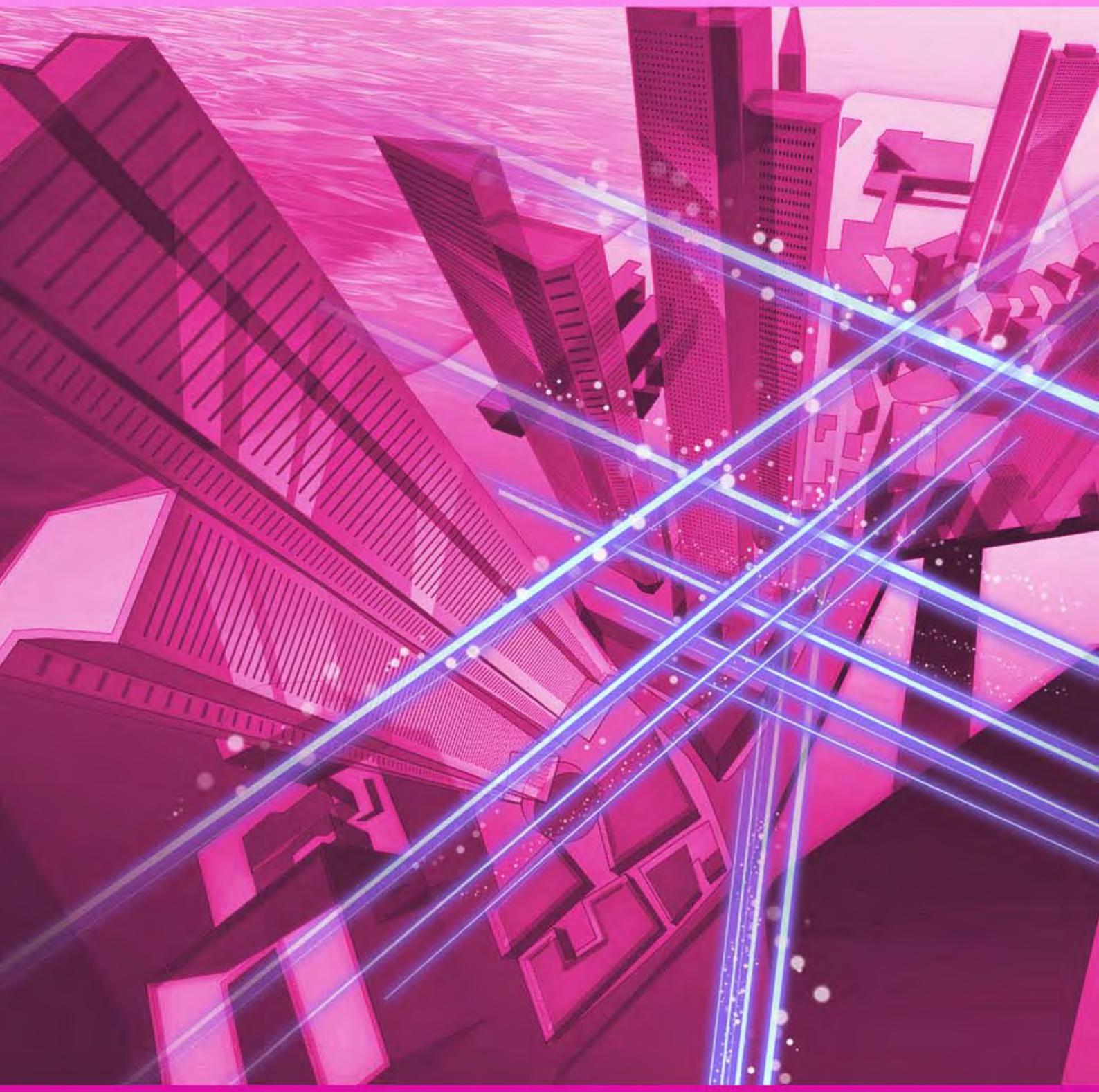
NTT Technical Review

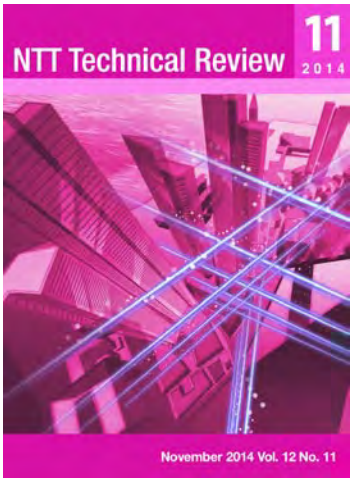
11
2014



November 2014 Vol. 12 No. 11

NTT Technical Review

November 2014 Vol. 12 No. 11



Global Standardization Activities

Information

Papers Published in Technical Journals and Conference Proceedings

Passing the Ball Forward Towards the Goal of Customer Satisfaction

Tetsuya Funabashi, Senior Executive Vice President, NTT Communications

Overview

Today, the information and communication technology (ICT) industry is witnessing increasingly intensifying competition. A company in the industry therefore must consider the customer's point of view at all times and steer a course through a *red ocean* market—highly competitive and with defined limits—by making full use of its core strengths. We asked Tetsuya Funabashi, NTT Communications Senior Executive Vice President, about the mindset and strategy that the company must adopt to navigate these rough waters.

Keywords: Vision 2015, innovation, cloud

Aggressively fostering innovation with execution and implementation in mind

—Mr. Funabashi, please accept our congratulations on assuming the office of Senior Executive Vice President. To begin with, please tell us about your goals in this new position.

I would like to aggressively foster innovation without being bound by the past. And what I mean by "innovation" is to create an atmosphere that results in people feeling, "Something has changed here—this is great!" with respect to all sorts of company endeavors such as processes, technology development, employee conduct, and customer relations.

I myself oversee technology, operations of sales and service, and information security. As for technology and operations, I feel that I can maximize the know-how I built up as head of Customer Services in making these areas of the company even better.

Information security, meanwhile, is a field that has a broad and deep range of applications, as you know. I would like to approach this field from a new perspective.

Finally, as a basic stance, I will continue to ask myself whether our customers are truly happy with our services, whether they are satisfied with the level of quality that we are offering, and whether there are any areas that we need to improve in our services.

—What does the past way of doing things mean to you? And what do you think needs to change going forward?

In the corporate world, having ideal objectives in expanding business does not guarantee that those objectives can be met given the limited amount of resources and the various conditions. I mean, in the past, we might have blamed a negative outcome on the external circumstances.

In services to date, I feel that there is still much room for improvement and enhancement. What I would like to do is to pool the knowledge and wisdom of all our employees to make improvements and to move forward, even if only by a little, by having the resolve to execute and implement various ideas.

No doubt you have heard the word "DevOps" before. As a portmanteau word of "development" and

"operations," it refers to an increasingly popular technique that aims to produce products and services with a high degree of perfection in a relatively short period of time through collaboration between development and operations personnel. I believe that unless we adopt this kind of style that can efficiently drive a project forward while leveraging the strong points of all departments involved, we will not be able to survive in this era of relentless change. I strongly feel the need for making changes to the status quo.

Five pillars of Vision 2015

—Based on the stance you just described, what specific forms of operation are you aiming for?

ICT and the telecommunications business are being forced into severely competitive conditions while amazing advances in technology continue to take place. I don't think a company can survive in such an environment unless it keeps the competition in mind and continuously reaches out to its customers.

Our financial results for the fiscal year 2013, which ended March 31, 2014, revealed an increase in consolidated revenues for the first time in six years and a record consolidated operating profit for three consecutive years. In fiscal year 2014, we want to see this increase in consolidated revenues continue and to maintain the profit level of fiscal year 2013.

With these objectives in mind, we announced our new corporate vision called "Vision 2015." To achieve this vision, we will promote our Global Cloud Vision (GCV) and work to provide our customers with low-cost, quickly constructed, and customer-tailored ICT environments, centered on the cloud. Specifically, we will promote services that strengthen the link between the network and the cloud in the form of a "carrier cloud." We will foster the evolution of services that use virtualization technology, including network virtualization techniques such as software-defined networking (SDN) and network functions virtualization (NFV), and we will enhance our efforts in automating operations. Our work here in achieving Vision 2015 will be based on five main pillars.

1. Operations

The first pillar is operations. Here, I would like to highlight integrated management that focuses on enhancing not just service quality—the obvious feature of a provided service—but also on operations quality that cannot be seen. In the end, the customer will set the standards that we must target. Naturally, we want to provide services that bring satisfaction to our customers. To this end, we will work to create services that meet and beat the level of quality that they expect by automating all sorts of operations to raise system efficiency and by incorporating new technologies such as machine learning.

2. Security

The second pillar is security. At NTT Communications, we make use of the Security Information and Event Management (SIEM) engine, one of the latest research and development (R&D) achievements to come out of NTT laboratories in order to provide managed security services and integrated risk management services under the WideAngle service brand. Going forward, we plan to expand the scope of applications of these technologies and services and to include the know-how obtained from daily operations with the aim of providing even better services for our customers.

Security, by the way, is not just a matter of technology—educating personnel is also a very important element. The development of security engineers is essential, I believe, for the entire NTT Group and not just NTT Communications.

3. Big data

The third pillar is big data. In this area, I would like to provide products and services that cannot be imitated by our competitors and that can expand our revenues. I intend to do this by using not only the management, sales, and systems information we have as a service provider but also real-time information on user experiences. In this way, I think that we can quickly provide information and environments

tailored to customer needs and desires.

For example, functions that customers are finding difficult to use can be identified from error information generated during customer operations or from comments received via Twitter or other social networking services by using the "050 Plus" service provided by NTT Communications. Whenever customers feel that something is strange, we check the network and its facilities as well as the state of the functions that are being talked about, and we make revisions to those functions as needed. In this way, we have been able to provide our customers with services that are even easier to use. Skillfully absorbing people's feelings such as, "Something is strange here" can lead to interesting developments!

4. Virtualization technology

The fourth pillar is virtualization technology. We will promote network virtualization through SDN and NFV to enable customers to rapidly scale-up or scale-down their networks and to reduce costs while maximizing the efficiency of resource usage within the company. We will also expand the cloud, which is a hardware virtualization technology, so that an entire datacenter can be operated at high efficiency.

5. Datacenters

Finally, the fifth pillar is datacenters. Today, the cloud is expanding at a dramatic pace throughout the world. When contemplating the construction of a datacenter in Japan, it must be kept in mind that this is a country with high land costs and electricity rates as well as a high occurrence of earthquakes. We must therefore give serious thought as to how to provide stable services at low cost under these conditions.

Against this background, we are working diligently to develop an optimal strategy to provide our customers with even better datacenter services. This involves not only datacenter projects within Japan but also possible mergers and acquisitions with overseas enterprises who are already developing high-quality, low-cost datacenter businesses in locations with low environmental risk.

Taking a "rugby-like" stance—passing the ball, and moving forward at all times

—I see how all standards originate with the customer. It is necessary to change the mindset from the viewpoint of the provider to the viewpoint of the customer, isn't it?

We must radically change our way of thinking not only in terms of viewpoint but in terms of awareness too. The results of this shift in thinking will depend on whether we think of it as something difficult to achieve or something that is enjoyable to pursue. I would like everyone to think of it as a very enjoyable pursuit. Moreover, so that all of us can enjoy our work on a day-in and day-out basis, I would like everyone to have both a feeling of "We can do it!" and "We can learn from failure."

As you know, NTT Communications has a rugby team called the "Shining Arcs" that compete in the Top League, a corporate rugby league in Japan. There is also a rugby team called the "NTT Hibiya Sunshiners" whose members include players who have retired from the Shining Arcs. I routinely attend matches to offer my support, and once, when I was watching a match, it suddenly occurred to me that this stance that I would like us to adopt is much like the game of rugby. A common expression in rugby is "One for all, all for one." Each and every player is facing forward, yet each has fellow players to the side or behind who are fighting for the same objective, that is, victory. In short, all members are working together and helping each other out by passing the ball toward a common goal.

—Are specific measures being taken to get all employees to aim for a common goal?

After entering the company and getting accustomed to work, there is a time when every employee suddenly realizes that he or she cannot do anything alone. When we are young, there are perhaps some situations in which we can accomplish something on our own, but doing work as a team provides the experience of sharing a sense of achievement with all team members regardless of good or bad results.

To cultivate this mindset, efforts have been made over the years to reform the "corporate culture." These include schemes for getting people on the same hierarchical level—for example, department managers, section managers, and project managers—to share some time together and get to know each other, such as by talking about recent happenings over a meal. The idea here is to break down walls to prevent "big company disease" from setting in, in which neighboring teams don't know what the other is doing. There have also been occasions in which employees have been brought together at training camps away from the office to exchange ideas and share feelings about work. It can take a while for people to reveal their inner selves and thoughts at such gatherings, but the end result is effective. However, not all staff members move in the same direction at the same time, and such differences also need to be accepted. I believe that adopting this team mindset in one's work is extremely important.

Putting yourself in the customer's shoes: it's not just results that count, it's the process too!

—What does it mean to give customers a sense of ease, or a feeling of security?

In short, it means putting yourself in the customer's shoes. For example, if a person went to a clinic because of stomach pain, and the doctor abruptly said, "Why did you leave this until it got this bad?" I don't think that person would go to that clinic again. Obviously, if that person knew that something was wrong from the start, his or her condition would never have gotten to that point. However, if the doctor were to approach the patient's problem with a caring attitude by saying something like, "What's the matter? Let's take a look!" instead of obsessing on what or who to blame, I think the patient's degree of satisfaction would surely turn out to be different. Even if there is no guarantee that the patient's condition can be completely cured, having a caring attitude can generate a very positive impression on the patient.

Similarly, in our work, when a customer is faced with a problem or difficulty, it is vital that we stand in the customer's corner with a caring attitude.

—Mr. Funabashi, can you say a few words to our researchers?

It's a matter of course that the standard approach at NTT laboratories is to create R&D solutions for problems of immediate concern. From here on, however, I would like to see NTT researchers adopt a more medium- and long-term approach and work to develop cutting-edge technologies that can compete on the world stage like the SIEM engine that I mentioned earlier, as well as dream-inspiring technologies that only the NTT Group can provide.

—And finally, can you leave a message for all NTT Communications employees?

The strength behind NTT Communications is its great number of outstanding employees. There are about 7000 employees in Japan, all of whom are of serious mind and very dedicated to their work. Being of serious mind with regards to work means always being truthful and apologizing sincerely when you make a mistake.

We have been in an era in which everything is created within the company or group starting from zero. I feel that this approach, however, makes it difficult to keep up with fast-moving developments in today's world. There are already many convenient and easy-to-use tools that are readily available from outside the company. Going forward, we must be able to skill-

fully use these tools in service development. We must continuously strive to make improvements without losing hope. If we stick to this approach, we will be able to rise to an even higher level of excellence and come out a winner! Let's continue to pass the ball and work as a team with all our might without fear of failure. I have great faith in all of you!

Interviewee profile

■ Career highlights

Tetsuya Funabashi joined Nippon Telegraph and Telephone Public Corporation in 1982. At NTT Communications, he has served as General Manager of Ubiquitous Services, General Manager of Financial Innovation Systems, General Manager of Customer Services, and Senior Vice President, General Manager of Customer Services. He has been serving as Senior Executive Vice President since June 2014.

The Evolution of Basic Research

Eisaku Maeda

Abstract

Basic research gives rise to innovation through new discoveries and inventions, and this can lead to changes in the structure of industry and our lifestyles. However, it is also true that such success stories are rare, and basic research carries a high degree of risk. This article analyzes the historical evolution of technologies born from the activities of NTT Communication Science Laboratories and clarifies the future strategy for the promotion of basic research and some issues that should be addressed based on previous cases.

Keywords: basic research, innovation, research and development

1. Introduction

It is 23 years since the establishment of NTT Communication Science Laboratories (NTT CS Labs), and 14 years since the start of basic research under a new system after the reorganization of NTT. During this period, a definite framework was formed at NTT CS Labs, and technologies born out of our basic research have gradually made their way out into the world. I believe that we have now reached a major turning point on our way towards the next phase. This series of Feature Articles introduces the technical details and future prospects of seven research achievements presented at the NTT CS Labs Open House 2014 exhibition. I will start by discussing the nature and role of basic research in research and development (R&D), the way in which basic research leads to innovation, and our strategy for promoting innovation.

2. *Invisible* innovation in information technology

Innovations in information technology occupy quite a special position compared with innovations in mechanical technologies such as the steam engine or letterpress printing. Information specialists complain that even though people talk about innovations for certain products and services, and even though information technology is used in these products and services, the importance of information technology is not being conveyed to people in general. That is, information technology cannot be *seen* in the innovation. Even if we use Google's search engine or an Apple iPhone every day, we cannot necessarily tell precisely what sort of information technology lies at the core of products and services such as these. So what sort of role does information technology play behind innovations such as the Nobel prize-winning work on iPS (induced pluripotent stem) cells or the *Hatsune Miku* singing software that is creating a stir in the music industry? Most people probably have no idea. This is because by the time new scientific discoveries have been incorporated into products and services that affect our everyday lives, they are no longer visible from the outside.

Furthermore, even though major technological innovations occur every few years in each field of information technology, these innovations are not immediately reflected in products or services. Even among people that work in information technology, these technical innovations are only noticed by specialists in related fields. These are the invisible innovations that I am referring to.

For example, in recent years, speech recognition technology has been put to use in a wide range of practical services, such as a system that records transcripts of debates in the House of Representatives. One of the things that made this system possible was the introduction of speech recognition using a weighted finite state transducer (WFST) in 2003, which enabled the implementation of speech

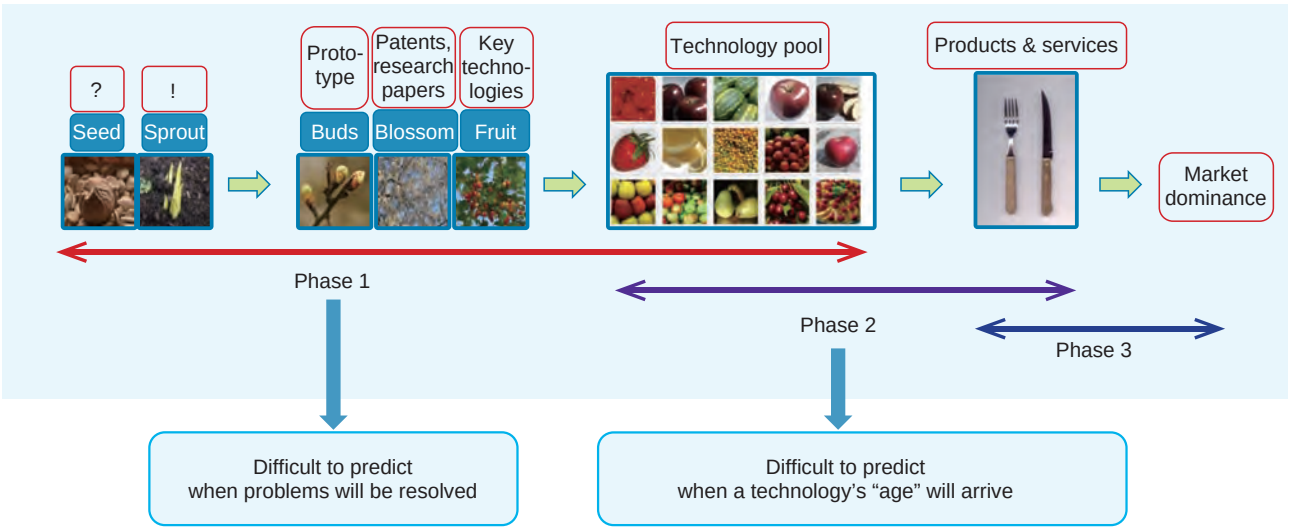


Fig. 1. The road from "basic research" to services.

recognition with a huge vocabulary of 2 million words [1]. It goes without saying that grasping this sort of hidden innovation as quickly as possible is the key to success in R&D, the development of new services, and the creation of new markets.

3. The cultivation and evolution of fruitful technologies

So when are hidden innovations produced in basic research, and how do we go about creating more of them? I have designed an illustration to show the workings of basic research (**Fig. 1**). The seeds of new research arrive as small flashes of inspiration, for example, the discovery of a new problem, or a new way of looking at a simple concept. If they are kept watered, some of these seeds will germinate and grow after a while, and given fertilizer and sunlight, they will blossom into research papers and patents. With a bit of luck, they will eventually bear fruit that is ripe for harvesting. Phase 1 of this process—from sowing the seeds to cultivating the plants and harvesting the fruit—forms the core of basic research and is the most important phase. Even plants that produce fine blossoms are sometimes of no value. There is also no guarantee that fine fruit will be produced by plants given plenty of water and fertilizer. While storing a range of different fruits, we obtain new seeds from these fruits, and once again, these are sown to promote the evolution of even more technologies.

The mission and value of basic research lies in tackling difficult problems without worrying about the risks. It is therefore usually impossible even for experts to predict when these problems will be solved. This is the biggest issue in Phase 1. One of the recent cases at NTT CS Labs is the invention of *Buru-Navi3*, which is a device that uses human sensory characteristics to create a tugging sensation. After we invented *Buru-Navi1* in 2004, we thought it would be difficult to miniaturize this device while maintaining the same tugging effect [2]. The solution to this issue suddenly came in 2014, when we discovered that it is possible to make a device 20 times smaller than the *Buru-Navi1* without impairing its tugging effect. This new device is introduced in the article "Buru-Navi3 Gives You a Feeling of Being Pulled" [3].

4. Fruit provides people with nourishment

In the same way that fruit provides nourishment only when eaten, the fruits of research have value only when they are used as technologies. The basic research achievements of NTT CS Labs are finding their way into technologies that are put to practical use in the real world. Some representative examples are listed in **Table 1**. By analyzing these examples, we can see that it can often take 10 years or more for seeds to grow to fruition, and that a lot of time is also needed to make this fruit available for consumption (Phase 2). Once a technology has been perfected, it will not be used unless it meets the needs of the current age. In most cases, it is difficult to predict when

Table 1. Examples of NTT CS Labs' achievements.

		Start of research	Representative results of basic research	Contributions to actual services
(1)	Robust media search [RMS]	1993	1995: Fast image search 1998: Fast time-series (audio/video) search 2004: Extremely robust media search with respect to severe noise and distortions 2013: Instance search	2004: Name-that-tune service via mobile phones 2007: Music copyright clearance for broadcasters 2008: Identification of known audio/video content embedded in media files in the Internet 2010: Second screen using smartphones
(2)	WFST-based speech recognition	2000	2004: Extremely large vocabulary (2 million words) recognition decoder 2010: Unified acoustic-language model learning	2011: Minute-taking system for House of Representatives
(3)	Reverberation control [REVTRINA]	2003	2005: Development of basic principles 2009: Surround technology based on reverberation control	2009: Pro reverberation control software 2012: Consumer surround speakers 2013: Pro surround equipment
(4)	Question answering	2001	2003: Japanese Q&A system SAIQA 2009: NAZEQA, capable of answering "Why?"	2012: Shabette Concier (talking concierge)
(5)	Statistical machine translation	2003	2009: Dependency parsing with semi-supervised learning 2010: Pre-ordering by head finalization	2011: Outperformed rule-based translation 2013: JAPIO starts Chinese patent translation service
(6)	Material perception information science	1998	2007: Proposal of a new theory of material perception (Published in *Nature*) 2012: Finding of a novel object after-effect	2010: Inauguration of Scientific Research on Innovative Areas "Shitsukan (material perception)"
(7)	Pseudo-Attraction Force [Buru-Navi]	2003	2004: Development of pseudo-attraction force haptic compass (Buru-Navi1) 2014: Miniaturization (Buru-Navi3)	—

JAPIO: Japan Patent Information Organization
SAIQA: System for Advanced Question Answering

the *age* of a technology will arrive, which is the main issue in Phase 2. Until such time, a technology must be maintained and protected in a technology pool, and a system must be established for bringing the technology out into the world as soon as its time has come.

The following paragraphs describe some specific examples of technologies listed in the table.

(1) Robust media search (RMS) began in 1993 with research into image search techniques that can quickly find an image fragment in a larger image (as in the "Where's Wally?" picture books). This technology was developed to include music search and video search functions, and in around 2008 it came to play a major role in the identification of uploaded video content and the protection of music copyright in broadcasting [4]. Today, it is still evolving to allow searching for specific instances of a specific object in video images, as discussed in the article "Instance Search Technology for Finding Specific Objects in Movies" [5].

(2) WFST-based speech recognition is a technology that was given a major boost by the introduction of WFST as mentioned above, but the introduction of a deep-learning technique has led to a new wave of development [1]. Also, the basic principles of (3) reverberation control technology (REVTRINA) were figured out about five years ago, and this technology is now being introduced into a wide variety of professional and consumer devices. This technology is introduced in the article "Enhancing Speech Quality and Music Experience with Reverberation Control Technology" [6].

(4) Question answering technology is a basic element of NTT DOCOMO's *Shabette Concier* service, and it grew out of SAIQA (System for Advanced Question Answering) that NTT CS Labs started researching back in 2001 [7, 8]. At the 2003 NTT CS Labs Open House exhibition, we connected it to a speech recognition system with a vocabulary of 2 million words to produce a speech-based question answering demonstration. However, it took 10 years for this technology to find its way into the real world. The fruits of this research in question answering technology led to further advances with the introduction of statistical machine learning methods into natural

language processing, which has progressed rapidly since the turn of the century. This introduction of statistical machine learning technology has also played a large role in the paradigm shift whereby (5) statistical machine translation is replacing conventional rule-based machine translation [9].

(6) Material perception information science came to the fore in 2010 as a new field of technical research promoted by the Ministry of Education, Culture, Sports, Science and Technology, and academic activity in this field is expanding rapidly. This came about through a joint study by NTT CS Labs and MIT (Massachusetts Institute of Technology) that started in 2000. Since the results of this study were published in Nature in 2007 [10], the research has been extended to include the other senses as well as sight [11]. More information can be found in the article "Recognizing Liquid from Image Motion and Image Deformation" [12]. This material perception information science and (7) Buru-Navi can be described as the fruits of research that are still waiting for their age to arrive.

A lot of work is also being done on technologies that are not yet complete. These technologies are introduced in the articles "Reading the Implicit Mind from the Body" [13], "Quantum Computing Beyond Integer Factorization—Exploring the Potential of Quantum Search" [14], and "Capturing Sound by Light: Towards Massive Channel Audio Sensing via LEDs and Video Cameras" [15].

5. Evolution of basic research

The pursuit of basic research will never be a special endeavor that takes place far away from the ordinary world. If we are to take on the challenges of the real world, we must be a part of it. John Pierce, the former executive director of Bell Laboratories, once said that, based on his own experience, "Ideas and plans are essential for innovation, but the time has to be right [16]." After leaving Bell, Pierce also became well known as one of the academic pioneers in computer music, but he is also known among speech recognition researchers as the person that pulled the plug on research on speech recognition at Bell Laboratories in the 1970s. In an article for the Acoustical Society of America, he wrote, "General-purpose speech recognition seems far away. Special-purpose speech recognition is severely limited. It would seem appropriate for people to ask themselves why they are working in the field and what they can expect to accomplish [17]." This had a large influence on research in the area of speech recognition in the US. This example shows that even a highly experienced research manager can sometimes make bad judgments, and it highlights the difficulties of basic research administration. In fact, at that time, the speech research at Bell Laboratories had been assigned to visiting researchers from overseas, including Dr. Fumitada Itakura, whose work at the time led to results in the field of line spectrum pairs (LSP). In 2014, LSP was confirmed as a milestone by IEEE (Institute of Electrical and Electronics Engineers), and you can read more about this pioneering research in the article "LSP (Line Spectrum Pair): Essential Technology for High-compression Speech Coding" [18].

In the 21st century, the information environment that surrounds us in our daily lives is changing rapidly, and speed has become an essential requirement for entry into the market, corresponding to Phases 2 and 3 in Fig. 1 [19]. Even in basic research, the choice of issues to study changes over time, and researchers need to contribute to the commercialization of their results with a sense of speed that matches the speed of the age. While NTT aims to create new markets through co-innovation with other industries, the value of basic research and the expectations of the fruit of this research are likely to grow in the future. Every one of the results from basic research is a valuable seed of innovation and is just waiting for its age to arrive. It is therefore necessary to always be on the lookout for hidden technologies with the potential to lead to new innovation, wherever it may be, and to put them to use as fast as we can. They will serve us well in the competition in R&D and in the creation of new services.

References

[1] Y. Kubo, A. Ogawa, T. Hori, and A. Nakamura, "Speech Recognition Based on Unified Model of Acoustic and Language Aspects of Speech," NTT Technical Review, Vol. 11, No. 12, 2013. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201312fa4.html

[2] T. Amemiya, H. Ando, and H. Ho, "Nonverbal Communication via a Five-senses Interface." NTT Technical Journal, Vol. 19, No. 6, pp. 35–37, 2007 (in Japanese).

[3] T. Amemiya, S. Takamuku, S. Ito, and H. Gomi, "Buru-Navi3 Gives You a Feeling of Being Pulled," NTT Technical Review, Vol. 12, No. 11, 2014. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa4.html

[4] K. Kashino, R. Mukai, K. Otsuka, H. Nagano, T. Izumitani, A. Kimura, T. Kurozumi, and J. Yamato, "Fast Media Search," NTT Technical Journal, Vol .19, No. 6, pp. 29–32, 2007 (in Japanese).

[5] M. Murata, H. Nagano, R. Mukai, K. Hiramatsu, and K. Kashino, "Instance Search Technology for Finding Specific Objects in Movies," NTT Technical Review, Vol. 12, No. 11, 2014.

https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa2.html
[6] K. Kinoshita, “Enhancing Speech Quality and Music Experience with Reverberation Control Technology,” NTT Technical Review, Vol. 12, No. 11, 2014. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa3.html
[7] E. Maeda, H. Isozaki, Y. Sasaki, H. Kazawa, T. Hirao, and J. Suzuki, “Question Answering System: SAIQA—A “Learned Computer” that answers any questions,” NTT R&D, Vol. 52, No. 2, pp. 122–133, 2003 (in Japanese).
[8] R. Higashinaka, K. Sadamitsu, K. Saito, and N. Kobayashi, “Question Answering Technology for Pinpointing Answers to a Wide Range of Questions,” NTT Technical Review, Vol. 11, No. 7, 2013. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201307fa4.html
[9] M. Nagata, K. Sudoh, J. Suzuki, Y. Akiba, T. Hirao, and H. Tsukada, “Recent Innovations in NTT’s Statistical Machine Translation,” NTT Technical Review, Vol. 11, No. 12, 2013. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201312fa2.html
[10] I. Motoyoshi, S. Nishida, L. Sharan, and E. Adelson, “Image Statistics and the Perception of Surface Qualities,” Nature, Vol. 447, No. 7141, pp. 206–209, 2007.
[11] J. Watanabe, “Communication Research Focused on Tactile Quality and Reality,” NTT Technical Review, Vol. 9, No. 11, 2011. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201111fa6.html
[12] T. Kawabe, M. Sawayama, K. Maruya, and S. Nishida, “Recognizing Liquid from Image Motion and Image Deformation,” NTT Technical Review, Vol. 12, No. 11, 2014. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa5.html
[13] M. Kashino, M. Yoneya, H.-I. Liao, and S. Furukawa, “Reading the Implicit Mind from the Body” NTT Technical Review, Vol. 12, No. 11, 2014. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa6.html
[14] S. Tani, “Quantum Computing Beyond Integer Factorization—Exploring the Potential of Quantum Search,” NTT Technical Review, Vol. 12, No. 11, 2014. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa7.html
[15] G. P. Nava, Y. Shiraki, H. D. Nguyen, Y. Kamamoto, T. G. Sato, N. Harada, and T. Moriya, “Capturing Sound by Light: Towards Massive Channel Audio Sensing via LEDs and Video Cameras,” NTT Technical Review, Vol. 12, No. 11, 2014. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa8.html
[16] J. Gertner, “The Idea Factory: Bell Labs and the Great Age of American Innovation,” Penguin Press, March 2012.
[17] J. R. Pierce, “Whither Speech Recognition?,” The Journal of the Acoustical Society of America, Vol. 46, No. 4B, pp. 1049–1051, 1969.
[18] T. Moriya, “LSP (Line Spectrum Pair): Essential Technology for High-compression Speech Coding,” NTT Technical Review, Vol. 12, No. 11, 2014. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411in1.html
[19] F. Vogelstein, “Dogfight: How Apple and Google Went to War and Started a Revolution,” Sarah Crichton Books, November 2013.

Eisaku Maeda
Director, NTT Communication Science Laboratories.
He received the B.E. and M.E. in biological science and the Ph.D. in mathematical engineering from the University of Tokyo in 1984, 1986, and 1993, respectively. He joined NTT in 1986. He was a guest researcher at the University of Cambridge, UK, in 1996–1997. His research interests are statistical machine learning, intelligence integration, and bioinformatics. He is a senior member of IEEE and a fellow of IEICE (Institute of Electronics, Information and Communication Engineers of Japan), and a member of IPSJ (Information Processing Society of Japan).

Instance Search Technology for Finding Specific Objects in Movies

Masaya Murata, Hidehisa Nagano, Ryo Mukai, Kaoru Hiramatsu, and Kunio Kashino

Abstract

Successful retrieval of multimedia such as images or videos often involves utilizing their textual metadata. However, adding such metadata information to all multimedia of interest is far beyond our capability. We are therefore pursuing an effective and efficient content-based search methodology based solely on the multimedia content itself. Our research group has been actively advancing multimedia identification technologies for more than 15 years, and our latest search method makes it possible to search for and find specific objects in many kinds of movies. A specific object, person, or place, is called an *instance*, and the retrieval of an object is called an *instance search*. In this article, we overview our instance search methodology.

Keywords: video retrieval, instance search, robust media search

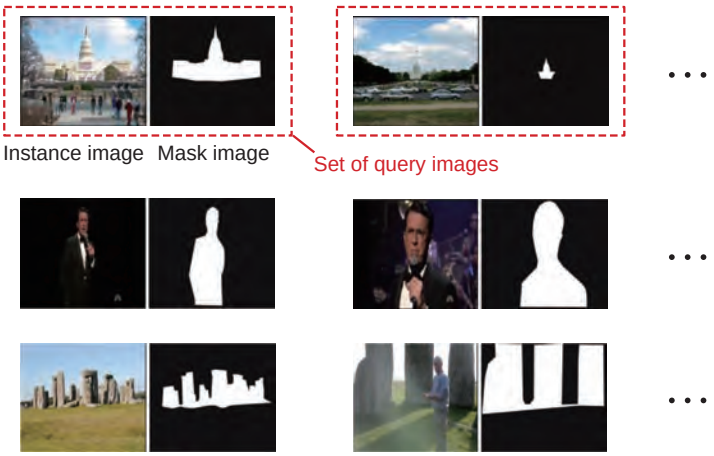
1. Introduction

1.1 Instance search task

Our research group has been conducting research on multimedia search technology for music, images, and video since the 1990s. Initially, we focused on identifying the same multimedia content, that is, on finding the same signal as the input signal. We believed that such technology would contribute to establishing the fundamentals for future media processing technology, in the same way that term (word) identification techniques have done in natural language processing. Our identification technology, called robust media search (RMS), enables the exact specification of multimedia content with high speed and high accuracy.

The research objective of our RMS team was to find the same signal as the input signal. Our BAM (binary area matching) and DAL (divide and locate) methods made this possible even if signals were corrupted with significant noise and distortion. This technical feature is known as robustness. With this innovative high-speed processing technique, RMS is now driving many services such as music search on mobile phones, music rights management for broadcasting and Internet distribution, World Wide Web (WWW) content identification, and content audience surveys. The upcoming instance search technology has enormous potential to advance RMS as a way to solve challenging real-world problems.

After completing the RMS product development, we focused next on not only finding the same signal, but also finding the same object in multimedia content. For example, the aforementioned instance search technology is aimed at searching for the same object such as a person, logo, or place, as the input instance in movies that have a different background and appearance (**Fig. 1**). It is expected that such search technology will become a powerful tool for organizing large-scale image and video data and for retrieving various kinds of information using real-world image queries. It is further expected that the instance search will realize automatic annotation of a specific person to any scene in video archives and that the information retrieval will be achieved by simply taking photos of unknown objects in, for example, a town or city. The potential of instance search is not limited to these kinds of applications. For example, it may contribute to the establishment of a multimedia dictionary. Such a dictionary would enable more



U.S. Capitol building (object instance, top panels), Stephen Colbert (person instance, center panels), Stonehenge (place instance, bottom panels). These examples are actual query images evaluated at an international workshop called TRECVID (see body of article for more details). At TRECVID, mask images showing the instance regions within the query images are also provided; the white regions indicate where the instance appears within the query images.

Fig. 1. Example sets of query images.

detailed semantic analyses of multimedia content.

1.2 Technical aspects of instance search

To find a specific object with a different appearance on a different background, we use local feature data extracted from a query image and a video keyframe. RMS takes a similar approach such as using local features of signals, but it is based on the fact that between the same signals, similar local features tend to occur at the same position in the spatiotemporal space. However, in cases where an object's appearance and background might differ, such consistency is not generally expected, which makes the instance search task very challenging. Our method tackles this problem by precisely matching local features extracted from a query image and a video keyframe in order to differentiate the same object from similar ones. We also consider the discriminative power of local features and assign larger weights to features that contribute to successful instance searching. In short, accurate matching of a highly discriminative local feature is the key to carrying out the instance search task.

2. Keypoint based search methodology

Our proposed instance search methodology is illustrated in **Fig. 2** [1]. We first detect characteristic keypoints from query images and video keyframes. These keypoints are respectively called query keypoints and video keypoints. We then describe each keypoint using a high-dimensional vector. We currently employ the Harris-Laplace detector for the keypoint extraction, and we use the scale-invariant feature transform (SIFT) and color SIFT for the feature vector description. Then we find the video keypoint closest to each query keypoint with the criterion that the cosine similarity value between the two vectors is greater than 0.9. The keypoint matching results enable us to count how many times a query keypoint occurs in a video. This number is called a *keypoint frequency*, and by taking all of the frequencies into account, we can determine whether the query instance appears in the video.

However, as mentioned in the previous section, there are two kinds of query keypoints: those with high discriminative power and those with low discriminative power, and we also consider these properties for ranking videos stored in the database.

Such kinds of discriminative power were originally proposed as inverse document frequency (IDF) in the text retrieval field, and the state-of-the-art ranking method called BM25 theoretically validates the use of the IDF weight for ranking documents of interest. The IDF weight quantifies the discriminative power of a keyword term in a document; the value is smaller when the term tends to occur in many kinds of documents. However, for the video retrieval task, we

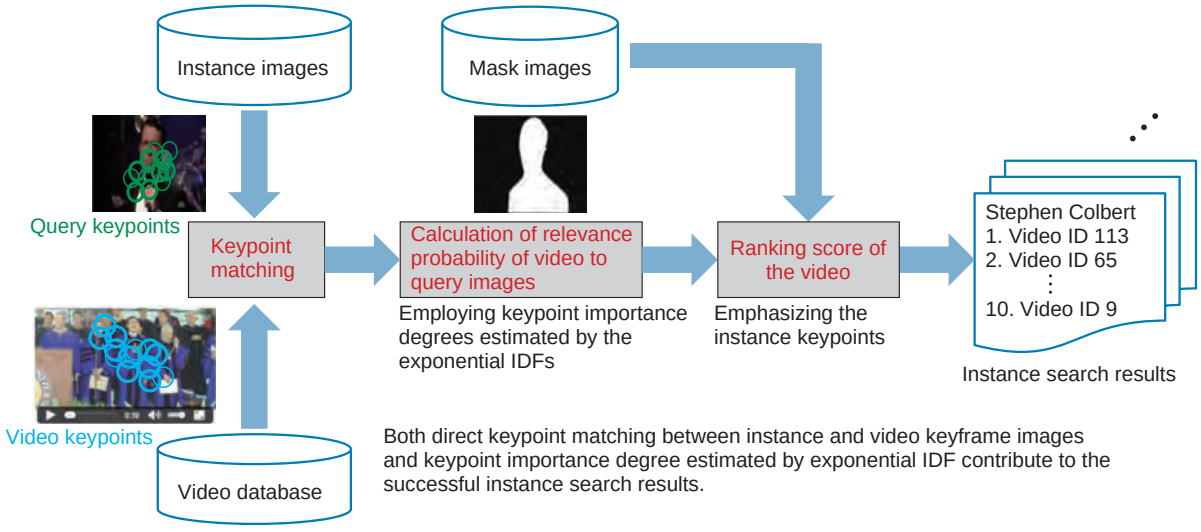


Fig. 2. Overview of instance search method.

found that the conventional IDF could not sufficiently lower the weights of lowly discriminative keypoints. The main reason for this is the inconsistency between the assumptions behind the IDF formulation and the actual properties/features of the images used for instance search.

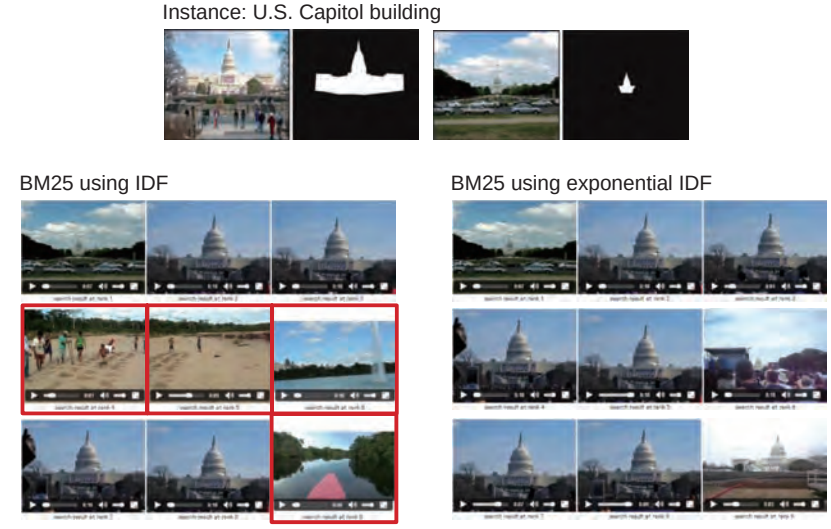
Therefore, we proposed a new keypoint weight that is suitable for the instance search task [2]. We call it exponential IDF, and it is designed to become sufficiently small when the keypoint shows a tendency to be a lowly discriminative feature. The BM25 using exponential IDF realizes accurate matching of highly discriminative local features and results in a high accuracy instance search. More accurate retrieval can generally be achieved by emphasizing keypoints extracted from the specified instance region in the mask image (as in Fig. 1), if the mask image is available. The videos in the database are ranked in decreasing order according to the BM25 scores with exponential IDF and shown to the search user as the instance search result.

The effect of exponential IDF is shown in **Fig. 3**. The uppermost panels correspond to the query and mask images, and the lower left and right search results are those obtained by using the standard IDF and exponential IDF. As seen in the figure, the standard IDF resulted in misdetection caused by the sky, sand, and trees in the query images, but the proposed method successfully suppressed such contributions from the lowly discriminative keypoints.

3. System and evaluation

An example of how our instance search system works is shown in **Fig. 4**. The user (top left) takes a photo of an object for a search using a tablet device. Here, the instance is the NTT logo on a white helmet. The user also specifies the region of interest within the query image (top right). Touching the search button initiates the instance search, and the video ranking results are provided on the right side of the tablet device (bottom left). A video showing the NTT logo can be previewed by tapping the thumbnail image on the video result rankings. If the video contains textual metadata information, the user can be directed to an outside information source such as the WWW and achieve the information retrieval by simply taking a photo.

The instance search task is attracting attention from video retrieval researchers, and it is indeed one of the important tasks at the TRECVID (TREC (Text Retrieval Conference) Video Retrieval Evaluation) workshop organized by the U.S. National Institute of Standards and Technology. Instance examples shown in Figs. 1 and 3 are the actual queries used in the previous TRECVID instance search tasks. At the TRECVID workshop held in 2013, our approach recorded the highest-level instance search accuracy among the 23 teams participating from all over the world [3]. This success is mainly because of our keypoint weighting technique described in section 2.



IDF often fails to estimate appropriate importance degrees of query keypoints, resulting in high rankings for incorrect videos (outlined in red).

Fig. 3. Comparison of instance search results.

(Top left) Taking a photo of the NTT logo on a white helmet. (Top right) Specifying the instance region within the query image; the video search results are shown on the right side of the tablet screen. (Bottom left) Viewing the top ranked video search result by tapping it. (Bottom right) The NTT logo appears in the video shown on the left part of the screen.

Fig. 4. Developing a demo system.

4. Future directions

We introduced our approach for an instance search task and focused in particular on the newly devised keypoint weighting method. We will continue to investigate the instance search task from various

viewpoints and dedicate ourselves to establishing robust media search technology.

Masaya Murata
Research Scientist, Media Recognition Research Group, Media Information Laboratory, NTT Communication Science Laboratories.
He received his B.S. and M.S. in physics from Waseda University, Tokyo, in 2005 and 2007, respectively. He joined NTT in 2007. In 2008, he received a best paper award at the 9th international conference on Web Information Systems Engineering (WISE2008). Since October 2014, he has also been a Ph.D. student at Graduate School of Information Science and Technology, the University of Tokyo. His area of research includes information retrieval, state filtering, and multimedia search applications.

Hidehisa Nagano
Senior Research Scientist, Media Recognition Research Group, Media Information Laboratory, NTT Communication Science Laboratories.
He received his B.Eng. and M.Eng. in information and computer sciences in 1994 and 1996, respectively, and his Ph.D. in information science and technology in 2005, all from Osaka University, Japan. He joined NTT in 1996. From 2011 to 2012, he was a visiting researcher at the Centre for Digital Music, Queen Mary University of London, UK. He has been working on audio and video analysis, search, retrieval, and recognition algorithms and their implementation. He is a senior member of the Institute of Electrical and Electronics Engineers (IEEE) and a member of the Institute of Electronics, Information and Communication Engineers (IEICE) and the Information Processing Society of Japan (IPSJ).

Ryo Mukai
Senior Research Scientist, Media Recognition Research Group, Media Information Laboratory, NTT Communication Science Laboratories.
He received his B.S. and M.S. in information science from the University of Tokyo in 1990 and 1992, respectively. He joined NTT in 1992. From 1992 to 2000, he was engaged in the research and development of processor architecture for network service systems and distributed network systems. Since 2000, he has been with NTT Communication Science Laboratories, where he is conducting research on media search technology. He is a senior member of IEEE and a member of the Association for Computing Machinery, the Acoustical Society of Japan, IEICE, and IPSJ.

Kaoru Hiramatsu
Senior Research Scientist, Supervisor, Media Recognition Research Group, Media Information Laboratory, NTT Communication Science Laboratories.
He received his B.S. in electrical engineering and his M.S. in computer science from Keio University, Tokyo, in 1994 and 1996, respectively, and his Ph.D. in informatics from Kyoto University in 2002. Since joining NTT Communication Science Laboratories in 1996, he has been working on the Semantic Web, sensor networks, and media search technology. From April 2003 to March 2004, he was a visiting research scientist at the Maryland Information and Network Dynamics Laboratory, University of Maryland, USA. He is a member of IPSJ and the Japanese Society of Artificial Intelligence.

Kunio Kashino
Distinguished Researcher, Group Leader of Media Recognition Research Group, Media Information Laboratory, NTT Communication Science Laboratories.
He received his Ph.D. from the University of Tokyo in 1995. Since joining NTT in 1995, he has been working on audio and video analysis, synthesis, search, and recognition algorithms and their implementation. He has received several awards including the Maejima Award in 2010, the Young Scientists' Prize for Commendation for Science and Technology from the Minister of Education, Culture, Sports, Science and Technology in 2007, and the IEEE Transactions on Multimedia Paper Award in 2004. He was a visiting scholar at the University of Cambridge, UK, in 2002 and has been a visiting professor at the National Institute of Informatics, Tokyo, since 2008. He is a senior member of IEEE.

Enhancing Speech Quality and Music Experience with Reverberation Control Technology

Keisuke Kinoshita

Abstract

Speech signals recorded with a distant microphone inevitably contain reverberation, which degrades speech intelligibility and automatic speech recognition performance. Even though reverberation is considered harmful for speech-oriented communication, it is essential for the music listening experience. For instance, when an orchestra plays music in a concert hall, the generated sound is enriched by the hall's reverberation, and it reaches the audience as an enhanced and more beautiful sound. In this article, we introduce a novel reverberation control technology that can be used to enhance speech and music experiences. We also describe how it is utilized in actual markets and how we expect it to open up new vistas for audio signal processing.

Keywords: speech, music, reverberation

1. Introduction

Speech signals captured in a room generally contain reverberation, which degrades speech intelligibility and automatic speech recognition (ASR) performance. Reverberation is considered harmful for speech-oriented communication, but it is essential for the music listening experience. For the purpose of recovering intelligible speech, improving ASR performance, and enriching the music listening experience, considerable research has been undertaken to develop a *blind*[*1] signal processing technology that separates and controls reverberation contained in such audio signals (**Fig. 1**).

2. Innovative audio signal processing technology: Reverberation control technology

2.1. History of research

Research on reverberation control technology, especially audio dereverberation technology, has a long history going back to the 1980s. However, despite the efforts of numerous researchers, it has remained a challenging problem, in contrast to the success achieved with noise reduction techniques. We tried many different approaches, most of which were just partly successful, before finally developing an efficient reverberation control technology for the first time in history. Our approach is an extension of a mathematical technique called linear prediction [1].

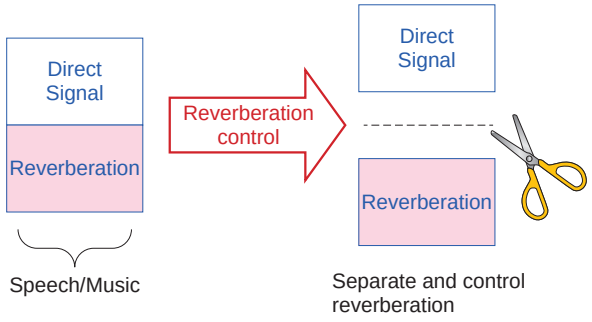


Fig. 1. Effect of reverberation control technology.

*1 The term "blind" refers to the fact that neither the true reverberation characteristics nor source (clean) signals are known in advance.

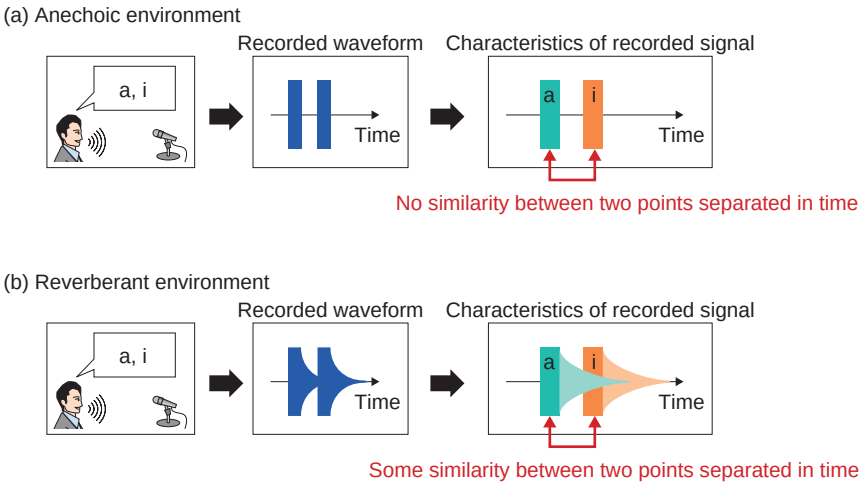


Fig. 2. Difference in temporal characteristics between non-reverberant signal and reverberant signal.

2.2. Key principle of our reverberation control technology

To achieve high-quality reverberation control, it is essential to develop a technique to accurately estimate the amount of reverberation contained in a target audio signal. A key principle of reverberation that we employ in our method is summarized in **Fig. 2**. A situation in which a speaker utters the letters "a" and "i" in an anechoic chamber is illustrated in Fig. 2(a). As can be seen from the figure, in this situation, there is no similarity (more specifically, correlation) between signals that are separated in time. This is because the original audio signal such as that for speech or music generally changes in time. In contrast, a situation in which the same sounds are generated in an echoic chamber is illustrated in Fig. 2(b). Unlike in Fig. 2(a), the signals that are separated in time do possess correlation, since reverberation stretches the preceding sound "a'' in time and lays it over the subsequent sound "i." These physical features of reverberation make it possible to accurately separate audio signals into reverberation and direct signals, assuming that the components that correlate well with the past signals are reverberation, and the components that do not are direct signals [1]. After the separation, we can control the volume of each component separately at our disposal to generate an appropriate output signal for different purposes.

3. Application of reverberation control technology to different market areas

Up until now, the proposed reverberation control technology has been applied to the following market areas and demonstrated its efficacy.

3.1. Recovery of intelligible speech

When making movies and television programs, many scenes have to be filmed and recorded in various environments including noisy and reverberant places. Quite naturally, the recorded signal often ends up containing too much reverberation. Dereverberation technology to deal with such situations has been in demand for decades by professional audio-post-production[*2] engineers. Our technology has been developed into commercial software for postproduction work and is now widely used in production studios all over the world.

Reverberation problems are also relevant in communication services such as teleconferencing systems. Our technology has been extended to real-time processing and will soon be employed in next-generation teleconferencing systems to help people communicate over networks with better speech quality.

Reverberation is known to be harmful not only for communication between people, but also for that between people and robots. Our dereverberation technology has been shown to improve the ASR performance by utilizing it as a preprocessing mechanism for the ASR system. In a speech recognition competition held in May 2014, we achieved the best score

*2 Audio-postproduction is the general term that refers to all stages of production that occur after the actual recording. It involves, for example, sound design, sound editing, audio mixing, and the addition of special effects.

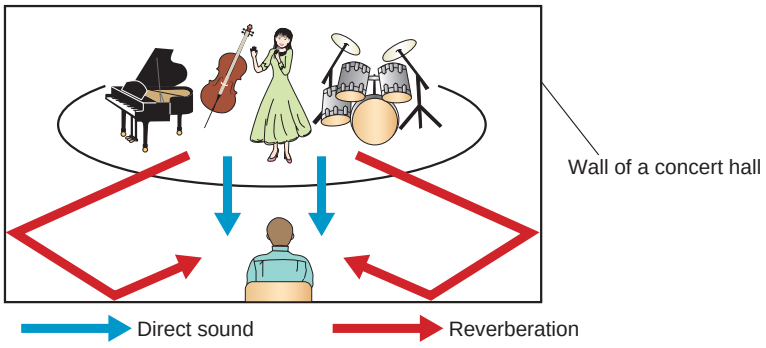


Fig. 3. Sound reaching the audience.

among submissions from 30+ research institutes by employing the dereverberation technique and a state-of-the-art ASR system [2].

3.2. Enrichment of music listening experience

The proposed technique can also be used as a technique to extract reverberation, i.e., three-dimensional (3D) spatial information, from a stereo (i.e., 2-channel (ch)) music signal, and convert the 2ch signal to a surround (e.g., 5.1ch) sound signal [3]. When we are listening to music from a seat in the audience of an actual concert hall, our ears receive two different types of signals from different angles. Considering the positional relationship between the band on the stage and the audience seat, we can expect to receive a direct signal from the front and reverberation from our surroundings, as shown in **Fig. 3**. By receiving these signals in such a manner, we can enjoy a 3D spatial feeling in a concert hall and a feeling of being *enveloped*.

However, once music is stored in the 2ch compact disc (CD) format, direct signals and reverberations are mixed, and consequently, we lose the 3D/front-rear information, which makes it difficult to reproduce an acoustic field similar to the original recording environment (i.e., audience seat in concert hall). We can utilize our technique to reproduce faithful surround sound from the 2ch music signal. Specifically, we can utilize it to separate the 2ch music signals into direct signals and reverberation, and then play them back separately from the appropriate speakers of a 5.1ch surround sound playback system, as depicted in **Fig. 4**. This technique has so far been used for remastering some famous old live albums and has earned a good reputation based on the results. It was also employed as a function of home audio systems that were launched as consumer products.

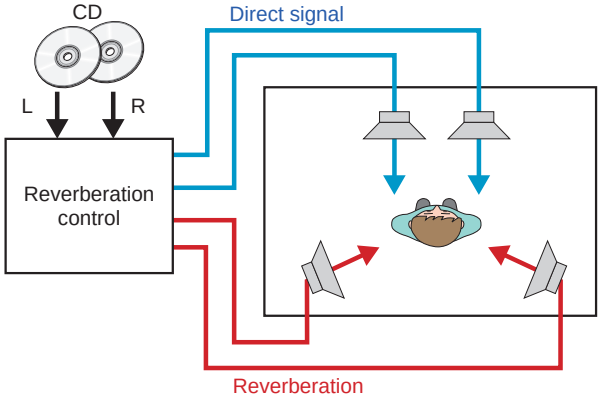


Fig. 4. Surround sound playback based on reverberation control.

4. Summary and future direction

Reverberation is an acoustic effect that becomes more and more remarkable as the distance between a microphone and a sound source increases. Reverberation control technology is an interesting process that can manipulate the perceived distance after the sounds are recorded. If the accuracy of reverberation control is improved in the future, it can open up new vistas of audio signal processing technologies. For instance, we may be able to use a phone or ASR system from a distance without explicitly being aware of the locations of the actual devices/microphones. The music listening experience will be greatly enhanced if we can regenerate an acoustic environment that is completely faithful to a concert hall where the original recordings were made.

References

[1] K. Kinoshita, M. Delcroix, T. Nakatani, and M. Miyoshi, "Suppression of Late Reverberation Effect on Speech Signal Using Long-term Multiple-step Linear Prediction," IEEE Trans. on Audio, Speech and Lang. Proc., Vol. 17, No. 4, pp. 534–545, 2009.

[2] M. Delcroix, T. Yoshioka, A. Ogawa, Y. Kubo, M. Fujimoto, N. Ito, K. Kinoshita, M. Espi, T. Hori, T. Nakatani, and A. Nakamura, "Linear Prediction-based Dereverberation with Advanced Speech Enhancement and Recognition Technologies for the Reverb Challenge," REVERB challenge, 2014.

[3] K. Kinoshita, T. Nakatani, and M. Miyoshi, "Blind Upmix of Stereo Music Signal Using Multi-step Linear Prediction based Reverberation Extraction," Proc. of 2010 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp. 49–52, March 2010, Dallas, TX, USA.

Keisuke Kinoshita

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

He received his M. Eng. and Ph.D. from Sophia University, Tokyo, in 2003 and 2010. Since joining NTT Communication Science Labs in 2003, he has been engaged in research on various types of speech, audio, and music signal processing, including speech enhancement such as 1ch/multi-channel blind dereverberation, noise reduction, distributed microphone array processing, and robust speech recognition. He was a Chief Coordinator of the REVERB challenge 2014. Presently he serves as an Associate Editor of IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences (since 2013). He was honored to receive the 2006 IEICE Paper Award, the 2009 ASJ Outstanding Technical Development Prize, the 2011 ASJ Awaya Prize, and the 2012 Japan Audio Society Award. He is a member of IEEE (Institute of Electrical and Electronics Engineers), ASJ (Acoustical Society of Japan) and IEICE (Institute of Electronics, Information and Communication Engineers).

Buru-Navi3 Gives You a Feeling of Being Pulled

Tomohiro Amemiya, Shinya Takamuku, Sho Ito, and Hiroaki Gomi

Abstract

We have been investigating a way to create a sensation of being pulled even when there is no actual physical pulling. We have found through our studies of the characteristics of human force perception that asymmetric oscillation can create this sensation. Prototype devices to achieve this have been developed, although some early prototypes were as bulky as the receiver of a conventional fixed-line phone and were therefore not suitable for mobile use. We have recently succeeded in creating a thumb-sized prototype device containing a linear actuator that can effectively provide a pulling sensation. We achieved this by focusing on the high tactile acuity of human fingers. This article introduces the innovative prototype and its possible applications.

Keywords: sensory illusion, haptics, force sensation

1. Introduction

What if a compass (or compass application in a mobile phone) actually pulled your hand in a certain direction instead of just showing the direction? The concept of *force stimulation* can be likened to the example of a parent leading a child by the hand, and it has the potential to be more intuitive and expressive than *cutaneous stimulation*, which refers to stimulation of nerves via skin contact. Unfortunately, the conventional force display systems are not suitable for mobile devices since the generation of low-frequency force requires a fixed mechanical ground, which mobile devices lack. Some mobile torque feedback devices that do not require grounding have also been proposed, but they produce neither a constant nor translational force; that is, they can generate only short-term rotational force.

We have been researching a way to create a force sensation in mobile devices and have developed a force display device called *Buru-Navi*, which generates a sensation of being pulled or pushed by exploiting the characteristics of human perception [1, 2]. Our objective is not to create a physical tugging force, but to create an illusory *sensation* of being pulled or pushed. We can do this by placing some kind of mass into a box and moving the mass back and forth with different acceleration patterns for the two directions. This asymmetric motion generates a brief and strong force in a desired direction (e.g., forward) and a weaker one over a longer period of time in the reverse direction (e.g., backward). Note, however, that the average magnitudes of the two forces are identical. Since the magnitude of the longer and weaker force is much smaller than the other force, people who hold the box feel as if they are being pulled in the desired direction rather than feeling only a discrete vibrating sensation that is common in conventional mobile devices today.

2. Miniaturization of force device

Over the past several years, we have been refining a method to create a sensory illusion of being pulled and have developed various prototypes to create such a sensation [3, 4]. Our previous prototypes were able to create a clear force illusion, but they were based on mechanical linkages, so they were generally too large and heavy to embed in mobile equipment. Thus, we have selected a linearly vibrating actuator[*1], which

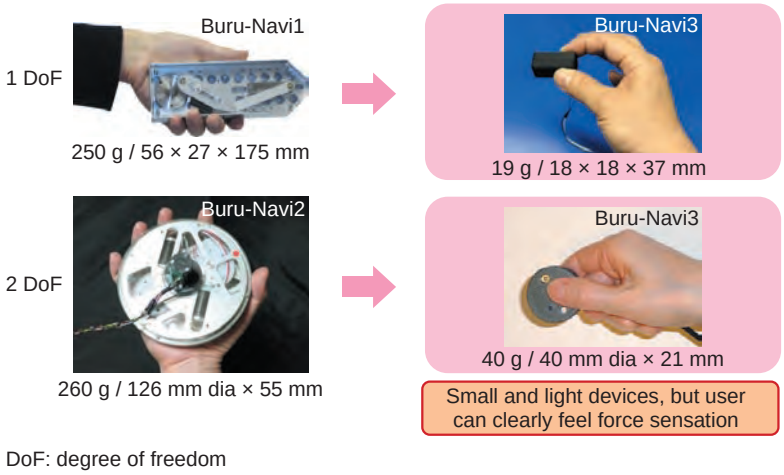


Fig. 1. Comparison of Buru-Navi3 devices.

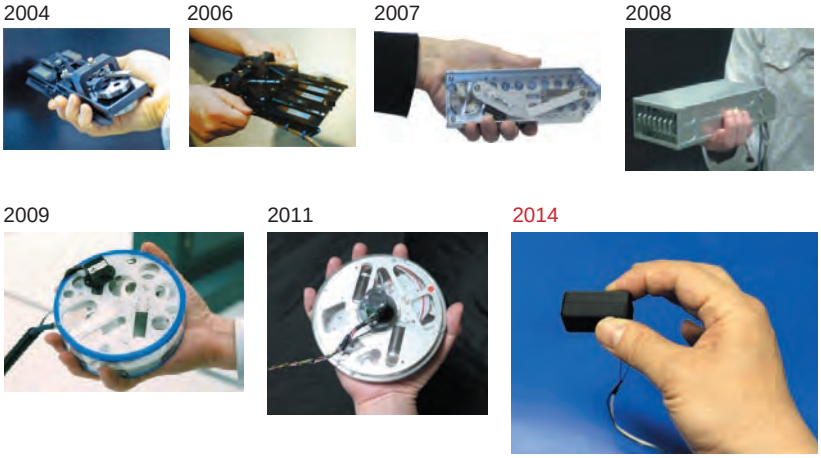


Fig. 2. Timeline of Buru-Navi prototypes.

does not require any mechanism to convert from rotational motion to translational motion, and used it in a thumb-sized prototype (**Figs. 1 and 2**). However, generally speaking, the power-to-weight of the actuator drastically decreases as the size of the actuator decreases, leading to an insufficient force illusion.

Accordingly, we have redesigned and optimized an asymmetric oscillation pattern by working to achieve the best combination of the amplitude of the small actuator and the sensitivity of the human fingerpad. Specifically, we have used a higher frequency range than that used before; the motion patterns of the actuator for forward and backward directions are shown in **Fig. 3**.

There are four tactile mechanoreceptors in the human fingerpad; of these, the Pacinian corpuscle is most sensitive to high temporal frequency vibrations (100–300 Hz) in the normal direction [5], but seems to have nearly the same sensitivity in sliding directions tangential to the skin [6, 7]. In contrast, SA1 and RA1 fibers, whose signals come mainly from Merkel and Meissner corpuscles, respectively, are sensitive to lower frequency (< 100 Hz) vibrations, and some of them can clearly code—or in other words, detect—the sliding or tangential force direction [6]. Thus,

*1 Actuator: A type of motor that moves or controls a mechanism or system

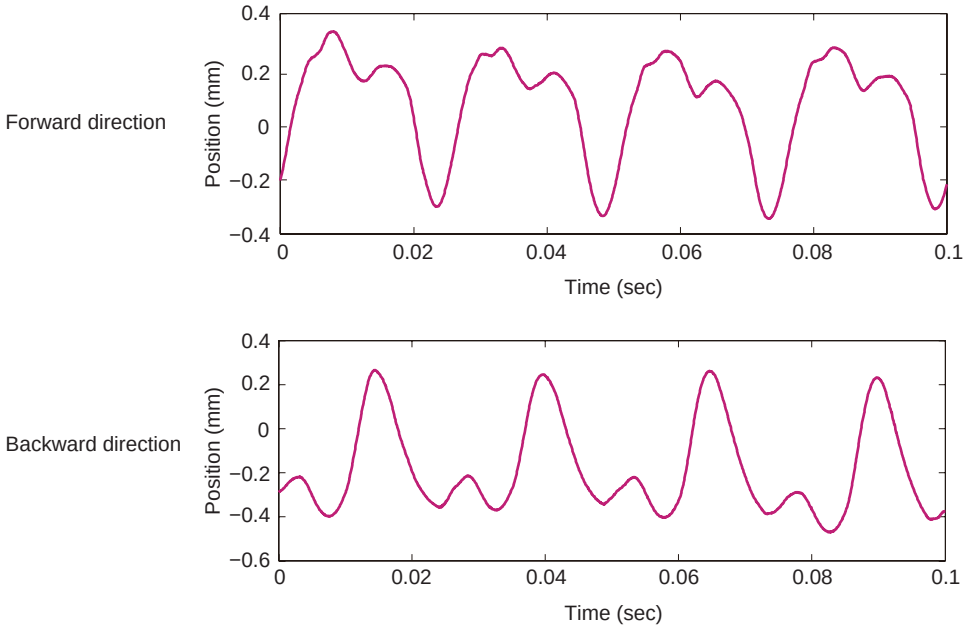


Fig. 3. Motion patterns of asymmetric oscillation.

asymmetrically oscillating stimuli that contain the frequency components that stimulate these receptors will induce a clearer force sensation of being pulled. As a result, we have succeeded in creating a sensation of being pulled with a thumb-sized force display, and the sensation is just as clear and strong as it was with our previous prototypes.

3. Comparison of clarity of force illusion

It is of no value to us if the effect of force illusion decreases when the size and weight of the device are decreased. To determine an effective profile of asymmetric acceleration to induce a clear force sensation, we conducted an experiment to compare the clarity of perceived force sensation between actuators and asymmetric driving patterns of an input signal in order to find the best combination of these for a miniaturized force display.

We selected two kinds of linear actuators; each actuator was covered with a cylinder that was 40 mm in diameter and 17 mm thick and made of ABS (acrylonitrile butadiene styrene) resin. A piece of sandpaper (#1000 grit) was pasted on its surface to control the surface roughness. Two driving patterns of 40 Hz (pulse alternation 7:18 ms) and 125 Hz (pulse alternation 2:6 ms) were used as optimal inputs to each actuator, which gave us four stimulus conditions (two actuators × two driving patterns). In the experiment, each participant was asked which stimulus condition gave a stronger sensation of being pulled/pushed between two of the four conditions. Each participant made six paired comparisons[*2].

All participants reported that they felt a strong force sensation of being pulled/pushed with a specific condition. There were no significant differences between the tendencies of the participants' evaluations ($\chi^2(40) = 25.5$, $p = 0.96$, not significant). Although a paired comparison provides ordinal data, the data can be converted into an interval scale by using the order statistics (Thurstone's method). The values of the clarity of the force sensation in the interval scale are shown in **Fig. 4**. In this scale, the four conditions were ordered along a continuum to represent the degree of difference of the clarity of perceived force. The figure shows that an acceleration pattern created by the combination of a specific actuator (actuator B) and a specific driving pattern was highly likely to be judged to create a clear force sensation. This suggests that a clear force sensation can be created if we select the appropriate asymmetric acceleration pattern of oscillation.

*2 Pairwise comparison: Any process of comparing objects in pairs to judge which of each has a greater amount of some quantitative property

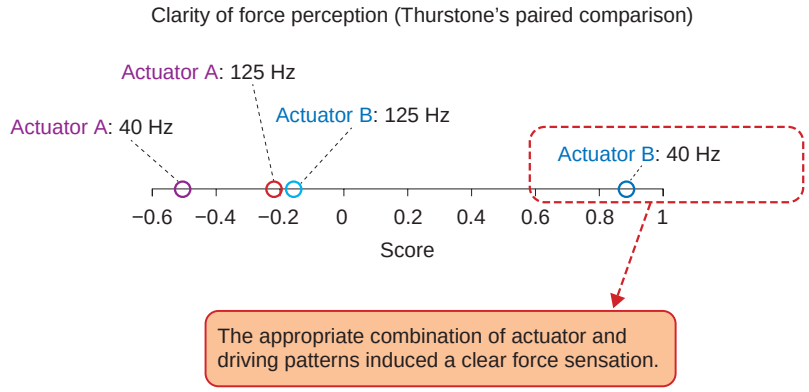


Fig. 4. Results of paired comparison.

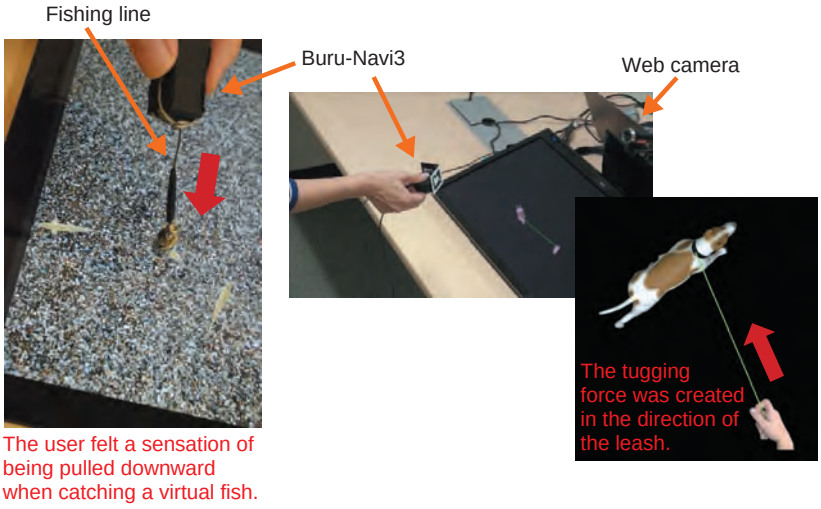


Fig. 5. Applications of Buru-Navi3.

4. Applications

Applications using Buru-Navi3 include a pedestrian navigation system and a new haptic experience in gaming. At NTT Communication Science Laboratories Open House 2014, we presented an angling game application and a virtual dog walking application. In the angling game, users can feel a virtual fish nibbling and pulling the hook. A fishing line with a small weight at its tip was attached to Buru-Navi3 (**Fig. 5(a)**). When the weight approaches the mouth of the virtual fish, the fish bites it. When the fish is caught, Buru-Navi3 generates the sensation of the fish pulling the hook. Many participants seemed to enjoy the novel sensation produced by the device, and some reported to be "addicted" to it. In the dog walking application (**Fig. 5(b)**), the participants felt as if a virtual dog that walks around was pulling on the leash. We used the 2-DoF[*3] version of Buru-Navi3 in combination with a motion tracking system. The amplitude and direction of the force sensation were altered dynamically according to the positions of the user's hand and the virtual dog.

In addition, we recently presented a pedestrian navigation system at the ACM SIGGRAPH[*4] 2014

*3 Degree of freedom (DoF): The number of independent parameters that define the state of a mechanical system

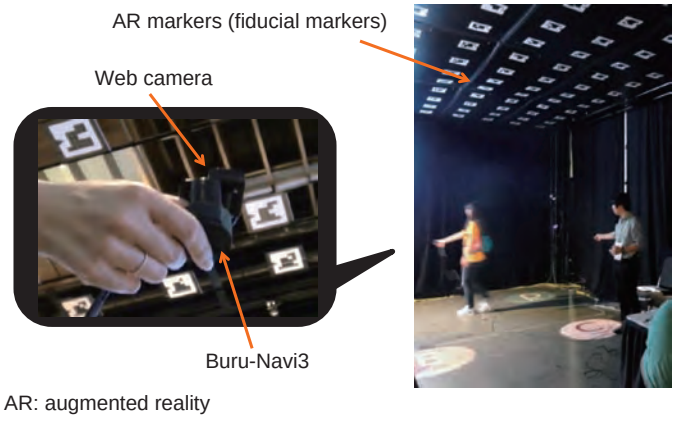


Fig. 6. Pedestrian navigation system using Buru-Navi3.

Emerging Technologies conference [8]. We have implemented a tracking system based on fiducial markers since marker-based tracking provides repeatable, robust, and reliable performance. We hung some polystyrene boards containing original printed fiducial markers on the ceiling. Users held a web camera so it was facing the ceiling and thus able to capture the markers. The user's position and orientation were estimated from each marker's unique ID and its distorted form in the captured image. A photo of the pedestrian navigation system using Buru-Navi3 is shown in **Fig. 6**.

It is worth mentioning that in the former study, we conducted an experiment to examine whether Buru-Navi2, the previous prototype, allowed people with visual impairments to walk safely along a predefined route at their usual walking pace without any previous training. We found that by delivering simple navigational information, Buru-Navi2 enabled more than 90% of participants with visual impairment to walk along a predefined route through a maze without any prior training [4]. This finding implies that the 2-DoF version of Buru-Navi3 will also be capable of leading people along a route—even people who might ordinarily have difficulty finding their way—just as well as Buru-Navi2 did since the fundamentals underlying these systems are essentially identical.

5. Conclusion and future work

Buru-Navi3 is the result of our efforts to investigate the mechanism underlying human perception and motion. In the development process, we have overcome issues such as a trade-off between miniaturization of the device and the effect of illusory sensations. Miniaturized devices are advantageous in that they can be used with mobile and wearable appliances. Such technology will open the door to providing rich haptic experiences in mobile applications, which is a relatively new field.

The characteristics of human perception have been considered deeply in developing conventional video and audio systems. In the future, various kinds of sensory information—not only for perception but also for action—will be considered in order to develop sophisticated interactive systems. Thus, we will continue to focus not only on understanding the mechanisms of sensorimotor processing in the brain, but also on finding prerequisites for developing interactive and natural user-friendly interfaces.

*4 ACM SIGGRAPH: Association for Computing Machinery, Special Interest Group on GRAPHics and Interactive Techniques

References

[1] T. Amemiya, H. Ando, and T. Maeda, "Virtual Force Display: Direction Guidance using Asymmetric Acceleration via Periodic Translational Motion," Proc. of the World Haptics Conference, IEEE Computer Society, pp. 619–622, Pisa, Italy, March 2005.
[2] T. Amemiya, H. Ando, and T. Maeda, "Lead-me Interface for Pulling Sensation in Handheld Devices," ACM Transactions on Applied Perception, Vol. 5, No. 3, pp. 1–17, 2008.
[3] T. Amemiya and T. Maeda, "Asymmetric Oscillation Distorts the Perceived Heaviness of Handheld Objects," IEEE Transactions on Haptics, Vol. 1, No. 1, pp. 9–18, 2008.
[4] T. Amemiya and H. Sugiyama, "Orienting Kinesthetically: A Haptic Handheld Wayfinder for People with Visual Impairments," ACM Transactions on Accessible Computing, Vol. 3, No. 2, Article 6, pp. 1–23, 2010.
[5] S. J. Bolanowski Jr., G. A. Gescheider, R. T. Verrillo, and C. M. Checkosky, "Four Channels Mediate the Mechanical Aspects of

Touch," Journal of the Acoustical Society of America, Vol. 84, No. 5, pp. 1680–1694, 1988.
[6] M. A. Srinivasan, J. M. Whitehouse, and R. H. Lamotte, "Tactile Detection of Slip: Surface Microgeometry and Peripheral Neural Codes," Journal of Neurophysiology, Vol. 63, No. 6, pp. 1323–1332, 1990.
[7] T. Maeno, K. Kobayashi, and N. Yamazaki. "Relationship between the Structure of Human Finger Tissue and the Location of Tactile Receptors," Bulletin of JSME International Journal, Vol. 41, pp. 94–100, 1998.
[8] T. Amemiya and H. Gomi, "Buru-Navi3: Behavioral Navigations Using Illusory Pulled Sensation Created by Thumb-sized Vibrator," Proc. of ACM SIGGRAPH 2014 Emerging Technologies, Article 4, Vancouver, Canada, August 2014.

Tomohiro Amemiya
Research Scientist, Human Information Science Laboratory, NTT Communication Science Laboratories.
He received the B.S. and M.S. in mechano-informatics from the University of Tokyo in 2002 and 2004, respectively, and the Ph.D. in information science and technology from Osaka University in 2008. He joined NTT in 2004 and has been conducting research on human-computer interfaces that apply sensory illusions. He is also an honorary research associate at the Institute of Cognitive Neuroscience, University College London, UK, for the year 2014–2015. He has received several academic awards including the Outstanding Paper Award (Virtual Reality Society Japan, 2005 and 2014), Best Presentation Award (Human Interface Society, 2005 and 2007), Best Paper Award in Human Communications Group (IEICE, 2009), "Grand Prix" du Jury at the Laval Virtual International Awards (2007), Best Paper Award (IEEE VSMM 2012), and Best Demonstration Award (Eurohaptics 2014). He is a member of the Institute of Electrical and Electronics Engineers (IEEE).

Shinya Takamuku
Researcher, Human Information Science Laboratory, NTT Communication Science Laboratories.
He received the B.E., M.E., and Ph.D. in engineering from Osaka University in 2003, 2005, and 2008, respectively. He joined NTT in 2009 and has been engaged in research on the computational basis of human motor awareness. He is a member of the Society for Neuroscience, the Japan Neuroscience Society, and The Robotics Society of Japan.

Sho Ito
Researcher, Human Information Science Laboratory, NTT Communication Science Laboratories.
He received the B.E. and M.E. in mechanical engineering from Kyoto University in 2008 and 2010, respectively. He joined NTT in 2010 and has been researching human sensorimotor mechanisms. He is a member of the Society for Neuroscience, the Japan Neuroscience Society, and The Robotics Society of Japan.

Hiroaki Gomi
Distinguished Senior Research Scientist, Group Leader of Sensory and Motor Research Group, NTT Communication Science Laboratories.
He received the B.E., M.E., and Ph.D. in mechanical engineering from Waseda University, Tokyo, in 1986, 1988, and 1994, respectively. He conducted research on computational biological motor control at ATR (Advanced Telecommunication Research Labs., Kyoto) from 1989 to 1994, in two JST-CREST projects from 1996 to 2003, and in a JST-ERATO project from 2005 to 2010. He was an adjunct professor at Tokyo Institute of Technology from 2000 to 2004. His current research interests include human sensorimotor control mechanisms and interactions between perception and action.

Recognizing Liquid from Image Motion and Image Deformation

Takahiro Kawabe, Masataka Sawayama, Kazushi Maruya, and Shin'ya Nishida

Abstract

Our research confirms that the human visual system recognizes a liquid and its viscosity on the basis of image motion and image deformation. This scientific understanding enables us to alter the impression of materials in movies and prints by manipulating several characteristics of image motion and image deformation.

Keywords: material recognition, liquid, projection mapping

1. Introduction

Humans can easily recognize an object such as a chair, a desk, or a book; we call this *object recognition*. In addition, we are also fairly proficient at recognizing materials such as metal, wool, plastic, and glass (*material recognition*). The visual mechanism of object recognition was clarified in the field of visual science over a decade ago. Meanwhile, the mechanism for material recognition had been virtually ignored [1]. Researchers working in various scientific areas such as psychophysics, computer graphics, and computer vision have recently started investigating how humans (or machines) recognize materials [2].

Still, most of the research that has been done has explored how human observers recognize static materials whose shape does not change over time. Very little research has focused on how human observers are able to recognize nonrigid materials such as rubber that easily deform, or how the visual system can recognize liquids and their viscosity. Here we focused on liquids and sought to clarify the underlying mechanism of liquid recognition. In this article, we first describe our scientific investigations of how humans recognize liquid. Then we describe image presentation techniques that can alter the impression of materials through the manipulation of several characteristics of image motion and image deformation, which strongly affect the recognition of liquid.

2. Recognition of liquid viscosity from image motion speed

We can easily discriminate a low-viscosity liquid such as water from a high-viscosity liquid such as honey. How can the visual system recognize a liquid's viscosity? In everyday life, it is easy for us to recognize the viscosity of liquids in still images. However, it may also be intuitively known that a low-viscosity liquid flows more quickly and smoothly than a high-viscosity liquid. We investigated how image motion generated by a flowing liquid contributed to the recognition of liquid viscosity [3].

We first created computer-rendered movies that simulated various scenes containing the flows of opaque liquids with different viscosities (**Fig. 1**). Next, we calculated the optical flow fields of liquid flows to quantify to what extent each pixel in the simulation movies moved between frames. On the basis of the calculated optical flow fields, we moved low-pass white noise within each of the 15 × 15 matrices of noise patches (noise motion movies). The noise motion movies lacked the static form information contained in the original movies. Thus, by using the noise motion movies, we were able to investigate how image motion information affected the recognition of a liquid's viscosity. We asked observers to

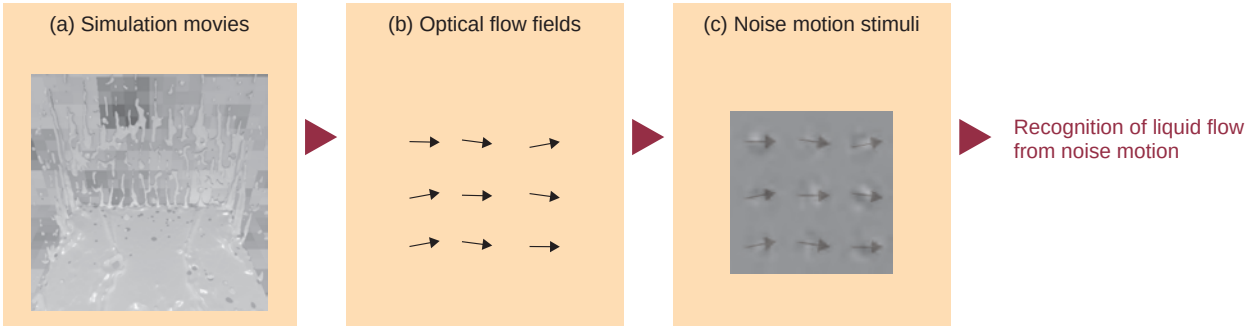


Fig. 1. (a) Computer-rendered movies simulate the flow of an opaque liquid. (b) Optical flow fields of the liquid flow were calculated. (c) Band-pass white noise within each patch in noise motion stimuli was moved based on the optical flow fields. Human observers view the noise motion stimuli and can recognize the flow of liquid from image motion only.

view the noise motion movies and rate the viscosity. Surprisingly, despite the fact that the noise motion movies contained no static form information, the observers were able to judge the viscosity with reasonably high accuracy. The results indicate that human observers can recognize a liquid's viscosity solely from image motion. Moreover, we also found that artificially increasing or decreasing image motion speed, respectively decreased or increased the rated viscosity values. This indicates the possibility that the visual system uses image motion speed to judge liquid viscosity.

3. Recognition of liquidness from smooth image motion flow

However, not all liquids that move slowly are always perceived to be highly viscous. Thus, an additional constraint is clearly necessary in order to recognize viscosity on the basis of image motion speed. In general, a liquid flows smoothly. Thus, we suggested that the additional constraint might be the smoothness of image motion flow. To investigate this idea, we scrambled the positions of the patches in the noise-motion movies to produce a more discrete, or discontinuous, image motion flow (**Fig. 2**), and we found that the spatial scramble of patches strongly reduced the impression of liquidness. Importantly, it became difficult to rate the viscosity under the scenario with a discontinuous image motion flow. We further investigated the image statistics that could characterize the smoothness of the image motion flow (and liquidness impression) and found that the discrete Laplacians (Laplace operators, i.e., the sum of the second derivatives of image motion vectors)

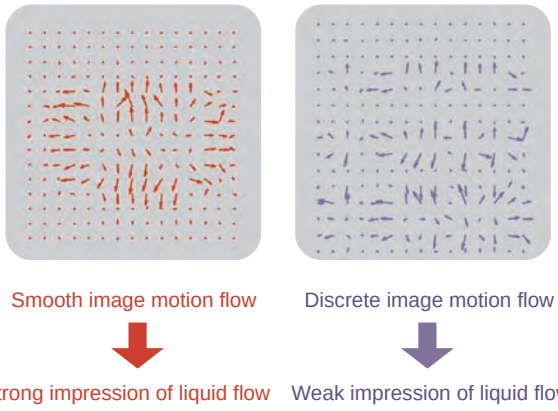


Fig. 2. Smooth and discrete image motion flows create respectively strong and weak impressions of liquid flow.

were negatively correlated with the rated impressions of liquidness. These results indicate that the visual system can use image motion speed to recognize liquid viscosity only when the image motion flow is spatially smooth and, resultantly, when observers interpret the image motion flow as liquid.

4. Recognition of a transparent liquid from image deformation

We have thus far summarized our investigations of how humans see an opaque liquid and its viscosity. However, many liquids such as water, honey, and oil are transparent. When light hits the surface of a transparent liquid, some portions of the light are reflected at the surface of the liquid (**Fig. 3**) and produce

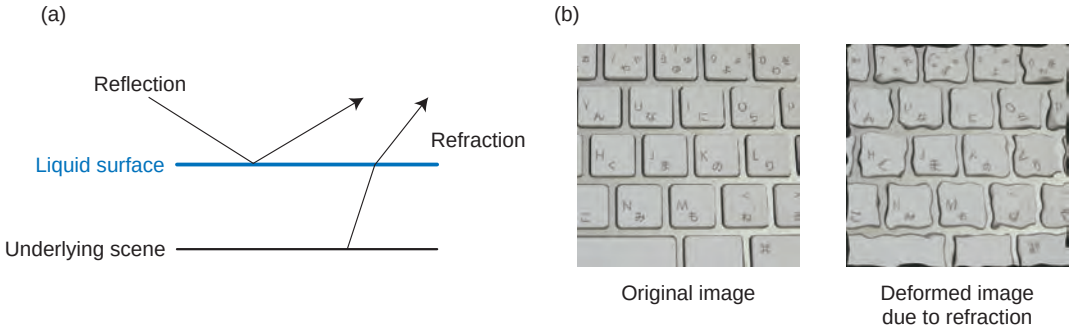


Fig. 3. (a) Light reflection and refraction at the surface of a transparent liquid. (b) Left: original image of an underlying scene. Right: deformation of the left image due to refraction at the surface of a transparent liquid.

highlights in the resulting image. Other portions of the light penetrate the body of the liquid, are reflected at the underlying surface, and are refracted at the liquid's surface. The refraction tends to cause image deformation of the scene underneath the liquid, and the pattern of image deformation depends critically on the three-dimensional structure of the liquid's surface.

We reported in Kawabe, Maruya, & Nishida (2013) that the visual system uses image deformation to see a transparent liquid [4]. We simulated the flow of a transparent liquid that generated the refraction-based dynamic image deformation of an underlying scene. In rendering, we removed reflection components at the surface of a transparent liquid to investigate the pure effect of dynamic image deformation on the recognition of a transparent liquid. We asked observers to view simulation movies and rate their impression of a transparent liquid. The observers were able to recognize a transparent liquid from dynamic image deformation even when the liquid lacked the reflection components. In further studies, we found that human observers often reported the presence of a transparent liquid when dynamic image deformation contained high-amplitude spectra in specific bands of the image deformation's spatiotemporal frequency. These results indicate that human observers recognize a transparent liquid from dynamic image deformation due to refraction.

5. Altering the impression of liquid materials in movies and printed materials

We recently succeeded in extending the obtained scientific knowledge to information presentation techniques that can alter the impression of liquids in movies and prints. For example, our finding that image motion speed can alter the recognition of liquid viscosity led to the idea that manipulating the amplitude spectra of the spatiotemporal frequency can alter the impression of liquid viscosity. Specifically, when we reduced the amplitude spectra of a high temporal frequency in a movie containing a thin liquid, the perceived viscosity of the liquid increased. In another example, we demonstrated that when dynamic image deformation was added to an online movie captured with a camera, the viewers were able to get the impression that a scene captured by the camera was behind the flow of a transparent liquid.

Recently, we developed a new light projection technique that was inspired by our understanding of how people see transparent liquids (**Fig. 4**). When we see a transparent liquid, we simultaneously perceive both an underlying static scene and a transparent layer. Thus, the brain possibly has two sets of representations, one corresponding to the static scene and the other to the dynamic liquid layer. However, what image features are causing these representations? Here, we tentatively hypothesize that the static scene representation is obtained from the direct current (DC) components of the temporal frequency, while the representation of the dynamic liquid layer comes from residual motion components. When we projected residual motion components onto the printed version of the DC components of the temporal frequency, we found that the printed DC components were perceptually deformed. Moreover, when the DC and residual motion components were obtained from the same movies of a liquid flow, the light projection of the residual motion components onto the printed DC components added the impression of a liquid flow to the printed DC components. With this light

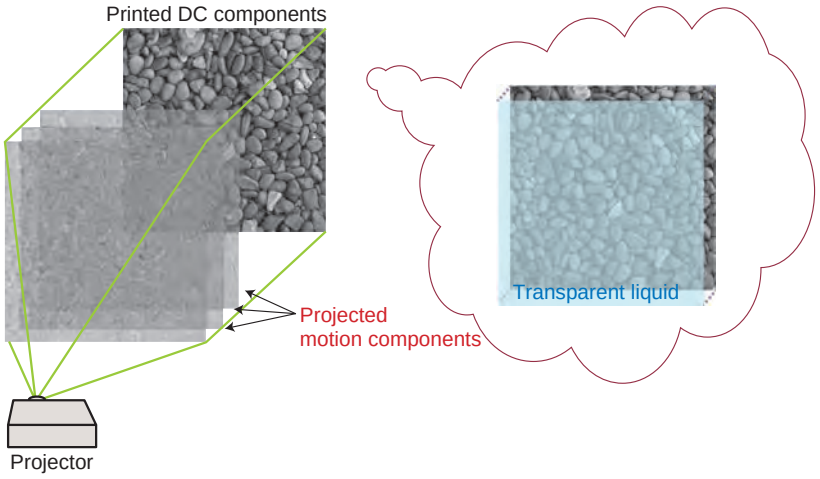


Fig. 4. Schematic illustration of our novel light projection technique. We first extract DC components of temporal frequency and residual motion components separately from a movie of a liquid flow. Then we print out the DC components of the temporal frequency. Finally, we project the residual motion components onto the printed DC components. When observers see this combination of the projected motion components and printed DC components, they see a transparent liquid layer in front of the printed DC components.

projection technique, it is possible to add the impression of a liquid flow to a floor, a wall, or paper.

6. Concluding remarks

We investigated how the brain works when a person sees materials. We plan to accumulate further scientific knowledge about what happens in the brain when we see various materials and explore the development of other novel and useful information presentation techniques that can communicate realistic and vivid material impressions.

References

[1] E. H. Adelson, "On Seeing Stuff: The Perception of Materials by Humans and Machines," Proc. of SPIE 4299, Human Vision and Electronic Imaging VI, pp. 1–12, 2001.
[2] R. W. Fleming, "Visual Perception of Materials and Their Properties," Vision Research, Vol. 94, pp. 62–75, 2014.
[3] T. Kawabe, K. Maruya, R. W. Fleming, and S. Nishida, "Seeing Liquids from Visual Motion," in Vision Research, accepted for publication.
[4] T. Kawabe, K. Maruya, and S. Nishida, "Seeing Transparent Liquids from Dynamic Image Distortion," Abstract presented at the 13th Annual Meeting of the Vision Sciences Society, Journal of Vision, Vol. 13, No. 9, p. 208, 2013.
[5] O. Braddick, "A Short-range Process in Apparent Motion," Vision Research, Vol. 14, No. 7, pp. 519–527, 1974.

Takahiro Kawabe
Research Scientist, Sensory Representation Group, Human Information Science Laboratories, NTT Communication Science Laboratories.

He received the Ph.D. in psychology from Kyushu University, Fukuoka, in 2005. In 2011, he joined NTT Communication Science Laboratories, where he studies human material recognition and cross-modal perception. He received the 2013 JPA (Japanese Psychological Association) Award for International Contributions to Psychology: Award for Distinguished Early and Middle Career Contributions. He is a member of the Vision Sciences Society and the Vision Society of Japan.

Masataka Sawayama
Research Associate, Sensory Representation Group, Human Information Science Laboratories, NTT Communication Science Laboratories.

He received the M.A. and Ph.D. in psychology from Chiba University in 2010 and 2013, respectively. In 2013, he joined NTT Communication Science Laboratories, where he studies human material processing. He received the Best Presentation Award in 2009 from the Japanese Psychonomic Society and in 2011 from the Vision Society of Japan. He is a member of the Vision Sciences Society, the Vision Society of Japan, the Japanese Psychonomic Society, and the JPA.

Kazushi Maruya
Senior Research Scientist, Sensory Representation Group, Human Information Science Laboratories, NTT Communication Science Laboratories.

He received the M.A. and Ph.D. in psychology from the University of Tokyo in 2001 and 2005, respectively. He joined NTT Communication Science Laboratories in 2008. He was a researcher in the Intelligent Modeling Laboratory, the University of Tokyo, from 2004 to 2005, and a visiting scientist at the Department of Psychology, Vanderbilt University, Nashville, TN, USA, from 2006 to 2008. He received best presentation awards at the 2001 and 2004 meetings of the Vision Society of Japan and at the 2003 and 2004 meetings of the Japanese Psychonomic Society. He is currently engaged in research on visual motion processing. He is also interested in human computer interface design for reading digital texts. He is a member of the Vision Sciences Society and the Vision Society of Japan.

Shin'ya Nishida
Senior Distinguished Scientist and Group Leader, Sensory Representation Group, Human Information Science Laboratories, NTT Communication Science Laboratories.

He received the M.A. and Ph.D. in psychology from Kyoto University in 1987 and 1996, respectively. He joined NTT Basic Rsearch Laboratories in 1992 and moved to NTT Communication Science Laboratories in 1999.

He is an expert in psychophysical research on human visual processing, particularly motion perception, cross-attribute/modality integration, time perception, and material perception. He is the current president of the Vision Society of Japan, and an editorial board member of the Journal of Vision and Vision Research.

Reading the Implicit Mind from the Body

Makio Kashino, Makoto Yoneya, Hsin-I Liao, and Shigeto Furukawa

Abstract

Recent studies in cognitive science have repeatedly demonstrated that human behavior, decision making, and emotion depend heavily on the *implicit mind*, that is, automatic, involuntary mental processes that even the person herself/himself is not aware of. We have been developing diverse methods of decoding the implicit mind from involuntary body movements and physiological responses such as pupil dilation, eye movements, heart rate variability, and hormone secretion. If mind-reading technology can be made viable with cameras and wearable sensors, it would offer a wide range of usage possibilities in information and communication technology.

Keywords: man-machine interface, physiological signals, eye movement

1. Introduction

In daily life, people often infer other people's feelings and intentions to some extent—even if they are not expressed explicitly by language or gesture—by taking the appearance of the person and the situation into account. This ability, often referred to as *mind reading*, is an essential characteristic that supports smooth communication among people. A majority of current information and communication technology (ICT) devices, on the other hand, do not function without receiving explicit commands, which are entered using predetermined methods such as typing, pressing buttons, using one's voice, and making specific gestures. If ICT devices had a mind-reading ability, the relationship between such devices and users would be more flexible and natural. Ultimately, users would not be aware of the existence of ICT devices. In other words, mind-reading technology would make ICT devices transparent to users.

2. Reading the mind from the body

Mind-reading technology has been a topic of extensive research in recent years. In fact, remarkable progress has been made in brain-computer interfaces (BCIs), which decode a person's brain activity to identify the content of the person's consciousness, such as a category of a perceived object, or a button to be pressed among multiple alternatives. Our approach, however, is essentially different from BCI in two aspects.

The first difference concerns the method of measurement. In BCI, brain activity is measured using such technologies as electroencephalograph or functional magnetic resonance imaging, whereas in our mind reading approach, we measure physiological changes on the body surface, including eye movements, pupil diameter changes, heart rate variations, and involuntary body movements. These signals can be measured with relatively simple devices such as a camera or surface electrodes. In contrast, BCI requires large-scale specialized measurement equipment. At present, measurement of body surface signals is not completely unconstrained, meaning that the movement of the person being measured is somewhat restricted. However, it will be even less constrained and more transparent in the near future, when sophisticated wearable sensors are developed.

The second difference concerns the decoding target. BCI tries to categorize the content of consciousness, such as types of perceived visual objects (e.g., a

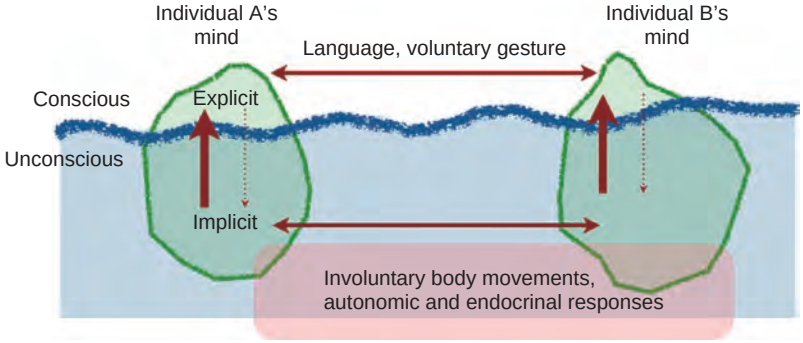


Fig. 1. Illustration depicting implicit and explicit minds.

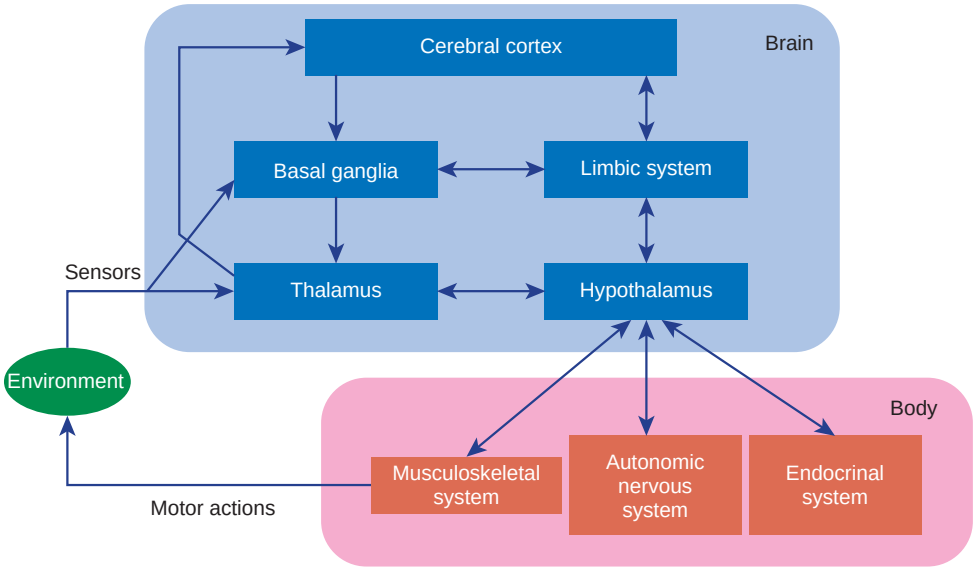


Fig. 2. Loops of the brain, body, and environment.

face vs. a house) and a button to be pressed (e.g., left or right). The system learns the statistical correspondence between the categories and brain activity patterns in advance. Based on the learning, it judges which category the observed pattern belongs to. However, it should be noted that consciousness is only a fraction of the whole *mind*. Recent findings in the field of cognitive science repeatedly demonstrated that human behavior, decision making, and emotion depend not only on conscious deliberation, but they also depend heavily on the implicit mind, that is, automatic, fast, involuntary mental processes that even the person herself/himself is not aware of (**Fig. 1**) [1]. This implicit mind is the target of our mind reading.

Many experimental studies have shown strong interactions between the implicit mind and the body. In a sense, they are inseparable, like two sides of a coin. Such a tight relationship reflects the complex loops of the brain, body, and environment (**Fig. 2**). If some event happens in the environment, the states of one's body such as the autonomic, endocrinal, and musculoskeletal systems change so that the person can react to the event appropriately. The information about the event is also sent to the cerebral cortex, where the event is recognized through complicated information processing. These changes in body states start prior to, and thus deeply affect, the processing in the cerebral cortex. This is why people often *lose control* of their mind and body, no matter how well

they understand what to do. For example, if they had to make a speech in front of some eminent people, they might inadvertently make some uncharacteristic mistakes due to extreme nervousness.

Moreover, in interpersonal communication, unconscious body movements of partners interact with one another, creating a kind of resonance. The resonance, in addition to explicit language and gesture, may provide the basis for understanding and sharing emotions (Fig. 1). An experiment we conducted recently demonstrated that the unconscious synchronization of footsteps between two people who had met for the first time and were walking side by side for several minutes enhanced the positive impressions that they had of each other. The mind is, in one aspect, a dynamic phenomenon that emerges through the interaction mediated by bodies. Thus, measuring the body surface instead of the brain not only has practical merit, but is also essential in mind reading. The next section introduces an example of our experiments.

3. Reading the familiarity and preference for music from the eyes

3.1 Measuring microsaccades

As the saying goes, "The eyes are more eloquent than the mouth." In the context of mind-reading technology, gaze direction has been used extensively as an index of visual attention or interest. However, what is reflected in the eyes is not limited to mental states directed to or evoked by visual objects. We are studying how to decode mental states such as saliency, familiarity, and preference for sounds based on the information obtained from eyes, namely, a kind of eye movement called a microsaccade as well as changes in pupil diameter. In the experiment, the measurement is conducted using a high-precision eye camera (sampling rate = 1000 Hz, spatial resolution < 0.01°), installed in a downward frontal position to the participant, whose head movement is restricted by a chin rest (**Fig. 3**). Although technical problems remain to be solved before we can achieve completely unconstrained measurement in the real world, it has been shown that the eyes can provide more diverse information than previously thought.

Microsaccades are small, rapid, involuntary eye movements that typically occur once every second or two during a visual fixation task (**Fig. 4(a)**). We have revealed a previously unexplored relationship between auditory salience and features of microsaccades by introducing a novel model of eye-position control [2]. In short, the presentation of a salient (i.e., unusual or prominent, and therefore easily noticeable) sound among a series of less salient sounds induced a temporal decrease in the damping factor of microsaccades, which is an indicator of the accuracy in position control, and a temporal increase in the natural frequency of microsaccades, which is an indicator of the speed of position control (**Fig. 4(b)**).

Fig. 3. Experimental setup to decode familiarity and preference for music from the eyes. Eye movements and pupillary responses of participant while she listens to a tune are measured using an eye camera.

3.2 Studying pupillary responses

One more thing we focus on is pupillary responses. The primary function of the pupils is to control the amount of light entering the retina, just as the diaphragm of a camera does. However, pupil diameter is also modulated by emotional arousal and cognitive functions such as attention, memory, preference, and decision making (**Fig. 5(a)**). This is because pupil diameter is controlled by the balance of sympathetic and parasympathetic nervous systems, and reflects, to some extent, the level of neurotransmitters that control cognitive processing in the brain. We have demonstrated that pupil dilation, as well as microsaccades, can be used as a physiological marker for certain aspects of auditory salience [3]. Temporary pupil dilation occurs when a salient sound is presented among a series of less salient sounds (**Fig. 5(b)**). Such pupil dilation responses depend on various factors including acoustic properties, context, and presentation probability, but not critically on voluntary attention to sounds.

We applied these basic findings in some attempts to estimate a listener's familiarity and preference for a tune based on the features of microsaccades and pupillary responses while listening to the tune. In

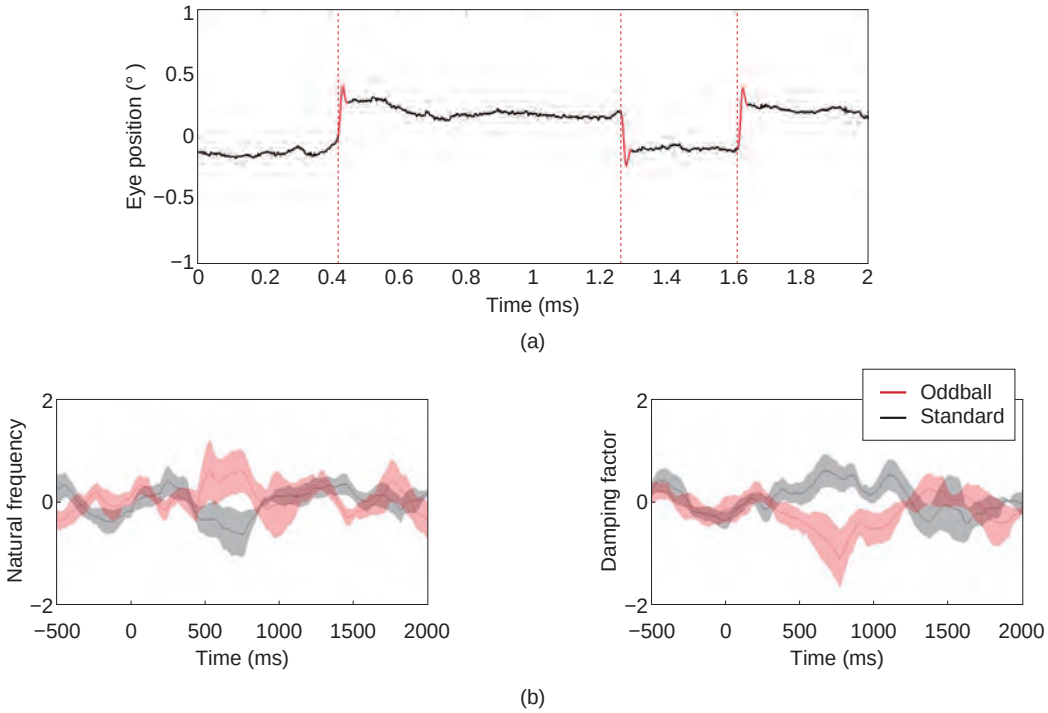


Fig. 4. (a) Example of microsaccade measurement. Microsaccades are indicated by red lines. (b) Time evolution of microsaccade parameters (natural frequency and damping factor) in response to oddball (rare and thus salient) and standard (common and not salient) sounds. The horizontal axis represents time from onset of sound. Each sound was 50 ms in duration.

addition to the physiological signals obtained from the eyes, we also analyzed the acoustic/musical properties of the tune. For example, we have developed a novel *surprise* index, which represents the extent of unpredictability of the musical data at a given moment in a tune, given the data up to that point. We consider this index useful because a typical tune consists of regularity (predictability) and deviation from it (unpredictability or surprise), and the balance between the two seems to be one of the critical factors that contribute to familiarity and attractiveness of the tune.

At NTT Communication Science Laboratories Open House 2014, we conducted a demonstration in which each participant listened to one of 15 tunes from various music genres including classical, rock, and jazz for 90 seconds. The participant's familiarity and preference ratings for the tune were then estimated based on 12 features of microsaccades and pupillary responses, together with several features of the tune including surprise. Prior to the demonstration, the decoding system had learned the mapping between those features and subjective ratings (7-point scale each) of familiarity and preference for 23 participants. In the demonstration, the differences between the actual and estimated ratings were 2 or smaller in more than 80% of the trials for nearly 200 participants. (Note that this was only an informal demonstration and not a rigorous test.) The demonstration was designed to estimate subjective (that is, not implicit) ratings, but could be extended to implicit mind (or behavior), once objective behavioral data are available, such as which tune each participant decided to buy and when and how many times the tune was played back.

4. Future directions

In addition to the work described in this article, we are conducting diverse lines of research concerning the responses of the brain and autonomic nervous systems, hormone secretion, and body movements. In one such project, we developed a method to measure the concentration of oxytocin (a hormone considered to promote trust and attachment to others) in human saliva with the highest accuracy to date (in

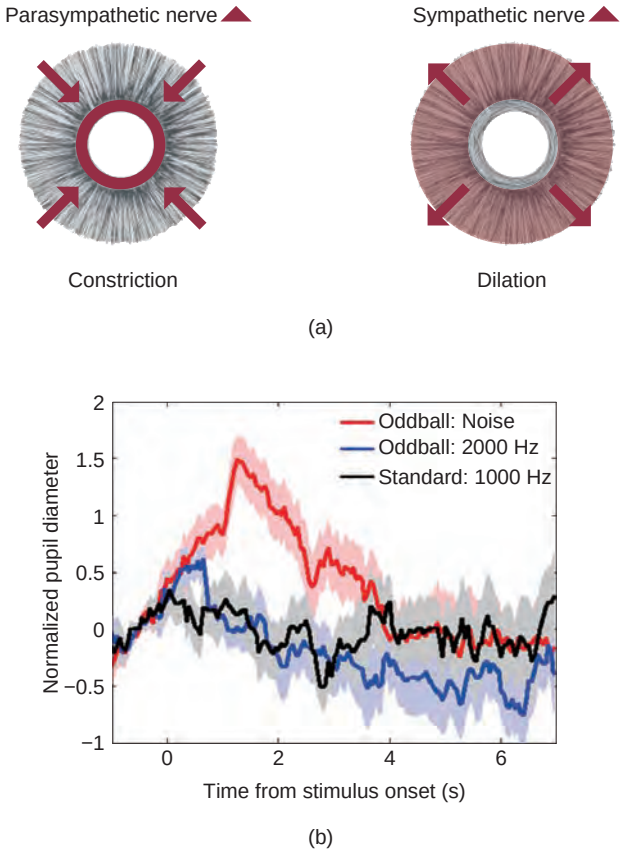


Fig. 5. (a) Relationship between autonomic nervous system activity and change in pupil diameter. (b) Pupil diameter changes in response to oddball and standard sounds. Duration of each sound was 50 ms.

collaboration with Prof. Suguru Kawato of the University of Tokyo). This has enabled us to identify a physiological mechanism underlying relaxation induced by music listening. Listening to music with a slow tempo promotes the secretion of oxytocin, which activates the parasympathetic nervous system, resulting in relaxation [4]. Analysis of hormone concentration is not fast enough for direct use in ICT devices. However, when the relationship between physiological signals that are quickly measurable and the relevant hormone concentration is revealed by laboratory experiments, the knowledge will be beneficial for mind reading. In evaluating the quality of video or audio, for example, it would be possible to capture the differences in quality of experience that are not apparent in subjective ratings.

As wearable sensor technology advances, mind reading will be applied more widely. It is especially attractive in sport-related areas. For example, measurement of heart rate, respiration rate, and muscle potential at various parts of the body using sensors woven into underwear would make it possible to report mental and physical states of players during a game, or to develop effective training methods that combine monitoring and sensory feedback.

The study of reading the implicit mind has just begun. For the moment, basic research is necessary to understand the mechanisms of the complex loops of mind and body. Needless to say, careful consideration should be given to ethical issues such as privacy and safety.

Acknowledgments

Part of this research is supported by CREST from Japan Science and Technology Agency and SCOPE (121803022) from Ministry of Internal Affairs and Communications.

References

[1] D. Kahneman, “Thinking, Fast and Slow,” Farrar, Straus and Giroux, 2011.
[2] M. Yoneya, H.-I. Liao, S. Kidani, S. Furukawa, and M. Kashino, “Sounds in Sequence Modulate Dynamic Characteristics of Microsaccades,” Proc. of the 37th Annual MidWinter Meeting of the Association for Research in Otolaryngology, PS-606, San Diego, CA, USA, February 2014.
[3] H.-I. Liao, M. Yoneya, S. Kidani, M. Kashino, and S. Furukawa, “Human Pupil Dilation Responses to Auditory Stimulations: Effects of Stimulus Property, Context, Probability, and Voluntary Attention,” Proc. of the 37th Annual MidWinter Meeting of the Association for Research in Otolaryngology, PS-599, San Diego, CA, USA, February 2014.
[4] Y. Ooishi, H. Mukai, K. Watanabe, S. Kawato, and M. Kashino, “The Effect of the Tempo of Music on the Secretion of Steroid and Peptide Hormones into Human Saliva,” Proc. of the 35th Annual Meeting of the Japan Neuroscience Society, P2-h01, Nagoya, Japan, September 2012.

Makio Kashino
Senior Distinguished Scientist/Executive Manager of Human Information Science Laboratory, NTT Communication Science Laboratories.

He received the B.A., M.A., and Ph.D. in psychophysics from the University of Tokyo in 1987, 1989, and 2000, respectively. He joined NTT in 1989. From 1992 to 1993, he was a Visiting Scientist at the University of Wisconsin (Prof. Richard Warren's laboratory), USA. Currently, he is a Visiting Professor in the Department of Information Processing, Tokyo Institute of Technology (2006–), and PI (principal investigator) of a JST CREST project on implicit interpersonal information (2009–). He has been investigating functional and neural mechanisms of human cognition, especially auditory perception, cross-modal and sensorimotor interaction, and interpersonal communication through the use of psychophysical experiments, neuroimaging, physiological recordings, and computational modeling.

Makoto Yoneya
Researcher, Sensory Resonance Research Group, Human Information Science Laboratory, NTT Communication Science Laboratories.

He received the B.E. and M.Sc. in engineering from the University of Tokyo in 2010 and 2012, respectively. He joined NTT Communication Science Laboratories in 2012 and has been studying biological signal processing, especially of eye movements. He is also interested in decoding people's thoughts based on brain or neural activity using machine learning methods and has researched decoding of 'internal voice' using magnetoencephalography signals and multi-class SVM (support vector machine). He is also studying auditory signal processing and is currently developing a mathematical model of auditory salience. He received the 2011 Best Presentation Award from the Vision Society of Japan. He is a member of the Acoustical Society of Japan (ASJ), the Japan Neuroscience Society (JNSS), and the Association for Research in Otolaryngology (ARO).

Hsin-I Liao
Research Associate, Sensory Resonance Research Group, Human Information Science Laboratory, NTT Communication Science Laboratories.

She received the B.S. and Ph.D. in psychology from National Taiwan University in 2002 and 2009, respectively. She joined NTT Communication Science Laboratories in 2012 and has been studying auditory salience, music preference, and preference of visual images. She has also explored the use of pupillary response recording to correlate human cognitive functions such as auditory salience and preference decision. During 2007–2008, she was a visiting student at California Institute of Technology, USA, where she studied visual preference using recorded eye movements and visual awareness using transcranial magnetic stimulation. She received a Best Student Poster Prize of the Asia-Pacific Conference on Vision (APCV) in 2008, a Travel Award of the Association for the Scientific Study of Consciousness (ASSC) in 2011, and a Registration Fee Exemption Award of the International Multisensory Research Forum (IMRF) in 2011. She is a member of the Vision Sciences Society (VSS) and JNSS.

Shigeto Furukawa
Senior Research Scientist, Supervisor, Group Leader of Sensory Resonance Research Group, Human Information Science Laboratory, NTT Communication Science Laboratories.

He received the B.E. and M.E. in environmental and sanitary engineering from Kyoto University in 1991 and 1993, respectively, and a Ph.D. in auditory perception from University of Cambridge, UK, in 1996. He conducted postdoctoral studies in the USA between 1996 and 2001. As a postdoctoral associate at Kresge Hearing Research Institute at the University of Michigan, USA, he conducted electrophysiological studies on sound localization, specifically the representation of auditory space in the auditory cortex. He joined NTT Communication Science Laboratories in 2001. Since then, he has been involved in studies on auditory-space representation in the brainstem, assessing basic hearing functions, and the salience of auditory objects or events. In addition, as the group leader of the Sensory Resonance Research Group, he is managing various projects exploring mechanisms that underlie explicit and implicit communication between individuals. He is a member of the Acoustical Society of America, ASJ (member of the Executive Council), ARO, and JNSS.

Quantum Computing Beyond Integer Factorization—Exploring the Potential of Quantum Search

Seiichiro Tani

Abstract

Quantum computers are novel computing devices that make effective use of quantum-mechanical phenomena, and they are expected to emerge within the next few decades. Extensive studies on them have revealed that they will be able to solve many problems significantly faster than current computers or their possible extensions. This article briefly outlines the theory and applications of quantum search to show the potential wide applicability of quantum computing.

Keywords: quantum computers, quantum algorithms, quantum communication

1. Introduction

Significant progress in computer technology has been made since Alan M. Turing, a British mathematician, invented the mathematical model of computers. State-of-the-art computers still follow Turing's model in principle. The term *classical computers* refers to all current computers and their possible extensions, which are based on Turing's model, and *classical computation/algorithms* refers to the computation/algorithms performed on classical computers.

Quantum computers are novel computing devices that make effective use of quantum-mechanical phenomena, so their mechanism is quite different from Turing's model (**Table 1**). It is therefore expected that quantum computers, which are being extensively studied all over the world, will be able to solve a variety of problems significantly faster than the intolerably long time it takes to solve them on classical computers. To solve problems on quantum computers, we need to write a series of explicit operations to be performed on the hardware of quantum computers, as in the case of current computers. Such a series of operations is called a *quantum algorithm*. The computation time on quantum computers heavily depends on the quantum algorithms, just as the computation time of current computers depends on the classical algorithms. It is therefore necessary to research and develop fast quantum algorithms as well as scalable and robust hardware for quantum computers.

2. Fast quantum algorithms developed by Shor and Grover

The most famous quantum algorithm is arguably the one developed by Peter W. Shor in 1994 [1] that factors integers exponentially faster than any known classical algorithm. The integer factoring problem is a basic mathematical problem that is very difficult in the sense that a long history of research on this problem has not yet succeeded in finding a fast classical algorithm. In fact, the security of the RSA (Rivest-Shamir-Adelman) cryptosystem, used in practice on the Internet for secure data transmission, is based on the hardness of this problem. Shor's discovery thus has the potential to affect a great number of people, even those outside academic communities, since it implies that quantum computers would be capable of breaking the cryptosystem. This is obviously an unhappy scenario.

Extensions of Shor's algorithm can solve hidden

Table 1. Comparison between classical and quantum computers.

	Classical computers	Quantum computers
Unit of information	Bit	Quantum bit (qubit)
Mathematical expression of information	Logical values (True/False)	Complex vectors
Elementary operations	Logical operations (AND, OR, NOT)	Linear operators (unitary operators)
Model of computation	Turing machine, logic circuits	Quantum Turing machine, quantum circuit

subgroup problems extremely quickly. In the research field of quantum algorithms, subsequent studies extended Shor's algorithm to a collection of more general mathematical problems. Although these extensions are highly important in a theoretical sense, they seem to have little relation to the problems that are familiar to people outside the theory field, and it would thus be unlikely that those people would appreciate the extended algorithms.

The quantum search algorithm (*quantum search* in short), developed by Lov K. Grover [2] in 1996, is also very well known. This algorithm solves the problem of finding a desired piece of data from among N pieces of data. Classical computers clearly need roughly N accesses to the data in the worst case, even if we allow a small error probability. However, the quantum search can find a desired piece of data with high probability only with approximately $\sqrt{N}$ accesses. Although this cannot achieve the exponential speedups over classical algorithms, in contrast to the case of factoring integers, the quantum search can still yield a significant speedup whenever N is very large.

One of the advantages in considering the search problem is that the definition of the problem is so simple that it can often emerge as a subproblem of various other problems that we want to solve. In other words, the algorithms for the problem potentially have wide applicability. One can imagine the following straightforward scenario; if one finds a search problem as a subproblem of some other problem, then one can solve the search problem significantly faster with the quantum search, which may imply a fast algorithm for the whole problem. In many cases, however, it is not an easy task to carve a search problem out of the original problem; even if one succeeds in doing so, it may be necessary to adapt the algorithm appropriately, often in a non-trivial way. This is why the search problem and its generalizations are still major topics in quantum computing research (**Fig. 1**).

3. Quantum bits and superposition

The unit of information on quantum computers is called a quantum bit, or a *qubit*. A qubit represents any superposition of 0 and 1, including 0 or 1 as a special case (**Fig. 2**). Similarly, two qubits represent any superposition of 00, 01, 10, and 11. Moreover, n qubits represent any superposition of 2^n bit-strings: 0…0 through 1…1. The superposition is called the *quantum state* over the qubits. In terms of linear algebra, we can describe it as follows: Consider a two-dimensional complex Euclidean space with two orthogonal basis vectors, $\vec{0}$ and $\vec{1}$, of unit length, and identify $\vec{0}$ and $\vec{1}$ with "0" and "1," respectively. Then, a superposition of 0 and 1 means simply a linear combination of the unit vectors $\alpha\vec{0} + \beta\vec{1}$, where α,β are complex numbers such that $|\alpha|^2 + |\beta|^2 = 1$. Similarly, a superposition over n qubits is a linear combination of 2^n orthogonal vectors of unit length. Since qubits are identified with vectors, the operations over qubits should be mappings over vectors. More concretely, the operations must be unitary operators[*1] or orthogonal projectors, which are linear transformations over complex vectors. In particular, applying a set of orthogonal projectors summed to the identity is called *measurement*, which produces a classical outcome and a quantum state. To get a classical value as the output of a quantum algorithm, we thus need to take measurements over a quantum state generated by the algorithm.

4. The search problem and quantum algorithms

To solve a problem such as the search problem that depends on N input data $X_1, \cdots, X_N$, we need to access the input data. Formally, we can think of this as follows:

*1 A unitary operator is a linear operator that maps a vector to another such that it does not change the inner product of any two vectors. Intuitively, the operator does not change the angles between any two vectors and the length of any vector.

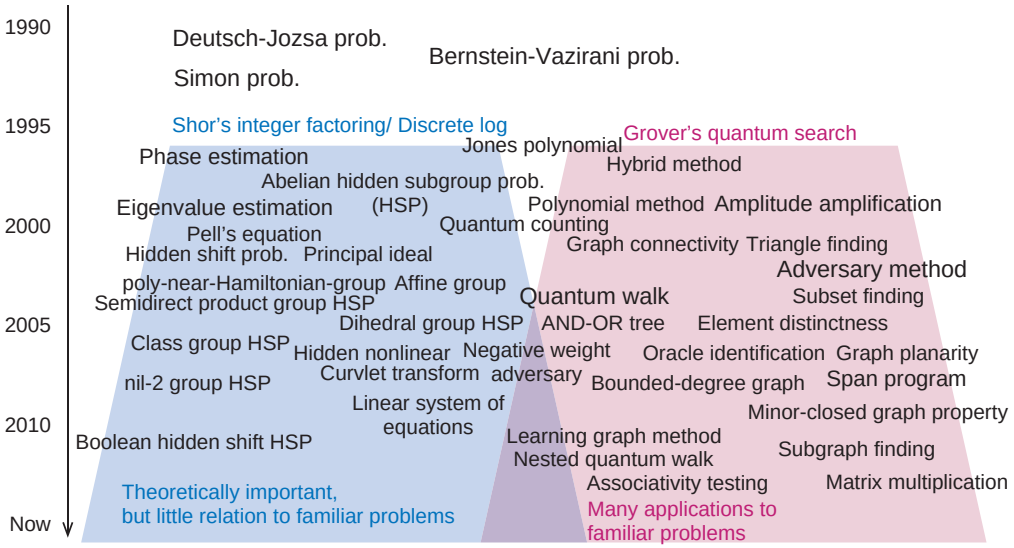


Fig. 1. Quantum algorithms.

	Superpositon	Mathematical expression
one qubit	Superposition of {0,1}	A vector in 2-dimensional complex Euclidean space
two qubits	Superposition of {00,01,10,11}	A vector in 4-dimensional complex Euclidean space
n qubits	Superposition of 2^n strings of {0…00,0…01,…,1…11}	A vector in 2^n-dimensional complex Euclidean space

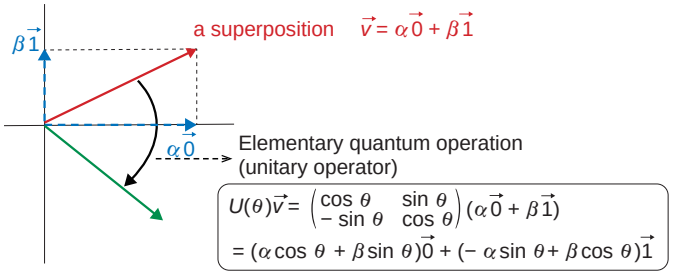


Fig. 2. Qubits and superposition.

We receive X_k by making a query with the index k. We call such a query a *classical query*. On quantum computers, the index associated with a query is expressed with qubits, and thus a query, in this case called a *quantum query*, will be a superposition of classical queries over all indices k ranging from 1 to N; accordingly, the answer to the quantum query will be the corresponding superposition of all X_k, where k ranges from 1 to N. With the ability to make quantum queries, the quantum search may be stated as follows (we assume for simplicity that each piece of input data is either 0 or 1, but this is not essential to the quantum search).

Theorem (Quantum Search) *Given N input data $X_1, \cdots, X_N \in \{0,1\}$, there exists a quantum algorithm that finds an index i with $X_i = 1$ with high probability by making approximately $\sqrt{N}$ data accesses (i.e., quantum queries).*

To estimate the total number of steps required to solve a problem, it is necessary to count the number of steps taken to process the input data obtained via queries, as well as the number of accesses to the input data. However, we will focus only on the number of accesses to the input data, since it is a dominant factor in the search problem and the other problems dealt with in this article.

As an application of the theorem, let us consider the problem of testing the planarity of a given graph.

Example 1: Graph planarity testing

We say that a graph is planar if the graph can be drawn without crossing any pair of edges on a two-

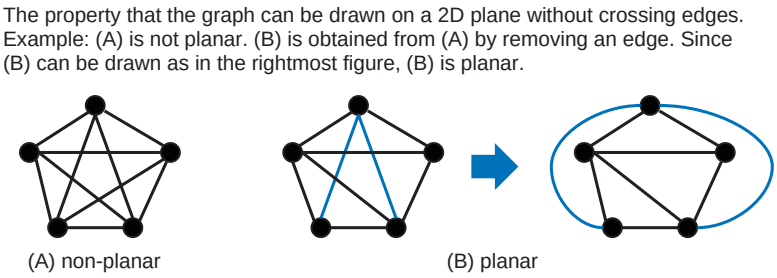


Fig. 3. Planarity of a graph.

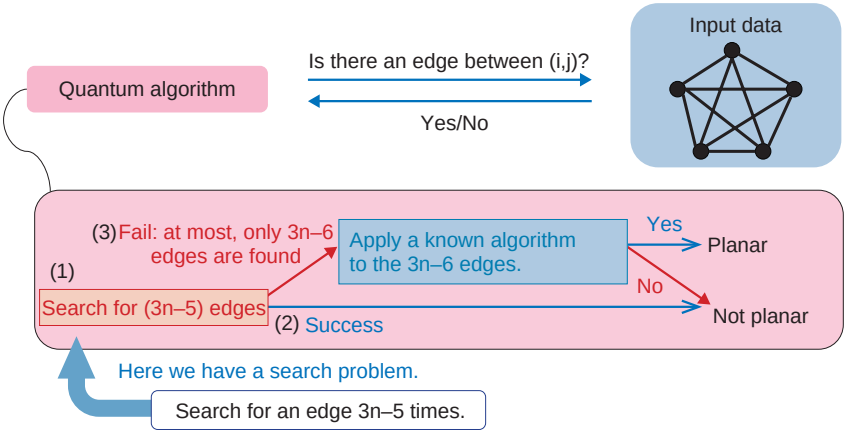


Fig. 4. Quantum algorithm for testing planarity of a graph.

dimensional plane (**Fig. 3**). It is well known that a lot of graph-theoretic problems that are hard for general graphs can be solved efficiently on planar graphs. Testing the planarity of graphs is hence one of the particularly important problems in testing graph properties. Let us define our problem more formally. We assume that graph G is an undirected graph[*2] consisting of n vertices. The input data are given as the information consisting of whether an edge exists between each pair of vertices, more concretely, the set of X_{ij} defined as follows; For each $i = 1, \cdots, n$ and $j = 1, \cdots, n$ with $i < j$, define $X_{ij} = 1$ if there is an edge between (i, j), and $X_{ij} = 0$ otherwise (that is, X_{ij} is the (i, j)-element of the adjacency matrix[*3] of G). The problem is determining with high probability if the graph represented by these $\frac{1}{2}n(n-1)$ pieces of data, X_{ij}, is planar. The key tool for carving out a search problem is a very old theorem discovered by Euler in the 18th century.

Theorem: *If a graph with n vertices is planar, then the graph has at most* $3n - 6$ *edges.*

With this theorem, we divide the problem into the following subproblems (**Fig. 4**).

(1) Identify $3n - 5$ edges from among the $\frac{1}{2}n(n-1)$ candidates of edges (in particular, this means identifying all edges, if the graph has at most $3n - 6$ edges).
(2) If Step (1) identifies exactly $3n - 5$ edges, the graph is not planar according to Euler's theorem, since $3n - 5 > 3n - 6$.
(3) If Step (1) identifies at most $3n - 6$ edges, all edges have been identified; then, without further queries, determine whether the graph is planar with a known algorithm.

Note that only Step (1) makes queries. With a little

*2 An undirected graph is a graph whose edges have no directions.
*3 For a graph with n vertices, the adjacency matrix A of the graph is an n-by-n matrix such that each element A[i,j] represents the existence of an edge between the vertices i and j.

more thought, it is possible to see that Step (1) can be realized by solving $3n - 5$ times the problem of searching for an edge. We can hence speed up Step (1) with the quantum search (plus some modifications). Consequently, we can prove that roughly $n^{1.5}$ quantum queries are sufficient to test the planarity of a graph on quantum computers, while nearly n^2 classical queries are required on classical computers [3]. We should emphasize that the key to carving out a search problem is the effective use of a theorem in the field of graph theory. There are many other important properties for which quantum speedups can be achieved, examples of which include testing graph connectivity[*4] and determining the existence of Hamiltonian paths[*5].

5. A generalization of quantum search

Let us sample a piece of input data at random on a classical computer. In other words, we choose a piece of data from among $X_1, \cdots X_N$ uniformly at random and access it. If there is exactly one desired piece of data among the N input data, the probability that a randomly chosen piece is the desired one is only $1/N$. In fact, very roughly speaking, what the quantum search does is to iterate this random sampling *in superposition* $\sqrt{N}$ times. Note that amplifying the success probability $1/N$ of the random sampling nearly to one requires roughly N times on classical computers. In this sense, we would say that the quantum search amplifies the success probability nearly to one with substantially fewer iterations. This viewpoint can be stated more generally as follows.

Theorem (Quantum Amplitude Amplification [4])
Consider a classical algorithm that solves a problem with probability p by accessing input data c times. Then, it is possible to construct a quantum algorithm that solves the problem with a probability close to one by accessing input data roughly $c/\sqrt{p}$ times. Moreover, if the exact value of p is known, it is possible to amplify the probability to one.

In the case of random sampling for the search problem, the parameters will be $c = 1$, $p = 1/N$. The resulting quantum algorithm thus accesses the input data (roughly) $c/\sqrt{p} = \sqrt{N}$ times. This means that a special case of the theorem is Grover's algorithm. Intuitively, the theorem says that roughly $1/\sqrt{p}$ iterations of the base classical algorithm *in superposition* can amplify the success probability p nearly to one. A similar idea can be applied to even distributed algorithms on a quantum network[*6] consisting of multiple quantum computers.

Example 2: The leader election problem

The leader election problem is a fundamental problem in the distributed computing field in that it often appears as a subproblem of various distributed computing problems. The goal of the problem is to elect a unique leader from among all nodes in a network (**Fig. 5**). The problem is seemingly easy, but in the most general case where each node does not necessarily have a unique identifier, it is mathematically proved that the problem cannot be solved within any bounded time with probability one by using classical computation and communication (i.e., sending/receiving *bits* through communication channels). In contrast, we proved that, if quantum communication and computation can be used, the problem can be solved within a certain bounded time with probability one [5]. This means that classical distributed computing and its quantum counterpart are *qualitatively* different. The first proof of our result did not use the quantum amplitude amplification, but an idea of another proof that employs the quantum amplitude amplification is given as follows.

First consider the following simple classical algorithm: (1) Every node flips a coin; (2) the algorithm succeeds only when there is exactly one node that sees *heads*, in which case, the node will be elected as a leader. We can easily see that the success probability of this algorithm is $p = n/2^n$, where n is the number of nodes. Since we know the exact value of the success probability, we can construct a quantum algorithm that amplifies the success probability to one by repeating the coin flipping roughly $1/\sqrt{p} = \sqrt{2^n/n}$ times together with the quantum amplitude amplification. In fact, a much more sophisticated version of this idea flips coins roughly n times.

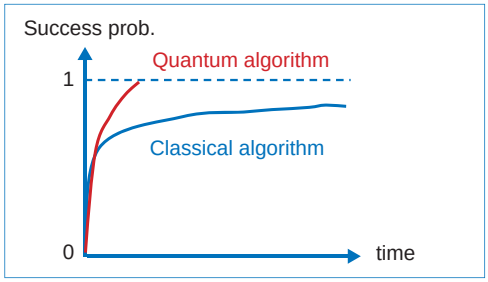
*4 The connectivity of a graph is the property that, for any two vertices of the graph, there exists a path, i.e., a sequence of edges, from one vertex to another.

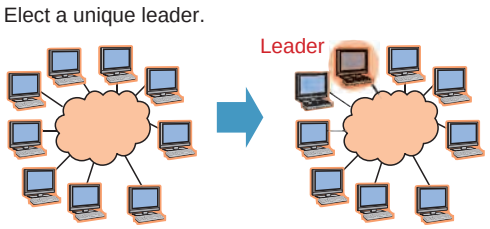
*5 A Hamilton path of a graph is a sequence of edges of the graph, such that the sequence visits every vertex exactly once.

*6 A quantum network consists of multiple quantum computers and quantum communication links between them. Quantum communication is the communication of qubits, which has already been achieved for practical use—for example, QKD (quantum key distribution)—by transmitting photons in a quantum state through optical fibers.

Leader election problem (LE)*
The goal is for all nodes in a network to collaborate in electing a unique leader from among the nodes.

* A fundamental distributed computing problem that emerges as a subproblem of various distributed computing problems.





Classical computation and communication	Cannot solve LE within any bounded time with probability one §.
Quantum computation and communication	Can solve LE within a certain bounded time with probability one.

§ Without the assumption that every party has a unique identifier.

Fig. 5. Quantum algorithm for electing a leader.

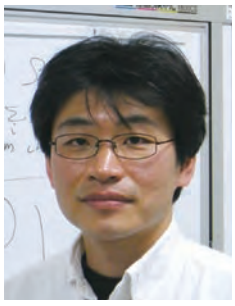
6. Conclusion

The development of quantum computation theory has progressed significantly in the last two decades. Recent research on quantum algorithms shows rich connections with classical computer science; state-of-the-art quantum search algorithms stem from (the quantum version of) random walk[*7] or semidefinite programming[*8], which are major technical tools in classical computer science. These connections will make it possible to make progress in both quantum and classical computer science in a collaborative way. Moreover, it is obviously important to study how to transform the high-level description of a quantum algorithm into a low-level description, i.e., a quantum circuit. A lot of fundamental techniques for this transformation have been intensively studied as well (e.g., [6]). Nevertheless, there are still a lot of fundamental problems to be solved before we can gain a deeper understanding of the potential power of quantum computing. To solve these problems and acquire new knowledge, we need to develop more novel techniques by making the best use of various ideas in related fields such as mathematics, physics, and classical computer science.

References

[1] P. W. Shor, "Algorithms for Quantum Computation: Discrete Logarithms and Factoring," Proc. of 35th Annual IEEE Symposium on Foundations of Computer Science, pp. 124–134, Santa Fe, NM , USA, November 1994.
[2] L. K. Grover, "A Fast Quantum Mechanical Algorithm for Database Search," Proc. of 28th Annual ACM Symposium on Theory of Computing, pp. 212–219, Philadelphia, PA, USA, May 1996.
[3] A. Ambainis, K. Iwama, M. Nakanishi, H. Nishimura, R. Raymond, S. Tani, and S. Yamashita, "Quantum Query Complexity of Boolean Functions with Small On-sets," Proc. of 19th International Symposium on Algorithms and Computation, Vol. 5369 of Lecture Notes in Computer Science, pp. 907–918, Springer, 2008.
[4] G. Brassard, P. Høyer, M. Mosca, and A. Tapp, "Quantum Amplitude Amplification and Estimation," Quantum Computation and Information, Vol. 305 of Contemporary Mathematics, pp. 53–74. American Mathematical Society, 2002.
[5] S. Tani, H. Kobayashi, and K. Matsumoto, "Exact Quantum Algorithms for the Leader Election Problem," ACM Transactions on Computation Theory, Vol. 4, No. 1, Article 1, 2012.
[6] Y. Takahashi and S. Tani, "Collapse of the Hierarchy of Constant-Depth Exact Quantum Circuits," Proc. of 28th IEEE Conference on Computational Complexity, pp. 168–178, Stanford, CA, USA, June 2013.

*7 Random walk on a graph is a stochastic process such that a particle continues to move one vertex to one of its neighbors chosen at random. It is a major technical tool for building efficient probabilistic algorithms.
*8 A semidefinite program (SDP) is a mathematical program that is described with positive-semidefinite matrices. There are efficient general algorithms that approximately solve SDPs.



Seiichiro Tani

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

He received the B.E. in information science from Kyoto University in 1993, and the M.S. and Ph.D. in computer science from the University of Tokyo in 1995 and 2006, respectively. He joined NTT in 1995 and studied algorithms for LSI testing, multi-point communication, and distributed computing problems. He is currently working on algorithmic theory and complexity of quantum computing. From 2004 to 2009, he was also a researcher of the ERATO Quantum Computation and Information Project of Japan Science and Technology Agency (JST). From 2010 to 2011, he was a visiting researcher at the Institute for Quantum Computing, University of Waterloo, Ontario, Canada. He received the 2000 IEICE ISS Best Paper Award. He is a member of IEICE (Institute of Electronics, Information and Communication Engineers), IPSJ (Information Processing Society of Japan), IEEE (Institute of Electrical and Electronics Engineers), and ACM (Association for Computing Machinery).

Capturing Sound by Light: Towards Massive Channel Audio Sensing via LEDs and Video Cameras

Gabriel Pablo Nava, Yoshifumi Shiraki, Hoang Duy Nguyen, Yutaka Kamamoto, Takashi G. Sato, Noboru Harada, and Takehiro Moriya

Abstract

We envision the future of sound sensing as large acoustic sensor networks present in wide spaces providing highly accurate noise cancellation and ultra-realistic environmental sound recording. To achieve this, we developed a real-time system capable of recording the audio signals of large microphone arrays by exploiting the parallel-data transmission feature offered by free-space optical communication technology based on light-emitting diodes and a high speed camera. Typical audio capturing technologies face limitations in complexity, bandwidth, and the cost of deployment when aiming for large scalability. In this article, we introduce a prototype that can be easily scaled up to 120 audio channels, which is the world's first and largest real-time optical-wireless sound acquisition system to date.

Keywords: microphone array, free-space optical communication, beamforming

1. Introduction

Imagine the TV broadcast of a live event taking place in a noisy, wide open environment. At the user end, it is often desired to have not only high quality image reproduction but also highly realistic sound that gives a clear impression of the event [1]. This can be achieved by the use of microphone arrays. According to the theory of sensor array signal processing [2], it is possible to listen to a particular sound from a desired location, and also to suppress the noise from the surroundings, by properly aligning and mixing the audio signals recorded by a microphone array. The theory also indicates that large microphone arrays produce remarkable sound enhancement. Examples have been demonstrated with arrays of 100 microphones [3]. However, to accurately record a three-dimensional sound field at a rate of up to 4 kHz and impinging from every direction on a 2-m^2 wall of a room, an array of about 2500 microphones would be needed. Moreover, if the space under consideration is larger, for example a concert hall, tens of thousands of microphones might be necessary. Unfortunately, typical wired microphones and audio recording hardware have limitations in terms of complexity and cost of deployment when the objective is large scalability. Furthermore, the use of multiple wireless microphones is constrained by radio frequency (RF) bandwidth issues.

To overcome these difficulties, we developed a prototype that allows the simultaneous capture of multichannel audio signals from a large number of microphones (currently up to 120). In contrast with existing RF wireless audio interfaces, the proposed system relies on free-space optical transmission of digital signals. Such technology allows the parallel transmission of multiple data channels, each with full bandwidth capacity regardless of the number of channels transmitted.

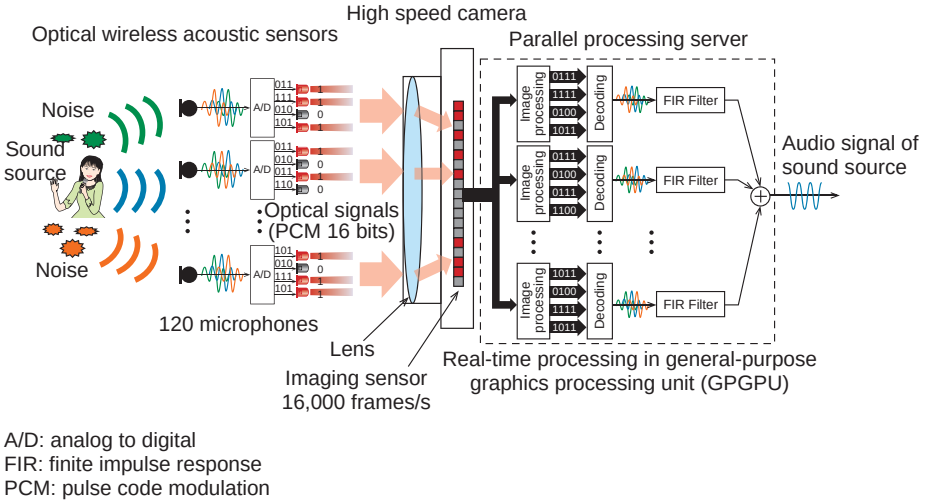


Fig. 1. Architecture of the multichannel audio acquisition system.

2. System description

Our system is composed of three main parts: 1) an optical wireless acoustic sensor (OWAS), 2) a high speed camera, and 3) a parallel processing server. The overall architecture of the system is illustrated in **Fig. 1**.

2.1 OWAS device

An OWAS device is shown in **Fig. 2**. The microphone picks up samples of the acoustic waves at a rate of 16 kHz and outputs a delta-sigma modulated digital stream. Then, a microcontroller converts that serial data into binary symbols of 16-bit pulse code modulation (PCM). The PCM symbols are used to light up an array of 16 light-emitting diodes (LEDs). An LED in the ON or OFF state means a binary 1 or 0, respectively. Because the camera can observe several OWASs simultaneously, the sound field can be sensed with a large array of OWASs such as the one shown in **Fig. 3**, where 200 OWAS devices have been arranged in a 5 × 40-node grid. With this array, our current experimental setup can acquire signals from 120 OWAS devices as allowed by the maximum image size of the camera. Each OWAS device is also equipped with an infrared photosensor that enables it to receive the master clock signal emitted by the pulse generator shown in **Fig. 4**. Therefore, the synchronization between the OWAS devices and the high speed camera is maintained through the master clock generator.

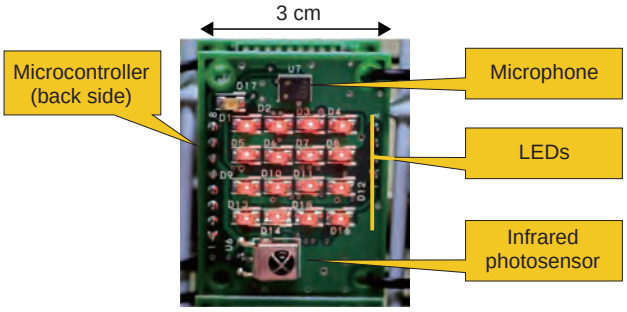


Fig. 2. Photo of OWAS device.

2.2 High speed camera

The imaging sensor of the high speed camera observes the optical signals from the OWAS and records them into intensity images at the rate of 16,000 frames per second (fps). An example of the actual images captured by the camera is shown in **Fig. 5**. To transfer the image data from the camera to the processing server, the camera is connected to a frame grabber card installed on the PCI (Peripheral Component Interconnect) bus of the server. Thus, the flow of the image data can be seen in **Fig. 6**.

2.3 Parallel processing server

The server is equipped with dual CPU (central processing unit) support and a standard general-purpose graphics processing unit (GPGPU), which provides enough massive parallel computing power to process 16,000 images and 120 audio channels within a

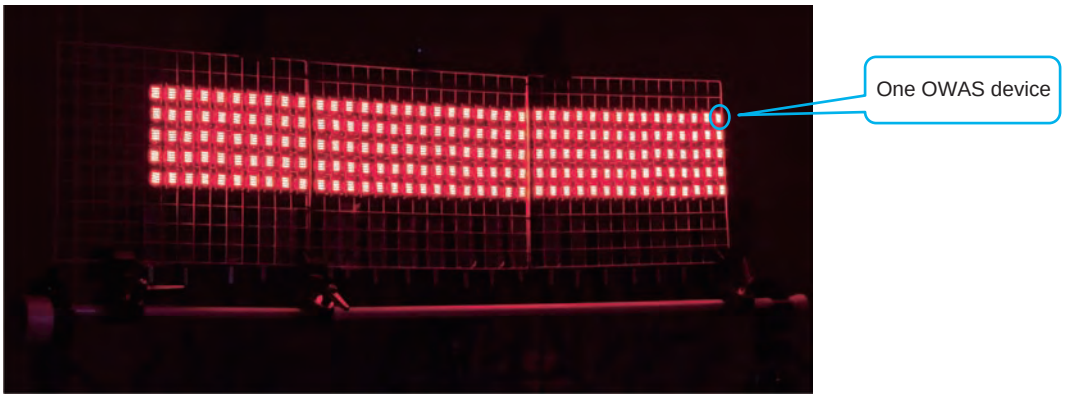


Fig. 3. Array of 200 acoustic sensors (5 × 40).

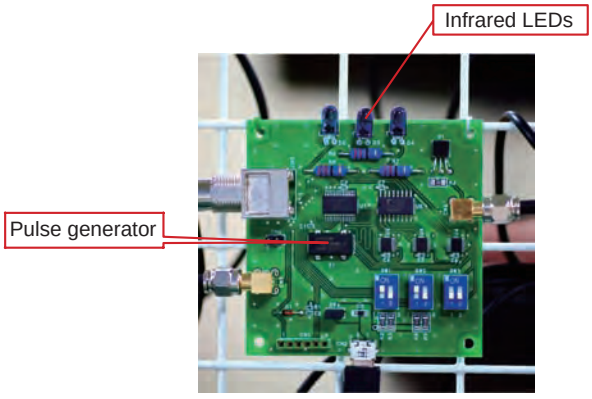


Fig. 4. Infrared transmitter for camera-sensor synchronization.

real-time factor of 0.75. The process to decode the audio signals from the images starts with the detection of the LEDs on the images. Several image processing algorithms have been proposed to accomplish this [4], and it was suggested that the optical transmission channel can be modeled as a MIMO (multiple input multiple output) port as shown in **Fig. 7**. With this model, the pixels of the images can be organized into clusters $\boldsymbol{C}$ by analyzing the spatiotemporal correlation of their intensity signals $\boldsymbol{s}$ across a block of captured frames, as shown in **Fig. 8**. Each cluster of pixels represents a detected LED on the images. Once the pixels for each LED have been identified, their intensity signals are optimally thresholded to convert them back to binary symbols (see Fig. 7). Finally, the binary data is decoded to obtain the originally transmitted audio signals from all the OWAS devices.

These audio signals are further processed with digital filters and mixed down to produce a single output channel (see Fig. 1). This multichannel signal processing, often known as *beamforming* [2], enhances the sound from the desired direction (with respect to the OWAS array), while the noise from other directions is reduced at the output audio signal. In other words, the OWAS array can be acoustically focused in any desired direction. In our preliminary experiments, we have been able to focus on the sound from targets placed as far as 10 m away from the OWAS array while suppressing the noise from the surroundings. An online demonstration video showing the experimental setup is available [6]. Furthermore, we have also achieved the optical transmission of multichannel signals at a distance of 30 m from the receiver camera [5]. Our system also has large scalability. Our numerical simulations indicated that our algorithms can receive and decode the optical signals of as many as 12,000 OWAS devices simultaneously within a single GPGPU card, therefore maintaining real-time processing, as can be seen in **Fig. 9**.

3. Future work

The current limitation we face in expanding our prototype involves the high speed camera. The resolution of existing commercial cameras must be considerably reduced in order to achieve high-speed frame rates (tens of thousands of fps). Nevertheless, the accelerated advances in imaging sensor and parallel computing technologies motivate us to carry out further development of our prototype. In the near future, we expect to build OWAS arrays over large

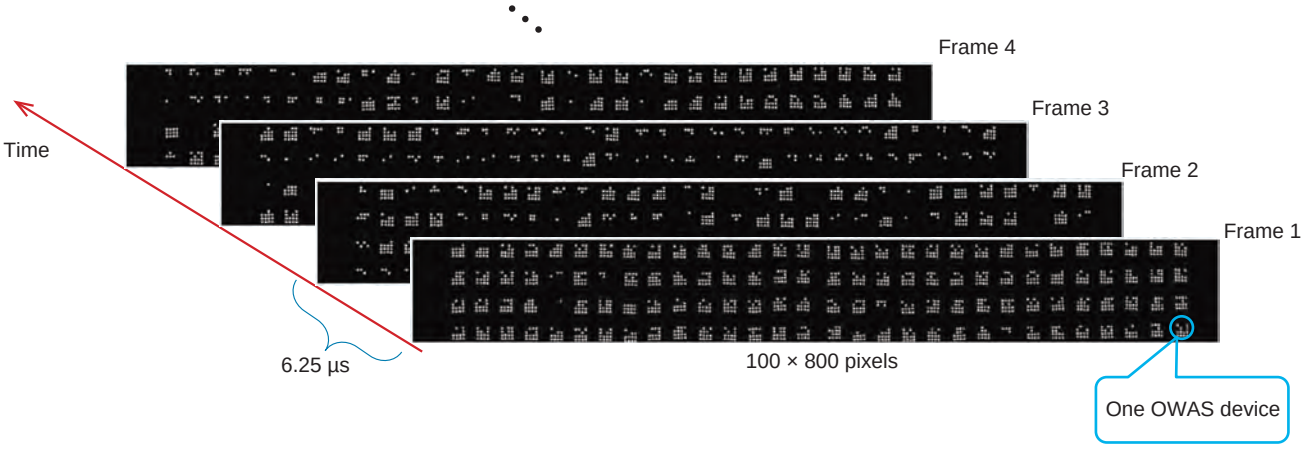


Fig. 5. Images streamed from the high speed camera.

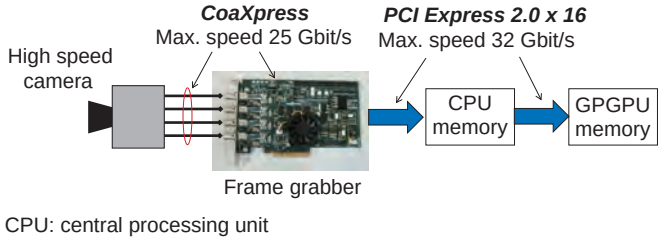


Fig. 6. Data flow from camera to GPGPU.

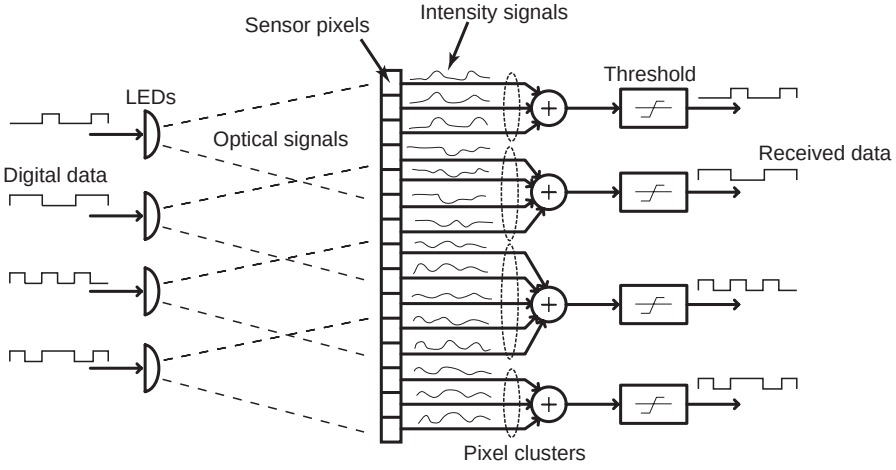
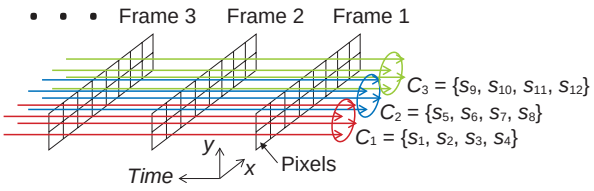
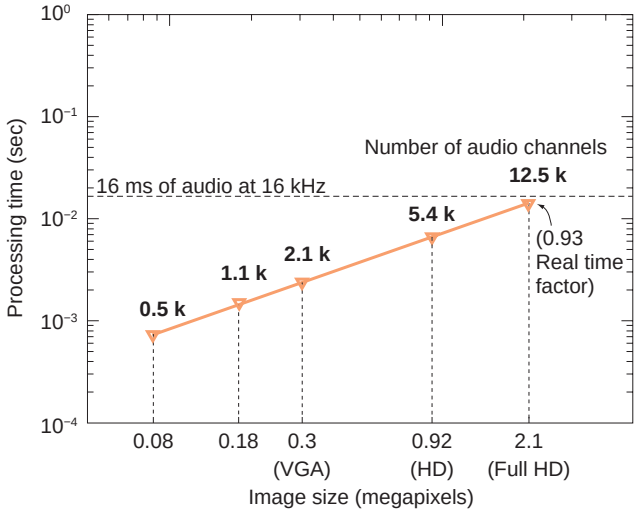


Fig. 7. Optical transmission channel as a MIMO port.



Each cluster *C* represents a detected LED on the images.

Fig. 8. Pixels clustered according to their spatiotemporal correlation.



HD: high definition
VGA: video graphics array

Fig. 9. System scalability in terms of parallel processing power.

spaces such as stadiums or concert halls, and we may achieve real-time position tracking of portable/wearable OWAS devices. Such progress will pave the way for novel applications and high quality services.

References

[1] S. Koyama, Y. Hiwasaki, K. Furuya, and Y. Haneda, "Inverse Wave Propagation for Reproducing Virtual Sources in Front of Loudspeaker Array," Proc. of the European Signal Processing Conference (EUSIPCO) 2011, pp. 1322–1326, Barcelona, Spain, 2011.

[2] H. L. Van Trees, "Optimum Array Processing: Part VI of Detection, Estimation, and Modulation Theory," Wiley-Interscience, New York, 2002.

[3] K. Niwa, Y. Hioka, S. Sakauchi, K. Furuya, and Y. Haneda, "Sharp Directive Beamforming Using Microphone Array and Planar Reflector," Acoustical Science and Technology, Vol. 34, No. 4, pp. 253–262, 2013.

[4] G. P. Nava, Y. Kamamoto, T. G. Sato, Y. Shiraki, N. Harada, and T. Moriya, "Image Processing Techniques for High Speed Camera-based Free-field Optical Communication," Proc. of the IEEE International Conference on Signal and Image Processing Applications (ICSIPA) 2013, pp. 384–389, Melaka, Malaysia, October 2013.

[5] G. P. Nava, Y. Kamamoto, T. G. Sato, Y. Shiraki, N. Harada, and T. Moriya, "Simultaneous Acquisition of Massive Number of Audio Channels through Optical Means," Proc. of the 135th Audio Engineering Society (AES) Convention, New York, USA, October 2013.

[6] Demo video (media player required), mms://csflash.kecl.ntt.co.jp/cslab/mrl/pablo/vasdemo1.wmv

Gabriel Pablo Nava
Research Scientist, Moriya Research Laboratory, NTT Communication Science Laboratories.

He completed his B.E. studies in Mexico City at the Instituto Politécnico Nacional in 1999. He received his M.S. and Ph.D. in 2004 and 2007, respectively, from the Department of Information Science and Technology of the University of Tokyo. He also held a postdoctoral position at the University of Tokyo from March 2008 to April 2014. He joined NTT Communication Science Laboratories in April 2008. He is currently conducting research and development on room acoustics, multichannel audio, microphone array signal processing, and digital image processing for next-generation ultra-realistic teleconferencing systems.

Yoshifumi Shiraki
Research Scientist, Moriya Research Laboratory, NTT Communication Science Laboratories.

He received his B.A. in natural science from International Christian University, Tokyo, in 2008 and his M.S. in computational intelligence and systems science from Tokyo Institute of Technology in 2010. Since joining NTT Communication Science Laboratories in 2010, he has been studying signal processing, particularly distributed compressed sensing. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) and the Institute of Electrical and Electronics Engineers (IEEE).

Hoang Duy Nguyen
Intern, Moriya Research Laboratory, NTT Communication Science Laboratories.

Mr. Nguyen was born in Brussels, Belgium, in 1990. He graduated summa cum laude with an M.S. in electronics and telecommunications engineering from the University of Brussels, Belgium, in 2013. From 2012 to 2013 he worked as a research engineer in the Smart Systems and Energy Technology (SSET) department of the Interuniversity Micro-Electronics Centre (IMEC), Leuven, Belgium. Since 2014 he has been in the Special Research Group of NTT Communication Science Laboratories, Atsugi City, Japan, where he carries out experimental research and development involving free-field optical communication. His research interests include MIMO detection algorithms and implementation aspects of acoustic beamformers on GPGPUs.

Yutaka Kamamoto
Research Scientist, Moriya Research Laboratory, NTT Communication Science Laboratories.

He received his B.S. in applied physics and physico-informatics from Keio University, Kanagawa, in 2003 and his M.S. and Ph.D. in information physics and computing from the University of Tokyo in 2005 and 2012, respectively. Since joining NTT Communication Science Laboratories in 2005, he has been studying signal processing and information theory, particularly lossless coding of time-domain signals. He also joined NTT Network Innovation Laboratories, where he worked on the development of the audio-visual codec for online digital stuff (ODS) from 2009 to 2011. He has contributed to the standardization of coding schemes for MPEG-4 Audio lossless coding (ALS) and ITU-T Recommendation G.711.0 Lossless compression of G.711 pulse code modulation. He received the Telecom System Student Award from the Telecommunications Advancement Foundation (TAF) in 2006, the IPSJ Best Paper Award from the Information Processing Society of Japan (IPSJ) in 2006, the Telecom System Encouragement Award from TAF in 2007, and the Awaya Prize Young Researcher's Award from the Acoustical Society of Japan (ASJ) in 2011. He is a member of IPSJ, ASJ, IEICE, and IEEE.

Takashi G. Sato
Research Scientist, Moriya Research Laboratory, NTT Communication Science Laboratories.

He received his M.S. and Ph.D. in information science and technology from the University of Tokyo in 2005 and 2008, respectively. He is currently a researcher at NTT Communication Science Laboratories. His research interests include bioengineering and tactile and audio interfaces using psychophysiological and neural measurement as feedback information. He is a member of IEEE.

Noboru Harada
Senior Research Scientist, Moriya Research Laboratory, NTT Communication Science Laboratories.

He received his B.S. and M.S. from the Department of Computer Science and Systems Engineering of Kyushu Institute of Technology in 1995 and 1997, respectively. He joined NTT in 1997. His main research area is lossless audio coding, high-efficiency coding of speech and audio, and their applications. He was also with NTT Network Innovation Laboratories, where he worked on the development of the audio-visual codec for ODS, from 2009 to 2011. He is an editor of ISO/IEC 23000-6:2009 Professional Archival Application Format, ISO/IEC 14496-5:2001/Amd.10:2007 reference software MPEG-4 ALS, and ITU-T G.711.0. He is a member of IEICE, ASJ, the Audio Engineering Society (AES), and IEEE.

Takehiro Moriya
NTT Fellow, Moriya Research Laboratory, NTT Communication Science Laboratories.

He received his B.S., M.S., and Ph.D. in mathematical engineering and instrumentation physics from the University of Tokyo in 1978, 1980, and 1989, respectively. Since joining NTT laboratories in 1980, he has been engaged in research on medium- to low-bit-rate speech and audio coding. In 1989, he worked at AT&T Bell Laboratories, NJ, USA, as a Visiting Researcher. Since 1990, he has contributed to the standardization of coding schemes for the Japanese public digital cellular system, ITU-T, ISO/IEC MPEG, and 3GPP. He is a member of the Senior Editorial Board of the IEEE Journal of Selected Topics in Signal Processing. He is a Fellow member of IEEE and a member of IPSJ, IEICE, AES, and ASJ.

Approaches to Technological Fusion of "ICT × Energy × Building" and New R&D Bases

Tsuneo Uekusa, Kenji Yokoyama, and Yosuke Mino

Abstract

NTT FACILITIES has had a research and development (R&D) framework ever since it was established, and we have honed our fusion technology by crossing information and communication technology with energy technology and building technology. We give an overview in this article of our R&D activities and describe the opening of our new R&D base in July 2014. We also explain our objectives in relocating our R&D Headquarters and our efforts underway toward fusing the above technologies at our new base.

Keywords: research and development, technological fusion, NTT FACILITIES Shin-Ohashi Building

1. Introduction

NTT FACILITIES was spun off from NTT in December 1992 and celebrated its 21st anniversary as an independent company at the end of 2013. We have been proactively implementing design and maintenance of power supply systems and building construction projects for the NTT Group, while also expanding the provision of services to the open market by making use of the technologies that have been continually improved over a long period, including when we were with the Nippon Telegraph and Telephone Public Corporation and NTT. We have recently been providing integrated facility services based on consideration of the global environment in all phases from planning, through design and construction, to operation, with the slogan "Smart & Safety®."

For example, in datacenter-related business operations, we have applied our know-how in the design, construction, and maintenance of large numbers of communication buildings; in fact, we have been involved in constructing 30% of the datacenters in Japan, covering 240,000 m^2 of space. We have also simultaneously expanded our business operations that involve the maintenance of overseas datacenters.

We have also contributed to the mega-solar construction business in response to recent increases in the demand for renewable energy. Since the start of the Feed-in Tariff scheme in 2012, we have constructed solar power generation systems totaling over 400 MW at 200 locations all over Japan, and we also own and run solar power generation plants at 50 locations that produce a total of 150 MW of power.

From the start of its foundation, this company has had a research and development (R&D) structure, and we have honed our strengths in information and communication technology (ICT), energy technology, and building technology to realize the "ICT × Energy × Building" fusion (**Fig. 1**). The mission of our R&D Headquarters is to create technology that differentiates us from other companies and that serves as a source of our competitive strength.

2. Locations and organization of NTT FACILITIES R&D Headquarters

Up until now, our R&D facilities were based at three locations: the G.H.Y. Building (offices) in Toshima-ku, Tokyo, the NTT Musashino Research and Development Center (laboratories) in Musashino City, Tokyo, and the main NTT Ochiai Building

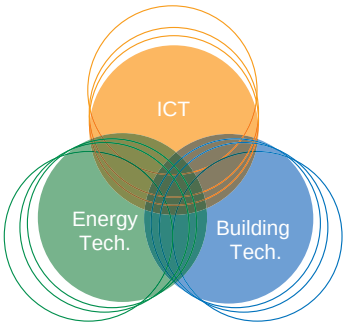


Fig. 1. Technological fusion of “ICT × Energy × Building”.

FM: facility management

Fig. 2. Composition of R&D Headquarters.

(laboratories) in Toshima-ku, Tokyo. We have now created a two-base structure by constructing the NTT FACILITIES Shin-Ohashi Building in Koto-ku, Tokyo, which brings together the functions of the G.H.Y. Building and the NTT Ochiai Building.

The Shin-Ohashi Building contains offices, laboratories for conducting tests on power supply systems aimed at datacenters, testing facilities for verifying our wide-area monitoring system, and shielded rooms for conducting tests related to electromagnetic compatibility (EMC), among other things. At the NTT Musashino Research and Development Center, we continue to make use of facilities such as a three-dimensional (3D) vibration table used for replicating earthquakes and testing the earthquake-resistance of the structural components of buildings and ICT equipment, a simulated datacenter facility for conducting full-size experiments on datacenter air conditioning systems, artificial climate chambers capable of implementing tests on air conditioning units in environments ranging from –10ºC to 50ºC regardless of the climate outside, and temperature-controlled rooms for performing accelerated life testing of storage batteries over long periods.

The R&D Headquarters consists of seven departments (**Fig. 2**). The R&D Strategy Division and the Intellectual Property Office are indirect departments, with the former being responsible for areas such as planning, general affairs, and financial matters of the headquarters, and the latter being responsible for intellectual property management and export control for the NTT FACILITIES Group. The remaining five departments are primarily involved in various R&D activities. We introduce here the fields and strengths of these five departments.

2.1 Environment & Energy Systems Development Division

This division is staffed by air conditioning engineers and monitoring and control engineers, who mainly develop air conditioning units for ICT equipment, integrated control technology for cold water-use refrigerating systems, airflow control technology for datacenter rooms, and technology for the air conditioning systems of datacenters. To date, a cumulative total of more than 50,000 units of the FMACS®-V series and FTASCL® series systems, which are packaged air conditioning units designed for datacenters, have been delivered. These systems make it possible to achieve high energy efficiency throughout the year. NTT FACILITIES has been the first in the industry to develop and bring to market cold aisle containment technology, which is intended to save energy by inhibiting the mixing of supplied cold air and emitted warm air within the datacenter room.

2.2 Power Systems Development Division

The staff of this division consists of electrical power engineers whose area of expertise is primarily power feed technology that focuses on reliability. They develop technology related to power reception, power transformation, shutdown, DC/AC (direct current/alternating current) conversion, power storage, and power distribution for ICT equipment. Their high-voltage direct current (HVDC) power supply system implements higher levels of reliability and power feed efficiency than ordinary AC power systems. Their bidirectional interruption-free switching technology is applied to equipment for supply-and-demand management that freely switches electrical

power from various sources, including power generated by solar power systems and power stored in batteries. This division has also developed and marketed a large flame-retardant lithium-ion battery that is suitable for backup use.

Additionally, they are working on R&D of design, construction, diagnosis, and maintenance technology for renewable energy systems such as solar power generation. They are engaged in the evaluation of solar power generation panels from various vendors, the development of design support tools for in-house use, and the development of malfunction and deterioration diagnosis technology.

2.3 Risk Management Systems Development Division

The engineers in this division specialize in the areas of building structures and the development of countermeasures against EMC problems and thunderstorm damage.

The ICT buildings that form our social infrastructure are of course required to have sufficient earthquake resistance. However, as they are often connected to steel towers, much effort has been put into the R&D of building structural engineering. This division's unique earthquake observation network, 3D vibration testing facilities, and structure simulation program that was developed in-house provides an edge in the area of building structural engineering.

After the Great East Japan Earthquake, there were reports of incidents involving the collapse of air conditioning units suspended from the ceilings of office buildings in the Kanto region around Tokyo, which is some distance from the earthquake center, so the engineers in this division clarified the mechanism of the occurrence through replication experiments using 3D vibration testing equipment. The expertise obtained was immediately reflected in building design.

The building structural engineers are also developing inexpensive frames for solar panels that are easily constructed and highly durable.

EMC and thunderstorm damage countermeasure technology that was primarily intended to protect ICT equipment is now not limited to ICT-related applications only; there is increasing demand for this technology in various fields such as those related to solar power generation and broadcasting. This division provides a consulting service and also applies the knowledge gained to develop countermeasure hardware, design methods, and design tools.

2.4 Advanced FM (Facility Management) Systems Development Division

This division is made up of engineers specializing in the areas of architectural planning and maintenance, facility management, and system development.

Because NTT FACILITIES has been responsible for the planning, design, and maintenance of the substantial number of buildings of the NTT Group, this division has accumulated a great deal of experience in fields such as overall design techniques, design and construction technology specializing in ICT infrastructure buildings, environmentally friendly *green* building design technology, and office design technology that fosters high intellectual productivity.

In addition, in FM practices today, it is necessary to collect information on all facilities owned by a company, make full use of ICT to perform quantitative analyses, and reflect the results in maintenance activities based on sound business judgment. For that reason, this division is developing technology and systems related to checking, measuring, information management, and analysis of buildings, and developing building management systems and entrance/exit management systems that effectively use ICT.

2.5 Total Operation Systems Development Division

The engineers in this division handle the monitoring and maintenance of ICT equipment and facilities, and also the development of related systems. The two bases in Japan perform round-the-clock monitoring of the 20,000 buildings and 200,000 units of ICT equipment of the NTT Group, and this division is in charge of developing facilities for these monitoring activities. Maintenance staff members are also stationed throughout Japan, and this division is also developing technology for supporting them.

3. Technological fusion of "ICT × Energy × Building"

Up until now, this company has accumulated many kinds of technology from diverse fields and provided it in services and products to our customers. For example, datacenter construction requires technological elements of all five departments described above.

However, the demands of customers have been rising recently, and there is also a need to reinforce our competitive strengths, so it has become necessary to apply a fusion of technology not only at the products

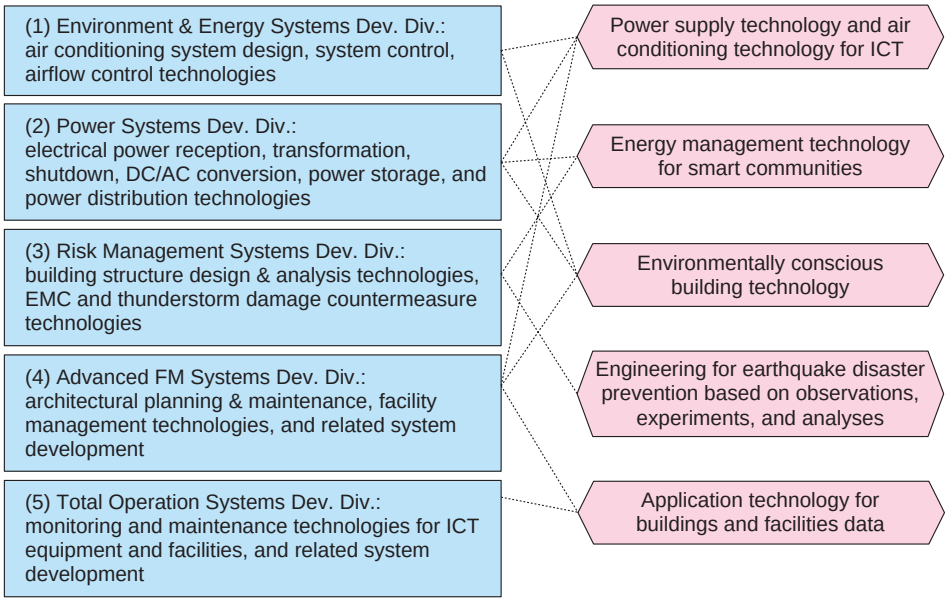


Fig. 3. Elemental technologies and fusion areas.

and services level, but also at the level of technology itself. We describe here an overview of the technological fusion areas introduced in these Feature Articles (**Fig. 3**).

3.1 Power supply technology and air conditioning technology for ICT [1]

There is a strong demand for energy conservation in ICT buildings, especially datacenters, because in addition to the energy consumed by the ICT equipment itself, the power supply and air conditioning equipment also consume a significant amount of energy.

In the past, we have sought out energy-conservation technology for power supplies and for air conditioning equipment. However, we are now conducting R&D on technology that not only increases the energy-conservation capability of both of these types of equipment, but also integrates and controls them to further conserve energy. We are proceeding with plans to provide our customers with this technology by incorporating it into DCIM (datacenter infrastructure management) tools, which are tools for managing assets in datacenters.

3.2 Energy management technology for smart communities [2]

FITBEMS®, NTT FACILITIES' building and energy management system (BEMS) product, applies cloud computing technology and is easy to introduce into even medium- and small-sized buildings due to its low initial cost. This is because FITBEMS does not require the installation of a server for collecting and analyzing data within the buildings. Collecting data in a cloud makes it possible to manage the data collection for a number of buildings and to compare the energy-conservation state of one building with that of a large number of other buildings to obtain a relative evaluation. This will become the basis for constructing a community energy management system (CEMS).

3.3 Environmentally conscious building technology [3]

One example of technological fusion in environmentally friendly building technology is the building information management system (BIS) that our company has proposed. As of now, some companies have started using a dedicated system or BEMS to implement features such as fine control in area units over lighting and air conditioning within offices or to control the intensity of lighting in response to changes in outdoor lighting. However, our company's BIS is targeted at not only conserving energy, but also improving security and increasing user friendliness by creating links to entrance/exit management systems and office automation hardware. For example, the system provides simple implementation of

controls that enable the surrounding lighting and air conditioning to be switched on by logging into a personal computer (PC); it can also disable PC login if there is no record of entry into the room. When there are a number of such controls to be set up, simply connecting the individual systems can make the system extremely complicated. This can be avoided by using BIS as the information distribution platform within the building.

3.4 Engineering for earthquake disaster prevention based on observations, experiments, and analyses [4]

Immediately after the Great East Japan Earthquake, many owners of large buildings received a flood of enquiries from their tenants asking "Is this building safe?" Our Yure-Moni® building safety determination support system makes it possible to estimate the location of damage in a building immediately after an earthquake and to display it on a screen. The system comprises multi-function sensor units in which micro-electro-mechanical systems (MEMS) sensors are installed on each floor. The MEMS sensors are smaller and lighter than previous types of micro-movement sensors, so large numbers of sensors can be used comparatively inexpensively. We have determined the attributes of these sensors in detail through repeated testing on the 3D vibration table and have created systems that use these sensors.

3.5 Application technology for buildings and facilities data [5]

We are subjecting the data we have collected through the monitoring of 200,000 units of ICT equipment and management of 20,000 buildings to multiple analyses and are using the knowledge we have obtained to streamline maintenance tasks and to create new businesses.

In the maintenance management of buildings and equipment, there is a trend to move away from the older style of *preventative* maintenance to *after-the-fact* maintenance in order to meet the demands for cost reductions, but our view is that in the future, we will move towards *predictive* maintenance that applies measurement technology and big data analysis.

4. Opening of new R&D base

We constructed the new NTT FACILITIES Shin-Ohashi Building as a milestone representing 20 years since our foundation and opened our new R&D base in July 2014. The steel-frame building has four floors above ground and one below (a reinforced concrete basement), providing a total floor area of approximately 4300 m^2. (**Fig. 4**)

Fig. 4. NTT FACILITIES Shin-Ohashi Building.

The second and third floors contain offices, and the basement is where the laboratories are located, as shown in **Fig. 5**. There are a number of conference rooms on the west side of the fourth floor, with the space configured so that the layout can be converted using movable partitions. The largest conference room on this floor faces on to a terrace overlooking the Sumida River and can be used for presentations using a multi-screen setup or even for simple receptions. There is a library on the east side of the fourth floor, where people can work and do research using books or stored documents, and there are also central work booths as well as individual desks.

In addition to an entrance hall and waiting area in the lobby on the first floor, there is a dedicated space for collaborative work done between people within the company and with people from other companies, and a "MI-SE-RU Server Room" that visitors can actually visit as an example of an in-house server room that provides a full datacenter solution for this company.

The NTT FACILITIES Shin-Ohashi Building is a highly environmentally friendly building that we are proud to show off to everyone as an example of building office design that provides integrated facility services that consider the global environment. Note that at the end of 2014, we expect to be awarded a Gold certification under the LEED (Leadership in Energy and Environmental Design) program in the New Construction category, which is intended for newly constructed buildings and issued by the US Green Building Council.

We focused on three objectives in this project in the relocation of our facilities: Activate, Assemble, and Unlock (**Fig. 6**).

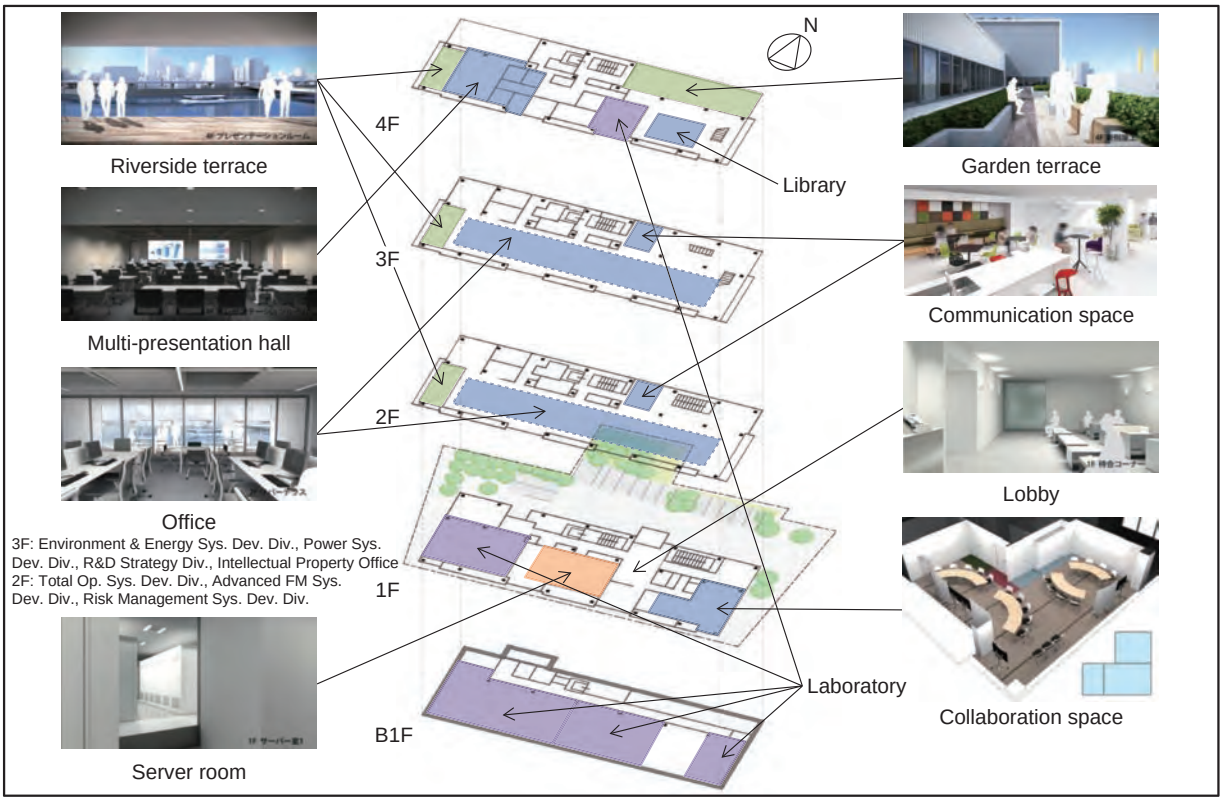


Fig. 5. Floor map of NTT FACILITIES Shin-Ohashi Building.

	Before opening	After opening	Expected effectiveness
Activate	With leased buildings, there are many constraints on building conversion, making them difficult to adapt for research and development.	If we have our own buildings: • We are free to remodel them. • We are free to operate experimentally. • We can also apply data freely.	In a highly flexible R&D field, we can achieve R&D that is: (1) feasible, (2) faster, and (3) more advanced, by using our own buildings and our own staff as test samples.
Assemble	Bases are distributed in three locations. Offices are distributed over five floors. Other companies' technology is dispersed.	• Consolidation of Otsuka base and Ochiai base • Consolidation of offices onto two floors • Provision of outside collaboration space	• Reduced costs and increased efficiency of movement between bases • Promotion of collaborative development between departments, promotion of technological fusion • Collaborative development with our own bases, accelerated use of other companies' technology
Unlock	• Insufficient implementation of in-house technology • Insufficient space to show off in-house technology • Conflict with development information confidentiality	• Incorporation of in-house technology • Environmentally responsive buildings we can be proud of • Designing routes for visitors in advance.	• We can implement our in-house technology ourselves and carry out public relations on it from the user's viewpoint. • Effective public relations exhibiting in-house technology in operation • CASBEE S class, 44% reduction of CO_2 emissions (57% cut in office areas) • We can maintain confidentiality of development data and also achieve openness to visitors.

Fig. 6. Objectives for relocation project.

4.1 Activate

Because this new base is our own building, we have configured it to enable us to freely remodel it and incorporate different technologies. We *activate* the flexibility of having our own building as strengths and aim to improve service quality by having developers act as users to implement their own products within the building.

4.2 Assemble

We have implemented a setup at our new base that stimulates communication between users. We stimulate innovation through technological fusion by promoting interaction between engineers from different fields. We expect that this will enable us to generate new technologies that could not have been achieved before.

We have also set up an *outside collaboration space* that was designed on the premise that collaborative work should be encouraged between our employees and people from other companies in order to promote open innovation. We are attempting to develop innovative services and technologies by gathering not just our own knowledge, but also the wisdom of the world.

4.3 Unlock

The NTT FACILITIES' Shin-Ohashi Building incorporates a great deal of our in-house technology that we implement as users ourselves, and we would also like to show customers a *living* face in which we use our technology in practice. We intend to disseminate information more proactively than in the past and will also combine it with the open innovation described previously.

5. Future plans

With the dismantling and privatization of national telecom carriers taking place around the world, for example, AT&T in the US and BT in the UK, there is currently no other company in the global marketplace that is continuing to research and develop technology centered on ICT infrastructure at a scale greater than ours. We have created a new R&D base and will work to promote the fusion of technology in various fields within and outside our company, and to stimulate R&D into services and technology that will please our customers around the world.

References

[1] Y. Udagawa and H. Matsuo, "Power Supply Technology and Air Conditioning Technology for ICT," NTT Technical Review, Vol. 12, No. 11, 2014.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa10.html

[2] T. Watanabe and T. Shimakage, "Energy Management Technology for Smart Communities," NTT Technical Review, Vol. 12, No. 11, 2014.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa11.html

[3] H. Tanno, T. Watanabe, T. Matsuoka, Y. Kohata, and T. Nakata, "Environmentally Conscious Building Technology," NTT Technical Review, Vol. 12, No. 11, 2014.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa12.html

[4] D. Chiba, K. Yoshida, and M. Suzuki, "Engineering for Earthquake Disaster Prevention Based on Observations, Experiments, and Analyses," NTT Technical Review, Vol. 12, No. 11, 2014.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa13.html

[5] M. Yamazaki, T. Matsuoka, T. Tsujikawa, and S. Nakashima, "Application Technology for Buildings and Facilities Data," NTT Technical Review, Vol. 12, No. 11, 2014.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa14.html

Tsuneo Uekusa
Senior Executive Manager, Research and Development Headquarters, NTT FACILITIES, INC.

He received the B.E., M.E., and Ph.D. in mechanical engineering from Waseda University, Tokyo, in 1983, 1985, and 1999, respectively. He joined NTT in 1985 and has been with NTT FACILITIES since 1992. He has long been engaged in R&D of air-conditioning systems for ICT equipment, including the FMACS® series. During 2000–2004, he worked for ENNET Corporation as a senior sales manager. During 2009–2011, he was the head of the real estate planning office of NTT WEST.

He is a Professional Engineer (Mechanical Engineering). He is a fellow of the Society of Heating, Air-Conditioning and Sanitary Engineers (SHASE) of Japan, and a fellow of the Japan Society of Mechanical Engineers (JSME).

Kenji Yokoyama
Senior Manager, R&D Strategy Division, Research and Development Headquarters, NTT FACILITIES, INC.

He received a Ph.D. in process engineering from Osaka University in 1990 and joined NTT Opto-Electronics Laboratories the same year. He has researched and developed chemical sensors, hologram materials, and optical communication devices. From 2000 to 2010, he was a general manager at NTT, where he established new business endeavors in the fields of wireless networks, SaaS, datacenters, network interconnection, monetary claims systems, and energy systems. From 2010 to the present, he has been working on creating new businesses and technologies based on energy, buildings, and ICT at NTT FACILITIES.

Yosuke Mino
Senior Research Engineer, Environment and Energy Systems Development Division, Research and Development Headquarters, NTT FACILITIES, INC.

He received the B.E. and M.E. in architecture from Tokyo Institute of Technology in 1996 and 1998, respectively. He joined the R&D Headquarters of NTT FACILITIES in 1998, where he began studying and developing air conditioning systems for ICT equipment. He was posted to a subsidiary company for three years and then went back to the R&D Headquarters. He is now again involved in R&D of air conditioning systems for ICT equipment. He is a Professional Engineer (Environmental Engineering) and a qualified architect of the first class. He is a member of the Architectural Institute of Japan (AIJ) and SHASE.

Power Supply Technology and Air Conditioning Technology for ICT

Yosuke Udagawa and Hidenori Matsuo

Abstract

To address the problem of reducing the power consumption of information and communication technology (ICT) buildings, NTT FACILITIES has been developing strategic technology for power and air conditioning systems from the equipment level such as highly efficient rectifiers and air conditioners, to the system level such as high voltage direct current power supply systems and task ambient air conditioning systems. Here, we introduce our power supply technology, air conditioning technology, and integrated control technology for ICT equipment, power supplies, and air conditioning systems that have been developed as part of our efforts toward reducing power consumption.

Keywords: datacenter, power supply, air conditioning

1. Introduction

In recent years, information and communication technology (ICT) equipment has come to play a bigger role in our daily lives, as it supports our social and economic activities. Communication buildings and datacenters manage ICT equipment centrally, and as ICT equipment continues to permeate our society, the importance of these facilities as social infrastructure increases. The amount of power consumed is increasing in parallel with the growing importance of these ICT buildings and facilities, and reducing power consumption has therefore become an important issue.

Communication buildings that house ICT equipment also contain power equipment that drives the ICT equipment and air conditioners that cool the ICT equipment. NTT FACILITIES has been contributing to reducing ICT power consumption for over 30 years by conducting research and development on power and air conditioning technology and producing practical results. Our efforts in developing such technology have resulted in the stepwise development of technology from the equipment level, for example, highly efficient rectifiers and air conditioners, to the system level, including high voltage direct current (HVDC) power supply systems and task ambient air conditioning systems.

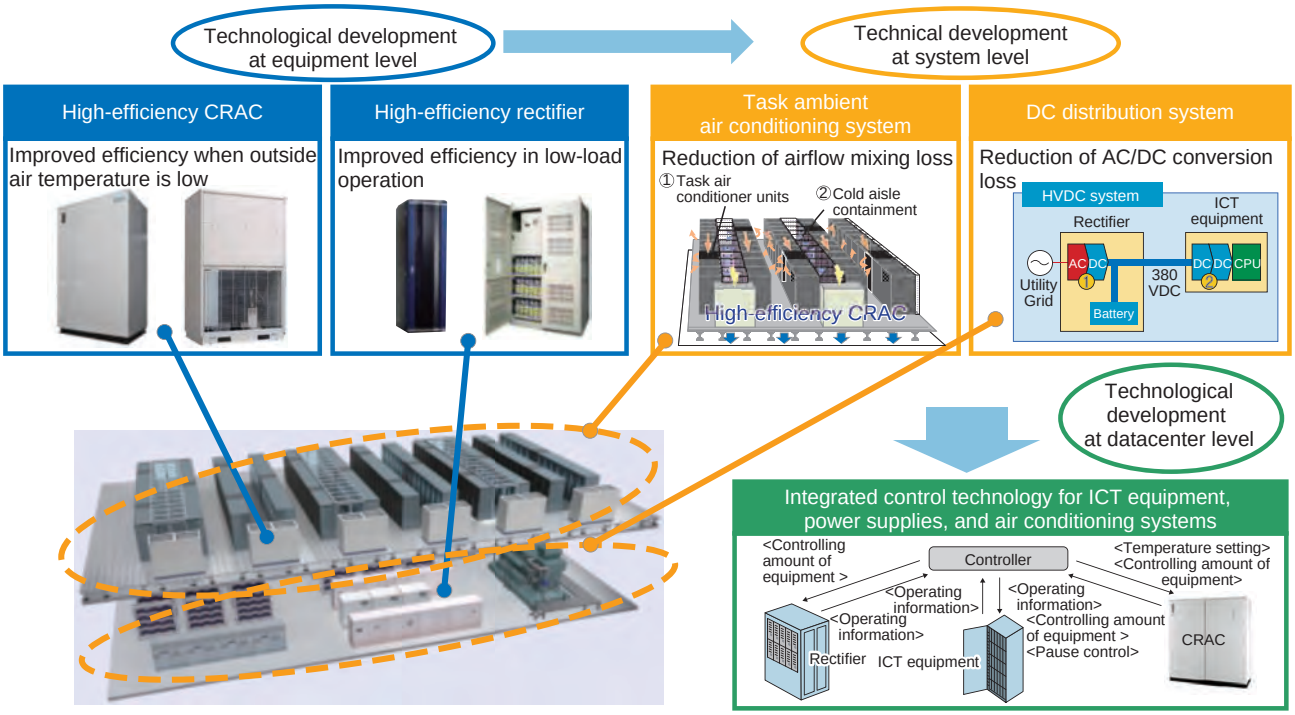
However, there are limits to achieving further reductions of the overall power consumption of ICT buildings through improved power and air conditioning technology alone. Therefore, as we develop technology for controlling power supplies and air conditioning systems, we are also focusing on ways to include the ICT equipment in the control technology (**Fig. 1**).

Our work in this area has been received well by scientific associations in Japan and other countries and by various organizations outside the NTT Group, and it has received various awards.

In this article, we describe our power supply and air conditioning technology as well as our integrated control technology for ICT equipment, power supplies, and air conditioning systems, which is designed to further reduce power consumption.

2. Power supply technology

ICT equipment operates continuously—24 hours a day, 365 days a year—and the power it consumes is increasing. For these reasons, systems that supply power to the ICT equipment are required to have both high efficiency (low loss up to when the power is supplied to the ICT equipment) and high reliability (no interruption of the power supplied to the ICT



AC/DC: alternating current/direct current
CPU: central processing unit
CRAC: computer room air conditioner

Fig. 1. Efforts to improve power supply and air conditioning technology.

equipment). Since the days of the Nippon Telegraph and Telephone Public Corporation, NTT FACILITIES has been working on equipment development, system design, and facilities maintenance technology in parallel with the evolution of network services.

In our past work, we have had success in reducing power consumption in ICT, but while the power consumption of conventional switches and transmitters has been less than 1.8 kW per rack, the most recent high-capacity routers and servers may consume as much as 15 kW per rack. There is thus a need to develop a way to supply power with both high reliability and high efficiency.

3. HVDC power supply systems

Datacenters conventionally supply alternating current (AC) power to ICT equipment via an uninterruptible power supply (UPS). The UPS is connected to storage batteries to provide backup in the event of an interruption of the commercial power supply. The charging current for storage batteries is DC. AC power supply systems convert the AC power to DC and then convert the DC back to AC. It is then necessary to again convert the AC to DC inside the ICT equipment. There is thus a total of four power conversion steps in an AC power system, and power loss occurs at each step.

In the DC power system that we recommend as a solution to this problem, AC power is converted to DC and is then supplied directly to the ICT equipment, thus reducing the power conversion to two steps. To achieve even higher efficiency, we are working toward changing the conventional DC power supply voltage of 48 V or so to a higher voltage of 380 V in an HVDC power supply system (**Fig. 2**). Doing so results in a current reduction of one-eighth for the same transmitted power and a reduction in wiring loss between the highly efficient rectifier and the ICT equipment. In parallel with that work, we have also been working on the development and standardization of DC power sockets and plugs to maintain the safety and simplicity that are needed in the operation of ICT equipment.

Storage batteries are needed for a highly reliable power supply, and we have developed flame-retardant

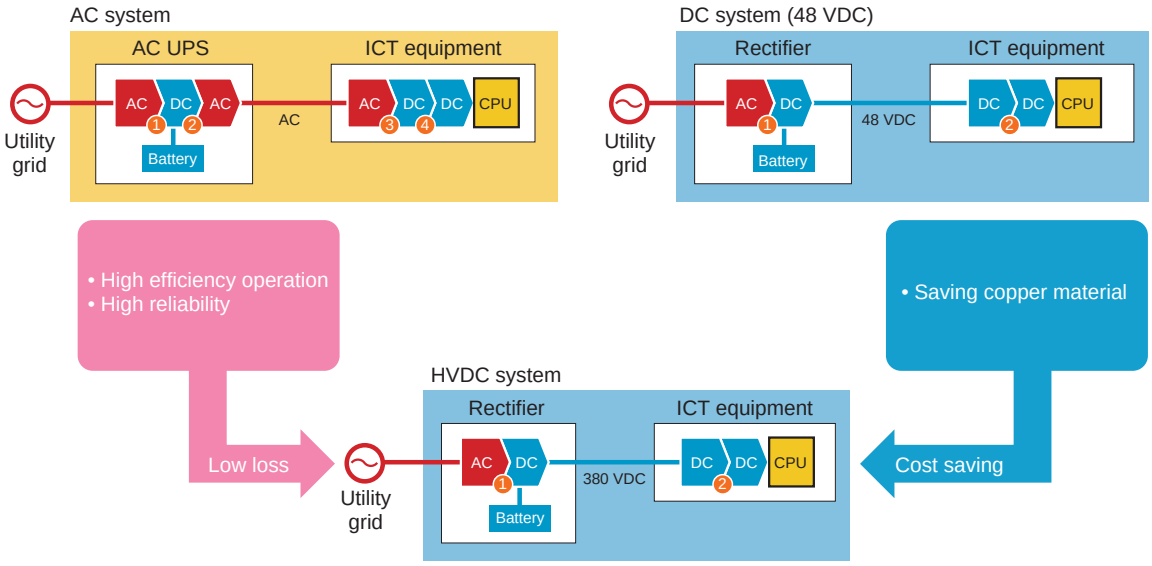
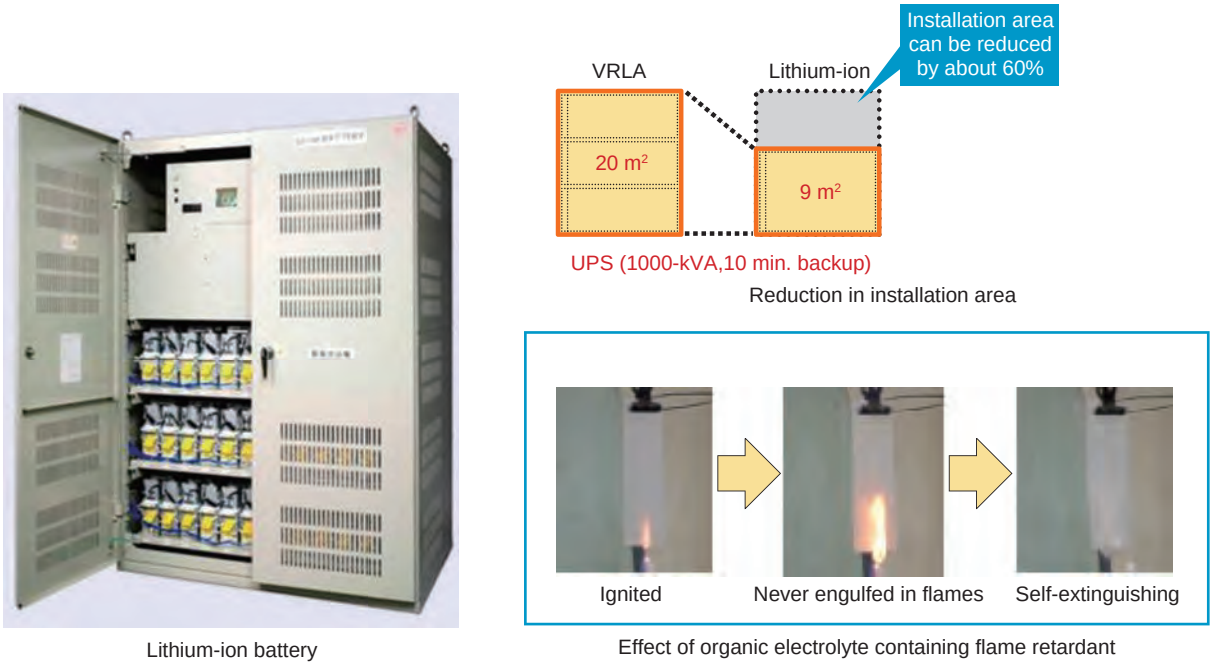


Fig. 2. Comparison of power systems.



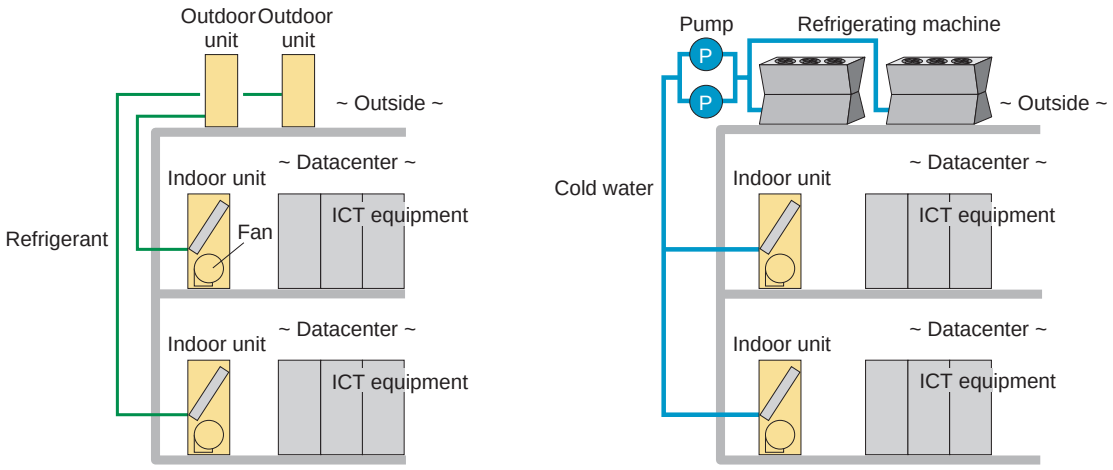
Lithium-ion battery

VRLA: valve-regulated lead acid

Effect of organic electrolyte containing flame retardant

Fig. 3. Flame-retardant lithium-ion battery.

lithium-ion batteries that are lighter and require up to 60% less space than lead-acid storage batteries (**Fig. 3**). The lithium-ion batteries have a flame-retardant additive in the electrolyte to prevent fire or rupture in abnormal conditions for increased safety. Adopting the flame-retardant lithium-ion batteries can both save space in electrical rooms and conserve material resources used for base frame by reducing



(a) Air-cooled package system (CRAC)

(b) Cold water-use system

Fig. 4. Air conditioning systems.

the weight of the equipment.

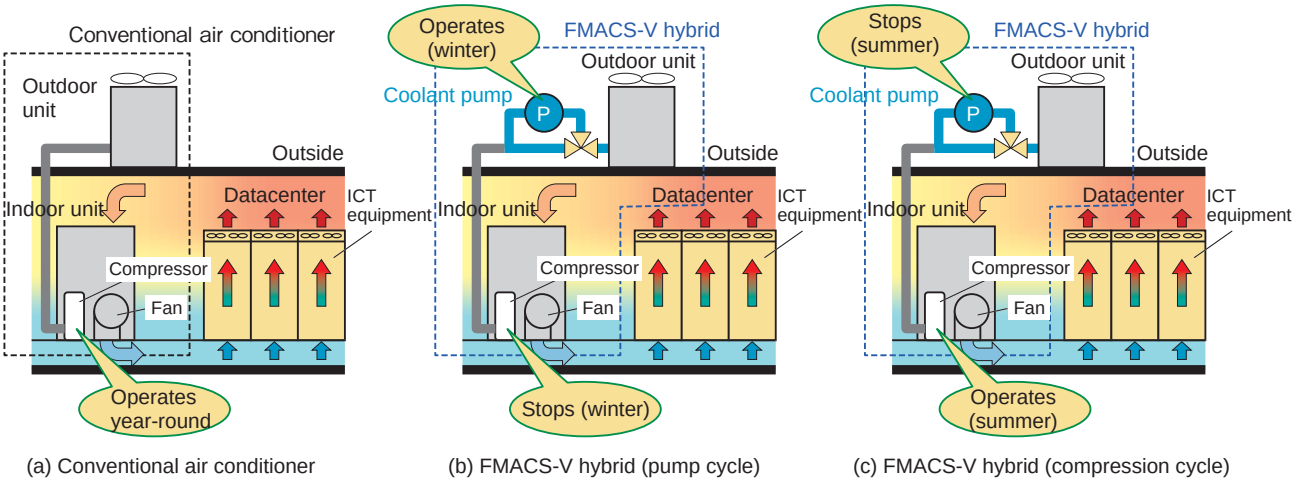
4. Air conditioning technology

ICT equipment generates a large amount of heat during operation, and consequently, it is cooled by special air conditioning systems. Generally, ICT equipment is cooled by placing air conditioning units along a wall and using the area below a raised floor as an air flow space for the cooling air. To cope with the higher heat density within the room that has come with the change from analog to digital ICT equipment and conversion to IP (Internet protocol), we developed special air conditioning technology for ICT equipment. Here, we describe the FMACS®-V (facilities multiunit air conditioning system-5th generation) hybrid, which is an indirect outdoor air-cooled package air conditioner for ICT equipment. We also describe a task ambient air conditioning system.

The FMACS-V and other conventional mainstream air conditioning systems are air-cooled package-type air conditioners that can easily accommodate successive expansion of ICT equipment and can also be backed up by other air conditioning units when a failure occurs, thus providing high reliability (**Fig. 4(a)**). However, customer needs have been diversifying in recent years, and more large scale datacenters now demand cold water air conditioning systems, which can hold down the initial costs (**Fig. 4(b)**). To respond to the wider customer needs, we developed our Unified Cooling system to a practical level. Unified Cooling involves control of the overall system to achieve high efficiency in a cold water air conditioning system.

5. FMACS-V hybrid

ICT equipment generates a large amount of heat in normal operation, and to cool that equipment, conventional air conditioning units use compressors throughout the year (**Fig. 5(a)**). The FMACS-V hybrid system, on the other hand, uses a coolant pump together with a compressor. During the winter and other times when the outdoor air temperature is low, air conditioning units perform cooling by using the coolant pump and the cold outdoor air with the indirect outdoor air cooling method (**Fig. 5(b),(c)**). The coolant pump can operate on about one-eighth the power needed by a compressor, so the energy saving is proportional to the coolant pump operating time. Thus, the effect in reducing power consumption is high in regions where the outdoor air temperature is low throughout the year. The annual power consumption of the FMACS-V hybrid is up to 54% less than that of conventional air conditioning units in Sapporo and up to 42% less in Tokyo. Therefore, this system can contribute greatly to reducing the power consumption of air conditioning systems.



(a) Conventional air conditioner (b) FMACS-V hybrid (pump cycle) (c) FMACS-V hybrid (compression cycle)

Fig. 5. Air conditioner operating conditions.

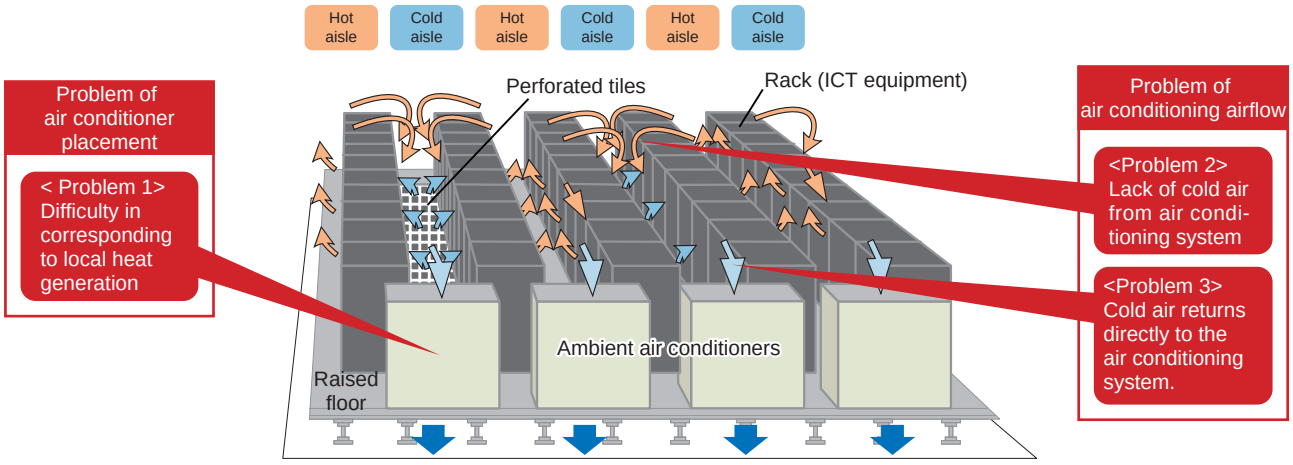


Fig. 6. Problems with ambient air conditioning systems.

6. Task ambient air conditioning systems

A task ambient air conditioning system combines ambient air conditioning, in which wall-mounted air conditioning units cool the entire room, and task air conditioning, in which a local area is cooled. Because air conditioning systems for ICT equipment produce about 20 times as much cold air flow as office air conditioning systems, an effective way to reduce power consumption is to reduce the power needed to produce the air flow. Ambient air conditioning facilitates successive expansion through rearrangement of the ICT equipment and the perforated panels through which cold air is drawn up from the floor. However, the following problems make it difficult to reduce the power needed to move air (**Fig. 6**).

6.1 Air conditioning unit placement

With ambient air conditioning, generation of high local heat requires that air be moved from wall-mounted air conditioning units to distant ICT equipment via a raised floor. A large amount of power is needed to drive the air movement.

6.2 Air conditioning air flow

With ambient air conditioning, cool air from the air conditioning units passes through the space beneath a raised floor and then through a perforated panel to

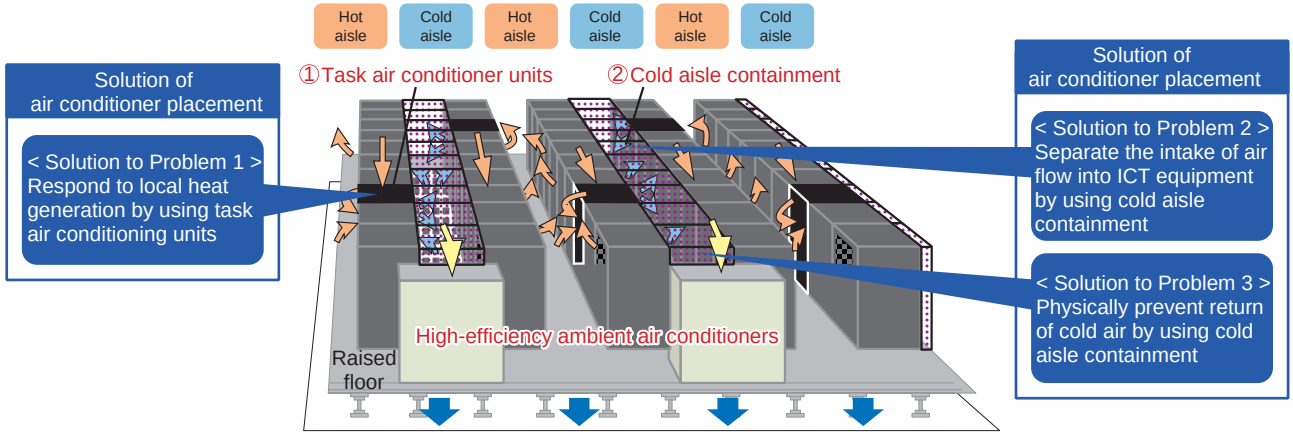


Fig. 7. Solving cooling problems with a task ambient air conditioning system.

reach the front side of the ICT equipment. It then passes through the inside of the ICT equipment where it is heated and then returned to the air conditioning units. However, if insufficient cool air is supplied from the perforated panels in the cold aisle (the corridor from which cool air is supplied), the high-temperature exhaust recirculates from the hot aisle (the air exhaust corridor for the ICT equipment) to the cold aisle, and the target temperature for the air drawn into the ICT equipment cannot be maintained. To maintain the ICT equipment air intake temperature, ambient air conditioners lower the temperature of the entire room to the set temperature. In that case, it is often not possible to deal with excessive air flow, resulting in inefficient operation. Also, some of the cool air from the air conditioning units is returned directly to the air conditioning units without cooling the ICT equipment, which is a further waste of power for the air flow.

6.3 Solutions to problems

A task ambient air conditioning system that can solve the two problems of ambient air conditioning systems (Fig. 6) is shown in **Fig. 7**. The task ambient system uses 30% less power than the ambient system. Specific solutions include those described below.

The problem of air conditioning unit placement is solved by using indoor air conditioning units that have the same form factor as the ICT equipment racks. This makes it possible to install FTASCL®-RS/C (facilities task cooler rack-shaped/compressor type) task air conditioner units near the ICT equipment that generates relatively high heat. Although wall-mounted air conditioning units must move a large volume of cool air a long distance through the raised floor space to deal with hot spots, the FTAS-RL-RS/C unit takes air in from the back side of the air conditioning unit and expels it to the front side, so the air flow distance is shorter, and less power is used for air transport.

The problem of the air flow of air conditioning can be addressed by capping the cold aisle to physically prevent recirculation of hot exhaust air to the cold aisle. It is also possible to prevent the cool air from being returned directly to the air conditioning units. Solving the air flow problem with cold aisle containment means that power can be saved by reducing the power needed to produce the air flow and by optimizing the temperature settings.

7. Unified Cooling

The main type of air conditioning systems used in datacenters has been the air-cooled package system, which comprises one indoor unit and one outdoor unit (Fig. 4(a)). In recent years, however, the diversification of customer needs has led to an increase in the use of cold water systems. Air conditioning systems that use cold water consist of a refrigerating system that produces the cold water, a water supply system for transporting the water, and an air conditioning system for cooling the room with the transported cold water. Each of these systems comprises multiple units of equipment (Fig. 4(b)). Conventional cold water air conditioning systems use separate controls for the refrigerating system, water supply system, and air conditioning system (**Fig. 8**), and there is no overall energy-conserving operation of the

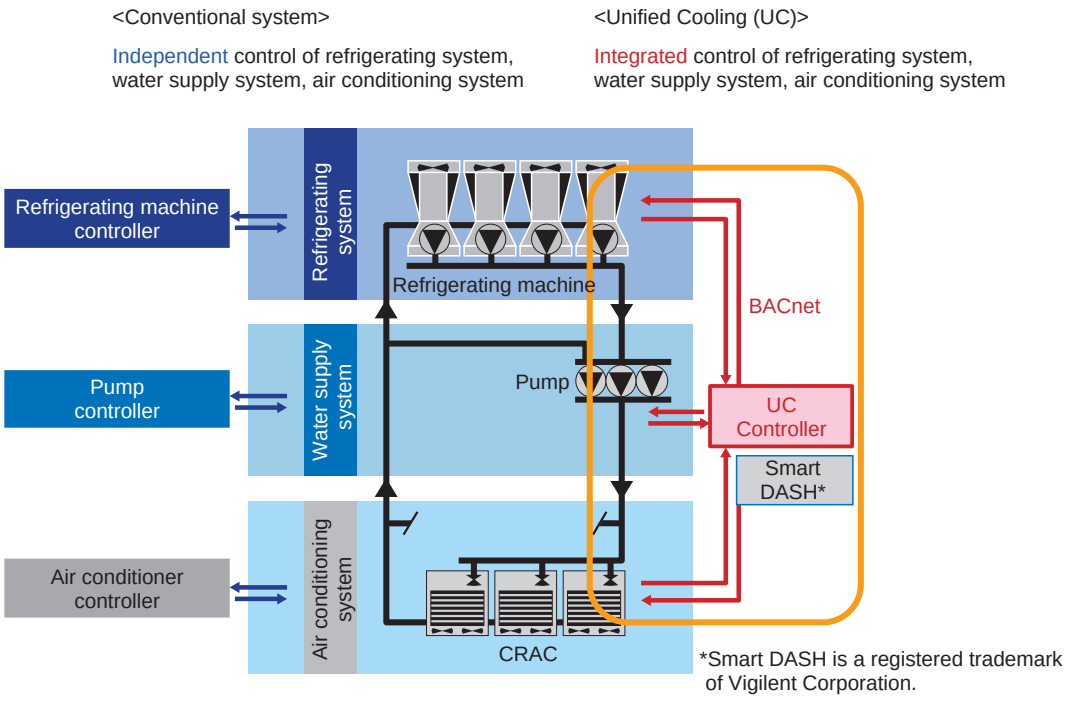


Fig. 8. Depiction of Unified Cooling system.

system. We therefore developed Unified Cooling (Fig. 8), which provides unified control of all three subsystems to achieve overall energy efficiency for the entire system. In maintaining the specified room environment, the Unified Cooling system adjusts the settings for the output temperature of the thermal system and the pressure of the water transport system. Integrating this control with Smart DASH®, which controls the air flow of the air conditioning system, makes it possible to reduce power consumption by up to 30%.

8. Integrated control technology for ICT equipment, power supply, and air conditioning systems

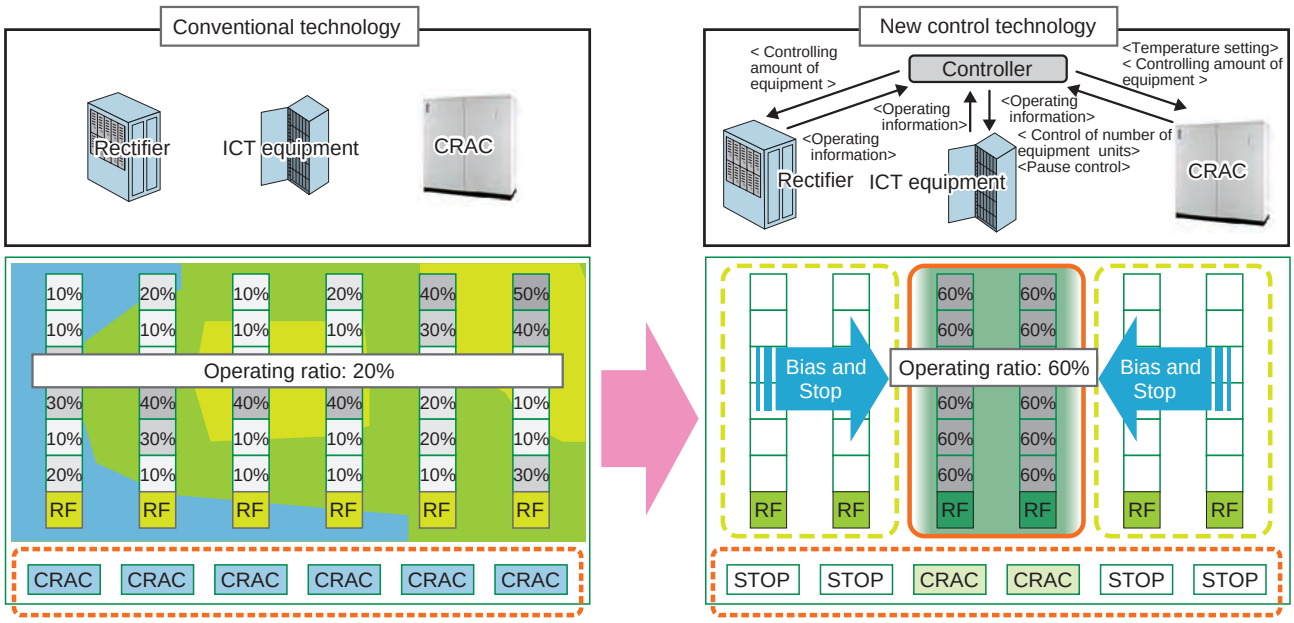
In recent years, large variations have been observed in the amount of data processing done within certain time periods by datacenter ICT equipment. ICT equipment generally has high energy efficiency when operated under high-load conditions, but low-load operation occurs routinely. Also, the correlation between ICT equipment power consumption and the amount of data processing is not proportional; a constant amount of power is being consumed even in standby states when no data processing is being done. Thus, power is being wasted when data is not being processed.

An efficient operating state can be maintained by shifting the processing to particular ICT equipment in time periods when the amount of data processing is low. Doing so makes it possible to turn off the power to ICT equipment that is in the standby state, thus reducing power consumption. Reducing the power consumption also reduces the sources of heat generation, so it may be possible to halt the operation of some air conditioning units. A sequence of controls of this kind (**Fig. 9**) is expected to reduce the total power consumption by the ICT equipment, the rectifier, and the air conditioning units by at least 50% compared to ordinary datacenter systems.

This work was conducted in collaboration with the NTT Energy and Environment Systems Laboratories.

9. Conclusion

In this article, we reviewed power supply and air conditioning technology. We also described the integrated control technology we developed to control



RF: rectifier

Fig. 9. Integrated control technology for ICT equipment, power supply, and air conditioning systems.

ICT equipment, power supplies, and air conditioning systems to further reduce ICT power consumption. We will continue working on technological development of such systems so as to continue contributing to the reduction of power consumption in the future.

The FMACS-V hybrid and FTASCL-RS/C are being developed with the cooperation of Hitachi Appliances.

Yosuke Udagawa
Research Engineer, Research and Development Headquarters, NTT FACILITIES, INC.
He received the B.E. and M.E. in mechanical engineering from Waseda University, Tokyo, in 2003 and 2005, respectively. He joined NTT FACILITIES in 2005. He worked on the development of the FMACS-V hybrid system and is currently studying air-conditioning systems for datacenters. He is a member of the Japan Society of Mechanical Engineers, the Japan Society of Refrigerating and Air Conditioning Engineers, the Society of Heating, Air-conditioning and Sanitary Engineers of Japan, and the Architectural Institute of Japan.

Hidenori Matsuo
Research Engineer, Research and Development Headquarters, NTT FACILITIES, INC.
He received the M.E. in human and environmental studies from Kyoto University in 2000 and joined NTT FACILITIES the same year. He worked on the development of HVDC rectifiers that reduce annual power consumption. He is currently studying 48 VDC rectifiers for telecommunication sites. He is a member of the Institute of Electrical Engineers of Japan.

Energy Management Technology for Smart Communities

Takeshi Watanabe and Toyonari Shimakage

Abstract

Various energy-saving efforts are now being made throughout Japan not just to put a brake on global warming but also to stabilize the supply and demand of electric power and to reduce the escalating energy prices in the wake of the Great East Japan Earthquake. In particular, the construction and arrangement of smart communities is now being promoted as an effective energy-saving solution. A smart community must be able to perform integrated management of electric power and other energy sources including thermal energy and untapped energy for various types of consumer facilities such as homes and office buildings. This article introduces key issues surrounding the smart community and essential elemental technologies and presents a case study of an actual implementation.

Keywords: smart community, cloud-based xEMS, supply and demand optimization

1. Introduction

Since the Great East Japan Earthquake of 2011, extensive efforts have been made in Japan to save power in order to cope with the severe situation regarding the supply and demand of electricity. This situation not only hinders everyday life but also corporate productivity and business activities, and there are concerns that it may even lower the quality of human life. The need for power-saving measures during times of peak power consumption such as in the summer and winter seasons has consequently become an issue of concern throughout society.

A *smart community* is a social infrastructure that combines information and communication technology (ICT) and various types of energy-related technologies to enable people to lead safe and comfortable lives while considering the needs of the global environment. Through the use of ICT, smart communities aim to provide a stable supply of power from solar power generation, wind power generation, and other renewable energy sources that fluctuate in output. They also seek to optimally manage the demand generated by ordinary households and office buildings through the use of storage batteries and other measures. A smart community is a social system that can provide an adequate supply of energy without waste through the formation of a network of diverse energy-related devices and the visualization of the power generation and power consumption of those devices.

2. Steps toward achieving smart communities

The concept of the smart community has spread throughout the world from developed to emerging nations, and a variety of technical proposals and standards have been formulated from diverse sectors in society including the government, public institutions, and private enterprises in each country. Against this background, NTT FACILITIES seeks to contribute to the creation of smart communities and a future society that can provide a safe, comfortable, and environmentally conscious supply of energy. The company plans to accomplish this by bringing together a wide range of technologies combining ICT, power, and construction that have been accumulated over many years at NTT from the early days of Nippon Telegraph and Telephone Public Corporation to the present (**Fig. 1**).

The smart community forms a society that targets not only independent, scattered facilities such as

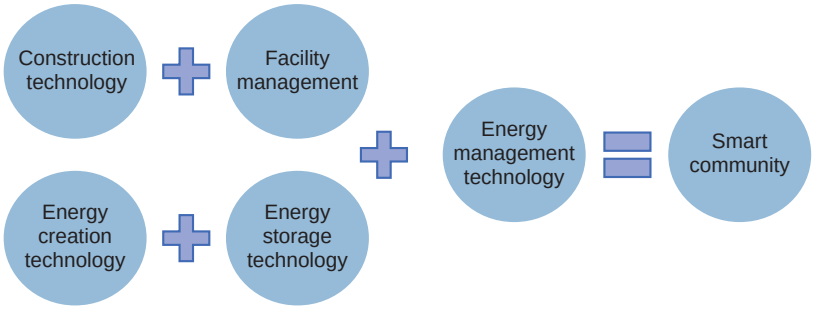
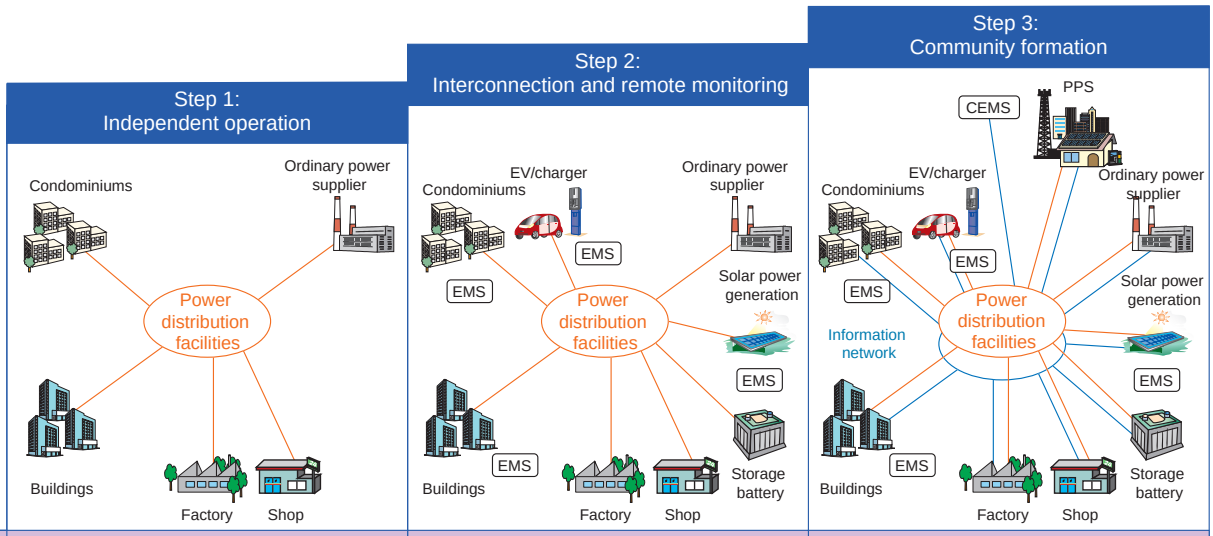


Fig. 1. Smart community approach at NTT FACILITIES.



	Step 1: Independent operation	Step 2: Interconnection and remote monitoring	Step 3: Community formation
Monitoring & control	Independent monitoring and control in units of buildings and systems	Monitoring and control by EMS constructed for each building group and system	Connects and integrates the entire community and performs monitoring and control by CEMS linked to power companies
Service	• Achieves optimal operation and energy savings in units of buildings and systems	• Achieves optimal operation and energy savings for all buildings/systems in a group • Efficient operation through remote monitoring and group control	• Achieves optimal operation and energy savings for both the demand side and supply side • Remote monitoring and group control targeting the entire community

CEMS: community energy management system
EMS: energy management system
EV: electric vehicle
PPS: power producer and supplier

Fig. 2. Steps in achieving a smart community.

residences and buildings but also an entire area for the deployment of a smart system. Achieving a smart community takes several steps, as shown in **Fig. 2**.

The first step is to perform monitoring and control that can achieve energy savings and optimal operation for each building, system, and device. The next step is to interconnect those nodes via an Internet protocol (IP) network to perform remote monitoring and control and achieve optimal operation for each group of buildings and systems. The final step is to build a smart community that integrates the elements of the entire community.

3. Elemental technologies of a smart community

A smart community consists of many elements including consumer premises such as homes and office buildings, consumer facilities such as various types of equipment (air conditioners, lighting, etc.) and electric vehicles (EVs), and energy-supply facilities such as solar power generators and power generating stations. These elements need to be managed in an integrated manner. In the following, we introduce NTT FACILITIES' approach to achieving a smart community.

3.1 Position of office buildings in a smart community

Consumer premises in a smart community must be connected to a network and be active in dispatching information to the outside while being appropriately controlled in accordance with requests from the supply side. Energy management systems such as the home energy management system (HEMS) and building and energy management system (BEMS) can provide the technology needed to meet these requirements. Among consumer premises, it is extremely important that office buildings in particular be controlled owing to the following features that set them apart from homes.

- They are concentrated in urban areas with high individual energy consumption.
- They frequently include power/heat storage facilities and power generation equipment that can act as a buffer in controlling energy supply and demand such as in facility-capacity control.
- They will have a major impact on society if their energy supply runs out.

Despite these features, it has not been very common in the past to connect office buildings to the outside world and to control their energy demands for the following reasons.

- Information obtained from measurements can differ depending on the source, making it difficult for the supply side to identify and manage.
- Energy consumption characteristics can be complicated and nonlinear, making demand forecasting difficult.
- Information networks have had a closed format.

Recently, however, it has become possible to make BEMSs compatible with IP networks and open-system technologies. In this way, BEMSs will be able to contribute to the stabilization and optimization of energy supply and demand in buildings and in smart communities.

3.2 Cloud-based BEMS and automatic demand response (ADR)

The usual practice in the BEMS field was to sell maker-specific systems. As a consequence, the operation of these systems differed from building to building, and the systems themselves tended to be complicated and expensive. The end result was that systems of this type came to be deployed only in buildings of a certain scale or larger. NTT FACILITIES took this into consideration and decided to focus its efforts on providing a low-cost, highly functional building management service that separates the BEMS server function and makes it available for common use. This scheme negates the need for one company to install all system equipment in each building as long as highly reliable network facilities exist. The total service image of this system is shown in **Fig. 3**.

The deployment costs of this system are reduced, and advanced monitoring and control capabilities are provided because in each building, only the minimally required amount of equipment for operating that building (such as switches, detectors, and sensors) is installed, and the functions for monitoring and controlling the equipment (servers, centralized monitoring functions, etc.) are placed on the cloud for execution over an IP network. This is called a *cloud-based BEMS*. The conversion of BEMS to a cloud-based format has the following advantages in addition to reducing costs:

- All information can be uniformly collected and managed, enabling advanced data analysis to be carried out.
- BEMS applications and services can be easily updated or added.
- Operations for each building are the same, making building management more efficient.

Demand response (DR) is a technique for temporarily suppressing the demand for power during times of tight electric-power supply-and-demand conditions; it is implemented by notifying consumers of changes in the price of electricity and offering them incentives for reducing power usage during peak periods designated by power companies or aggregators and shifting that use to other time periods. Efficient execution of DR requires prompt response to DR requests and reduction of the operations needed for DR execution. An effective mechanism for meeting these requirements is to provide an online connection between power-companies/aggregators and consumers and to execute DR automatically.

In fiscal year 2013, NTT FACILITIES teamed up with ENNET Corporation to construct a system in

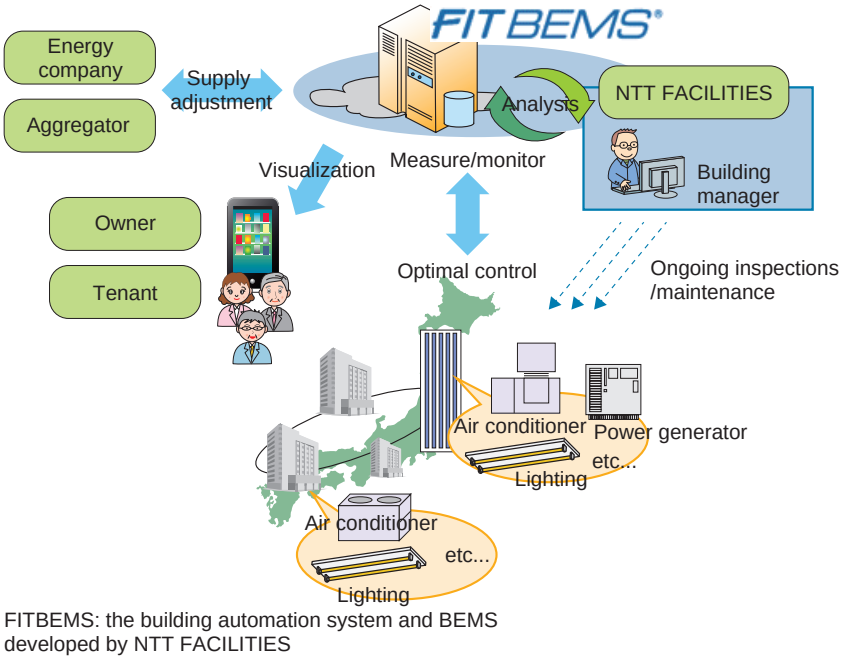


Fig. 3. Total service image of cloud-based BEMS.

which our FITBEMS® was connected online to the ENNET system (**Fig. 4**) and to hold a field trial of automatic demand response (ADR) using an actual building. This field trial was conducted to evaluate the results of prior simulations that quantified the amount of power demand that could be reduced by applying ADR and any subsequent effects on the indoor environment. The results revealed that halting the operation of air conditioners and other facilities in response to 35 ADR requests made over a one-year period reduced the power demand by 2074 kWh without affecting the indoor environment.

3.3 Composite renewable energy system

It is necessary to provide a stable power supply even during times of commercial power outage as a disaster prevention measure. Likewise, from the viewpoints of saving energy and preserving the environment, there is a need for efficient energy-use techniques that can make maximum use of the output from clean distributed generation including solar power generation. As a forerunner to such a system, composite renewable energy systems have been developed, such as the one shown in **Fig. 5**. During normal, everyday operation, a composite renewable energy system reduces or shifts peak power demand through prioritized use of distributed generation, and at the time of a commercial power outage, it shifts immediately to automatic operation mode to deal with a long-term power outage through the use of distributed generation and storage batteries. Accordingly, if composite renewable energy systems are to be linked with xEMS* in the future, not only can they provide a stable and highly reliable power supply, but they also hold the possibility of becoming a key device in smart communities as a consumer-side facility that can coordinate energy supply and demand in a building or area.

3.4 Mega-solar systems

The deployment of mega-solar systems has been expanding greatly in recent years, as they are an important source of power in smart communities. The collection of measurement data in a mega-solar system is generally done with the aim of monitoring the system's output, which fluctuates according to the weather, and checking the health of the system. In the future, we can expect mega-solar systems to be increasingly reflected in the regional coordination of energy supply and demand and used in predicting power-generation levels in the system planning stage.

* xEMS: An energy management system for electric power and gas using ICT including BEMS, HEMS, etc.

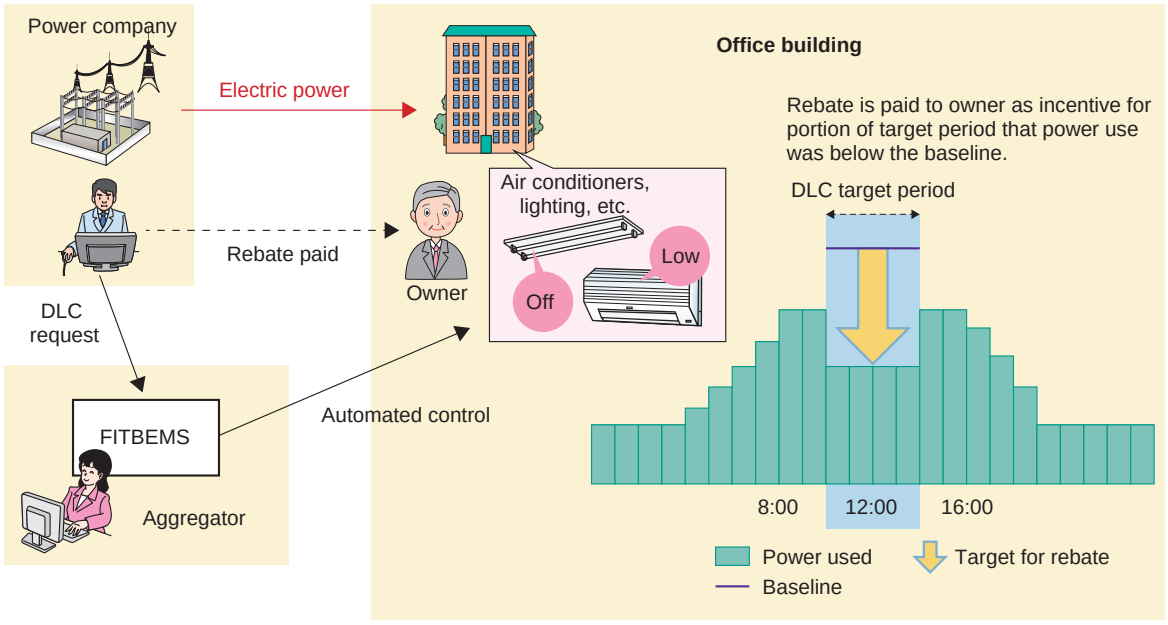


Fig. 4. Configuration of automatic demand response system.

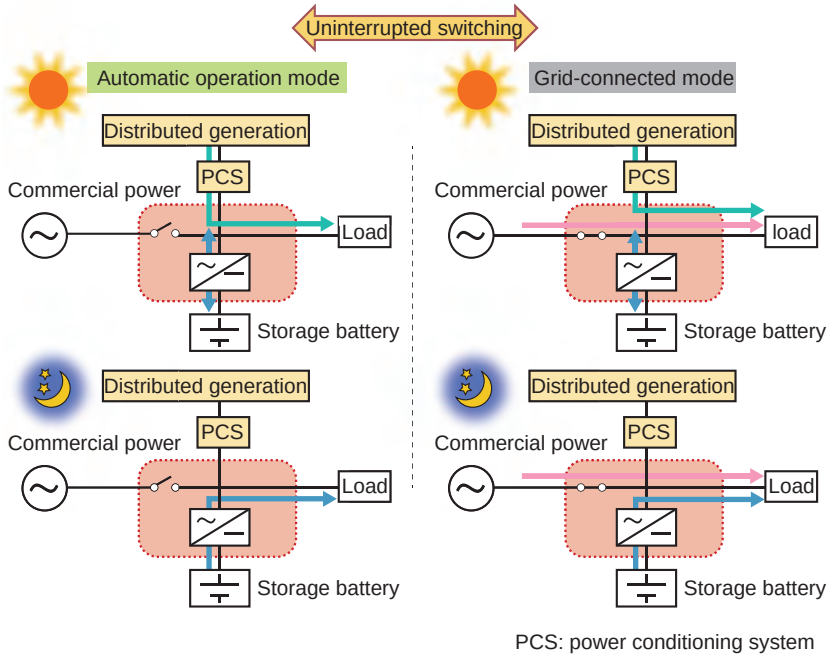
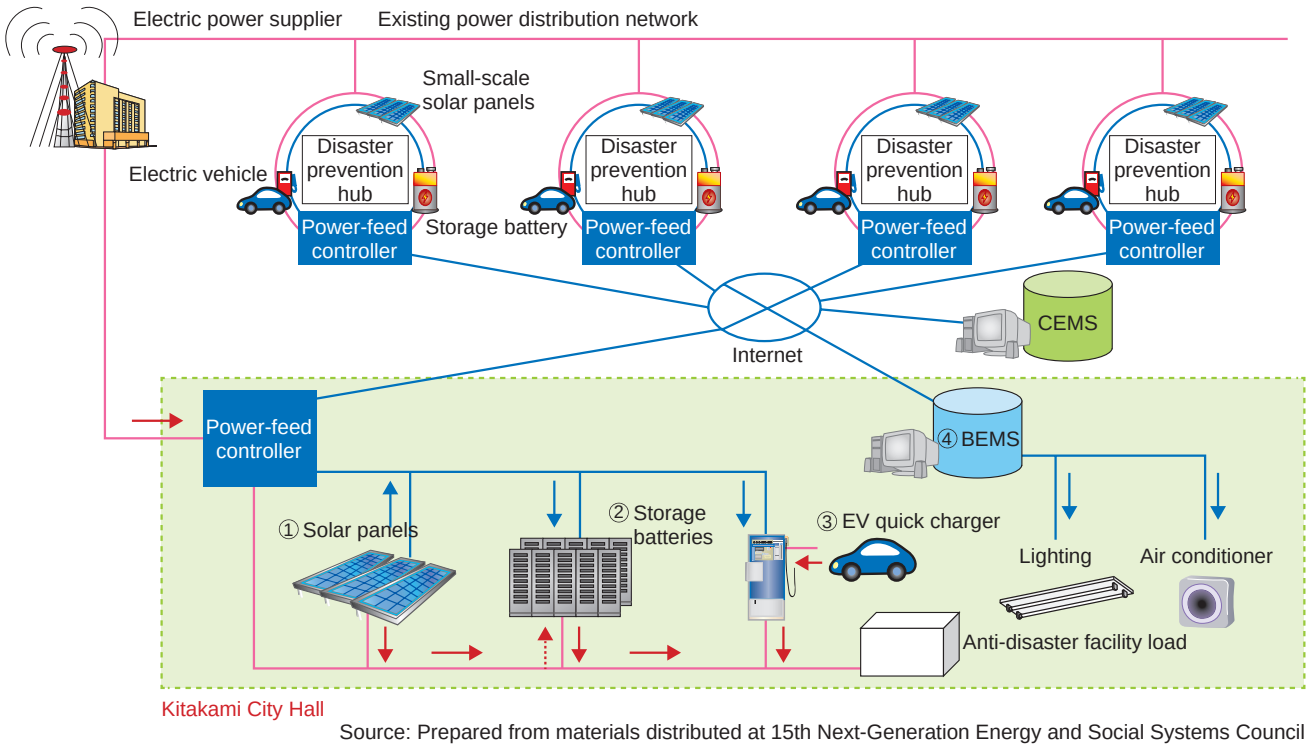


Fig. 5. Composite renewable energy system.

3.5 Cloud-based CEMS

The target of conventional energy management using ICT has been facility groups in specific areas such as BEMS in office-building areas, a mansion/condominium energy management system (MEMS) in condominium areas, and HEMS in home areas.



Source: Prepared from materials distributed at 15th Next-Generation Energy and Social Systems Council

Fig. 6. CEMS in Kitakami City Smart Community Promotion Project.

However, to achieve comprehensive energy management across an entire community, there will be a need for a community energy management system (CEMS) that can collect demand-related information across such EMSs on the consumer side and absorb information from the power-grid and distributed-generation side.

A smart community aims to promote the introduction of renewable energy, improve energy conservation through effective use of local energy, and achieve urban development that is robust against disasters, so we can expect CEMS to play a centralized role in optimizing energy supply and demand across the entire community. Furthermore, a CEMS implemented on the cloud provides a number of benefits; not only is it flexible to the addition or removal of power-consuming and power-generating facilities within the target area, but it also enables high-speed and close information exchange with other cloud-based consumer-side EMSs.

4. Case study

NTT FACILITIES is currently involved in constructing a smart community in Kitakami City, Iwate Prefecture. The purpose of this project is twofold: (1) introduce distributed sources of renewable energy in stages to existing facilities owned by the local government in order to increase the ratio of renewable energy used to power those facilities, and (2) achieve urban development that is robust against disasters (**Fig. 6**).

Within this smart community, the CEMS will be used to control the charging/discharging of storage batteries according to supply-and-demand conditions in the target area through the use of supply-and-demand management equipment. This approach will enable the provision of a high-quality, uninterruptible power supply. The CEMS will also be used to achieve load control via a BEMS. The development of this CEMS is now in progress and is scheduled to come online in fiscal year 2015.

5. Future activities

Integrated management of consumers and suppliers to achieve a safe and comfortable smart community is a pressing social problem. Looking to the future, we can expect the smart community to become an important technology domain that Japan will be able to take

pride in as a global leader. NTT FACILITIES aims for early deployment of smart communities by coordinating the advanced information technologies, energy management technologies, and facility management technologies developed to date by the NTT Group and by analyzing and visualizing various types of information to achieve appropriate system control. A transition is already taking place from the trial phase to the actual business phase.

In future research and development activities at NTT FACILITIES, we plan to incorporate more accurate forecasting techniques, develop supply-and-demand management and control techniques as well as diverse applications based on those forecasting techniques, and expand the types of facilities and equipment that can be added or connected to the system. We also plan to research and develop behavior-recognition and behavior-prediction techniques targeting community residents and building tenants using smart devices and sensor networks. NTT FACILITIES is committed to making ongoing contributions to the construction and optimal operation of smart communities.

Takeshi Watanabe
Senior Research Engineer, Advanced FM Development Division, Research and Development Headquarters, NTT FACILITIES, INC.
He received the B.S., M.S., and Ph.D. in architecture from Nagoya University in 1993, 1995, and 2011, respectively. He joined NTT FACILITIES in 1995 and engaged in designing an HAVC (hybrid automatic voltage control) system and sanitary system for buildings. Since 2002, he has been engaged in research and development on a co-generation system and BACS (building automation and control system). He is a member of various groups including ISO/TC205/WG3, the Architectural Institute of Japan, and the Society of Heating, Air-Conditioning and Sanitary Engineers (SHASE) of Japan. He is a technological fellow of SHASE.

Toyonari Shimakage
Research Engineer, Power System Development Division, Research and Development Headquarters, NTT FACILITIES, INC.
He received the B.S., M.S., and Ph.D. in electrical engineering from Nagoya University in 2001, 2003, and 2010, respectively. He joined NTT FACILITIES in 2003 and has been engaged in research and development on distributed energy supply systems interconnected with electric power systems. From 2013, he has also worked as Associate Professor of the Solutions Research Laboratory, International Research Center of Advanced Energy Systems for Sustainability, Tokyo Institute of Technology. He is a member of various groups including the Institute of Energy Economics of Japan and the Japan Solar Energy Society.

Environmentally Conscious Building Technology

Hiroyuki Tanno, Takeshi Watanabe, Tatsuro Matsuoka, Yuji Kohata, and Tatsuya Nakata

Abstract

Environmentally conscious efforts in corporate activities are a natural consequence of a society and government whose environmental awareness is growing. Such efforts are particularly important in the architecture and construction industries, whose activities and processes can have a major environmental impact. Ongoing efforts are being made to develop energy-saving designs for the diverse elements that make up buildings and to reduce carbon dioxide emissions in business operations and manufacturing processes. This article introduces the environmentally conscious activities based on advanced cross-domain technologies developed by NTT FACILITIES, which is known for its expertise in architecture and power technologies from design to operation.

Keywords: environmentally conscious design, BIM, energy saving

1. Introduction

It is commonly accepted that the architecture and construction industries can have an impact on the environment in diverse ways over the life cycle of a building, from the production of building materials to the actual construction, operation, maintenance/repair, and eventual demolition. Most greenhouse gases that give rise to global warming originate in the emission of carbon dioxide (CO_2) caused by the burning of fossil fuels, and about one third of Japan's total CO_2 emissions originate in the architecture and construction industries.

Environmentally conscious design can be summed up as an effort to reduce the load on the environment over the entire life cycle of a product, from planning and design to production, market distribution, and disposal or recycling. It has been recognized the world over as a viable approach since the 1990s, but it has recently been attracting renewed attention with the emergence of a global economy and the strengthening of environmentally conscious laws and regulations in many countries.

In Europe and the United States, for example, the move toward new and tougher regulations covering chemical substances and energy saving measures continues. Some of those regulations, moreover, have become international standards. It is consequently said that corporate enterprises cannot afford to ignore international trends in environmentally conscious design if they wish to be competitive in overseas markets.

2. GreenITy Building®: Environmentally conscious design at NTT FACILITIES

NTT FACILITIES is actively engaged in environmentally conscious design based on the concept of *GreenITy Building*. GreenITy is a coined word combining *green* and *IT* (information technology), and it signifies the integration of environmentally conscious technologies and IT (as opposed to treating them separately) and the construction of long-life buildings taking into account both environmental and user needs. We introduce some key examples of this development in this article.

NTT FACILITIES finished construction on a new research and development (R&D) center (NTT FACILITIES Shin-Ohashi Building) in July 2014. This building is intended to be a showcase of the

GreenITy Building concept. The plan is to deploy a variety of GreenITy technologies and to conduct field trials of operations.

3. Enhancement of Live-Link Design® Office

Today, environmental consciousness has become an essential approach even in the design of offices. However, the objective in office design is to enable management resources to be used as effectively as possible. To achieve this objective, an office must be designed so that office work and resource usage are efficient and intellectual productivity is improved. As a result, the corporate world needs to satisfy this objective while being environmentally conscious in order to contribute to society.

At NTT FACILITIES, we have been proposing the Live-Link Design Office for some time as an approach to designing offices that can contribute to the effective use of management resources, including power-saving measures. In this approach, we assess current conditions, analyze where office improvements can be made based on seven types of work activities (scenes) and five types of elements (schemes), propose an office plan that includes necessary improvements combined with environmentally conscious considerations, and construct the actual office (**Fig. 1**).

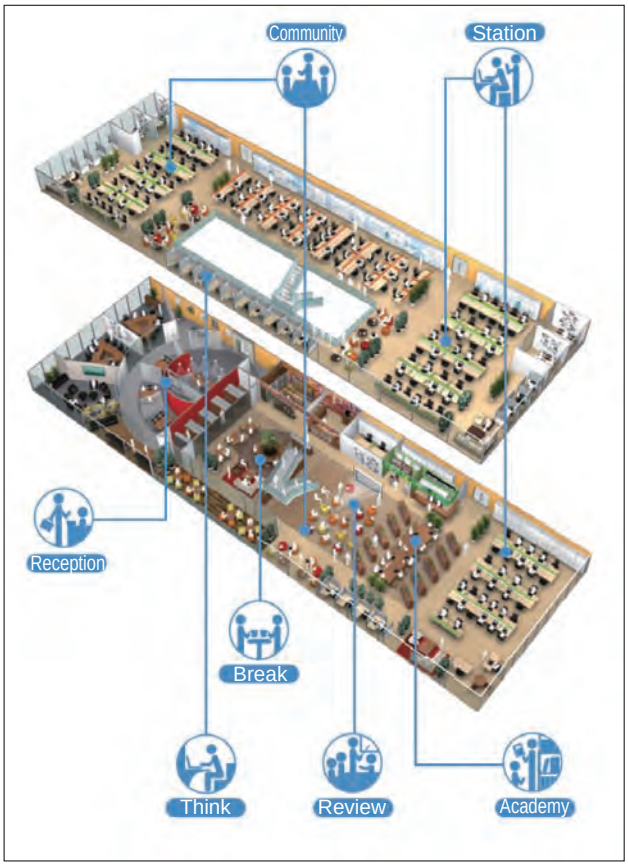


Fig. 1. Live-Link Design Office.

At present, we are working on improving the Live-Link Design Office concept even further. To give some background, it can be said that the management target in recent approaches to office design has been shifting from the way in which the workplace is used to the way in which work is performed (work style). That is, the need has arisen for an office design and management approach that enables work to be done anywhere in the office and enables workers to select where and how they want to work. It goes without saying that it is *people* who work in an office and who support corporate productivity. An increasingly popular idea is that providing a choice of office scenes to workers with individual differences so that each of them can achieve a way of working that promotes intellectual productivity can help improve the overall intellectual productivity of the company.

Being part of this trend and promoting further improvements in intellectual productivity requires focusing attention on work (labor). Specifically, worker behavior (movement, area use, tool use, etc.) needs to be analyzed and reflected in zoning and planning that stimulates communication and worker motivation; workers must be able to move freely within an office and to work as desired. Designing such a new type of office requires more than solving problems that come to light and preparing work zones that are easy to use. Worker behavior needs to be observed and analyzed so that space that enhances communication and motivation can be planned, and the effects of the office environment and worker behavior on improvements in intellectual productivity need to be visualized through simulations. This is what we are proposing with our enhanced version of Live-Link Design Office.

4. Testing the effect of linking BIM to FM

Building information modeling (BIM) is a project management technique based on virtual construction of a building on a computer using three-dimensional (3D) shape data and attribute data such as component costs and types of materials. BIM enables building design and other data to be shared among those concerned and facilitates information sharing and early

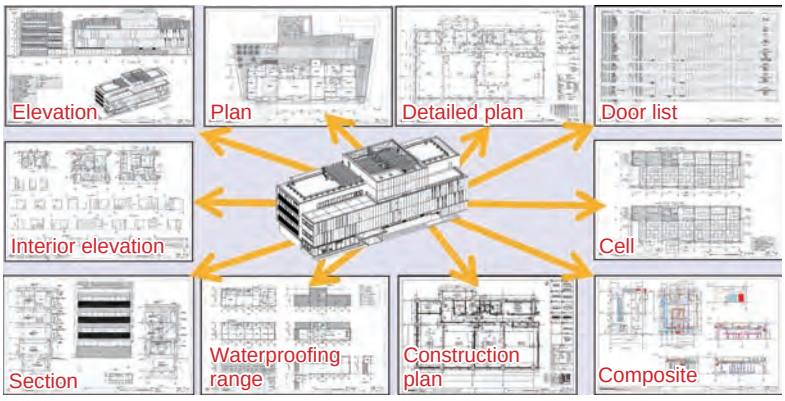


Fig. 2. Facility drawings based on BIM.

consensus building among individuals as a type of front-loading for making decisions in the early stages of a project. It makes it possible to identify problems such as drawing inconsistencies and interference between structural components and piping before the actual construction begins. It therefore prevents an escalation in building costs and project delays caused by starting over and redoing some work.

In the United States, the use of BIM began to spread in 2006, and at present, it is being increasingly applied to the life cycles of buildings from construction to operation. In Japan, meanwhile, BIM was introduced only as recently as 2009, but its implementation and use in the construction of new buildings has already entered a growth period; it is by no means rare in the industry. However, while studies have been performed from many viewpoints on the use of BIM data in facility management (FM) after construction has been completed, case studies are, for the most part, unknown. Here, using BIM data in FM is not just a matter of inheriting that data once construction has been completed. Rather, it is necessary to achieve a link between BIM and FM with a view to front-loading in FM too. NTT FACILITIES provides an integrated set of building services from design to FM and is therefore committed to introducing and effectively applying BIM. The building construction project for the new NTT FACILITIES Shin-Ohashi Building mentioned above is one of the first examples of applying BIM in Japan. The use of BIM from design and construction to FM is progressing well. In the following subsections we describe the effects of linking construction and FM through BIM.

4.1 Effectiveness of BIM in design and construction work

In this building construction project, all data were consolidated as BIM data, and about 65% of the design documents needed for construction were generated automatically using integrated BIM data. This made the process of preparing drawings much more efficient and greatly reduced the time required for checking drawings (**Fig. 2**). Specifically, a check for inconsistencies among drawings that would normally take about two weeks was performed in 3D space and completed in only one day. Another benefit is that when making revisions to drawings, the fact that all data are integrated as BIM data means that all drawings can be changed without generating inconsistencies.

Furthermore, in actual construction work, the use of 3D composite drawings is shortening the time needed to examine and check the construction process. Additionally, since BIM enables a virtual building to be completed on computer using BIM data before the real building is completed, information can be shared among concerned parties on issues such as color decisions for building components and renderings of on-site construction work in progress, and a consensus can be formed before the building is actually completed.

4.2 Effectiveness of introducing BIM in FM

This BIM feature of being able to provide a virtual building in the form of data before the building's actual completion enables a virtual building handover. In this project as well, we were able to hand over virtually completed BIM data to achieve a link with FM. We have verified that such a smooth transition of

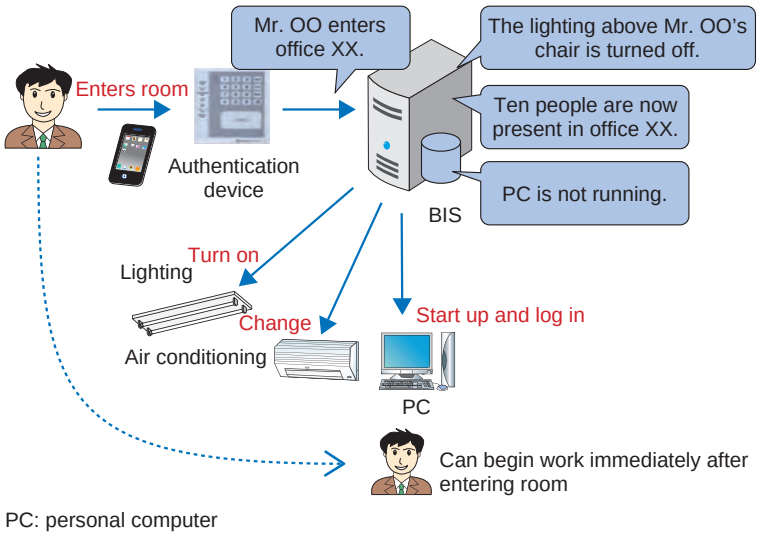


Fig. 3. Example of using BIS.

virtually completed BIM data into the FM system makes it possible to clarify the life cycle cost (LCC) before the building is completed in order to achieve both an accurate design and the appropriate maintenance and management, and to reduce the cost of maintenance and management over a long-term life cycle.

A virtual building created by BIM can be used in various types of simulations. It also has a high affinity with the field of energy management and can be used to evaluate energy-saving performance. Thus, BIM can be considered an important element in advancing environmentally conscious design.

NTT FACILITIES has been engaged in building design and FM as a business, and we have strived to achieve efficient and advanced FM by developing and introducing various types of FM tools and providing FM services linked to building design. Going forward, we plan to expand upon the knowledge gained from this building construction project to achieve a level of BIM-FM integration that can contribute to making FM even more efficient and advanced.

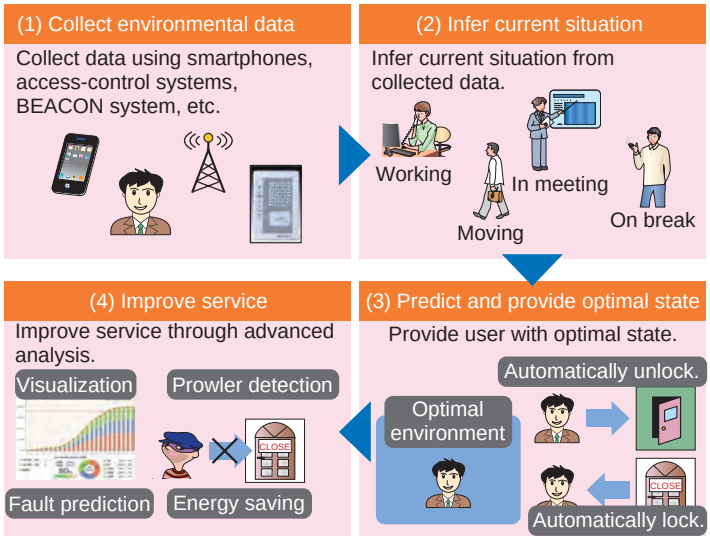
4.3 In-building information management system

A variety of systems and applications exist for achieving a safe and comfortable working environment within a building. These include security systems for managing access to the building itself as well as access to rooms inside it, a building energy management system (BEMS) for managing facilities and devices such as lighting and air conditioners, an information and communications system such as for managing internal and outside telephone lines, and more recently, wireless LAN (local area network) and thin-client systems. Such systems normally exist independently of each other in providing services to users, but such a scheme has its drawbacks too, such as:

- the same information may have to be input multiple times in different formats, and
- authentication (user name, password, etc.) may have to be performed separately, and different levels of security may exist.

A complaint that is often heard is that more advanced functions and services could be provided if such individual systems could be closely linked. One example of an advanced means of linking separate systems is the building information management system (BIS), which was formulated to interconnect diverse systems such as a building's security system, BEMS, information and communications system, and FM system. An example of applying BIS is shown in **Fig. 3**.

The act of connecting individual systems to BIS brings various types of data together. The BIS processes the data into a form that each system can use and stores the data accordingly. The system side can use the data to reduce energy consumption, enhance security, improve work styles, and raise productivity. In addition, BIS can be made to support the information-communication specifications of each system so that existing systems can be used in their present



BEACON: Short-range radio communication

Fig. 4. Role targeted by BIS.

configuration. This approach enables BIS to be deployed in a building in a relatively short time at low cost. The role that BIS aims to perform is shown in **Fig. 4**. At NTT FACILITIES, we aim to achieve the following by introducing BIS:

(1) Control building facilities according to the number of people in each room (achieve 35% energy savings)
(2) Change security level according to situation (provide stress-free environment, improve information security)
(3) Provide an optimal room environment (raise productivity)

5. Applying unused energy to air conditioning

Air conditioning systems account for a large percentage of the power consumed by all building facilities, and thus, there is a great need for energy-saving measures in this area. One effective technique is to make positive use of unused energy sources such as by using outdoor air or recovering heat that has traditionally been discarded as waste.

NTT FACILITIES is developing technologies for systems that can maximize the use of unused energy as air conditioning energy with the aim of achieving environmentally conscious buildings. The following introduces some examples of this system development.

5.1 Incorporating unused energy in buildings

We have developed a heat-interchange control system for maximizing the use of unused energy in the form of ground-source heat at about 17°C and server exhaust heat at about 35°C (**Fig. 5**).

This system features the following mechanisms:

(1) Uses exhaust heat recovered from a server by a heat exchanger for office heating
(2) Uses cold energy extracted from a ground pile for cooling in server rooms or offices
(3) Uses cold outdoor air to cool rooms depending on the indoor environment

The system can automatically select any of the above mechanisms to configure a heat-interchange pattern that saves the most energy and maximizes the use of unused energy throughout the building. Taking full advantage of temperature differences that exist within a building as well as inside and outside a building and making positive use of heat in this way can reduce the amount of energy used by the building's air conditioning system.

This new technology will be deployed at the new NTT FACILITIES Shin-Ohashi Building. Trial calculations indicate that this technology will provide an annual power-consumption reduction (energy-saving) effect of 31% relative to no deployment.

6. New approaches to office air conditioning

Office air conditioning must provide both an

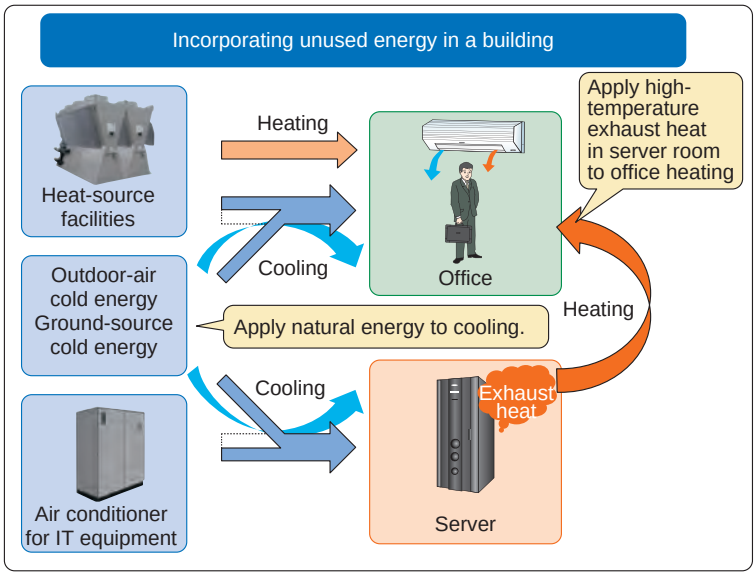


Fig. 5. Conceptual diagram of heat-interchange control system.

energy-saving effect and a comfortable environment at the same time. However, these are often conflicting requirements, so to achieve good performance in both of these categories, we have undertaken the development of advanced office air conditioning systems. Some of the technologies that we plan to use in these systems are introduced in this section.

6.1 Membrane-radiation air conditioning

In this type of air conditioning, a membrane consisting of fibrous material is stretched across the ceiling and used as an air outlet. Driving air through this membrane at low speed can suppress airflows that feel uncomfortable to occupants of the room. Cooling or warming the membrane itself can also improve heat radiation in a room and provide a more comfortable environment. Improving the heat-radiation environment in this way can help relax the room's temperature setting and contribute to energy savings. Key features of this system include the capability to obtain a strong radiation effect with the larger area of the membrane and flexibility regarding the shape of the membrane.

6.2 Task ambient air conditioning

Task ambient air conditioning consists of a system that provides both ambient air conditioning, which targets the entire space in question, and task air conditioning, which targets partial areas. In a conventional office, air conditioning is uniform within a room even though the occupants of the room have various individual preferences. This new technology, however, includes an arrangement of personalized air vents for task air conditioning to allow for individual preferences in addition to ambient air conditioning using membranes. The user can control the amount of air flowing through his or her personalized air vent from a smartphone or computer, and the end result should be a more comfortable experience tailored to personal preferences. Being able to adjust the airflow in an occupied area can likewise help relax the room's temperature setting and contribute to energy savings.

6.3 Air conditioning by separate sensible and latent cooling

A typical air conditioning system will simultaneously adjust temperature and humidity, which means that a temperature that is lower than the blowout temperature must be set in order to dehumidify the room when cooling. In contrast, an air conditioning system based on separate sensible and latent cooling has an air conditioner for adjusting temperature and an air conditioner for adjusting humidity so that the temperature and humidity can be independently controlled. This scheme makes it unnecessary for the temperature-adjusting air conditioner to cool the room to a temperature lower than the blowout temperature. It also enables ground-source cold energy to be used as energy for air conditioning, thereby contributing to energy savings. Furthermore, it enables

temperature and humidity to be controlled in a much finer manner than that possible in current systems, thereby improving the level of comfort.

This system can reduce the energy required for air conditioning by 20% compared with ordinary systems without sacrificing comfort. In short, it can achieve both an energy-saving effect and a comfortable environment at the same time. The plan is to deploy this technology at the new NTT FACILITIES Shin-Ohashi Building and to perform ongoing testing while using visualization techniques to assess its effects.

7. Future developments

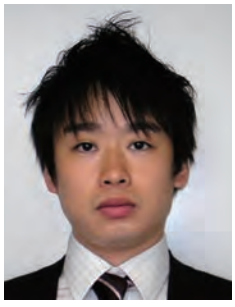
In addition to the technologies introduced in this article, NTT FACILITIES is examining and assessing a wide range of technologies that can contribute to the construction of environmentally conscious buildings in combination with energy technologies. At NTT FACILITIES, we are acutely aware that one of our roles is to firmly establish each of these technologies and to effectively integrate them to implement the GreenITy Building concept based on environmentally conscious design intertwined with safety, reliability, energy conservation, and flexibility. In this way, we seek to contribute to the business operations of the environmentally conscious NTT Group in the fields of architecture and energy.

Hiroyuki Tanno
Senior Research Engineer, Research and Development Headquarters, NTT FACILITIES, INC.
He received the B.E. and M.E. in architecture from Tokyo University of Science in 1991 and 1993, respectively. He joined NTT FACILITIES in 1993 and moved to R&D Headquarters in 2013. His research interests include design and construction techniques of communication buildings and evaluation of the working style of people in offices. He is a member of the Architectural Institute of Japan (AIJ), Japan Facility Management Association (JFMA).

Takeshi Watanabe
Senior Research Engineer, Advanced FM Development Division, Research and Development Headquarters, NTT FACILITIES, INC.
He received the B.S., M.S., and Ph.D. in architecture from Nagoya University in 1993, 1995, and 2011, respectively. He joined NTT FACILITIES in 1995 and engaged in designing an HAVC (hybrid automatic voltage control) system and sanitary system for buildings. Since 2002, he has been engaged in research and development on a co-generation system and BACS (building automation and control system). He is a member of various groups including ISO/TC205/WG3, AIJ, and the Society of Heating, Air-Conditioning and Sanitary Engineers (SHASE) of Japan. He is a technological fellow of SHASE.

Tatsuro Matsuoka
Senior Research Engineer, Research and Development Headquarters, NTT FACILITIES, INC.
He received the M.E. in intelligent engineering from Nihon University, Tokyo, in 1989. He joined NTT in 1989 and moved to NTT FACILITIES in 1994. His team is responsible for information management, database construction and operation, and R&D of FM systems. Since 2008, his team has been researching and studying techniques for using building information modeling (BIM) in FM. He is a member of AIJ. He is licensed as an Architect of the First Class, a Professional Engineer (for computer science sector), and a Certified Facility Manager Japan (CFMJ).

Yuji Kohata
Research Engineer, Research and Development Headquarters, NTT FACILITIES, INC.
He received the M.E. in intelligent engineering from Hokkaido University in 2007. He joined NTT FACILITIES in 2007 and moved to the Research and Development Headquarters in 2012. His research interests include air conditioning systems for datacenters.



Tatsuya Nakata
Staff member, Research and Development Headquarters, NTT FACILITIES, INC.
He received the B.E. and M.E. in architectonics from Waseda University, Tokyo, in 2008 and 2010, respectively. He joined NTT FACILITIES in 2010. He was engaged in the design of sanitation and air conditioning of the building until 2012. He is now engaged in developing air conditioning systems for telecom central offices and datacenters. He is a member of SHASE and AIJ.

Engineering for Earthquake Disaster Prevention Based on Observations, Experiments, and Analyses

Daisuke Chiba, Ken'ichi Yoshida, and Mikio Suzuki

Abstract

NTT FACILITIES has established three unique earthquake resistant techniques used for observations, experiments, and analyses in order to provide a reliable telecommunications infrastructure. We aim to realize a resilient society and are therefore working on developing new technology for earthquake disaster prevention by further advancing and linking these existing techniques. We introduce in this article an overview of these techniques and also describe the concept of structural health monitoring that has been a recent focus of development.

Keywords: earthquake disaster prevention, resilience, structural health monitoring

1. Introduction

NTT FACILITIES has been developing unique earthquake resistant techniques for buildings and equipment that constitute the telecommunications infrastructure. These techniques have been continuously developed since the days of the Nippon Telegraph and Telephone Public Corporation. Because Japan is a country with considerable seismic activity, NTT FACILITIES has established three methods of seismic observation, vibration testing, and structural analysis. These techniques have also been improved over time in order to provide and operate a strong and reliable telecommunications infrastructure. We are working on not only improving the technology for the telecommunications infrastructure, but also developing society-wide earthquake disaster prevention technology by further advancing and linking the three techniques. Major earthquakes are still a primary concern to society as a whole, so one of the missions of NTT FACILITIES is to realize a resilient society through development and improvement of earthquake disaster prevention technology (**Fig. 1**).

2. Observation of seismic shaking in buildings

The development of earthquake resistant engineering originated from a desire to learn as much as possible from seismic damage that occurred in the past. Damage factors need to be clarified scientifically in order to learn from such damage, and seismic observation serves as a foundation for this. Seismic observation involves the use of a measurement instrument called a strong-motion seismograph. This instrument is fixed on the ground or on the floor surface of a building, and it records the shaking caused by earthquakes. Strong-motion seismographs to observe only ground motion have been deployed by the national and local governments in many places [1]. In contrast, the number of seismographs to observe building shaking by earthquakes is much smaller, although it is gradually increasing.

The NTT Group has installed its own seismographs to measure the shaking of buildings and has been conducting seismic observations in NTT buildings nationwide. The history of seismic observation in the NTT Group goes back more than half a century, when two seismographs were installed at the Koto Telephone Office Building (Tokyo) in 1960. Since then, seismic observation systems have been installed in

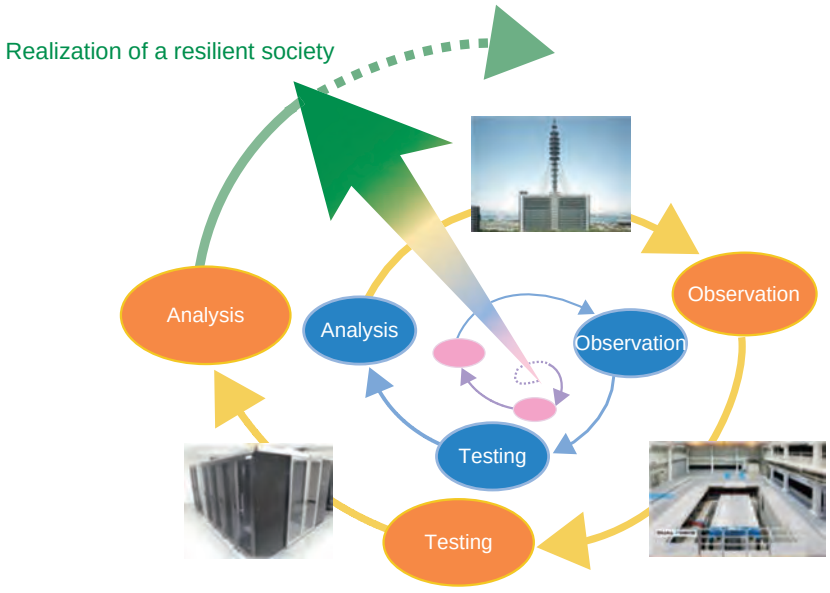


Fig. 1. Conceptual diagram of development of earthquake disaster prevention technology at NTT FACILITIES.

various types of buildings such as high-rise buildings and base-isolated buildings (buildings whose bases are not fixed to the ground). Seismic observation is now being conducted in 71 buildings nationwide (**Fig. 2**). During the 2011 Off the Pacific Coast of Tohoku Earthquake* (main shock), strong shaking motions of buildings caused by the earthquake were recorded at observation points in Sendai and other areas across the country.

NTT FACILITIES has determined how buildings shake during an earthquake by analyzing the strong-motion records obtained from the observation points over the years and has utilized this information in evaluating the seismic performance of telecommunications equipment installed in the buildings. In particular, in the 1995 Hyogo-ken Nanbu Earthquake, valuable strong-motion records were obtained at the NTT Kobe Ekimae Building (**Fig. 3**). At that time, there were only a few cases of seismic observation of building motion in Japan; consequently, the successful acquisition of strong-motion records of an entire building in the disaster area in Kobe attracted a lot of attention in academia. The seismic records were reflected in the document "Seismic Testing Method for Telecommunications Equipment," which is now a standard for evaluating the seismic performance of telecommunications equipment installed in NTT buildings [2], and these records have contributed to improving the reliability of NTT's telecommunications services. Furthermore, the seismic records have contributed to the development of seismic technology in Japan as data used to understand the dynamic characteristics of buildings during large earthquakes, and have also been utilized as seismic waveforms for structural design such as high-rise buildings and base-isolated buildings.

3. Vibration testing to evaluate seismic performance

Vibration testing is an effective means of understanding the dynamic characteristics of telecommunications equipment in a real setting during an earthquake and for confirming the seismic performance of equipment, for example, whether physical damage or functional failures occur in the equipment. Moreover, it is also effective for examining whether developed seismic countermeasures are truly efficient and whether any unexpected situations occur when they are applied.

Telecommunications equipment installed in NTT buildings must satisfy a prescribed level of seismic performance. The seismic performance is evaluated using vibration testing with a shaking table in

* The name designated by the Japan Meteorological Agency that refers to the main earthquake shock. On April 1, 2011, the Japanese Government officially named the disaster (including the earthquake and tsunami) "The Great East Japan Earthquake."

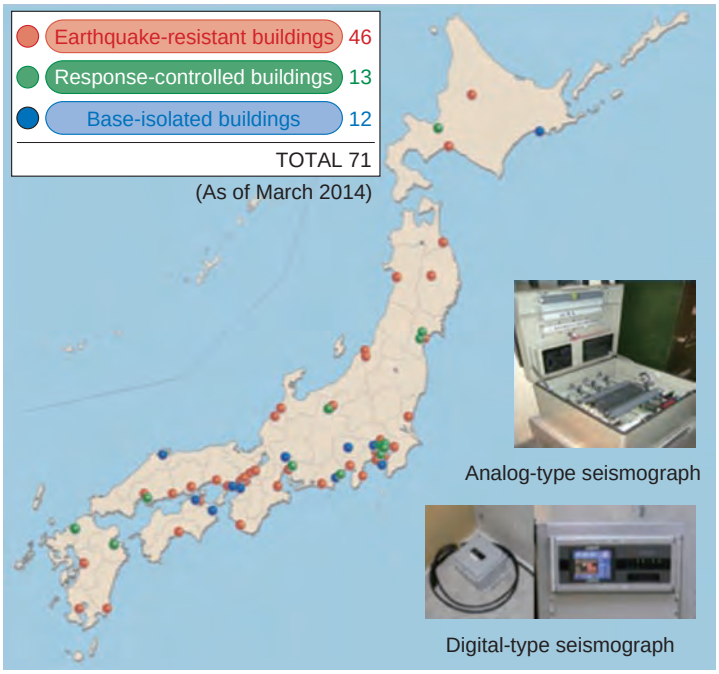


Fig. 2. Seismic observation points of NTT Group.

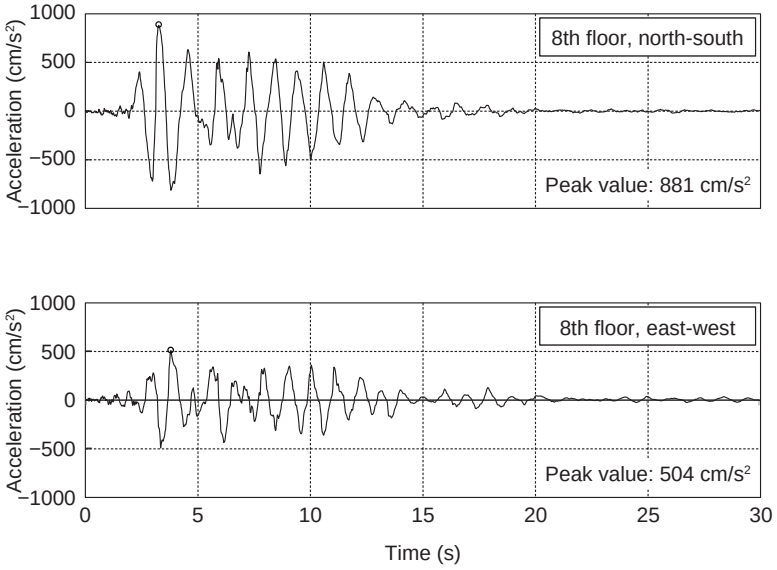
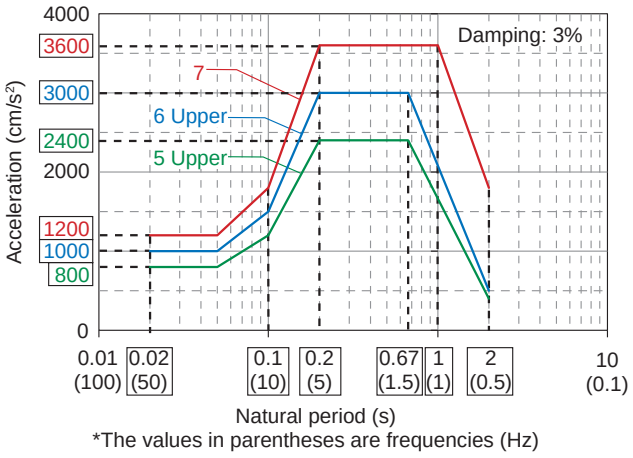


Fig. 3. Strong-motion records from NTT Kobe Ekimae Building obtained during the 1995 Hyogo-ken Nanbu Earthquake.

accordance with the standard described previously [2]. In the vibration testing, actual equipment is installed on the shaking table, and the equipment is shaken by reproducing simulated seismic waveforms with periodic characteristics, as shown in **Fig. 4(a)**. This makes it possible to determine whether the equipment has been installed adequately and whether its performance satisfies the criteria as shown in



(a) Periodic characteristics of simulated waveforms

Seismic scale by the JMA	Physical damage	Functional failure
5 Upper	No damage to main structural parts.	No functional failures
6 Upper	- No permanent deformation in main structural parts - No cracks or ruptures in welded parts - No falling off, detachment, or damage to any housed devices - Covers on housed devices do not fall off; door locks on cabinets remain securely closed. - Maximum horizontal deformation at the top: 50 mm or less	Temporary functional failure only during testing (Recovery to normal functions without requiring part replacement or manual restart after testing)
7	- No fatal damage to main structural parts - No falling off, detachment, or damage to any housed devices - Covers on housed devices do not fall off; door locks on cabinets remain securely closed.	

(b) Acceptance criteria for seismic performance

JMA: Japan Meteorological Agency

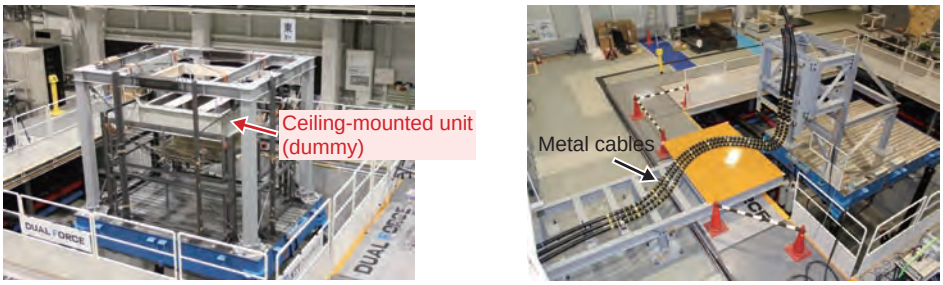
Fig. 4. Outline of NTT's "Seismic Testing Method for Telecommunications Equipment."

Fig. 4(b). The characteristics of the simulated waveforms shown in Fig. 4(a) are defined on the basis of technical knowledge obtained by analyzing strong-motion records, which reflect the dynamic characteristics of telecommunications buildings of NTT [3, 4].

The first shaking table was introduced in the NTT Group in 1971, and ever since then, the seismic performance of telecommunications equipment has been evaluated in order to ensure the quality of NTT's telecommunications services. NTT FACILITIES introduced its newest shaking table in 2010, a three-dimensional (3D) vibration testing system called DUAL FORCE [5]. The size of the shaking table is 4 m × 3 m, and it is possible to reproduce a broad range of short- to long-period vibrations. In addition, since the shaking table can move horizontally up to 1.1 m (zero to peak), it can reproduce vibration phenomena with large displacement such as shaking on an upper floor of a high-rise building subjected to long-period ground motion. It is utilized to evaluate the performance of equipment and fixtures installed in various environments and to develop seismic countermeasures.

We describe here some recent examples of tests conducted using DUAL FORCE. A photograph of a test on a ceiling-mounted air-conditioning unit is shown in **Fig. 5(a)**. This test was conducted to investigate what caused the unit to fall in the 2011 Great East Japan Earthquake and to verify countermeasures to prevent it from falling. A photograph of a test done to verify the flexibility of metal cables drawn into a base-isolated telecommunications building is shown in **Fig. 5(b)**.

When conducting vibration tests where various environments are simulated, we utilize not only seismic records obtained by observation, but also techniques for simulating the behaviors of various structures obtained by structural analysis. The combination of observation and simulation expands the range of techniques that can be used to evaluate seismic performance by vibration testing.



(a) Ceiling-mounted air-conditioning unit (b) Metal cables drawn into base-isolated building

Fig. 5. Images of recent tests conducted using DUAL FORCE.

4. Simulation of seismic behavior of structures by structural analysis

In order to construct a reliable telecommunications infrastructure that can withstand a severe seismic risk, it would be ideal, and perhaps even technically possible, to design building structures that would not be damaged at all during the largest imaginable earthquakes. However, this would require an excessive investment, which is not always desirable in practice. It would be more realistic to set a seismic level to aim for zero damage and to design structures to allow some minimal damage from earthquakes beyond the set level. However, a design that allowed some damage would require a dynamic analysis program that could simulate precisely the process whereby a structural member of a building changes from elastic to plastic under a seismic load. NTT FACILITIES developed DIAS (Dynamic Inelastic Analysis System) in 1991 before spinning-off from NTT. Utilizing DIAS makes it possible to precisely simulate the 3D behaviors of structures in the event of dynamic and non-stationary disturbances such as earthquakes, and the results can be checked using animation (**Fig. 6**).

Seismographs have been installed in four places in a high-rise building of the NTT Group in Sendai, both inside the building and at the top of its tower. When the main shock of the 2011 Off the Pacific Coast of Tohoku Earthquake occurred, strong motion was recorded at each observation point on the building. A comparison of the observed motion recorded at the top of the tower during the earthquake and the result of the response analysis by DIAS indicated that they were well matched (Fig. 6). This shows that DIAS can simulate very precisely the extent of damage that might occur in buildings [6].

5. Health monitoring for structures

When building structures are damaged in a large earthquake, a technique known as structural health monitoring is used to examine the state of damage and to evaluate the necessity of reinforcing the structures or taking further action. Building structures usually have various materials covering them on both the interior (e.g., sheetrock, wallpaper, ceiling boards) and exterior (e.g., curtain walls, tiles), and the parts that can be inspected directly are limited. Therefore, it is necessary to remove the materials in order to visually confirm the state of the structural elements of an entire building. However, this is rarely done in view of the prohibitive cost and time required for removing the materials and re-applying them after.

Structural health monitoring is a technique for identifying which components in a structure have been damaged or degraded by an earthquake, grasping the degree of damage, and evaluating the health of the structure without having to visually check it, by installing sensors such as seismographs in the structures and analyzing the measured data. Structural health monitoring thus involves analyzing and examining a structure using data obtained by sensors and may therefore be described as a technique that was developed from seismic observation. There are two key elements in this technique: the sensor and the damage detection method. It is necessary to accurately identify the damaged elements of a structure and the degree of damage by using an efficient detection method with the minimum necessary number of low-cost sensors.

The damage predictions of large earthquakes occurring along the Nankai Trough and under the Tokyo metropolitan area were reviewed after the Great East Japan Earthquake. The most serious results were announced, which consequently increased the

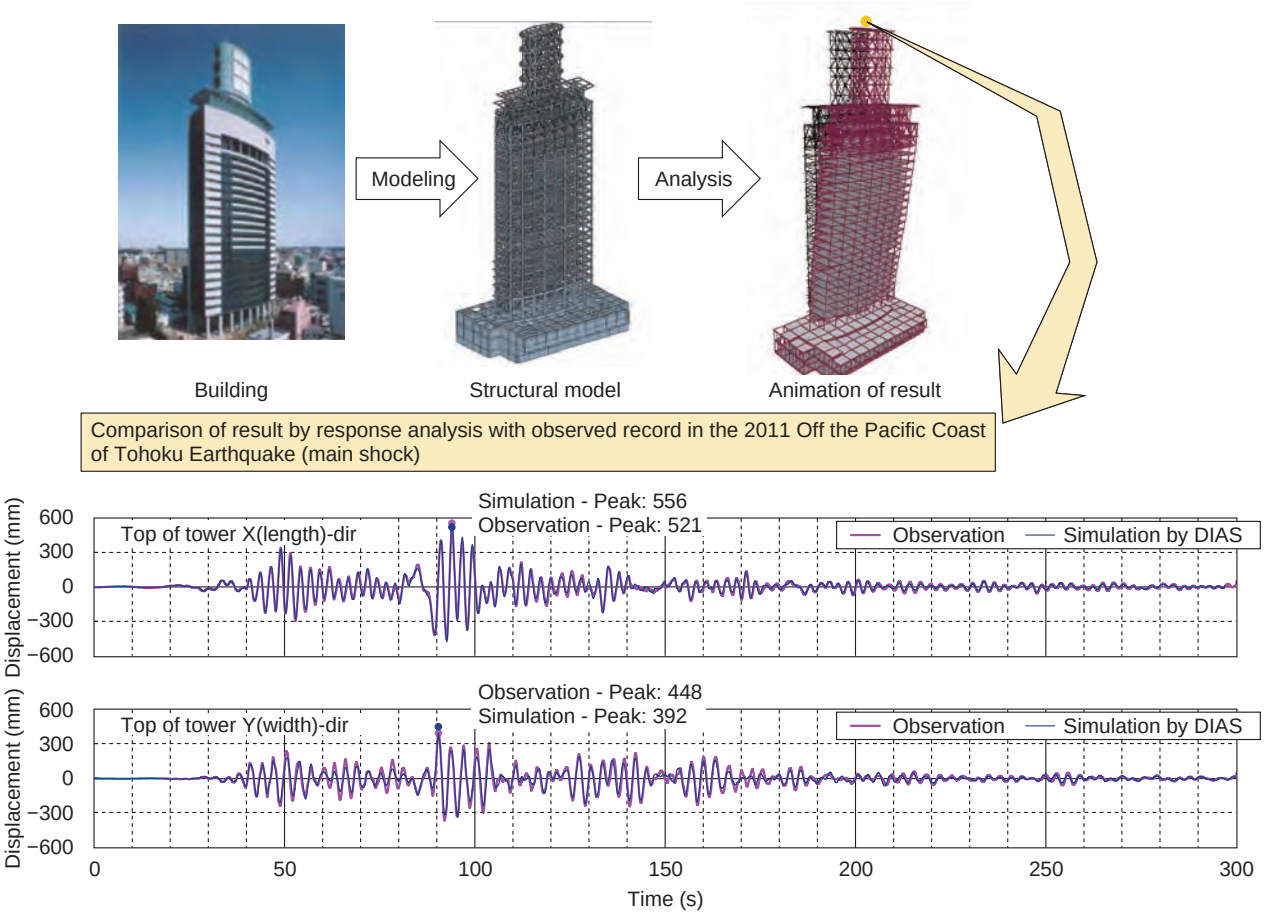


Fig. 6. Example of high-accuracy simulation by DIAS.

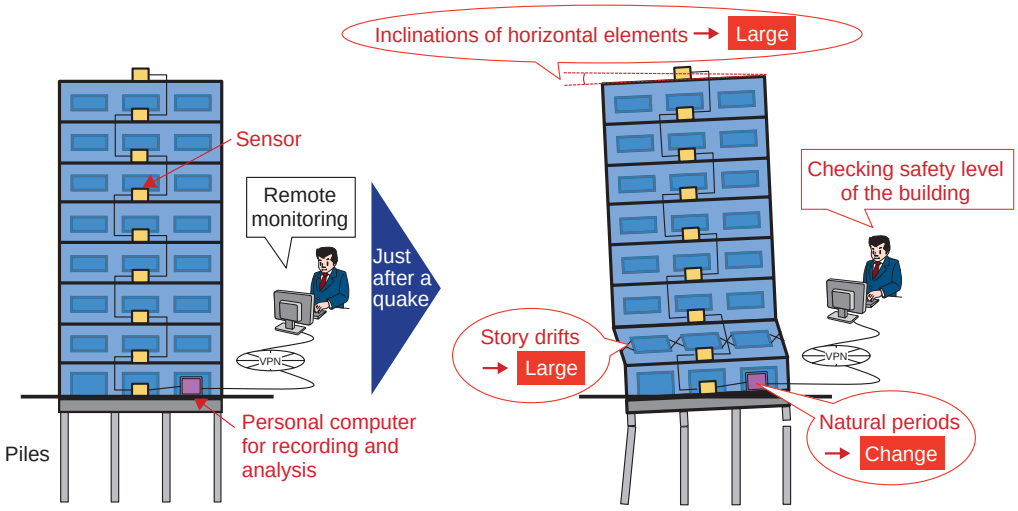
demand for a system that can reliably determine whether a building will be able to be used safely immediately after an earthquake.

NTT FACILITIES developed a low-cost sensor and a method for evaluating the safety level of a building that can be applied for structural health monitoring and also started selling the Yure-Moni® support system to determine building safety levels on October 1, 2013 (**Fig. 7**). Conventionally, only story drifts (inclinations of vertical elements) has been used as an index to evaluate the damage state of a building. In contrast, Yure-Moni uses three indices: inclinations of horizontal elements, the natural periods of a building, and story drifts, to evaluate the damage state of a building. Checking the damage state from various aspects in this way makes it possible to provide more accurate and precise information about the safety level of a building just after an earthquake, and even building managers who are not professional engineers can make a proper determination.

Currently, in order to enhance the functions of Yure-Moni as a tool for disaster control before and after an earthquake, we are focusing on adding earthquake early warning functions and unified building management functions to it (**Fig. 8**). Furthermore, in the future, we are planning to increase the precision of damage detection methods using structural analysis by DIAS and to expand the field of applications of Yure-Moni to civil engineering structures such as road bridges.

6. The present and future of earthquake disaster prevention

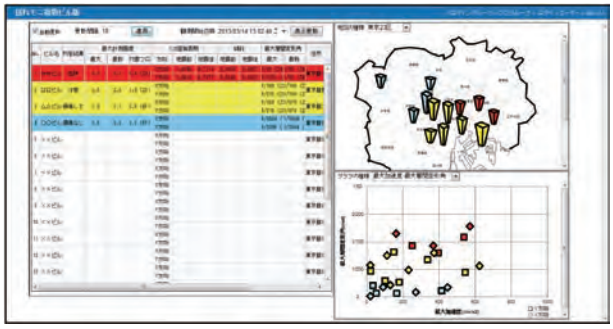
The Tohoku region has experienced many large earthquakes, including the 1978 Miyagi-ken-oki Earthquake and the 2008 Iwate-Miyagi Nairiku Earthquake. More than three years have passed since the Great East Japan Earthquake occurred. The number of buildings that collapsed or were structurally



The use of three indices (inclinations of horizontal elements, natural periods of a building, and story drifts), makes it possible to provide more precise information about the safety level of a building.

Fig. 7. Yure-Moni support system to determine building safety level.

(a) Example of screen display when receiving an earthquake early warning



(b) Example of screen display for unified management of buildings

Fig. 8. Yure-Moni as a tool for disaster control.

damaged in that earthquake was small and limited. However, examples of damage caused by the collapse and fall of nonstructural elements in buildings far from that earthquake's epicenter were confirmed over a wide area from the Tohoku region to the Kinki region. In the Tokyo metropolitan area in particular, damage resulting in human fatalities occurred because of a collapsing ceiling. NTT FACILITIES is working to prepare for a major earthquake that will undoubtedly occur in the near future by developing countermeasures to problems revealed by the Great East Japan Earthquake. We will continue to develop new technology for earthquake disaster prevention in order to realize resilient societies in other countries where earthquakes occur as frequently as they do in Japan.

References

[1] "Seismic observation points," Japan Meteorological Agency (in Japanese), http://www.data.jma.go.jp/svd/eqev/data/seis_st/

[2] "Seismic Testing Method for Telecommunications Equipment," NTT (in Japanese), http://www.ntt-east.co.jp/info-st/info_dsl/taisin.pdf

[3] Y. Sato, T. Fuse, and H. Akagi, "Building Appendage Seismic Design Force Based on Observed Floor Response," Proc. of 8th World Conference on Earthquake Engineering (WCEE), pp. 1167–1174, San Francisco, CA, USA, 1984.

[4] S. Tsukuda, K. Hiramatsu, T. Ozawa, and K. Okuta, "Evaluation of Seismic Wave for Designing Equipment Installed in Buildings," Proc. of 13th WCEE, No. 2435, Vancouver, Canada, 2004.

[5] K. Yoshida, S. Nagashima, T. Ema, and H. Nakamura, "Engineering of NTT FACILITIES Supporting BCP," NTT Technical Journal, Vol. 22, No. 8, pp. 38–41, 2010 (in Japanese).

[6] T. Motohi, Y. Sugimura, W. Goto, M. Suzuki, S. Nagashima, and K. Saito, "Evaluation of Earthquake Response of High-rise Building with Dampers in the 2011 Off the Pacific Coast of Tohoku Earthquake – Part 2: Dynamic Analysis by Observed Strong Motion," Summaries of Technical Papers of Annual Meeting of the Architectural Institute of Japan, Vol. B-2, pp. 317–318, 2011 (in Japanese).

Daisuke Chiba

Engineer, Risk Management System Development Division, Research and Development Headquarters, NTT FACILITIES, INC.

He received the B.E. and M.E. in architecture from Tohoku University, Miyagi, in 2003 and 2005, respectively. He joined NTT FACILITIES in 2005. He studied evaluation methods of seismic performance for various equipment installed in buildings and was involved with improving seismic observations in the NTT Group and constructing DUAL FORCE, the 4th 3D vibration testing system. He was also engaged in various architectural projects as a structural engineer in the design division. He is currently developing the functions of Yure-Moni as a tool for earthquake disaster preparedness and expanding its fields of application. He is qualified as a 1st-class Architect and is a member of the Architectural Institute of Japan (AIJ).

Ken'ichi Yoshida

Senior Research Engineer, Risk Management System Development Division, Research and Development Headquarters, NTT FACILITIES, INC.

He received the B.E. in architecture from Fukuoka University in 1991 and the M.E. in architecture from Tokyo Institute of Technology in 1993. He joined NTT FACILITIES in 1993. He was engaged in many architectural projects as a structural engineer in the design division. His main duties involve research and development of earthquake-resistant structures. He is also studying structural health monitoring, and is the manager of the development projects on Yure-Moni. He is qualified as a 1st-class Architect and as an APEC Engineer in structural engineering. He is a member of AIJ.

Mikio Suzuki

Executive Manager, Risk Management Systems Development Division, Research and Development Headquarters, NTT FACILITIES, INC.

He received the B.E. in architecture from Tokyo Metropolitan University in 1985 and the M.E. in architecture from the University of Tokyo in 1989. He joined NTT in 1989 and has been engaged in researching seismic technology. He has worked on the structural design of many important telecommunications buildings as the senior structural engineer in the architectural design division of NTT and NTT FACILITIES. He is qualified as a 1st-class Architect and an APEC Engineer in structural engineering. He is a member of AIJ.

Application Technology for Buildings and Facilities Data

Masahiro Yamazaki, Tatsuro Matsuoka, Tomonobu Tsujikawa, and Sohei Nakashima

Abstract

NTT FACILITIES has accumulated a massive amount of data through the longstanding implementation of planning, designing, construction, and maintenance management services for telecommunications power systems, air conditioning systems, and buildings. In recent years, we have been developing technologies for applying this data to establish faster and more efficient ways of responding to malfunctions, deterioration, and fault prediction of buildings and facilities. In this article, we describe how we intend to use this collection of data on buildings and facilities to strengthen our services.

Keywords: power supply system, buildings and facilities, big data

1. Introduction

In maintaining the information and communication technology (ICT) infrastructure, NTT FACILITIES strives to fulfill the critical mission to "Never let telecommunications be interrupted" to our customers as well as NTT Group companies. This maintenance process involves managing approximately 20,000 telecommunications and office buildings at around 12,000 locations all over Japan (as of March 2014), and conducting monitoring and maintenance of telecommunications power systems and air conditioning systems amounting to approximately 200,000 devices. To accomplish these tasks smoothly and reliably, we have been developing and operating maintenance support and remote monitoring systems such as the Building Emergency Call service and an integrated maintenance system for facility operation. These systems are helping to improve the reliability of ICT services by making it possible to quickly implement the appropriate maintenance services (**Fig. 1**).

NTT FACILITIES has been accumulating a massive amount of data on buildings and facilities over the years via these maintenance and monitoring systems and has been utilizing the data for various building maintenance and facility management services.

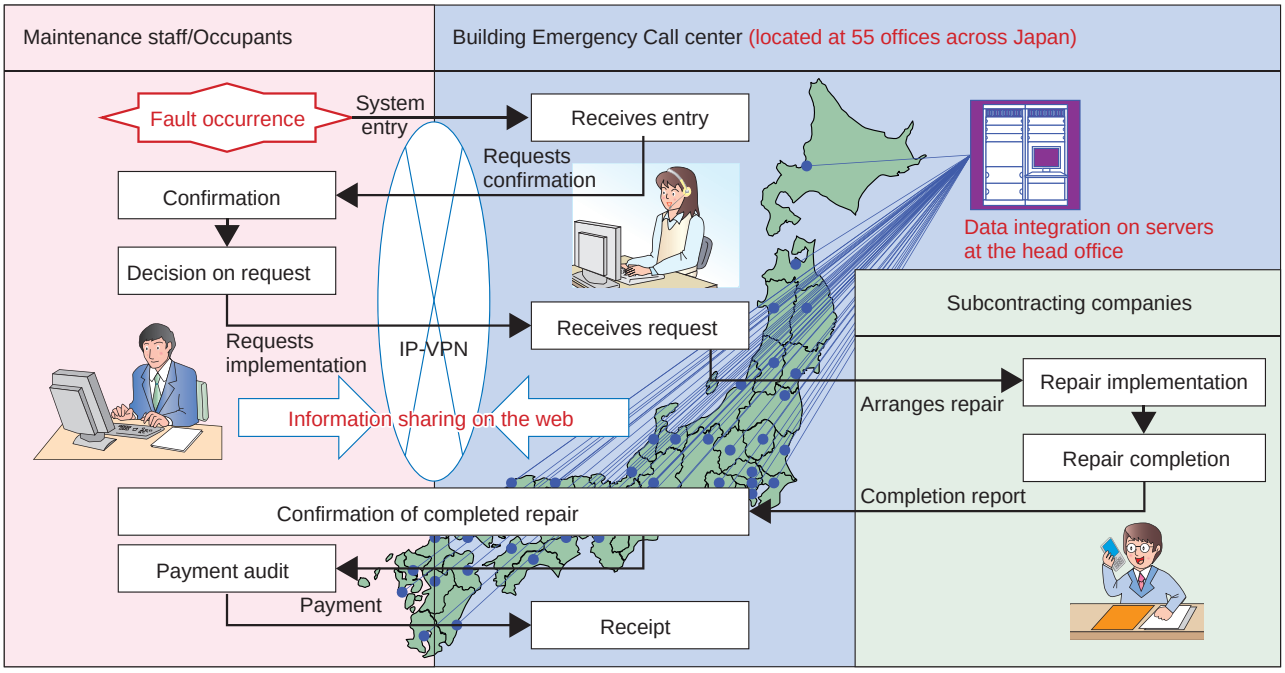
In recent years, big data technology has been attracting a lot of attention, and this technology is making it possible to take advantage of the vast amount of varied information that was once difficult to deal with. Currently, NTT FACILITIES is also working on ways to generate new added value by taking advantage of the accumulated data. As part of these efforts, we would like to introduce our new approaches to applying the accumulated data in building maintenance management and power supply systems for telecommunications. These approaches will be useful for monitoring and maintaining buildings and facilities nationwide.

2. Application of building maintenance management data

2.1 Versatility and potential of building management data

In addition to the vast amount of data that has been acquired, building data also includes a wide variety of figures and photographs. Because of this, while optimization of individual items on a task-by-task basis has been achieved, overall optimization in terms of ensuring the soundness of telecommunications buildings has not yet been realized. Applying big data related technologies to the area of building maintenance management allows us to find tendencies in

◆Reports are received on buildings nationwide (approximately 20,000 buildings at around 12,000 locations).
◆Data has been accumulated on approximately 50,000 entries a year, over 20 years of operation.



IP-VPN: Internet protocol-virtual private network

Fig. 1. Overview of Building Emergency Call system.

deterioration and the occurrence of faults that were previously unknown. To this end, we are working on achieving lower cost and higher precision in predicting deterioration and in planning, and also in realizing the overall optimization of corporate real estate (CRE[*1]) for the future.

2.2 Application of fault and complaint records

NTT FACILITIES receives fault and complaint reports from associated facilities all over Japan, and we operate the Building Emergency Call service to assist in dealing with them (Fig. 1). This emergency call system receives about 50,000 entries a year, and data has been accumulated over 20 years of operation.

Specific keywords are extracted from these records upon their reception, confirmation, and handling, and then sorted by facility and type of equipment. By examining the combination of keywords and the frequency of occurrence, the system can predict inherent risks as well as potential faults in the future and visualize them as an *alert*. It may also be possible to predict fault occurrences and the time between occurrences by scrutinizing the interrelation between keywords and chronological databases (**Fig. 2**). In the future, we will use regional features and source building or facility information to try to recognize whether a problem involves an individual building, or if it is a matter to be wrestled with in terms of general measures, and reflect it into building maintenance plans.

2.3 Application of foundation and construction records

The existing telecommunications buildings were built mainly in the peak construction period of Japan's postwar economic boom, which spans several decades. These building construction records have been used to create a model cycle of periodically required construction work and costs based on the building type, scale, and construction period. This is being utilized to predict when construction work might be necessary as well as the cost (**Fig. 3**).

*1 CRE: Corporate real estate. All the properties used by corporations for business purposes are positioned as business tools, including rental and lease.

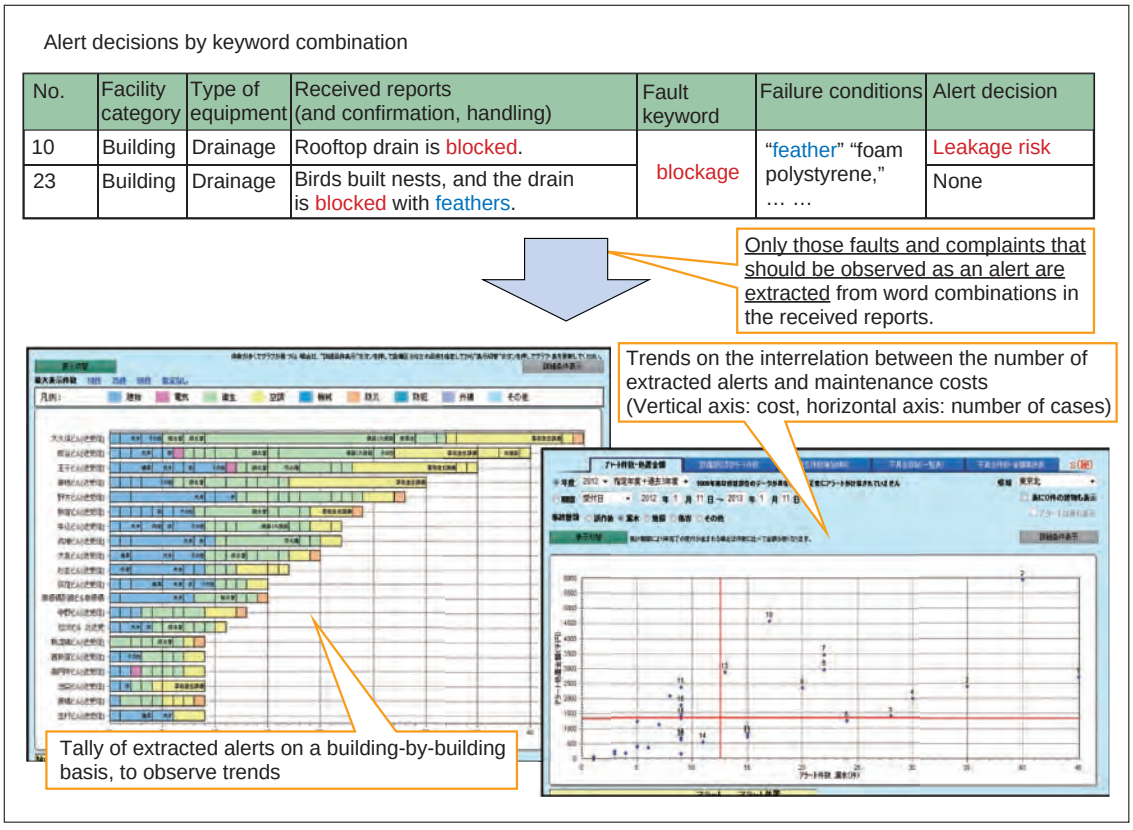


Fig. 2. Extracting alerts from fault and complaint records.

In the event of faults such as leaking roofs or peeling of external walls, simple repairs might be advantageous financially, but the time that passes before the same faults occur again also becomes shorter. On the other hand, although a large-scale repair prevents a fault from recurring for a longer period, it also increases the repair costs. In the event of roof leakage, for example, we can make an accurate and immediate decision on whether to do a simple repair or to carry out a large-scale restoration by considering the right time for construction work with respect to the life cycle of the building. This will lead us to the right conclusion in terms of total cost.

2.4 Information analysis of reports and proposals regarding building maintenance management

Over the last 20 years, NTT FACILITIES has implemented maintenance and management of telecommunications buildings and has also been reporting on maintenance issues and proposing measures to deal with them. The data acquired during this time has been compiled into a massive volume of text and presentation files consisting of various types of data such as figures and photos as well as text.

Until now, it has been possible to refer to each type of data individually, but it has been difficult to look at the entirety of the data chronologically. These files were created by maintenance technicians based on inspection and diagnosis results, and they include various kinds of information and know-how as *tacit knowledge*[*2]. Making this information *explicit knowledge*[*3] and ensuring that it is systematically applied is a challenge for the future.

Handling this information as big data allows us to clarify interrelations between tasks that have previously been performed separately and the correlation between faults or disasters such as typhoons or

*2 Tacit knowledge: Instinct or experience based knowledge passed on within an organization, which is difficult to express in words.
*3 Explicit knowledge: Explainable and expressible knowledge mainly by means of textual, graphical, and numerical formulation.

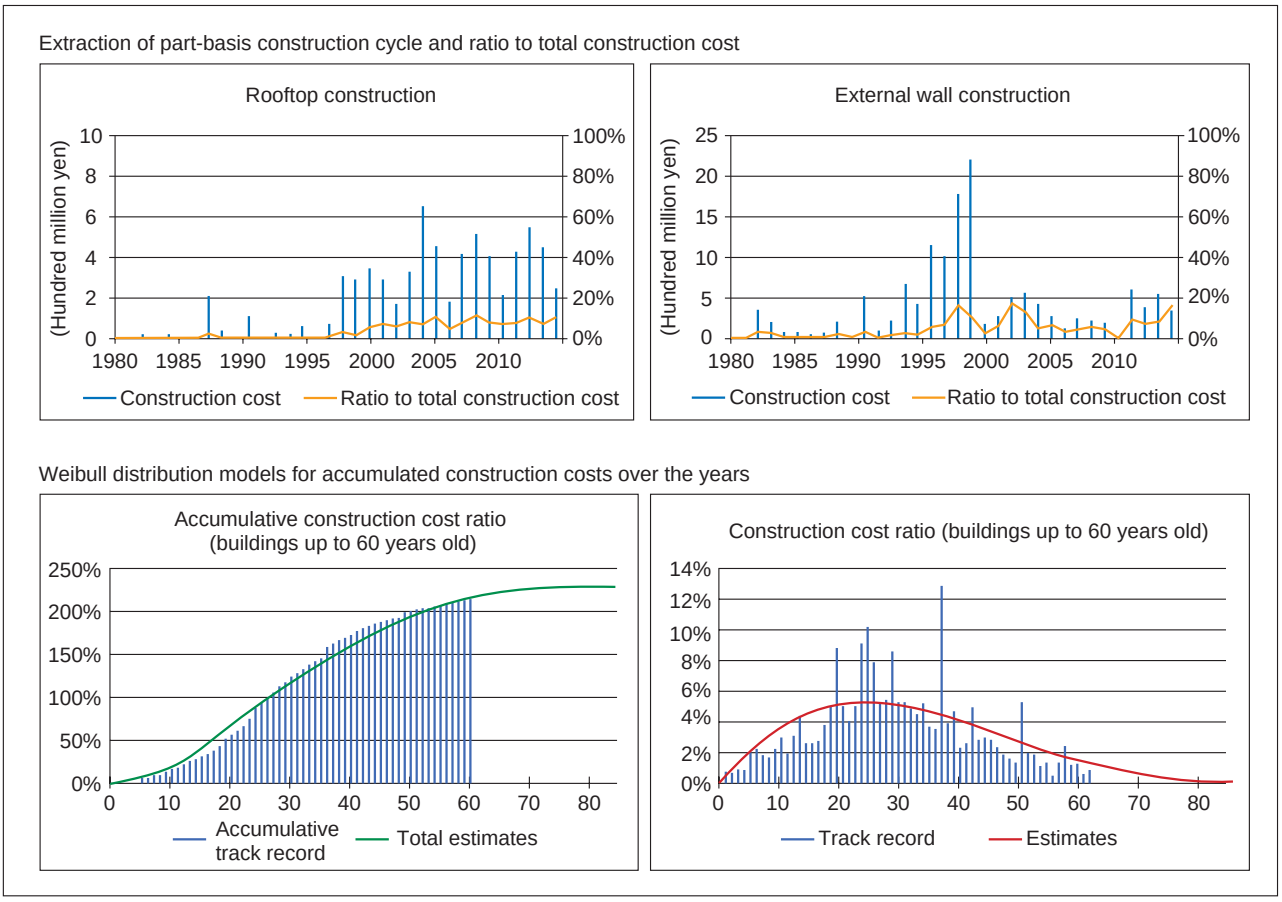


Fig. 3. Construction requirement cycle pattern and cost prediction.

earthquakes and the effectiveness of countermeasures. This encourages us to strive to improve planning measures (**Fig. 4**).

In general, it is said that the ratio of the life cycle cost (LCC[*4]) of a building from its foundation to its demolition and removal is about 25% for construction and 75% for operation. Future building management practices will have to minimize the operational cost that accounts for 75% of LCC by improving task adjustment performance during building operations. In the future, we are going to work systematically on the following objectives by promoting the application of big data to building maintenance and facility management.

(1) Establishment of maintenance prediction technology by developing more efficient and lower-cost inspection and diagnosis processes and higher precision of deterioration prediction.

(2) Optimization of building maintenance and capital expenditure plans by identifying interrelations and trends that were previously unknown.

(3) Overall optimization by considering buildings and facilities as CRE, in relation to real estate utilization. Continuous improvement is aimed for in buildings and facilities in their functions as telecommunications business tools, and practices will be implemented for sustaining them.

In order to optimize the maintenance and capital expenditure plans for decades of building operations, it is necessary to relate various kinds of data such as information obtained in the telecommunications business, technical trends, legal or economic shifts in

*4 LCC: Life cycle cost. For buildings, this refers to the costs of the "life" of the building from design and construction through operation and removal.

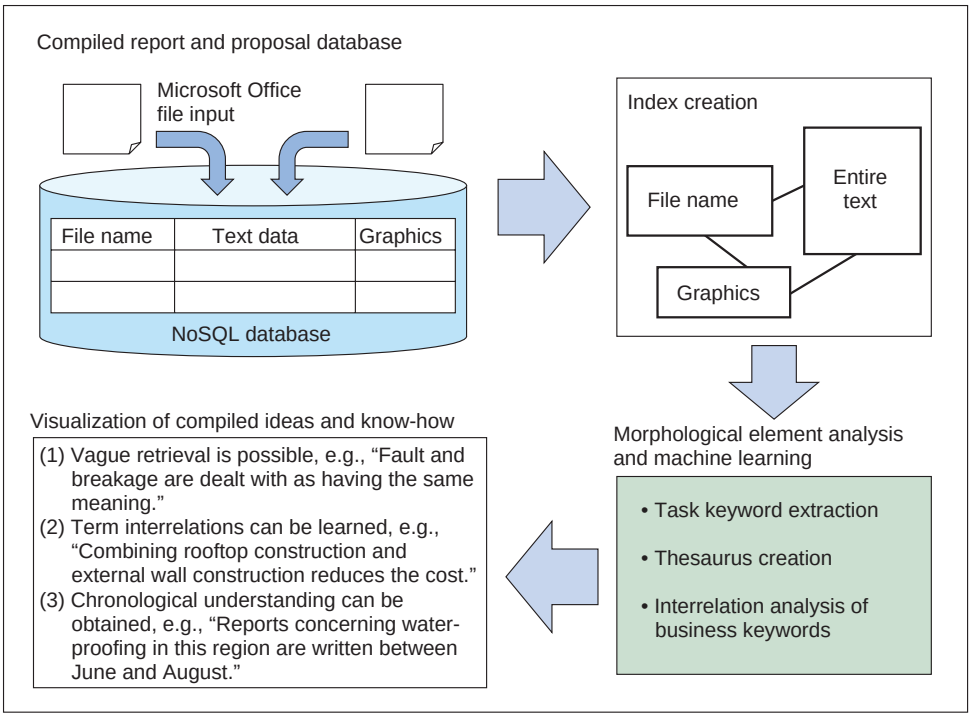


Fig. 4. Analysis of reports and proposals.

social situations, trends in construction technology, as well as conventional building-related data. Another necessary task is to organize and expand a team of data scientists to apply the compiled data.

3. Application of facility maintenance data

3.1 Integrated management system for facility operation

NTT FACILITIES operates an integrated management system it developed for facility operation called MaRIA (**M**aintenance and **R**emote monitoring, **I**ntegrated **A**dvanced management system) that monitors and maintains telecommunications power systems and air conditioning systems in a unified manner (**Fig. 5**).

MaRIA monitors alerts issued due to output abnormalities or power outages at facilities as well as alerts regarding measurement values collected by sensors such as those for voltage, current, and temperature. This system immediately informs the Facilities Operation Center of any measurement values that deviate from a threshold range; the appropriate staff then rush into the field from service centers at around 180 locations across the nation. Information on alerts and measurement values is automatically accumulated in the system database. Also, source equipment, fault details, and maintenance records are compiled by maintenance staff at service centers as facility maintenance information. The amount of information the system has accumulated is huge, amounting to 700 million alerts, 30 billion pieces of measurement data, and information on 300 thousand faults. The maintenance data has been utilized to implement monitoring and maintenance work. We are also currently working on other ways to apply the data, as follows.

3.2 Application of fault information

Every time a fault occurs in an electronic or air conditioning system, a fault card that describes the fault is registered in MaRIA via a web system. The fault card includes information on the *five Ws and one H* that describe the entire event: when and where (i.e., the geographical location), which equipment, what happened, why it happened, and how it was repaired. This has been utilized for fault management and evidence for replacing equipment in the system concerned.

By adding tallying and analysis procedures to this accumulated information, we have developed a tool

◆Remote monitoring of alerts from 200,000 devices and values measured at 450,000 points 24 hours a day/365 days a year
◆Management and accumulation of approximately 20,000 instances of facility faults a year

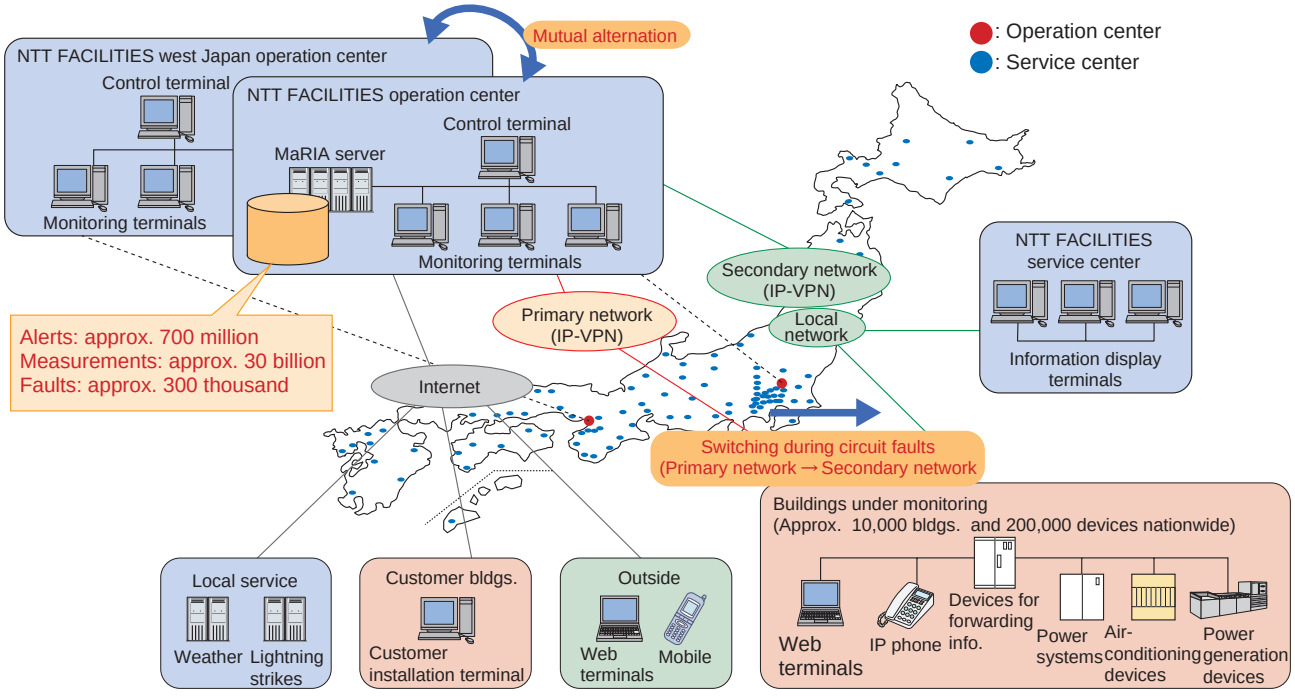


Fig. 5. Overview of MaRIA.

for multidimensional data analysis that applies various approaches to analyze the information on the compiled fault cards. This contributes to generating new added value (**Fig. 6**).

This tool enables any information registered using these cards to be analyzed in a variety of ways. As shown in the figure, general and seasonal trends can be grasped from the shift in the number of faults in each device, and a simultaneous comparison can be done based on a certain manufacturer or a defective component. This allows us to evaluate the situation comparatively using various approaches and to determine, for example, which component of which manufacturer is likely to have more defective parts.

Other kinds of comparative approaches can also be done such as a comparison on a cause basis or an installation environment basis, for example, whether it occurred in a power receiving room or a telecommunications room. A fault rate based comparison is also possible by calculating the operational time of the equipment based on information about the initiation and completion date of operation. This enables us to do a comparative evaluation based on the region or on the year of manufacture, even when the number of installations differs. The example in Fig. 6 confirms a general tendency of facilities; the older ones are more likely to fail. In the future, these types of analyses will help to establish the optimum cycles for conducting detailed analyses and for carrying out the appropriate actions for facility maintenance. We will also employ further approaches using statistical analysis in our continued efforts to improve the quality of facility maintenance.

3.3 Prediction of deterioration trends of storage batteries

NTT FACILITIES monitors and maintains storage batteries that are essential backup power systems for telecommunications equipment. Some of our maintenance methods include conducting thermal management procedures to extend the life of storage batteries by managing the temperatures of the environment surrounding the storage batteries and conducting discharge circuit tests to confirm operational normalcy while storage batteries are being discharged. Also, a backup battery for telecommunications equipment

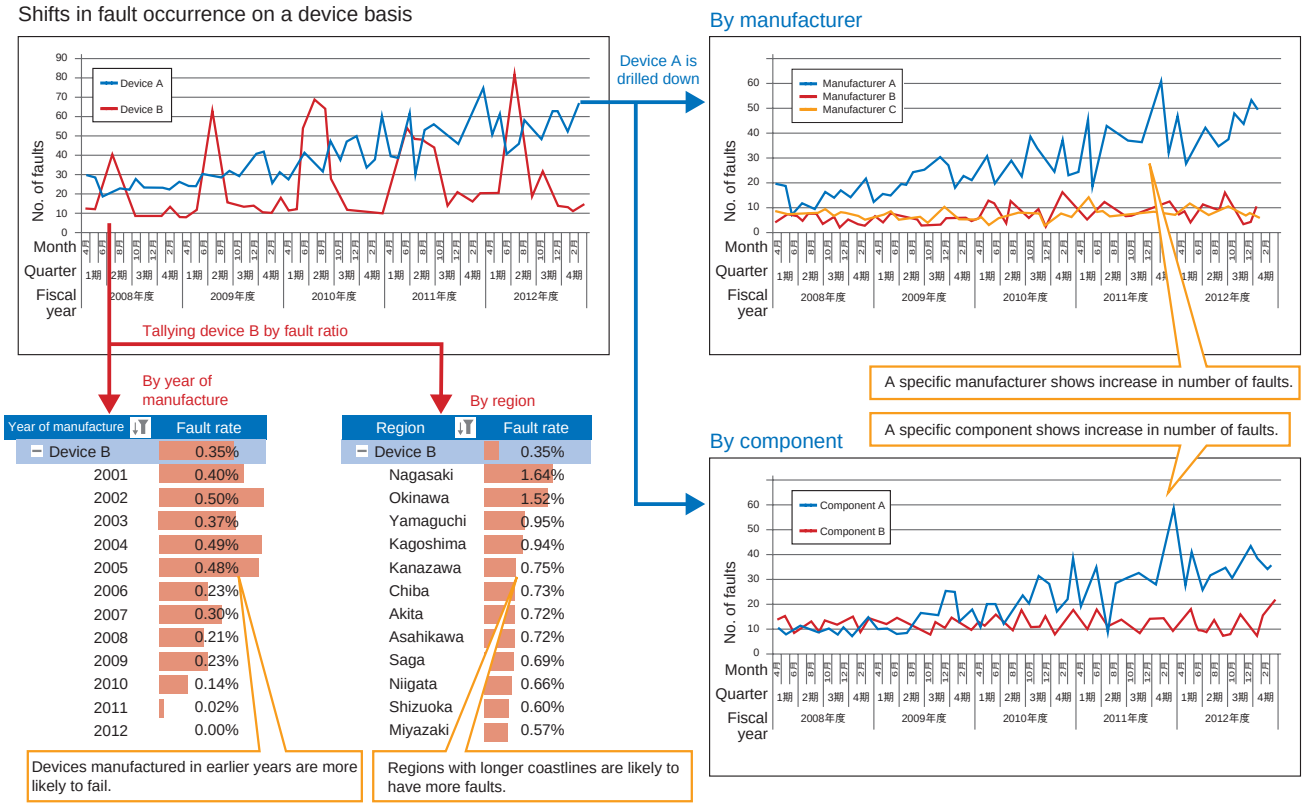


Fig. 6. Multidimensional data analysis of fault information.

consists of several single cells with series connections. The internal resistance of these batteries is also being measured for quality management purposes. Furthermore, the voltage of 1.3 million cells across Japan is being remotely monitored by a storage battery management unit that can measure and manage the voltage of each single cell.

We are currently working on predicting trends in deterioration in order to further improve storage battery management by making use of cell voltages that are continuously measured by the storage battery management unit. Conventionally, information on the measured cell voltages has been utilized to determine whether the voltage is normal or abnormal based on upper and lower limit thresholds. We are attempting to predict deterioration trends by compiling and analyzing this information as data values (**Fig. 7**).

Operational normalcy of the storage batteries is confirmed through the aforementioned examinations such as the discharge circuit test, where storage batteries are periodically discharged. The measured discharge voltages are utilized to create discharge curves with relation to discharge duration and the resulting voltage decrease tendencies. These techniques enable us to evaluate the deterioration state now. The discharge curves of normal storage batteries and ones that might have potential deterioration are shown in Fig. 7, which graphically depicts the voltages of all cells within the same combined battery. The storage batteries with potential deterioration show steeper voltage decreases compared to other cells. In the future, by compiling voltage information on cells during discharge and by analyzing their shift with regard to age, we aim to achieve accurate prediction of deterioration of storage batteries.

4. Future plans

In this article, we described the approaches of NTT FACILITIES in applying field data on buildings and facilities that has been accumulated in the course of maintaining ICT infrastructures. We are fortunate in not only having this vast amount of field data on buildings and facilities but also engineers in each field who can apply and analyze it. With the foresight obtained through these data analyses, we will

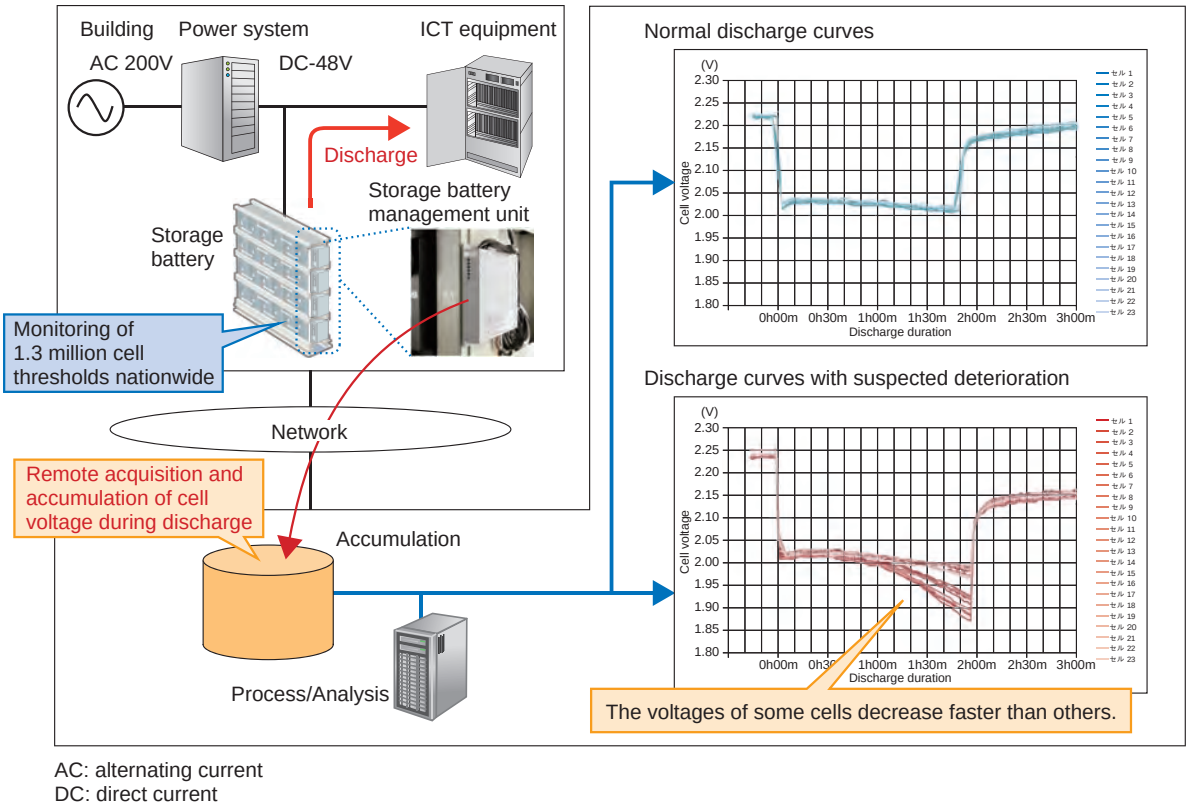


Fig. 7. Voltage shifts in storage batteries during discharge.

contribute to improving the reliability of ICT infrastructures and at the same time will strive to create new maintenance services for buildings and facilities.

Masahiro Yamazaki
Group Leader, Research and Development Headquarters, NTT FACILITIES, INC.
He received the B.E. in electronic engineering from Muroran Institute of Technology, Hokkaido, in 1992. He joined NTT in 1992 and moved to NTT FACILITIES in 2012. His research interests include establishing a means of predicting problems in power supply facilities and air conditioner for communication buildings.

Tatsuro Matsuoka
Senior Research Engineer, Research and Development Headquarters, NTT FACILITIES, INC.
He received the M.E. in intelligent engineering from Nihon University, Tokyo, in 1989. He joined NTT in 1989 and moved to NTT FACILITIES in 1994. His team is responsible for information management, database construction and operation, and R&D of FM systems. Since 2008, his team has been researching and studying techniques for using building information modeling (BIM) in FM. He is a member of the Architectural Institute of Japan. He is licensed as an Architect of the First Class, a Professional Engineer (for computer science sector), and a Certified Facility Manager Japan (CFMJ).

A 10G/1G Dual-rate Burst-mode Receiver for Next Generation Optical Access Networks

Masafumi Nogawa, Hiroaki Katsurai, and Hiroshi Koizumi

Abstract

This article describes a burst-mode receiver that is a key component of 10 Gigabit Ethernet passive optical network (10G-EPON) systems. In addition to coping with a burst signal, the receiver can operate at data rates of both 10.3125 and 1.25 Gbit/s to meet the need for coexistence of future 10G-EPON and existing Gigabit Ethernet PON (GE-PON) optical network units. The receiver sensitivities are –30.3 and –35.6 dBm at data rates of 10.3125 and 1.25 Gbit/s, respectively, when the receiver settling time is 400 ns. These results meet the 10G-EPON and GE-PON specifications with sufficient margins.

Keywords: burst-mode receiver, 10G-EPON, transimpedance amplifier

1. Introduction

The number of subscribers of fiber-to-the-home (FTTH) services exceeds 25 million in Japan [1]. The Gigabit Ethernet passive optical network (GE-PON) [2] is now widely used for FTTH. However, since the data traffic of mobile, cloud computing, and ultra-high-definition (4K/8K) video services is expected to expand rapidly, the demand for bandwidth for optical access network services, including FTTH, is continuously increasing. Thus, an optical access network with higher bandwidth than that of GE-PON is needed. The 10 Gigabit Ethernet PON (10G-EPON) [3], which has ten times the bandwidth of GE-PON, is a promising candidate for the next generation optical access networks.

The PON system is illustrated in **Fig. 1**. The main feature is that multiple customers share one optical fiber, which reduces the cost. An optical network unit (ONU) is installed on each customer's premises, and multiple ONUs connect to an optical line terminal (OLT) in the central office. Downstream data are transmitted as a continuous signal. However, upstream data consist of asynchronous burst signals, and these bursts have different optical power levels and timings. Thus, it is essential that the OLT have a burst-mode receiver that can receive such optical burst signals and instantaneously amplify them to a fixed amplitude in an electrical signal.

Such a burst-mode receiver requires high sensitivity, a wide dynamic range, and a short settling time in order to meet PON specifications. Another important issue is dual-rate operation. When upgrading the GE-PON to 10G-EPON, we have to be able to connect the 10G-EPON OLT to both 10G-EPON ONUs and existing GE-PON ONUs. Therefore, a burst-mode receiver for the 10G-EPON system, particularly in Japan, is strongly required in order to be able to receive 1.25-Gbit/s burst signals from GE-PON ONUs and 10.3125-Gbit/s burst signals from 10G-EPON ONUs as well.

We describe in this article a 10G/1G dual-rate burst-mode receiver containing a burst-mode transimpedance amplifier (TIA) and limiting amplifier (LA). Careful and optimized design of automatic gain control (AGC) with a dummy circuit in the TIA and a 10G/1G dual-path configuration in the LA makes it possible to achieve sensitivities of –30.3 and –35.6 dBm

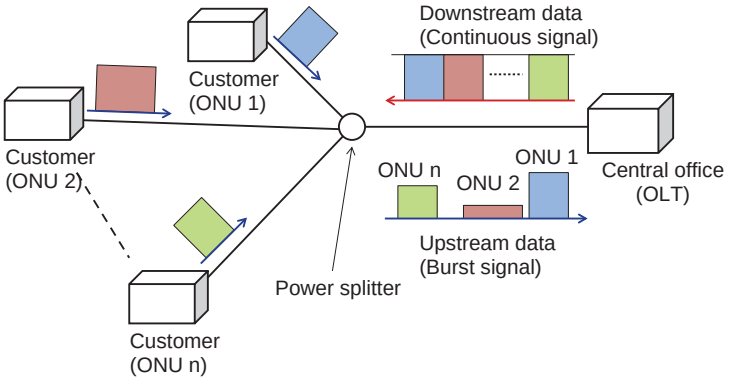


Fig. 1. PON system.

at data rates of 10.3125 and 1.25 Gbit/s, respectively, and a settling time of 400 ns at both data rates.

2. Target performance of dual-rate burst-mode receiver

Several burst-mode receivers have been studied in order to meet the requirements described above. To achieve wide dynamic range of more than 22 dB, which is specified in the 10G-EPON standard, the receiver gain must vary in accordance with the input optical power. That is, the gain should be large when the power is small and small when the power is large. Receivers are generally classified into two groups: those with continuous AGC and those with step AGC [4, 5]. The gain in continuous AGC varies continuously, whereas it varies discontinuously in step AGC. Generally, step AGC has two or three fixed gains.

In a receiver with continuous AGC, a gain control signal tracks the average of the input signal power so that it can tolerate fluctuations in the power. Moreover, a reset signal that initializes the AGC circuit and is generally required in step AGC is not required in many cases. This leads to a simple receiver configuration. However, it requires a relatively longer settling time than that of step AGC. This is because it becomes difficult to tolerate long consecutive identical digits (CIDs) since the gain control signal tracks the average of the input power when the response time of continuous AGC is shortened in order to obtain a short settling time. A continuous AGC burst-mode receiver [6] achieved high sensitivities of –30.8 and –35.5 dBm in 10G and 1G operations, respectively, but needed a settling time of 800 ns in 10G operation. The receiver developed by Takahashi et al. [7] reduced the settling time to 400 ns, although the sensitivity for 1G operation stayed at around –30 dBm because the receiver architecture was not optimized for 1G operation.

By contrast, in a receiver with step AGC, the settling time can be shortened to less than 200 ns [8, 9]. The receiver gain in step AGC is fixed in each step, so the receiver can obtain robustness to long CIDs even when the time constant of the AGC is short. However, step AGC requires a reset signal between each burst packet to initialize the gain. This reset signal should be a short pulse with a width of less than a few tens of nanoseconds, and the reset timing should be precise enough to correctly apply the reset at a gap of two bursts. This leads to a complex receiver configuration.

A recently developed continuous AGC receiver achieved a settling time of 240 ns with high sensitivity by switching the time constant of the AGC [10]. The time constant in this receiver remains short before the reception of a burst signal in order to achieve a short settling time. After a burst signal arrives, the receiver detects the signal and then switches the time constant from short to long in order to handle long CIDs. However, the reset signal has to initialize the time constant of the AGC between each burst packet.

The 10G/1G dual-rate burst-mode receiver presented in this article is a continuous AGC without a reset signal. We selected the time constant of the AGC carefully in order to achieve a settling time of 400 ns, which is half that specified in the 10G-EPON standard and the same as that specified in the GE-PON standard. Furthermore, the targets of the receiver sensitivity are at least –30 and –32 dBm at bit error ratios (BERs) of 10^{-3} and 10^{-12} at data rates of 10.3125 and 1.25 Gbit/s, respectively. These values are 2 dB better

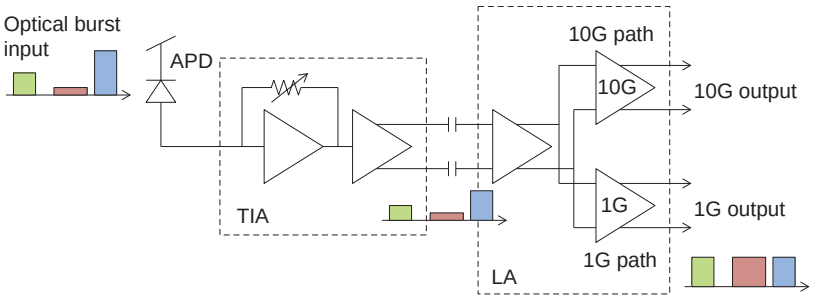


Fig. 2. Block diagram of burst-mode receiver.

than that of the standards. To achieve high sensitivity, we optimized circuit parameters for both 10G and 1G operations.

3. Dual-rate burst-mode receiver architecture and circuit

3.1 Receiver architecture

A simplified schematic of our dual-rate burst-mode receiver, which consists of an avalanche photodiode (APD), TIA, and LA, is shown in **Fig. 2**. The APD converts an input optical burst signal into an electrical current signal. The TIA converts the current signal into a voltage signal and amplifies it. The LA amplifies various voltage amplitudes for each burst packet to a constant level.

The TIA uses continuous AGC, which, as mentioned above, does not require a reset signal. The details are described in section 3.2. The LA has two signal paths: a 10G path for a rate of 10.3125-Gbit/s and a 1G path for a rate of 1.25-Gbit/s. The –3-dB bandwidth of the 1G path was designed to be 1 GHz to improve sensitivity at a data rate of 1.25 Gbit/s, while that of the 10G path was designed to be 8 GHz. The TIA and the LA are connected via alternating current (AC) coupling capacitors to cancel the signal offset.

3.2 Burst-mode TIA

The dual-rate burst-mode TIA circuit is shown in **Fig. 3**. Continuous AGC is used in the TIA to achieve a wide dynamic range and make a reset signal unnecessary. The first stage of the TIA has two identical circuit blocks: a TIA core and TIA dummy. The TIA core receives the input current signal. The TIA dummy generates a reference voltage equal to that of the output of the TIA core with no input current. The reference voltage from the TIA dummy is employed

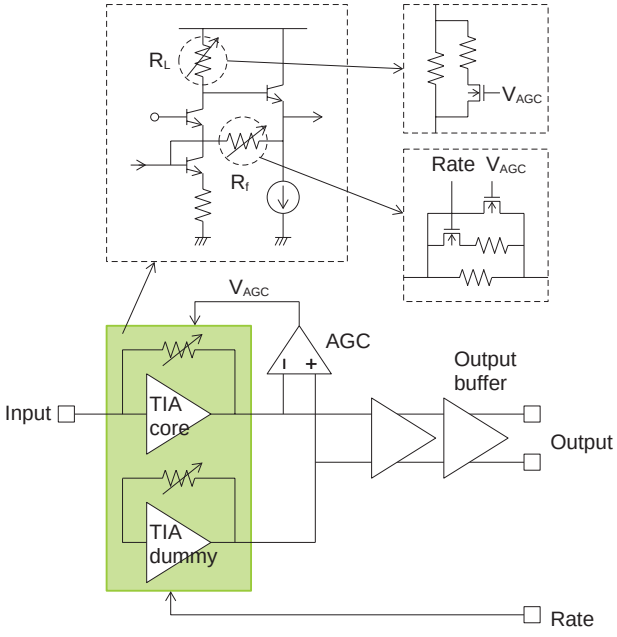


Fig. 3. TIA circuit.

to cancel process, voltage, and temperature (PVT) variations, to generate a precise AGC voltage, and to cancel common mode noise. The AGC circuit averages the difference between the output voltages of the TIA core and dummy and then outputs an AGC signal, V_{AGC}. When the input current is small, V_{AGC} is low. This makes feedback resistance, R_f, and load resistance, R_L, high, which results in a high TIA gain. This enables the TIA to receive a small input current. When the input current is large, V_{AGC} becomes high. This makes R_f and R_L low, which results in a low gain. In this way, the linear operation of the TIA is maintained, and the output signal is prevented from saturating and distorting even when the input current

is large.

A *rate* signal is employed to optimize the sensitivity at each data rate. To obtain a high-sensitivity receiver, it is important to reduce the equivalent input noise current, i_{eq}, of a TIA. Most of the i_{eq} comes from R_f, and i_{eq} becomes low as R_f becomes large. Thus, to obtain high sensitivity, large R_f is desired. However, large R_f limits the –3-dB bandwidth of a TIA because the bandwidth is inversely proportional to R_f. In other words, the optimum value of R_f is different at different data rates. Therefore, we employed a rate signal to change R_f according to the data rate.

One of the most important parameters for burst-mode receivers is the time constant for AGC. A study showed that the time constant should be shorter than one-fifth of the receiver settling time and more than eight times the maximum number of CIDs [11]. Since the target for the settling time is 400 ns, the time constant should be less than 80 ns. The maximum number of CIDs in GE-PON and 10G-EPON is 5 and 66 bits because GE-PON and 10G-EPON use 8B/10B and 64B/66B coding, respectively. Consequently, the longest CID period is 4 ns in GE-PON and 6.4 ns in 10G-EPON, so the time constant should be larger than 50 ns. In other words, the –3-dB low-cutoff frequency of the frequency response of the receiver should be between 2 and 3 MHz. In our TIA, we designed the AGC loop to have a cutoff frequency of 2 MHz because we considered it important to have a tolerance to CIDs longer than 66 bits.

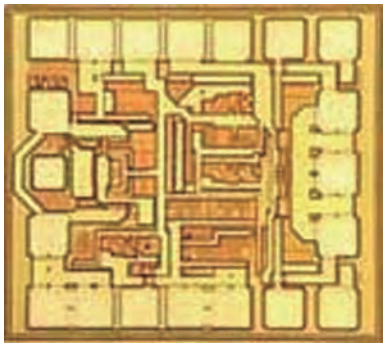


Fig. 4. Micrograph of TIA IC.

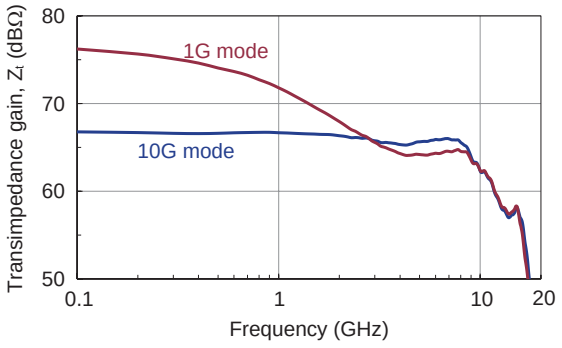


Fig. 5. Measured frequency responses of TIA.

4. Experimental results

A micrograph of the prototype TIA integrated circuit (IC) is shown in **Fig. 4**. The IC was fabricated by using a 0.25-μm SiGe BiCMOS (bipolar complementary metal-oxide semiconductor) process. The chip size is 1.1 × 0.99 mm^2.

4.1 Frequency response of the TIA

The measured frequency responses of the TIA are shown in **Fig. 5**. The transimpedance gain, Z_t, of the TIA was calculated from S (scattering) parameters obtained by on-wafer measurement. When we calculated Z_t, we also took into consideration the parasitic capacitance of an APD and the inductance of the bonding wire between the APD and the TIA. The 10G and 1G modes indicated in Fig. 5 denote that the feedback resistance, R_f, is set small for 10.3125-Gbit/s operation and large for 1.25-Gbit/s operation, respectively, by using a rate input as described in section 3.2. As shown in Fig. 5, we obtained Z_t of 66 and 76 dB, and a –3-dB bandwidth of 9 and 0.8 GHz in the 10G and 1G modes, respectively. These are sufficient gain values and bandwidths for 10G/1G dual-rate operation.

4.2 Performance of dual-rate burst-mode receiver

The TIA was mounted in a transistor outline (TO)-can package with an APD. A flexible printed circuit was soldered to the lead pins of the TO-can package as an electrical interface. The package was connected to the LA through the flexible printed circuit and AC-coupling capacitors, as shown in Fig. 2.

Burst responses of the dual-rate burst-mode receiver from loud to soft bursts are shown in **Fig. 6**. The loud burst was used as an interference burst, and its optical power was –6 dBm, which is the same as the overload specified in the standards. This is the most severe condition for evaluating the settling time and sensitivity of burst-mode receivers. The soft burst was used as a measured burst consisting of a preamble and payload. The preamble pattern defined in the

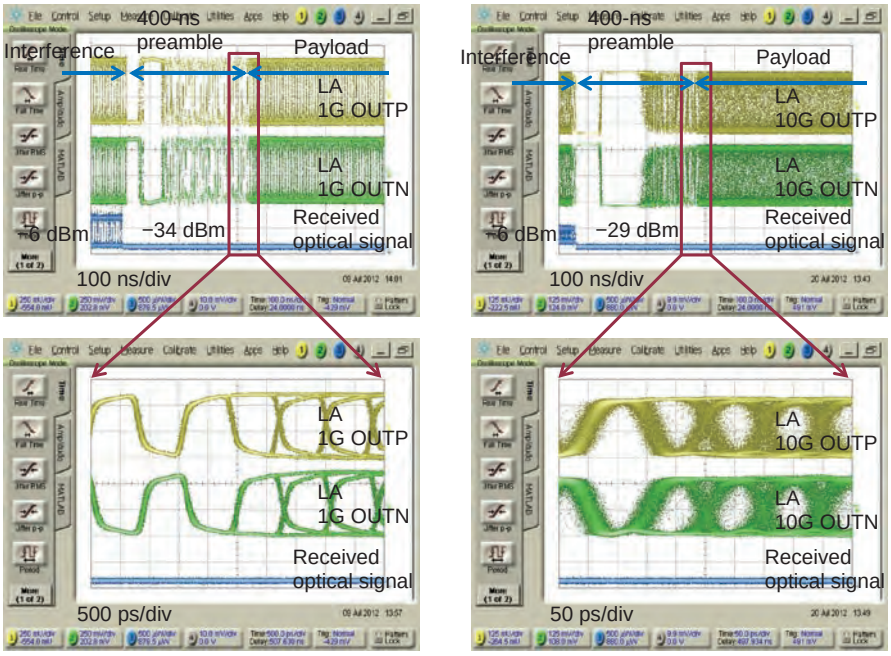


Fig. 6. Received optical and LA output waveforms for burst-mode operation from loud to soft burst signal with 400-ns preamble. (a) 1.25-Gbit/s burst response from −6 to −34 dBm. (b) 10.3125-Gbit/s burst response from −6 to −29 dBm.

existing standards was used, and its period was 400 ns to confirm a receiver settling time of 400 ns. The payload patterns were 2^{31}-1 and 2^{7}-1 PRBS (pseudo-random binary sequence), and the input optical powers of the measured bursts (soft bursts) were –29 and –34 dBm at 10.3125 and 1.25 Gbit/s, respectively. Since these are lower powers than the receiver sensitivities specified in the standards, these values are sufficient for evaluating the settling time. In addition, the guard time between interference and measured bursts was set at less than 10 ns. As shown in Fig. 6, the receiver can respond in less than 200 and 300 ns for 1.25- and 10.3125-Gbit/s operations, respectively, which means the receiver has sufficient margin to achieve a settling time of 400 ns.

The measured BER is shown in **Fig. 7**. The data patterns used in the BER measurement are the same as described above. The measured sensitivities are –30.3 dBm at a BER of 10^{-3} and –35.6 dBm at a BER of 10^{-12} at 10.3125 and 1.25 Gbit/s, respectively. The sensitivity is more than 2 dB higher than that in the standards. The measured results and the standard specifications are summarized in **Table 1**. We confirmed that the receiver successfully meets the 10G-EPON/GE-PON standards.

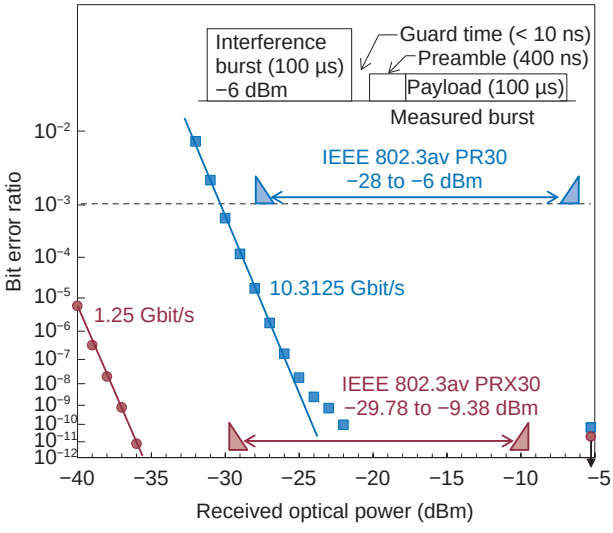


Fig. 7. Measured BER of dual-rate burst-mode receiver.

Table 1. Measured results for dual-rate burst-mode receiver.

	Unit	IEEE 802.3av (10G-EPON)	IEEE 802.3ah (GE-PON)	Measured results
10G operation				
Power budget class	—	PR30	N/A	—
Data rate	Gbit/s	10.3125	N/A	10.3125
Sensitivity	dBm	−28	N/A	−30.3
Overload	dBm	−6	N/A	−5
T_receiver_settling	ns	< 800	N/A	< 400
1G operation				
Power budget class	—	PRX30	PX20	—
Data rate	Gbit/s	1.25	1.25	1.25
Sensitivity	dBm	−29.78	−27	−35.6
Overload	dBm	−9.38	−6	−5
T_receiver_settling	ns	< 400	< 400	< 400

5. Conclusion

We have developed a dual-rate burst-mode receiver that complies with 10G-EPON and GE-PON standards. The receiver sensitivities are –30.3 and –35.6 dBm at 10.3125 and 1.25 Gbit/s, respectively, with a settling time of 400 ns and without a reset signal by using a continuous AGC. These values are more than 2 dB higher than that in the standards.

References

[1] Ministry of Internal Affairs and Communications (in Japanese), http://www.soumu.go.jp/johotsusintokei/field/data/gt010103.xls

[2] IEEE Standard 802.3ah, 2004, http://www.ieee802.org/21/doctree/2006_Meeting_Docs/2006-11_meeting_docs/802.3ah-2004.pdf

[3] IEEE Standard 802.3av, 2009, http://standards.ieee.org/findstds/standard/802.3av-2009.html

[4] Y. Ohtomo, H. Kamitsuna, H. Katsurai, K. Nishimura, M. Nogawa, M. Nakamura, S. Nishihara, T. Kurosaki, T. Ito, and A. Okada, "High-speed Circuit Technology for 10-Gb/s Optical Burst-mode Transmission," Proc. of Conference on Optical Fiber Communication (OFC) 2010, OWX1, San Diego, CA, USA, Mar. 2010.

[5] X. Z. Qiu, X. Yin, J. Verbrugghe, B. Moeneclaey, A. Vyncke, C. V. Praet, G. Torfs, J. Bauwelinck, and J. Vandewege, "Fast Synchronization 3R Burst-mode Receivers for Passive Optical Networks," Journal of Lightwave Technology, Vol. 32, No. 4, pp. 644–659, Feb. 2014.

[6] M. Noda, S. Yoshima, K. Ishii, S. Shirai, M. Nogami, and J. Nakagawa, "Dual-rate Optical Transceiver Incorporating Fully Optimized Burst-mode AGC/ATC Functions for 10G-EPON Systems," Proc. of European Conference and Exhibition on Optical Communication (ECOC) 2010, Mo.2.B.2, Torino, Italy, Sept. 2010.

[7] S. Takahashi, K. Shiba, E. Mizuki, K. Makita, and A. Tajima, "Over 25-dB Dynamic Range 10-/1-Gbps Optical Burst-mode Receiver Using High-power-tolerant APD," Proc. of OFC 2009, NME2, San Diego, CA, USA, Mar. 2009.

[8] M. Nogawa, H. Katsurai, M. Nakamura, H. Kamitsuna, and Y. Ohtomo, "10-Gbit/s Burst-mode Receiver Integrated Circuits for Broadband Optical Access Networks," NTT Technical Review, Vol. 9, No. 3, Mar. 2011. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201103fa5.html

[9] T. Ito, T. Kurosaki, M. Nakamura, S. Nishihara, Y. Ohtomo, A. Okada, and M. Yoneyama, "A Burst-mode APD-ROSA Using Reset Signal with Less Than 100 ns Response for 1G/10G-EPON Dual-rate Optical Transceivers," Journal of Lightwave Technology, Vol. 29, No. 14, pp. 2089–2101, Jul. 2011.

[10] S. Yoshima, M. Noda, N. Suzuki, S. Shirai, D. Mita, S. Ihara, M. Nogami, H. Aruga, and J. Nakagawa, "Latest Progress of Burst-mode Transceiver for 10G-EPON," Proc. of ECOC 2012, Tu.1.B.3, Amsterdam, the Netherlands, Sept. 2012.

[11] H. Ben-Amram, "10GEPON Burst Receiver Simulation," http://www.ieee802.org/3/av/public/2008_01/3av_0801_benamram_2.pdf.

Masafumi Nogawa

Senior Research Engineer, Supervisor, Metro-Access Network Device Project, NTT Device Innovation Center.

He received his B.E. and M.E. in electronic engineering from Tohoku University, Miyagi, in 1988 and 1990, respectively. He joined NTT LSI Laboratories in 1990 and has been engaged in research on high-speed CMOS/BiCMOS circuit technology. His current research interests include high-speed analog front-end circuits for optical communications. He served as an Associate Editor of IEICE Transactions on Electronics and as a member of the Technical Program Committee of ISSCC (International Solid-State Circuits Conference). He is a member of the Institute of Electrical and Electronics Engineers (IEEE) and the Japan Society of Applied Physics (JSAP) and a senior member of the Institute of Electronics, Information and Communication Engineers (IEICE).

Hiroaki Katsurai

Researcher, Metro-Access Network Device Project, NTT Device Innovation Center.

He received his B.E. in systems innovation and his M.E. in precision engineering from the University of Tokyo in 2005 and 2007, respectively. He joined NTT Microsystem Integration Laboratories in 2007. He is currently working on the development of high-speed LSIs for optical communications. He is a member of IEICE.

Hiroshi Koizumi

Senior Research Engineer, Supervisor, Social Device Technology Laboratory, NTT Device Technology Laboratories.

He received his B.S. and M.S. in materials engineering from Waseda University, Tokyo, in 1990 and 1992, respectively. He has been engaged in research and development of high-speed LSIs for optical communications. His current research interests include optical biosensing devices. He is a member of IEEE, JSAP, and IEICE.

Cooperation between Telecommunication Enterprises in Japan and Other Countries in Smart Home Standardization Efforts—Activities of the HGI/ZigBee Alliance

Takefumi Yamazaki, Nobuyuki Akazawa, Akihiro Otaka, Nobuhiro Takagi, Kenji Muto, Takenao Nakagawa, and Kazuyuki Terao

Abstract

The technical issues that are important for the business development of smart home technology are being addressed through cooperation and solution-sharing with telecom companies around the world in international standardization activities. These activities include the sharing of technical requirements by telecom companies in the Home Gateway Initiative and the use of requirements documents for standardization in the ZigBee Alliance. This article reports the current states of these activities.

Keywords: smart home, HGI, ZigBee

1. Introduction

The term *smart home* refers to devices in the home that are connected to services via the Internet. These services are being developed by telecom companies around the world, for example, AT&T's Digital Life [1] and Deutsche Telkom's QIVICON [2]. An important factor in developing smart home business is the in-home wireless communication protocol for connecting a home gateway and other customer premises equipment with various devices such as home appliances, sensors, and actuators.

NTT is working to create an environment that facilitates smart home business development and to expand the market by cooperating with other telecom companies through the Home Gateway Initiative (HGI) [3], which is an influential organization in the telecommunications industry. The HGI members are focusing on improving ZigBee and other important in-home wireless communication protocols.

2. HGI activities

2.1 Overview of HGI

HGI was founded mainly by European telecom companies in 2004 for the purpose of writing specifications for applications and technology with respect to the requirements of home gateways. The members are telecom companies and vendors of systems, chips, software, and hardware. The members from Japan include NTT, NEC, Oki Electric Industries, Sumitomo Electric Industries, Hitachi, and Mitsubishi Electric. The basic plan is to effectively apply existing technology and to collect and organize common specifications for home gateways on the business development and technical fronts rather than

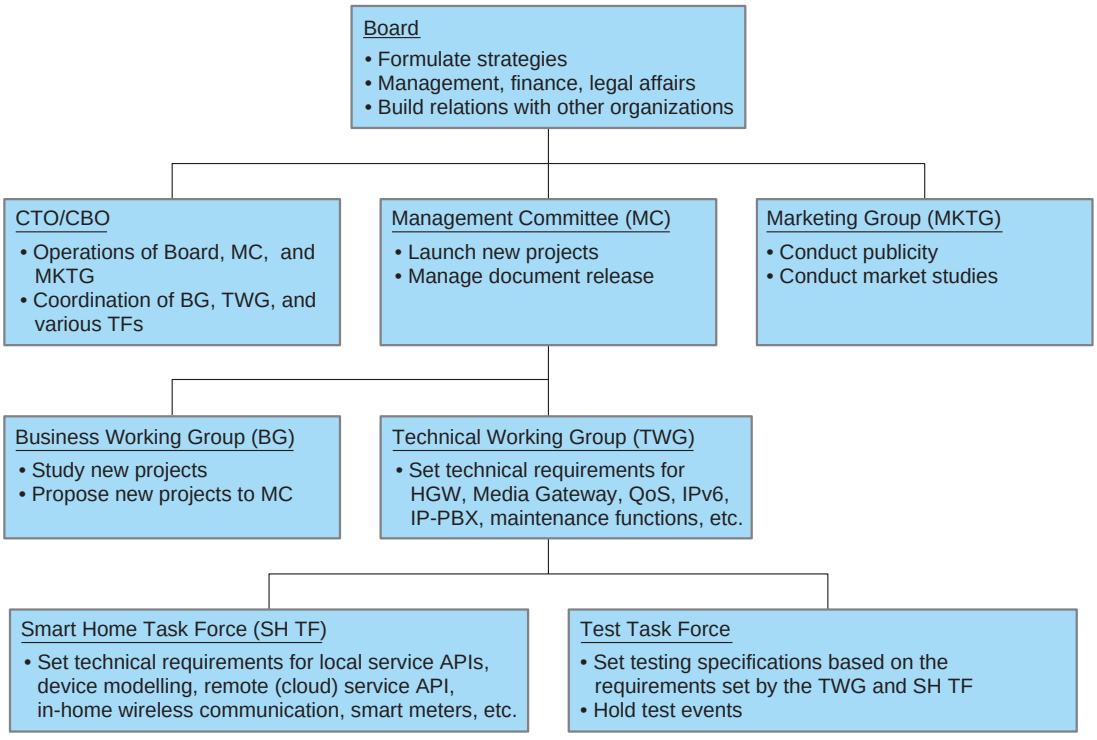


Fig. 1. HGI organization and roles.

writing new specifications.

The HGI organization and roles are shown in **Fig. 1**. Work on the technical requirements documents is mainly being done by the Technical Working Group. The completed specification documents concerning home networks and M2M (machine-to-machine) are submitted to the relevant standards organizations for feedback via liaisons. The Technical Working Group utilizes the feedback from the standards organizations to set new specifications.

2.2 Activities of Smart Home Task Force

The technical requirements for realizing smart home services are being discussed by the Smart Home Task Force, which produces a requirements specifications document (**Table 1**). The main technical fields and topics include use cases, architecture and system requirements, the operating environment of home gateways, in-home wireless communication protocols, and device models.

The home gateway operating environment is being specified mainly by Deutsche Telekom and Orange (formerly France Telecom) as sets of software modules and performance requirements for running in a Java/OSGi (Open Service Gateway initiative) software environment. HGI periodically holds test events to validate the specified performance requirements. Those events are held to determine whether or not the technology can meet the requirements and to build relationships among members.

2.3 Setting technical requirements for in-home wireless communication (RD039)

In cooperation with Deutsche Telekom, Orange, NEC, and other members, we have been working in the Smart Home Task Force to set the technical requirements for existing in-home wireless communication protocols from the viewpoint of providing smart home services (RD039) [4].

Document RD039 sets the requirements that

Table 1. HGI Release Documents.

Number	Title
HGI-RD001-R2.01	Home Gateway Technical Requirements: Residential Profile V1.01
HGI-GD002-R2.01	Remote Access Guidelines
HGI-GD003-R2	Parental Control in the Home
HGI-GD004-R2	Performance Metrics
HGI-GD006-R2	IMS (Internet protocol Media Subsystem) Enabled HG
HGI-RD007-R2	Requirements for HG Interworking with an External NT (Network Termination)
HGI-RD008-R3	HG Requirements for Software Execution Environment
HGI-RD009-R3	Requirements for an Energy Efficient Home Gateway
HGI-RD010-R3	Home Gateway Requirements for Multiple Session Support
HGI-GD013-R2	QoS Whitepaper
HGI-RD015-R3	Requirements for Common Power Supply for Home Networking Equipment
HGI-RWD016-R3	HG and Home Network Diagnostics Module Requirements
HGI-GD017-R3	Use Cases and Architecture for a Home Energy Management Service
HGI-RD024	Requirements for an NGA (Next Generation Access) (Active Line Access) Capable NT
HGI-RD026	IP-PBX Module Requirements
HGI-RD027-R3	Home Gateway QoS Module requirements
HGI-RD039	Requirements for Wireless Home Area Networks (WHANS) Supporting Smart Home Services
HGI-RD048	HG Requirements for HGI Open Platform 2

in-home wireless communication protocols must satisfy for usability (e.g., simple setup, setup support), communication reliability (e.g., averting interference, coverage), maintenance functions (e.g., testing functions, remote support), and security (e.g., authentication, encryption). NTT has proposed technical requirements for simplified pairing, optimized power consumption, appropriate addressing on restart, and isolation for when network problems occur, based on assumed improvements in the ZigBee IP (Internet protocol) expected to be used in Japan.

The technical requirements were completed in May 2014 and formally released as RD039. Through collaboration with the Broad Band Forum (BBF) [5], another standards organization in the smart home field that was established in May 2013, HGI made RD039 available to BBF members as well.

3. ZigBee Alliance Activities

3.1 ZigBee Alliance

The ZigBee Alliance [6] was established in October 2002 to set wireless sensor network standards and ensure interconnectability. This alliance follows the IEEE (Institute of Electrical and Electronics Engineers) 802.15.4 standard for the physical layer and MAC (media access control) layer, and is proceeding with the standardization for the higher layers including network layers and application interfaces. The members include telecom companies and vendors of systems, chips, software, and hardware. Companies from Japan that are participating include NTT, Oki Electric Industries, NEC, Toshiba, and others. In Japan, ZigBee SIG (special interest group) Japan cooperates with the ZigBee Alliance and other regional organizations to conduct information campaigns and carry out market studies and market development in collaboration with member corporations.

The ZigBee Alliance organization and roles are shown in **Fig. 2**, and the process of setting the technical requirements is shown in **Fig. 3**. The ZigBee Marketing Steering Committee (ZMSC) and the Market Working Groups compile use cases and market requirements in the Market Requirements Document (MRD), which is submitted to the ZigBee Architecture Review Committee (ZARC). ZARC passes the MRD on to the appropriate Working Groups (WGs), which produce Technical Requirements Documents (TRDs). The TRDs define the technical scope, and specifications are written based on them.

3.2 Activities in the Technical WG

NTT explained the in-home wireless technical requirements using the HGI RD039 document described in section 2.3 at the ZigBee Alliance in order to clarify to the Alliance members that ZigBee

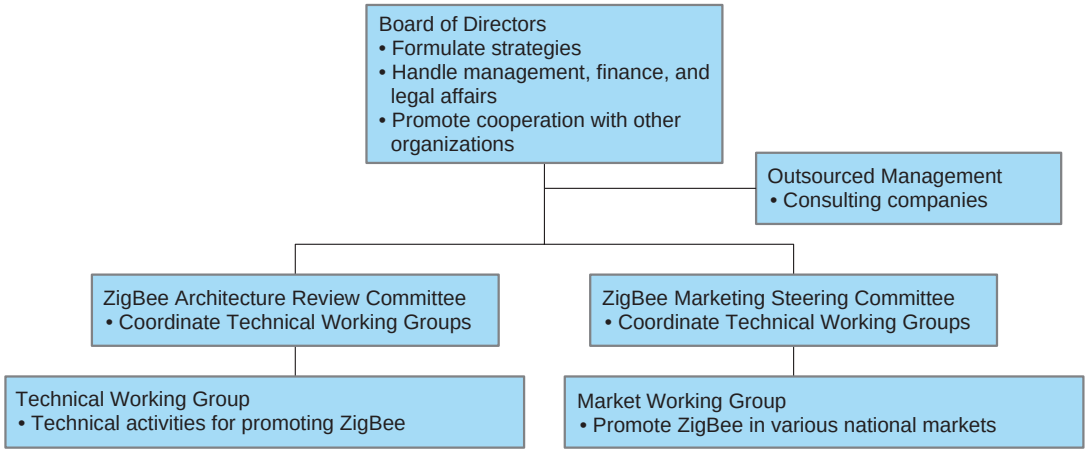


Fig. 2. ZigBee Alliance organization and roles.

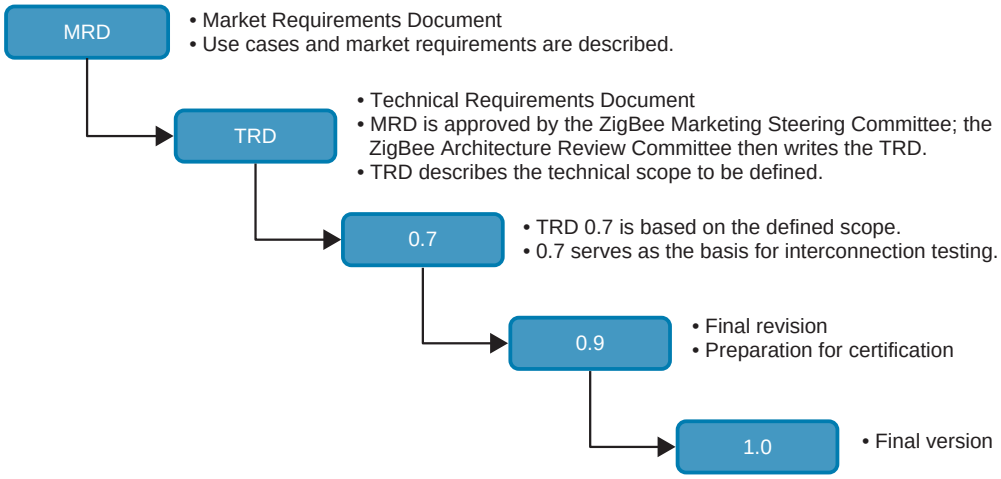


Fig. 3. Process of setting technical requirements in ZigBee Alliance.

specifications should also follow such requirements. Also, as a member of the ZigBee Technical WG, we regularly compare the HGI in-home wireless technical requirements and the ZigBee technical specifications to check for correspondence. Simple connection functions for which there are no corresponding ZigBee technical specifications were proposed as new functions that are now being studied. The work on the MRD and TRD described in Fig. 3 had been completed as of June 2014.

4. Publicity activities at exhibitions

In addition to using HGI RD039 to build relationships with the ZigBee Alliance and other wireless standards organizations, we are promoting improvements of in-home wireless communication protocols through presentations at international exhibitions. So far, we have made presentations at the Broad Band World Forum (BBWF; October 2013) in the Netherlands and at Wireless Japan 2014 (May 2014) in Japan (**Fig. 4**).

5. Conclusion

We have described our work toward developing smart home business through cooperation with telecom companies in other countries in the arena of

Fig. 4. Presentation at Wireless Japan 2014.

international standardization, including HGI and the ZigBee Alliance. We will continue to participate in international standardization activities in the relevant technological fields and to collaborate with other telecom companies in solving problems and promoting services for smart home systems.

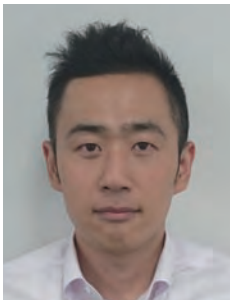
References

[1] Website of AT&T Digital Life, http://www.att.com/shop/digital-life.html
[2] Website of Qivicon, https://www.qivicon.com/
[3] Website of HGI, http://www.homegateway.org/
[4] HGI RD039, http://www.homegateway.org/publis/RD039-Req-for-Wireless-home-area-networks.pdf
[5] Website of Broad Band Forum, http://www.broadband-forum.org/
[6] Website of ZigBee Alliance, http://www.zigbee.org/

Takefumi Yamazaki
Senior Research Engineer, Supervisor, NTT Service Evolution Laboratories.
He received the B.E. in physical engineering from the University of Tokyo in 1986 and joined NTT the same year. He moved to NTT Service Evolution Laboratories in 2006. He is involved in standardization activities related to home networking. He is a member of the board of directors of both OSGi Alliance and HGI. He is also Chair of the Next Generation Home Network System WG in the Telecommunication Technology Committee.

Nobuyuki Akazawa
Senior Manager, 1st Division, New Business Development HQ, NTT EAST.
He received the B.E. from Nagoya Institute of Technology in 1987 and joined NTT the same year. He has worked in the areas of telecommunication equipment, corporate sales, and corporate planning. He is currently responsible for developing home gateways.

Akihiro Otaka
Executive Manager, NTT Access Network Service Systems Laboratories.
He received the B.S. and M.S. in physics from the University of Tokyo in 1989 and 1991, respectively. He joined NTT in 1991 and engaged in developing optical lithography technologies for LSI fabrication. In 1998, he began working on the development and standardization of optical access systems such as Gigabit and 10 Gigabit EPON. From 2010 to 2014, he was with NTT EAST R&D Center, where he worked on optical access, wireless access, and wireless home networks. He joined NTT Access Network Service Systems Laboratories in 2014.



Nobuhiro Takagi
Engineer, NTT Broadband Platform.
He received the M.E. in electrical engineering and computer science from Shibaura Institute of Technology, Tokyo, in 2006. He joined NTT EAST in 2006 and engaged in the development of Wi-Fi and various wireless devices. He moved to NTT Broadband Platform in 2014 and works in the maintenance section.

Kenji Muto
Senior Manager, Strategic Business Development, Global Business Office, NTT EAST.
He received the B.Ec. in economics from Rikkyo University, Tokyo, in 1997 and an MBA from Emory University, Atlanta, GA, USA, in 2008. He is currently responsible for global carrier relations and strategic business development. Before moving to the Global Business Office, he was in charge of corporate sales strategies for Yamagata Prefecture.

Takenao Nakagawa
Manager, Strategic Business Development, Global Business Office, NTT EAST.
He received the M.E. in global environment and energy engineering from the University of Tokyo in 2006 and joined NTT the same year. His work has focused on customer support, network management, and home gateway development. He is currently involved in global carrier relations and standardization.

Kazuyuki Terao
Senior Manager, Alliance Development Office, Marketing Department, NTT WEST.
He received his M.S. in electrical engineering from Kyoto University in 1996 and an MBA from Indiana University, Bloomington, IN, USA, in 2005. He is currently responsible for designing the "HIKARI" box that enables users to access the Internet and various applications on TV and also for managing its sales and marketing. Before moving to his current department, he was in charge of planning technology strategies and budgeting for all research activities at the R&D Center. He also led research projects related to smart communities such as HEMS and M2M at the R&D Center.

LSP (Line Spectrum Pair): Essential Technology for High-compression Speech Coding

Takehiro Moriya

Abstract

Line Spectrum Pair (LSP) technology was accepted as an IEEE (Institute of Electrical and Electronics Engineers) Milestone in 2014. LSP, invented by Dr. Fumitada Itakura at NTT in 1975, is an efficient method for representing speech spectra, namely, the shape of the vocal tract. A speech synthesis large-scale integration chip based on LSP was fabricated in 1980. Since the 1990s, LSP has been adopted in many speech coding standards as an essential component, and it is still used worldwide in almost all cellular phones and Internet protocol phones.

Keywords: LSP, speech coding, cellular phone

1. Introduction

On May 22, 2014, Line Spectrum Pair (LSP) technology was officially recognized as an Institute of Electrical and Electronics Engineers (IEEE) Milestone. Dr. J. Roberto de Marca, President of IEEE, presented the plaque (**Photo 1**) to Mr. Hiroo Unoura, President and CEO of NTT (**Photo 2**), at a ceremony held in Tokyo. The citation reads, "Line Spectrum Pair (LSP) for high-compression speech coding, 1975. Line Spectrum Pair, invented at NTT in 1975, is an important technology for speech synthesis and coding. A speech synthesizer chip was designed based on Line Spectrum Pair in 1980. In the 1990s,

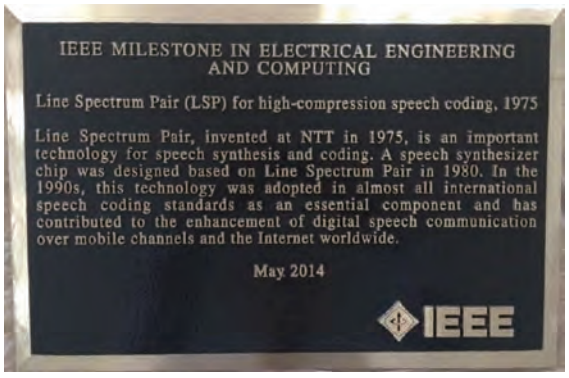


Photo 1. Plaque of IEEE Milestone for Line Spectrum Pair (LSP) for high-compression speech coding.

Photo 2. From IEEE president to NTT president.

this technology was adopted in almost all international speech coding standards as an essential component and has contributed to the enhancement of digital speech communication over mobile channels and the Internet worldwide."

IEEE Milestones recognize technological innovation and excellence for the benefit of humanity found in unique products, services, seminal papers, and patents, and they have so far been dedicated to more than 140 technologies around the world.

2. Properties of LSP

LSP is an equivalent parameter set of LP (linear prediction) coefficients $a[i]$. Among the various types of linear prediction, AR (auto-regressive) or all-pole systems have mainly been used in speech signal processing. In an AR system, the current sample is predicted by summation (from 1 to p, e.g., 16) of i past sample multiplied by each associated coefficient $a[i]$. A prediction error signal $\hat{x}[n]$ at time n is obtained by the difference between the current sample $x[n]$ and the predicted values of the term $-\sum_{i=1}^{p}$ as

$$\hat{x}[n] = x[n] + \sum_{i=1}^{p} a[i]x[n-i] \; . \qquad (1)$$

The preferable set of $a[i]$ can be adaptively determined to minimize the average energy of prediction errors in a frame. This relation can be represented by the polynomial of z as

$$A(z) = 1 + \sum_{i=1}^{p} a[i]z^{-i} \; , \qquad (2)$$

while $1/A(z)$ represents the transform function of the synthesis filter. The frequency response of $1/A(z)$ can be an efficient approximation of the spectral envelope of a speech signal or that of a human vocal tract. This representation, normally called linear prediction coding (LPC) technology, has been widely used in speech signal processing, including for coding, synthesis, and recognition of speech signals. Pioneering investigations of LPC were started independently, but simultaneously, by Dr. F. Itakura at NTT and Dr. M. Schroeder and Dr. B. Atal at AT&T Bell Labs, in 1966 [1].

For the application to speech coding, bit rates for LP coefficients need to be compressed. In 1972, Dr. Itakura developed PARCOR[*1] coefficients to send information equivalent to LP coefficients with low bit rates while keeping the synthesis filter stable. A few years later, he developed LSP [2]–[4], which achieved better quantization and interpolation performance than PARCOR. A set of pth-order LSP parameters is defined as the roots of two polynomials $F_1(z)$ and $F_2(z)$, which consists of the sum and difference of $A(z)$ as

$$F_1(z) = A(z) + z^{-(p+1)} A(z^{-1}) \qquad (3)$$
$$F_2(z) = A(z) - z^{-(p+1)} A(z^{-1}) \; . \qquad (4)$$

The LSP parameters are aligned on the unit circle of the z-plane, and the angles of LSP, or LSP frequencies (LSFs), are used for quantization and interpolation. An example of 16th-order LSF values $\theta(1),\ldots, \theta(16)$ and the associated spectral envelope along the frequency axis are shown in **Fig. 1**. The synthesis filter is stable if each root of $F_1(z)$ and $F_2(z)$ is alternatively aligned on the frequency axis. It has been proven that LSP is less sensitive to the shape of a spectral envelope; that is, the influence of distortion due to quantization in LSP on the spectral envelope is smaller than it is with other parameter sets, including PARCOR and some variants of it. In addition, LSP has a better interpolation property than others. If we define LSP vector $\Theta_A = \{\theta(1),\ldots, \theta(P)\}$ corresponding to spectral envelope A, the envelope approximated by $envelope((\Theta_A + \Theta_B)/2)$ with LSP Θ_A and Θ_B can be a better approximation of the interpolated spectral envelope $(envelope(\Theta_A) + envelope(\Theta_B))/2$ than that with other parameter sets. These properties can further contribute to efficient quantization when they are used in combination with various compression schemes, including prediction and interpolation of LSP itself. These properties of LSP are beneficial for the compression of speech signals.

3. Progress of LSP

After the initial invention, various studies were carried out by Dr. N. Sugamura, Dr. S. Sagayama, Mr. T. Kobayashi, and Dr. Y. Tohkura [5] to investigate the fundamental properties and implementation of LSP. In 1980, a speech synthesis large-scale integration (LSI) chip (**Fig. 2**), was fabricated and used for real-time speech synthesis. Until that time, real-time synthesizers had required large equipment consisting of as many as 400 circuit boards. Note, however, that the complexity of the chip was still 0.1 MOPS (mega operations per second), less than 1/100 of the complexity of chips used for cellular phones in the 1990s.

*1 PARCOR (partial auto correlation): Equivalent parameter set of LP coefficients. PARCOR is advantageous in terms of its easy stability checks and better quantization performance than LP coefficients.

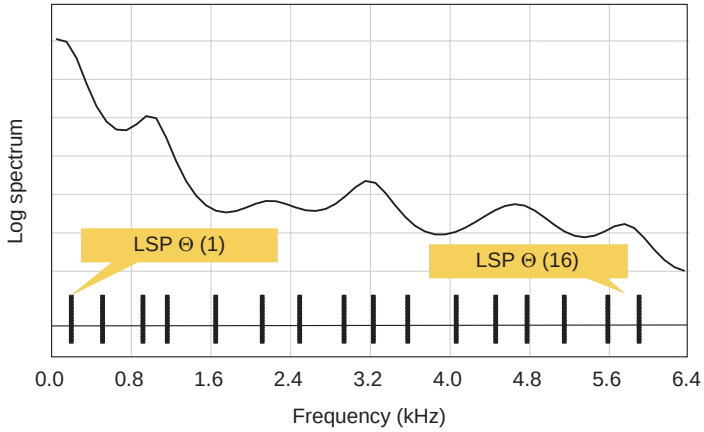


Fig. 1. A set of LSP frequency values and the associated spectral envelope in the frequency domain.

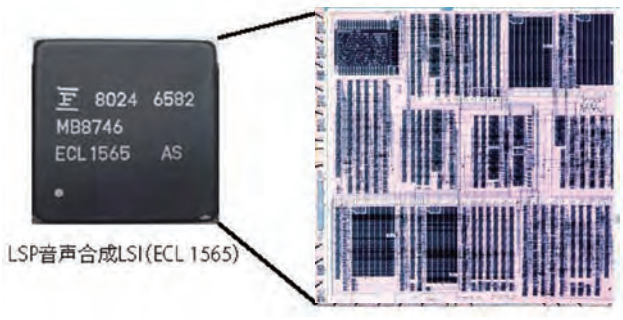


Fig. 2. LSI speech synthesis chip based on LSP in 1980.

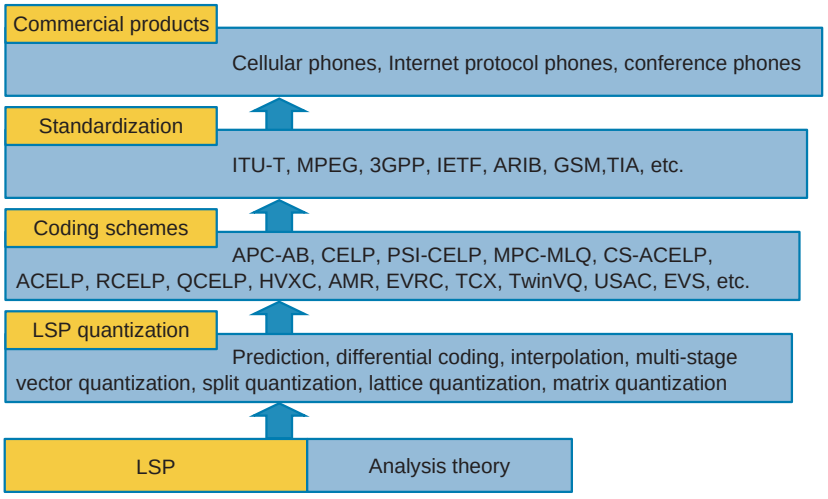
Around 1980, low-bit-rate speech coding was achieved with a vocoder scheme that used spectral envelope information (such as LSP) and excitation signals modeled by periodic pulses or noise. These types of coding schemes were able to achieve low-rate (less than 4 kbit/s) coding, but they were not applied to public communication systems because of their insufficient quality in practical environments with background noise. Another approach for low-bit-rate coding was waveform coding with sample-by-sample compression. However, it also could not provide sufficient quality below 16 kbit/s. In the mid 1980s, hybrid vocoder and waveform coding schemes, typically CELP[*2], were extensively studied; these schemes also need an efficient method for representing spectral envelopes such as LSP.

4. Promotion of LSP in worldwide standards

During the 1980s, however, the general consensus was that compression of speech signals would probably not be useful for fixed line telephony, and there was some doubt as to whether digital mobile communications, which requires speech compression, could easily be used in place of an analog system in the first generation. Just before 1990, however, new standardization activities for digital mobile communications were initiated because of the rapid progress being made in LSI chips, batteries, and digital modulation, as well as in speech coding technologies. These competitive standardization activities focusing on commercial products accelerated the various investigations underway on ways to enhance compression, including extending the use of LSP, as shown in **Fig. 3**. These investigations led to the publication of some insightful research papers, including one on LSP quantization by the current president of IEEE, Dr. Roberto de Marco [6].

In the course of these activities, LSP was selected for many standardized schemes to enhance the overall performance of speech coding. The major standardized speech/audio coding schemes that use LSP are listed in **Table 1**. To the best of our knowledge, the federal government of the USA was the first to adopt LSP as a speech coding standard in 1991. The Japanese Public Digital Cellular (PDC) half-rate

*2 CELP (code-excited linear prediction): Among large numbers of sets of excitation signals, the encoder selects the most suitable one that minimizes the perceptual distortion between the input and the synthesized signal with LP coefficients. This was initially proposed by AT&T in 1985 and has been widely used as a fundamental structure of low-bit-rate speech coding.



APC-AB: adaptive predictive coding with adaptive bit allocation
ARIB: Association of Radio Industries and Businesses
CS-ACELP: conjugate structure algebraic CELP
EVRC: Enhanced Variable Rate Codec
EVS: Enhanced Voice Service
GSM: Global Standard for Mobile Communications
HVXC: Harmonic Vector Excitation Coding
IETF: Internet Engineering Task Force
MPC-MLQ: Multipulse LPC with Maximum Likelihood Quantization
PSI-CELP: pitch synchronous innovation CELP
QCELP: Qualcomm CELP
RCELP: relaxed CELP
TCX: transform coded excitation
TwinVQ: transform-domain weighted interleave vector quantization
USAC: Unified Speech and Audio Coding

Fig. 3. Steps of technologies towards commercial products.

standard in 1993 may have been the first adoption of LSP for public communications systems; the USA and Europe soon followed suit. In 1996, two ITU-T (International Telecommunication Union-Technology Sector) recommendations (G.723.1 and G.729) were published with LSP as one of the key technologies. Both, but especially G.729, have been widely used around the world as default coding schemes in network facilities for Internet protocol (IP) phones. In 1999, speech coding standards for the third generation of cellular phones, which are still widely used around the world, were established by both 3GPP[*3] and 3GPP2[*4] with LSP included.

Furthermore, LSP has proven to be effective in capturing spectral envelopes not only for speech but also for general audio signals [7] and has been used in some audio coding schemes defined in ISO/IEC (International Organization for Standardization/International Electrotechnical Commission) MPEG-4 (Moving Picture Experts Group) in 1999 and MPEG-D USAC (Unified Speech and Audio Coding) in 2010.

5. Future communication

In the VoLTE[*5] service introduced in 2014 by NTT DOCOMO, 3GPP adaptive multi-rate wideband (AMR-WB) is used for speech coding, and it provides wideband speech (16-kHz sampling, the same speech bandwidth as mid-wave amplitude modulation (AM) radio broadcasting). For the next generation of VoLTE, the 3GPP Enhanced Voice Service (EVS) standard is expected to be used, which can

*3 3GPP (3rd Generation Partnership Project): Joint project for third-generation mobile communications by ETSI (European Telecommunications Standards Institute) and Japanese, Korean, and Chinese standardizing bodies. The activities are continuing and are focused on a fourth-generation system.
*4 3GPP2: Joint projects for third-generation mobile communication by the TIA and Japanese, Korean, and Chinese standardizing bodies.
*5 VoLTE: IP-based speech communication system over LTE mobile networks.

Table 1. Major standards with LSP.

Standardization body	Coding scheme	Bit rate (kbit/s)	Applications	year
Federal govt. of USA	FS1016 CELP	4.8	Govt. communication	1991
Federal govt. of USA	FS1017 MELP	2.4	Govt. communication	1995
Japan RCR (now ARIB)	STD-T27 PSI-CELP	3.4	2nd generation half-rate	1993
USA TIA/EIA	IS-95 RCELP	2,4,8	2nd generation half-rate	1995
Europe GSM	GSM-EFR	12.2	2nd generation enhanced full-rate	1997
ITU-T	G.723.1 MLP-MLQ/ACELP	5.3/6.3	TV (television) phone, IP phone	1996
ITU-T	G.729 CS-ACELP	8	IP phone Cellular phone (PDC)	1996
3GPP	AMR	12.2	3rd generation cellular phone	1999
3GPP2	EVRC	9.6	3rd generation cellular phone	1999
ISO/IEC MPEG-4	14496-3:2009 CELP/HVXC/TwinVQ	2–16	Speech/audio coding	1999
ISO/IEC MPEG-D	23003-3:2012 USAC	8–256	Speech/audio coding	2010
3GPP	AMR-WB	8–23	VoLTE	2001
3GPP	AMR-WB+	6–48	Speech/audio coding	2004
3GPP	EVS	5.9–96	VoLTE	2014

AMR-WB: adaptive multi-rate wideband
EIA: Electronic Industries Alliance
GSM EFR: GSM Enhanced Full Rate
MELP: mixed-excitation linear prediction
RCR: Research and Development Center for Radio Systems
TIA/EIA: Telecommunications Industry Association/Electronic Industries Alliance
VoLTE: voice over Long Term Evolution

handle a 32-kHz sampling rate signal and general audio signals. LSP or a variant of LSP is incorporated in both AMR-WB and EVS. In the near future, it may be possible to achieve all speech/audio coding functions with downloadable software. Even in such a case, we expect that LSP will still be widely used.

In this way, LSP may be a good example of technology that has contributed to the world market. The NTT laboratories will continue to make efforts to enhance communication quality and the quality of services by meeting challenges in research and development.

References

[1] B. S. Atal, "The History of Linear Prediction," IEEE Signal Processing Magazine, Vol. 23, No. 2, pp. 154–157, March 2006.
[2] F. Itakura, "Line Spectrum Representation of Linear Predictive Coefficients of Speech Signals," J. Acoust. Soc. Am., Vol. 57, S35, 1975.
[3] F. Itakura, "All-pole-type Digital Filter" Japanese patent No. 1494819.
[4] F. Itakura, "Statistical Methods for Speech Analysis and Synthesis From ML Vocoder to LSP through PARCOR," IEICE Fundamentals Review Vol. 3, No. 3, 2010 (in Japanese).
[5] F. Itakura, T. Kobayashi, and M. Honda, "A Hardware Implementation of a New Narrow to Medium Band Speech Coding," Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 1982, pp. 1964–1967, Paris, France, May 1982.
[6] J.R.B. de Marca, "An LSF Quantizer for the North-American Half-rate Speech Coder," IEEE Trans. on Vehicular Tech., Vol. 43, No. 3, pp. 413–419, August 1994.
[7] N. Iwakami, T. Moriya, and S. Miki, "High-quality Audio-coding at Less Than 64 kbit/s by Using TwinVQ", Proc. of ICASSP 1995, pp. 3095–3098, Detroit, USA, May 1995.

Takehiro Moriya
NTT Fellow, Moriya Research Laboratory, NTT Communication Science Laboratories.

He received his B.S., M.S., and Ph.D. in mathematical engineering and instrumentation physics from the University of Tokyo in 1978, 1980, and 1989, respectively. Since joining NTT laboratories in 1980, he has been engaged in research on medium- to low-bit-rate speech and audio coding. In 1989, he worked at AT&T Bell Laboratories, NJ, USA, as a Visiting Researcher. Since 1990, he has contributed to the standardization of coding schemes for the Japanese PDC system, ITU-T, ISO/IEC MPEG, and 3GPP. He is a member of the Senior Editorial Board of the IEEE Journal of Selected Topics in Signal Processing. He is a Fellow member of IEEE and a member of the Information Processing Society of Japan, the Institute of Electronics, Information and Communication Engineers, the Audio Engineering Society, and the Acoustical Society of Japan.

Report on NTT Communication Science Laboratories Open House 2014

Kazuo Aoyama, Hidehisa Nagano, Shin Mizutani, Takaaki Tanaka, and Naotoshi Abekawa

Abstract

Open House 2014 was held at NTT Communication Science Laboratories in Keihanna Science City, Kyoto, on June 5 and 6. Over 1000 people visited the facility to enjoy 6 talks and 29 exhibits introducing our latest research activities and efforts in the fields of information and human sciences.

Keywords: information science, human science, big data

1. Overview

At NTT Communication Science Laboratories (NTT CS Labs), we aim to build a new technical infrastructure connecting *people* and *information*, and we are studying aspects of both human and information sciences to create innovative technologies and discover new principles. NTT CS Labs deals with the most fundamental research targets in the fields of human and information sciences of NTT laboratories; our facilities are located in Kansai Science City (Seika-cho, Kyoto) and Atsugi City, Kanagawa.

NTT CS Labs Open House has been held annually with the aim of introducing the results of the labs' basic research and innovative leading-edge research to both NTT Group employees and visitors from industries, universities, and research institutions who are engaged in research, development, business, and education.

This year, Open House was held at the NTT Keihanna Building in Kyoto on the afternoon of June 5 and all day on June 6, and a total of around 1070 visitors attended it over the two days. At this event, we prepared many hands-on exhibits to allow visitors to intuitively understand our latest research results and to share a vision of the future where new products based on the research results are widely used. We also organized an invited talk on the theme of philosophy, which differs from our research fields, so that people in various fields could take an interest in our activities. This report summarizes the event's research talks and exhibits.

2. Keynote speech

Open House started with a speech by the Director of NTT CS Labs, Dr. Eisaku Maeda, entitled "Basic research: Defining our age and the future—The origin of ideas and the seeds of innovation" (**Photo 1**).

It has been 14 years since NTT CS Labs was

Photo 1. Dr. Eisaku Maeda, Director of NTT CS Labs, giving keynote speech.

Photo 2. Research presentation by Dr. Makio Kashino.

Photo 3. Research presentation by Dr. Keisuke Kinoshita.

reorganized as the sole basic research institution for human and information sciences under the new NTT research and development (R&D) scheme instituted by the reorganization of NTT in 1999. During this time, NTT CS Labs has established a framework that has enabled the gradual penetration into the world of the technologies developed from their basic research. Dr. Maeda considered this time to be a major turning point for the next step of NTT CS Labs, and he talked about how important it is that basic research tackles problems that have to be selected and solved in line with the times. He also stressed the importance of launching products onto the market based on a sense of urgency required by the times. He then introduced two representative products, i.e., the seeds of innovation: an extremely large vocabulary recognition decoder that is a core part of a minute-taking system used at the House of Representatives, and the Robust Media Search (RMS) technology for music copyright clearance that has been adopted by broadcasters. He also demonstrated Buru-Navi3, the newest version of a haptic compass, which utilizes human sensory properties to induce a pseudo-attraction force. It was successfully miniaturized through a consistent research effort, which is the "origin of ideas" he refers to in the title of his speech.

3. Research talks

There were three talks, which highlighted recent significant research results and high-profile research themes.

- "Mind changes body, body leads mind—Feasibility and potential of mind-reading technology," by Dr. Makio Kashino, Human and Information Science Laboratory. Dr. Kashino introduced the basic concepts and some examples of mind-reading technology that is based on methods of decoding the implicit mind from involuntary body movements and physiological responses (**Photo 2**).
- "Enhancing speech quality and music experience—Opening up new vistas for audio experience with reverberation control technology," by Dr. Keisuke Kinoshita, Media Information Laboratory. Dr. Kinoshita clarified both positive and negative aspects of reverberation in audio signal processing, presented his lab's novel reverberation control technology that can be used to enhance speech and music experiences, and explained how they are actually utilized in the market (**Photo 3**).
- "Quantum computing beyond integer factorization—Exploring the potential of quantum search," by Dr. Seiichiro Tani, Innovative Communication Laboratory. Dr. Tani reviewed the theory and applications of quantum search to show that the potential advantages of quantum computing are not limited to integer factorization, and he emphasized the importance of research on quantum algorithms (software) as well as quantum computers (hardware), by comparing them to our conventional computing systems (**Photo 4**).

Each presentation introduced some of the latest research results and provided some background and an overview of the research. All of the talks were very well received.

Photo 4. Research presentation by Dr. Seiichiro Tani.

4. Research exhibits

Open House featured 23 exhibits displaying NTT CS Labs' latest research results. They were classified into three categories: computer science, media intelligence, and communication and human science. Open House also included a special category on big data science that consisted of three exhibits from NTT CS Labs and three from other NTT laboratories.

Each exhibit was housed in a booth and employed techniques such as slides on a large-screen monitor or hands-on demonstrations, with researchers explaining the latest results directly to visitors. The following list summarizes the research exhibits in each category.

4.1 Big Data Science

- Finding latent relationships between different data sets—Unsupervised object matching
- Extracting common patterns from multiple data sets—Non-negative multiple matrix factorization: NM2F
- Knowledge discovery from large-scale graph data—Efficient graph clustering and distributed query optimization
- Smarter and instant analysis for huge amounts of video—Jubatus, a scalable big data real-time analysis framework
- Optimizing network operation through NW data analytics—Inferring latent network status through machine learning
- Monitoring remote habitats of endangered species—Online environment monitoring with a wireless sensor network

4.2 Computer Science

- Formal evaluation of network security—Verification of cryptographic protocols using formal methods
- Generating a common secret based on bounded observability—Secret key distribution using broadband random light
- Opening the possibility of realizing quantum computers—Constant-step quantum circuits can compute the OR function
- Programming for everyone—Introduction to computer programming in VISCUIT
- Let's chat with a computer!—Dialogue system with various utterance generation methods
- Reordering Japanese for better translation—Translation using Japanese predicate argument structure
- Generating coherent summaries from documents—Document summarization by discourse tree trimming

4.3 Media Intelligence

- Single frame level detection from dailies—Media search-based collaboration system for movie production
- Retrieving video immediately with camera shots—Instance search for specific objects in movies
- You may know the lion by its TWO claws—Image matching based on affine-invariant spatial context
- Understanding multimedia content without seeing it—User behavior reveals meaning of multimedia content
- Capturing sound by light—Towards massive-channel audio sensing via LEDs and a camera
- Making computers listen to desired sounds anywhere—Probabilistic modeling and integration for speech enhancement
- Defeat reverberation: enemy of speech recognition—Advanced speech enhancement and recognition
- How accurate are speech recognition results?—Estimating speech recognition accuracy without references

4.4 Communication and Human Science

- Feeling conversation in motion—Recreating a conversation space with augmented body motions
- You may feel your eyes meet with others—Determining factors related to the eye-contact

Photo 5. Exhibit: "Capturing sound by light."

Photo 6. Exhibit: "I've got the knack!"

perception area
- Reading mind from body—Body movements and physiological responses reveal emotions
- I've got the knack!—Visualization and sonification of action in sports
- Seeing materials from image movements—Motion-based liquid perception by human vision
- What determines a person's hearing performance?—Exploring sources of inter-individual variation
- Texture integration in touch—Integration process of tactile perception
- Buru-Navi3: tiny but powerful sensation of being pulled—Asymmetric oscillation induces clear kinesthetic illusion

The exhibit "Capturing sound by light" introduced a new multichannel system that is capable of capturing audio signals of huge microphone arrays via LEDs (light-emitting diodes) and a high speed video camera. The system achieves real-time performance by massive parallel processing with a single graphics processing unit. Visitors to this exhibit were able to experience demonstrations in which an observed sound space was quickly rebuilt with realistic sensations (**Photo 5**).

The "I've got the knack!" exhibit introduced effective ways of providing motor information, i.e., motor feedback techniques, to facilitate motor learning in sports. These feedback techniques were designed to visualize and convert to sound certain key features of an action. They can be applied not only in the field of sports but also in the rehabilitation and entertainment fields (**Photo 6**).

5. Invited talk

This year's event also featured two invited talks by Associate Prof. Masaya Chiba, Graduate School of Core Ethics and Frontier Sciences, Ritsumeikan University, and Prof. Shin-ichi Minato, Graduate School of Information Science and Technology, Hokkaido University. The titles of their respective talks were "Considering the information society in terms of 'After post-structuralism'" and "Science and engineering of discrete structure manipulation based on 'Power of enumeration.'"

Associate Prof. Masaya Chiba began with a brief review of his research field. Then he described the concepts of structuralism and post-structuralism stemming from 20th-century French philosophy, and provided a simple explanation of the discussions and considerations arising after post-structuralism in relationship to social networks and communications. He suggested that in contemporary society, we need to appropriately control connection and disconnection with other people since we have lapsed into a state of excess connection due to the always-on Internet connection.

Prof. Shin-ichi Minato talked about his concept of discrete structure manipulation. In the first half of his talk, he simply explained the technical aspects of the concept using some examples of set expressions and set operations based on the zero-suppressed decision diagram (ZDD) that he developed. In the second half, he explained how discrete structure manipulation was applied to actual problems. He focused on enumeration problems and discussed detailed applications of ZDD as shown in a popular and interesting YouTube video called "Don't count naively" produced by his group. He also discussed the Graphillion software

(a) Official home page (b) Webpage of an exhibit

Fig. 1. Website of NTT CS Labs' Open House 2014.

package for search, optimization, and enumeration for large-scale graphs, and optimization of power distribution networks.

6. Information transmission using web

NTT CS Labs has made continuous efforts to inform a large number of people both domestically and internationally about their research activities and results. As part of the ongoing effort, we simultaneously released both Japanese and English websites [1, 2] for Open House 2014, which included a booklet, exhibition posters, and reference information (**Fig. 1**). We plan to upload photos of the presentation and exhibition halls and videos of the director's keynote speech and the three research talks.

The members of NTT CS Labs consider it important to transmit high-quality and attractive information on a timely basis. They also have a clear awareness that it is important to convey information on the professional research activities that are underway at the basic research laboratories to a wide range of people in ways that are easy to understand. Consequently, we simply tweeted about the exhibition contents via Twitter through NTT's Public Relations Office, and we gathered feedback by putting up a questionnaire page on the Open House 2014 websites. NTT CS Labs is continuously trying to improve the ways of disseminating our research activities and results.

7. Concluding remarks

Just as they did last year, many visitors came to NTT CS Labs' Open House 2014 and engaged in lively discussions on the research talks and exhibits and provided many valuable opinions on the presented results. In closing, we would like to offer our sincere thanks to all of the visitors and participants who attended this event.

References

[1] Website of NTT Communication Science Laboratories Open House 2014 (in Japanese). http://www.kecl.ntt.co.jp/openhouse/2014/index.html
[2] Website of NTT Communication Science Laboratories Open House 2014 (in English). http://www.kecl.ntt.co.jp/openhouse/2014/index_en.html

Kazuo Aoyama

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

He received a B.E. in applied physics from Waseda University, Tokyo, in 1986 and an M.E. from Tokyo Institute of Technology in 1988. He joined NTT in 1988 and worked on device modeling of various types of MOSFETs for circuit simulation and studied reconfigurable computer architectures and circuit design methodology. He is currently engaged in research on data structures and algorithms. He is a member of the Institute of Electrical and Electronics Engineers (IEEE), the Institute of Electronics, Information and Communication Engineers (IEICE), the Information Processing Society of Japan (IPSJ), and the Japan Society of Applied Physics.

Hidehisa Nagano

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

He received the B.Eng. and M.Eng. in information and computer sciences in 1994 and 1996, respectively, and the Ph.D. in information science and technology in 2005, all from Osaka University, Japan. He joined NTT in 1996. From 2011 to 2012, he was a visiting researcher at the Centre for Digital Music, Queen Mary University of London, UK. He has been working on audio and video analysis, search, retrieval, and recognition algorithms and their implementation. He is a senior member of IEEE and a member of IEICE and IPSJ.

Shin Mizutani

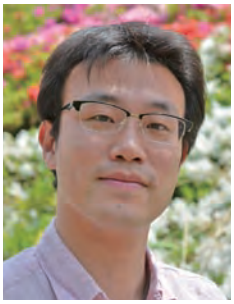
Research Scientist, Learning and Intelligent Systems Research Group, Innovative Communication Laboratory, NTT Communication Science Laboratories.

He received the B.E. in physics from Tokyo University of Science in 1990, the M.E. in biophysical engineering from Osaka University in 1992, and the Ph.D. from Osaka University in 1999. He joined NTT Human Interface Laboratories in 1992 and studied neural networks and nonlinear science to understand the information processing mechanisms in the nerve system. He has been a researcher at NTT Communication Science Laboratories since 1998.

Takaaki Tanaka

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

He received the B.E. and M.E. from Osaka University in 1994 and 1996, respectively. He joined NTT Communication Science Laboratories in 1996 and studied natural language processing including bilingual knowledge acquisition for machine translation. He stayed at the Center for the Study of Language and Information (CSLI) at Stanford University as a guest researcher in 2002 to study multiword expressions such as idioms and compound nouns. He is currently studying deep parsing in Japanese, which integrates syntactic and semantic information. He is a member of the Association for Computational Linguistics, the Association for Natural Language Processing, and the Japanese Society for Artificial Intelligence.



Naotoshi Abekawa

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

He received the B.E. from Tokyo Metropolitan University in 2003, the M.E. from Tokyo Institute of Technology in 2005, and the Ph.D. from Kyoto University in 2014. He joined NTT in 2005 and has been engaged in research on human information processing. His research interests include human sensorimotor mechanisms, especially visuomotor control properties. He is a member of the Society for Neuroscience, the Japan Neuroscience Society, and the Japanese Neural Network Society.

Papers Published in Technical Journals and Conference Proceedings

Visual SyncAR: Augmented Reality which Synchronizes Video and Overlaid Information

S. Yamamoto, H. Tanaka, S. Ando, A. Katayama, and K. Tsutsuguchi

The Journal of the Institute of Image Electronics Engineers of Japan, Vol. 43, No. 3, pp. 397–403, July 2014.

We developed video synchronized augmented reality technology called Visual SyncAR. The novel technology is characterized by the ability to superimpose three-dimensional computer graphic (CG) animations that are synchronized with the timing of the video captured by the smartphone camera. By embedding time codes into the video's digital watermarks in advance and detecting these watermarks at a reliably high speed, it is possible to synchronize the video and the CG animation. This enables unprecedented visual expression for mobile devices that jumps out beyond the *fourth wall*, a term used in the theatrical world.

Programming Education by Non-professionals in the Extra-curricular Activities of Public Elementary Schools

Y. Harada, N. Katsunuma, and Y. Kuno

Journal of Information Processing Society of Japan, Vol. 55, No. 8, pp. 1765–1777, August 2014.

Recently, computer programming has been recognized as an important subject for elementary schools in the next decade. Even in Japan, various activities that involve teaching programming to elementary school children are in progress. However, most of those activities are taught by professional programming educators, with support from many assistants; such activities cannot simply be ported as-is to public schools. In this paper, we introduce our experiences with the *Viscuit* educational programming language in Midorikko-club, a club for after-school activities held at Midori elementary school, in Sumida-ku, Tokyo. The activities are held for groups of children and have been led by a small number (one or two) of local non-professional volunteers for more than three years. Thanks to the novel teaching method and the support functions of the Viscuit system, both described in this paper, the activities have been very successful. We evaluated the results by conducting interviews with some of the children who took part. The results indicated that (1) children have come to recognize that computers are not just boxes, but are something on which they can create interesting things, and (2) they have come to enjoy teaching various programming techniques to each other, thus realizing cooperative learning.

Predicate-argument Structure Analysis with Zero-anaphora Resolution for Dialogue Systems

K. Imamura, R. Higashinaka, and T. Izumi

Proc. of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, pp. 806–815, Dublin, Ireland, August 2014.

This paper presents predicate-argument structure analysis (PASA) for dialogue systems in Japanese. Conventional PASA and semantic role labeling have been applied to newspaper articles. Because pronominalization and ellipses frequently appear in dialogues, we base our PASA on a strategy that simultaneously resolves zero-anaphora and adapts it to dialogues. By incorporating parameter adaptation and automatically acquiring knowledge from large text corpora, we achieve a PASA specialized for dialogues that has higher accuracy than that for newspaper articles.

Flexible and Robust Optical Network Technologies for SDN and Network Virtualization

T. Tanaka

Proc. of the 12 International Conference on Optical Internet (COIN 2014), FA3-2, Jeju, Korea, August 2014.

This paper reviews the progress of flexible and robust technologies for a reliable optical network infrastructure that can be used for software-defined networking (SDN) and network virtualization. We focus on the elastic optical network (EON), which will play an important role in achieving both flexibility and robustness, and introduce our current works on fault tolerance.

Flow-based User Pairing Scheme for Multi-user Transmissions over WLANs

Y. Inoue, S. Shinohara, M. Mizuguchi, and M. Morikura

Proc. of the 11th IEEE Vehicular Technology Society Asia Pacific Wireless Communications Symposium (APWCS 2014), Ping Tung, Taiwan, August 2014.

To enhance system capacity and spectrum efficiency, multi-user transmission techniques such as MU-MIMO (multi-user multiple-input multiple-output) and OFDMA (orthogonal frequency-division multiple access) have been extensively studied and adopted in recent wireless standards such as IEEE 802.16e and IEEE 802.11ac. In order to achieve high efficiency of multi-user transmissions, it is important to consider characteristics of the traffic flow when determining the users group for the multi-user transmissions. In this paper, a flow based user pairing scheme for multi-user transmissions over WLANs (wireless local area networks) is proposed. The proposed scheme collects characteristics of traffic flows by using an existing admission control mechanism. Those parameters are considered in order to determine the users group for multi-user transmissions to achieve higher efficiency. The effect of the multi-user transmissions with the proposed pairing scheme was evaluated by computer simulations. We also discuss the effect of pairing policies of the user selection process.

Adaptive Sampling for Estimating True Values in Participatory Sensing Environment

H. Kurasawa, H. Sato, A. Yamamoto, H. Kawasaki, M. Nakamura, H. Matsumura, Y. Yamashita, M. Suzuki, and H. Morikawa

Journal of Information Processing Society of Japan, Vol. 55, No. 9, pp. 2190–2202, September 2014.

Participatory sensing causes huge network traffic in spatiotemporally dense areas where participants upload large numbers of sensor values. Furthermore, sensor values of participatory sensing are often inaccurate due to both measurement error and variations in the sensing accuracy of devices. Thus, if we control the network traffic

regardless of the accuracy of sensor values, the measurement results will be unreliable. We propose a method to control the sampling rate of devices while maintaining high reliability of the measurement results. This method involves calculating a sufficient number of sensor values to estimate the true value on the basis of the confidence interval for a population mean of sensor values from a spatiotemporal area. The same sampling rate is then set for all sensing devices in the area based on that amount, and network traffic is reduced.

Privacy-preserving Statistical Analysis Method by Splitting Server Roles for Distributed Real-world Data

J. Ishii, H. Maeomichi, A. Tsutsui, and I. Yoda

IEICE Transactions on Communications, Vol. E97-B, No. 9, pp. 1779–1789, September 2014.

This paper proposes a novel method for obtaining statistical results such as averages, variances, and correlations without leaking any raw data values from data-holders by using multiple pseudonyms. The authors split the roles of servers into publishing pseudonyms and collecting answers. Splitting these roles enables different entities to more easily join as pseudonym servers than in previous secure multiparty computation methods, and there is less chance of collusion between servers. We also estimated a typical problem that occurred with our method and added a pseudonym availability confirmation protocol to prevent the problem. We report our evaluation of the effectiveness of our method through implementation and experimentation. Finally, we explain how our method can obtain averages, variances, and correlations from 5000 data holders within 50 seconds.

An Access Control Model Based on Feature Structure

K. Fujita and Y. Tsukada

Journal of Information Processing Society of Japan, Vol. 55, No. 9, pp. 2081–2091, September 2014.

In this paper, we report on access control because it is the underlying core technology to enforce security and privacy. Access control determines whether packets are permitted or denied according to access control policies. Since the notations of policies are specialized in each system, it is difficult to ensure consistency of policies that have different notations. In this paper, we propose a descriptive notation for policies by adopting the concept of feature structures, a technique which has mainly been used for parsing in natural language processing. Our proposed notation is also logically well-founded, which guarantees strict access control decisions, and expressive in that it returns not only a binary value of permit or deny but also various result values through the application of partial order relations of the security risk level. We illustrate the effectiveness of our proposed method using examples from P3P (Platform for Privacy Preferences).
