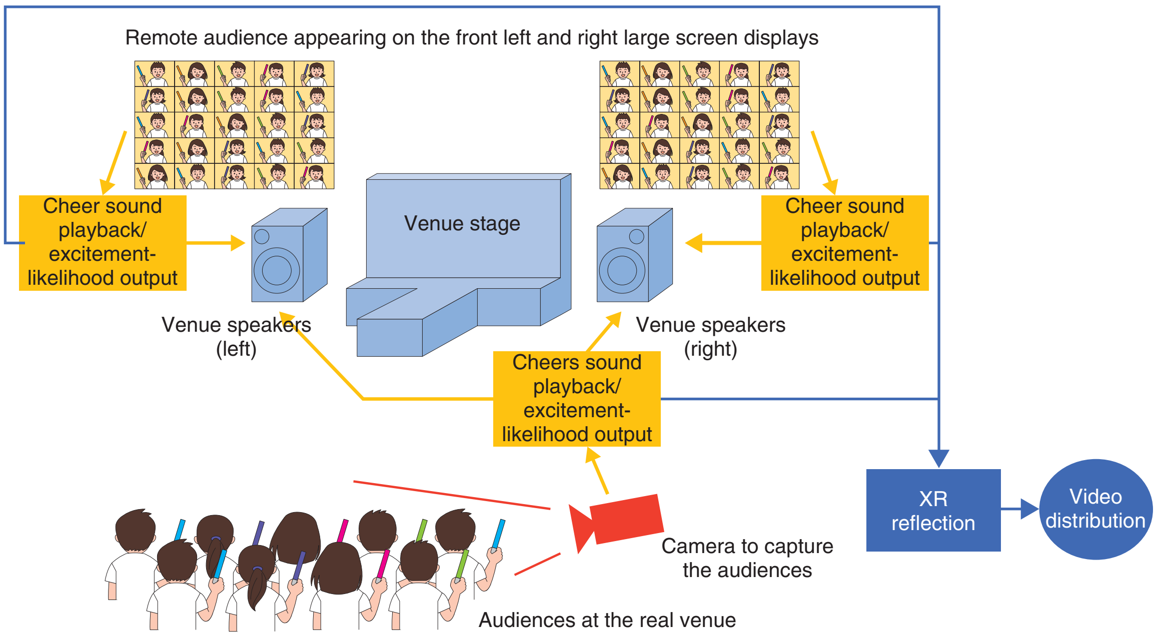


Fig. 1. Conceptual diagram of the demonstration experiment.

The real venue may also require harmonious reproduction. The COVID-19 pandemic has restricted people's activities, and even at live concerts, audiences are prohibited from cheering, even if they are wearing masks, to prevent infection. It is an uncomfortable situation for the audience at live music concerts since they cannot cheer for the popular songs that drew cheers before the pandemic, even though there are many people in attendance. It is also difficult for the performers to interact with the audience if they cannot hear the cheers, because it is difficult to understand the audience's reactions. Therefore, we studied the possibility of reproducing the harmony of images and sound so that the venue audience can feel the same excitement from the cheers as before the pandemic.

2. Joint experiment to enable two-way video communications for the home

IMAGICA EEX, NTT Communications, and NTT conducted a joint experiment to enable interactive high-resolution video viewing for home users at the 34th Mynavi TOKYO GIRLS COLLECTION 2022 SPRING/SUMMER [2] held on March 21, 2022. The experiment aimed to support the excitement of the venue audience, who were unable to cheer due to the COVID-19 pandemic, and create a sense of participation for the remote audience through interaction with the real venue. We constructed a system to reproduce cheer sounds in accordance with the excitement of the venue and remote audiences using low-latency video communication technology and cross-modal sound retrieval technology [3] and verified the reproduction of harmony.

Figure 1 shows a conceptual diagram of the entire experiment. Remote audience members remotely participated via NTT Communications' two-way low-latency communication systems (Smart vLive® [4] and ECLWebRTC SkyWay [5]) using a personal computer (PC) with a camera and appeared on the left and right large-screen displays in front of the stage at the real venue. The cheer sounds were estimated from the left and right images of the remote audience members, as described below, and the corresponding left and right speakers played the cheers in response to the excitement of the event. For the venue audience, the cheer sounds were estimated from the images taken by the camera aimed at the audience seats and played from the venue speakers. The audience could control the cheer sounds so that the cheers became louder when the audience shook their penlights faster and quieter when they shook them slower. The volume of excitement was reflected in the cheer sounds as well as in the extended reality (XR) expression of the live-streamed video, with IMAGICA EEX

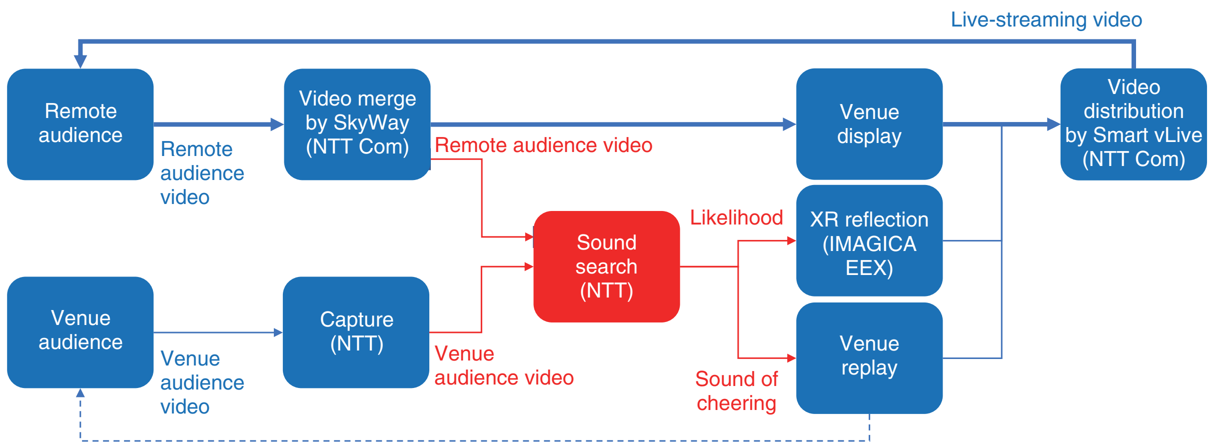


Fig. 2. System configuration and flow of processing and information.

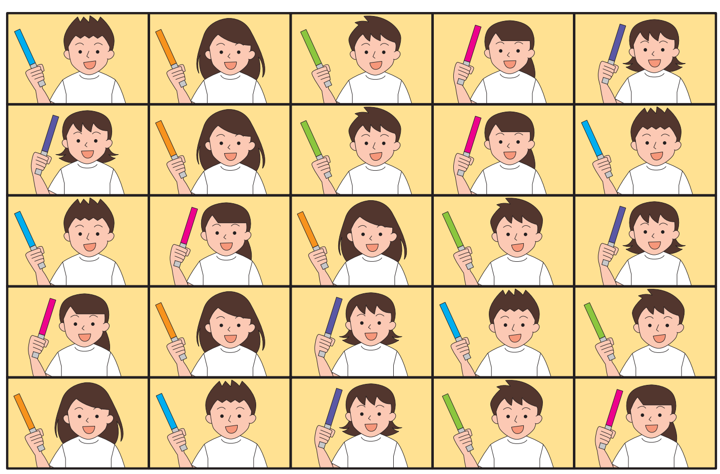


Fig. 3. Videos of remote audience waving penlights are arranged in a grid and aggregated.

creating a production in which the amount of light particles changes in accordance with the volume of excitement from the venue audience and remote audience members in the streamed video, allowing the remote audience members to enjoy the excitement of both audiences with sound and images.

2.1 System configuration

The overall system configuration is shown in **Fig. 2**. The experimental system consists of a function that lays out images of remote audience members in a tiled format, one that displays the images on a large display at the real venue, one that searches for and plays the cheer sounds from the tiled images and audience images at the real venue, one that expresses the search results in XR, and one that distributes images of the real venue reflecting these results.

The sound-retrieval system used in this project is based on machine learning technology using NTT Communication Science Laboratories' cross-modal sound retrieval technology [3] to estimate the cheer sounds from video images of audience members waving penlights. To map the images of the audience waving penlights to the cheer sounds, training data were prepared in advance by pairing the images of the venue and remote audiences waving penlights with the cheer sounds, and a model for estimating the sound from the images was trained. For the remote audience, we input aggregated video images laid out in a 5 × 5 tiled pattern, as shown in **Fig. 3**, and

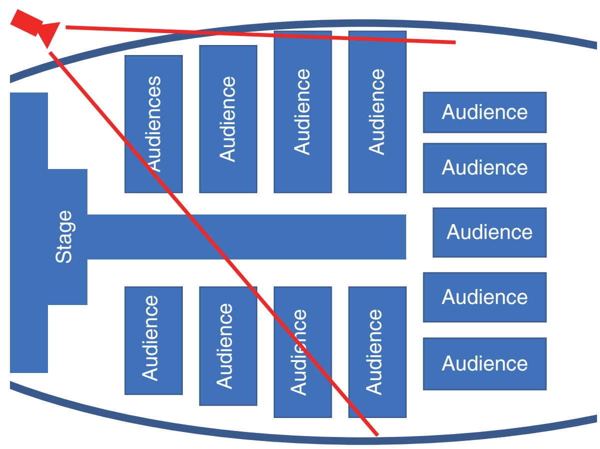


Fig. 4. Arrangement of cameras capturing the venue audience.

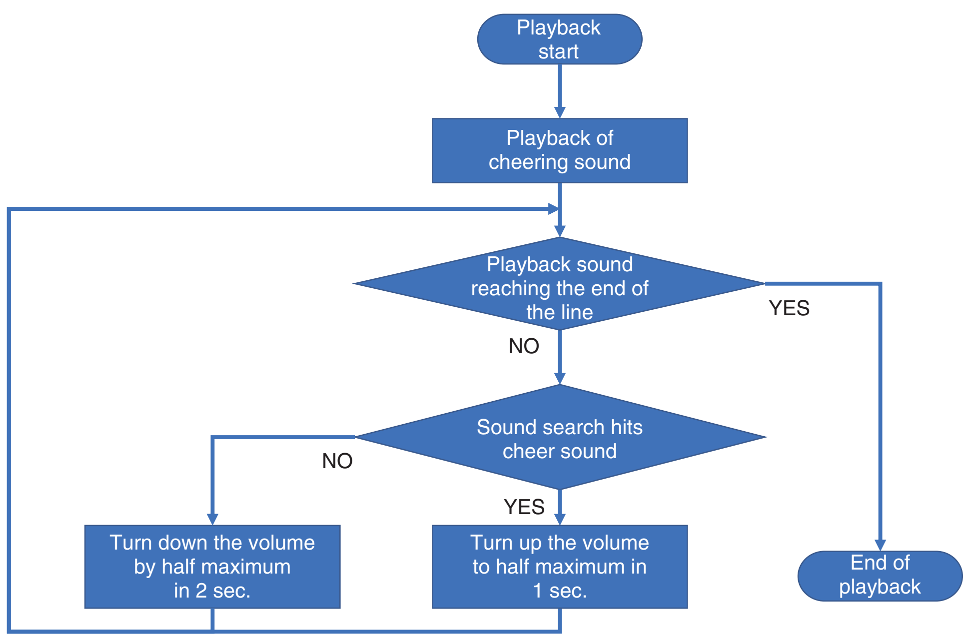


Fig. 5. Flowchart for volume determination.

retrieved and played back the corresponding cheer sounds on the basis of the penlight waving. For the venue audience, as shown in **Fig. 4**, a camera was set up in the venue to take a video of the audience, and the corresponding cheer sounds were retrieved and played back on the basis of the images of the audience waving penlights in the same way. The sound source for playback was a pre-recorded cheer sound.

In addition to searching for the cheers using the cross-modal sound retrieval technology, a method of determining the volume of the cheers using the flowchart shown in **Fig. 5** was implemented to smoothly change the volume of the cheers. An example of a hit point in the cheer-sound search and corresponding volume change is shown in **Fig. 6**. With this mechanism, it is possible to control the volume so that it increases when the penlights are continuously waved and decreases when they stop waving, thus enabling

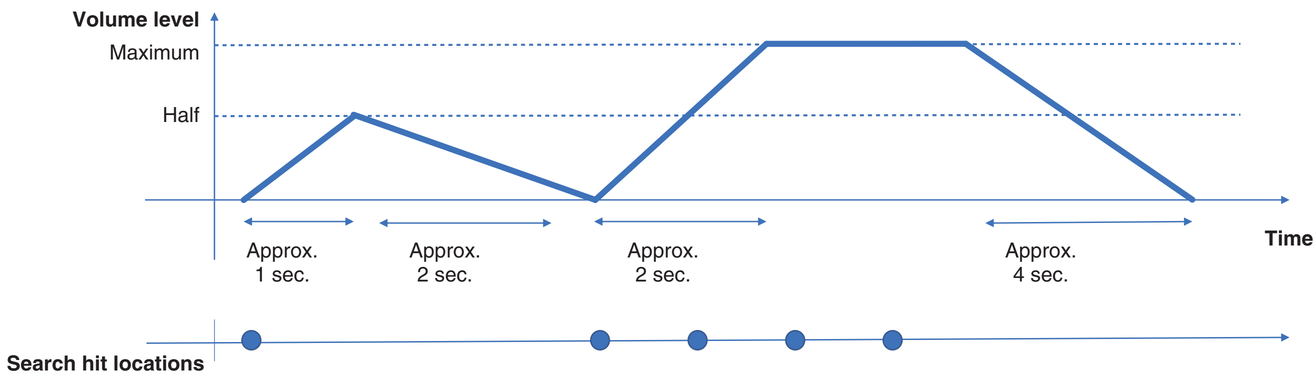


Fig. 6. Examples of search-hit locations and volume changes.

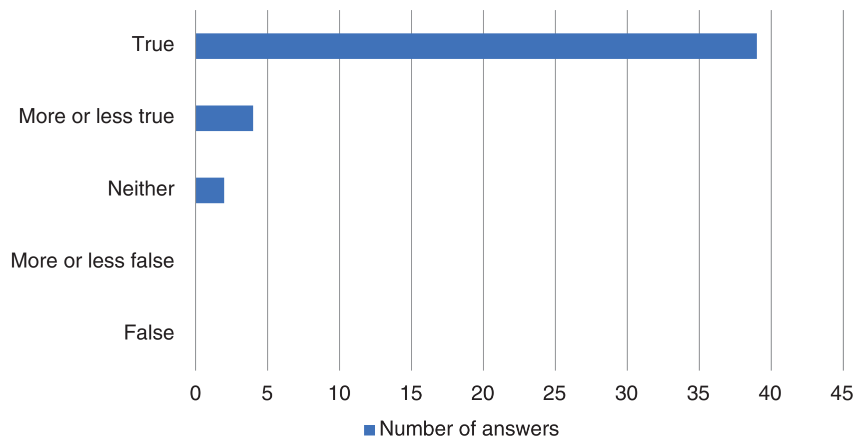


Fig. 7. Questionnaire item "I felt that it would be better to have a mechanism to produce cheers compared with the usual delivery of just watching the show" (45 respondents).

intuitive volume changes.

2.2 Evaluation by questionnaire to participants

To confirm the effectiveness of the harmonic reproduction, in which the cheer sounds are estimated from the audience images and played back at the venue, a questionnaire was sent to those participated as remote audience members during the experiment. In response to the statement, "I felt that it would be better to have a mechanism to produce cheers compared to the usual delivery of just watching the show," 86.6% of the participants responded positively on a 5-point scale, i.e., "true," "more or less true," "neither," "more or less false," and "false" (**Fig. 7**). Thus, the majority of participants had a favorable view of the harmonization reproduction system, confirming the effectiveness of delivering responses from the remote audience to the real venue.

3. Future developments

Remote audience members were asked to connect one by one from their homes, and a survey was conducted to determine what type of viewing environment they would prefer for remote participation, i.e., "In which of the following viewing situations would you prefer to watch a live webcast remotely?". More than half (68.8%) of the respondents answered, "Gather at home or a friend's house and participate with a friend from a single smartphone, PC, or monitor" (**Fig. 8**). This result may suggest that a viewing style in which good friends gather to participate in a remote-viewing environment and many such viewing environments are connected to the real venue to

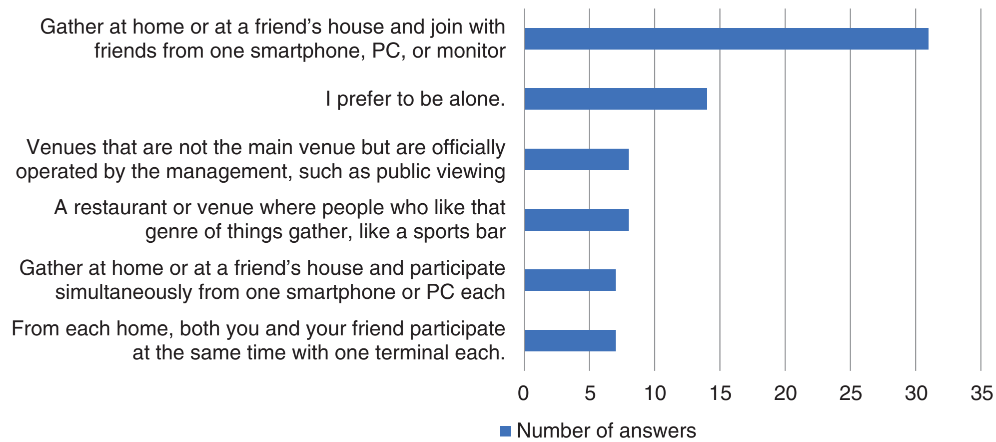


Fig. 8. Questionnaire item "Which of the following viewing situations would you prefer for watching a live webcast remotely?" (45 respondents, multiple responses allowed).

watch together will be preferred in the future. In a well-developed environment, such as a public-viewing event, there may be no problem in pursuing a high level of realism and rich transmission, including the atmosphere of the venue. However, for home-use, as mentioned earlier, if the video images captured with cameras in the home environment and the sound captured with microphones are transmitted and played back at the venue as they are, problems are expected to arise in terms of production. Therefore, NTT will continue researching and developing two-way video communications that can reproduce images and sounds in harmony so as not to interfere with live transmission by selecting information in the pursuit of reality as well as consideration of what information is to be emphasized or suppressed.

References

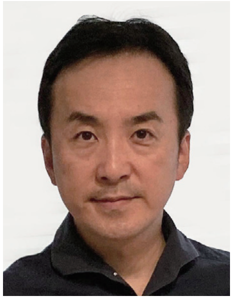
[1] S. Usui, S. Fukatsu, E. Matsumoto, M. Imoto, D. Shirai, and S. Kinoshita, "Marathon × Ultra-low-latency Communication Technology," NTT Technical Review, Vol. 19, No. 12, pp. 78–84, 2021. https://doi.org/10.53829/ntr202112fa10
[2] Mynavi TOKYO GIRLS COLLECTION 2022 SPRING/SUMMER (in Japanese), https://tgc.girlswalker.com/22ss/
[3] M. Yasuda, Y. Ohishi, Y. Koizumi, and N. Harada, "Crossmodal Sound Retrieval Based on Specific Target Co-occurrence Denoted with Weak Labels," Proc. of INTERSPEECH 2020, pp. 1446–1450, Virtual, Oct. 2020.
[4] Smart vLive® (in Japanese), https://www.ntt.com/business/services/voice-visual-communication/business-support/smartvlive.html
[5] ECLWebRTC, https://webrtc.ecl.ntt.com/en/

Takayuki Kurozumi
Senior Research Engineer, Cyber-World Laboratory, NTT Human Informatics Laboratories.

He received a B.S. in physics from the Tokyo Metropolitan University in 1997 and M.S. and Ph.D. in information science from the Japan Advanced Institute of Science and Technology, Ishikawa, in 1999 and 2007. In 1999, he joined NTT Communication Science Labs, where he was engaged in research and development (R&D) on multimedia information retrieval. From 2015 to 2017, he was with NTT Media Intelligence Labs, where he researched and developed image retrieval technology. He was with NTT DATA Corporation between 2017 and 2019, where he engaged in the planning and product development of multimedia information retrieval systems. From 2019 to 2021, he was with NTT Media Intelligence Labs, where he researched and developed compressed sensing and modeling technology. Since 2021, he has been with NTT Human Informatics Labs, where he is researching and developing distributed ultra-reality viewing technology. He was awarded the IEEE Transactions on Multimedia Paper Award in 2004.

Keisuke Hasegawa
Research Engineer, NTT Human Informatics Laboratories.

He received a B.E. and M.E. in informatics from Kyoto University in 2012 and 2014. He joined NTT WEST in 2014 and engaged in the maintenance of network facilities. Since joining NTT Service Evolution Laboratories in 2016, he has been engaged in research of media processing for the super realistic telecommunication technology Kirari!. He is a member of the institute of image information and television engineers (ITE).



Eiichiro Matsumoto
Senior Research Engineer, Cyber-World Laboratory, NTT Human Informatics Laboratories.

He received a B.E. in science engineering (information science) from Meiji University in 1999 and joined NTT the same year. From August 1999 to 2008, he was engaged in the development of geographic information systems and voice/video communication systems at NTT EAST. From 2008 to 2010, he was involved in the R&D of Internet Protocol television (IPTV)/mobile broadcasting standards and the formulation of technical standards at NTT. From 2010 to 2020, he was engaged in the planning and development of network services and systems for video distribution systems, such as IPTV and radio-frequency TV and researching video-transmission-system technologies for adding high value to relay networks and access networks at NTT EAST. He has been at his current position since 2020.

Toshihiko Eura
Senior Research Engineer, Cyber-World Laboratory, NTT Human Informatics Laboratories.

He received an M.E. from the Graduate School of Engineering at the University of Tokyo in 2005. He joined NTT EAST in 2005, where he was engaged in corporate user business development. From 2007 to 2019, he was engaged in network business development. From 2019 to 2021, he was engaged in the promotion project of the major international sporting event held in Tokyo in 2021.

Shinji Fukatsu
Senior Research Engineer, Supervisor, Cyber-World Laboratory, NTT Human Informatics Laboratories.

He received a Ph.D. in engineering from Osaka University in 2002 and joined NTT the same year. He has been engaged in R&D of human interfaces and video streaming services as well as in planning and development of video streaming services at NTT Plala and in promoting standardization and international development of information and communication technologies at the Ministry of Internal Affairs and Communications. He is currently engaged in R&D of remote-world infrastructure technology.

Emotional-perception-control Technology for Estimating and Leading Human Emotions

Rika Mochizuki, Motohiro Makiguchi, Masahiro Kohjima, Masanori Yokoyama, and Ryuji Yamamoto

Abstract

The number of people who participate in live music and sporting events in remote environments has been increasing; however, participants must sacrifice the emotional experiences that can only be obtained on site, such as enthusiasm, a sense of unity, and contagious enthusiasm of the audience. Toward the creation of a world where people's emotions are actively amplified and resonate even in remote environments, we introduce an emotional-perception-control technology for generating a personal virtual site where participants can experience the unique pleasures of virtual reality through estimating their emotional-expression characteristics and leading (guiding) their emotions based on these characteristics.

Keywords: emotion modeling, emotion leading, personal virtual site

1. Emotional-perception-control technology

Online participation in live music, sports, and other events is expected to become essential along with onsite participation as a means of easily gathering people from around the world without worrying about location or distance. Various efforts are being made to improve the sense of presence, such as increasing the resolution of live images, creating multiple viewpoints by arranging multiple cameras, and increasing the wide viewing angle by using 360-degree cameras. However, the emotional experiences that can only be experienced on site, such as enthusiasm, a sense of unity, and contagion of enthusiasm felt in stadiums and live venues, are being lost. Therefore, we aim to create a world in which people's emotions are actively amplified and resonated even in remote environments by generating a personal virtual site optimized for each person's unique way of enjoying themselves, such as feeling a sense of unity in conjunction with other spectators or being absorbed in a space alone, by estimating each person's emotional expression characteristics and leading (guiding) their emotions based on these characteristics.

Emotional-perception-control technology leads to desirable emotions for users, such as enhancing enthusiasm and a sense of unity, through two core technologies: emotion estimation for estimating and understanding human emotional characteristics through sensing and data analysis and emotion leading for leading emotions through perceptual stimuli tailored to human emotional characteristics. By combining the estimation and leading of emotions including those of individuals, crowds, and interactions between the two, we can generate a personal virtual site for providing an optimal experience for each person (**Fig. 1**).

Emotion-estimation technology is used to quantitatively understand and model the emotions of individuals and crowds on the basis of sensed biometric, image, sound, and content data. Emotion-leading technology is an interaction technology for inducing

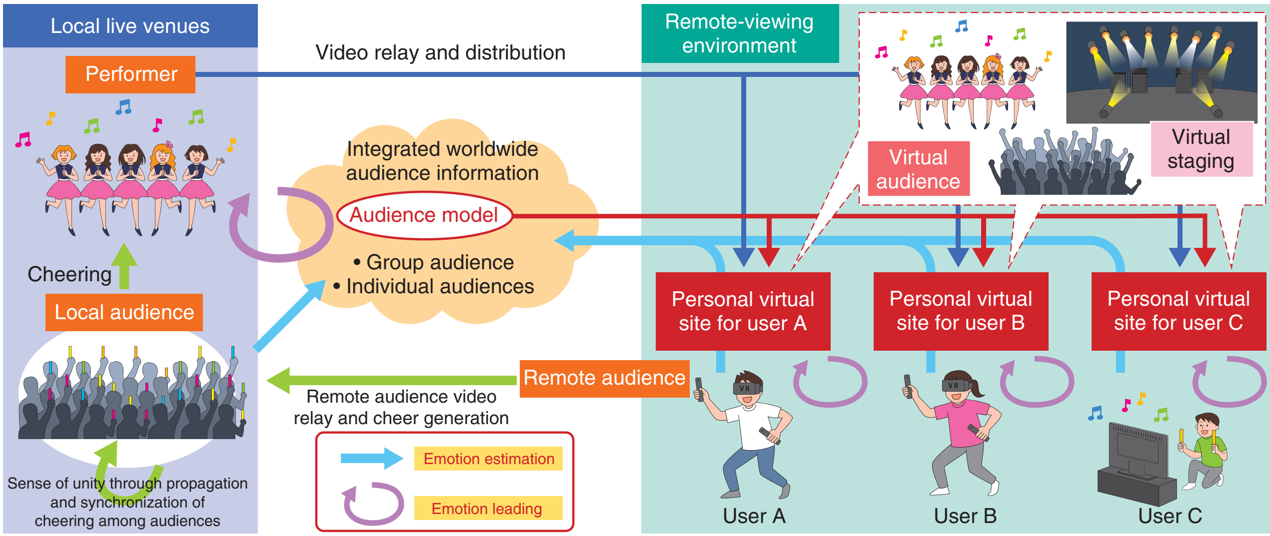


Fig. 1. Personal virtual sites generated with emotion-estimation and emotion-leading technologies.

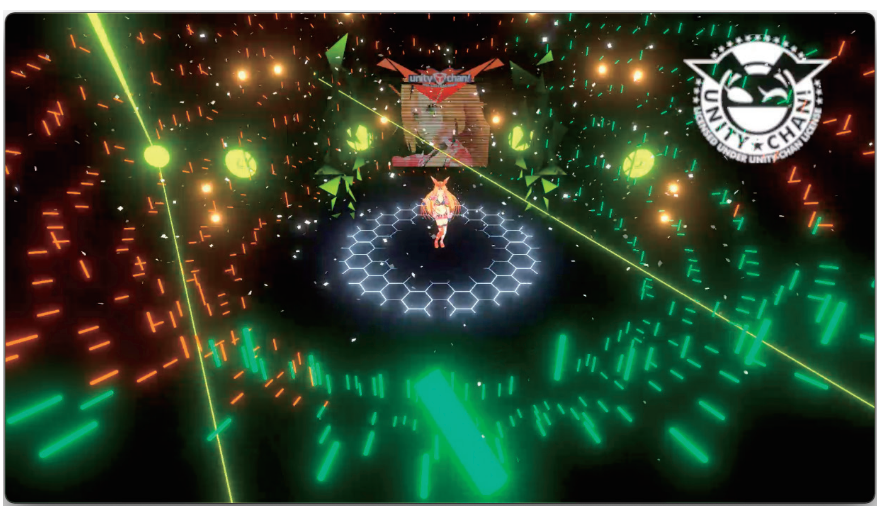


Fig. 2. Image of a personal virtual live site.

the same or even more emotions in a remote environment as in a local venue using knowledge of perception, cognitive psychology, and human-computer interaction. Emotion-estimation technology estimates and models the states of local and remote audiences and reflects them in the virtual audience, while emotion-leading technology optimizes the behavior of the virtual audience and the staging of the venue according to the characteristics and emotional state of each remote audience member, as shown in Fig. 1.

This article introduces an emotional-perception-control technology for generating a personal virtual live site (**Fig. 2**) for events such as concerts. By optimizing the presentation and interaction methods of the virtual audience for each remote audience member, the emotional experience is enhanced by perceptually enhancing the sense of unity and enthusiasm caused by the propagation and synchronization of cheering that occurs among audiences at different locations during live concerts.

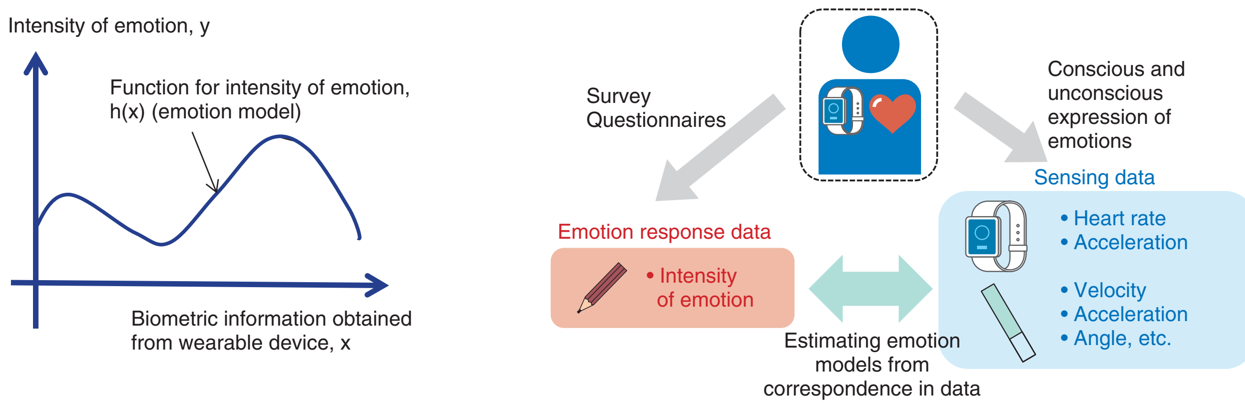


Fig. 3. Emotion-model estimation.

2. Emotion-estimation technology

In live music and sporting events, the point of excitement differs depending on the individual/entire audience's preference for music and staging of live events and their knowledge of sports. Suppose we can estimate the emotion of individual spectators and entire audiences without explicitly asking them. The magnitude of the effects of the music, staging, and camera work on emotional responses can then be measured, which helps evaluate and improve the music and staging. Estimated emotion can also be used for generating a personal virtual live site, as illustrated in Fig. 2. We can adaptively change the type of venue (such as a house and arena), staging, and movements of the surrounding virtual audiences for each site to enhance or suppress emotions. With these application scenarios in mind, we are developing methods for estimating (i) the emotions of individual audience members using wearable devices and (ii) the emotions (collective characteristics) of the entire audience using live videos of the actual event venue.

2.1 Estimating the emotions of individual audience members using biometric signals

This is a method that uses artificial intelligence (AI) to estimate the intensity of emotions, such as pleasure, discomfort, high arousal, and low arousal, and types of emotions, such as joy and sadness, felt by the user. This involves using biometric information obtained from wearable devices such as smartwatches and hitoe™*, which have become widespread, as input. The advantage of this method is that it can be applied to any environment where there are no cameras or microphones, such as the living room or outside the home, because it uses a wearable device that the user wears in daily life. The method requires the training of AI, i.e., estimating an unknown function (*emotion model*), the input of which is sensed biometric information and output is the intensity or classification result of an emotion from the user's biometric information and an emotion response (e.g., a five-point rating of pleasure) at a certain point in time (**Fig. 3**). This training requires the user's subjective responses since emotions that are not observable are the model's output. However, such a simple five-point rating format has problems such as differences in interpretation of each point between participants [1], making it more difficult than the standard setup for training AI using non-subjective data such as whether a certain image is a cat or dog, and collection of large amounts of data. We are currently developing a method for handling such difficulties, including a data-collection format suitable for handling subjective data and an AI training method tailored to that format [2, 3].

2.2 Estimating the collective characteristics of the entire audience using live video from an actual event venue

This method estimates the collective characteristics of a group based on the observed group behavior (such as shouting, waving penlights, clapping, and hand-signing) in the live video of the actual event venue, including the audience seats. For example,

* hitoe™: A functional fabric that is made by coating a conductive polymer on a cutting-edge fiber material called nanofiber fabric. A clothing-type vital sensor using hitoe™ can measure biosignals with high sensitivity without burden on the wearer.

how the audience's behavior is affected (on average) when the performer shouts out to excite them or how well the audience as a whole behaves in a synchronized and united manner. Unlike the above method involving wearable devices, this method does not require users to own a device and can be used at any event venue. However, the estimation target is not the emotion of an individual but the characteristics of the group as a whole. By using the estimated group characteristics, it is possible to understand which songs in a music concert generate a unified behavior of the entire audience. We can also enable the virtual audience in the personal virtual live site behave similar to that of the audience at the actual venue.

3. Emotion-leading technology

Emotion-leading technology is an interaction technology that naturally leads remote audience members to the desired emotional state when they participate in an event online. In many cases, online participation, such as live music streaming, differs from on-site participation in many ways such as participation from a computer or smartphone at one's home. Therefore, we are conducting research to enhance the emotional experience from two aspects: psychological and behavioral [4], to provide similar or better emotional experience in a remote environment as in a local environment. An example of the psychological aspect is the excitement from the enthusiasm of the surrounding audience, and an example of the behavioral aspect is the excitement of cheering, clapping, and waving penlights.

To enhance the emotional experience from the psychological aspect, we are researching an audience-presentation-optimization method for optimizing the presentation of audience members other than oneself to remote audience members.

As a concrete example of optimization, we can consider reducing the display of audiences that behave too differently from the target audience or are not excited at all and present audiences that behave and are excited in the same way as the target audience. Another example would be to highlight the behavior of audiences who are similarly cheering for a band member that the target audience likes. To make such optimization possible, it is first necessary to obtain data on the behavior of the audience. We focused on penlights as a typical item at music events and developed a sensing penlight equipped with multiple sensors such as acceleration (**Fig. 4**).

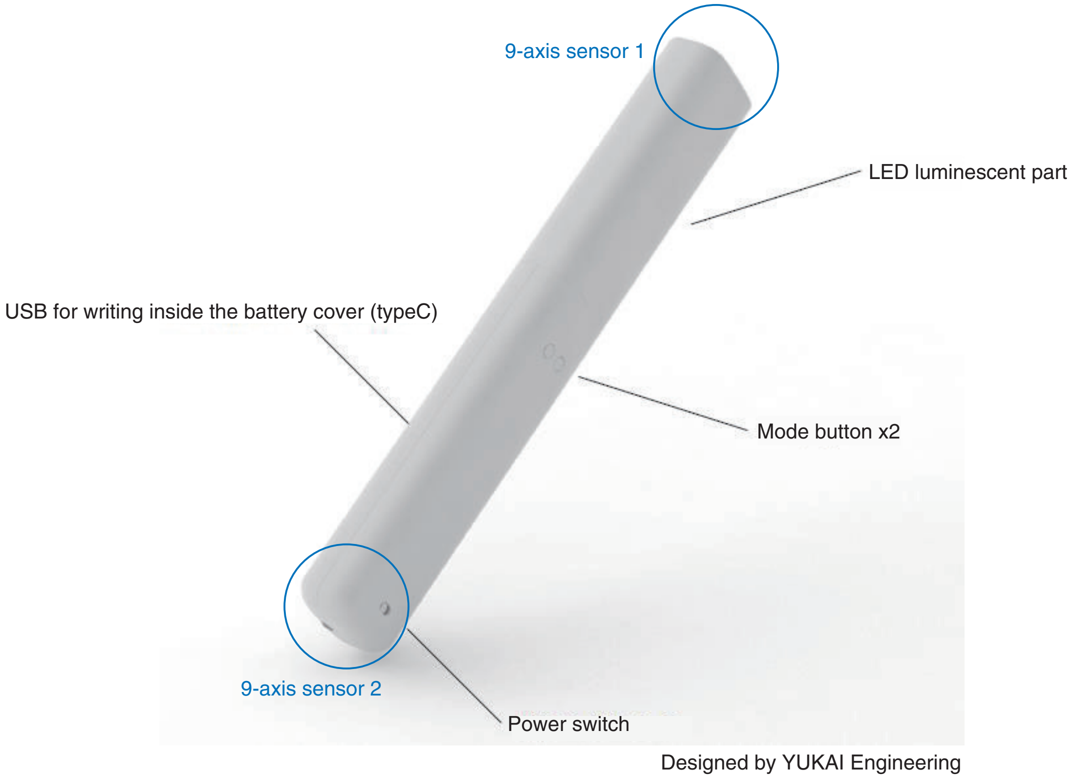
The sensing penlight can detect the presence or absence of swing, swing scale, swing cycle, color conversion, and other data. The data obtained can be used to generate the optimal audience-presentation pattern for each remote audience. The behavior of the audience in response to the generated pattern is obtained again to tune the presentation pattern, thus improving optimization accuracy.

Our approach to enhancing the emotional experience from a behavioral aspect is based on the fact that in many online viewing environments, such as at home, it is difficult for the audience to cheer loudly or move as much as they would at a local venue. We are investigating a method of adding multimodal feedback stimuli to give the remote audience the illusion that their actions are larger than they actually are. This method aims to provide an experience in which the remote audiences feel as if they are performing the same physical actions as in the actual venue. By studying the methodologies for incorporating an on-site experience into the remote viewing experience, multiple emotion-leading methods and application patterns of such methods for each user can be developed. The selection scheme, intensity, and timing of application of emotion-leading methods can be optimized on the basis of the emotions of individual users or groups of users using emotion-modeling methods.

By combining emotion-leading technology and emotion-estimation technology, it will be possible to design a personalized virtual space and venue and generate a personalized virtual live site where each individual remote audience member can share their enthusiasm as well as be influenced by others'.

4. Future work

Hybrid events held in both real and remote environments have become more common, and it is important to enhance the experience that can only be had in a remote environment. Our emotional-perception-control technology aims to provide new experiences through the generation of a personalized virtual site, where each user can enjoy the experience of being involved in a space optimized for him or her by taking advantage of the effects of online performances that are impossible in reality, such as unlimited seating for audience members from around the world, freely changing seating arrangements, and the spread of one's actions throughout the venue. We will also consider extending this technology to fields other than live music performances.



LED: light-emitting diode
USB: Universal Serial Bus

Fig. 4. Developed sensing penlight.

References

[1] G. N. Yannakakis and H. P. Martinez, "Ratings Are Overrated!", Frontiers in ICT, Vol. 2, No. 13, 2015.

[2] Y. Nanbu, M. Kohjima, and R. Yamamoto, "Learning of Regression Models from Pairwise Data and Target Variable Quantiles," Transactions on Mathematical Modeling and its Applications (TOM), Vol. 15, No. 3, pp. 19–28, July 2022.

[3] M. Kohjima, "Uncoupled Nonnegative Matrix Factorization with Pairwise Comparison Data," Proc. of 2022 ACM SIGIR International Conference on the Theory of Information Retrieval (ICTIR), pp. 103–107, Madrid, Spain, July 2022.

[4] D. Yamashiro, M. Aihara, C. Ono, H. Kanemura, K. Aoyagi, Y. Goto, Y. Iwadare, and S. Nakazawa, "Sympathetic Skin Response and Emotional Changes of Visual Stimuli," No to Hattatsu (Brain and Development), Vol. 36, No. 5, pp. 372–377, 2004 (in Japanese).

Rika Mochizuki

Senior Research Engineer, Symbiotic Intelligence Laboratory, NTT Human Informatics Laboratories.

She received an M.E. and Ph.D. in engineering from Chuo University, Tokyo, in 2009 and 2011. She joined NTT in 2009, and her research interests include color vision compensation, life-log processing, and their applications for communication support. She received the Minister of Education, Culture, Sports, Science and Technology Award of the 23th Grand Prize for Frontier Technology: Pave a New Way for Creativity, and the Best Paper Award from the Institute of Electronics, Information and Communication Engineers (IEICE) in 2009. She is a member of IEICE, Human Interface Society, the Computer Software Association of Japan, and the Information Processing Society of Japan (IPSJ).

Motohiro Makiguchi

Research Engineer, Symbiotic Intelligence Laboratory, NTT Human Informatics Laboratories.

He received an M.E. and Ph.D. in information science from Hokkaido University in 2012 and 2020. He joined NTT in 2012 and has been researching three-dimensional display technology and virtual-reality technology.

Masahiro Kohjima

Research Engineer, Symbiotic Intelligence Laboratory, NTT Human Informatics Laboratories.

He received a B.E., M.E., and Ph.D. in computer science from Tokyo Institute of Technology in 2009, 2012, and 2019. He joined NTT in 2012, and his research interests lie in the area of machine learning with emphasis on probabilistic models, Bayesian methods, and reinforcement learning. He is a member of IPSJ.

Masanori Yokoyama

Research Engineer, Symbiotic Intelligence Laboratory, NTT Human Informatics Laboratories.

He received an M.Sc. from Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, Kanagawa, in 2010. He joined NTT in 2010, and his research interests include nonverbal media, telepresence interface, and their applications for communication support. He received the 3rd Best Paper Award of ACM Augmented Human 2016, and Young Researcher Award of the HIS Symposium on Human Interface 2017. He is a member of IEICE and IPSJ.

Ryuji Yamamoto

Senior Research Engineer, Supervisor of Symbiotic Intelligence Laboratory, NTT Human Informatics Laboratories.

He received an M.E. from Kyushu Institute of Technology, Fukuoka, in 1998 and joined NTT the same year. His research interests include media processing and human-computer interaction.

Initiatives toward Open RAN

Daisuke Hiratsuka, Keiko Kuriu, Anil Umesh, and Haruki Mori

Abstract

The fifth-generation mobile communications system (5G) network will need to support a wide range of services compared with past networks. To meet this need, it will be necessary to construct a flexible network that can accommodate all types of services. The Open Radio Access Network (RAN) can make this a reality. This article provides an overview of Open RAN, presents the current state of the O-RAN ALLIANCE that promotes Open RAN standards, and describes the 5G Open RAN Ecosystem, a new NTT DOCOMO Open RAN initiative.

Keywords: vRAN, virtualization, Open RAN

1. Introduction

To respond in a more suitable manner to customer needs, all mobile operators are required to add new network equipment and replace equipment to expand network functions on a continual basis. The network will therefore have to be capable of such scalability in a flexible and agile manner. To this end, NTT DOCOMO has been promoting open interfaces that will enable operators to freely select and adopt optimal solutions from a variety of vendor products.

This article provides an overview of the Open Radio Access Network (RAN)[*1] and touches on the current state of the O-RAN ALLIANCE. It also describes the 5G (fifth-generation mobile communications system) Open RAN Ecosystem launched by NTT DOCOMO to globally accelerate Open RAN and the approach that it takes to solve problems related to performance, integration[*2], and interoperability testing in Open RAN.

2. Open RAN

2.1 Three elements of Open RAN

Open RAN consists of the following three elements (**Fig. 1**):

(1) Open interfaces that combine RAN equipment from a variety of vendors
(2) Virtualization (i.e., virtualized RAN (vRAN)) that enables hardware and software inside RAN equipment to be separated
(3) Intelligent control that optimizes and automates RAN operation

2.2 Effects of Open RAN

The base-station components for Open RAN can be separated into three units—radio unit (RU)[*3], distributed unit (DU)[*4], and central unit (CU)[*5]—each of which can be connected via standardized interfaces. This means that telecom operators can free themselves from vendor lock-in, shorten the time to commercial deployment, and adopt a better equipment configuration to provide optimized services to consumers. The virtualization of a RAN means that costs can be decreased through the use of general-purpose hardware and that flexibility and scalability can be enhanced. Envisioning an increasingly complex mobile network, manual operation of the network as has been the practice will become all the more

*1 RAN: A network consisting of radio base stations and other equipment situated between a core network and mobile terminals to control the radio layer.

*2 Integration: The process of incorporating equipment and systems into a network managed by an operator.

*3 RU: The radio unit of a wireless base station.

*4 DU: A component of a base station, the node that processes radio signals and transmits and receives radio waves.

*5 CU: Equipment that connects to the baseband unit and controls radio resources.

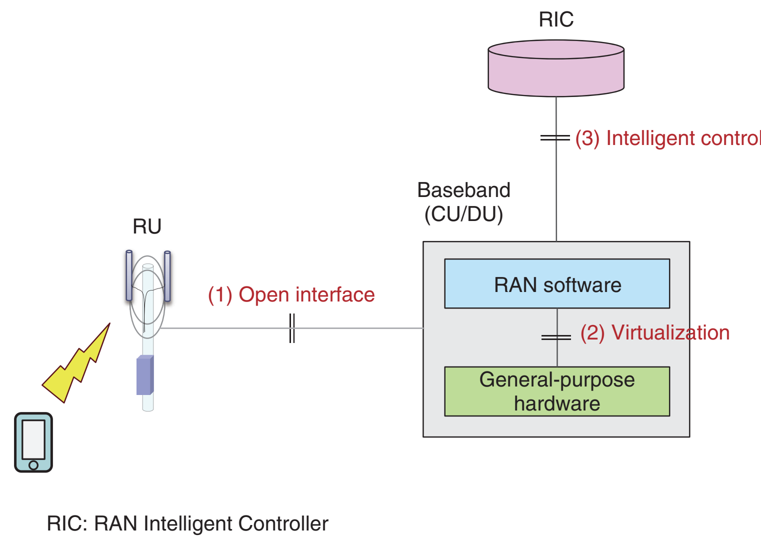


Fig. 1. Main elements of Open RAN.

difficult, but this problem can be solved through the intelligent control of a RAN.

3. Current status of Open RAN standardization

In February 2018, NTT DOCOMO joined with other industry operators to establish the O-RAN ALLIANCE with the aim of promoting open and intelligent RANs [1]. We introduce new initiatives of this organization from 2019 to the present.

There were 19 operators and 55 vendors and institutions serving as members of the O-RAN ALLIANCE in 2019, but this has expanded to 31 operators and 294 vendors and institutions as of February 22, 2022.

While only fronthaul[*6] specifications from Work Group 4 (WG4) had been released in 2019, new specifications had been released from each WG and Focus Group (FG) as of February 22, 2022. We introduce newly released specifications from WG4 and WG5 for which NTT DOCOMO has been serving as co-chair.

In WG4, in addition to fronthaul specifications consisting of the control, user and synchronization plane (CUS-Plane) and management plane (M-Plane), the following three types of specifications have been released with version upgrades already being made.

Test specifications:

- Open Fronthaul Conformance Test Specification
- Fronthaul Interoperability Test (IOT) Specification

Cooperative transport interface (CTI) specifications:

- Fronthaul CTI Transport Control Plane Specification

In constructing Open RAN, testing must be conducted to confirm whether equipment from different vendors conforms to the interface specifications of the O-RAN ALLIANCE and whether multi-vendor connections can be made.

In WG5, the following specifications were released with version upgrades being made.

X2[*7] specifications:

- New Radio (NR)[*8] control plane (C-plane) profile
- NR user data plane (U-plane) profile

Transmission path specifications:

- Transport Specification

Monitoring control specifications:

- O1 Interface specification for O-CU-UP and O-CU-CP

*6 Fronthaul: In a radio base station, the interface for the optical-fiber link connecting the baseband processing section that executes digital signal processing and the RUs that transmit and receive radio waves.

*7 X2: A reference point between eNodeB (base station for Long-Term Evolution (LTE) radio access) defined by the 3rd Generation Partnership Project (3GPP).

*8 NR: A radio system standard formulated for 5G. Compared with 4G, it enables faster communication by using high-frequency bands (e.g., 3.7- and 28-GHz bands), and low-latency and highly reliable communication for achieving advanced Internet of Things.

Table 1. O-RAN ALLIANCE technology study groups.

	WG/FG	Scope of study
WG1	Use Cases and Overall Architecture	Architecture, use cases, slicing, demonstrations
WG2	Non-real-time RIC and A1 Interface	Non-real-time RIC, A1, rApp
WG3	Near-real-time RIC and E2 Interface	Near-real-time RIC, E2, xApp
WG4	Open Fronthaul Interfaces	Fronthaul
WG5	Open F1/W1/E1/X2/Xn Interface	Interoperability profiles such as X2, Xn, and F1; O1 to DU and CU
WG6	Cloudification and Orchestration	O-Cloud, vDU/vCU, AAL, O2
WG7	White-box Hardware	Mainly RU-hardware reference design
WG8	Stack Reference Design	DU/CU software-architecture reference design
WG9	Open X-haul Transport	Transport equipment, transport-network control/maintenance protocol
WG10	OAM for O-RAN	SMO, O1 (overall coordination)
SFG	Security	Security risk analyses and countermeasure studies for Open RAN
TIFG	Test & Integration	Compile test specifications, plugfest, OTIC, certification and badging processes
OSFG	Open Source	O-RAN Software Community
SDFG	Standard Development	Standardization strategies, interface to other standard development organizations

AAL: Acceleration Abstraction Layer
OAM: operation and maintenance
OTIC: Open Testing and Integration Centre
rApp: Non-real-time RIC application
SMO: Service Management and Orchestration
xApp: Near-real-time RIC application

- O1 Interface specification for O-DU

IOT specifications:

- Interoperability Test Specification

To make intelligent Open RAN a reality, a number of new WGs and FGs have been established. **Table 1** summarizes the WG/FG activities.

The WGs and FGs that have been added since the founding of the O-RAN ALLIANCE in 2018 are WG9, WG10, Security FG (SFG), Test & Integration FG (TIFG), Open Source FG (OSFG), and Standard Development FG (SDFG).

Security concerns, in particular, have been raised as a key issue in Open RAN, and at the O-RAN ALLIANCE, SFG is conducting security risk analyses and countermeasure studies.

4. NTT DOCOMO Open RAN initiatives

4.1 NTT DOCOMO multi-vendor network initiative

In conventional networks, a single vendor provides base stations for both baseband signal processing units (BBUs) and RUs in a single-vendor configuration. One advantage of having a single vendor is that the operator can rely on that vendor for a range of services from deployment to maintenance in a one-stop manner. Since the interfaces between equipment are vendor-proprietary interfaces, a disadvantage is that the operator cannot upgrade to equipment from other vendors, decreasing flexibility.

NTT DOCOMO took the lead in achieving a multi-vendor network before the coming of 5G. Specifically, NTT DOCOMO enabled connections between different vendors by prescribing original interfaces between the BBU and RU, achieving a flexible base-station configuration (**Fig. 2**). Since there were now multiple options for selecting base-station-equipment vendors, it became possible to select the most optimal vendors in terms of cost and performance, lowering the cost of deploying equipment.

4.2 NTT DOCOMO Open RAN initiatives in 5G

When launching 5G commercial services in 2020, NTT DOCOMO was the first in the world to achieve Open RAN in a commercial network using interfaces conforming to O-RAN ALLIANCE standards. All 5G base stations that NTT DOCOMO is now rolling out conform to O-RAN ALLIANCE fronthaul and X2 specifications. Since this is Open RAN, NTT DOCOMO has been gradually expanding equipment vendors and the variation in equipment ever since the 5G pre-commercial service period.

Specifically, NTT DOCOMO has come to support millimeter-wave (mmW)*9 capabilities, sub-6*10 inter-band carrier aggregation (CA)*11, and stand-alone

*9 mmW: Radio signals in the frequency band from 30 to 300 GHz as well as the 28-GHz band targeted by 5G all of which are customarily called “millimeter waves.”

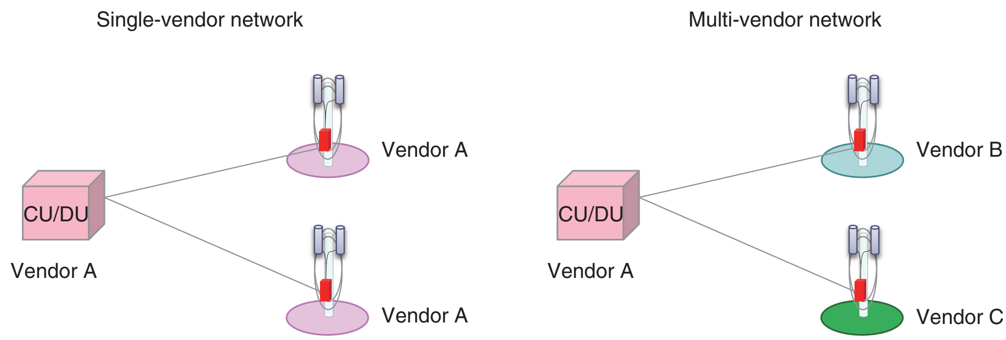


Fig. 2. Single-vendor and multi-vendor networks.

(SA)*12 operation. While only small RUs (SRUs) for small-cell use of both the antenna-separated and antenna-integrated types were deployed at the launch of 5G commercial services, NTT DOCOMO has since expanded the equipment lineup to include regular power RUs (RRUs) for macrocell use and 5G fronthaul multiplexers (FHMs)*13. NTT DOCOMO is therefore adopting new vendors for CU/DU and RU.

NTT DOCOMO considers its multi-vendor initiative between equipment through open interfaces that it has been implementing as being its first step in Open RAN deployment. As the next step, it will promote studies on achieving the remaining two elements of Open RAN described above, i.e., virtualization and intelligent control (**Fig. 3**). This will be achieved through NTT DOCOMO's 5G Open RAN Ecosystem initiative described in the next section.

5. Promoting Open RAN across the globe

5.1 Founding of 5G Open RAN Ecosystem

In February 2021, NTT DOCOMO established the 5G Open RAN Ecosystem together with 12 other companies with the aim of accelerating the adoption of Open RAN by operators. Through this ecosystem, NTT DOCOMO aims to accelerate vRAN testing. It also intends to make state-of-the-art RAN into a commercial package based on the requirements of mobile operators that are studying the introduction of Open RAN and to offer Open RAN deployment, operation, and maintenance services. By making use of the Open RAN knowledge it has built up over many years, NTT DOCOMO will work to promote the 5G Open RAN Ecosystem and provide high-quality and flexible networks.

5.2 Open RAN problems and solutions

As described above, Open RAN has advantages but also has a number of problems that must be solved. We give an overview of the main problems in deploying Open RAN and describe how the 5G Open RAN Ecosystem plans to solve these problems.

(1) Performance

In vRAN, one of the elements of Open RAN, general-purpose servers are used as hardware, but when running RAN applications on such servers, radio characteristics, accommodation capacity, etc. may deteriorate. As a solution to this problem, end-to-end (E2E) vRAN testing using an accelerator is being promoted within the 5G Open RAN Ecosystem to achieve performance 2–3 times that of current performance. The 5G Open RAN Ecosystem Whitepaper including these performance targets was released in June 2021, so we ask the reader to refer to that material as well [2].

(2) Integration

In Open RAN, base-station components can be separated, but an issue that arises is how to integrate those components. In vRAN, hardware and software can be separated, so it is believed that components from different vendors can be integrated and provided accordingly. In this case, the number of interfaces needed for integration increases compared with conventional RAN. As a solution to this problem, an Open RAN testing environment, which is described

*10 Sub-6: A radio signal with the frequency band between 3.6 and 6 GHz.
*11 CA: A technology for increasing bandwidth by simultaneously transmitting/receiving signals using multiple component carriers.
*12 SA: Stand-alone format. A form of mobile communication network on which terminals connect using a single wireless technology.
*13 FHM: Equipment that multiplexes multiple fronthaul lines between the baseband processing section and radio equipment.

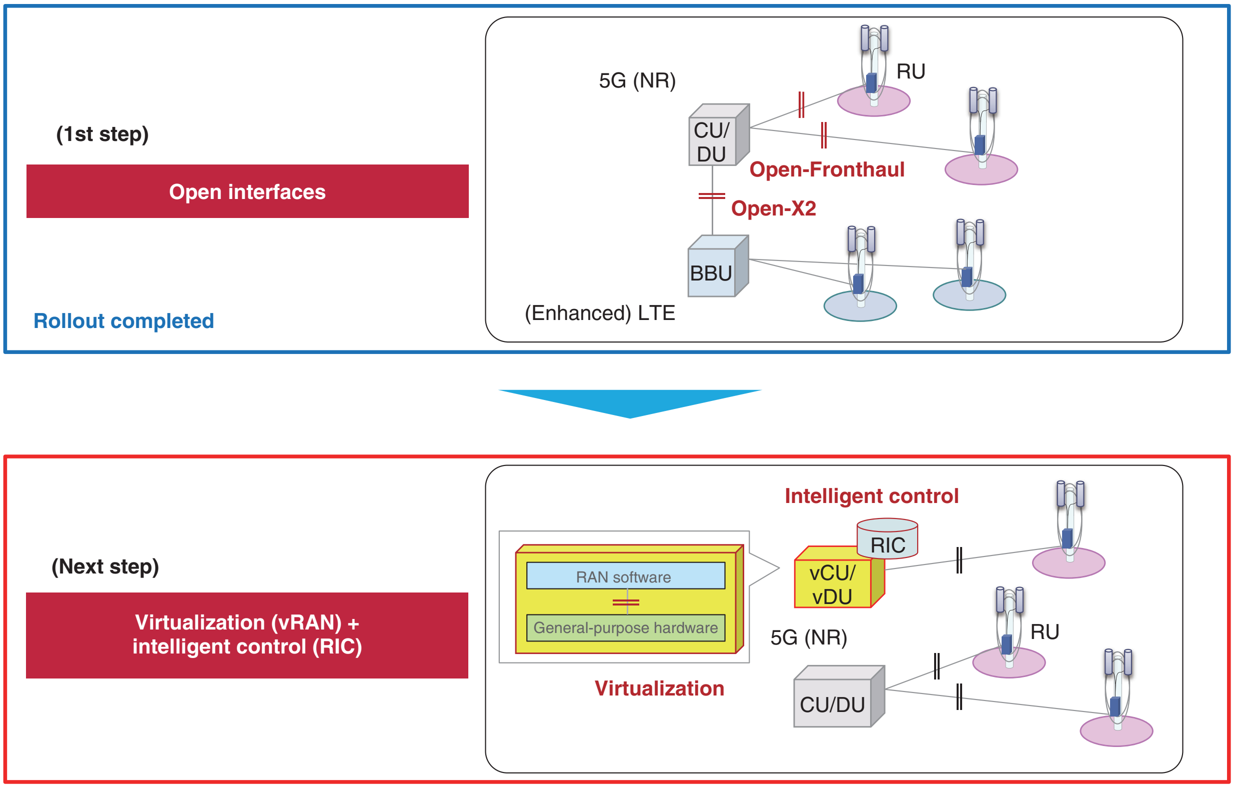


Fig. 3. NTT DOCOMO Open RAN deployment steps.

in the next section, is being set up within the 5G Open RAN Ecosystem to provide multi-vendor integration testing to overseas operators who are testing and operating RAN.

(3) Other issues

Costs, automation, and equipment rollout are also problems in Open RAN, and NTT DOCOMO solutions to these problems are described in other feature articles in this issue [3, 4].

5.3 Sharing of Open RAN testing environment

As described above, using multi-vendor products means that products must be appropriately selected and integration testing conducted. However, test cases increase as the number of products and interfaces increases, and having each operator prepare an environment tailored to its testing needs is inefficient.

Against the above background, NTT DOCOMO set up an open testbed in Japan as part of the 5G Open RAN Ecosystem. One of the main functions of this testbed is enabling overseas mobile operators to remotely control it as if it was a testbed in their own laboratories. This testbed can be connected to an operator's core network*14, making it easy to conduct tests of vRAN equipment using products from multiple vendors. Thus, an operator can dramatically decrease the time and costs incurred by testing, so it is thought that this testbed can contribute to the timely deployment of Open RAN. The testbed began operation in summer 2021. Products from 5G Open RAN Ecosystem partners have been provided since October 2021 and testing has begun. This testbed was released in February 2022 as a shared open laboratory providing an environment for operators around the world to test vRAN while based overseas. NTT DOCOMO is using this shared open laboratory to deepen its ties with a wide range of stakeholders including other operators and to contribute to the establishment of technologies and expertise toward the early diffusion of an open network, especially Open RAN and vRAN, that can respond flexibly and rapidly to diverse needs.

*14 Core network: A network comprising switching equipment, subscription information management equipment, etc. A mobile terminal communicates with a core network via a RAN.

6. Conclusion

This article presented an overview of Open RAN, described the current state of the O-RAN ALLIANCE that is responsible for Open RAN standardization, and described the 5G Open RAN Ecosystem, a new NTT DOCOMO Open RAN initiative. As an Open RAN pioneer, NTT DOCOMO will continue to promote Open RAN that should be implemented for its own network as well as globally.

References

[1] S. Abeta, T. Kawahara, A. Umesh, and R. Matsukawa, "O-RAN Alliance Standardization Trends," NTT DOCOMO Technical Journal, Vol. 21, No. 1, pp. 38–45, July 2019.

[2] NTT DOCOMO, "5G Open RAN Ecosystem Whitepaper," June 2021.
https://ssw.web.docomo.ne.jp/orec/5g_open_ran_ecosystem/whitepaper/OREC_WP.pdf

[3] S. Mizuta, A. Umesh, Y. Nakajima, and Y. Kuno, "Initiatives toward Virtualized RAN," NTT Technical Review, Vol. 20, No. 11, pp. 52–63, Nov. 2022.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr202211fa7.html

[4] T. Katsuragawa, A. Kawana, Y. Inoue, T. Tateishi, E. Hashimoto, and T. Fujitsuka, "Initiatives toward Intelligent RAN," NTT Technical Review, Vol. 20, No. 11, pp. 64–72, Nov. 2022.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr202211fa8.html

Trademark notes

All company names or names of products, software, and services appearing in this article are trademarks or registered trademarks of their respective owners.

Daisuke Hiratsuka
Global Business Department, NTT DOCOMO, INC.

He received a B.E. and M.E. from Yokohama National University in 2015 and 2017. He joined NTT DOCOMO in 2017 and has since been involved in research and development (R&D) activities related to wireless access technologies for 5G. He contributed to the O-RAN Alliance Working Group 4 and 5 open fronthaul management-plane topics as a delegate from NTT DOCOMO. He has also been involved in business development of the 5G Open RAN Ecosystem since 2021.

Anil Umesh
Senior Manager, Radio Access Network Development Department, NTT DOCOMO, INC.

He received a B.S. in engineering from Case Western Reserve University, USA, in 2002. He joined NTT DOCOMO the same year and has since been involved in R&D activities related to wireless access technologies for HSPA (High Speed Packet Access), LTE, LTE-Advanced, and 5G. He has been the O-RAN Alliance Working Group 4 co-chair since the start of O-RAN ALLIANCE in 2018 and has chaired open fronthaul topics.

Keiko Kuriu
Radio Access Network Development Department/Global Business Department, NTT DOCOMO, INC.

She joined NTT in 2012, engaged in R&D activities related to virtualization infrastructure for telecommunications, and contributed to the OpenStack project community. In 2021, she joined NTT DOCOMO and was involved in the launch of the 5G Open RAN Ecosystem and has since been involved in the vRAN/Open RAN verification project and marketing for global operators in the 5G Open RAN Ecosystem.

Haruki Mori
Manager, Global Business Department, NTT DOCOMO, INC.

He received a B.E. and M.E. from Okayama University in 2002 and 2004. He joined NTT DOCOMO in 2004, where he was involved in R&D activities related to wireless access technologies for LTE. He was also involved in the launch of sports-related new business. He has been involved in the promotion of the 5G Open RAN Ecosystem since 2021.

Initiatives toward Virtualized RAN

Shinji Mizuta, Anil Umesh, Yoshihiro Nakajima, and Yuya Kuno

Abstract

There is growing demand for high processing performance in wireless base-station equipment as bit rates and capacity continue their upward trend in Long-Term Evolution and fifth-generation mobile communications system networks. To meet this demand, NTT DOCOMO has been using purpose-built hardware and software. There has also been remarkable technical innovation in cloud computing for web services, as reflected by ongoing improvements in hardware performance and increasing decoupling of hardware and software provided via virtualization and cloud-computing technologies. These innovative technologies should lead to superior wireless base-station design and solutions so that a more disaggregated and efficient radio access network (RAN) can be achieved. The development and commercialization of such equipment is moving forward as RAN virtualization is becoming predominant. This article describes NTT DOCOMO initiatives toward RAN virtualization technology.

Keywords: vRAN, virtualization, Open RAN

1. Introduction

NTT DOCOMO began providing services under the fifth-generation mobile communications system (5G) in March 2020 and has been expanding the 5G coverage area ever since. To roll out 5G while making maximum use of existing Long-Term Evolution (LTE) base-station equipment, NTT DOCOMO has developed the bare minimum amount of hardware together with software to run on that hardware and added both to existing LTE base-station equipment to achieve 5G services.

There has also been tremendous technical innovation in the information technology (IT) field. Thus, there have been rapid gains in the performance of general-purpose hardware, in the effective use of hardware resources by decoupling hardware and software using virtualization technology, and in the omnipresent use of common cloud technology by leveraging standard and de-facto cloudification and orchestration platforms. To achieve high-performance data encryption, artificial intelligence (AI), and machine-learning technologies, specialized hardware accelerators for various types of computing are now being developed and marketed.

General-purpose hardware as used in the IT field and virtualization technology based on hardware accelerators customized for wireless processing are starting to be applied to wireless base stations in a network transformation toward the virtualization of a radio access network (RAN), or virtualized RAN (vRAN) for short. The use of general-purpose hardware in virtualization and cloudification and the effective use of vRAN technology are expected to reduce infrastructure investment. With the idea of maximizing the benefits of vRAN, in February 2021 NTT DOCOMO jointly launched with other partners having various vRAN-related technologies an initiative for collaborative co-creation solutions called the 5G Open RAN Ecosystem (OREC) [1]. Since NTT DOCOMO had already introduced virtualization technology for its core network, such development and operational experience is becoming a key asset and input for the introduction of vRAN.

In this article, we describe the technical issues that must be addressed to achieve vRAN and NTT DOCOMO's approach to addressing those issues. We also describe the importance of standardization and initiatives toward standardization of vRAN.

2. RAN virtualization technology

2.1 Network functions virtualization technology

Network functions virtualization (NFV)*1 relies on the offering of a virtualization layer on top of general-purpose hardware based on IT virtualization and orchestration technology*2. The virtualization layer exposes virtual resources that can be composed and used to run applications, including network function software applications. NFV transforms the way network operators can deliver communication services; such communication services had conventionally been provided by dedicated hardware and software optimized for satisfying the high-reliability and high-performance requirements of systems operated by telecom operators. NFV supports the decoupling of hardware and software, which in turn enables the rapid deployment and provision of new network functions simply by deploying and/or updating the software that implements the network function behavior. The use of open source software and agile development techniques cultivated in the IT field can also provide additional benefits such as shortening the time-to-service to service launch.

NTT DOCOMO has been researching and developing NFV technology and promoting its standardization by the European Telecommunications Standards Institute (ETSI) NFV Industry Specification Group since the first half of the 2010s. In parallel, NTT DOCOMO has been introducing virtualization technology into the core network of the commercial network since FY2015. By the end of FY2020, the virtualization rate of core network equipment for LTE and beyond exceeded 50%, and the 5G core network deployed in FY2021 became fully virtualized [2]. NTT DOCOMO has reaped the following four benefits through NFV in the core network [3].

(1) Improved economic performance of network function equipment

There had been equipment and telecom software for each equipment vendor and network function, and it had been necessary to perform maintenance for each piece of equipment and each package of telecom software. However, NFV technology enabled NTT DOCOMO to run telecom software from multiple vendors on a unified virtualization platform, which improved cost performance, unified and simplified operations and maintenance, and enabled the use and sharing of low-cost general-purpose hardware.

(2) Quick deployment of new services

NFV shortened the time-to-service by eliminating the preparation and installation of dedicated hardware when introducing a new service.

(3) Improved communication service connectivity during times of congestion

In the event of congestion, e.g., caused by a natural disaster or a sudden concentration of traffic, NFV enabled the network-functions capacity to be automatically increased in a short period, thereby improving connectivity for consumer communication services.

(4) Improved reliability of communication services

The use of low-cost hardware enabled the construction of redundant hardware configurations on which telecom software could be automatically loaded to maintain normal operations upon detecting hardware failures. This enabled the automatic recovery of network functions in a short time (auto-healing), and such a capability eliminated the need to dispatch personnel for immediate maintenance work on site while achieving high maintainability and high reliability of consumer communication services.

2.2 Expected effects of introducing vRAN

Recent advances in IT virtualization technology, general-purpose hardware, and hardware accelerators have broadened the domains in which virtualization can be applied. It has even become possible to apply virtualization to baseband*3 processing on the radio layer, where high-level service requests such as high performance and sensitive real-time requirements are even more demanding than those of the core network. With this in mind, a number of domestic and overseas operators have begun initiatives toward the deployment and commercialization of vRAN using virtualization technology. The expected effects of introducing vRAN are summarized as follows (**Fig. 1**).

(1) Optimal combination of RAN solutions by decoupling hardware and software

- Improved cost performance of network functions through the use of general-purpose hardware
- Improved performance and reduced power consumption by using up-to-date hardware
- Extension of RAN capabilities and services by focusing on software updates only

---

*1 NFV: Principle of separating network functions of a telecommunications carrier network from general-purpose hardware through virtualization technology.

*2 Orchestration technology: Technology for managing and mediating the connectivity of essential resources and networks to automate the operation and management of applications and services.

*3 Baseband: The band of information signals before/after converting to/from the radio frequency band on the transmitting/receiving side in wireless communications. It is usually a low frequency band achieved through digital signal processing.

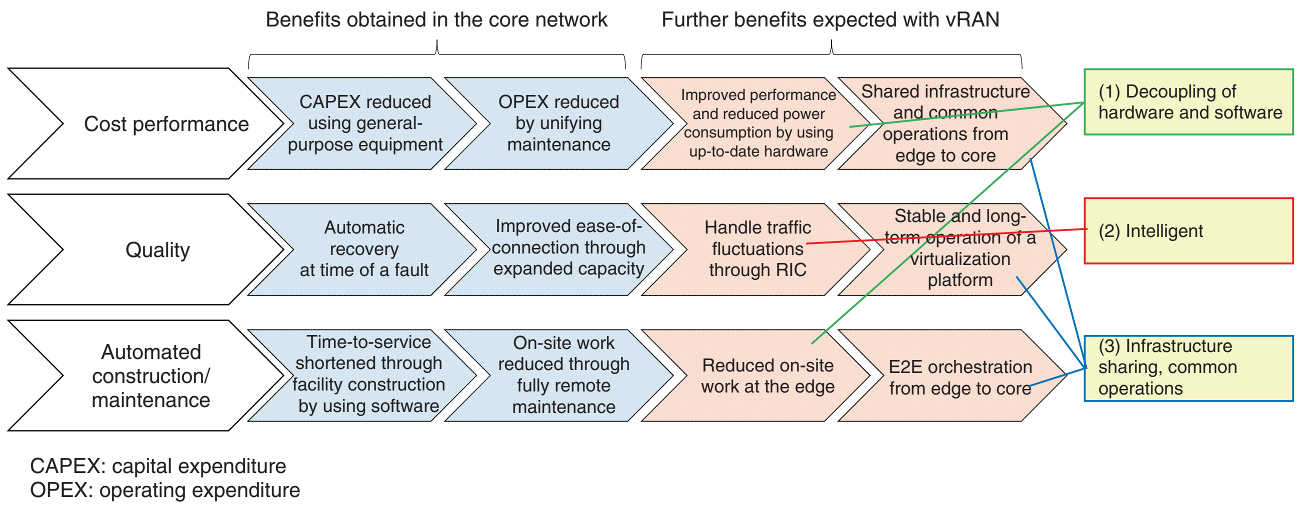


Fig. 1. Expected effects of introducing vRAN.

- Easy introduction of new vendors

(2) Simple and intelligent RAN operation and maintenance through virtualization and automation

- Facility construction with software shortens the time-to-service and improves flexibility of deployment by making maximum use of an already installed hardware and software resource pool*4
- Expanded scope of remote maintenance work and reduced on-site work by leveraging remote software management capabilities
- Improved connectivity of communication services by expanding their capacity to flexibly address fluctuating traffic demand
- Improved reliability by automatically recovering from hardware and software failures in a short period
- Handling of traffic fluctuations by using a RAN Intelligent Controller (RIC)

(3) Infrastructure sharing and unified operations from the edge to core network

- Achieving a unified virtualization platform and unified operations wherein the virtualization of RAN applications bridges the edge and core network and enables end-to-end (E2E) 5G network deployment
- Stable and long-term operation of a virtualization platform
- Provision of orchestration on an E2E basis

2.3 Issues in meeting vRAN requirements through NFV technology

To maximize the benefits of vRAN described above, it is necessary to resolve the technical issues that arise from specific RAN requirements in addition to those that arose in the virtualization of the core network.

Issues 1: Issues in achieving an optimal combination of solutions by decoupling hardware and software

(1) Support of high-performance and real-time processing executed using general-purpose hardware and software

In conventional base-station equipment, computationally intensive processing on the radio layer requiring real-time characteristics became feasible through dedicated hardware such as an optimized field-programmable gate array and application-specific integrated circuit devices and dedicated software. Instead, in vRAN, conventional base-station-equipment architecture featuring such tightly coupled hardware and software must be achieved using general-purpose hardware and software running on that hardware. It is necessary to reduce fluctuations in the time sensitivity of processes on the radio layer that affect base-station performance and achieve low latency. To this end, optimized general-purpose

*4 Resource pool: A grouping mechanism for bundling and managing infrastructure resources such as computers, central processing units, and memory, and arranging them for subsequent use.

hardware must be selected, high-performance and time-sensitive software must be implemented to meet such requirements, and the necessary execution environment fulfilling all these requirements must in turn be supported by the virtualization platform. For functions and components having particularly severe processing requirements, studies must also be conducted on optimizing software processing even further or on offloading some of the radio-layer processing to a hardware accelerator. Finally, high-performance and high real-time sensitive processing must be achieved while dealing with the problem of tightly coupled hardware and software that has been (until now) typically assumed for fulfilling extreme performance requirements.

(2) Construction of a mechanism supporting the installation of base-station equipment

In virtualizing the core network, it has been possible to consolidate and install relevant core network functions at few centralized sites (datacenters). However, in the case of RAN, many more compact and distributed base-station sites (e.g., placed in a room within a building or outdoors) are necessary to cover the required radio coverage. Therefore, required performance, installation space, and power limitations greatly differ depending on the involved network functions to deploy at a target location. General-purpose hardware also needs to be able to withstand severe conditions such as high dust environments and extreme temperatures. In short, a mechanism is needed that can support scalable and flexible installation formats under such a variety of installation environments while ensuring an efficient decoupling of software and hardware.

Issues 2: Issues in achieving simple and intelligent RAN operation and maintenance through virtualization and automation

(1) Management of virtual resources provided to RAN applications

In conventional base-station equipment, base-station functions are coupled to specific hardware, so hardware and software can be managed as a set and maintenance can be conducted on that set. In vRAN, in which virtual resources are provided from multiple hardware devices, the RAN application management system dynamically requests the virtual resources needed and runs RAN applications on those virtual resources. In the latter, software and virtual resources are managed independently. To maintain base-station functions under these conditions, there is a need to associate software, virtual resources (logical), and hardware (physical). There is also a need to handle new hardware components, such as hardware accelerators, in a unified manner together with other virtual resources.

(2) Support of maintenance and operation use cases

With conventional base-station equipment, operation methods such as hardware configuration, installation, and maintenance differ among vendors. With virtualization, if these different operation methods were to be simply implemented without making any changes, and if an attempt would be made to support them all, the volume of development and amount of testing would certainly increase and not be bearable by the network operator. Therefore, there is a need to both promote the integration of more automation into the operations support systems and standardize solutions that clearly identify the functional splits and boundaries of systems to integrate. This is a necessary step to ensure unifying operational and maintenance procedures for all systems involved in vRAN deployment.

(3) Reduction in on-site work and manual operations

An entire workflow must be designed to achieve automation in processes ranging from the construction of base stations to be distributed nationwide and creation of virtualization platforms. It is particularly important to minimize on-site work and expand the scope of automation by expanding the scope of software control when migrating to vRAN. For this reason, it is essential that construction work include the kitting*5 of general-purpose equipment, provision of spare parts to be standardized to reduce on-site work, and various types of interfaces to be set up to operate and manage this equipment remotely.

Issues 3: Issues in achieving infrastructure sharing and unified operations from edge to core network

(1) Support of diverse applications by using a unified virtualization platform

To support the deployment of new network functions in the future, a vRAN virtualization platform will have to support the deployment and execution of base-station functions, core-network functions, and higher-level layer services. The execution-environment capability of the platform will require both supporting virtual machine (VM)-based network

*5 Kitting: In this article, it describes the act of installing applications in equipment such as servers and making various types of settings and registrations to put those applications in a state ready to use.

function deployments as well as container[*6]-based network function deployments, the use of which has increased. To make operations more efficient, there will also be a need for a virtualization environment that can support application requirements in each network domain.

(2) Achieving long-term operation of virtualization platforms distributed nationwide

With vRAN commercialization, we can expect many small-scale virtualization sites to be distributed nationwide, so how to go about maintaining and managing those platforms is an important issue. It will be necessary to upgrade and maintain a virtualization platform periodically to cope with the introduction of new general-purpose hardware and new hardware accelerators, need for security updates, evolution of applications, etc. All version upgrades of virtualization platforms must also be carried out without interrupting any communication services. Taking into account recent development trends in container platforms, it will also be necessary to make short-term upgrades including those to the platforms, but it is unrealistic to expect an operator to carry out large-scale upgrades and development, etc. continuously. A mechanism that can execute long-term and stable platform operations would therefore be desirable.

3. Addressing open issues: approaches

At NTT DOCOMO, to address the issues described above and achieving its vRAN goals, two approaches will be adopted: vRAN standardization activity in the O-RAN ALLIANCE[*7] and ETSI NFV and interoperability testing in a multi-vendor environment through OREC.

For Issues 1, there will be a need to study, in cooperation with RAN-application vendors and general-purpose hardware vendors, the optimal implementation of applications and optimization of their performance. NTT DOCOMO will promote these activities within OREC, which are described later in this article.

For Issues 2 and 3, recognizing the prime importance of standardizing interfaces and information models[*8] between vRAN and associated systems and in achieving unified operation and maintenance of vRAN, NTT DOCOMO will promote the standardization of vRAN. Standardization will enable various types of products provided by multiple vendors to be combined. If the support period for a particular product in a product configuration is coming to an end, a different product conforming to the same standard on the market can be used to minimize the impact on other products in that configuration and on surrounding systems. To avoid fragmentation of specifications related to NFV with the same objective among standardization organizations within the industry and make maximum reuse of the NFV-related specifications that NTT DOCOMO has been actively pushing for standardization, further coordination among standardization organizations will be promoted.

Standardization activities related to the virtualization and orchestration support for vRAN are progressing in the O-RAN Work Group 6 (WG6). The standardization items that NTT DOCOMO must take up on the basis of the overall architecture of O-RAN (**Fig. 2**) are summarized below.

- Interface between Service Management and Orchestration (SMO) and multi-vendor applications implementing the O-RAN central unit - control plane (O-CU-CP), O-RAN central unit - user plane (O-CU-UP), Near-real-time (Near-RT) RIC, and O-RAN distributed unit (O-DU) functions (O1 in Fig. 2)
- Interfaces between SMO and a multi-vendor virtualization platform, referred to as O-Cloud, which contains deployment management services (DMS)/infrastructure management services (IMS) functions (O2dms/O2ims in Fig. 2)
- Unified operations through SMO and multiple virtualization platforms (instantiation[*9], scaling[*10], healing[*11], termination[*12])
- Common information models for SMO-managed vRAN applications, virtualization platforms, and packages toward automated operations

*6 Container: A type of computer virtualization technology in which a dedicated area (the container) is created on a host operating system (OS), and the necessary application software is run in the container.

*7 O-RAN ALLIANCE: An industry and standardization group founded by NTT DOCOMO and leading overseas operators in February 2018 for achieving an open and intelligent RAN for the 5G era.

*8 Information modeling: The modeling of a system to make it easier to handle by corresponding management systems.

*9 Instantiation: Refers to the generation of a VM. Indicates a sequence of operations from installing a VM on a physical machine, making virtual network settings, and booting the VM to start up the software and putting it into a usable state through various settings.

*10 Scaling: The optimization of processing power by increasing or decreasing VMs that configure communications software whenever processing power of that software is insufficient or excessive in accordance with hardware or VM load conditions.

*11 Healing: A procedure for restoring communications software to a normal state in the event of a hardware or VM failure by moving the VM to (or recreating the VM on) hardware operating normally.

*12 Termination: The process of bringing a VM to a close.

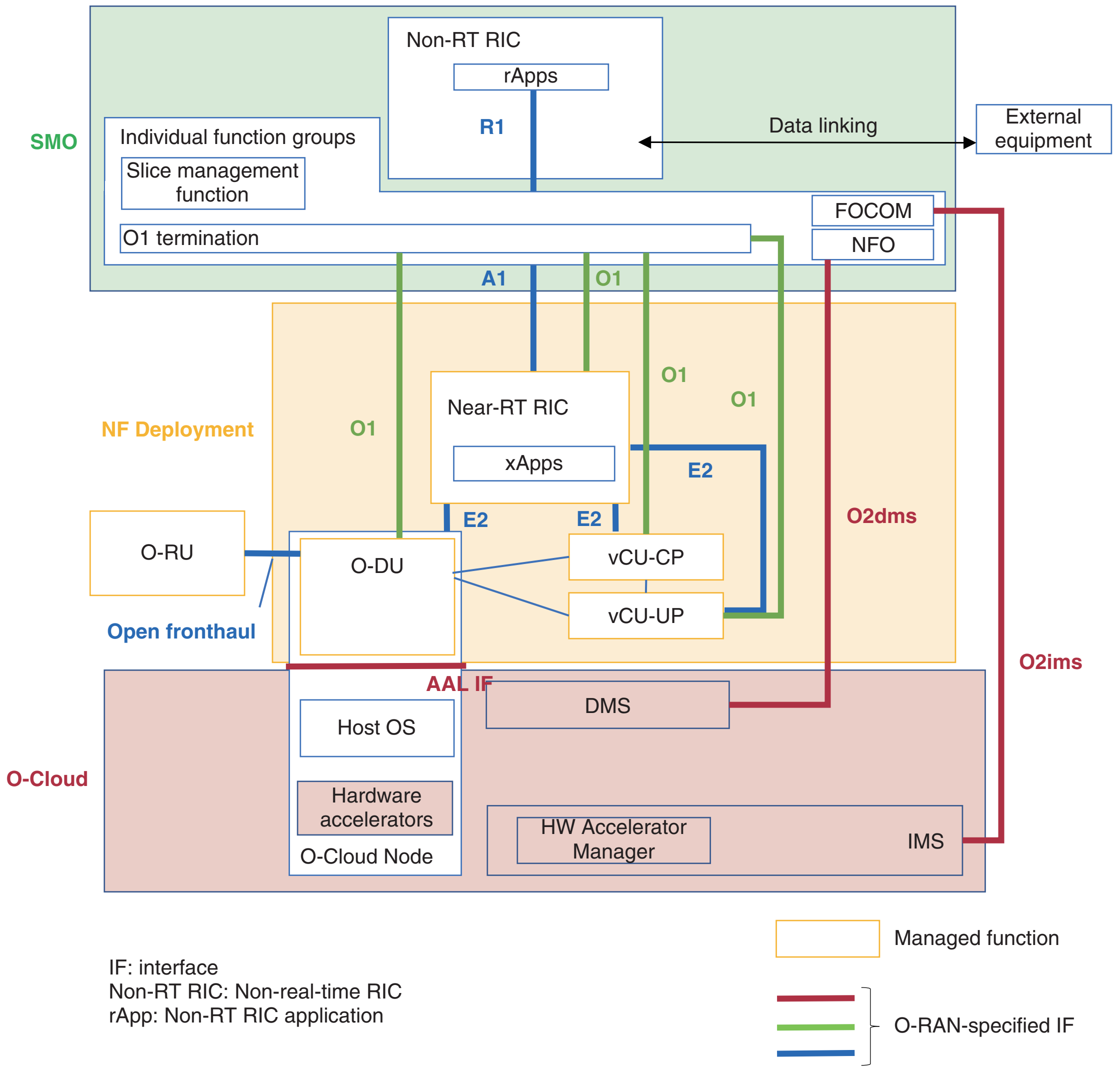


Fig. 2. Overall O-RAN architecture.

- A unified E2E virtualization platform that can support core network functions, vRAN, multi-access edge computing (MEC)*13, etc.
- Stable interfaces maintaining backward compatibility with current operation systems and related interconnected systems and information flow between systems
- Functional splits and interface specifications that show high extensibility while making maximum reuse of the operator's assets

4. Latest developments in O-RAN WG6 standardization

At O-RAN, standardization related to virtualization and orchestration are mainly being conducted in the Cloudification & Orchestration WG (O-RAN WG6). O-RAN WG6 specifications and reports consist of General Aspects and Principles (GAP) defining overall concepts; Orchestration Use Cases (ORCH-UC), which defines use-case and information flow between the specified functions; and other interface specifications that specify data models and protocols. Details on each of the above specifications and reports are given as follows.

*13 MEC: An approach for installing servers, storage devices, etc. at locations nearer to users within a mobile communications network to provide low-latency services with high real-time characteristics.

(1) GAP

In WG6, function names such as O-Cloud are defined and the relationships among those functions are described. Two types of GAP documents are delivered: O2GAP, specific to concepts related to the O2 interface, and AAL-GAP, specific to concepts about the Acceleration Abstraction Layer (AAL)[*14].

(2) ORCH-UC

For O-Cloud, Network Function (NF), and Near-RT RIC application (xApp), typical use cases related to provisioning[*15], fault management, and performance management are specified.

(3) Interface specifications

As specified by O-RAN WG6, O2 is the reference point between SMO and O-Cloud and AAL is the abstraction layer for the provision of hardware accelerators by O-Cloud to NF Deployment. The interface specifications cover the reference point and abstraction layer on various further detailed aspects. For instance, the O2ims interface document specifies the interfaces exposed by the O-Cloud IMS on the O2; the AAL forward error correction (FEC) specification and AAL High-PHY detail an AAL profile for FEC and AAL profile for the hardware acceleration of High-PHY functions, respectively. The specification of the protocol and data models for the O2dms interface exposed by the O-Cloud DMS on the O2 is in progress.

The constituent elements of vRAN are NF Deployment that specifies the virtual resources of vRAN applications, O-Cloud that serves as the virtualization platform providing virtual resources to NF Deployment, and SMO that operates and manages NF Deployments and O-Cloud (Fig. 2).

4.1 SMO

The SMO functional block consists of Federated O-Cloud Orchestration and Management (FOCOM) and Network Function Orchestration (NFO), which operate, manage, and orchestrate O-Cloud and NF Deployment as logical functions, respectively. FOCOM is capable of inventory management and alarm management for O-Cloud, while NFO is capable of lifecycle management, alarm management, and performance management of NF Deployments on O-Cloud.

4.2 O-Cloud

O-Cloud consists mainly of the following logical functions: the O-Cloud Node, which provides virtual resources to NF Deployment; DMS, responsible for the management of NF Deployments; IMS, which manages O-Cloud Node and other O-Cloud infrastructure components including DMS instances; and HW (Hardware) Accelerator Manager that manages hardware accelerators. The O-Cloud Node, in turn, consists of computing resources, network resources, and storage resources, where computing resources include AAL functions (described later). Possible solutions of DMS are assumed to be, for instance, Kubernetes® or OpenStack®. As previously introduced, DMS exposes its services and connects to SMO via the O2dms interface, while IMS connects to the SMO over the O2ims interface. IMS mainly provides inventory management, alarm management, and performance management for the O-Cloud Node and deployment of DMS instances.

4.3 Hardware accelerator

One of the key specifications O-RAN WG6 is focusing on is the support of hardware accelerators, which are typically required to implement vRAN. In this context, AAL includes a hardware accelerator as physical equipment, AAL profile for achieving various types of processing using the hardware accelerator, and AAL interface (AALI) between AAL profile and vRAN application. The HW Accelerator Manager manages AAL profiles supported by hardware accelerations that are part of O-Cloud.

Two major acceleration types for offloading radio processing to the hardware accelerator are look-aside acceleration and inline acceleration. The former offloads parts of radio layer 1 and particularly coding/decoding having high processing load, while the latter offloads all of radio layer 1. For either type, specifications are progressing to facilitate virtualization and orchestration.

Going forward, the plan at O-RAN WG6 is to complete O2dms interface specifications, AAL use-case specifications, and a vRAN application package and to deliver specifications with clear functional splits and function definitions, which are key for network operators who consider fully interoperable multi-vendor vRAN deployment.

*14 AAL: A layer decoupling the hardware and software of hardware accelerators installed on the O-Cloud. It is being defined by O-RAN to enable hardware accelerators from different vendors to be combined with software and accessible by specifying standard formats, interfaces, etc.

*15 Provisioning: The work of setting hardware and software for running network resources deemed necessary for providing services.

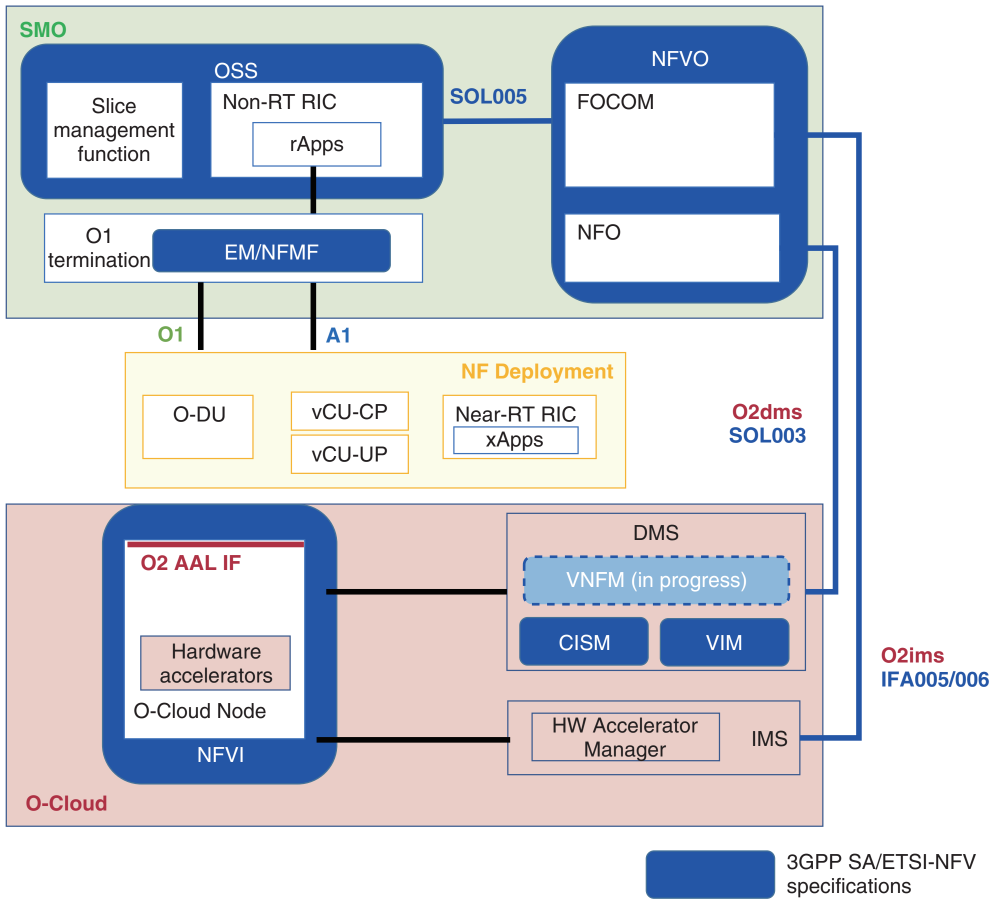


Fig. 3. Relationship between O-RAN and ETSI-NFV specifications.

5. NTT DOCOMO's standardization activities

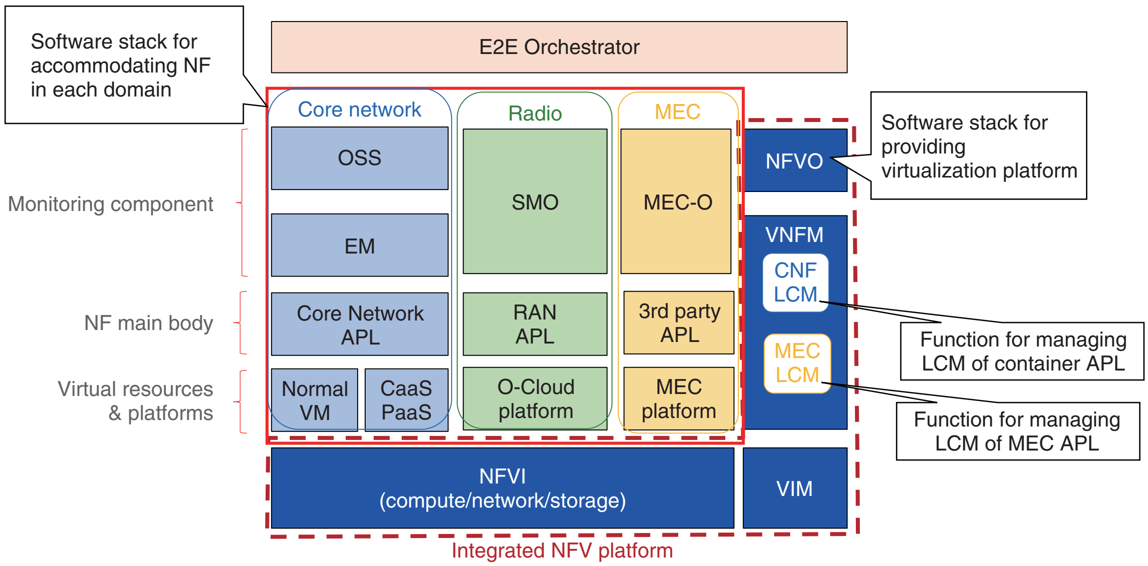
NTT DOCOMO has been actively promoting the standardization of core network virtualization at ETSI NFV. To maximize the reuse of existing ETSI NFV specifications and avoid fragmentation of NFV standards and specifications with the same objective within the industry, NTT DOCOMO has been studying and comparing O-RAN specifications and ETSI NFV specifications, as shown in **Fig. 3**. Specifically, O-RAN WG6 is specifying the O2dms interface while ETSI NFV is analyzing any gaps between O-RAN specifications and ETSI NFV specifications on the basis of ETSI GR NFV-IFA 046.

The goal is to standardize an integrated and unified NFV platform able to cover from RAN to the core network, as shown in **Fig. 4**. For the distributed equipment making up a base station, requirements and features that are unique to each functional block need to be considered, such as the limited length and constrained delay in the fronthaul between the base station and radio unit (RU)*16. NTT DOCOMO will therefore continue to study optimal deployment scenarios taking into account such requirements and features. NTT DOCOMO considers the following points to be important in promoting O-RAN and ETSI-NFV standardization.

5.1 Positioning of ETSI-NFV-defined NFV Orchestrator in SMO

In O-RAN specifications, SMO becomes a huge functional block that includes the functionality specified by the 3rd Generation Partnership Project Service & Systems Aspects WG5 (3GPP SA5) regarding

*16 RU: As one unit of equipment making up a base station, it converts digital signals to be transmitted to radio signals and received radio signals to digital signals. It also amplifies transmission power, transmits/receives radio signals at antenna elements, and executes processing necessary to generate beams in massive multiple-input and multiple-output.



Software stack: An accumulation of protocols and software having interoperability that serves as a single system or function

APL: application
CaaS: Container as a Service
CNF: cloud native network function
LCM: lifecycle management
PaaS: Platform as a Service

Fig. 4. Integrated NFV platform from RAN to core network.

an operations support system (OSS) and Element Manager (EM)/Network Function Management Function (NFMF) as well as the NFV Orchestrator (NFVO) (Fig. 3) functionality defined by ETSI NFV. Therefore, to achieve an SMO supporting a multi-vendor environment, dividing up the SMO framework into more granular functional blocks in the standardization process is considered essential. Since the OSS functional blocks differ between operators in terms of existing assets, operational workflow, and external equipment to be connected to, current discussions are pointing out the difficulty of handling all the relevant functionality only by SMO; as a result, there exists the possibility that typical OSS functions will not be included in SMO solutions.

While O1 termination is assumed to be EM/NFMF, as currently defined by 3GPP SA5, the deployment scenario has been to provide NF and EM as a single set from a vendor, but simplifying this process by deploying general-purpose EM (Generic EM) is desirable. In terms of virtualization, SMO would then take on the role of a Generic EM functional block that terminates O1 and a NFVO functional block that terminates O2.

5.2 Relationship between O-Cloud and Container Infrastructure Service Manager/Virtual Infrastructure Manager of ETSI NFV

The O2dms interface is currently being specified in O-RAN, and the finalization of detailed specifications is forthcoming. In terms of ETSI NFV specifications, two platforms having a deep relationship with DMS are being specified: Container Infrastructure Service (CIS) and CIS Manager (CISM) as a container platform and Virtual Infrastructure Manager (VIM) and NFV Infrastructure (NFVI) as a VM platform. The management entities (i.e., CISM and VIM) are considered to correspond to DMS (Fig. 3) functionality. For O2dms, studies are now being conducted on the use of ETSI NFV specifications such as ETSI GS NFV-SOL 018 [4], which specifies the interfaces produced by the CISM; ETSI GS NFV-SOL 014 [5], which specifies the data model of virtualized resource descriptors that can be used over the interfaces produced by the VIM; and ETSI GS NFV-SOL 003 [6], which specifies the interfaces produced by the Virtual Network Function Manager (VNFM) regarding VNF management, and that can be leveraged as an abstraction above the other aforementioned interfaces. These studies will determine

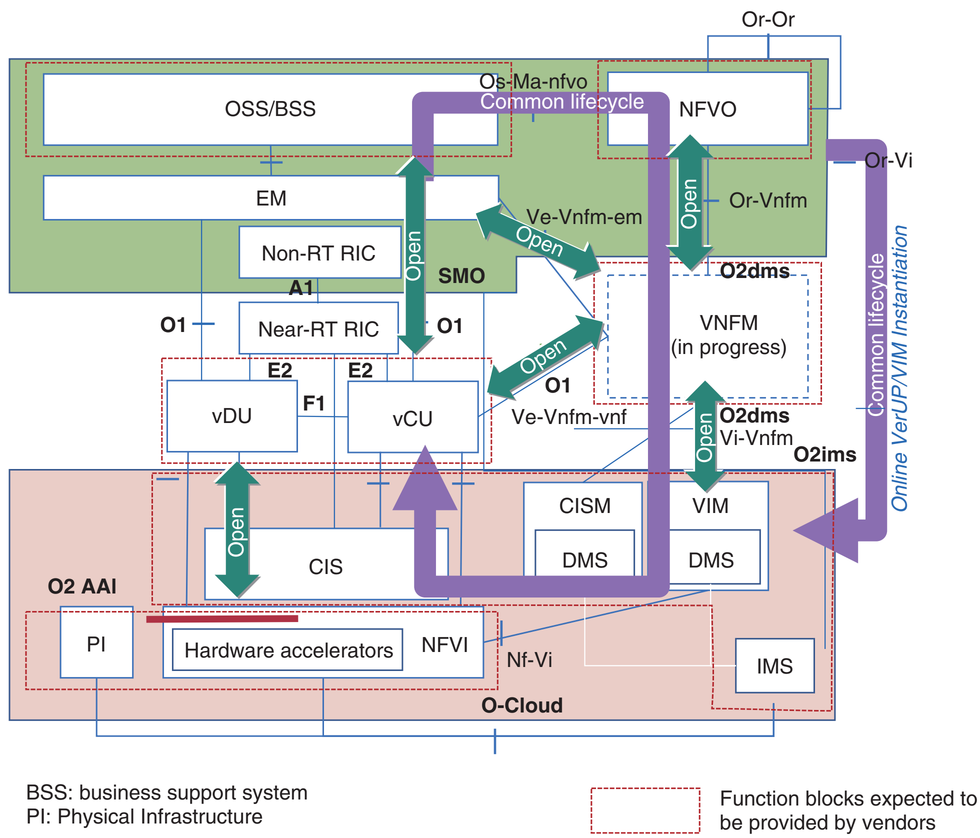


Fig. 5. Open networking targeted by NTT DOCOMO.

whether VNFM functionality will belong to SMO or O-Cloud. The ETSI NFV will study the entire functionality of IMS and extend the ETSI NFV specifications so that NFV can support IMS or a related function.

6. Activities at OREC

Since there are many implementation-dependent functions not specified in standards and optional functions that are specified in standards, there is a need to align options and interfaces to combine a variety of products and conduct testing to develop vRAN solutions that can provide 5G services on a commercial level. In addition to promoting standardization at O-RAN and ETSI NFV, NTT DOCOMO is focusing its attention on the following points at OREC.

(1) Creation of guidelines for operational workflow and network design toward the provisioning of vRAN solutions

Considering that a variety of products can be combined from diverse vRAN application vendors, hardware accelerator vendors, and O-Cloud vendors, NTT DOCOMO is creating guidelines on achieving multi-vendor interoperability to cover various types of operational workflow and network designs. While a redundant configuration for the network and management nodes is essential in a cloud environment, it may not be possible to achieve suitable redundancy at a base-station site due to the installation environment, which would mean many cases in which the network disconnects or only a management node goes down. Due to such issues, the need arises for a variety of options, so studies in design guidelines are also proceeding from this viewpoint. The architecture of specific open-network initiatives targeted by NTT DOCOMO is shown in **Fig. 5**. With a focus on VNFM, NTT DOCOMO will work to maximize the use of ETSI NFV specifications by making various interfaces open with the end goal to achieve interoperability between vRAN applications and virtualization platforms delivered by multiple vendors.

(2) Provisioning of testing environments for end solutions and E2E function/performance testing

To use a variety of new technologies in vRAN, operational testing in terms of both functionality and maintenance is essential. To facilitate the use of hardware accelerators in new ways, performance testing will also be necessary. Testing of optimal hardware configurations at small-footprint base-station sites will likewise be necessary. NTT DOCOMO is now working on providing these testing environments to cover E2E testing from terminals to core networks.

Three combinations of vRAN applications, virtualization platforms, general-purpose servers, and hardware accelerators are being tested at the OREC Shared Open Lab. Verification of communications in a 5G stand-alone configuration has been completed, for look-aside hardware accelerators, the commercialization of which is in progress, as well as for inline hardware accelerators. NTT DOCOMO plans to accelerate these testing activities toward commercial-level quality testing.

7. Conclusion

This article provided an overview of NFV, described the effects of introducing virtualization in RANs, and the issues surrounding implementation, and presented NTT DOCOMO's approach to addressing these issues. The article also described the current status of standardization at O-RAN, NTT DOCOMO's standardization activities, and initiatives at OREC.

NFV has already produced positive effects in the deployment of NTT DOCOMO's core network. To adapt such a vision and assets to make it extensible to also cover the virtualization of RAN, NTT DOCOMO is working and promoting standardization and OREC activities as two key approaches. Both standardization and OREC are evolving and aiming for a world that can achieve optimal combinations of solutions applicable to a variety of future network requirements. As a means to this goal, NTT DOCOMO will promote vRAN development together with its partners and conduct testing at the OREC Shared Open Lab. It will also contribute to standardization by expanding use cases and specifying interfaces so that functions that are key to operators can be provided and integrated easily, yet guaranteeing stability of the system for long-term telco support.

References

[1] D. Hiratsuka, K. Kuriu, A. Umesh, and H. Mori, "Initiatives toward Open RAN," NTT Technical Review, Vol. 20, No. 11, pp. 46–51, Nov. 2022.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr202211fa6.html

[2] H. Tamura, Y. Kuno, and Y. Suzuki, "Migration to ETSI NFV Stage 3 Specification-compliant Multivendor MANO Configuration on Network Virtualization Platform," NTT DOCOMO Technical Journal, Vol. 23, No. 4, pp. 74–86, Apr. 2022.

[3] T. Kamada, Y. Kuno, H. Tamura, and H. Iwamiya, "Practical Implementation of Virtualization Platform in NTT DOCOMO Network," NTT DOCOMO Technical Journal, Vol. 18, No. 1, pp. 20–28, July 2016.

[4] ETSI GS NFV-SOL 018: "Profiling specification of protocol and data model solutions for OS Container management and orchestration."

[5] ETSI GS NFV-SOL 014: "YAML data model specification for descriptor-based virtualised resource management."

[6] ETSI GS NFV-SOL 003: "RESTful protocols specification for the Or-Vnfm Reference Point."

Trademark notes

All company names or names of products, software, and services appearing in this article are trademarks or registered trademarks of their respective owners.

Shinji Mizuta
Manager, Radio Access Network Development Department, NTT DOCOMO, INC.

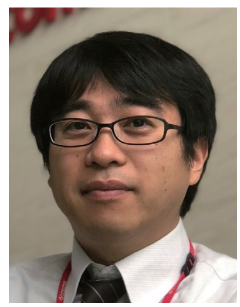
He received a B.E. and M.E. from Kyushu University, Fukuoka, in 2000 and 2002. He joined NTT DOCOMO in 2002 and has been involved in the research and development of 3G, 4G, and 5G radio base station. He is currently working on the vRAN project.

Anil Umesh
Senior Manager, Radio Access Network Development Department, NTT DOCOMO, INC.

He received a B.S. in engineering from Case Western Reserve University, USA, in 2002. He joined NTT DOCOMO in 2002 and has since been involved in R&D activities related to wireless access technologies for HSPA (High Speed Packet Access), LTE, LTE-Advanced, and 5G. He has been the O-RAN Alliance Working Group 4 co-chair since the start of O-RAN ALLIANCE in 2018 and has chaired open fronthaul topics.

Yoshihiro Nakajima
Manager, Core Network Development Department, NTT DOCOMO, INC.

He received a B.E., M.E., and Ph.D. from University of Tsukuba in 2003, 2005, and 2008. He was a research fellow of the Japan Society for the Promotion of Science from 2005 to 2008. He was a researcher at NTT Network Innovation Laboratories from 2008 to 2017. Since joining NTT DOCOMO in 2017, he has been involved in the development and standardization of network virtualization for mobile core networks. He served as an Industry Specification Group (ISG) vice chair of ETSI NFV from 2019 to 2022. He has been the ISG chair of ETSI NFV since September 2022.



Yuya Kuno
Developer, Core Network Development Department, NTT DOCOMO, INC.

He received a B.E. from Keio University, Kanagawa, in 2004. He was a developer of OSS in NTT DOCOMO to support operation of 2G/3G/4G core and radio network functions from 2004 to 2014. He was a developer of MANO (Management and Orchestration) in NTT DOCOMO to virtualized EPC (Evolved Packet Core) and IMS network. He has been involved in ETSI NFV standardization since 2014 and served as rapporteur of several NFV SOL specifications in ETSI ISG NFV SOL WG.

Initiatives toward Intelligent RAN

Taichi Katsuragawa, Akihiro Kawana, Yoshio Inoue, Takahiro Tateishi, Eina Hashimoto, and Takumi Fujitsuka

Abstract

The network of the fifth-generation mobile-communications-system era must be able to support many and varied applications, which means ever increasing complexity. It will therefore not be easy to control and optimize network operations by manual means as done in the past. NTT DOCOMO is developing radio access network (RAN) Intelligent Controller (RIC) technology now being standardized at the O-RAN ALLIANCE with the aim of achieving more autonomous and automatic RAN operations, that is, intelligent RAN, by using machine learning and other forms of artificial intelligence, big data, etc. This article describes the state of RIC standardization at the O-RAN ALLIANCE and presents NTT DOCOMO's approach to use cases and a roadmap toward intelligent RAN.

Keywords: RIC, AI/ML, automation

1. Introduction

The mobile network of the fifth-generation mobile communications system (5G) is expected to provide services that can satisfy diverse requirements and conditions such as high speeds and large capacity, low latency, and massive connectivity. To meet these advanced service requirements, mobile network operators have been working continuously to enhance radio access network (RAN) functions and expand the scale of the network; however, RAN design and operation have become increasingly complex. In response, a technology called self-organizing network (SON)*1 has been standardized at the 3rd Generation Partnership Project (3GPP) as a means of automatically constructing networks, optimizing coverage areas and operation parameters, and recovering from faults to ease the burden of RAN construction and operation on operators. The evolution of artificial intelligence (AI) technology is spurring demand for automation achieved through intelligent systems using big data and AI with machine learning (AI/ML). From the viewpoint of RAN control, AI/ML will enable proactive control, which is expected to improve RAN performance and the customer's degree of satisfaction.

At the O-RAN ALLIANCE, standardization of architecture and various control interfaces is underway to achieve operation and control using big data and AI/ML toward intelligent RAN. NTT DOCOMO is actively conducting technical studies of this technology.

In this article, we describe the architecture and functions and control procedures of various interfaces for achieving intelligent RAN, the specifications of which are now being studied at the O-RAN ALLIANCE. We also give an overview of NTT DOCOMO's approach toward a roadmap for achieving intelligent RAN, discuss issues that need to be addressed, and present a future outlook.

2. Overview of RIC standardization at O-RAN ALLIANCE

2.1 RIC architecture

In RAN architecture at the O-RAN ALLIANCE, a RAN Intelligent Controller (RIC) is defined as a

*1 SON: A network installed with functions to self-configure and self-optimize its parameters.

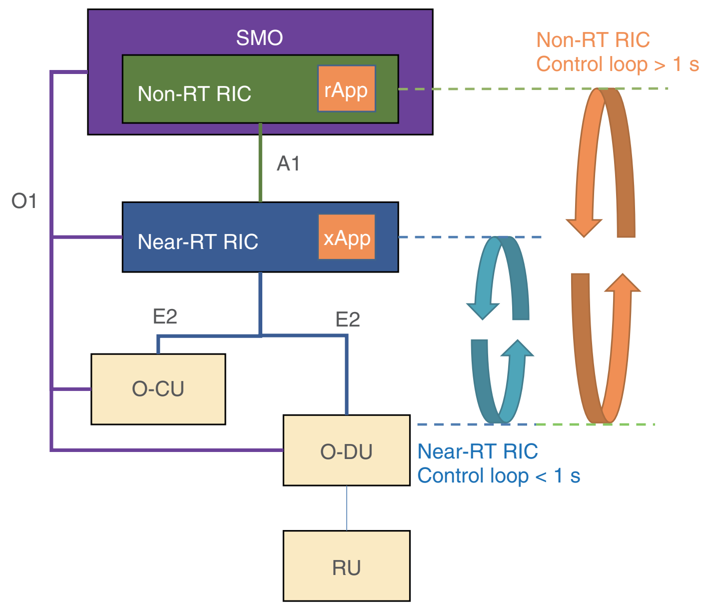


Fig. 1. RAN architecture at O-RAN ALLIANCE.

logical node for executing base-station parameter design and settings and automating and optimizing operations.

As shown in **Fig. 1**, there are two types of RICs: the Non-real-time (Non-RT) RIC and Near-real-time (Near-RT) RIC. The Non-RT RIC is situated within the Service Management and Orchestration (SMO) block that executes RAN monitoring/maintenance and orchestration. It connects to the Near-RT RIC via an A1 interface, and the Near-RT RIC connects to E2 nodes, such as the O-RAN central unit (O-CU)*2 and O-RAN distributed unit (O-DU)*3, via an E2 interface. These E2 nodes and the Near-RT RIC connect to SMO via an O1 interface. Based on RAN architecture now under study at the O-RAN ALLIANCE, intelligent control can be achieved in a variety of formats by combining the control interfaces to be used with an appropriate arrangement of AI/ML functions.

(1) Non-RT RIC

The Non-RT RIC links with the function that provides operation administration and maintenance (OAM) services within SMO and uses the O1 interface to collect various types of data from the E2 nodes such as performance management (PM) counter, fault management (FM) data, and trace management (TM) data accumulated within those nodes. It can optimize parameter settings through advanced analysis using AI/ML and reflect those settings optimized to the radio environment, traffic load, etc. in the E2 nodes via the O1 interface. The Non-RT RIC can also generate policies governing RAN control and notify the Near-RT RIC of those policies via the A1 interface. These control processes are executed in a relatively long control loop greater than 1 s.

(2) Near-RT RIC

The Near-RT RIC collects information from the E2 nodes via the E2 interface and can reflect the results of internal analysis in the control of those nodes in accordance with the policies notified by the Non-RT RIC. It executes high-speed control in a control loop of from 10 ms to 1 s by directly connecting with the E2 nodes via the E2 interface.

(3) rApp

The Non-RT RIC uses an application called the Non-RT RIC application (rApp) for analyzing various types of information and generating policies. The rApp takes on architecture independent of the Non-RT RIC framework*4 and connects with that framework via the R1 interface.

*2 O-CU: Equipment that controls radio resources of the radio base station.

*3 O-DU: A function executing real-time Layer 2 functions, etc. of the radio base station.

*4 Framework: Software that encompasses functionality and control structures generally required for software in a given domain. In contrast to a library in which the developer calls individual functions, code in the framework handles overall control and calls individual functions added by the developer.

(4) xApp

The application that analyzes various types of information and executes control processes on the Near-RT RIC framework is called the Near-RT RIC application (xApp). In the Near-RT RIC as well, the framework and application are separated but interconnected by Near-RT RIC application programming interfaces (APIs) being standardized at the O-RAN ALLIANCE.

2.2 Interfaces specified at O-RAN ALLIANCE

(1) O1 interface

The O1 interface is used by SMO to provide E2 nodes and the Near-RT RIC with OAM functions such as fault, configuration, accounting, performance, and security (FCAPS) management, software management, and file management. The Non-RT RIC links up with a function that provides OAM services within SMO to obtain PM counter data generated by E2 nodes using the O1 interface. It can also reflect configuration settings, which are optimized by the rApp in the Non-RT RIC, in the E2 nodes. For cases that apply ML in the Near-RT RIC, the O1 interface can potentially be used to deploy an ML model.

(2) A1 interface

The A1 interface connects the Non-RT RIC and Near-RT RIC. Three functions are specified for this interface: A1 policy management service (A1-P), A1 enrichment information service (A1-EI), and A1 ML model management service (A1-ML).

(3) E2 interface

The E2 interface connects the Near-RT RIC and E2 nodes. It executes the functions of exposing information on E2 node control functions and control history to the Near-RT RIC and notification of control commands to E2 nodes. The E2 interface can control E2 nodes through radio resource control (RRC)[*5] handover (HO) control[*6] and control of S1[*7]/X2[*8]/NG[*9]/Xn[*10]/F1[*11]/E1[*12] procedures. Control can be specified in units of cells, slices, or user equipment (UE).

(4) R1 interface

The R1 interface sends and receives data and control information between the rApp and Non-RT RIC framework. Its main functions are service management and exposure (SME) services and data management and exposure (DME) services. A1-related services, O1-related services, O2-related services, and AI/ML workflow services are also being specified.

The functions in SME include BootStrap that conveys endpoints of various services provided by the Non-RT RIC framework, Registration that registers services provided by the rApp, Discovery that searches for services provided by the Non-RT RIC framework and rApp, Heartbeat that monitors the state of the rApp, and Authorization that authenticates and authorizes services provided by the rApp.

DME functions include data registration that registers data that can be provided by the Non-RT RIC framework and rApp, data discovery that obtains data that have completed registration (data catalog), data request/subscription that requests data, data collection that collects data, and data delivery that transmits data.

Currently specified for O1-related services are a network information service that obtains the configuration, status, etc. of a network function (NF)[*13] that SMO can obtain and an FM/PM service that obtains FM and PM information. Details on A1-related, O2-related, and AI/ML workflow functions have not yet been specified. Standardization of the R1 interface at the O-RAN ALLIANCE has just begun, so details have not yet been specified. Function updates and changes can therefore be expected.

(5) Near-RT RIC APIs

Near-RT RIC APIs lie between the xApp and Near-RT RIC framework. In addition to A1-related APIs and E2-related APIs, these include management APIs that handle the xApp and AI/ML model management (registration, modification, deletion, and configuration), logging, tracing, and metric collection, shared data layer (SDL)[*14] APIs that provide access to SDL-related functions, and enablement APIs that execute authentication, registration, etc. to enable the xApp to use APIs.

2.3 rApp/xApp

Control algorithms running on the Non-RT RIC and Near-RT RIC are implemented by the rApp and xApp, respectively, as described above, and separated from the Non-RT RIC and Near-RT RIC frameworks by the R1 interface and Near-RT RIC APIs, respectively. This scheme enables a mobile network operator

*5 RRC: Layer 3 protocol that controls radio resources in the radio network.

*6 HO control: Changes the base station that a UE is connected to.

*7 S1: Interface between Evolved Packet Core and evolved NodeB (eNB).

*8 X2: Interface between eNB equipment.

*9 NG: Interface between 5GC (5G core network) and next generation NodeB (gNB).

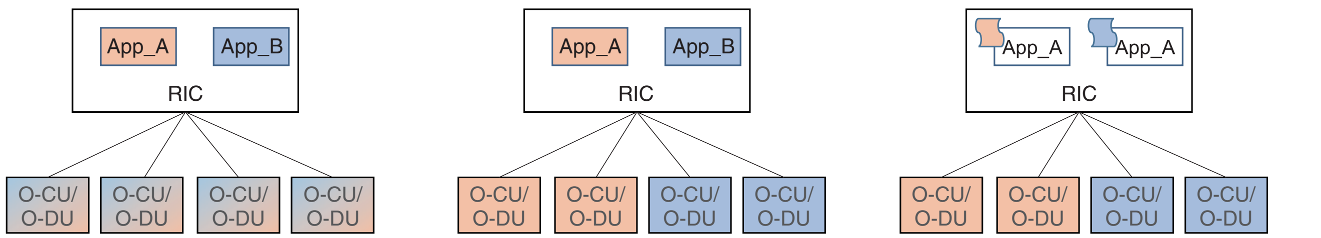
*10 Xn: Interface between gNB equipment.

*11 F1: Interface between O-CU and O-DU.

*12 E1: Interface between O-CU-CP (control plane) and O-CU-UP (user plane).

*13 NF: Logical unit for identifying individual network functions.

*14 SDL: Function that simplifies database access.



(a) Case 1: Different Apps (App_A and App_B) control all O-CUs/ O-DUs

(b) Case 2: Different Apps (App_A and App_B) control different O-CUs/ O-DUs

(c) Case 3: Two instances of the same App having different configuration settings control different O-CUs/ O-DUs

Fig. 2. rApp/xApp deployment scenarios.

to freely select applications. In other words, it is possible to adopt not only an rApp from a vendor that provides the Non-RT RIC framework but also an rApp provided by a third party. An operator can also reflect RAN control policies in rApp algorithms on the basis of RAN operation experience and knowledge, thereby providing services that satisfy a variety of requirements and conditions.

An rApp can share data with another rApp via DME. For example, rApps having different functions, such as an rApp specialized in collecting and analyzing data, an rApp that takes the results of that analysis to generate a ML model, and an rApp that executes inference using that ML model and generates control commands and control policies for E2 nodes on the basis of inference results, could be linked to achieve a single use case.

It is also possible to apply different rApps or xApps (hereinafter referred to as "Apps") for each automation/optimization objective (use case). For example, a flexible automated service can be applied by running App_A and App_B in parallel, as shown in **Fig. 2(a)**. Alternatively, different Apps can be applied to different areas, as shown in **Fig. 2(b)**, or configuration settings can be changed for the same App to change automation or optimization operations, as shown in **Fig. 2(c)**.

2.4 Application of ML

Recent progress in cloud technology has made it easy to accumulate large amounts of data; thus, the application of ML to a variety of fields is attracting attention. At the O-RAN ALLIANCE, which aims to achieve intelligent RAN, there are expectations that network performance will improve by applying ML to the RAN field, so architecture for making that possible is being prepared.

The application of ML requires training and inference processes. In the training process, network-performance data stored in a data lake[*15] are used by ML training hosts situated inside or outside the RIC architecture to train an ML model and save it on the RIC. In the inference process, the ML model is loaded into the rApp or xApp on the RIC and used by ML inference hosts situated inside or outside the RIC architecture to infer optimal values for target parameters. Optimized parameters are then set in O-CU and O-DU via the A1, O1 and E2 interfaces.

3. Implementation scenario for intelligent RAN

3.1 Intelligent RAN roadmap

On implementing intelligent RAN by the RIC, the functions and control interfaces that will be needed, the data collection items needed for optimization analysis will differ depending on the use cases adopted, and the interfaces and functions needed in RAN equipment (next generation NodeB (gNB)[*16]) will differ. The formulation of an implementation plan that takes the above into consideration is therefore important. It is also necessary to draw up a plan for enhancing intelligent RAN in a step-by-step manner taking into account the status of formulating specifications related to use cases at the O-RAN ALLIANCE and the time that those specifications can be expected to mature.

As shown in **Fig. 3**, NTT DOCOMO considers the automating of operations implemented via maintenance personnel in the conventional RAN operation

*15 Data lake: A data storage repository.

*16 gNB: A radio base station providing NR (New Radio) signals in RAN.

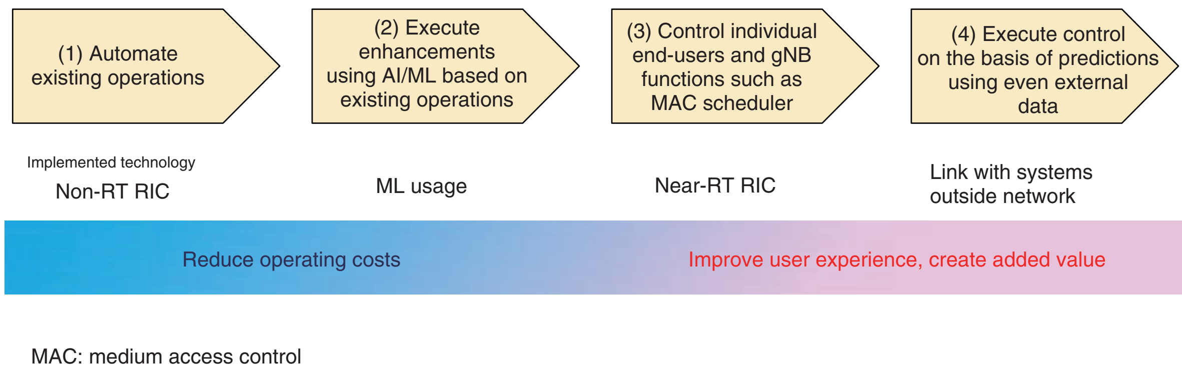


Fig. 3. Implementation scenario for intelligent RAN.

system as an initial use case with the aim of reducing operating costs (Fig. 3 (1)). This phase envisions, for example, the automating and optimizing of base-station operation parameters, antenna directivity, and base-station sleep control for power-saving purposes on the basis of predictions of traffic load. The aim is to provide optimized network settings adaptively in control loops of several hours to several days on the basis of changes in the radio environment or fluctuations in traffic load. It is also possible to achieve such use cases by control processes using the O1 interface by the Non-RT RIC, and since the impact on functions on the RAN equipment (gNB) side should be relatively small, costs at the time of implementation should be minimized. The plan is to increase the application domain of AI/ML in a stepwise manner to make RAN operations increasingly intelligent (Fig. 3 (2)).

In the next phase, NTT DOCOMO will target use cases connected to improving RAN performance and the degree of customer satisfaction by enhancing control schemes (for example, by shifting from low-speed to high-speed control or from individual-cell control to individual-user control). Specifically, it envisions "traffic steering" or "quality of service/quality of experience optimization" that optimizes resource control according to service requirements for each end-user or network slice*17 (Fig. 3 (3)). Achieving these use cases will require support of A1/E2 interfaces in addition to implementing the Near-RT RIC. In RAN equipment (gNB) as well, it will be necessary to support various functions prescribed by E2 service model RAN control (E2SM-RC), the specifications of which are being drafted in O-RAN ALLIANCE Working Group 3 (WG3). It will also be necessary to add functions to RAN equipment or undertake a medium-term or long-term migration that eyes the possibility of upgrades. With a view to linking with external systems and applying prediction technology in the future, there is the goal of creating new value through the mobile network (Fig. 3 (4)).

3.2 Examples of use cases in each phase

The following introduces (1) optimization of HO control parameters as a use case for the initial phase and (2) traffic steering as a use case for the following phase.

(1) Optimization of HO control parameters

Executing an HO between a base station and UE either too early or too late will cause an HO failure and the UE to be temporarily disconnected from the network. To prevent such a failure from occurring, The Non-RT RIC can adjust thresholds, timing, etc. used in HO control by analyzing information on the cell environment and disconnection events. The procedure for optimizing HO control parameters by the Non-RT RIC is shown in **Fig. 4**. An optimal HO environment can always be provided to the UE by autonomously and automatically repeating steps 1 to 5 in Fig. 4 in units of cell or time.

(2) Traffic steering

Mobile communication networks can support combinations of various access networks such as New Radio (NR), Long-Term Evolution (LTE), and Wi-Fi. These combinations feature a radio environment with multiple frequency bands and traffic fluctuations due to diverse user applications, so to provide a stable network, advanced traffic management beyond what

*17 Network slice: A format for achieving next-generation networks in the 5G era that entails logically dividing a network into service units for use cases, business models, etc.

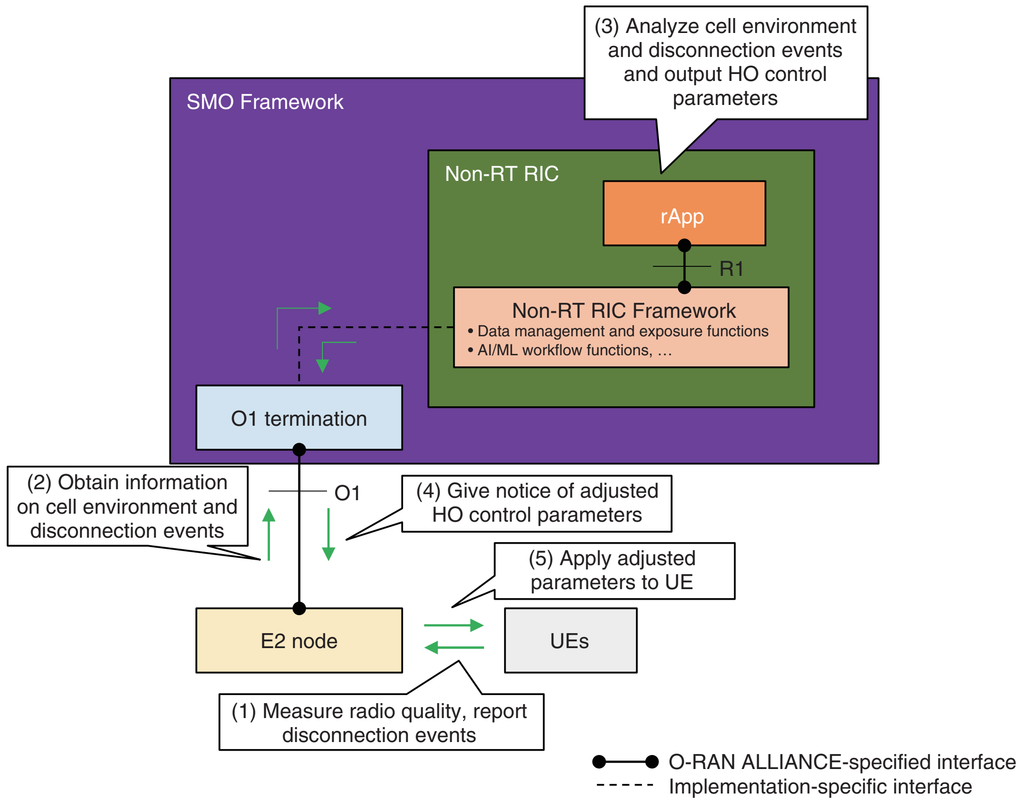


Fig. 4. Example of optimizing and controlling HO control parameters by Non-RT RIC.

had been required is deemed necessary, as summarized below.

- Radio resource management (RRM) covering the range from conventional cell units to UE units having diverse requirements
- Load balance through a multi-access network and UE performance prediction
- Application of traffic control with appropriate timing

To meet the above requirements, a mobile network operator can use RIC, which will configure optimal policies flexibly in accordance with the purpose of network operation, execute appropriate network and UE performance measurements in real time, and manage traffic proactively. The procedure for traffic steering by the Non-RT RIC and Near-RT RIC is shown in **Fig. 5**. A load-balanced smooth-running network can always be provided by autonomously and automatically repeating steps (1) to (9) in Fig. 5.

4. Summary

4.1 Future issues

The following problems related to multi-vendor operation are taken up as future issues.

Activities are progressing in O-RAN ALLIANCE WG5 to achieve multi-vendor interoperability in RAN equipment interfaces, but interoperability is also needed for RIC interfaces (R1 interface, A1 interface, E2 interface, Near-RT RIC APIs, and external interfaces to outside services and applications). Specifications for these interfaces are currently being drafted in WG2 and WG3, but clarification of control and operation with respect to E2 nodes for each use case is deemed necessary so that parameter interpretation among vendors does not differ.

Various issues surrounding operation can be considered, such as management of AI/ML models provided by Apps (rApp, xApp), conflict management between Apps, and operation management of applicable/non-applicable areas of RIC functions at the time of system implementation. It is thought that maintenance personnel will evaluate and manually

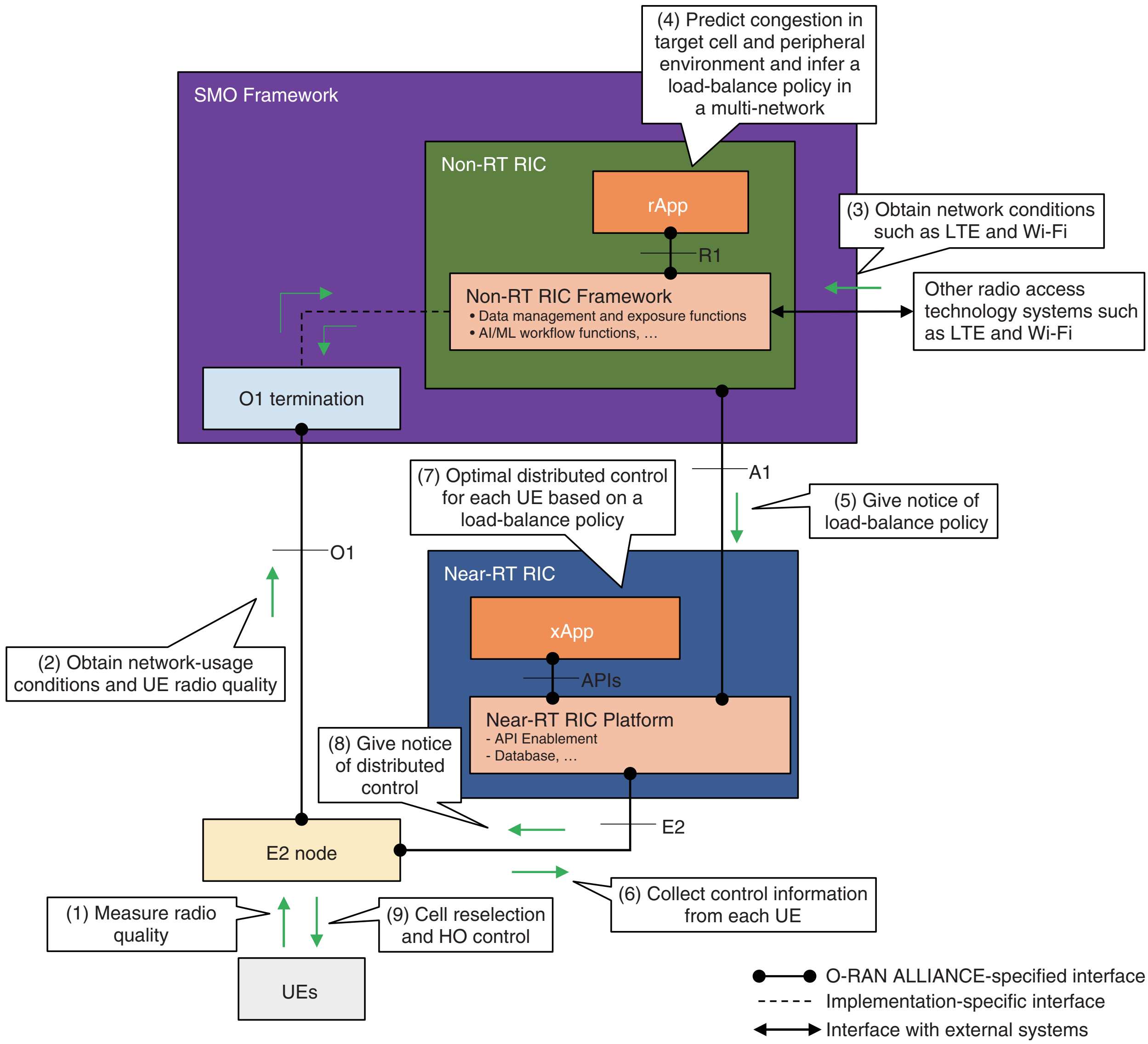


Fig. 5. Example of traffic-steering control by Non-RT RIC and Near-RT RIC.

carry out these operation management tasks in the initial implementation phase, but studies will be needed on incorporating RIC functions to automate these operations.

4.2 Future use cases

Looking to the future, RAN parameter optimization using data outside the RAN domain can also be considered.

We can consider the example of mobility optimization on an expressway. In this example, the RIC optimizes base-station parameters in accordance with the speed of UE movement and UE density using base-station data in the vicinity of the expressway as well as real-time congestion information or congestion-prediction data provided by the expressway management company. We can also consider parameter optimization in accordance with current conditions by having the RIC obtain data collected from in-vehicle sensors from an intelligent in-vehicle terminal such as a smartphone via the Internet. Specifically, congestion in the local mobile network could be predicted and measures for eliminating that congestion taken, or the speed of UE movement or predictions of the vehicle's destination could be used to optimize HO control or secure resources at that destination.

To carry out network operations appropriate to the amount of power available in a country or region with an unstable power supply, power-off and power-on plans from the power provider could be used to automatically shut down base stations in the power-off area and expand the coverage of base stations in the power-on area. In another case, a stable power supply may be available, but mobile network operators are free to select the power-supply source they wish to use from among multiple power retailers, solar power generation equipment or storage batteries near base stations, etc. In this case, data on the amount and cost of power available from each power supply can be used to minimize power expenses while maintaining communications quality.

We can also consider events where a large number of people come together such as large-scale sports events, music festivals, and fireworks displays. In such cases, postings on social media could be used to determine the location and time of such an event, and that information could then be passed to the RIC. Using that information, the RIC could then increase the number of simultaneous UE connections by adjusting the parameters of the base stations near that event beforehand, which should eliminate the difficulty in making a connection due to base-station congestion.

If the public release of such data related to the urban infrastructure expands, intelligent RAN should be able to contribute to the creation of smart cities.

5. Conclusion

This article described the RIC now being standardized at the O-RAN ALLIANCE, a roadmap and various use cases as a scenario for implementing intelligent RAN, and upcoming issues and future use cases.

Going forward, NTT DOCOMO plans to continue contributing to the drafting of specifications for intelligent RAN at the O-RAN ALLIANCE. It will also promote studies on enhancing intelligent RAN and achieving multi-vendor interoperability at the 5G Open RAN Ecosystem that is evolving.

Trademark notes

All company names or names of products, software, and services appearing in this article are trademarks or registered trademarks of their respective owners.

Taichi Katsuragawa

Assistant Manager, Radio Access Network Development Department, NTT DOCOMO, INC.

He received a B.E. from Ritsumeikan University, Kyoto, in 2001. Since joining Panasonic Corporation in 2001, he had been involved in software development of the base station equipment for 3G, LTE/LTE-Advanced, and 5G. In 2021, he joined NTT DOCOMO and is currently engaged in the research and development of the RAN Intelligent Controller.

Takahiro Tateishi

Engineer, Radio Access Network Development Department, NTT DOCOMO, INC.

He received a B.E. and M.E. from Yokohama National University, Kanagawa, in 2014 and 2016. In 2016, he joined NEC Corporation, where he was engaged in the development of operation support systems. In 2021, he joined NTT DOCOMO, where he is currently engaged in the research and development of the RAN Intelligent Controller.

Akihiro Kawana

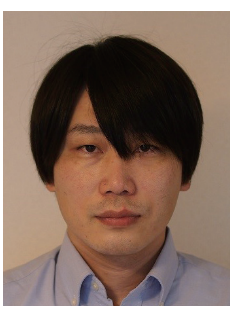
Assistant Manager, Radio Access Network Development Department, NTT DOCOMO, INC.

He received a B.E. from Waseda University, Tokyo, in 2000 and M.S. from Japan Advanced Institute of Science and Technology, Ishikawa, in 2002. He joined NTT in 2002 and NTT DOCOMO in 2010. His research interests include data science and data utilization. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).

Eina Hashimoto

Engineer, Radio Access Network Development Department, NTT DOCOMO, INC.

She received a B.E., M.E., and Ph.D. from the University of Electro-Communications, Tokyo, in 2015, 2017, and 2021. Her research interests include big data analysis and machine learning. Since joining NTT DOCOMO in 2017, she had been a data scientist of the big data group with the Service Innovation Department, where she was involved in big data analytics. She is currently engaged in the automatic optimization of wireless networks.



Yoshio Inoue

Engineer, Radio Access Network Development Department, NTT DOCOMO, INC.

He received a B.E. and M.E. from Oita University in 2009 and 2011. Since joining Fujitsu Limited in 2011, he had been involved in the development of the private RAN system and the contribution of open source software for the software-defined RAN. In 2021, he joined NTT DOCOMO and is currently engaged in the research and development of the RAN Intelligent Controller.

Takumi Fujitsuka

Engineer, Radio Access Network Development Department, NTT DOCOMO, INC.

He received a B.E. and M.E. from Toyohashi University of Technology, Aichi, in 2017 and 2019. Since joining NTT DOCOMO in 2019, he has been involved in the development of 4G and 5G radio access technologies. He is currently engaged in the research and development of the RAN Intelligent Controller, including standardization at the O-RAN ALLIANCE.

Extreme-broadband Analog ICs Based on InP HBT Technology toward Multi-Tbit/s Optical Communications

Munehiko Nagatani, Teruo Jyo, Hitoshi Wakita, Tsutomu Takeya, and Hiroyuki Takahashi

Abstract

To cope with the rapid growth of communications traffic, transmission capacity per channel (wavelength) in optical core networks is required to reach the multi-Tbit/s level in the next decade. In this article, we review our latest extreme-broadband analog integrated circuits (ICs), e.g., 110-GHz-bandwidth analog multiplexer/analog de-multiplexer ICs and 200-GHz-bandwidth amplifier, mixer, and combiner ICs, based on our in-house indium phosphide (InP) heterojunction bipolar transistor (HBT) technology that can break through the bandwidth bottlenecks in optical transceivers and be key enablers toward future multi-Tbit/s/ch optical communications systems.

Keywords: analog IC, InP HBT, optical communications

1. Introduction: challenges toward multi-Tbit/s/ch optical communications

Communications traffic has been rapidly increasing due to the spread of broadband applications. In addition, the COVID-19 pandemic has dramatically changed our lifestyles, and various activities and services have migrated to remote and virtual platforms. These changes are further accelerating the growth of communications traffic. Communications networks must continue to increase their capacity to cope with such rapid growth in traffic. Optical core networks require huge-capacity and long-haul transmission systems to accommodate client data and link metropolitan areas as the backbone of the communications infrastructure. Digital coherent technology, which combines coherent detection and digital signal processing, was introduced into optical core networks in the 2010s. Thus far, 100-Gbit/s/ch systems based on 32-GBaud polarization-division-multiplexed (PDM) quadrature phase shift keying (QPSK) and 400-Gbit/s/ch systems consisting of 64-GBaud PDM 16-ary quadrature amplitude modulation (16QAM) have been deployed [1, 2].

To sustain ever-increasing traffic, the transmission capacity per channel (per wavelength) is expected to exceed 1 Tbit/s in the near future and reach multi-Tbit/s in the 2030s. Channel capacity is defined as the product of the symbol rate and modulation order, so it can be improved by increasing the symbol rate and/or modulation order, as illustrated in **Fig. 1**. Increasing the symbol rate while maintaining a low modulation order is advantageous for constructing long-haul transmission systems because higher-order modulation faces optical signal-to-noise-ratio limitation due to amplified spontaneous emission noise and fiber-nonlinearity in the optical links, shortening transmission distance. Therefore, a symbol rate of over 200 GBaud is necessary to construct future multi-Tbit/s/ch long-haul optical transmission systems. This means that each building block in an optical transceiver is required to have an analog bandwidth of at

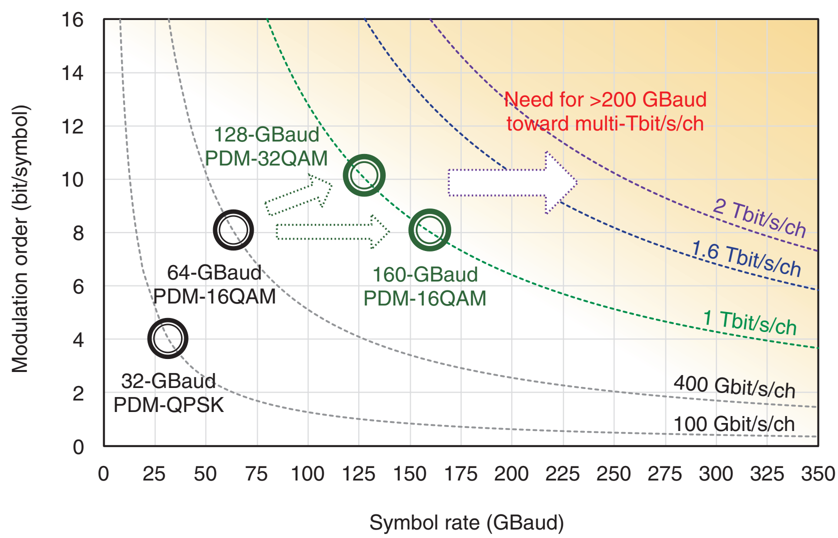


Fig. 1. Channel capacity: symbol rate versus modulation order.

least 100 GHz, which is a Nyquist frequency of 200 GBaud. One of the most significant challenges in constructing multi-Tbit/s/ch systems is implementing extreme-broadband digital-to-analog converters (DACs) and analog-to-digital converters (ADCs) in the optical transceiver. To address this challenge, we previously proposed and intensively investigated bandwidth doubler technology [3, 4] based on analog multiplexer (AMUX) and analog de-multiplexer (ADEMUX) integrated circuits (ICs) that can double the bandwidth of DACs and ADCs, respectively, as illustrated in **Fig. 2**.

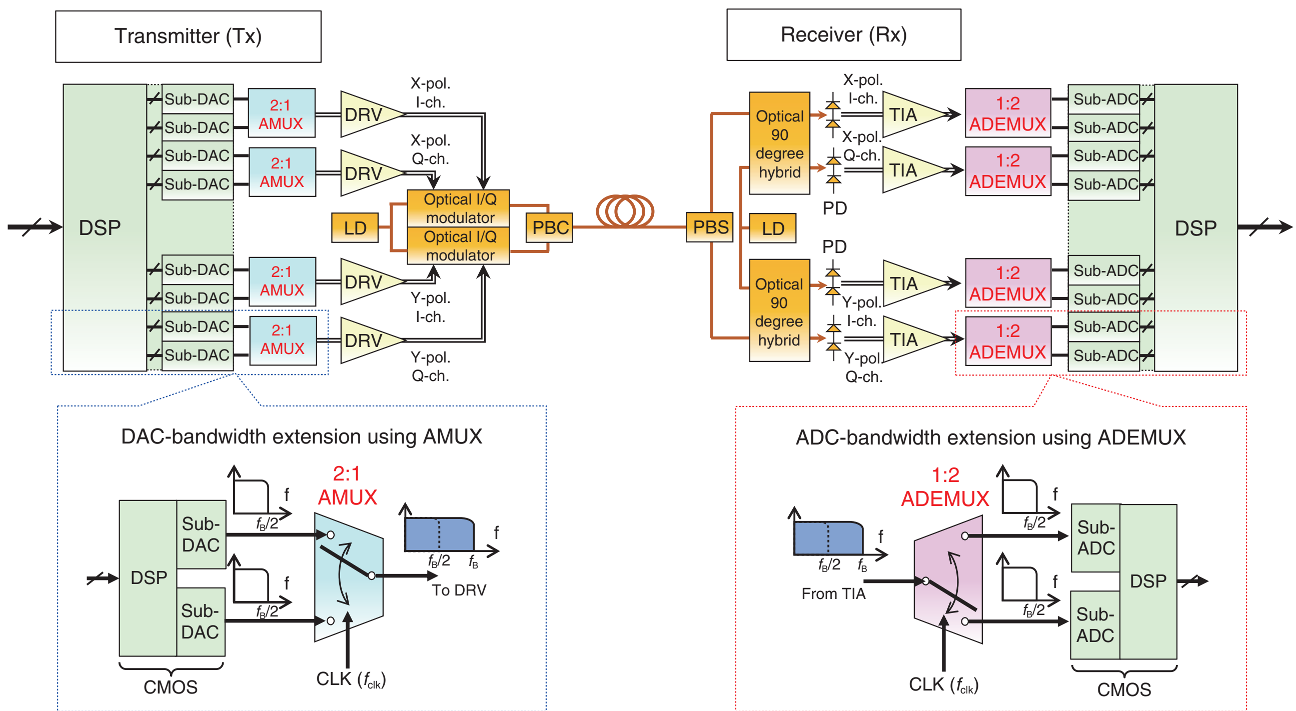
This article reviews our latest extreme-broadband analog ICs, e.g., 110-GHz-bandwidth AMUX/ADEMUX ICs [5, 6] and 200-GHz-bandwidth amplifier, mixer, and combiner ICs [7–10], based on our in-house 250-nm indium phosphide (InP) heterojunction bipolar transistor (HBT) technology [11] that enable further bandwidth extension.

2. 110-GHz-bandwidth AMUX and ADEMUX ICs

We designed and fabricated AMUX and ADEMUX ICs using our in-house 250-nm InP HBT technology [11]. **Figures 3(a)** and **(b)** show a scanning electron microscopy (SEM) image and cross-section of our InP HBT. The HBT layer structure was grown on a 3-in InP substrate and consists of a degenerately doped indium gallium arsenide (InGaAs) emitter contact, 20-nm-thick InP emitter, 25-nm-thick compositionally graded InGaAs base, 100-nm-thick InGaAs/indium aluminum gallium arsenide (InAlGaAs)/InP collector, and InGaAs/InP subcollector. The HBTs have a peak cutoff frequency (f_T) and maximum oscillation frequency (f_{max}) of 460 and 480 GHz, respectively, as shown in **Fig. 3(c)**. They also exhibit a common-emitter breakdown voltage (BV_{CEO}) of 3.5 V. Thus, the InP HBTs have both high-speed and high breakdown voltage characteristics and suitable for implementing broadband and high output power circuits.

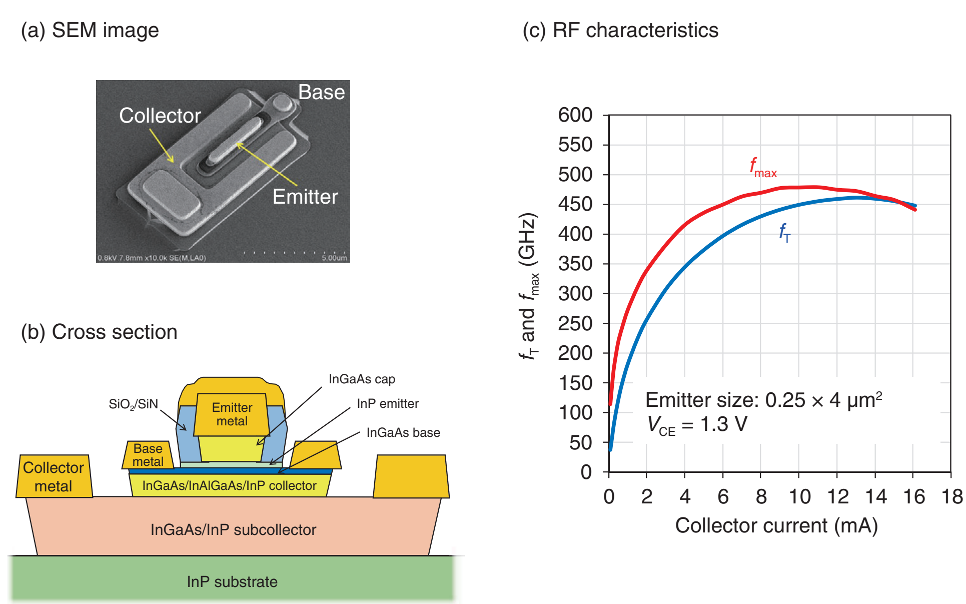
The 2:1 AMUX IC is composed of two input linear buffers, a clock limiting buffer, AMUX core based on Gilbert-cell selector circuitry, and output linear buffer, as shown in **Fig. 4(a)**. The detailed transistor-level structures are described in a previous study [5]. All the building blocks in the AMUX IC have fully differential configurations. **Figure 4(b)** shows a microphotograph of the 2:1 AMUX IC. The chip is 2 mm^2 and contains 210 HBTs and consumes a total power of 0.90 W with a single power-supply voltage of –4.5 V. **Figure 4(c)** shows the measured and simulated S-parameters for the analog signal path in through mode. The measured 3-dB bandwidth was over 110 GHz and agreed well with the simulation. These results indicate that this AMUX IC enables us to generate 200-GBaud-class baseband signals.

A block diagram of the 1:2 ADEMUX IC is shown in **Fig. 5(a)**. This IC consists of input linear buffers, clock limiting buffers, an ADEMUX core based on two parallel track-and-hold (T/H) circuits, and output



CLK: clock
CMOS: complementary metal-oxide semiconductor
DRV: optical modulator driver
DSP: digital signal processor
I/Q: in-phase and quadrature
LD: laser diode
PBC: polarization beam combiner
PBS: polarization beam splitter
PD: photodiode
TIA: trans-impedance amplifier

Fig. 2. Our bandwidth doubler technology using 2:1 AMUX and 1:2 ADEMUX.



SiO_2: silicon dioxide
SiN: silicon nitride

Fig. 3. Our InP HBT: (a) SEM image, (b) cross section, and (c) radio-frequency (RF) characteristics.

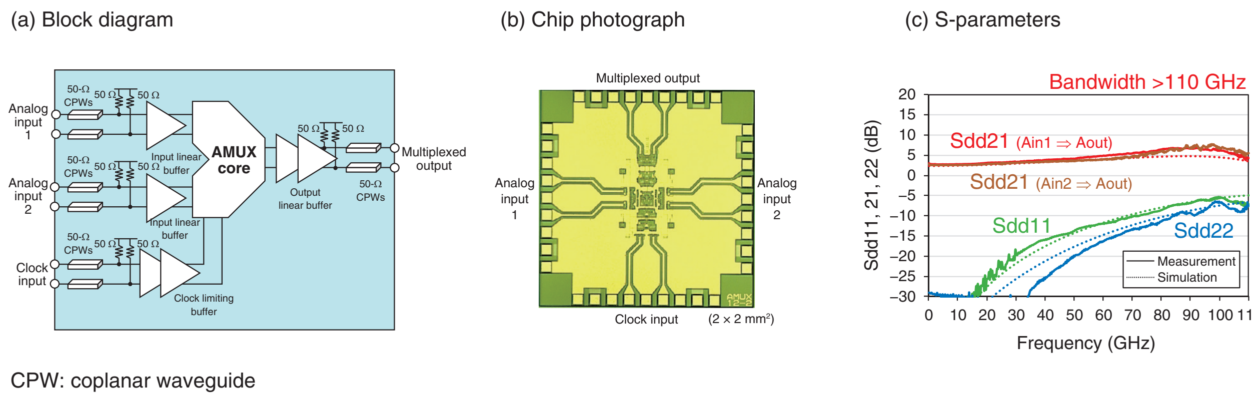


Fig. 4. 2:1 AMUX IC: (a) block diagram, (b) chip photograph, and (c) S-parameters.

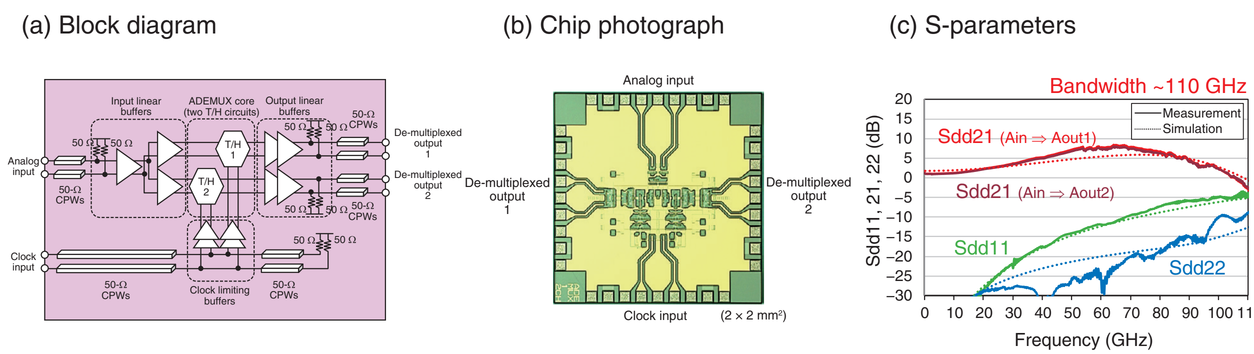


Fig. 5. 1:2 ADEMUX IC: (a) block diagram, (b) chip photograph, and (c) S-parameters.

linear buffers. The T/H circuit in the ADEMUX core is composed of switched emitter followers and hold capacitors [6]. The other blocks have similar topologies with the blocks used in the AMUX IC. **Figure 5(b)** is a microphotograph of the 1:2 ADEMUX IC. The chip size is also 2 mm^2 and has 320 integrated HBTs. The ADEMUX IC dissipates a power consumption of 1.14 W with a –4.5-V power supply. The measured and simulated S-parameters for the analog signal path in track mode are shown in **Fig. 5(c)**. The measured 3-dB bandwidth was 110 GHz and agreed with the simulations, as expected. These results also indicate that this ADEMUX IC can be used to receive 200-GBaud-class baseband signals.

We have already succeeded in demonstrating various record ultrahigh-speed optical transmission experiments using an AMUX-integrated optical transmitter front-end module [5], such as a 200-GBaud-class optical modulation [12] and beyond-1-Tbit/s/ch over-1000-km long-haul optical transmissions [13]. Thus, the AMUX and ADEMUX ICs can be key enablers for next generation 200-GBaud-class long-haul optical transmission systems.

3. 200-GHz-bandwidth amplifier, mixer, and combiner ICs

Toward future multi-Tbit/s/ch systems with a symbol rate of over 200 GBaud, we proposed and investigated bandwidth-extension techniques [7–10], as illustrated in the conceptual block diagrams (**Fig. 6**). On the transmitter side, the signal bandwidth can be doubled using a mixer, amplifier, and combiner. On the receiver side, however, the signal bandwidth can also be doubled using an amplifier (splitter) and mixer. We have designed and fabricated amplifier, mixer, and combiner IC prototypes that have a bandwidth of 200 GHz, which is twice that of the AMUX and ADEMUX ICs described in the previous section,

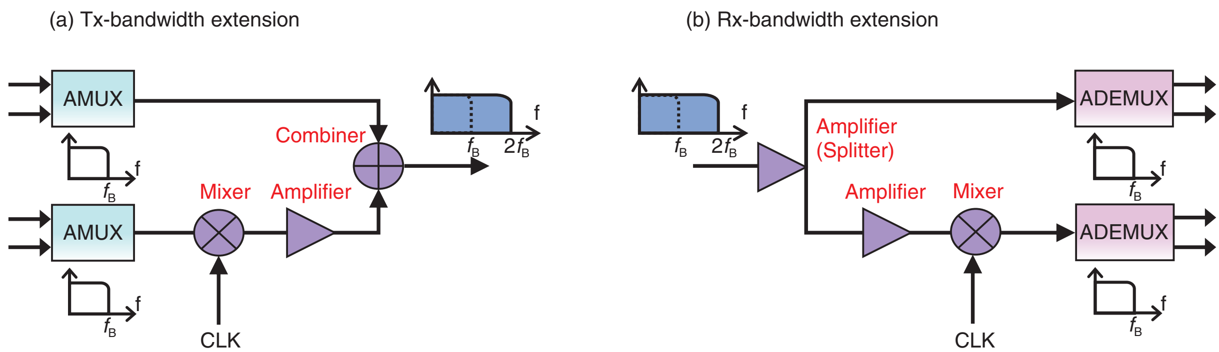


Fig. 6. Bandwidth-extension techniques using amplifier, mixer, and combiner: (a) Tx-bandwidth extension and (b) Rx-bandwidth extension.

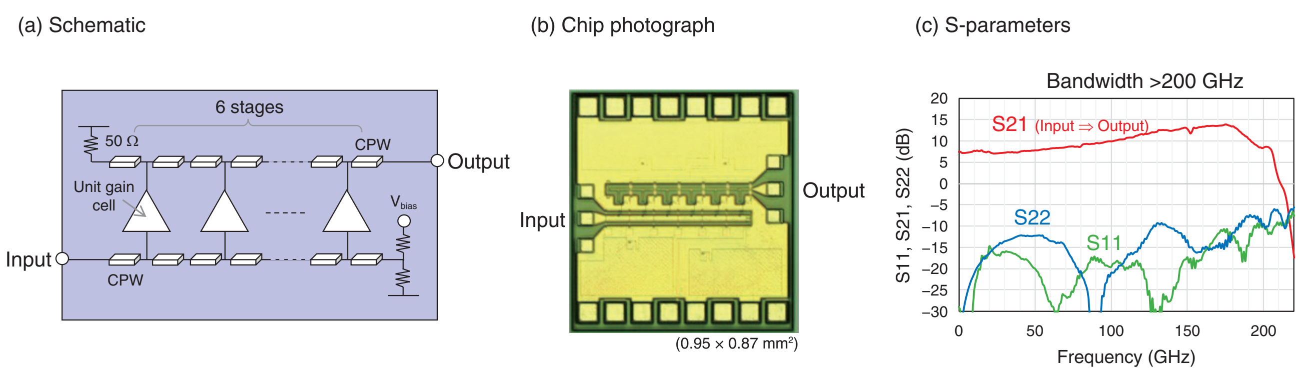


Fig. 7. Amplifier IC: (a) schematic, (b) chip photograph, and (c) S-parameters.

using the same InP HBT.

Figure 7(a) shows a schematic of the amplifier IC. We applied a distributed circuit topology to the amplifier to achieve a broad gain profile and impedance matching. The distributed amplifier consists of six unit gain cells, coplanar waveguide (CPW) transmission lines, and termination resistors. Multi-peaking techniques were proposed to make broad-peaking characteristics to compensate for loss derived from the packaging and subsequent components [7, 8]. A microphotograph of the fabricated amplifier IC is shown in **Fig. 7(b)**. The IC is 0.95 × 0.87 mm and consumes 0.54 W of power with –4.5-V supply voltage. **Figure 7(c)** shows the measured S-parameters. The direct current (DC) gain and 3-dB bandwidth were 7.5 dB and 208 GHz, respectively.

Figure 8(a) is a schematic of the mixer IC. The mixer is based on a distributed topology and composed of six unit cells, each of which is a single-balanced mixer circuitry [9]. One of the differential outputs is terminated inside this prototype for testing. **Figure 8(b)** shows a microphotograph of the fabricated mixer IC. The chip size is approximately 0.8 mm^2. The power consumption is 0.13 W with a supply voltage of –3.8 V. The measured conversion gain (CG) and S-parameters are shown in **Fig. 8(c)**. The clock signal was set to 97.2 GHz for CG measurement. The CG is –3 dB in low frequency range. The output covers the frequency range from DC to approximately 200 GHz. This mixer has the broadest bandwidth of any previously reported mixers.

A schematic of the combiner IC is illustrated in **Fig. 9(a)**. It consists of two parallel distributed amplifiers and a subsequent distributed resistive combiner [10]. This architecture helps in achieving both broadband characteristics and good isolation between two input ports. **Figure 9(b)** shows a microphotograph of the combiner IC, which is 0.8 mm^2. The power dissipation is 1.1 W from a supply of –4.5 V. The measured S-parameters are shown in **Fig. 9(c)**. The

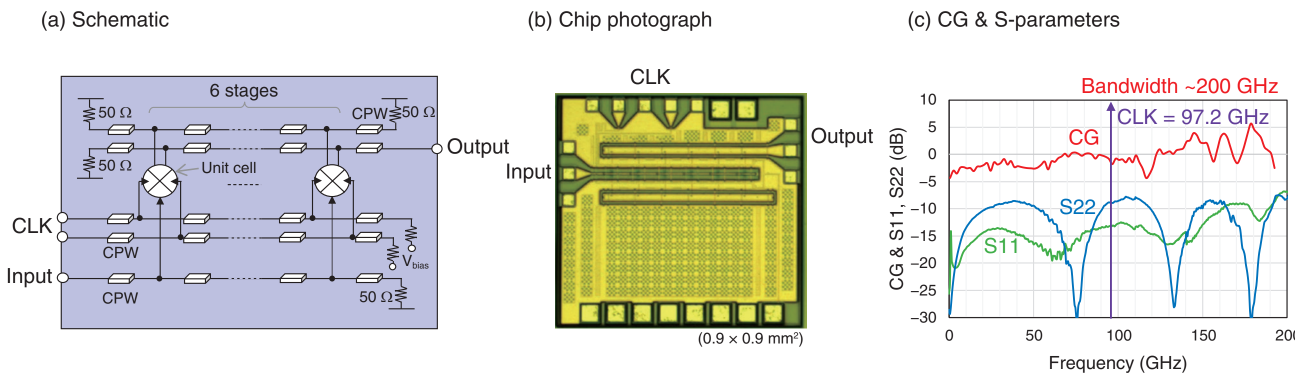


Fig. 8. Mixer IC: (a) schematic, (b) chip photograph, and (c) CG and S-parameters.

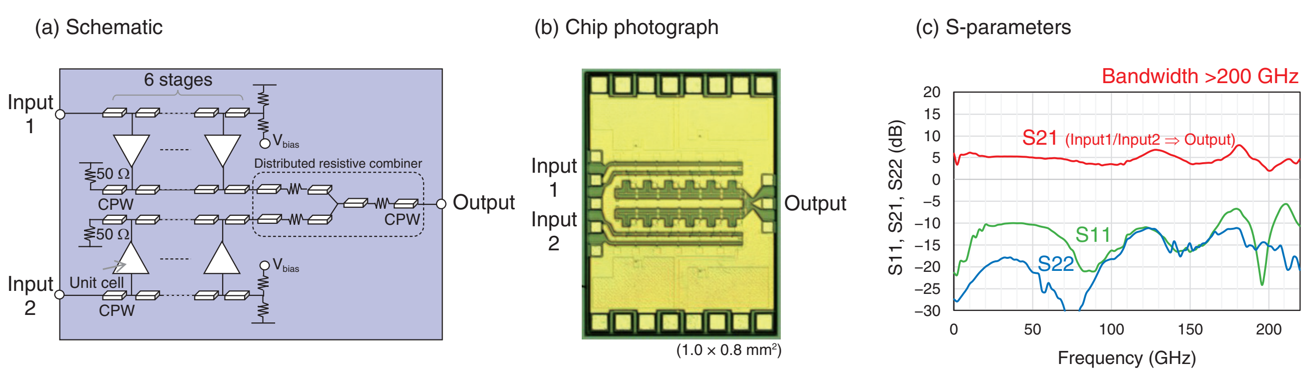


Fig. 9. Combiner IC: (a) schematic, (b) chip photograph, and (c) S-parameters.

measured 3-dB bandwidth is over 220 GHz, which is the broadest reported to date.

We have succeeded in fabricating extreme-broadband amplifier, mixer, and combiner ICs that have a bandwidth of 200 GHz. They enable us to handle 400-GBaud-class signals. Thus, these extreme-broadband ICs could be key to future multi-Tbit/s/ch systems.

4. Summary

We reviewed our latest extreme-broadband analog ICs that can break through the bandwidth bottlenecks in transceivers toward multi-Tbit/s/ch optical communications. We will accelerate investigations for the practical application of these analog ICs and pursue further extreme performance to ensure sustainable progress of communications technologies.

References

[1] Y. Miyamoto, K. Watanabe, and K. Nakajima, "Research and Development of Scalable Optical Transport Technologies," NTT Technical Review, Vol. 20, No. 7, pp. 24–29, 2022. https://doi.org/10.53829/ntr202207fa3

[2] Y. Kisaka, E. Yamazaki, H. Nishizawa, and T. Saida, "Future Development of Digital Coherent Optical Transmission Technology," NTT Technical Review, Vol. 20, No. 7, pp. 17–23, 2022. https://doi.org/10.53829/ntr202207fa2

[3] M. Nagatani, H. Wakita, Y. Ogiso, H. Yamazaki, M. Ida, and H. Nosaka, "Ultrahigh-speed Optical Front-end Device Technology for Beyond-100-GBaud Optical Transmission Systems," NTT Technical Review, Vol. 17, No. 5, pp. 27–33, 2019. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201905fa4.html

[4] Press release issued by NTT, "World's first experimental long-distance WDM optical transmission at 1 Tbps per wavelength—Expected to provide large-capacity network communications for IoT and 5G services," Mar. 7, 2019. https://group.ntt/en/newsrelease/2019/03/07/190307a.html

[5] M. Nagatani, H. Wakita, H. Yamazaki, Y. Ogiso, M. Mutoh, M. Ida, F. Hamaoka, M. Nakamura, T. Kobayashi, Y. Miyamoto, and H. Nosaka, "A Beyond-1-Tb/s Coherent Optical Transmitter Front-end Based on 110-GHz-bandwidth 2:1 Analog Multiplexer in 250-nm InP DHBT," IEEE J. Solid-State Circuits, Vol. 55, No. 9, pp. 2301–2315, 2020.

[6] M. Nagatani, H. Wakita, T. Jyo, T. Takeya, H. Yamazaki, Y. Ogiso, M. Mutoh, Y. Shiratori, M. Ida, F. Hamaoka, M. Nakamura, T. Kobayashi, H. Takahashi, and Y. Miyamoto, "110-GHz-bandwidth InP-HBT AMUX/ADEMUX Circuits for Beyond-1-Tb/s/ch Digital Coherent Optical Transceivers," Proc. of IEEE Custom Integrated Circuits Conference (CICC), Newport Beach, CA, USA, Apr. 2022.

[7] T. Jyo, M. Nagatani, M. Ida, M. Mutoh, H. Wakita, N. Terao, and H. Nosaka, "A 241-GHz-bandwidth Distributed Amplifier with 10-dBm P1dB in 0.25-μm InP DHBT Technology," Proc. of IEEE MTT-S International Microwave Symposium (IMS), pp. 1430–1433, Boston, MA, USA, June 2019.

[8] T. Jyo, M. Nagatani, M. Mutoh, Y. Shiratori, H. Wakita, and H. Takahashi, "An Over 130-GHz-bandwidth InP-DHBT Baseband Amplifier Module," Proc. of IEEE BiCMOS and Compound Semiconductor Integrated Circuits and Technology Symposium (BCICTS), Monterey, CA, USA, Dec. 2021.

[9] T. Jyo, M. Nagatani, M. Ida, M. Mutoh, H. Wakita, N. Terao, and H. Nosaka, "A DC to 194-GHz Distributed Mixer in 250-nm InP DHBT Technology," Proc. of IEEE MTT-S International Microwave Symposium (IMS), pp. 771–774, Virtual, Aug. 2020.

[10] T. Jyo, M. Nagatani, M. Ida, M. Mutoh, H. Wakita, N. Terao, and H. Nosaka, "An over 220-GHz-bandwidth Distributed Active Power Combiner in 250-nm InP DHBT," Proc. of BCICTS 2020, Virtual, Nov. 2020.

[11] N. Kashio, K. Kurishima, M. Ida, and H. Matsuzaki, "Over 450-GHz f_t and f_{max} InP/InGaAs DHBTs with a Passivation Ledge Fabricated by Utilizing SiN/SiO_2 Sidewall Spacers," IEEE Trans. Electron Devices, Vol. 61, No. 10, pp. 3423–3428, 2014.

[12] M. Nakamura, F. Hamaoka, M. Nagatani, Y. Ogiso, H. Wakita, H. Yamazaki, T. Kobayashi, M. Ida, H. Nosaka, and Y. Miyamoto, "192-Gbuad Signal Generation using Ultra-broadband Optical Frontend Module Integrated with Bandwidth Multiplexing Function," Proc. of 2019 Optical Fiber Communications Conference (OFC 2019), San Diego, CA, USA, Mar. 2019.

[13] M. Nakamura, T. Sasai, K. Saito, F. Hamaoka, T. Kobayashi, H. Yamazaki, M. Nagatani, Y. Ogiso, H. Wakita, Y. Kisaka, and Y. Miyamoto, "1.0-Tb/s/λ 3840-km and 1.2-Tb/s/λ 1280-km Transmissions with 168-GBaud PCS-QAM Signals Based on AMUX Integrated Frontend Module," Proc. of OFC 2022, W3C.1, San Diego, CA, USA, Mar. 2022.

Munehiko Nagatani

Distinguished Researcher, NTT Device Technology Laboratories and NTT Network Innovation Laboratories.

He received an M.S. and Ph.D. in electrical and electronics engineering from Sophia University, Tokyo, in 2007 and 2021. He joined NTT Photonics Laboratories in 2007, where he engaged in the research and development (R&D) of ultra-high-speed analog and mixed-signal ICs for optical communications systems. He is currently with NTT Device Technology Laboratories and NTT Network Innovation Laboratories, where he is involved in R&D of extreme-broadband ICs for communications and emerging applications. He received the Institute of Electronics, Information and Communication Engineers (IEICE) Young Researcher's Award in 2011, the 67th Maejima Hisoka Encouragement Award in 2022, and the Commendation for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology, the Young Scientists' Award in 2022. He served as a technical program committee (TPC) member for Institute of Electrical and Electronics Engineers (IEEE) Compound Semiconductor Integrated Circuits Symposium (CSICS) from 2014 to 2017. He has been serving as a TPC member for IEEE BiCMOS and Compound Semiconductor Integrated Circuits and Technology Symposium (BCICTS) since 2018 and IEEE International Solid-State Circuits Conference (ISSCC) since 2019. He is a member of IEICE and IEEE.

Teruo Jyo

Researcher, NTT Device Technology Laboratories.

He received a B.E. and M.E. in electronic engineering from Keio University, Kanagawa, in 2012 and 2014. In 2014, he joined NTT Microsystem Integration Laboratories. He is currently a researcher at NTT Device Technology Laboratories. His research interests are high-speed analog circuits for optical fiber communications and terahertz-wave wireless communications and imaging. He received the 2013 IEEE Radio and Wireless Symposium (RWS) Best Student Poster Award and the 2019 IEEE Microwave Theory and Techniques Society (MTT-S) Japan Young Engineer Award.

Hitoshi Wakita

Researcher, NTT Device Technology Laboratories.

He received a B.S. in organic and polymeric materials in 2008 and M.S. in chemistry and materials science in 2010 from Tokyo Institute of Technology. In 2010, he joined NTT Photonics Laboratories, where he has been engaged in R&D of ultrahigh-speed mixed-signal ICs and modules for optical communications systems. He is currently with NTT Device Technology Laboratories. He is a member of IEICE and IEEE.

Tsutomu Takeya

Researcher, NTT Device Technology Laboratories.

He received a B.S., M.S., and Ph.D. in electrical engineering from Keio University, Kanagawa, in 2008, 2010, and 2013. He joined NTT Microsystem Integration Laboratories in 2013, where he engaged in the research of analog and mixed-signal ICs for optical communications. From 2016 to 2020, he was with NTT Electronics, where he engaged in the development of digital coherent DSP for optical communication. He is currently with NTT Device Technology Laboratories researching high-speed analog ICs. He is a member of IEICE and IEEE.

Hiroyuki Takahashi

Senior Research Engineer, Supervisor, NTT Device Technology Laboratories.

He received a B.S. and M.S. in applied physics from Nagoya University in 2001 and 2003. He also received a Ph.D. in advanced electronics and optical science from Osaka University in 2014. In 2003, he joined NTT Microsystem Integration Laboratories. He is engaged in R&D of millimeter-wave/terahertz monolithic microwave ICs. His other research interests include ultrahigh-speed wireless technologies. He was the recipient of the 2008 Young Engineers Prize presented at the European Microwave Integrated Circuits Conference, the 2009 Radio Achievement Award of the Association of Radio Industries and Businesses, the 2010 Asia-Pacific Microwave Conference (APMC) Prize, the 2012 IEEE MTT-S Japan Young Engineer Award, and the 2012 APMC Prize. He is a member of the IEEE MTT-S and IEICE.

Report of the World Telecommunication Development Conference 2022 (WTDC-22)

Memiko Otsuki and Hideo Imanaka

Abstract

The eighth World Telecommunication Development Conference (WTDC-22) of the International Telecommunication Union (ITU) was held in Kigali, Rwanda, from 6 to 16 June 2022. This article gives an overview of the conference and its major deliberations.

Keywords: ITU, WTDC, development

1. Introduction

The International Telecommunication Union (ITU), through its Telecommunication Development Bureau (BDT), organizes a World Telecommunication Development Conference (WTDC) every four years in the period between two Plenipotentiary Conferences to consider topics, projects, and programmes relevant to telecommunication development. WTDC-22 was held in Kigali, Rwanda, from 6 to 16 June 2022. It was held in a hybrid format and attended by 2152 participants from 150 countries. Forty representatives from Japan attended including the Ministry of Internal Affairs and Communications, NTT, NTT DOCOMO, NTT-AT, NEC, Softbank, KDDI, and the ITU Association of Japan.

2. Overview of WTDC

WTDCs set the strategies and objectives for the development of telecommunication/information and communication technology (ICT), providing future direction and guidance to the ITU Telecommunication Development Sector (ITU-D). WTDC-22 established five committees (COMs) and a Working Group of the Plenary (WG-PL) under the Plenary. COM1 was a Steering Committee, COM2 dealt with budget control, COM3 with objectives (including regional initiatives, action plans, study group Questions, and related resolutions), COM4 with working methods, COM5 with editing, and WG-PL with the ITU Strategic Plan and WTDC Declaration. Ms. Paula Ingabire, Minister of ICT and Innovation of Rwanda, was appointed the chair of the conference.

3. Major results and deliverables

WTDC-22 took place just over three months after the invasion of Ukraine by the Russian Federation, which gave the conference a very intense political atmosphere. During the opening session, Ukraine issued a statement condemning Russia's actions and objecting to the appointment of Russian representatives to the chair and vice-chair positions of WTDC-22 and ITU-D. While the statement was supported by many member countries, no compromise was reached, resulting in a secret vote in the Plenary in the second week (see below for details). In addition, WTDC-22 discussed 215 proposals, approved 14 study group Questions, 28 regional initiatives, and 4 new resolutions. The following is a summary of the main discussions and outcomes.

3.1 The Kigali Declaration

The Kigali Declaration outlines the main conclusions and priorities established by the conference and reinforces the political support for the ITU development mission and strategic goals. Recognizing that

Table 1. Questions to be addressed by SG1 and SG2 during the 2022–2025 study period.

SG1: Enabling environment for meaningful connectivity
• Q1/1: Strategies and policies for the deployment of broadband in developing countries • Q2/1: Strategies, policies, regulations and methods of migration to and adoption of digital technologies for broadcasting, including to provide new services for various environments • Q3/1: The use of telecommunications/ICTs for disaster risk reduction and management • Q4/1: Economic aspects of national telecommunications/ICTs • Q5/1: Telecommunications/ICTs for rural and remote areas • Q6/1: Consumer information, protection and rights • Q7/1: Telecommunication/ICT accessibility to enable inclusive communication, especially for persons with disabilities
SG2: Digital transformation
• Q1/2: Sustainable smart cities and communities • Q2/2: Enabling technologies for e-services and applications, including e-health and e-education • Q3/2: Securing information and communication networks: Best practices for developing a culture of cybersecurity • Q4/2: Telecommunication/ICT equipment: Conformance and interoperability, combating counterfeiting and theft of mobile devices • Q5/2: Adoption of telecommunications/ICTs and improving digital skills • Q6/2: ICTs for the environment • Q7/2: Strategies and policies concerning human exposure to electromagnetic fields

some 2.9 billion people remain unconnected in 2021, the WTDC representatives, delegates, and participants committed to accelerating the expansion and use of efficient and up-to-date digital infrastructures, services, and applications; mobilizing financial resources for providing universal, secure, and affordable broadband connectivity; promoting investments in broadband infrastructure deployment; creating alliances and partnerships among stakeholders; urgently mitigating the impact of disasters and the COVID-19 pandemic; and strengthening efforts to tackle environmental and climate-change issues.

3.2 The Kigali Action Plan

The Kigali Action Plan describes the ITU-D priorities and scope of activities and associated Outcomes and Outputs for the next period. It has identified the ITU-D priorities as the following five pillars: (1) Affordable Connectivity, (2) Digital Transformation, (3) Enabling Policy and Regulatory Environment, (4) Resource Mobilization and International Cooperation, and (5) Inclusive and Secure Telecommunications/ICTs for Sustainable Development.

3.3 Regional initiatives

The regional initiatives are intended to identify the principal telecommunication/ICT areas of concern to the six regions (Africa, the Americas, Arab States, Asia-Pacific, the Commonwealth of Independent States (CIS), and Europe), and were adopted by the preparatory meetings of each region. The regional initiative for Asia-Pacific consists of the following five priorities.

(1) Addressing special needs of the least developed countries, small island developing states, including pacific island countries, and landlocked developing countries
(2) Harnessing ICTs to support the digital economy and inclusive digital societies
(3) Fostering development of infrastructure to enhance digital connectivity and connecting the unconnected
(4) Enabling policy and regulatory environments to accelerate digital transformation
(5) Contributing to a secure and resilient ICT environment

3.4 Questions, chairs and vice-chairs of Study Groups (SG)

The Questions to be addressed by SG1 and SG2 during the 2022–2025 study period are summarized in **Table 1**. The conference also elected the chairs and vice-chairs of Telecommunication Development Advisory Group (TDAG) and SGs for the new study period, with the chairs elected from Cote d'Ivoire for SG1, Egypt for SG2, and the US for TDAG. Two vice-chairs were elected from Japan, with the authors of this article, Ms. Memiko Otsuki as vice-chair of SG1 and Mr. Hideo Imanaka as vice-chair of SG2.

As mentioned above, the Russian delegation had candidates for the vice-chair positions in each of TDAG, SG1, and SG2, as well as COM5 of WTDC-22, and the Plenary discussed whether to approve the appointment of these four persons. At the suggestion of Ghana, a secret ballot was conducted, and the appointment of the Russian candidates was rejected.

3.5 Major resolutions discussed at WTDC-22

(1) Resolution 37: Bridging the digital divide

The Asia-Pacific region proposed to add a reference to a high altitude platform station (HAPS) as a means of reducing the digital divide. Although some countries pointed out that no specific technology should be mentioned, HAPS was included as an example of a stratospheric service, which clarified that HAPS is useful for bridging the digital divide.

(2) Resolution 45: Mechanisms for enhancing cooperation on cybersecurity, including countering and combating spam

Arab States and Africa proposed to encourage the ITU to provide information on the existing cooperative framework for data protection (GCA: Global Cybersecurity Agenda). In addition, Europe proposed to revise the resolution to encourage the sharing of information on cybersecurity threats and vulnerabilities. Japan, in cooperation with the United States and other like-minded countries, opposed the inclusion of information provision to the ITU in the resolution because security-related information is highly sensitive. After lengthy discussions in the ad hoc meeting, a compromise was agreed to limit the GCA to only the description of capacity building.

(3) Resolution 66: Information and communication technology, environment, climate change and circular economy

In Resolution 66, there was a proposal from Europe and Africa to include a specific technology, "scientific monitoring and reliable telecommunications (SMART) cables." Japan opposed the description of the specific technology as the resolution should be a high-level document. After deliberations, the general description "undersea sensing technology, including SMART submarine telecommunication cables" was adopted.

(4) New resolution: Digital transformation for sustainable development

Africa and the Arab States had submitted a new draft resolution. It instructed the BDT Director to promote digital transformation in developing countries in line with the Kigali Action Plan. An informal group was formed, and after adjustments, such as excluding references to specific technologies, the resolution was agreed upon as a new resolution.

(5) New resolution (not adopted): Promotion of global development and adoption of Open Radio Access Networks (RAN)

It was proposed by Arab States to conduct research on the development and implementation of Open RAN. Canada, Sweden, and China opposed the new resolution, citing concerns that Open RAN has not been standardized in the ITU Radiocommunication Sector (ITU-R) and the ITU Telecommunication Standardization Sector (ITU-T), that it is a premature technology, and that a resolution should not be created on a specific technology. India, Vietnam, and Arab States and Africa supported the creation of the new resolution because Open RAN would be useful in bridging the digital divide. After the drafting session led by the United States, it was agreed to insert the promotion of information sharing into Resolution 37 (Bridging the digital divide) instead of making it a stand-alone new resolution.

(6) New resolution (not adopted): Use of information and communication technologies to combat pandemics

In addition to the submission by Arab States and the CIS on the new resolution to combat pandemics, Japan proposed to include the promotion of e-health cooperation in the event of a pandemic. However, as the World Telecommunication Standardization Assembly (WTSA-20) agreed to discuss the new resolution on pandemics at the Plenipotentiary Conference (PP-22) in September 2022, it was agreed to withdraw the proposals and consider preparing an input document to PP-22 by the interested members.

4. Election campaign for PP-22

At PP-22, which was held in Bucharest, Romania, in September 2022, elections were scheduled for the council members, secretary general, deputy secretary general, directors of three bureaus, and the members of the Radio Regulations Board. Receptions and coffee breaks were held every day at the WTDC-22 to call for support for the elections. Japan also held an evening reception to request support for the election of the council member and of the director of the Telecommunication Standardization Bureau, for which Mr. Seizo Onoe, the chief standardization strategy officer of NTT, was running.

5. Conclusion

The ITU-D's enthusiasm for pioneering initiatives was well demonstrated at WTDC-22, where subevents, such as the Youth Summit, Partner2Connect (P2C) Digital Development Roundtable, and the Network of Women, were held in conjunction with the conference. The ITU-D is expected to be even more effective this year, as it has made clear its policy to focus on connectivity and digital transformation with the goal of "Connecting the unconnected." The NTT Group has secured two vice-chair positions and will continue to engage in activities related to the development of telecommunications and ICTs and their deployment in developing countries.

Memiko Otsuki
Manager, Tariff and Regulatory Affairs Office/Spectrum Planning Office, NTT DOCOMO, INC.

She received an MPP from the Australian National University (Canberra, Australia), and B.A. in business management from the International Christian University (Tokyo, Japan). She is currently working on her doctorate degree in economics at Kyushu University (Fukuoka, Japan). Ms. Memiko Otsuki currently serves as a manager of the Tariff and Regulatory Affairs Office and the Spectrum Planning Office at NTT DOCOMO, where she has been actively involved in policy and standardization work at ITU and Asia-Pacific Telecommunity (APT). Most recently, she was appointed the vice-chair of the ITU-D SG1 and the rapporteur of Question 12 of ITU-T SG3. Prior to her current position, she was the assistant director of the Global Strategy Division at the Ministry of Internal Affairs and Communications, Japan. In her role, she had been contributing to a number of ITU activities including the Council, the Council Working Group, the Expert Group on the International Telecommunication Regulations, the Informal Experts Group on World Telecommunication/ICT Policy Forum (WTPF-21), the World Summit on the Information Society (WSIS) Forum, the ITU Telecom, the Broadband Commission, and preparatory meetings for WTSA-20, WTDC-22 and PP-22. She has over 15 years of professional experience in the telecommunications/ICT field and has expertise in policy and regulatory issues. She received the ITU-AJ Encouragement Award from the ITU Association of Japan in 2017 for her contributions over the years. She is a member of the Japan Society of Information and Communication Research (JSICR).

Hideo Imanaka
Managing Expert, Standardization Promotion Office, National Institute of Information and Communications Technology.

He received a B.E., M.E., and Ph.D. in electrical engineering from Mie University in 1985, 1987, and 2001. He joined NTT Telecommunication Network Laboratories in 1987. In 2004, he started international standardization activities in ITU-T and the APT standardization program (ASTAP) in NTT Service Integration Laboratories. He was in charge of standardization strategy planning of the entire NTT Group in NTT Research and Development Planning Department from 2010 to 2015. Since joining NTT Advanced Technology Corporation (NTT-AT) in 2015, he has been engaged in standardization consulting. He is currently with National Institute of Information and Communications Technology. He was the rapporteur of Question 1 of SG13 from 2006 to 2010 for the Next-Generation Network and Internet Protocol television standardization work in ITU-T, the chairman of the Internet of Things (IoT) working group in ASTAP from 2012 to 2016 for IoT and smart grid standardization work, and the vice-chairman of Focus Group on IMT-2020 in 2015 for pre-standardization of 5G (fifth-generation mobile communication) networks. He is currently the rapporteur of Question 8 of SG16 in ITU-T for standardization of Immersive Live Experience and the vice-rapporteur of Question 5 of SG2 in ITU-D for international deployment of emergency telecommunications and disaster relief solutions. He was recently appointed the vice-chairman of ITU-D SG2. He received the ITU-AJ Award from the ITU Association of Japan in 2009. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) and the Society of Instrument and Control Engineers (SICE).

Event Report: NTT Communication Science Laboratories Open House 2022

Takaaki Tanaka, Sadao Hiroya, Xiaomeng Wu, Masakazu Ishihata, Hiromi Narimatsu, Yuya Chiba, Naohiro Tawara, Toyomi Meguro, and Tomoki Ookuni

Abstract

NTT Communication Science Laboratories Open House 2022 was held online, the content of which was published on the Open House 2022 web page at noon on June 2nd, 2022. The videos of 6 talks and 29 exhibits presented our latest research efforts in information and human sciences and were viewed more than 5000 times in the first month after the event.

Keywords: information science, human science, artificial intelligence

1. Overview

NTT Communication Science Laboratories (CS Labs) aims to establish cutting-edge technologies that enable heart-to-heart communication between people and between people and computers. We are thus working on a fundamental theory that approaches the essence of human beings and information science, as well as on innovative technologies that will transform society. CS Labs' Open House is held annually to introduce the results of our basic research and innovative leading-edge research with many hands-on intuitive exhibits to those who are engaged in research, development, business, and education.

Open House 2022 was held online, as done last year, considering the situation against the spread of COVID-19. The latest research results were published with recorded lecture videos on the Open House 2022 web page at noon on June 2nd [1]. The content attracted many views in a month from NTT Group employees as well as from businesses, universities, and research institutions. The event content is still available.

In the afternoon of the day before the online event, a pre-event was held at an NTT WEST's open innovation facility QUINTBRIDGE in Osaka as the first on-site exhibition in three years. Visitors showed great interest in the eight exhibits, and the significance of the hands-on exhibits were reaffirmed.

This article summarizes the online event's research talks and exhibits.

2. Keynote speech

Dr. Futoshi Naya, vice president and head of CS Labs, presented a speech entitled "Communication science that adapts to the changing present and creates a sustainable future – Aiming to create technology that brings harmony and symbiosis with people, society, and the environment –," in which he looked back upon the history of NTT and establishment of CS Labs then introduced present and future cutting-edge basic research and technologies (**Photo 1**).

Over the past 30 years since its founding on July 4, 1991, CS Labs has been promoting world-class basic research with the mission of constructing fundamental

Photo 1. Keynote speech (Dr. Futoshi Naya). "Communication science that adapts to the changing present and creates a sustainable future – Aiming to create technology that brings harmony and symbiosis with people, society, and the environment –."

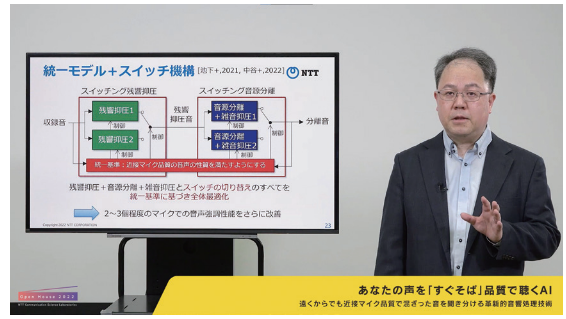


Photo 2. Research talk (Dr. Tomohiro Nakatani). "AI hears your voice as if it were 'right next to you' – Audio processing framework for separating distant sounds with close microphone quality –."

theories on the essence of human beings and information science and creating innovative technologies that will bring about communication that "reaches the heart."

After introducing the latest research results of CS Labs, Dr. Naya declared that CS Labs will promote synergies between different research areas to create new values other than research within each specific area, and explore new frontiers through interdisciplinary collaboration with various fields such as physics, basic mathematics, medicine, biology, social sciences, and philosophy. He also stated that CS Labs will engage in basic research in communication science to bring about natural harmony and symbiosis among people, society, and the environment.

3. Research talks

The following four research talks highlighted recent significant research results and high-profile research themes. Each talk introduced some of the latest research results and provided background and an overview of the research. After each talk, we received many questions from participants in real time during the question-and-answer (Q&A) session, and we could see that the viewers had a high interest in our researches.

3.1 "AI hears your voice as if it were 'right next to you' – Audio processing framework for separating distant sounds with close microphone quality –," Dr. Tomohiro Nakatani, Media Information Laboratory

Dr. Tomohiro Nakatani introduced advanced speech-enhancement techniques for extracting high-quality speech signals from degraded recordings with microphones distant from speakers, as if close microphones recorded them. When distant microphones capture a speech signal, they also record reverberation, background noise, and voices from extraneous speakers, degrading the quality of the captured speech signal. The techniques include a unified model for executing joint processing of dereverberation, denoising, and source separation, as well as a switching mechanism for enabling high-quality processing with a small number of microphones. He also discussed integration with deep-learning-based speech enhancement, e.g., SpeakerBeam (**Photo 2**).

3.2 "Making digital twin of your heart – Personal heart modeling with mobile sensing and its potential application –," Dr. Kunio Kashino, Media Information Laboratory/Bio-Medical Informatics Research Center

Dr. Kunio Kashino introduced the basic research he and co-researchers are conducting to actualize a future of healthcare using digital technology. Expectations for healthcare in modern life using digital technology are increasing every day. To support a more comprehensive approach, they are attempting to digitally model and replicate certain states and functions of the human body. In this talk, he introduced their initial trials focusing on the heart's electrical and

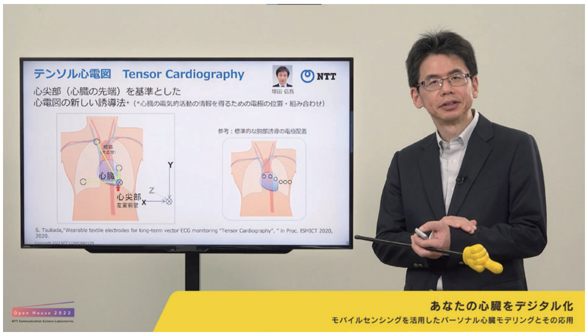


Photo 3. Research talk (Dr. Kunio Kashino). “Making digital twin of your heart – Personal heart modeling with mobile sensing and its potential application –.”

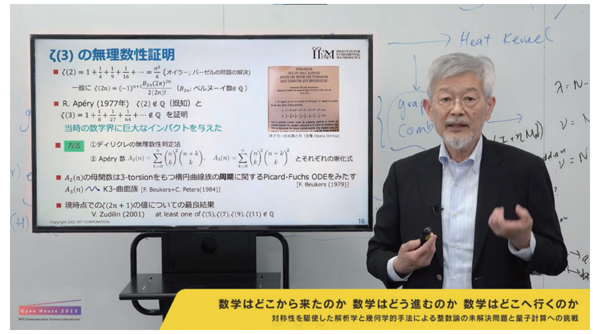


Photo 5. Research talk (Dr. Masato Wakayama). “Where does mathematics come from? How does it proceed? Where it goes? – Challenges to unsolved problems in number theory and quantum computation by symmetry-based analysis and geometry –.”

Photo 4. Research talk (Dr. Kenta Niwa). “Smart traffic coordination via digital twins – Future possibilities of distributed deep learning –.”

mechanical activities. He also discussed possible applications for their approach, including early detection of diseases and rehabilitation after treatment (**Photo 3**).

3.3 “Smart traffic coordination via digital twins – Future possibilities of distributed deep learning –,” Dr. Kenta Niwa, Innovative Communication Laboratory

Dr. Kenta Niwa introduced his latest research on collective intelligence formation and distributed cooperative control via digital twins. For the next innovation in the machine learning field, his research aims to optimally control large-scale systems, e.g., Internet of Things devices in the overall city, by learning collective intelligence from accumulated data on distributed devices. In this talk, Dr. Niwa introduced several research projects associated with smart traffic coordination via learnable digital twins (**Photo 4**).

3.4 “Where does mathematics come from? How does it proceed? Where it goes? – Challenges to unsolved problems in number theory and quantum computation by symmetry-based analysis and geometry –,” Dr. Masato Wakayama, Institute for Fundamental Mathematics

Dr. Masato Wakayama introduced arithmetic geometry, automorphic representations, and graph theory and representation/number theory, which are the mathematicians’ specialties of the Institute for Fundamental Mathematics, which was inaugurated in October 2021. In this talk, with several important historical examples in mathematics, he spoke on what kind of mathematics we have studied, how mathematics is progressing, and where mathematical research is heading in the future. He introduced the research connecting quantum interaction and number theory, and mentioned its position in the stream of modern mathematics (**Photo 5**).

4. Research exhibition

The Open House featured 29 exhibits displaying CS Labs’ latest research results. We categorized them into four areas: *Science of Machine Learning, Science of Communication and Computation, Science of Media Information*, and *Science of Humans*. Each exhibit prepared videos explaining the latest results and published them on the event web page (**Photo 6**).

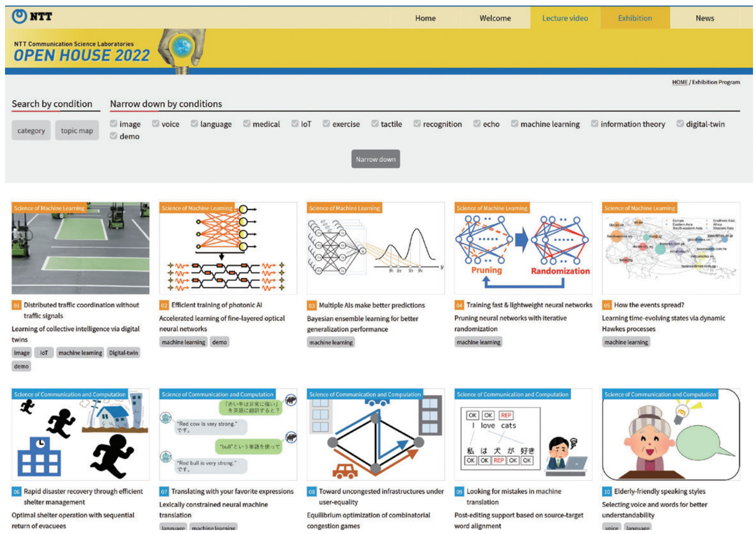


Photo 6. Exhibition web page.

Several provided online demonstrations or demo videos to make them similar to direct demonstrations. We also introduced four Q&A sessions for each of the four areas and received many questions from participants in real time during the sessions. As in the previous year, we also had a Q&A system in which page visitors could freely post questions and comments and CS Labs' researcher answer them. There were more than 50 public questions, and some made showed the enthusiasm of the visitors, such as long question comments and professional questions.

The following list, taken from the Open House website, summarizes the research exhibits in each category.

4.1 Science of Machine Learning

- Distributed traffic coordination without traffic signals
  – Learning of collective intelligence via digital twins –
- Efficient training of photonic AI
  – Accelerated learning of fine-layered optical neural networks –
- Multiple AIs make better predictions
  – Bayesian ensemble learning for better generalization performance –
- Training fast & lightweight neural networks
  – Pruning neural networks with iterative randomization –
- How the events spread?
  – Learning time-evolving states via dynamic Hawkes processes –

4.2 Science of Communication and Computation

- Rapid disaster recovery through efficient shelter management
  – Optimal shelter operation with sequential return of evacuees –
- Translating with your favorite expressions
  – Lexically constrained neural machine translation –
- Toward uncongested infrastructures under user-equality
  – Equilibrium optimization of combinatorial congestion games –
- Looking for mistakes in machine translation
  – Post-editing support based on source-target word alignment –
- Elderly-friendly speaking styles
  – Selecting voice and words for better understandability –
- Talking with AIs about views from a vehicle
  – Casual-dialog system based on scenery and nearby information –
- Toward secure cryptography against quantum attacks
  – Quantum algorithm for finding collisions of hash functions –
- Where does the wonder of numbers come from?
  – Finding new arithmetic phenomena via generalized motives –

4.3 Science of Media Information

- "Huh? What do you mean?" Summarize a long story short
  – Robust speech summarization against speech recognition errors –
- Flexible bokeh renderer based on predicted depth
  – Deep generative model for learning depth and bokeh effects only from natural images –
- Heart health monitoring with sounds and electric signals
  – Estimating heart activities from multichannel sounds and ECG signals –
- Controlling facial expressions in face image from speech
  – Crossmodal action unit sequence estimation and image-to-image mapping –
- Maintain comfortable visibility anytime, anywhere
  – Image blending with content-adaptive visibility predictor –

4.4 Science of Humans

- Gazing and talking help infants learn
  – Elucidating effects of social cues on infants' object learning –
- Why do people hesitate to use contact tracing apps?
  – Social factors influencing adoption of COCOA –
- Is the rising fastball a perceptual illusion?
  – Modifying pitched ball perception by VR –
- Mental skills of esports experts revealed by brain measurement
  – The relationship between frontal neural oscillation and performance –
- Unveiling the auditory system with a neural network
  – Approaches to cochlear implant and binaural processing –
- How does mindfulness meditation reduce stress?
  – Autonomic and endocrinological variation by meditation style –
- Measuring well-being through diverse aspects
  – Well-being in terms of mental states, values, and idea of self –
- Faster walking by moving the wall forward
  – Vision-based speedometer regulates human walking –
- Fingertip illusions direct the mind
  – How the brain decodes pulling sensations –
- What do we want to touch?
  – Understanding of desire to touch using large-scale Twitter data –
- Eyes as a window of our mind
  – Pupil size tracks subjective perceptual changes –

5. Special lecture

We asked Professor Kazuo Okanoya, professor of Advanced Comprehensive Research Organization, Teikyo University, to give a special lecture entitled "Animal mind and communicative behavior," and conduct a panel discussion with CS Labs' researchers Makio Kashino, Tessei Kobayashi, and Sadao Hiroya.

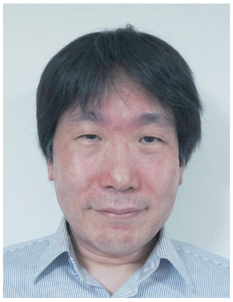
In the lecture, he talked about his research on biopsychology of communication to speculate on the mental experiences of animals and elucidate the evolution and neural mechanisms of such mental experiences by measuring neural activities that correlate with behavior. He explained the need for a method of studying the "animal mind" as well as the science of the "human mind," which is based on anthropomorphism, although the study of the "animal mind" has been discouraged as anthropomorphism. He explained that in communicative behavior, sending a certain signal changes the behavior of the receiver, which in turn allows the sender to gain some benefit, and he introduced videos of experiments with examples of animal behavior for such communication. He concluded by stating that the following are necessary for the mind of artificial intelligence: "communication," similar to what animals do; "theory of mind," a system for predicting the behavior of others; and "mirror neurons," a system for converting the behavior of others into its own behavior.

6. Concluding remarks

Open House 2022 was held as an event to present our latest results on a website. The lecture videos were viewed more than 4000 times in June by various groups of users. Participants provided many valuable opinions that will encourage us in further research activities through interactive means of communication with CS Labs' researchers such as the Q&A system and Q&A session at the exhibition and lectures. In closing, we would like to offer our sincere thanks to all the participants of this online event.

Reference

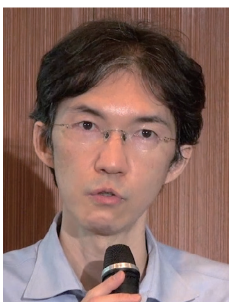
[1] Website of NTT Communication Science Laboratories Open House 2022, http://www.kecl.ntt.co.jp/openhouse/2022/index_en.html



Takaaki Tanaka

Senior Research Scientist, Linguistic Intelligence Research Group, Innovative Communication Laboratory, NTT Communication Science Laboratories.

He received a B.S., M.S., and Ph.D. from Osaka University. He joined NTT laboratories in 1996 and has been engaged in research on natural language processing including machine translation and language analysis. His current research interests include natural language generation and crosslingual semantic textual similarity based on syntactic and semantic analysis. He is a member of Association for Computational Linguistics, Association for Natural Language Processing, and the Japanese Society of Artificial Intelligence (JSAI).



Sadao Hiroya

Senior Research Scientist, Sensory and Motor Research Group, Human Information Science Laboratory, NTT Communication Science Laboratories.

He received an M.E. and Ph.D. from the Tokyo Institute of Technology in 2001 and 2006. He joined NTT Communication Science Laboratories in 2001. From 2001 to 2003, he was also a researcher on the CREST project of the Japan Science and Technology Agency. He was a visiting scholar at Boston University from 2007 to 2008. He received the first Itakura Prize Innovative Young Researcher Award and the ASJ Activity Contribution Award by the Acoustical Society of Japan (ASJ). His current research interests include the links between production and perception of speech, functional brain imaging, and musical performance. He is a member of the Institute of Electrical and Electronics Engineers (IEEE), the Institute of Electronics, Information and Communication Engineers (IEICE), and ASJ.

Xiaomeng Wu

Senior Research Scientist, Recognition Research Group, Media Information Laboratory, NTT Communication Science Laboratories.

He received a B.S. from the University of Shanghai for Science and Technology, China, in 2001, and M.S. and Ph.D. from the University of Tokyo, Japan, in 2004 and 2007. He joined NTT in 2013 and has been engaged in research on image retrieval and recognition. His current research interests include computer vision, image processing, and multimedia. He is a member of IEEE and IEICE.

Masakazu Ishihata

Senior Research Scientist, Learning and Intelligent Systems Research Group, Innovative Communication Laboratory, NTT Communication Science Laboratories.

He received a B.E., M.E., and Ph.D. in engineering from Tokyo Institute of Technology in 2008, 2010, and 2013. He joined NTT laboratories in 2013 and has been engaged in research on artificial intelligence (AI) and machine learning (ML). His current research interests include discrete structure manipulation systems and their application to AI and ML. He received the JSAI Incentive Award in 2010 and the JSAI Best Paper Award in 2013.

Hiromi Narimatsu

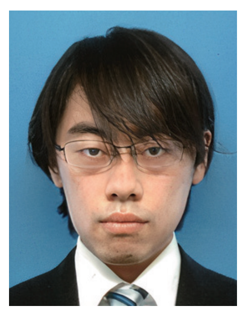
Senior Research Scientist, Human Information Science Laboratory and Innovative Communication Laboratory, NTT Communication Science Laboratories.

She received an M.E. and Ph.D. in engineering from the University of Electro-Communications, Tokyo, in 2011 and 2017. She joined NTT in 2011. Her research interests include natural language processing, spoken dialogue systems, and mathematical modeling. She is a member of IEEE, IEICE, Information Processing Society of Japan (IPSJ) and JSAI.

Yuya Chiba

Research Scientist, Interaction Research Group, Innovative Communication Laboratory, NTT Communication Science Laboratories.

He received a B.E., M.E., and Ph.D. in engineering from Tohoku University, Miyagi, in 2010, 2012, and 2015. He was an assistant professor at the Graduate School of Engineering, Tohoku University, from 2016 to 2020. He is currently a research scientist at NTT Communication Science Laboratories. His research interests include spoken dialogue systems, multimodal dialogue systems, and human-centric interfaces.



Naohiro Tawara

Research Scientist, Media Information Laboratory, NTT Communication Science Laboratories.

He received a Ph.D. in engineering from Waseda University, Tokyo, in 2016. After serving as a lecturer at Waseda University, he joined NTT Communication Science Laboratories in 2019. He has been engaged in research on speech processing including speaker recognition, age estimation, and language modeling. His current research interests include profiling human attributes such as age, height, and emotion from voice. He is a member of IEEE, IEICE, IPSJ, and ASJ. He received the Awaya Prize Young Researcher Award from ASJ in 2018 and Yamashita SIG Research Award from IPSJ in 2019.

Toyomi Meguro

Research Scientist, Interaction Research Group, Innovative Communication Laboratory and Research and Planning Section, NTT Communication Science Laboratories.

She received an M.E. in engineering from Tohoku University, Miyagi, in 2008 and joined NTT Communication Science Laboratories the same year. Her research interests include spoken dialogue systems and social skill training. She received the COLING 2010 Best paper finalist from the COLING in 2010 and the JSAI Annual Conference Award from JSAI in 2014.

Tomoki Ookuni

Senior Research Scientist, Research and Planning Section, NTT Communication Science Laboratories.

He received a B.A. in economics from Kyoto University in 1990 and M.A. in marketing science from Osaka Prefecture University in 2004. He joined NTT in 1990 and has been engaged in managing research and development at both NTT WEST and NTT since 1998. He joined NTT Communication Science Laboratories in 2019.

External Awards

Certificate of Appreciation

Winner: Minoru Inomata, NTT Access Network Service Systems Laboratories
Date: August 31, 2022
Organization: The Institute of Electrical and Electronics Engineers (IEEE)

For serving as a Technical Program Committee member at the 2022 IEEE International Workshop on Electromagnetics: Applications and Student Innovation Competition (iWEM 2022).

Distinguished Contributions Award

Winner: Shuto Yamamoto, NTT Network Innovation Laboratories
Date: September 7, 2022
Organization: The Institute of Electronics, Information and Communication Engineers (IEICE) Communications Society

For contribution as an associated editor in the editorial committee of IEICE Communications Express.

Contribution Award

Winner: Junko Takahashi, NTT Social Informatics Laboratories
Date: September 7, 2022
Organization: IEICE Technical Committee on Hardware Security

For contribution to the operation and activities of the Technical Committee on Hardware Security.

Papers Published in Technical Journals and Conference Proceedings

Nonlinear Time-frequency Analysis of Lightning Strike Surge Current Waveforms Recorded at Gasing Hill, Kuala Lumpur

F. Ishiyama and M. Maruyama

IEEE 18th International Colloquium on Signal Processing & Applications (CSPA 2022), Selangor, Malaysia, May 2022.

We are investigating countermeasure technique against electromagnetic noise on telecommunication-related equipment, and we have developed a method of nonlinear time-frequency analysis for the purpose. We applied this method to the analysis of lightning strike surge current waveforms. Such waveforms produce highly nonlinear transient time series, and we have revealed their nonlinear characteristics. The relaxation process of these waveforms begins with an oscillating region with an extremely large decay rate, and is followed by a non-oscillating region with a nonlinearly increasing decay rate.

Nonlinear Time-frequency Analysis of Nonlinearly Coupled Electromagnetic Noise Sources

F. Ishiyama and M. Maruyama

CSPA 2022, Selangor, Malaysia, May 2022.

We are investigating countermeasure techniques against electromagnetic noise on telecommunication-related equipment, and we have developed a method of nonlinear time-frequency analysis for the purpose. We are looking for a noise fingerprint signature, which is stable under nonlinearly coupled conditions with surrounding noise sources through power line. We considered such a case, and found that the time-frequency characteristics obtained with our method are fairly stable. Therefore, the characteristics represent a possible candidate for a noise fingerprint under such interdependent condition.

Visually Equivalent Light Field 3-D for Portable Displays

M. Date, S. Shimizu, and S. Yamamoto

IEEE Transactions on Industry Applications, Vol. 58, No. 5, pp. 5659–5666, September/October 2022.

Highly realistic 3-D displays that can reproduce object images to look like physical objects are utilized for natural and correct remote operation in industrial scenes. Therefore, we developed a visually equivalent light field 3-D (VELF3D) display that can produce highly realistic, accurate images with a high resolution and a smooth, accurate motion parallax. However, the observation distance is slightly long, and users cannot reach the displayed images. Therefore, we aim to develop a tablet-computer-type VELF3D display that enables users to touch the displayed objects. The display viewpoint density has been increased to achieve a shorter observation distance, while maintaining the display depth range. Because higher resolutions are required for a close observation distance and increased display viewpoints, we aimed to improve the effective resolution using almost the same pixel pitch display panel. Therefore, we built a prototype that combines a vertical red, green, blue stripe display panel and a parallax barrier with subpixel width slits. We confirmed effective resolution improvement by tiny subjective tests. This method also helps increase the depth range of the display when it is observed from a normal distance.