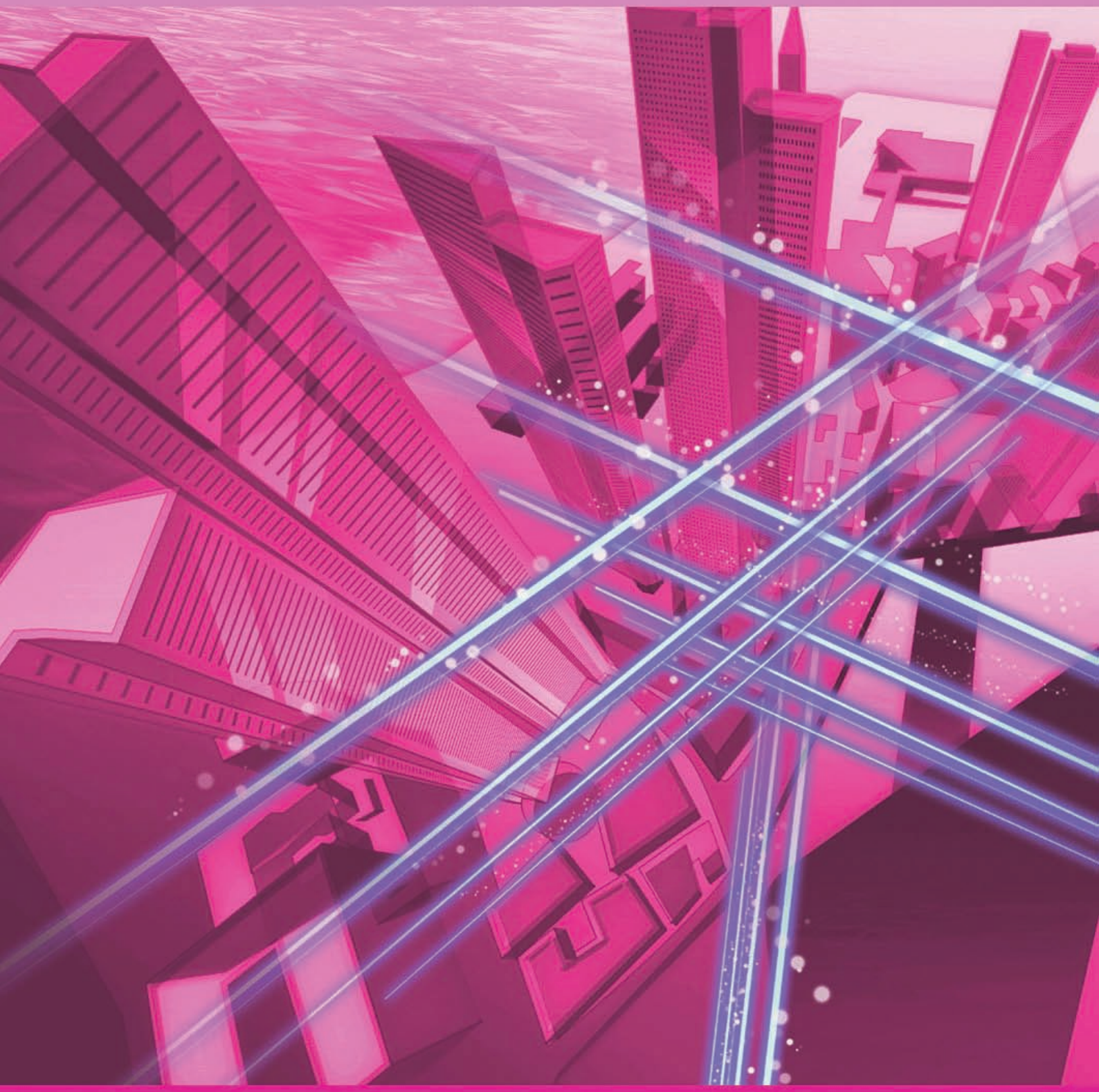
NTT Technical Review

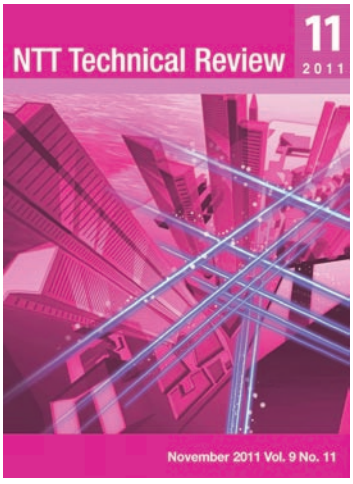
11
2011



November 2011 Vol. 9 No. 11

NTT Technical Review

November 2011 Vol. 9 No. 11



Refining Our Business, Technical, and Personnel Capabilities during Normal Times to Build Reliable Facilities Robust to Disasters

Kouichi Takahatake
Executive Vice President, NTT WEST

Overview

Just as NTT WEST had begun to formulate a new network vision for the future, the Great East Japan Earthquake of 2011 struck the Tohoku region of Japan. What light has this unprecedented disaster shed on the mission of NTT WEST? We sat down with Kouichi Takahatake, Executive Vice President of NTT WEST, to hear how personnel development and a new perspective on network facilities are essential to building a network robust to disasters.

Personnel development as a foundation for the future

—Mr. Takahatake, could you first brief us on the business conditions surrounding NTT WEST?

Of course, but first let me start with some historical background. When NTT was reorganized into NTT Holding Company, NTT EAST, NTT WEST, and NTT Communications in 1999, only NTT WEST was losing money. I suspect that at that time many people had no wish to work at a company engulfed in red figures, but it's precisely because NTT WEST was in the red that I wanted to work there. I have worked there since then, and I feel that I have grown in parallel with the company's growth.

NTT WEST continued to be in the red from its founding in 1999 until 2002, and while business conditions then recovered owing to a restructuring, they gradually headed downward again with the trend toward decreasing revenues, and profits hit bottom about two years ago. In the last couple of years, however, we have been able to turn things around and show an increase in earnings since 2010. For fiscal year 2011, our objective is to achieve a net increase of 850,000 FLET'S Hikari subscribers and make a single-year profit in our optical connection services.

—What would you say were the main reasons for this turnaround in the last few years?

Well, the first recovery that I mentioned came about because of restructuring, but I would attribute this recent turnaround to nothing short of corporate abilities. Specifically, the construction of an optical infrastructure and sales of optical services have had a great impact in addition to the fact that all employees have played their part in reducing costs. I believe that fiscal year 2011 will mark a new chapter in the history of NTT WEST.

—What were your specific strategies for making a single-year profit in optical connection services?

My primary strategy was to enhance personnel development. In this regard, I mention here that my

duties in a previous position as manager of Personnel Department No. 2 included the duties of a personnel-development manager as well as the ordinary duties of a personnel manager. At that time, NTT WEST had two major objectives: completely migrating from the telephone era to the Internet protocol (IP) era and achieving a turnaround from red to black in company earnings. Out of these objectives came a simple, straightforward business policy.

Let me explain. Achieving these objectives was directly connected to people and skills. In other words, I felt that these objectives could be achieved by enriching our human resources. For example, making the shift from conventional telephone services to IP required personnel and engineers skilled in both technologies. Thinking that we had to raise the existing skills of our personnel to a higher level, I therefore embarked on a path of personnel development. Specifically, I created a system for training engineers, a mechanism for educating employees, by analyzing and organizing the skills that would be needed and assembling them into a booklet as a training program.

Advanced IP engineers currently number more than 3700. This figure in itself reflects the extent of personnel training at NTT WEST. My plan is to continue with these educational and training activities so that we can continue to produce highly knowledgeable and professional personnel.

Looking back, NTT WEST has come to grow on the basis of three strategies: (1) business strategy, or the company's business policy; (2) technology strategy, or the technology required to execute the business strategy; and (3) human resources strategy, or the training of personnel to execute the technology strategy. Specifically, we established the telephony-to-IP transition as a business policy, clarified the technologies needed to execute that policy, and proceeded to develop personnel knowledgeable in those technologies and assign them appropriately. NTT WEST has grown by creating a well-balanced tripod formed by these three strategies. This process took patience, but taking one step at a time led to the present results.

Up-close communication training

—What is the feeling that drives your step-by-step approach to leadership?

As a general rule, the bigger an organization becomes, the slower it moves. Needless to say, NTT

WEST is not a small company. With such a large staff, I believe that nothing will change unless I take it upon myself to get the ball rolling. To this end, I think along the lines of *my future lies in continual change, and now is the time to decide my future.* Doing something on the basis of what someone else says is not interesting at all; what's important is to take action on one's own.

To give an example, I hold a morning meeting every day. When I first became a section chief about 25 years ago, I made it a practice to hold a face-to-face meeting with my staff every morning except when I could not meet them because of a business trip or other reasons. Since we are a communication service company that places great importance on communication, it made sense that we ourselves should have superb communication skills.

Accordingly, to stimulate communication as well as to provide practical training in it, I arranged that the meeting's participants could exchange information effectively by feeling each other's presence up close. To be sure, a video conference can also provide a means to communicate, but only after one has mastered direct, face-to-face communication, will it become possible to recognize those aspects of video conferencing that prevent participants from developing a real feel for each other. Without this sense of contact, I think that communication can be difficult.

However, no matter how intimate communication may be or how confident speakers may be that they have communicated well, it is not unusual that their intentions are still not conveyed exactly as they would wish. That is why I have made it a practice to hold a meeting every morning. Some people might think that during very busy periods, this simply adds an additional burden, but I believe that it works. The

participants become able to sense if someone has a problem just by sitting together at the meeting.

Migrating existing facilities using lessons learned from the Great East Japan Earthquake

—This year, Japan experienced an unprecedented disaster in the Great East Japan Earthquake. Wasn't this an event of great significance for NTT WEST too?

It certainly was—this event provided many lessons for us. About two years ago, we began to consider what we could do to enhance *people-to-people* communications, to promote the evolution of *people-to-things* communications, and to broaden and deepen the business possibilities of *people-to-neighborhood* communications. In any case, it's people-to-people communications that all forms of communication revolve around. This has been NTT's core axis up to now, but to promote the evolution of communications, we felt the need to modify somewhat our approach to business and the way we configured our project teams. For example, we had been thinking that the need to form alliances with companies in a variety of fields was becoming stronger than ever—and then the earthquake struck. Given the unexpected scale of the resulting damage, this event revealed our strengths and weaknesses for dealing with emergencies.

The Great East Japan Earthquake was of the plate-boundary (ocean) type as opposed to an inland type.

It featured slow oscillations and generated a tsunami as well. As a result, the damage it caused was of a different type to that of the Great Hanshin-Awaji Earthquake of 1995. In Iwate prefecture, Rikuzentakata was one city that suffered much damage from the earthquake, but the NTT building somehow survived. Amid wide-spread devastation, the building remained standing! When I visited this building, the feeling that arose in me was that all of our past research efforts at NTT had not been made in vain. I walked up to the roof, and while all I could do was put my hands together in a gesture of prayer as I surveyed the surrounding damage, I instantly realized that having strong NTT buildings was the starting point for migrating to future facilities robust to disasters.

It goes without saying that NTT buildings should be constructed in safe places. But there are some situations in which all requirements cannot be met. In these cases, it is necessary to consider measures appropriate for the environment in question, such as installing covers on windows. On assessing the damage from this disaster, we were able to clearly determine the weak points of each type of facility and region. It is my desire to reflect all of the lessons that we learned here in future facilities.

—Could you tell us how you plan to achieve this migration of existing facilities?

The plans that we conceived before the earthquake and the lessons we learned from it indicate that we should convert the network to a full-IP system that can provide a wide array of services. We should also aim for all-optical access and further cost reductions. We also feel that an NTT exchange (building) should be more than just a datacenter; it should also provide value in the form of a *smart community center* that acts as a base for the entire district.

This does not mean that new facilities will be created while completely ignoring existing ones. Though easy to talk about, determining how to make good use of facilities that we have deployed to date, how to merge them with new facilities, and what to change about them form a very difficult problem.

There is another valuable lesson that we learned here: it's important that we refine our skills as much as possible during normal times. In this way, every employee will be able to demonstrate his or her full capabilities at the time of a crisis such as the one we recently experienced. It's inevitable that crisis management will tend to slacken off when peaceful conditions continue. But the skills and strengths that

can be manifested fully only at the time of an emergency can also be tapped in everyday work, and it is with this in mind that I wish to be mentally prepared through ongoing refinement of our individual skills. Now is the time to decide our future.

—Mr. Takahatake, could you leave us with some words on what you expect from the R&D department?

First of all, I would like to express my deep respect for NTT's research and development (R&D). Although I did not come out of the laboratories, I have always been interested in people-to-people communications and have often thought that perhaps that was a technology that I should have taken up. NTT researchers study the five human senses and research many interesting and useful themes such as how to read emotions. While I was manager of the NTT WEST Kumamoto Branch, I sometimes took important customers to NTT laboratories to show them some of this cutting-edge, creative research and introduce NTT's technical expertise to them. Thanks to the energetic efforts of everyone at NTT laboratories, we have been able to increase the number of NTT fans. I also remember how NTT researchers stayed up all night at the time of a major problem with the Hikari Denwa optical phone service to find the cause. I sincerely appreciate the way that NTT laboratories support the daily business activities of NTT Group companies.

Finally, I would say to everyone in NTT's laboratories that I expect you to continue being world leaders in a variety of R&D fields.

Interviewee profile

■ Career highlights

Kouichi Takahatake graduated from the Department of Electrical Engineering and Computer Science, School of Engineering, Kyushu University in 1978 and entered Nippon Telegraph and Telephone Public Corporation (now NTT) in the same year. At NTT WEST, he served as Manager of the NTT WEST Kumamoto Branch, Manager of Network Operations, Senior Vice President and Executive Manager of the Service Management Department, and Senior Vice President and Executive Manager of the Network Department before taking up his present position in June 2011.

Building a Theory of Communication

Naonori Ueda†

Abstract

At NTT Communication Science Laboratories, we are conducting basic research in communication science with the aim of making telecommunication as accurate and satisfying as face-to-face communication between people. The Feature Articles in this issue describe our research efforts aimed at building a theory of communication as a foundation for communication science and introduce our latest achievements.

1. Introduction

What is information? A well-known answer to this question was provided by the information theory proposed by Claude Shannon in 1949. Shannon laid the foundations of information theory by devising mathematical definitions for quantities of information and entropy. Needless to say, the subsequent expansion of this theory led to the astonishing speed and accuracy of today's information and communications technology. But what is communication? There are limits to how far information theory can currently go towards answering this question. This is because the quality of communication is evaluated only in terms of the probability of the occurrence of an event, which defines the amount of information that it conveys. When humans communicate with each other, the important thing is not how much information is conveyed, but the intrinsic value of this information. In other words, instead of quantitative objective measures such as speed and accuracy, greater importance is attached to qualitative subjective measures like satisfaction. In situations such as job interviews, greater importance has recently been attached to more abstract qualities such as the ability to gain a deep understanding of the other person's intentions and provide accurate answers rather than the ability to speak fluently.

To deliver a rich communication environment, it seems that simply researching information itself is not enough; we must also study the humans who are actually sending and receiving this information. At NTT Communication Science Laboratories (CS Labs), in order to pursue the essential qualities of communication via the dual aspects of computer science and human science, we are conducting basic research on three fronts—creating the communication environment of the future, establishing an intelligent computing platform, and delivering a rich quality of life (QoL) for humans. Specific examples of the work that we have recently been doing in these areas are described below.

2. Creating the communication environment of the future

We are constructing a remote communication system called the t-Room with the aim of implementing more natural video communication in the future, and we are developing middleware to achieve a sense of *co-presence* and analyzing and evaluating remote collaborative work in this environment. Here, co-presence refers to the perception and recognition of people and objects that exist in the room being mutually shared with people in the remote room in order to provide video communication in a more natural way in the future. We are also conducting research aimed at implementing a system that allows people to communicate in a natural way. Recently, we have been developing technology that automatically learns interaction control (back-channeling and questioning) from interactions between humans and from interactions between humans and systems.

We are also building a system called the s-room, which recognizes people's actions if they simply

† NTT Communication Science Laboratories
Soraku-gun, 619-0237 Japan

wear a camera and acceleration sensor on the wrist. We intend to use this system in applications such as analyzing human actions and watching over the elderly.

3. Establishing fundamental technologies for intelligent computing

Most of the computer science research at the CS Labs is aimed at establishing fundamental technologies for intelligent computing and extends across a wide range of research fields. The main technical fields are introduced below. Techniques for finding specific information from vast quantities of information will become increasingly important in the future.

At the CS Labs, we have spent over ten years working on media search techniques that can rapidly find a desired item of audio or video content from a large collection of material. As far back as 1998, we had developed technology capable of searching six hours of video content in just two seconds (allowing us to find a particular TV commercial from videos containing thousands of commercials). By 2010, we were able to search 60,000 hours of content in one second, and we developed a robust media search technique capable of performing searches very accurately even if the audio and video media are degraded by noise. This technique has been put to practical commercial use in a system for detecting copyright-infringing videos on video-sharing websites.

In our speech recognition research, our technology was capable of recognizing words read out directly into a microphone by the late 1990s, but the technology has recently advanced to the stage where it can recognize conversational speech from more than one person via a remote microphone. In conversations involving more than one person, there are inevitably problems due to effects such as sound blending and reverberation inside the room. To address these problems, we are researching speech and audio processing techniques such as an audio source separation technique that can resolve multiple audio sources in a real environment and a reverberation control technique.

We are also developing technology for recognizing the circumstances of a conversation (identifying who is talking to whom and who is holding the attention of the others and recognizing emotions such as sympathy and antipathy) in order to automatically analyze scene attributes such as the atmosphere of a conversation. Furthermore, we are working hard to develop lossless coding techniques that are free of distortion and to implement international standards for these techniques.

In recent years, as a result of the explosive growth of the Internet, the range of applications for natural speech processing has diversified to include information retrieval, detection of illegal or harmful information (e.g., spam filtering), and analysis of sentiments on websites. As this trend continues, the diversity of languages has become a significant problem. Documents such as blog posts and emails often contain neologisms and non-standard grammar. To tackle this problem, we are developing a natural language processing technique based on a machine learning approach in which a language model is automatically trained from a large volume of language data by using the power of computers. We have already made a number of key achievements using this technique. For example, in the syntax analysis technique that automatically analyzes the dependence relationships of words (a core technique in natural language processing), we utilized a semi-supervised learning technique[*1] to obtain the best analysis precision ever achieved in an international standard benchmark test. Moreover, in Japanese-English machine translation, we have not only developed translation technology but also proposed an automatic evaluation method called RIBES[*2] (rank-based intuitive bilingual evaluation score) that is closer to human evaluation than the conventional evaluation metric of BLEU[*3] (bilingual evaluation understudy). We are also utilizing these cutting-edge language processing technologies to develop a medical information access system that allows non-native speakers of English to obtain the latest medical information from English-language medical literature. As for machine learning technology itself, we have proposed a topic model based on Bayesian statistics[*4], which we have applied to data mining, and we are also engaged in cutting-edge research of topics such as non-parametric Bayesian theory[*5].

*1 Semi-supervised learning: A learning technique that uses both supervised (correctly labeled) and unsupervised (unlabeled) data.

*2 RIBES: A method for evaluating a translation system by focusing on words that appear in common between a machine translation and a correct translation and awarding scores on the basis of the correlation coefficient of the sequence in which these words appear.

*3 BLEU: A conventional measure of translation quality for machine translation.

*4 Bayesian statistics: Statistics derived from an inverse probability calculation method based on Bayes' theorem.

*5 Non-parametric Bayesian theory: A generator model where the number of models increases according to the complexity of data and can theoretically increase to infinity.

Research aimed at making a quantum computer, which will deliver ultrafast processing by means of massively parallel computation based on the principles of quantum mechanics, is being conducted by NTT Basic Research Laboratories. At the CS Labs, we are working on a study of the quantum information processing performed in quantum computers. By using a quantum algorithm, it has recently become possible to solve the anonymous leader election problem—a problem that is impossible to solve with classical algorithms. Other topics of theoretical research include the world's fastest generation of physical random numbers using a semiconductor laser and a theory of privacy verification. Details of the latter can be found in the Feature Article "Mathematical Duality between Anonymity and Privacy and Its Application to Law" [1].

4. Delivering a rich QoL for humans

As mentioned in the Introduction, in order to construct a theory of communication, we need to study not only media processing and information processing, but also human information processing mechanisms. At the CS Labs, we are taking a psychophysical and neuroscientific approach to the study of mechanisms that process information in the brains and bodies of humans, such as human sensory and perceptual mechanisms, the actions of emotions, and the mechanisms of movement. We have recently clarified the brain's mechanism for seeing textures, and we have devised a novel method for using the discomfort and illusion of a stopped escalator to generate the sensation of traction forces, which are used by the *Buru-Navi* traction force device [2]. With regard to illusions, we have established the Illusion Forum [3] on the web where people can experience visual and auditory illusions for themselves. We have also made progress in our research related to tactile senses: further details of this can be found in the Feature Article "Communication Research Focused on Tactile Quality and Reality" [4].

References

[1] K. Mano, "Mathematical Duality between Anonymity and Privacy and Its Application to Law," NTT Technical Review, Vol. 9, No. 11, 2011.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201111fa5.html

[2] Buru-Navi haptic interface.
http://www.ntt.co.jp/RD/OFIS/active/2007pdfe/hot/ct/05.html

[3] http://www.brl.ntt.co.jp/IllusionForum/

[4] 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

Naonori Ueda
Director, NTT Communication Science Laboratories.

He received the B.S., M.S., and Ph.D. degrees in communication engineering from Osaka University in 1982, 1984, and 1992, respectively. He joined the Yokosuka Electrical Communication Laboratories of Nippon Telegraph and Telephone Public Corporation (now NTT) in 1984. In 1994, he moved to NTT Communication Science Laboratories, Kyoto, where he has been researching statistical machine learning, Bayesian statistics, and their applications to web data mining. From 1993 to 1994, he was a visiting scholar at Purdue University, Indiana, USA. He is a guest professor at the National Institute of Informatics and the Nara Advanced Institute of Science and Technology. He is a Fellow of the Institute of Electronics, Information and Communication Engineers and a member of the Information Processing Society of Japan and IEEE.

Statistical Dialogue Processing for Ambient Intelligence

Yasuhiro Minami† *and Toyomi Meguro*

Abstract

In this article, we introduce statistical dialogue control methods that are essential for the implementation of ambient intelligence and present the results of experiments performed to assess the validity of statistical dialogue control.

1. Introduction

As social values have shifted away from attaching too much importance to convenience and efficiency, NTT Communication Science Laboratories has been working on projects related to ambient intelligence, which aims to enrich information and communications technology and improve the quality of life [1]. Towards this aim, we propose a new style of living that is achieved through technological developments in communication science, and we have made it our mission to perform interdisciplinary research around the theme of *integrated intelligence*.

As shown in the concept depicted in **Fig. 1**, the ecology of ambient intelligence can be expressed using phrases such as: there are a lot of intelligent entities, they can be discreet, they answer you, and they are hidden [2]. We are currently trying to realize a form of ambient intelligence by utilizing engineering. One of these trials is introduced below.

In the project, we first demonstrated diverse future life styles with ambient intelligence by presenting them in the form of various scenarios. A simplified version of one of these scenarios is presented below.

Scenario: Some people are meeting in a room. A number of ambient intelligences are also scattered around the room.

- Ambient intelligence 1: Analyzes the room's atmosphere. Is it friendly or serious? Also figures out the discussion topic from what is being said.
- Ambient intelligence 2: Presents hints relevant to the discussion by engaging in dialogues with humans such as "Could you remind me what ... is?" to collect information.

† NTT Communication Science Laboratories
Soraku-gun, 619-0237 Japan

2. Interactive processing techniques

How far has this scenario actually been realized so far? To answer this question, we will talk about research on dialogues, which is currently a very active subject. A multi-party prototype dialogue system that we are working on is shown in **Fig. 2**. We aim to develop a system that can understand the environment of a meeting. In addition, by using the system, we are conducting research into ambient intelligence that acts as an attentive *presence* that works without making a nuisance of itself [3], [4]. To implement natural dialogue control, we are also researching dialogue processing technology that satisfies the desire for someone to talk with, or for someone who is willing to listen [3], [4]. Below, we focus mainly on our dialogue processing research.

3. Statistical dialogue control methods

In a wide variety of environments, a huge effort is needed for the manual programming of systems that can dialogue flexibly with humans. Furthermore, complex programs of this sort are impossible to write without inconsistencies. We have therefore developed a statistical method that automatically learns dialogue control from dialogues among humans and between humans and systems, as shown in **Fig. 3**. This method is based on reinforcement learning: when the system performs an action in a particular situation, a *reward*

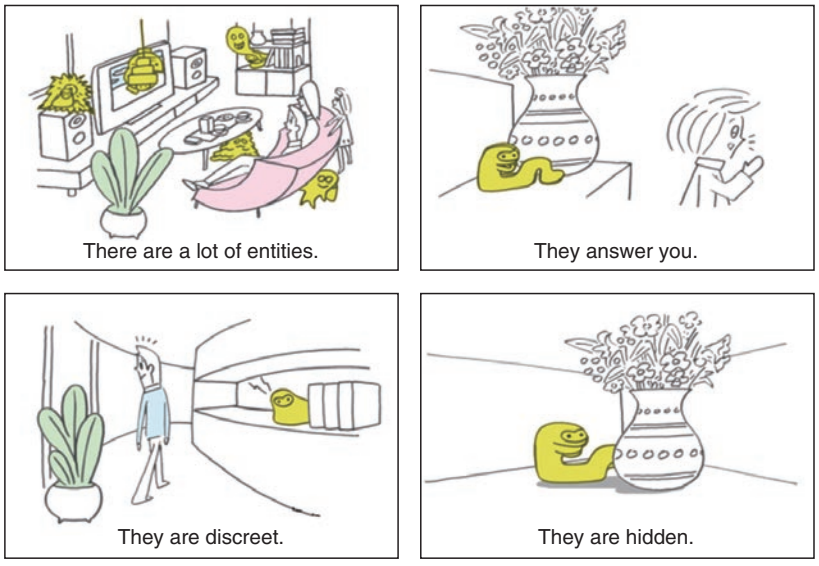


Fig. 1. Ecology of ambient intelligence.

Fig. 2. Multi-party dialogue system.

value is set to determine the action's degree of appropriateness. Reinforcement learning is a technique in which dialogue control for action selection to maximize the average reward obtained in the future is determined through a process of repeated trial and error. This technique is used in many different fields such as mechanisms that enable systems to learn actions, but when it is used in dialogue systems, a problem arises: how should the reward values be set to determine the system behavior? We prescribed two types of reward in order to implement better dialogue interactions [5].

(1) Rewards related to satisfaction

An annotator is asked to look at dialogues between a human and another human and between a human and a system and evaluate them in terms of their degree of satisfaction. For example, the dialogues are graded according to a 7-point scale by asking questions such as "Did the system listen properly to what the user was saying?"

(2) Rewards related to naturalness

Initially, we expected that with satisfaction rewards,

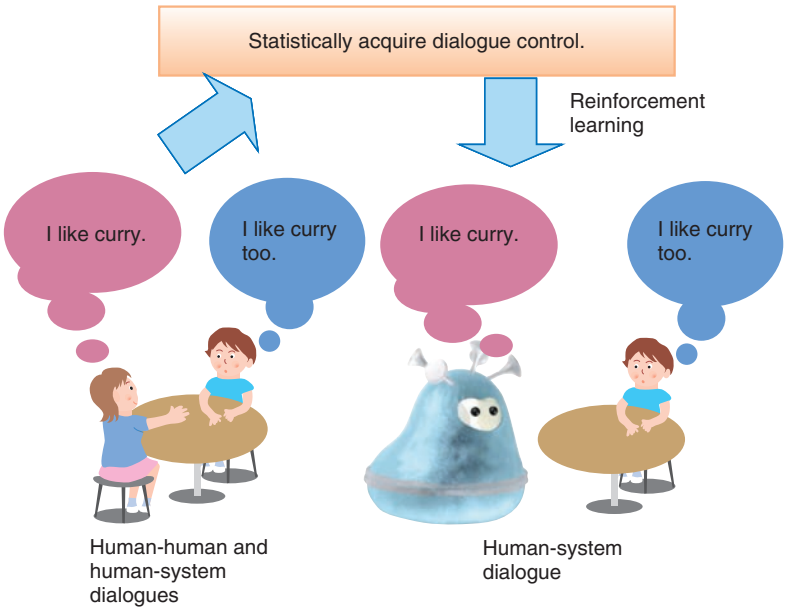


Fig. 3. Statistical interactive control method based on reinforcement learning.

the system would generate actions that satisfy users. However, in actual experiments, we found that the dialogue system ignored user actions that are naturally included in dialogues between humans, such as greetings, and instead suddenly launched into dialogues that gained higher satisfaction scores. To improve upon this, we considered that dialogues that are performed frequently but are not highly rated are also essential, and we also gave high rewards for such dialogues.

By simultaneously setting these two rewards, we implemented dialogue control in a listener interactive system that achieves both high satisfaction and natural interaction.

4. Evaluation experiment

Next, we experimentally evaluated the statistical dialogue control method [6]. With current technology, it is extremely difficult for a computer to automatically generate complex sentences such as those used in dialogues between humans. Therefore, in our proposed method, we implemented dialogue control in action units such as *asking questions* and *indicating acknowledgment*. Here, we refer to these action units as *dialogue actions*. To check the performance of the statistical dialogue control method, we performed an evaluation experiment with systems based on rules and a hidden Markov model. The evaluation consisted of three steps.

- Step 1: We constructed a statistical dialogue control system and five comparative systems (described below) and output dialogue action sequences for evaluation from each system by performing a simulation.
- Step 2: It is difficult to perform a direct evaluation by looking at generated dialogue action sequences. Therefore, the dialogue sentences were first generated manually from the sequences to make them more natural and then evaluated. Before the sentences were generated, an ordinary everyday scenario was presented to each participant. Specifically, participants were presented with fictitious information about what they had eaten, when, and with whom (e.g., last night, meat-and-potato stew, family). The 16 participants were then asked to create dialogue sequences by imagining themselves to be in those circumstances.
- Step 3: Three additional participants (different from those in Step 2) evaluated the dialogues produced in Step 2 on a 7-point scale in response to the question “If you were the speaker, would you think that the listener is doing a good job?”.

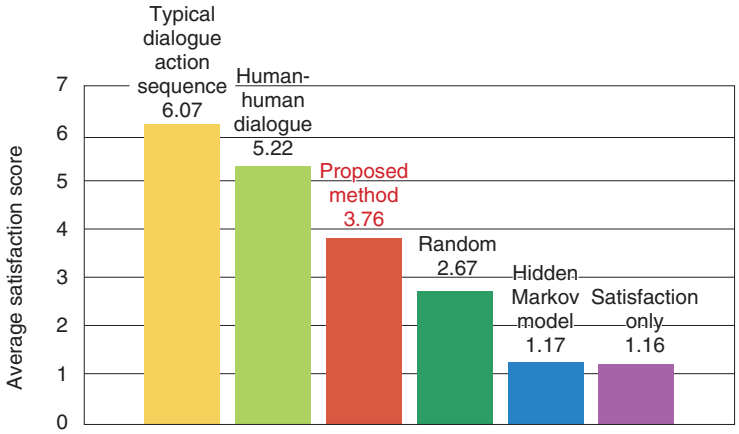


Fig. 4. Results for satisfaction values.

4.1 Comparative system

Statistical dialogue control was learned from the dialogue data. In this evaluation, we introduced a user simulator trained from dialogue data that mimics the user's action. The system's dialogue actions were produced using the learned dialogue control. That is, the actual users did not interact, but instead a simulator performed dialogues automatically instead of a computer and user. The user and computer generated 33 dialogue actions. The five systems provided for comparison are described below.

(1) Satisfaction only

This system uses exactly the same method as the statistical dialogue control method except that it does not use rewards for naturalness: it uses rewards only for satisfaction. This system was provided to evaluate whether naturalness rewards are necessary.

(2) Hidden Markov model

A hidden Markov model is a type of statistical model that can capture the flow of statistical time sequences. The dialogue actions of the system and user were generated from this model.

(3) Typical dialogue action sequence

Using rules generated from knowledge obtained in previous studies, interaction sequences were produced for the interaction of both systems and users.

(4) Human-human dialogue action sequence

This sequence was made by extracting actual dialogue action sequences from data for dialogues between humans.

(5) Random

Dialogue action sequences were generated randomly.

4.2 Evaluation results

The evaluation results are shown in **Fig. 4**. Our statistical dialogue control method ranked third best, meaning that it came immediately after the human-human dialogue action sequence. Consequently, we found that we were able to build a system that made people feel as if they were being listened to. The figure also shows that rewards for naturalness are important because the system that used only satisfaction was poorly rated.

An example dialogue from the system based on the statistical dialogue control method is shown in **Fig. 5**. It shows that the listener asked questions and performed self-disclosure. The typical actions of a listener were reproduced to some extent, and this helped to draw out the speaker's conversation. However, the third section was found to contain an inappropriate (unnatural) dialogue action and *thanks* was generated.

If the frequency of dialogue actions generated in a human-human dialogue is similar to the frequency of dialogue actions generated by the system, then the system dialogue actions should be close to the human dialogue actions. To study this similarity, we calculated the correlation coefficient between the frequency of dialogue actions in human-human interaction and the frequency of dialogue actions in each method (1 represents the highest correlation and 0 represents the lowest). The results are shown in **Fig. 6**. The highest correlation (closest frequencies) was obtained for the typical dialogue action sequence, and we think that the use of a rule that suitably reproduces human-human dialogue led to the high evaluation value in Fig. 4. The statistical dialogue control method had the

Utterance	Dialogue action
L: Where and whom did you have dinner with last night?	Question (sub-action-type: fact)
S: I had niku-jaga (meat-and-potato stew) with my family at home.	Self-disclosure (sub-action-type: fact)
L: Thank you. How many brothers and sisters do you have? My brother and his wife will visit me soon. I'm going to make *niku-jaga* for us all.	Thanks Question (sub-action-type: fact) Self-disclosure (sub-action-type: fact) Self-disclosure (sub-action-type: fact)
S: I see	Sympathy

Fig. 5. Excerpt from a typical listening-oriented dialogue.

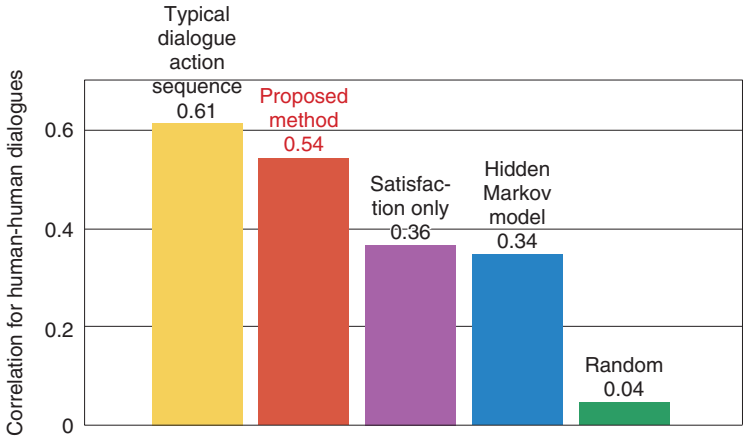


Fig. 6. Correlations between the five methods and human-to-human dialogues.

next highest correlation, and compared with the other statistical methods, it was able to reproduce the frequency of dialogue actions between humans. These results show that the statistical dialogue control method can acquire a suitable dialogue control method from the dialogue data.

5. Future prospects

So far, we have implemented a dialogue control that generates only dialogue actions. In the future, we intend to work at deepening our understanding of dialogues and the generation of dialogue content in order to implement a smarter dialogue system.

References

[1] E. Maeda, Y. Minami, and K. Dohsaka, "Resurgence of Fairies and Goblins: A Proposal for the New Vision of "Ambient Intelligence"," IPSJ, Vol. 47, No. 6, pp. 624–640, 2006.
[2] E. Maeda and Y. Minami, "Steps towards Ambient Intelligence," NTT Technical Review, Vol. 4, No. 1, 2006. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr200601050.pdf
[3] Japanese website for NTT Communication Science Laboratories OPEN HOUSE 2011 (in Japanese). http://www.kecl.ntt.co.jp/openhouse/2011/download.html
[4] English poster for NTT Communication Science Laboratories OPEN HOUSE 2011. http://www.kecl.ntt.co.jp/openhouse/2011/oh2011_poster_en.pdf
[5] Y. Minami, A. Mori, T. Meguro, R. Higashinaka, K. Dohsaka, and E. Maeda, "Dialogue Control by POMDP Using Dialogue Data Statistics," Spoken Dialogue Systems Technology and Design, Springer-Verlag, pp. 163–186, 2010.
[6] T. Meguro, R. Higashinaka, Y. Minami, and K. Dohsaka, "Controlling Listening-oriented Dialogue Using Partially Observable Markov Decision Processes," Proc. of the 23rd International Conference on Computational Linguistic (Coling 2010), pp. 761–769, Beijing, China, 2010.

Yasuhiro Minami

Senior Research Scientist, NTT Communication Science Laboratories.

He received the M.Eng. degree in electrical engineering and the Ph.D. degree in electrical engineering from Keio University in 1988 and 1991, respectively. He joined NTT in 1991. He had worked in the field of robust speech recognition. He was a visiting researcher at MIT, USA, from 1999 to 2000. Since February 2000, he has been with NTT Communication Science Laboratories. He is interested in modeling for speech recognition and speech dialogue systems.

Toyomi Meguro

Researcher, NTT Communication Science Laboratories.

She received the B.E. and M.E. degrees in electrical engineering from Tohoku University, Miyagi, in 2006 and 2008, respectively. She joined NTT in 2008. Her research interests are building listening agents and recommendation systems.

Communication Support Based on Interaction Analysis

Naomi Yamashita†

Abstract

This article describes the difficulties of human-computer interaction (HCI) research and the approach taken at NTT Communication Science Laboratories to alleviate them and establish findings for this research field. The main difficulty is that HCI research results are difficult to assess objectively owing to a lack of common yardsticks with other research fields.

1. Difficulties of HCI research

Human-computer interaction (HCI) research originates from the interactions between humans and information technology (IT), and it aims to develop technologies that bring greater comfort and convenience to our lives. HCI research is generally difficult; in particular, many researchers are currently struggling with evaluation methods. For example, if a system has been constructed with the expectation that it will improve the quality of interactions, it is surprisingly difficult to assess whether or not this expectation has been successfully met.

One of the reasons it is so difficult to evaluate an HCI-related technology is that no one has yet defined common yardsticks for interaction properties such as speed and correctness. Consequently, researchers themselves must find answers to questions such as what sort of state would indicate that an interaction has been improved, what would need to change and how before we can regard the objective as having been achieved, and what should things be compared with in order to justify the conclusion that there has been an improvement. The danger here is that different researchers can have very different views on these issues since there is no such thing as an absolutely correct value judgment.

As a result, each researcher uses his or her own value judgments as the basis for deciding when a particular interaction has been improved from the user's viewpoint and ends up creating systems based on those personal judgments. In this way, each researcher creates systems according to his or her own individual preferences, which results in many different systems whose similarities and differences are difficult to judge.

When a system is completed, it is then evaluated by setting tasks. Even when a system can be defined as an *improvement* according to everyone's values, it is not easy to assess whether or not this is actually the case. For example, suppose that someone has made a system that can be used to make communication more enjoyable. Since there are no yardsticks that can be used to measure enjoyment objectively, this is typically done through questionnaires—where users try out the system and then say whether or not they enjoyed using it. However, researchers generally have strong opinions about their own systems and tend have a strong belief in their usefulness. Consequently, even when users are surveyed by means of interviews or questionnaires, researchers are prone to phrase their questions so as to obtain results that favor their own systems. Even if there are no problems with the way the researchers phrase their questions, when users are asked to compare systems where a particular function is present and where it is not present, it is known that most of them will respond that the system with the function is better, regardless of whether or not they actually felt that it was better.

Through this sort of user testing, the expected results are obtained in most cases, and the researchers

† NTT Communication Science Laboratories
Soraku-gun, 619-0237 Japan

will report that their objective has been achieved. Unfortunately what normally happens next is that the system is not subjected to review and few further improvements are made.

My purpose here is not to belittle the diverse values of other researchers or the innovative systems that they have created. However, I feel that there is a problem whereby researchers can become so engrossed in the ingenuity of their own systems that they overlook or misapprehend the relevance of existing systems. Since only superficial evaluations are performed (in many cases), there is also a similar problem whereby improvements are overlooked or cannot be identified and are not connected with subsequent steps. For these reasons, the field of HCI unfortunately gives researchers in other fields the impression that it has gained little knowledge as a research field.

To allow HCI research to develop as a healthy research field, it is vital that evaluation and analysis are implemented in an accepted rational fashion and that systems are created with an awareness of the relevance of previous research, instead of superficial evaluations being performed on disparate systems created according to the vagaries of individual researchers. It is also necessary to have a common platform that all researchers can agree upon.

Of particular importance is the generalization of results obtained experimentally and by other means, and the accumulation of these results as knowledge. The results obtained from individual experiments are often limited and fragile, and it is not at all that rare for slight changes in experimental conditions and tasks to lead to different results. We thus need a way of generalizing experimental results by performing a comparative study with previous research so that they can be converted into meaningful knowledge. Through the accumulation of these results it should become possible to form a common platform by strengthening and storing knowledge covering the entire field of HCI research.

Below, I introduce the approach that my coworkers and I have taken so far. A characteristic of this approach is that we first observe and analyze how users use the system to understand how the system affects the actions of users, and we then consider the overall points of contact between humans and information technology. Since this approach makes it possible to suggest improvements on the basis of an understanding of the system and its role in collaboration, we can expect it to result in promising proposals.

Fig. 1. Two workers and two helpers at a remote site working on a mentoring task through the t-Room.

2. The t-Room video communication system

The t-Room is a video communication system that was developed at NTT Communication Science Laboratories [1]. It was designed for the purpose of reproducing live images of remote users and enabling them to engage in collaborative tasks as they move around a room. It can accommodate multiple users at a single location and transmit the positions in which remote users are standing almost exactly, wherever they are in the t-Room.

The t-Room hardware configuration is as follows. The system consists of multiple cameras arranged so as to surround the space where the users are, a number of vertical display devices, a central table consisting of two display devices, and a video camera that captures images of this table (**Fig. 1**). When multiple t-Rooms with the same configuration are set up in different locations and connected by audio and video communication links, it is possible for users in the t-Rooms at remote locations to share the same work space. A physical object on the table is also reproduced in two dimensions at the same position on the remote table, thereby making it possible to perform collaborative work while standing in positions around the table, not only for the users in the same t-Room but also for the remote users.

In a conventional system, the displays are connected and arranged so as to avoid any loss of video between the table display device and the tall display devices used to display the upper bodies of remote users, but in a t-Room system, a space is provided for people to enter between the table display and tall display devices, and the room is arranged so that people can move around freely in this space.

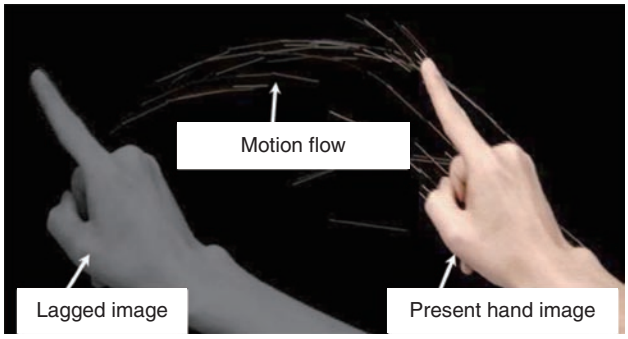


Fig. 2. Representation of t-Room.

3. Problem discovery through observation

Through our observations of preliminary tests, we have discovered a problem that can obstruct the gesture recognition of remote users in remote cooperative work using a t-Room. From related research and a more thorough analysis, we have identified two factors behind this problem. These two factors are not limited to the t-Room, but are also common to conventional remote table sharing systems.

- Diverted attention: When users are concentrating on work or paying attention to other things, they tend to miss the gestures of remote users. For example, in a remote instruction scenario, when workers are preoccupied with something in front of them, they will often fail to notice a gesture made by a remote user attempting to draw their attention to something else.
- Occlusion: When the table surface is covered by users' hands or by other objects, the video of a remote gesture displayed on the table surface may be buried underneath these objects and hidden from view. This phenomenon (occlusion) is observed particularly frequently in remote collaboration involving a real physical object, and it tends to have a detrimental effect on work efficiency. For example, in a remote instruction scenario, part of the gesture video of the remote user is buried by the object, and the worker may be unable to correctly apprehend the content of the instruction. A worker may sometimes attempt to recognize the instructor's gesture (pointing position) by lifting up the object, but by the time he or she has picked up the object, the gesture will most likely have finished and may no longer be visible.

4. Function proposed as a result of observation

In the above observation, we noticed that even if a user misses or is unable to see a remote gesture, he or she is often still aware of the gesture's existence and will take steps to search for the gesture. We therefore came up with a method called Remote Lag that delays the video of a remote user's gestures by a few hundred milliseconds when it is superimposed on the live video image (**Fig. 2**). That is, we thought that by making skillful use of delayed video of remote gestures, it should be possible to recover from states where gestures cannot be seen, even in cases where a remote gesture has been overlooked due to occlusion or a break in attention.

5. Experimental design

To evaluate the effects of Remote Lag in remote collaboration, we conducted a user study using the t-Room system. The research questions that we specifically wished to clarify in this experiment were as follows:

(1) Does Remote Lag alleviate the problem of being unable to see remote gestures?
(2) Does Remote Lag improve the efficiency of speech in remote instruction? For example, is there a reduction in the number of questions and confirmations uttered by the workers and in redundant instructions given by the instructors?
(3) What sort of cognitive load does Remote Lag place on users?

A total of 56 participants took part in this experiment (28 in a t-Room in Kanagawa and 28 in a t-Room in Kyoto). The participants were split into 14 groups of 4 individuals: 2 in Kanagawa and 2 in Kyoto. Each group worked on a group task related to remote work instruction on a t-Room table. This task was performed twice: once without Remote Lag and once with Remote Lag.

The group task consisted of remote work instruction using real objects, which is a common subject for remote collaborative work using a table. This task involved achieving an objective by means of a specialist (instructor) issuing instructions to a novice (worker) at a remote location. In this experiment, we performed remote collaboration tasks with four people by creating a situation where two specialists with different specialist knowledge who were located in Kanagawa provided instructions to two novices working in Kyoto.

Specifically, this collaboration task consisted of

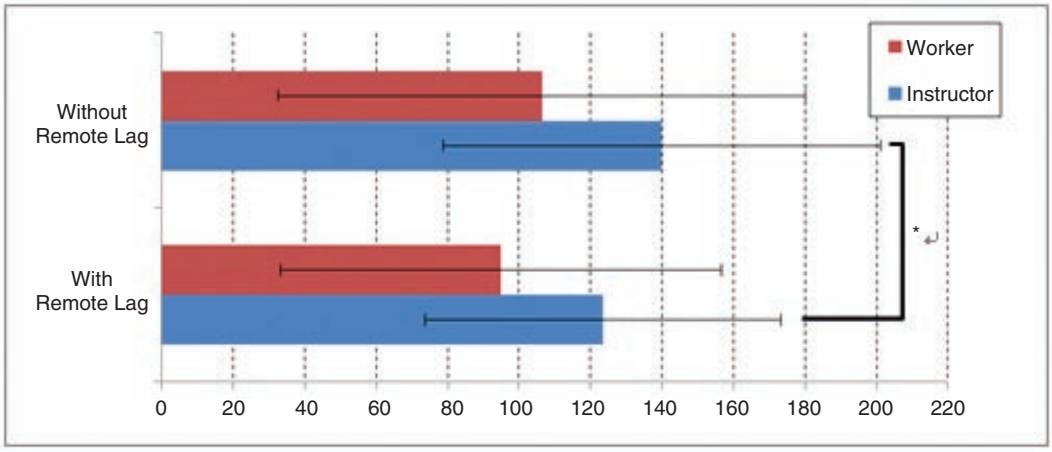


Fig. 3. Mean number of utterances per task for each experimental role.

assembling and arranging Lego blocks. The assembly of Lego blocks mainly involves having workers manipulate blocks by rotating and connecting them according to instructions from the instructors. This is often used in experiments involving collaboration on real objects. The requisite number of Lego pieces were placed on the table in the same t-Room as the workers, so that only the workers were able to manipulate them.

6. Evaluation of proposed function

To investigate in detail how the addition of Remote Lag affects the collaboration, we filmed the progress of the experiment by setting up video cameras in overhead positions at the entrances and opposite the entrances to both t-Rooms. First, we analyzed the video images to check whether or not the Remote Lag display alleviated the problem of being unable to see remote gestures. As a result, we discovered a number of situations in which Remote Lag was useful for alleviating the two problems mentioned above.

To verify the overall extent to which Remote Lag reduces redundant utterances in the remote collaboration instructions, we performed separate calculations of the average number of utterances made per task for each role and for both with and without Remote Lag (**Fig. 3**). As a result of comparing the number of utterances by using a paired t-test*, we found that compared with the tasks performed without Remote Lag, the tests performed with Remote Lag required significantly fewer utterances from the instructors ($t[27]=2.74$, $p<.05$). Next, to investigate whether or not Remote Lag helps to reduce the number of utterances associated with questioning and confirmation, we counted the number of question and confirmation utterances and compared the results obtained with and without Remote Lag. Here, utterances associated with questions and confirmations were those made for the purpose of requesting or confirming information that was insufficient or unclear during the task. As a result of investigating the ratio of such utterances among all of the utterances made during the task, we found that the ratio of questioning or confirming utterances made by the workers was 23% when Remote Lag was used, which was less than the ratio of 30% when Remote Lag was not used. By using a paired t-test, we found that the tasks conducted with Remote Lag were accomplished with significantly fewer question and confirmation utterances from the workers than ones without Remote Lag ($t[27]=2.71$, $p<.05$).

From these results, it appears that the Remote Lag display improves the visibility of gestures by the instructors and makes it easier for the workers to understand the instructions being given to them by the instructors, thereby reducing the frequency of questions and confirmations.

Finally, we evaluated the workload imposed on the workers by measuring the subjective workload due to the Remote Lag display. To perform these measurements, we used the NASA TLX (task load index)

* t-test: A general term for statistical hypothesis tests that make use of the fact that statistical quantities follow Student's t-distribution when the null hypothesis is assumed to be correct.

method. NASA TLX evaluates the overall workload in terms of six indicators (mental demand, physical demand, temporal demand, work performance, effort, and frustration). The overall workload is calculated in the range from 0 to 100 (lower scores are better) based on the weighted average of each indicator. This method is widely used for evaluating all manner of interfaces between humans and digital equipment as a method for evaluating workloads subjectively. By performing a corresponding t-test, we showed that the time pressure felt by the workers was significantly lower with Remote Lag than without Remote Lag ($t[27]=1.72$, $p=.097$). We can infer that this was because there is no need for extra effort in tracking realtime remote gestures because the delayed video provided by Remote Lag makes it possible to follow realtime gestures even if the remote gestures of the instructors and workers are missed.

7. Summary

In this article, I discussed the importance of making functional proposals based on an understanding of human actions, introducing our t-Room research as a case in point. In the future, I will continue making system designs and functional proposals on the basis of an understanding of human behavior, which I hope will be of use to others.

Reference

[1] K. Hirata, Y. Harada, T. Takada, S. Aoyagi, Y. Shirai, N. Yamashita, and J. Yamato, "The t-Room—Toward the Future Phone," NTT Technical Review, Vol. 4, No. 12, 2006.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr200612026.pdf

Naomi Yamashita
Communication Environment Research Group, Media Information Laboratory, NTT Communication Science Laboratories.
She received the B.Eng. and M.Eng. degrees in applied mathematics and physics and the Ph.D. degree in informatics from Kyoto University in 1999, 2001, and 2006, respectively. She joined NTT in 2001. Her primary interests lie in the areas of computer-supported cooperative work and computer mediated communication. Using a combination of quantitative and qualitative research methods, she aims to uncover the nature of human discourse/interaction and propose guidelines for designing new communication technologies.

Research on Social Network Mining and Its Future Development

Albert Ching-man Au Yeung and Tomoharu Iwata†

Abstract

In this article, we explore some core research problems in social network mining and discuss our latest research results. With the world becoming increasingly connected in this age of globalization, communications is no longer bounded by geographical location or languages; in particular, online social networking has become an important area of research.

1. Introduction

Social networks have been studied for many years by social scientists [1], who are particularly interested in understanding the roles of the people in a social network, how they are connected, and how information spreads among them. Social networks greatly influence how people interact and communicate with one another.

In recent years, online social networking has experienced explosive growth, transforming the World Wide Web into a platform for social interactions. Users share their opinions, photographs, music, and videos on social networking services (SNSs). Microblogging services like Twitter have also become an important tool for disseminating realtime information [2]. This phenomenon has motivated the development of social network analysis using computers and algorithms.

2. Social network mining

In social network mining, we apply data mining algorithms to study large-scale social networks. Social network mining has attracted a lot of attention for many reasons. For example, studying large social networks allows us to understand social behaviors in different contexts. In addition, by analyzing the roles of the people involved in the network, we can understand how information and opinions spread within the network, and who are the most influential people (**Fig. 1**). In addition, since social network users may receive too much information from time to time, social network mining can be used to support them by providing recommendations and filtering information on their behalf.

In social network mining, we generally ask three broad questions:

(1) What are the characteristics of the social network?
(2) How can we model the network?
(3) How can we support its users?

When trying to answer the first question, we aim to identify different properties of a given social network. For example, what do people do in this social network? Do they exchange messages, or do they share items among themselves? We can also ask, for any two persons, what is the probability distribution of the distance between them? Are there any clusters or communities within the network? Answering these questions enables us to understand how information flows and how social relations in the network evolve. For example, some research on trying to understand the social networks of Twitter [2] and Flickr [3] has been done.

After understanding the characteristics of a particular social network, we may want to construct a mathematical model that explains the processes in the network. A mathematical model lets us predict future changes in the network. For example, what is the probability of a new edge between two given persons?

† NTT Communication Science Laboratories
Soraku-gun, 619-0237 Japan

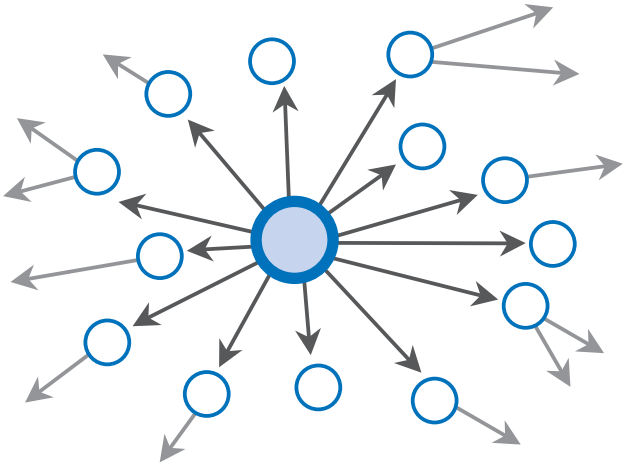


Fig. 1. Identifying the most influential people in a social network.

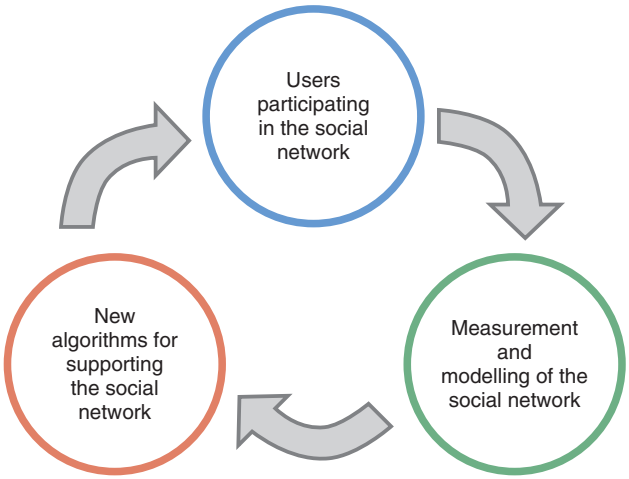


Fig. 3. Life cycle of social network mining.

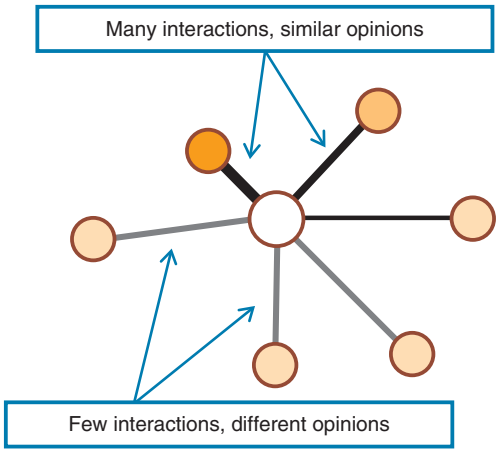


Fig. 2. Using social network mining to estimate the strengths of social relations

When a new person joins the network, who will he or she connect to? We may also want to model the behavior of the people in the network in order to explain when and why two persons interact.

Once we have some knowledge about a social network and its underlying mechanism, it would be good if we could make use of it to support communication among the people in the network. For example, on the basis of their past activities, can we predict who is most likely to become a friend of a given person? Can we estimate and infer the strengths of social relations among different people (**Fig. 2**)? And can we make better recommendations to the users on the basis of their social circles?

The above three steps form a cycle (**Fig. 3**) that one can travel along in order to continuously gain more and more insight into how people interact and then improve the experience of the social network's users, thus attracting more people to participate in it.

3. Our research on trust networks

We recently proposed a new method for analyzing trust networks on the web and generating more accurate predictions [4]. Trust networks are social networks in which a person is connected to others because the person believes that they are trustworthy.

To study trust networks, we collected data from a popular product review site called Epinions [5]. It lets users write comments and rate any product that they have brought, such as digital cameras, vacuum cleaners, and books. In addition, it lets users create a trust network. For example, if a user thinks that another user's comments and ratings are reliable, he/she can add that user to his/her own trust network (**Fig. 4**).

Using data collected from Epinions, our first step was to investigate how trust relations shape user opinions. To do this, we calculated the similarity between pairs of users who had established trust relations between them at some point. Similarity was calculated from which products they had rated and the ratings they had given to them. The results are shown in **Fig. 5**.

The graph shows that the similarity between users increases over time both before and after a trust

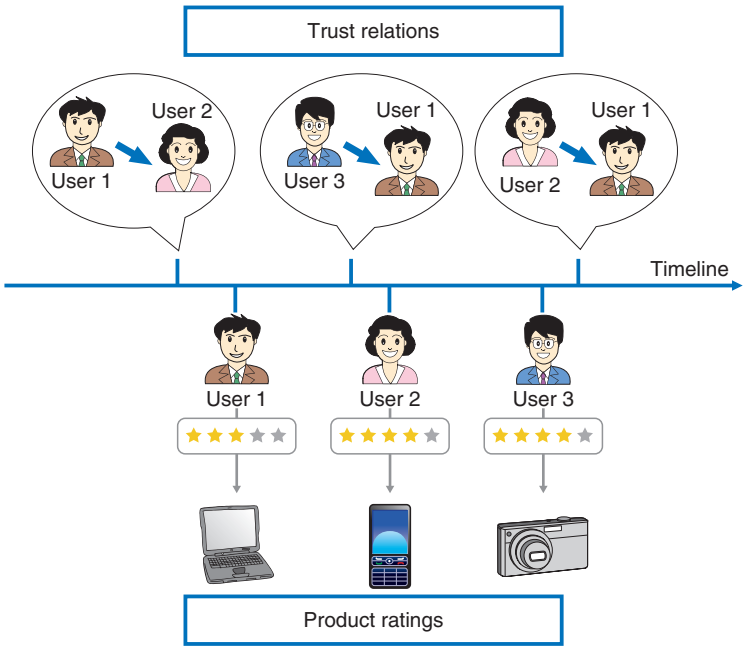


Fig. 4. Epinions user activities.

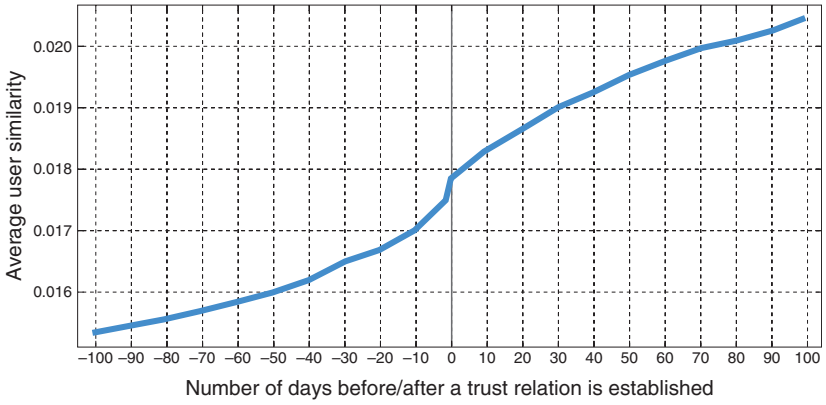


Fig. 5. Change in similarity over time between pairs of users connected by trust relations.

relation is established. This can be explained by two factors. First, the increase before trust is established can be explained by the theory of homophily, which states that people who are similar to each other tend to congregate (birds of a feather flock together). Therefore, when similarity increases to a certain level, it triggers one user to trust another. Second, the increase after trust is established can be explained by influence. That is, when a user trusts another, the former will be influenced by the later and his or her opinions will become more similar to the latter's.

On the basis of this finding, we further proposed an algorithm based on matrix factorization to predict the rating that a given user will give to a given product. More accurate predictions will let us generate better recommendations for users. In our method, we assume that a user's rating is determined by two factors: whether the user and the product are compatible with each other, which is determined using the basic matrix factorization technique, and the influence from other users in the trust network. If other users tended to give this product high ratings, then this user

is likely to be influenced by them to give a higher rating too. Our algorithm first estimates the strengths of influence between different users on the basis of past activities and then uses these strengths to make predictions.

Our experiments showed that this new algorithm can produce predictions with lower errors than standard matrix factorization techniques. This project has demonstrated that our understanding of a social network can be used to develop new algorithms for supporting the network's users.

4. Summary

Since more and more people are likely to interact and communicate with one another on the web, social network mining will become even more important. It not only holds the key to understanding social behavior and group dynamics on a huge scale, but also is crucial in developing new tools and functions to support communication in social networks.

References

[1] S. Wasserman and K. Faust, "Social Network Analysis: Methods and Applications," Cambridge University Press, 1994.

[2] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a social network or a news media?" Proc. of the 19th International World Wide Web Conference (WWW2010), pp. 591–600, Raleigh, NC, USA, 2010.

[3] M. Cha, A. Mislove, and K. Gummadi, "A Measurement-driven Analysis of Information Propagation in the Flickr Social Network," Proc. of the 18th International World Wide Web Conference (WWW2009), pp. 721–730, Madrid, Spain, 2009.

[4] C. M. Au Yeung and T. Iwata, "Strength of Social Influence in Trust Networks in Product Review Sites," Proc. of the 4th ACM International Conference on Web Search and Data Mining (WSDM2011), pp. 495–504, Hong Kong, China, 2011.

[5] Epinions. http://www.epinions.com/

Albert Ching-man Au Yeung

Research Associate, Innovative Communication Laboratory, NTT Communication Science Laboratories.

He received the M.Phil. degree in computer science and the B.E. degree in information engineering from the Chinese University of Hong Kong and the Ph.D. degree in computer science from the University of Southampton, UK, in 2001, 2004, and 2009, respectively. He joined NTT Communication Science Laboratories in 2009. His research interests include social media, social network analysis, user modeling, and recommendation systems.

Tomoharu Iwata

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

He received the B.S. degree in environmental information from Keio University, Tokyo, the M.S. degree in arts and sciences from the University of Tokyo, and the Ph.D. degree in informatics from Kyoto University in 2001, 2003, and 2008, respectively. His research interests include data mining, machine learning, information visualization, and recommender systems. He received the IPSJ (Information Processing Society of Japan) Best Paper Award, FIT (Forum on Information Technology) Young Researcher's Award, and Funai Best Paper Award. He is a member of the Institute of Electronics, Information and Communication Engineers and IPSJ.

Mathematical Duality between Anonymity and Privacy and Its Application to Law

Ken Mano†

Abstract

My colleagues and I have developed a formal verification method for the anonymity and privacy of information systems based on the idea that anonymity is hiding information about who performed a certain action and that privacy is hiding information about what action a certain person performed. In this article, I investigate the use of a formal representation of privacy in a legal context. As an example, I compare such a formal concept of privacy with a legal one. On the basis of a comparison with the *after the banquet* case, whose decision is widely used as a precedent, I point out a difference between the two concepts related to identifiability and investigate its meaning. To clarify the meaning of the difference, I use ambient intelligence as a technical context.

1. Introduction

Formal methods [1] are methods aiming at verifying the correctness of information systems in a logically rigorous manner. Such correctness is usually confirmed by putting test data into the system and checking the result. However, it is generally impossible to input all available data, so untested data could cause errors. In systems that require very high reliability such as mission-critical systems or security systems, such a possibility cannot be overlooked. Formal methods have been proposed as a way to solve this problem. By formally describing the system's correctness and by formally proving it, we can achieve a highly reliable system.

My colleagues and I are conducting research on privacy verification by using a formal method [2]. One of the characteristics of our method is the approach that we use for the formal formulation of privacy properties. Privacy is a somewhat ambiguous notion, and there are two key points in our approach regarding its formalization: (1) the use of epistemic logic and (2) the formulation of privacy as the hiding of information about what action a certain person performed or what states a certain person is in. The first point is important because security problems, including privacy, are problems concerning how much information can be known by an adversary of the system. Thus, we chose a formal framework appropriate for formalizing knowledge. The second point arose from comparison with anonymity, which is explained later.

In this article, I describe the use of a formalization method for anonymity and privacy in a legal context. Specifically, I compare formally and legally defined concepts of privacy. I do not intend to perform a thorough comparison. On the basis of a comparison with the *after the banquet* case [3], whose decision is widely used as a precedent, I point out a difference between the two concepts in terms of identifiability and investigate its legal meaning.

Identifiability is one of the requirements of a tort of privacy invasion, which would appear quite natural in a daily sense. Describing the problem in a logical formulation is expected to be useful for examining the meaning of such a *seemingly natural* requirement and for obtaining a beneficial legal consequence.

Not only should we devise the way of formulating using the logical formulation; we must devise

† NTT Communication Science Laboratories
Atsugi-shi, 243-0198 Japan

examples to clarify the difference between two concepts. For this purpose we use *ambient intelligence* as a technical context. Ambient intelligence is an ability that actively reacts to activities of people in a ubiquitous network environment by using electronic tags and sensors.

This article is organized as follows. Sections 2 and 3 explain privacy as formal and legal concepts, respectively. Section 4 compares the two concepts and points out a difference between them with respect to identifiability. Its meaning is investigated in section 5.

2. Privacy as a formal concept

There are several methods for formally formulating and verifying privacy. Here, I introduce our approach [2], which is based on anonymity research by Halpern et al. [4]. One of the characteristics of their research is that they utilize mathematical logic called epistemic logic. The expressions of epistemic logic are not only rigorous but also relevant to intuition, so epistemic logic is appropriate for legal issues.

In mathematical logic, the basic elements are expressed by symbols, and logical statements are expressed as formulae. We use the symbol $\wedge$ to express *and*, $\vee$ for *or*, $\Rightarrow$ for *implies*, and $\neg$ for *not*. The logical statement "If A or B holds, then C holds and D does not hold." is expressed by the following mathematical logic formula: $(A \vee B) \Rightarrow (C \wedge \neg D)$.

The original purpose of mathematical logic was to express the truth of logical statements, but there are some extensions for making it easy to express human epistemic activities such as "i knows that ….", "j thinks it is possible that ….", etc. Such an extension is called *epistemic logic*. In epistemic logic, K_i is an operator expressing i's knowledge (K is the initial letter of *know*). For instance, if *Beautiful*(*rose*) is a formula expressing "A rose is beautiful.", then K_i(*Beautiful*(*rose*)) means that "i knows that a rose is beautiful." By combining the knowledge operator with negation $\neg$, we can express possibility. For instance, if *Delicious*(*grasshopper*) expresses "a grasshopper is delicious", then $\neg K_i \neg$(*Delicious*(*grasshopper*)) means that "i does not know that a grasshopper is not delicious", in other words, "i thinks that it is possible that a grasshopper is delicious." Thus, using the initial letter of *possibility*, $\neg K_i \neg$ is abbreviated to P_i.

Halpern et al. proposed a general method for formulating anonymity properties of information systems by using epistemic logic [4]. For example, they call an anonymity property defined by the following epistemic formula *anonymity up to* I_A.

$$\theta(i, a) \Rightarrow \bigwedge P_j[\theta(i', a)]$$

Here, $\theta(i, a)$ is an abstract expression indicating that i has performed action a (or that i is in state a). In other words, θ expresses a linkage between people and actions (states). We regard anonymity properties as how well the information about *who performed a certain action* is hidden. This makes the method general, in the sense that it is independent of specific applications. Moreover, I_A is a subset of participants called the *anonymity set*, and j is an observer of the system. Thus, the meaning of the whole formula is "if i has performed a, then j thinks that a could have been performed by anybody in I_A." The demand for anonymity can vary from system to system, so some systems can require more complicated forms of anonymity. Even in such a case, the expressiveness of epistemic logic enables a flexible description of the required anonymity. By describing such an anonymity formula and choosing an appropriate interpretation of θ and I_A, we can define the specific anonymity required for a specific system.

On the basis of the above research, we proposed a method for formulating and verifying privacy properties formally [2]. In this method, we consider privacy properties that can be expressed as a relation between people and their actions. In other words, we regard privacy properties as how well the information about *what action a certain person performed* is hidden. Then, privacy can be formalized in a similar way to the formalization of anonymity. That is, we can define various privacy properties in terms of knowledge about $\theta(i, a)$, and as with anonymity, an epistemic-logic-based approach provides a good way of specifying them.

For example, we define a property called privacy up to A_I, whereby, from j's point of view, i could have performed any action A_I:

$$\theta(i, a) \Rightarrow \bigwedge_{a' \in A_I} P_j[\theta(i, a')]$$

Note that the above formulae for privacy and anonymity are symmetric[*1]. However, in general, there are no logical dependencies such as one holds if the other holds, or one holds if the other does not hold. They are logically independent of each other.

*1 This symmetry can be thought of as a kind of duality.

3. Privacy as a legal concept

In this section I introduce privacy as a legal concept in Japan. Here, I concentrate on the right of privacy in private law. There are other related legal concepts, e.g., the right to control personal information in public law or the Personal Information Protection Law, but they are beyond the scope of this article.

When we consider the right of privacy in private law, the most important decision is that of the *after the banquet* case. This decision recognized the legal guarantee and/or right that private life not be disclosed without reason so that personalities are mutually respected and private individuals are protected against unjustifiable interference. In its decision, the court presented the following three tort requirements of privacy right violation for the disclosed matter: (1) it is true, or can be taken to be true in intimate life, (2) it is offensive to a reasonable person from his or her viewpoint, and (3) it is not of concern to public.

For (1), there is an additional requirement, which is not explicitly stated in the above three requirements: the said person must be identifiable by the disclosed matter. This is essential, e.g., in the case of invasion of privacy by a novel based on real people and incidents.

In the following, I use the above requirements as the definition of privacy as a legal concept and call it *legal privacy*. That is, legal privacy is protected, or satisfied, when any of the above three conditions does not hold. In the following, I investigate the difference between legal and formal privacy.

First, a trivial difference is that legal privacy has the explicit conditions "offensive" and "not of concern to the public", while formal privacy does not, or at least such conditions are implicit. We regard this difference as trivial since it originates from the basic idea of formal privacy formulation, where we concentrate on privacy properties that can be expressed as a linkage between persons and actions. However, a specific privacy definition suitable for a specific case can be obtained by fixing the interpretations of I_A, A_I, and θ of formal privacy. The conditions of legal privacy should be reflected in the interpretation. On the other hand, for legal privacy, whether or not the matter is offensive and whether or not it is of concern to the public are judged according to the individual situation. There seems to be rather a natural correspondence between them.

As mentioned at the beginning of this article, it is not my objective to thoroughly compare formal and legal privacy. In the following, I concentrate on the difference as regards identifiability: the identifiability of the said person is a requirement in the definition of legal privacy, while there is no corresponding condition in the definition of formal privacy. Unlike in the trivial difference mentioned above, there is a reason for the lack of this condition in formal privacy. The negation of identifiability is valid as a necessary condition for anonymity. Actually, both properties concern hiding information about who performed a certain action. However, formal privacy hides the information about what action a certain person performed. They are generally independent, and there must be no relation between them such as one being a necessary condition for the other.

From a legal standpoint, the most plausible interpretation of this difference is that the definition of formal privacy is insufficient. However, this article poses the following question in order to examine the legal meaning of the difference: Is there any social situation or technical context where formal privacy is more suitable than legal privacy? This question can be generalized to whether or not identifiability is necessary. In the following, I call this the *identifiability problem*.

4. Identifiability problem

In this section, I investigate the problem posed in the previous section by using ambient intelligence [5] as the technical context.

4.1 Ambient intelligence

The ambient in ambient intelligence means the environment of a ubiquitous society, where electronic tags, biometric sensors, and networks connecting them are widespread. The intelligence is of course analogous to artificial intelligence and means the ability to actively react to human activities by processing information gathered by the ambient intelligence using natural language processing, knowledge and data processing, and other means [6].

If ambient intelligence is achieved, it will be convenient for users in various ways. On the other hand, some potential problems have been pointed out. One of the biggest criticisms is the potential loss of privacy. For example, let us consider the following case. Person A uses an ambient intelligence system at home. The system takes care of his family's health by using biometric technology and a healthcare service site on the Internet. One day, B visits A, and the ambient intelligence system measures B's biometric information and sends it outside the system. B could

potentially file a lawsuit against A for invasion of privacy.

There are two important points in this case. One is that the root of the problem lies in the incentive to actively circulate the obtained information, which is assumed to lead to improved convenience for individuals and society. This can lead to private information being communicated without foresight. This is not restricted to ambient intelligence; for instance, the privacy issues related to Google Street View can be thought of as a consequence of such an incentive. The other point is that this is not only a technical problem but also a social problem and in particular a legal problem.

Such potential problems with ambient intelligence have been extensively studied [7]. For instance, the aim of PRIAM [8], a research project in INRIA [9], is transversal and multidisciplinary research through the exchange of ideas between lawyers and experts in information and communications technology. One of the specific objectives of PRIAM is to develop a methodology for privacy policy specifications, and a formal method is expected to provide a promising approach for handling both technical and social problems.

4.2 Two cases

To investigate the difference mentioned in section 3, let us consider the following two cases.

(1) Biometric information

In ambient intelligence, biometric information is assumed to be recorded and exchanged in various situations. Concerning the handling of such information, there have been reports and discussions in relation to the Personal Information Protection Law. Among the arguments, those related to whether or not some information corresponds to personal information under the Personal Information Protection Law are closely connected with our problem. Shimpo [10] pointed out the importance of the existence of referable information for distinguishing individuals. Murakami [11] pointed out the importance of the owner of such referable information. A similar precise argument is needed for invasion of privacy.

Here, let us consider an extreme case to make the problem clear. Sato stated that a symbolic issue of privacy is that the analysis of a complete human genome has become possible [12]. Genome information can be regarded as the ultimate personal information. Now let us assume that your complete genome information is disclosed without reason. However, only the genome information itself is disclosed, and no information about your identity such as your name and address is provided. Is this an invasion of your privacy? A key point here is the possibility of identifying someone solely from their complete genome information.

Actually, the owner of a given genome is uniquely determined with very high probability, except in the case of an identical twin. However, it is generally difficult to determine a genome's owner by using only the genome information since that person could be anywhere in the world. So if we conclude that the identification of the said person is impossible, the disclosure of genome information would not in itself be an invasion of privacy.

There are various possible interpretations of this problem. For instance, it is possible to think that it is not a problem to disclose genome information about an unnamed person. However, today's biotechnology, which is enabling personalized medicine and genomic medicine, seems to be leading to a situation in which such a pastoral attitude is not allowed. For instance, genome medicine could be used to produce material that is harmful only to the genome's owner.

(2) Surveillance of a bank doorway

Let us consider the following situation. A publicly accessible surveillance camera is placed in front of a bank door. For privacy protection, the camera's image output is pixelated by realtime image processing so that people in the images cannot be identified by their faces. However, other parts of their bodies remain clear, and we can recognize features such as shirt color and the extent to which a bag is bulging. Moreover, a thermographic camera and a microwave-based motion sensor are also installed near the door and they are connected to the ambient intelligence and are publicly accessible.

Information about clothing is insufficient to identify a person, but sufficient to allow us to guess that *the person coming out now is the same person who went in three minutes ago*. Many public organizations in Japan are setting up infrared thermographic cameras in preparation for a pandemic of a new type of influenza. Automatic sliding doors often have a type of approach sensor called 2.4-GHz microwave Doppler radar. The same type of radar has been used in research aimed at remote sensing of the human pulse [13].

Thus, on the basis of the above knowledge, suppose that a third party observes that the bag of a man just coming out of the bank is bulging more than at the time he entered, that his pulse rate is elevated, and that his body temperature is lower. This third party

could presume that the man has withdrawn a large amount of money. Is this an invasion of the man's privacy?

What I want to stress here is the following. We are considering situations where the incentive to publish and circulate obtained information is dominant, that various types of information about a person are published since they do not identify that person, and thus that such information can easily be accumulated. Then, even if identification is impossible, some information can still be known, e.g., that a man in a red shirt will pass a certain place carrying a lot of money, and such knowledge can easily be used for malicious purposes. If we suppose, in the first case, that the disclosed genome information is known to be that of a rich person, the situation is similar to this case.

5. Self-information condition

It seems likely that a reasonable person would feel that the cases described in the previous section are invasions of privacy. However, it cannot be so according to the precedent that identifiability is required. Moreover, the disclosed information does not correspond to personal information under the Personal Information Protection Law since it is insufficient to distinguish individuals, so the cases do not conflict with the law.

The idea of requiring identifiability dates back to early legal arguments on the right to privacy. For instance, in the four types of invasion of privacy in the classic research by Prosser [14], the second type called *public disclosure of private facts* is closely related to the right recognized in the decision in the *after the banquet* case. The protected interest of this type is considered to be reputation[*2]. When the protected interest is reputation, identifiability is required.

Should we devise a new interest or right in order to avoid the situations highlighted in the previous section? The more widely a person's right to privacy is recognized, the more restricted the rights of others to know become. Thus, we should not devise new rights thoughtlessly. In fact, I think that it is possible to regard the above situations as invasions of the recognized interest that private life not be disclosed without reason.

In many privacy torts, including the *after the banquet* case, the problem was damage to reputation, that is, detriment that accrues from a person being blamed behind his or her back. Therefore, identifiability of the person to be blamed became a requirement. However, the cases in the previous section show that, in future technical contexts, private lives could be invaded for no reason, and serious detriment could be caused in a completely different way from blaming. It is obviously impossible to reduce this problem to that of the right not to be blamed behind one's back. Actually, one of the main criticisms of the Prosser's classification is that it constitutes such a reductionistic attitude to the right of privacy [15].

So, instead of reducing the problem of privacy to a legal guarantee of reputation or of not being criticized, let us return to the decision in the *after the banquet* case, which recognizes the legal guarantee and/or right for one's private life not to be disclosed without reason in order that personalities be mutually respected and that private individuals be protected against unjustifiable interference. If we consider the right itself as the interest to be protected, the genome information and surveillance information in the cases described in the previous section should not be disclosed without reason.

Now, recognizing the problem in the previous section as a legal one, what solution is possible? In the following, I consider how to solve this problem by making a minimal modification to the interpretation of the decision in the *after the banquet* case. Specifically, I consider how to relax the identifiability requirement.

First of all, it is nonsense to simply eliminate the identifiability requirement. If we did that, we would lose any condition that requires the disclosed matter to concern the said person, and then disclosure of matters related to other people could be an invasion of the said person's privacy. Thus, let us instead consider adding a requirement to avoid such a silly situation, that is, a requirement that the disclosed matter is a true fact about the person. In this article, I call this requirement *self-information*. The difference between identifiability and self-information is that the former concerns whether or not others can know who the person is from the disclosed matter, while the latter concerns whether or not the matter is about the person regardless of how it is seen by the others.

To fulfill self-information in the case of genome information disclosure, it is enough to prove that the disclosed genome information is that of the said person by using a DNA (deoxyribonucleic acid) test. With surveillance information, let us suppose the court recognizes that, even in a public area, there is no implicit consent to allow the disclosure of an

*2 There are criticisms concerning the idea; I refer to these later.

unnecessarily detailed image, body temperature reading, or pulse rate of a person. Then it is sufficient to prove that, for example, the sensor was operating when the person was there. Then we can prove these cases to be illegal regardless of identifiability.

However, it is impossible simply to substitute self-information for identifiability. As is mentioned in the first requirement, the disclosed matter is true, or *can be taken to be true* in intimate life. That is, the matter does not have to be true, whereas self-information must be true.

So such a simple replacement would not be a minimal modification since it would narrow the range of privacy protection. Therefore, we propose to interpret the disjunctive sentence as actually expressing a disjunction of two requirements as follows:

- "it is true in intimate life" expresses self-information and
- "it can be taken to be true in intimate life" expresses identifiability.

In the following, I call the legal privacies with the former and latter requirements *legal privacy based on self-information* and *legal privacy based on identifiability*, respectively.

Interestingly, this relaxation of identifiability is closely related to formal privacy. To clarify the relationship, let us ignore the conditions "offensive" and "not of concern to the public" that are abstracted from formal privacy. Then legal privacy is something concerning an identifiable and disclosed matter. Let us consider an invasion of privacy where a private matter "i performed action a" of i is true. If we formally interpret identifiability as the negation of anonymity up to I_A and disclosure as the negation of the conclusion part of privacy up to A_I, then legal privacy is interpreted as being expressed by the following formula.

$$\theta(i, a) \wedge \left\{ \neg \bigwedge_{i' \in I_A} P_j[\theta(i', a)] \right\} \wedge \left\{ \neg \bigwedge_{a' \in A_I} P_j[\theta(i, a')] \right\}$$

In this interpretation, formal versions are weaker than legal versions. For instance, identifiability implies failure of anonymity but the reverse does not hold in general. Nevertheless, the above formula is too restrictive, or too strong, as a formula expressing the invasion of formal privacy[*3]. So, to weaken the above formula, let us omit the second condition $\neg \bigwedge_{i' \in I_A} P_j[\theta(i', a)]$. The resulting formula

$$\theta(i, a) \wedge \left\{ \neg \bigwedge_{a' \in A_I} P_j[\theta(i', a)] \right\}$$

intuitively means that the matter is self-information $\theta(i, a)$ and is disclosed ($\neg \bigwedge_{a' \in A_I} P_j[\theta(i, a')]$), which corresponds to legal privacy based on self-information. And this formula is also acceptable as a formula expressing an invasion of privacy in a formal sense, since it is itself the negation of privacy up to A_I.

6. Concluding remarks

I compared formal privacy with legal privacy and investigated their difference and its meaning by using ambient intelligence as a technical context. This comparison yielded a problem with identifiability, and to resolve it I introduced the notion of self-information. Moreover, I showed that there is a close relationship between legal privacy based on self-information and formal privacy.

There are various issues related to self-information. For instance:

- What kind of self-information is the target of the protection? What is offensive self-information?
- How good a match is sufficient to judge that a certain matter is self-information? It is easy to change just the appearance of digital data.
- How about the case where the matter is not disclosed but exploited at someone's discretion?

These cases are always subjects of privacy whether or not they are based on self-information. It is not apparent whether a formal method can help solve these problems.

A problem specific to self-information is the following. Although a person can prove illegality regardless of identifiability, he or she is identified by the very fact of proving in court that the matter is self-information, which could be detrimental. First, a person may want to file a lawsuit even if it enables him or her to be identified, which should be admitted by law. Next, this problem can resolved if the privacy court is anonymized.

Moreover, it would be worthwhile applying the same perspective to investigations of the Personal Information Protection Law.

*3 This is proved by formally showing the existence of a model that formally satisfies both the negation of privacy and the negation of the above formula expressing legal privacy invasion. The cases in the previous section can be regarded as similar models in the context of ambient intelligence.

References

[1] Y. Tsukada (Eds.), "Special Feature: New Trends of Formal Methods," Information Processing Society of Japan, Vol. 49, No. 5, pp. 491–543, 2008 (in Japanese).

[2] K. Mano, Y. Kawabe, H. Sakurada, and Y. Tsukada, "Role Interchange for Anonymity and Privacy of Voting," Journal of Logic and Computation, Vol. 20, No. 6, pp. 1251–1288, 2010.

[3] After the Banquet. See e.g., http://en.wikipedia.org/wiki/After_the_Banquet

[4] J. Y. Halpern and K. R. O'Neill, "Anonymity and Information Hiding in Multiagent Systems," Journal of Computer Security, Vol. 13, No. 3, pp. 483–514, 2005.

[5] Y. Tonomura and E. Maeda (Eds.), "Proposing Ambient Intelligence—New Paradigm for Information-driven Society," Maruzen, 2008 (in Japanese).

[6] W. Weber, E. H. L. Aarts, and J. Rabaey (Eds.), "Ambient Intelligence," Springer, 2005.

[7] D. Wright, S. Gutwirth, M. Friedewald, E. Vildjiounaite, and Y. Punie (Eds.), "Safeguards in a World of Ambient Intelligence," Springer, 2010.

[8] PRIAM: Privacy Issues and Ambient Intelligence. http://priam.citi.insa-lyon.fr/

[9] INRIA. http://www.inria.fr/recherches/domaines-de-recherche/actions-de-recherche-collaboratives (in French).

[10] F. Shimpo, "Use of Biometrics Based on Personal Information Protection Law Journal of Information and Media Studies, Vol. 4, No. 1, pp. 55–76, 2006 (in Japanese).

[11] Y. Murakami, "Legal Issues Related to Biometrics," The Information Network Low Review, the Information Network Law Association Japan, Vol. 4, No. 2, pp. 74–92, 2005 (in Japanese).

[12] K. Sato, "Article 13 of the Constitution of Japan and the Rights of Controlling Self-information," Lecture Meeting, the Information Network Law Association Japan, 2009 (in Japanese).

[13] S. Suzuki, T. Matsui, and S. Fujie, "Trial of Contactless Measurement of Heart-rate-variability Index during Work Using Microwave Rader," Journal of Japan Ergonomics Society, Vol. 44, Special Issue, pp. 272–273, 2008 (in Japanese).

[14] W. L. Prosser, "Privacy," California Law Review, Vol. 48, No. 3, pp. 383–423, 1960.

[15] R. Gavison, "Privacy and the Limits of Law," The Yale Law Journal, Vol. 89, No. 3, pp. 421–471, 1980.

Ken Mano

Senior Research Engineer, Innovative Communication Laboratory, NTT Communication Science Laboratories.

He received the B.E. degree in applied physics and the M.E. degree in information engineering from Nagoya University, Aichi, in 1987 and 1989, respectively. He joined NTT Basic Research Laboratories in 1989 and studied term rewriting systems and process algebra. He is currently studying formal verification of security systems. He is also interested in legal aspects of security and privacy. He is a member of the Institute of Electronics, Information and Communication Engineers, the Information Processing Society of Japan, and the Japan Society for Software Science and Technology.

Communication Research Focused on Tactile Quality and Reality

Junji Watanabe†

Abstract

Although tactile sensation is increasingly being used in modern electronic devices and games, humans use tactile sensation not only to recognize information but also to richly perceive quality and confirm the reality of objects. This article describes current research on communication based on the unique features of tactile sensation.

1. Introduction

The presentation of tactile information has been used in mobile terminals and gaming devices in recent years. This has spurred research and development of information-presentation technologies based on tactile illusions. Illusion is a phenomenon that results when there is a discrepancy between what our brain perceives and what actually exists. It forms the basis for presenting information using limited resources and for creating a sensation that has not been perceived previously. Some examples of presenting information via tactile illusions are shown in **Fig. 1**. As they indicate, understanding the principle of how humans sense their environment is the key to developing new technologies for tactile information presentation and communication.

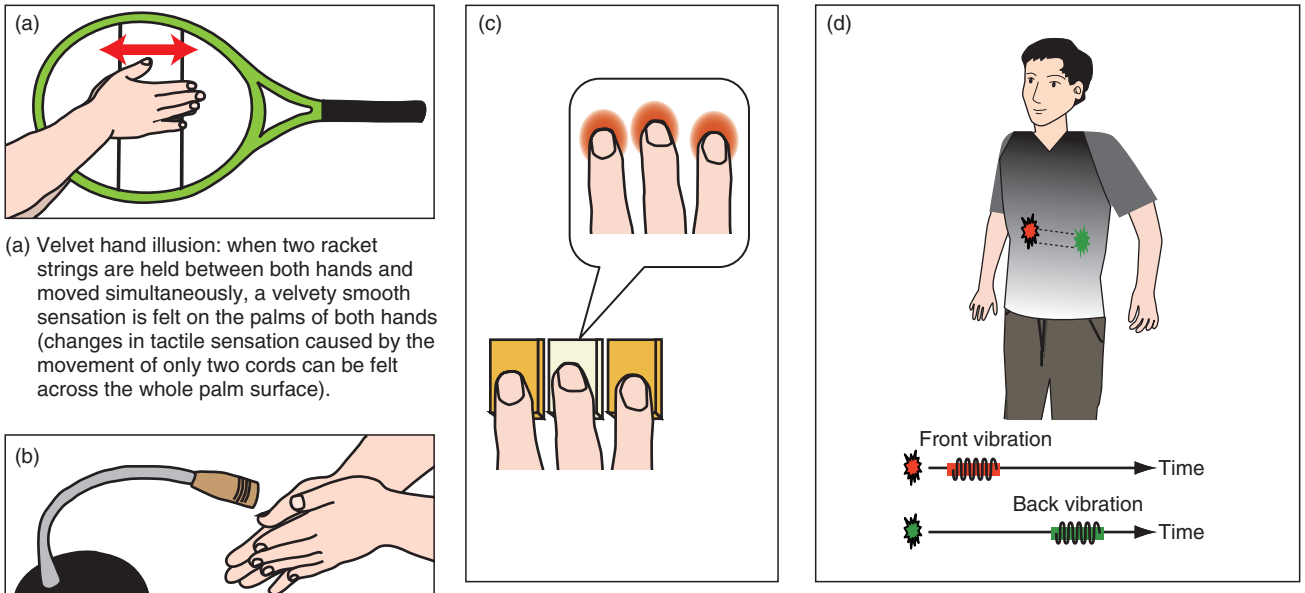
Humans perceive and differentiate various tactile qualities such as roughness, hardness, comfort, and fit from all aspects of daily life (clothing, furniture, one's own skin or another person's skin, etc.). In the same way as we feel comfort or discomfort with what we wear, we feel comfort or discomfort even in the simple act of pushing a button, depending on the button's material and how it feels when pushed. Toys based on only the sensation of touching a button have recently been created. As such, humans use tactile sensation not only for operating objects, but also for perceiving differences in tactile qualities and for feeling comfort, affection, and other emotions.

Likewise, we confirm the existence of objects by touching them. By touching something with our hands, we make sure that the object actually exists, and by touching a person's body, we ascertain the presence of the other person. Thus, the sense of touch is considered to be the most important sensation that humans use to confirm the reality of things around them. In this article, I discuss the unique features of touch, namely, tactile quality and reality, and introduce research and projects being conducted in NTT Communication Science Laboratories on these topics.

2. Tactile quality and onomatopoeia

First, I will explain about the qualities of tactile sensation and give an overview of our research on its classification. Analyzing the words for expressing sensations is one of the methods used in psychological analysis and sensation classification. For example, the colors red, green, and blue are the most commonly used categories of visual sensation. Studying the relationships among these color categories is one way to understand how humans classify visual sensations. For example, the color hue circle shows that green is opposite to red, and it provides intuitive understanding of the relationships among the colors that humans perceive. The relationships among tactile sensation categories, however, are not as clearly generalized as color category ones. Although categories pertaining to the material composition of objects,

† NTT Communication Science Laboratories
Atsugi-shi, 243-0198 Japan.



(a) Velvet hand illusion: when two racket strings are held between both hands and moved simultaneously, a velvety smooth sensation is felt on the palms of both hands (changes in tactile sensation caused by the movement of only two cords can be felt across the whole palm surface).

(b) Changing the sounds produced from rubbing both hands in front of a microphone causes a change in tactile sensations.

(c) Touching an unheated board positioned between two heated boards with three fingers causes all the three fingers to feel the heat (even partial increases in temperature cause all the boards to feel warm) [1].

(d) Placing vibrators on the abdomen and on the back and producing temporally shifted vibrations causes a sensation of having something "passing through" the body (the illusion causes a sensation that could not be sensed physically) [2].

Fig. 1. Excerpt from the exhibit on information presentation devices using tactile illusions entitled "Touchable Illusions: tools for understanding human tactile perception" (by Kitagawa, Ho, and Watanabe) in NTT Communication Science Laboratories Open House 2011.

such as metal, cloth, or paper, exist, there are currently no generally used categories for the sensations derived from touching objects.

To analyze tactile sensation categories, we have focused, in particular, on onomatopoeia (general term for mimetic words formed by imitation of a sound or synesthetic association between a sound and sensation) used to express tactile sensations. Onomatopoeia is a convenient way to express sensations in everyday life and is also widely used in Japanese manga and literary works. Compared with other languages, Japanese is known to have a large number of onomatopoeic words for tactile sensations. Thus, we chose tactile onomatopoeic words as tactile sensation categories and analyzed the impressions that they connote. In our experiment, we chose forty-two Japanese onomatopoeic words for tactile sensations and asked participants to numerically rate their perception of how the onomatopoeic words connote size, friction, and viscosity. For example, the word *zara-zara* may connote size at a level of 3, friction at 5, and viscosity at 2 for a certain participant, indicating how he or she perceives different onomatopoeic words in terms of these impression criteria. A two-dimensional distribution map of onomatopoeic words constructed from the responses of 20 participants is shown in **Fig. 2** (For details of the analysis, refer to [3]).

The distribution map shows a spatial diagram of how tactile sensations are categorized by Japanese people. In this diagram, onomatopoeic words that express closely related sensations are also located close to each other on the map. Mapping the onomatopoeic words spatially enables visualization of categories and grouping axes of tactile sensations. On the distribution map, words such as *jyari-jyari* and *jyori-jyori*, which express roughness, are grouped in the upper left, while words such as *tsuru-tsuru* and *sube-sube*, which express smoothness, are grouped in the lower right. The word *kori-kori*, which expresses hardness, is located in the lower left, while words such as *gunya-gunya* and *necho-necho*, which express softness, are grouped in the upper right. And words such as *nuru-nuru* and *nyuru-nyuru*, which express wetness, are grouped in the middle right, while words

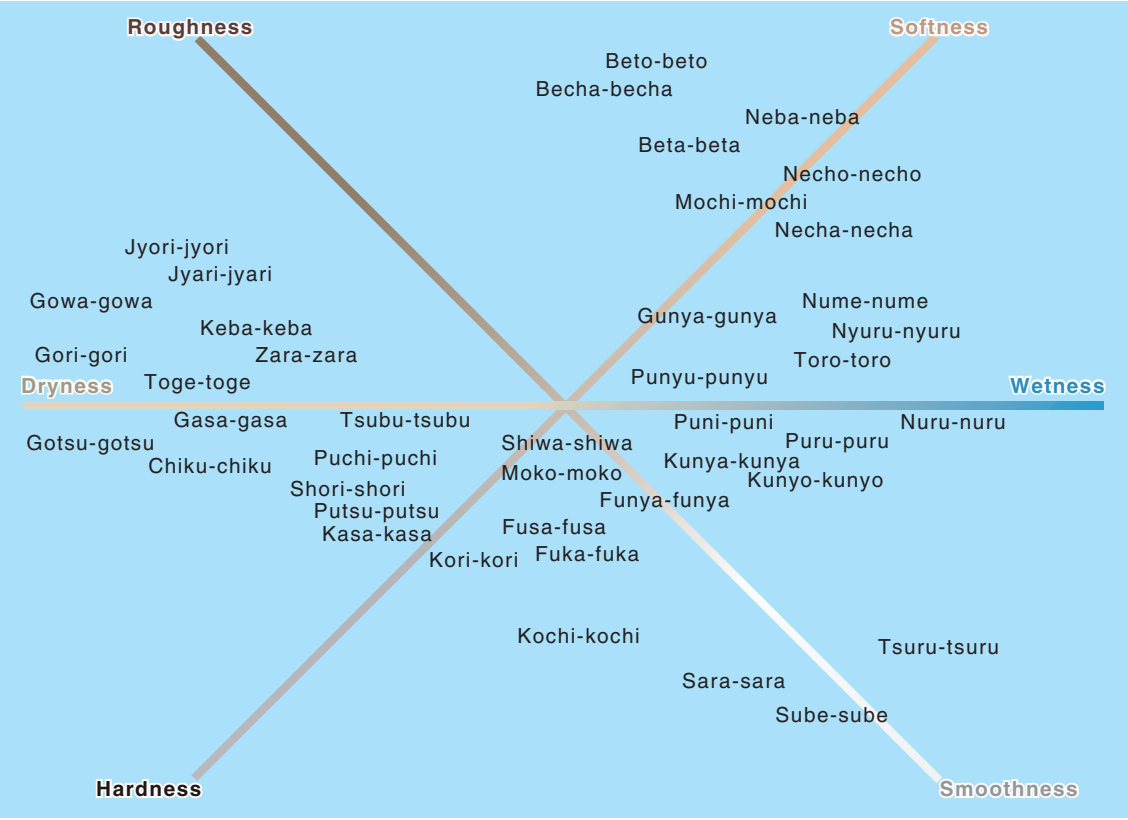


Fig. 2. Two-dimensional distribution map of onomatopoeic words for tactile sensation.

such as *gasa-gasa* and *kasa-kasa*, which express dryness, are grouped in the middle left. These groupings of onomatopoeic words, which also indicate their phonemic similarities, represent the major categories of tactile sensations. The distribution map also revealed that the sensations of smoothness/roughness, hardness/softness, and wetness/dryness serve as the basic classification criteria for tactile sensations.

3. Visualizing tactile material relationships

Next, we used the distribution map to visualize the relationships among tactile materials in accordance with the spatial distribution of tactile onomatopoeic words on the map. We used the distribution map in a workshop aimed at enabling participants to recognize how they perceive objects through touch [4]. Workshop participants were asked to touch ten different kinds of materials and determine their locations on the map on the basis of how they perceived those materials. Then, the participants were asked to choose a material whose tactile quality they liked and another that they did not like and to draw an arrow from the latter to the former. As shown in **Fig. 3**, the arrows on the map show the directions of personal preferences. The onomatopoeic words were superimposed on the pictures of the materials for easier viewing. On this map, most of the arrows pointed from the hard and dry sensations in the upper left to the smooth sensations in the lower right (preference for smoothness) and to the soft and wet sensations in the upper right (preference for softness). The onomatopoeic distribution map enabled a systematic discussion of personal preferences.

This research is based on the diversity of words (onomatopoeia) used to express tactile sensation in the Japanese language and is aimed at establishing the principles for evaluating the quality of tactile sensations. Although, as mentioned earlier, the theories for effectively determining visual relationships and auditory relationships, such as color and musical scales, are already more or less well established, through this distribution map, we hope to enrich communication through touch by providing a useful and

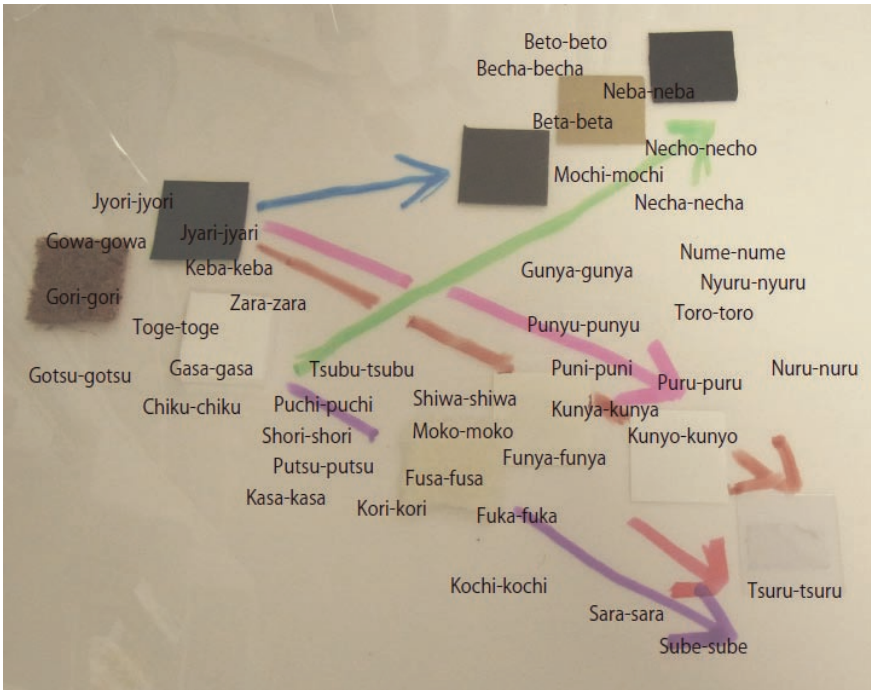


Fig. 3. Arrows showing personal preferences.

effective format for conveying the quality of tactile sensations.

4. Perceiving reality through the sense of touch

In this section, I discuss how the sense of touch provides us with a way to confirm the reality of objects. Visual and auditory modalities deal with the sensations of objects that are far from the body, while tactile sensation is derived from the act of directly touching an object. Therefore, tactile sensation provides a means of confirming the existence of the object as well as knowing what kind of object it is. We have attempted to extend this ability of confirming the existence of objects by touching them to objects that cannot be touched. We believe that *touching* invisible or conceptual objects (called haptization) will allow a deeper understanding of these objects through their physical reality.

5. Workshop on touching heartbeats

We developed a simple device, shown in **Fig. 4(a)**, for haptizing the heartbeat, a very basic phenomenon of life, in a workshop called Heartbeat Picnic aimed at reaffirming people's lives [5], [6]. The participants were asked to hold a stethoscope in one hand and a vibration speaker (referred to as the heart box) in the other, as shown in **Fig. 4(b)**. When the stethoscope was placed on the person's chest, his or her own heartbeat was output as both sounds and vibrations from the heart box, which enabled participants to not only hear their own heartbeats but also feel them as vibrations.

Since the workshop was held outdoors, participants were free to move about as if they were having a picnic, and they were able to feel the changes in their heartbeat with their own hands. Moreover, as shown in **Fig. 4(c)**, by exchanging their heart boxes with those of other participants, they were able to feel the differences between their own heartbeats and those of other people. The experience of touching heartbeats, which is usually impossible, provided each participant with the opportunity to appreciate the importance of his or her own life as well as those of others.

Life appreciation workshops are usually held in a natural environment where participants can directly experience the abundant life around them. In the highly digitized cities in which we live, however, there are almost no opportunities for us to appreciate our own lives. Although there are sometimes

Fig. 4. (a) Equipment used in the workshop: stethoscope, vibration speaker (heart box), signal processing circuit, and batteries. (b) Equipment in use. (c) Exchange of heart boxes among participants.

opportunities for us to face life-and-death situations through accidents, such as when we go to hospital, these are extraordinary and limited experiences. Everyday life in this modern digital world provides us with little opportunity to appreciate the dignity of life. This workshop, however, is aimed at enabling people to experience the *reality of life* without having to escape from their modern everyday environment, by evoking the imagination through the sense of touch.

We received various responses from the workshop participants regarding their experience of touching their own heartbeats. Some said that they felt a sense of endearment with their own heart boxes and a sense of affinity and kindness upon feeling other people's heart boxes. One participant commented that the only other time she had felt another person's heartbeat was when she had a baby in her womb. These responses indicate that the experience of touching the heartbeat, even if only simulated and artificial, provides an opportunity to appreciate the vitality of one's own life and the vitality of other people's lives.

In our modern world, we live surrounded by information in the form of digital code representing phenomena, but such representations are not the real phenomena themselves. Our digitized world enables us to use symbols to simultaneously convey phenomena to many people. Symbolic representations of phenomena, however, dilute the actual sensation of the environment in which phenomena occur and weaken our connection to them. We believe that using the sense of touch, as in the workshop, enables us to have a physical and realistic connection with the wealth of information about various phenomena that surround us every day. As the information era advances, touch-based communication technologies will provide the means for us to more realistically understand ourselves and people around us.

6. Future research prospects

In this article, I introduced research activities based on the unique features of tactile sensation, namely, tactile quality and reality. Using these experiences as a springboard, we plan to continue our efforts to elucidate human sensory mechanisms and carry out practical initiatives for exploring the role of tactile communication in society and in our daily lives.

References

[1] H.-N. Ho, J. Watanabe, H. Ando, and M. Kashino, "Mechanisms Underlying Referral of Thermal Sensations to Sites of Tactile Stimulation," The Journal of Neuroscience, Vo. 31, No. 1, pp. 208–213, 2011.

[2] J. Watanabe, Y. Fukuzawa, H. Kajimoto, and H. Ando, "Presentation of Feeling When Pierced Using Apparent Movement Passing thorough the Body," Transaction of Information Processing Society of Japan, Vol. 49, No. 10, pp. 3542–3545, 2008 (in Japanese).

[3] T. Hayakawa, S. Matsui, and J. Watanabe, "Classification Method of Tactile Textures Using Onomatopoeias," Transaction of Virtual Reality Society of Japan, Vol. 15, No. 3, pp. 487–490, 2010 (in Japanese).

[4] Tactile texture workshop. http://www.junji.org/texture/ (in Japanese, but includes videos).

[5] J. Watanabe, Y. Kawaguchi, K. Sakakura, and H. Ando, ""Heartbeat Picnic": Workshop for Touching Heartbeats," Trans. of Virtual Reality Society of Japan, Vol. 16, No. 3, pp. 303–306, 2011 (in Japanese).

[6] Heartbeat Picnic. http://www.junji.org/heartbeatpicnic/

Junji Watanabe

Research Specialist, NTT Communication Science Laboratories.

He received the Ph.D. degree in information science and technology from the University of Tokyo in 2005. From 2005 to 2009, he was a PRESTO researcher in the foundation of technology supporting the creation of digital media content in Japan Science & Technology Agency. From 2009 to 2011, he was a Research Fellow in the Japan Society for the Promotion of Science. He became a Research Specialist in NTT Communication Laboratories in 2011. He studies cognitive science and communication devices with applied perception. His fields of interests are visual and haptic perception and communications. He received honorary mentioned in Ars Electronica 2004, and 2011. His works were exhibited at the Japan Media Arts Festival in 2006, 2007, and 2008. He also works on stage design for the performing arts with the media performance unit cell/66b. Recent activities include Kobe Biennale 2009, "Sensory Circuit Collection" at Miraikan, the National Museum for Emerging Science and Innovation (2009–2010), and "Cyber Arts Japan Ars Electronica—30 years for Art and Media Technology" at the Museum of Contemporary Art Tokyo (2010).

Current Progress of ITU-T Speech Coding Standardization

Shigeaki Sasaki†

Abstract

Since speech coding is a key technology required for telecommunication, several recommendations for it have been approved in ITU-T (International Telecommunication Union, Telecommunication Standardization Sector), and new standardization work, e.g., G.711.1 and G.729.1, which further enhance the speech bandwidth and have interoperability with the conventional telephony standard, is currently progressing for the new products and services demanded by the market.

1. Introduction

Broadband networks based on Internet protocol (IP) over wired networks, e.g., FLET'S HIKARI NEXT, have spread widely in Japan and wireless ones over cellular phone networks are also available. One of the services on those IP networks, VoIP (voice over IP), such as HIKARI DENWA, is also becoming popular for use in both homes and offices. Even though the access networks have become broadband, meaning that they have high capacity, speech coding technologies are still required in order to provide telecommunication services to many customers with high quality simultaneously because the network has limited throughput. Moreover, the standardization of speech coding methods guarantees the speech quality and interoperability of telecommunication services that use methods conforming to the standards.

2. Standardization in ITU-T

ITU-T (International Telecommunication Union, Telecommunication Standardization Sector) develops global standards for telecommunication and has standardized speech coding algorithms for telephony and voice communication. During the current study period, 2009–2012, Study Group 16 (SG16) is responsible for standardization of multimedia coding, systems, and applications. As shown in **Fig. 1**, WP3, one of its Working Parties, coordinates overall standards regarding multimedia coding and handles study items, which are assigned to Questions. A lot of experts in the media coding field participate in those groups and contribute to the standardization. Work items related to speech media are being or have been studied in the following Questions.

- Q.8 (Generic sound activity detection): Q.8 is developing a method of detecting a voice period and classifying signals present in that period, e.g., speech, music, or noise. It was responsible for the G.720.1 standard aimed at pre-processing for speech coding.
- Q.9 (Embedded variable bit rate coding of speech signals): This Question studied a coding algorithm that can provide different coding bitrates and produced G.718 as a result; it was then terminated in Oct. 2010.
- Q.10 (Speech and audio coding and related software tools): In the previous study period, 2005–2008, the main role of Q.10 was maintenance for existing coding standards. Its scope of work has been extended to take on most of the responsibilities regarding speech coding from this study period. Maintenance of G.191, which is a package of tools required for standardization work, is also an important role of Q.10.

SG meetings are usually held every eight or nine months, but interim meetings, such as WP3 meetings, are organized once every two or three months in order

† NTT Cyber Space Laboratories
Musashino-shi, 180-8585 Tokyo

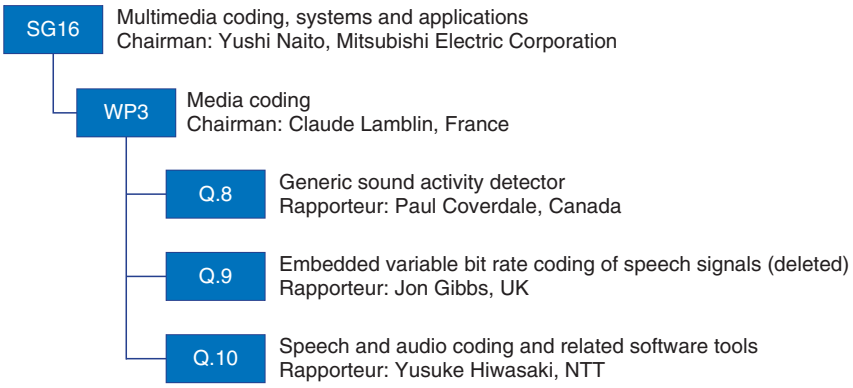


Fig. 1. SG16 structure related to speech coding.

to deal with market demands for new standards in a timely manner.

3. Hot trends in speech coding standard

Conventional standards for speech coding were designed to have speech bandwidth, e.g., telephone band or wideband, and bitrate to suit services for which the standard was expected to be used. However, since the last study period, before the start of standardization work, Requirements were defined to meet market needs for products and services. For the current market demands, the following features have been included in the recommendations.

- SWB and FB audio coding: Telecommunication systems with higher speech quality and higher presence, such as a telepresence system, need the capability to encode speech signals with bandwidths in the classes known as super-wideband (SWB, 50 Hz to 14 kHz) or full band (FB, 20 Hz to 20 kHz).
- Scalable audio coding: To give the bitstream a layered structure, both the bitrate and speech quality can be controlled according to the throughput of the transmission channel or some other conditions. Furthermore, in order to enhance the speech quality and the bandwidth of the conventional coding, enhancement bits can be layered onto the bitstream of the conventional structure. This structure also enables interoperability with the conventional coding through the extraction of the bits corresponding to those of the conventional structure.
- Lossless audio coding: Coding in which no information contained in the original signals is lost through compression and decompression. (In conventional speech coding, called lossy coding, distortion from the original is allowed in order to obtain a higher compression rate.)

Below, current recommendations with the above features are introduced in the order of their approval date.

(1) G.722.1 Annex* C

On the basis of a wideband speech standard, G.722.1, Polycom proposed its SWB extension as G.722.1 Annex C for videoconferencing and teleconferencing systems. G.722.1 Annex C was the first SWB speech coding standard in ITU-T and has very low computation for SWB coding. Its basic algorithm was common to G.722.1, but compatibility with G.722.1 was not provided.

(2) G.729.1

G.729.1 is a wideband scalable extension of G.729 (8 kbit/s), originally designed for VoIP. Since its bitstream contains that of G.729 as the core bitstream, terminals supporting G.729 can decode the core bitstream into speech signals by extracting it from the G.729.1 bitstream. Its bitrate can be controlled from 8 kbit/s to 32 kbit/s in 2-kbit/s steps. France Telecom, VoiceAge, Panasonic, and Siemens and others contributed its standardization.

(3) G.711.1

The proposal from NTT and four other organizations, ETRI, France Telecom, Huawei Technologies, and VoiceAge, was approved as G.711.1. The most important feature of this standard is interoperability

* Annex: An appendix recommendation that forms a part of the recommendation body. If related to the body, it might be made into a series as a part of the body.

with G.711, which is the most widely spread for speech coding. Scalable coding, for which the core layer consists of the G.711 bitstream, was introduced into it. 7-kHz wideband speech is obtained by decoding the whole bitstream and it is possible to extract only the core part and decode it by G.711. Because G.711.1 was based on G.711, which has a bitrate of 64 kbit/s, the bitrate of this new standard seems to be higher, up to 96 kbit/s, than that for speech compression, but it was targeted for use on the broadband network and no attention was paid to compression rate. Where legacy terminals that support only G.711 and new terminals that can handle G.711.1 co-exist, interoperability between the two standards is a great advantage. Moreover, its speech quality has been confirmed to be the highest among the wideband recommendations in ITU-T. G.711.1 is now used in the high-quality telephone service of HIKARI DENWA provided on NTT's broadband network FLET'S HIKARI NEXT.

(4) G.718

Using scalable coding, G.718 has also achieved both a bitrate that can range from 8 kbit/s to 32 kbit/s and wideband speech coding. The bitrate and basic structure of the bitstream are similar to those of G.729.1; however, the coding algorithm of the core layer at 8 kbit/s is a new design rather than the conventional one. Since there are no constraints with regard to interoperability with existing standards, a more efficient coding algorithm could be introduced. This recommendation was studied in the open consortium organized by Ericsson, Motorola, and Nokia and others.

(5) G.719

G.719 is the first FB speech coding standard in ITU-T. A joint proposal from Polycom and Ericsson, based on the algorithm of G.722.1 and G.722.1 Annex C, was approved as G.719. It has the advantage of much less computation like G.722.1 and G.722.1 Annex C.

(6) G.711.0

The concept of lossless audio coding was introduced for the first time in ITU-T in G.711.0. To be precise, this algorithm compresses the bitstream from G.711, not the input audio signals, without any degradation. The compression rate for lossless coding will be varied according to the redundancy of the input source to be encoded, but the rate has been confirmed to be 50% on average for a G.711 bitstream. Making new network devices, such as routers and gateways, that conform to this standard would save backbone network throughput that could be used for voice channels. A joint proposal from NTT, Huawei Technologies, and Cisco and others was approved as G.711.0.

(7) G.718-SWB/G.729.1-SWB

There are some common features between G.718 and G.729.1, such as the speech bandwidth, bitrate, and bitstream structure scalability. Their extensions for SWB capability were initially proposed independently, but additional layers for SWB enhancement, which are applicable to both standards, were developed considering their common features. As a result, the SWB extensions to G.718 and G.729.1 were approved as G.719 Annex B and G.729.1 Annex E, respectively.

(8) G.711.1-SWB and G.722-SWB

SWB extensions to G.711.1 and G.722 were developed in the same way as for the standardization of SWB extensions to G.718 and G.729.1. There are some features in common between G.711.1 and G.722, e.g., wideband capability, low delay, and low computation. Two additional enhancement layers at 16 kbit/s each consist of two types of sublayers: a common sublayer for SWB expansion and a sublayer specific to each core layer for wideband enhancement. G.711.1-SWB, which is composed of G.711.1 and these enhancement layers, was approved as G.711.1 Annex D and G.722-SWB was approved as G.722 Annex B. This work item was developed by five organizations: NTT, ETRI, France Telecom, Huawei Technologies, and VoiceAge.

Specifications of the above recommendations—usage, bitrate, delay, and computational complexity—are listed in **Table 1**.

The standards for stereo telecommunication are now being studied as the latest work items. Specifically, stereo extensions to the above extensions, (7) and (8), are aimed at videoconferencing systems and telepresence systems. The former is planned to be approved in 2012, and the latter was scheduled to be completed in 2011.

These recommendations are intended to be implemented in suitable digital signal processing chips and their algorithms are described in fixed-point arithmetic. Once the fixed-point recommendations have been completed, floating-point recommendations, which are guaranteed to be compatible with the fixed-point ones, would usually be planned for implementation on personal computers and for assistance in understanding the algorithms. For example, a floating-point implementation of G.711.1 was standardized as G.711.1 Annex A.

Table 1. Speech coding recommendations approved after previous study period.

Name	G.722.1 Annex C	G.729.1	G.711.1	G.718	G.719	G.711.0	G.718 Annex B	G.729.1 Annex E	G.711.1 Annex D	G.722 Annex B
Approval	May 2005	Apr. 2006	Mar. 2008	Jun. 2008	Jun. 2008	Sep. 2009	Mar. 2010		Nov. 2010	
Audio bandwidth	50 Hz to 14 kHz	50 Hz to 7 kHz	50 Hz to 7 kHz	50 Hz to 7 kHz	20 Hz to 20 kHz	300 Hz to 3.4 kHz	50 Hz to 14 kHz		50 Hz to 14 kHz	
Bitrate (kbit/s)	24, 32, 48	8–32	64, 80, 96	8–32	32–128	Not fixed (50% on average)	36–48	36–64	96, 112, 128	64, 80, 96
Frame length (ms)	20	20	5	20	20	5–40	20	20	5	5
Algorithmic delay (ms)	40	48.9375	11.875	43.875	40	Same as frame length	49.625	55.6875	12.8125	12.3125
Computational complexity (WMOPS)	11	35.8	8.7	57	21	1.67	80	63	21.498	22.76
Technology	MLT	G.729, MDCT, BWE, SVQ	G.711, MDCT, Interleave VQ	ACELP, MDCT, AVQ	Adaptive MDCT, FLVQ	Lossless compression	G.718, MDCT, Sinusoidal coding	G.729.1, MDCT, Sinusoidal coding	G.711.1, MDCT, BWE, AVQ	G.722, MDCT, BWE, AVQ
Usage	Teleconfer-encing and telepresence systems	VoIP	NGN and VoIP	VoIP	Teleconfer-encing and telepresence systems	Routers on backbone network	Teleconferencing and telepresence systems		Teleconferencing and telepresence systems	
Other features		Interoperable with G.729	Interoperable with G.711			Lossless coding of G.711 bitstream	Interoperable with G.718	Interoperable with G.729 and G.729.1	Interoperable with G.711 and G.711.1	Interoperable with G.722

ACELP: algebraic code excited linear prediction
AVQ: algebraic vector quantization
BWE: bandwidth extension
FLVQ: fast lattice vector quantization
MDCT: modified discrete cosine transform
MLT: modulated lapped transform
NGN: Next Generation Network
SVQ: spherical vector quantization
WMOPS: weighted millions of operations per second

4. Related speech processing standardization

As mentioned above, assuring interoperability and compatibility in telecommunication is the primary purpose of an ITU-T standard while keeping communication quality appropriate could be another purpose even if it is unrelated to the primary purpose. For example, methods such as packet loss concealment (PLC) and voice activity detection (VAD), which have no impact on interoperability, are usually integrated into the coding algorithm in order to qualify the performance requirements, but those methods could be standardized separately from the main body of the recommendation as Appendices.

(1) PLC

When packets containing the speech bitstream are lost on the IP network, the packet loss could be audible as an interruption in the speech. PLC makes the break inaudible by reproducing the lost speech. The original speech coding standard was made when the network was assumed to have no packet loss, so it did not have any PLC; therefore, it was newly developed as an Appendix. G.722 Appendices III and IV approved in November 2006 are examples of appendices that were established much later than the main body.

(2) VAD

The classification of input signals as speech, silence, etc. and use of a coding method appropriate to the signal's characteristics, such as not transmitting any information during silent periods, enable very efficient coding to be achieved. The input detection and classification method is called VAD. In the coding standards with VAD, e.g., G.729 Annex B and G.722.2, the algorithms, which are specifically designed to meet requirements, and the processing using the information, e.g., silence compression, are described together. In Q.8, the concept of a generic sound activity detector (GSAD), independent of the coding algorithm and applicable to general use, was newly defined; the first recommendation, which was based on this concept, was approved as G.720.1 in January 2010. It enables a 10-ms period of input to be classified into one of four categories: speech, music, noise, or silence.

Besides the speech processing methods, some other

recommendations were developed in order to support the standardization work. G.191 provides lots of tools, which are required for evaluating whether an algorithm meets the requirements, e.g., level adjustment, sampling rate conversion, filtering, error pattern generation, and fixed-point arithmetic library. They have been updated regularly and the current version of the set of tools is STL2009 (Software Tool Library 2009).

5. Use of ITU-T speech coding standard

This section describes how to obtain and use ITU-T speech coding standards. Most of the recommendations can be downloaded from the ITU-T website [1] except for preliminary versions. Since the current recommendations consist of recommendation text and source code, it is easy to evaluate the performance, e.g., speech quality and interoperability.

The recommendations are freely available for the purpose of standardization work and performance evaluation: however, for business use, a license must be obtained from the licenser holding the intellectual property rights, such as patents and software copyrights, included in the recommendations, which usually involves a fee set by the licenser. As mentioned above, several organizations jointly contributed to making some standards, so there are often multiple licensers per recommendation. To avoid the trouble of contacting all of them independently, the licensers have established a patent pool that provides a one-stop service for licensees.

6. Future activities

To meet market needs, new work items will progress in ITU-T such as speech and audio coding for multichannel telecommunication, NGN (Next Generation Network), FMC (fixed mobile convergence), and other audio processing for improving the quality of telecommunication.

NTT has participated in ITU-T standardization, such as contributing to the development of G.711.1 and providing the rapporteur for Q.10, and will continue to make further contributions to the standardization work in order to provide better telecommunication services to its customers.

Reference

[1] ITU downloads. http://www.itu.int/rec/T-REC/en

Shigeaki Sasaki
Senior Research Engineer, Speech, Acoustics and Language Laboratory, NTT Cyber Space Laboratories.
He received the B.E. degree in physics from Kyoto University in 1991. Since joining NTT in 1991, he has been engaged in the research field of wideband speech coding. He received the Achievement Award from the Institute of Electronics, Information and Communication Engineers (IEICE) in 2009 and the Maejima Award from the Teishin Association in 2010. He is a member of IEICE and the Acoustical Society of Japan.

ICC2011 was Held in Kyoto—With Wishes for Recovery from the Earthquake

Osamu Kamatani†, *Takashi Shimizu, Hisaya Hadama, Hirotaka Nakamura, Katsuhiro Ishimaru, and Takeshi Yamada*

Abstract

Despite the Great East Japan Earthquake (also known as the 2011 Tohoku Earthquake and Tsunami), the International Conference on Communications 2011 (ICC2011) was held on June 5–9, 2011, at Kyoto International Conference Center (ICC Kyoto) in Kyoto with wishes for recovery from the earthquake. The NTT Group had a booth and showed the latest R&D products in the co-located ICC Kyoto Exhibition. This article reports on ICC2011 and ICC Kyoto Exhibition.

1. Introduction

The International Conference on Communications 2011 (ICC2011) [1], [2] was held on June 5–9, 2011 at Kyoto International Conference Center (ICC Kyoto) in Kyoto, Japan. Following the tradition of Kyoto, which was the ancient capital of Japan for 1200 years, the theme of the conference was chosen as "Source of Innovation: Back to the Origin". With this theme, the conference was intended to provide opportunities to rethink traditions and conventions, which could lead to epoch-making innovations. Noritaka Uji, Representative Director and Senior Executive Vice President of NTT, acted as the General Chair of the conference, which was sponsored by the IEICE Communications Society, the IEEE Communications Society, and the Science Council of Japan, in cooperation with Kyoto Prefecture and Kyoto City (IEICE: Institute of Electronics, Information and Communications Engineers). This article describes the main topics of the conference.

† NTT Network Innovation Laboratories
Yokosuka-shi, 239-0847 Japan

His Imperial Highness Prince Akishino giving remarks at the commemorative ceremony.

Because of the Great East Japan Earthquake (also known as the 2011 Tohoku Earthquake and Tsunami) on March 11, 2011, the conference was almost called off. However, backed up by warm and cheerful supporters from around the world, the conference was regarded as a cornerstone of recovery, and a

Noritaka Uji, the General Chair and Senior Executive Vice President of NTT, giving greetings at the opening plenary.

Ichiro Kanazawa, Chairman of Science Council of Japan, giving greetings.

Prof. Byeong Gi Lee, President of the IEEE Communications Society, giving greetings.

Kazuo Hagimoto, President of the IEICE Communications Society and the Director of NTT Science and Core Technology Laboratory Group, giving greetings.

Prof. Kouichi Asatani, Executive Chair, leading the meeting.

commemorative ceremony was performed in the presence of His Imperial Highness Prince Akishino, resulting in powerful messages for recovery being sent out to the world.

2. Commemorative ceremony

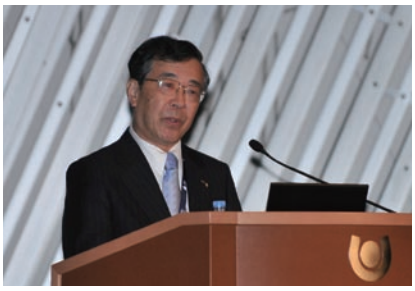
Wishing for recovery from the earthquake and regarding the conference as its cornerstone, the conference organizers arranged a special program called the commemorative ceremony for the morning of June 7. It was our great honor and privilege to have the presence of His Imperial Highness Prince Akishino at this ceremony. The distinguished guests also included the Vice Governor of Kyoto Prefecture, Shuichi Yamauchi; the Vice-Mayor of Kyoto City, Fumihiko Yuki; and the President of the Science Council of Japan, Ichiro Kanazawa. Reflecting on the progress of the recovery effort from the Great East Japan Earthquake, Prince Akishino offered special appreciation to the packed audience for participating in this conference and extended his heartfelt sympathy to all that were affected, while emphasizing the essential role of this conference in achieving an advanced information and communications technology (ICT) society. In the solemn atmosphere, the hall was filled with tremendous applause.

3. Opening plenary

ICC2011 officially commenced with the opening plenary on June 6. After being introduced by the Executive Chair, Prof. Kouichi Asatani, the General Chair, Noritaka Uji of NTT, gave the welcome address in which he expressed his sympathy to those affected by the earthquake as well as thanks to those who had supported Japan's recovery effort and to the participants for coming to this conference. Looking back at its history since 1987, when GLOBECOM was first held in Tokyo and the first mobile phone service started in Japan, he identified the valuable contribution of this community; it has played an essential role in the advancement of ICT. However, he also emphasized that this recent disaster reacquainted us with the challenges of the real world and

Ryuji Yamada, President and CEO of NTT DOCOMO, giving a keynote address.

Prof. Maurizio Decina, Politecnico di Milano, giving a keynote address.



Dr. Toshitaka Tsuda, Fujitsu Laboratories, giving a keynote address.

that further contributions were needed from this community to achieve sustainable societies.

In their remarks, the President of the IEEE Communications Society, Prof. Byeong Gi Lee, and the Technical Program Committee (TCP) Chair and President of IEICE Communications Society, Kazuo Hagimoto, also expressed their sympathy as well as their thoughts on future directions of research and development. It was an impressive occasion when all of the participants joined in a moment of silence at the request of Prof. Lee.

4. Technical program

Despite the concerns and worries created by the earthquake and tsunami and by the subsequent nuclear power plant problems, more than 1800 researchers and engineers from all over the world participated in the conference and actively discussed the latest developments in ICT. The majority of the participants were from Japan (31%) followed by the USA (15%), China (9%), and Korea (6%) (**Fig. 1**).

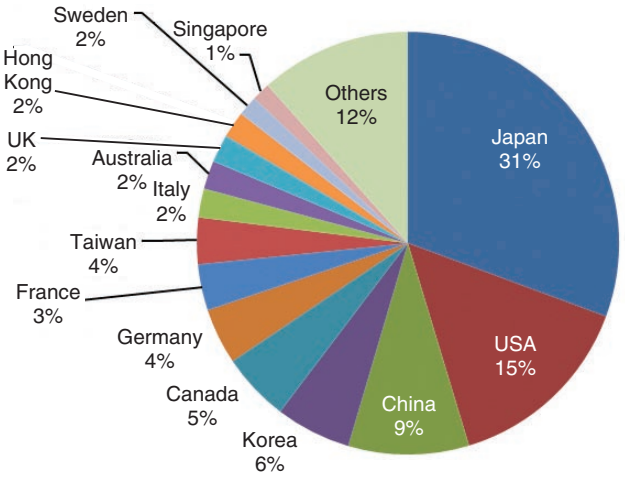


Fig. 1. ICC2011 participants by countries and regions.

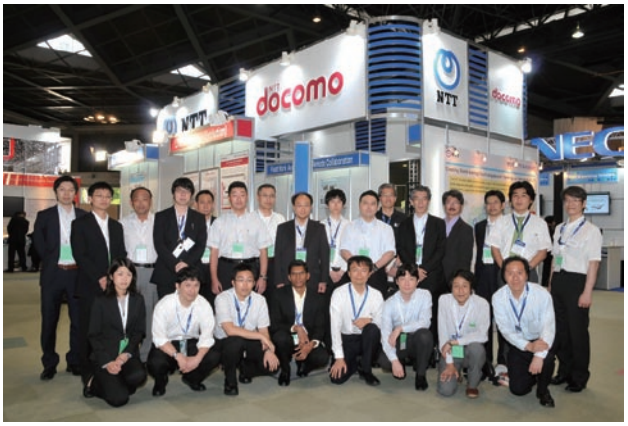
4.1 Keynote addresses

Three executive leaders from industry and academia were selected as the ICC2011 keynote speakers.

Ryuji Yamada, President and CEO of NTT DOCOMO, described his company's actions for new growth. First, he detailed the company's response to the earthquake, highlighting the fact that communications were restored to the pre-disaster level by the end of April through the use of large-zone base stations, satellite links, and high-performance antennas. Noting the proliferation of smartphones and the new Xi™ (pronounced crossy) service provided by NTT DOCOMO, which is Japan's first LTE (Long Term Evolution) mobile service, he introduced the concept of collaboration between the device and the network as an enabler of new services. The demonstration of the Translator Phone was impressive enough to convince the audience of the veracity of his remarks.

Prof. Maurizio Decina, Politecnico di Milano, described his vision of future networks and services by introducing the latest developments of enabling technologies for the Internet of things. He envisioned the need for a flatter and much more densely interconnected Internet for things. At the end of his talk, he eloquently appealed to the audience, shouting: "Ganbare, Nippon" (Stay Strong, Japan).

Dr. Toshitaka Tsuda, Fujitsu Laboratories, spoke of the ICT paradigm shift and communications technology trends. Showcasing innovative enabling technologies, he proposed a new paradigm called the



Booth of NTT Group at Kyoto Exhibition.

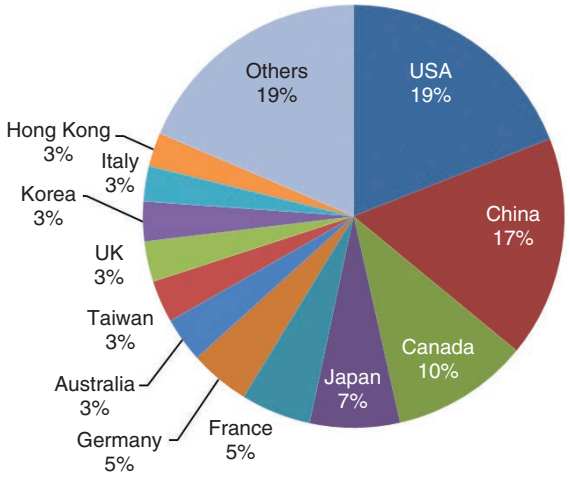


Fig . 2 Accepted papers by countries and regions.

human-centric intelligent society, which will create a new value chain for a more dependable, intelligent ICT society.

4.2 Technical symposia

The TPC received 2838 technical submissions from 64 countries and regions and selected 1092 papers for presentation. The overall acceptance rate was 38.5%. For this selection, each paper was carefully peer-reviewed by at least three reviewers, all of whom worked closely with the 47 symposium co-chairs and 2254 TPC members in 11 symposia and three tracks. The accepted papers were presented in 198 sessions: 180 lecture-style sessions (900 papers) and 18 poster-style sessions (192 papers) (**Fig. 2**).

4.3 Tutorials

As the ICC2011 tutorial program, the best 16 tutorials were selected out of the 59 proposals submitted. The latest topics by high-profile lecturers attracted many attendees. Topics included evolution to 4G (fourth generation) wireless, cooperative wireless communications, next-generation Internet, visible light communications, and vehicular networking. ICC2011 offered one free tutorial to each attendee with a full or limited registration and more than 900 people were registered.

4.4 Workshops

Nine full- or half-day workshops were organized for ICC2011. 81 papers were selected for presentation out of 171 submissions. The topics included heterogeneous networks, smart grid communications, future networks, and green communications. Those workshops were organized by small groups of researchers interested in brand new topics. Extensive discussion continued even after the scheduled program ended.

4.5 Business forum

The Business Forum was organized to create a venue for facilitating collaborations between attendees from academia and industry. The program consisted of two parts: executive and general sessions. In the executive sessions, high-profile leaders were invited to present their visions and strategies for corporate business or emerging research topics. General sessions were organized as panels on emerging topics, such as smart grid, next-generation access networks, wireless sensor networks, green ICT, and optical devices. The attendance at sessions was roughly 30–100 people, who had interesting discussions.

4.6 ICC Kyoto Exhibition

The ICC Kyoto Exhibition was organized by Hiromichi Shinohara, Director and Senior Vice President, Director of Research and Development Planning Department of NTT, and co-located with ICC2011 in the event hall of ICC Kyoto on June 6–8. For this exhibition, ten exhibitors and two publishers were registered. Nine of the world's leading companies, including the NTT Group, and a national research institute working in all areas of telecommunications and networking chose the largest available booths (floor space: 54 m^2) to exhibit their latest

products, services, and applications. The many demonstrations on next-generation optical network technologies, LTE technologies, augmented reality technologies, green ICT, and Home ICT attracted many researchers and engineers to ICC2011. The NTT Group exhibited its state-of-the-art ICT on Japanese advanced broadband networks. As application technologies on the Next Generation Network, it exhibited a Home ICT platform using OSGi [3] technology (NTT Cyber Solutions Labs.), broadband wireless home networks (NTT Network Innovation Labs.), intelligent cryptosystems (NTT Information Sharing Platform Labs.), and field work assistance by remote collaboration (NTT Service Integration Labs.). Ultrahigh-speed optical network technologies were introduced for ultrahigh-capacity optical transport networks (NTT Network Innovation Labs. and NTT Photonics Labs.) as well as research and development of holey fibers (NTT Access Network Service Systems Labs.). Xi™, Japan's first LTE mobile service (NTT DOCOMO), was also presented. The origami corner, the lounge with free Internet personal computer stations and free drinks station, the lucky draw, and the student travel grant award ceremony organized by ICC2011 were also useful in creating an enjoyable time. The ICC Kyoto Exhibition was wildly successful and the total number of participants over the three days came to 2600.

5. Social events

Kyoto was the capital of Japan for 1200 years and is known as the cultural heart of Japan. With its history and legacy of tradition and elegance, Kyoto was considered to be the best place for this conference. The venue, the Kyoto International Conference Center (ICC Kyoto), was the first national conference facility in Japan when it opened in 1966. Its unique design let participants enjoy Japanese traditional architecture as well as modern decoration and furnishings.

On June 5, attendees were welcomed by a performance of *koto*, a Japanese traditional musical instrument, played by ladies dressed in Japanese kimono. With traditional Japanese cuisine such as sushi served at a small wooden booth, attendees met and talked with friends and new acquaintances from all over the world until late in the evening.

Attendees who participated in the conference banquet at the Grand Prince Hotel Kyoto on June 7 enjoyed Kyoto's traditional dance and a karate demonstration by the national team of Japan. The performance was highly impressive and so breathtaking that the stage soon rang with sustained applause.

6. Conclusions

With the presence of His Imperial Highness Prince Akishino, the International Conference on Communications 2011 (ICC2011) was successfully held in Kyoto by overcoming the difficulties of the Great East Japan Earthquake. It turned out to be an unforgettable event in the history of Japan's ICT for the participants, organizers, and supporters including the NTT Group. It was headed by Noritaka Uji, Senior Executive Vice President of NTT, who was General Chair of the conference. The next GLOBECOM (GLOBECOM 2011) will be held in Houston, Texas, USA, in December and ICC2012 will be held in Ottawa, Canada, next June.

References

[1] ICC2011. http://www.ieee-icc.org/2011/
[2] IEEE conference report. http://dl.comsoc.org/livepubs/ci1/public/2011/jul/pdf/ciconfreport.pdf
[3] OSGi. http://www.osgi.org/

Osamu Kamatani
Senior Research Engineer, Media Innovation Laboratory, NTT Network Innovation Laboratories.

He received the B.E., M.E., and Ph.D. degrees in electrical engineering from the University of Tokyo in 1988, 1990, and 1993, respectively. He joined NTT Transmission Systems Laboratories in 1993 and worked on high-speed optical time division multiplexed transmission systems, optical networking architectures, and high-speed interactive data transfer applications. He moved to NTT Service Integration Laboratories in 2004 and worked on the NGN architecture and transport control protocols and was also involved in related standardization activities in ITU-T and ETSI TISPAN. His current interest is in terabit-per-second-class high-end computing systems based on multi-core processors and a future network architecture that can offer end-to-end terabit-per-second scalable data processing. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) Communications Society.

Takashi Shimizu
Executive Research Engineer, Research Laboratories, NTT DOCOMO.

He received the B.S., M.S., and Dr.Eng. degrees in electrical engineering from Tokyo Institute of Technology in 1991, 1993, and 1996, respectively. In 1996, he joined NTT Optical Network Systems Laboratories, (now NTT Network Innovation Laboratories), Kanagawa, Japan. Since then, he has been engaged in researching packet switching technology for high-speed computer communications. Since 2011, he has been with NTT DOCOMO, where he has been engaged in R&D of next-generation mobile networks. His research interests include a flow-based network architecture, hardware and software system design techniques, and reliability engineering for distributed systems. He is a member of the IEICE Communications Society, the IEEE Communications Society, and the Association of Computing Machinery (ACM) SIGCOMM.

Hisaya Hadama
Executive Manager, NTT Network Innovation Laboratories.

He received the B.S. and M.S. degrees from Kyushu University and the D.Eng. degree from Osaka University in 1985, 1987, and 1997, respectively. After joining NTT in 1987, he engaged in research on ATM virtual path transport networks and multimedia services over broadband networks. From 1994 to 1995, he researched multimedia network control techniques as a visiting researcher at CTR, Columbia University, USA. After returning to Japan, he worked to develop NTT's R&D strategy, which aimed to provide affordable broadband network services. From 2003 to 2005, he was engaged in research on ubiquitous network service systems, including RFID tag-based healthcare systems and the wide area ubiquitous network. From 2007 to 2009, he worked on the architectural design of future carrier networks in NTT Network Innovation Laboratories. From 2009 to 2011, his research activities focused on developing next-generation FTTH systems. He is a member of IEICE, the Information Processing Society of Japan (IPSJ), and IEEE.

Hirotaka Nakamura
Research Engineer, NTT Access Network Service Systems Laboratories

He received the B.E. degree in applied physics from the University of Tokyo and the Ph.D. degree from Hokkaido University in 1999 and 2011, respectively. In 1999, he joined NTT Access Network Service Systems Laboratories, where he engaged in research on WDM-based high-speed optical access systems. He has been participating in FSAN activities related to NG-PON2 since 2009. He is a member of the IEICE Communications Society.

Katsuhiro Ishimaru
Manager, Research Planning Department, NTT Science and Core Technology Laboratory Group.

He received the B.S. and M.S. degrees in electro-communication engineering from the University of Electro-communications, Tokyo, in 1987 and 1991, respectively. Since joining NTT Laboratories in 1991, he has been active in R&D of information theory, digital archive platforms for cultural heritage, and super-high-definition digital cinema systems. He is a member of IEEE and IEICE.

Takeshi Yamada
Senior Manager, Research Planning Department, NTT Science and Core Technology Laboratory Group.

He received the B.S. degree in mathematics from the University of Tokyo in 1988 and the Ph.D. degree in informatics from Kyoto University in 2003. He joined NTT in 1988. His research interests are in machine learning, data mining, and combinatorial optimization. He is a member of ACM, IEEE, IEICE, IPSJ, and the Scheduling Society of Japan.

Report on NTT Communications Science Laboratories Open House 2011

Kaoru Hiramatsu, Hitoshi Sakano†, Isamu Motoyoshi, Hideki Sakurada, Tessei Kobayashi, and Shinji Watanabe

Abstract

We report on Open House 2011 held at NTT Communications Science Laboratories in Keihanna Science City, Kyoto, which welcomed over 1100 people on June 9 and 10. Visitors attended excellent lectures and colorful demonstrations based on the latest research results.

1. Overview

Open House at NTT Communications Science Laboratories in Keihanna Science City, Kyoto (hereinafter referred to as the open house) is an annual event. This year, it welcomed over 1100 visitors on June 9 and 10 [1], [2]. The theme of this year's event was the 20th Anniversary of NTT Communications Science Laboratories (CS Labs). The CS Labs, which belongs to the Science and Core Technology Laboratory Group, is a source of new principles and concepts in fields including media and communications, information science, and human science. It is accommodated at two locations: Keihanna Science City in Seika-cho, Kyoto, and Atsugi City in Kanagawa Prefecture.

The open house receives many visitors every year. The visitors during the course of this year's open house included NTT Group employees and people from companies, universities, and research facilities engaged in research and development, business, and education. The presentations and exhibitions are introduced below. More details about some of the topics can be found in the Feature Articles in this issue.

† NTT Communications Science Laboratories
Soraku-gun, 619-0237 Japan

2. 20th Anniversary of CS Labs

Noritaka Uji, Representative Director and Senior Executive Vice President of NTT, opened the event with words of welcome (**Photo 1**). He thanked all those that have supported research at CS Labs leading up to its 20th anniversary and asked for their continued support. He then described his hopes for future basic research.

Makoto Nagao, the Chief Librarian at the National Diet Library, then gave a message in celebration of the 20th anniversary.

Finally, Naonori Ueda, the Director of CS Labs, offered a greeting and discussed some of CS Labs' history, recent research results, and future research directions, interwoven with overviews of some of the exhibits and presentations featured in the open house (**Photo 2**).

The history of CS Labs was described in a booklet commemorating the 20th anniversary that was given to each visitor, and a related presentation in a special booth showed a timeline displaying organizational transitions and major research results (**Fig. 1**).

Photo 1. Noritaka Uji, NTT Representative Director and Senior Executive Vice President, giving an address.

Photo 2. Naonori Ueda, Director of CS Labs, giving an address.

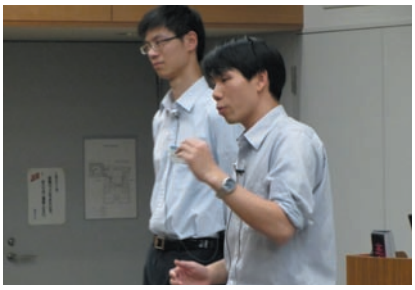


Photo 3. Research presentation by Dr. K. Duh and Dr. A. C. Au Yeung.

Photo 4. Research presentation by Dr. N. Yamashita.

Fig. 1. Booklet commemorating the 20th anniversary and related document.

3. Research presentations

Five research presentations were given, highlighting recent significant research results and high-profile research themes.

- "Supporting Communication in Global Networks—Machine Translation and Social Network Mining," Kevin Duh and Albert Ching-man Au Yeung (Innovative Communication Laboratory)
- "Formal Methods for Security and Privacy," Yasuyuki Tsukada (Innovative Communication Laboratory)
- "Human Interfaces Based on Applied Sensory-illusion Engineering," Tomohiro Amemiya (Human and Information Science Laboratory)
- "Achievement for Ambient Intelligence Project—Steps towards IT Environment Appropriately Communicating with Humans," Yasuhiro Minami (Media Information Laboratory)
- "Interaction: Invention? Intellection!," Naomi Yamashita (Media Information Laboratory)

Each presentation introduced some of the latest research results, including some background and an overview of the research. All the talks were very well received by the many participants.

"Supporting Communication in Global Networks" provided an overview and discussed some of the latest trends in communications analysis technology. Presented in English, it focused on technology utilizing the huge number of messages being exchanged over the Internet and statistical machine translation, which works on the basis of huge quantities of data (**Photo 3**).

"Interaction: Invention? Intellection!" introduced the latest results in interaction research at the CS Labs and raised questions regarding what contributions researchers need to make to the research community and what the research field needs in order to grow (**Photo 4**).

Photo 5. Researcher explaining the latest results directly to a visitor.

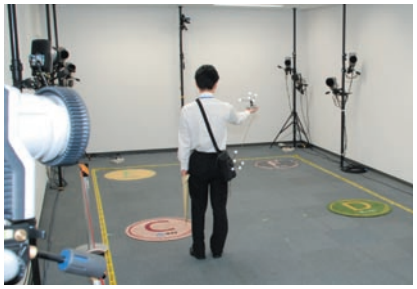


Photo 6. Visitor holding the Buru-Navi device and trying the indoor navigation technology.

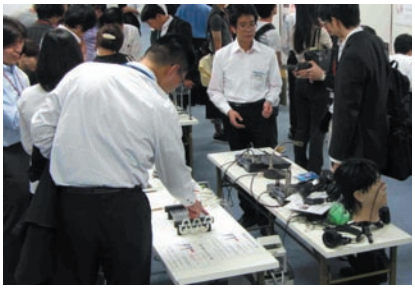


Photo 7. Visitors trying touchable illusions.

4. Research exhibits

The open house featured 30 exhibits displaying the latest research results and was divided into four zones: Moriya Research Laboratory, Innovative Communication Laboratory, Media Information Laboratory, and Human and Information Science Laboratory. Each exhibit was accommodated in a booth and used techniques such as slides on a large-screen monitor or hands-on demonstrations, with researchers explaining the latest results directly to visitors (**Photo 5**).

4.1 Moriya Research Laboratory

- Relationship between respiration and music: Possibility of music presentation that corresponds to respiration frequency

4.2 Innovative Communication Laboratory

- Enhancing the feasibility of a novel quantum computational model: Measurement-based method for quantum operations
- How to cut a cake fairly: Meta-envy-free cake-cutting protocols
- Aiming at promoting content circulation: A formal approach to interoperability between licenses
- Reading between the lines for extracting 5W1H[*1] from texts: Highly accurate analysis of subjects and objects for real-world texts
- More data are better data: High-performance dependency parsing using large quantities of text data
- How many points do you give for this translation? Why do you use BLEU[*2]?
- Early recognition methods for time sequence classification
- Efficient data gathering using singular value decomposition for sensor networks
- Recognizing human activities with magnetic fields: Finger-ring-type sensor device with magnetic sensors for healthcare and fitness monitoring

4.3 Media Information Laboratory

- Stereo 9-band camera system for accurately estimating object color and shape
- Link to media: Robust media search on smartphones and web browsers
- Instantaneous detection of identical image patterns: Fast image matching technique
- Detecting and annotating generic objects in the real world: Easy-to-use semi-supervised learning for generic object recognition
- How to train an agent using dialogue data: Statistical dialogue control for natural and satisfying interactions
- How conversational robots achieve smartly timed speech behavior: Adaptive timing control of robot behavior for smooth human-robot interaction
- Viscuit: Digital education tool for co-production and creation—Practice of visual programming language teaching on the network
- Bridging interaction analysis and system design: What has the t-Room solved, and what has it not yet solved?
- Optical dice: Compact and fast random bit generator using a laser chip
- Understanding empathy between people from

*1 5W1H: When, where, who, what, why, and how.

*2 BLEU: bilingual evaluation understudy

their smiles
- Automatic-speech-recognition-friendly speech enhancement: Noise-robustness techniques based on learning what natural speech is like
- It's OK to leave some noise! Unification of noise reduction technique and speech recognizer
- Speech recognition system that learns from errors: Discriminative model training for speech recognition
- Meeting assistance system that rapidly indicates who spoke when and what: Technology for analyzing the scene of a multi-speaker conversation

4.4 Human Information Science Laboratory
- The mechanism of visual motion integration over space
- Why complain when listening? Auditory tests that sense brain variability
- Touchable illusions: Tools for understanding human tactile perception
- Subliminal perception and underlying mechanism of mind
- Smooth writing by optic flow: How the implicit visuomotor system works
- Guide you to the desired place: Route navigation by pseudo-attraction force "Buru-Navi"

At the "Guide you to the desired place" exhibit, many visitors actually held the Buru-Navi device and tried out the indoor navigation technology (**Photo 6**). "Touchable illusions" introduced tactile illusions, which many visitors tried, and included a demonstration in which two strings attached to a tennis racquet were held in the hands. The strings moved up and down and produced a slippery feeling on the palms (**Photo 7**).

5. Invited talk

This year also featured an invited talk by Professor Mariko Hasegawa of the Graduate University for Advanced Studies. Professor Hasegawa spoke on the topic of "Birdsongs and human tweets: Evolution of human communication and ICT society" (ICT: information and communications technology) (**Photo 8**). She covered many topics, from the courtship behavior of peacocks and a comparison of language acquisition in humans and chimpanzees to the contribution of information technology (IT) to the evolution of human communications. She closed by raising the question of the role of communications technology in society and talking about her hopes for NTT in that regard.

Photo 8. Professor Mariko Hasegawa of the Graduate University for Advanced Studies giving an invited talk.

6. Concluding remarks

Many visitors from previous years came again to see this year's research results at CS Labs. They engaged in lively discussions regarding the research presentations and exhibits and provided many valuable opinions about the results. In closing, we would like to offer our sincere thanks to all the visitors and participants who attended this event.

References

[1] Open House website (in Japanese). http://www.kecl.ntt.co.jp/openhouse/2011/
[2] Open House poster. http://www.kecl.ntt.co.jp/openhouse/2011/oh2011_poster_en.pdf

Kaoru Hiramatsu
NTT Research and Development Planning Department.
He received the B.S. degree in electrical engineering and M.S. degree in computer science from Keio University, Kanagawa in 1994 and 1996, respectively, and the Ph.D. degree in informatics from Kyoto University in 2002. He joined NTT Communication Science Laboratories in 1996 and moved to the R&D Vision Group, NTT Research and Development Planning Department in 2011.

Hitoshi Sakano
Senior Research Scientist, Emergent Learning and System Research Group, NTT Communication Science Laboratories.
He received the B.S. degree in physics from Chuo University, Tokyo, in 1988, the M.S. degree in physics from Saitama University in 1990, and the Ph.D. degree in applied physics from Waseda University, Tokyo, in 2008. He joined NTT Communication Science Laboratories in 2008 and studied pattern recognition technology. He is a member of IEEE, the Information Processing Society of Japan (IPSJ), the Institute of Electronics, Information and Communication Engineers (IEICE), and the Physical Society of Japan (PSJ).

Isamu Motoyoshi
Research Scientist, Perception and Emotion Research Group, NTT Communication Science Laboratories.
He received the B.A., M.A., and Ph.D. degrees in psychology from Tohoku University, Miyagi, in 1994, 1996, and 1999, respectively. He joined NTT Communication Science Laboratories in 1999 and studied human visual perception. He was also a JSPS (Japan Society for the Promotion of Science) Fellow from 2000 to 2004 and a visiting scientist at McGill University, Canada, from 2002 to 2003. Since 2006, he has been a visiting associate professor of Tokyo Institute of Technology.

Hideki Sakurada
Research Scientist, Computing Theory Research Group, Innovative Communication Laboratory, NTT Communication Science Laboratories.
He received the B.Sc. degree in science and the M.Sc. degree in information science from Kyoto University in 1997 and 1999, respectively. He joined NTT Communication Science Laboratories in 1999. He has been working on formal methods and computer security.

Tessei Kobayashi
Research Scientist, Linguistic Intelligence Research Group, NTT Communication Science Laboratories.
He received the Ph.D. degree in psychology from the University of Tokyo in 2004 and joined NTT Communication Science Laboratories in 2005, studying children's linguistic and intelligence development. He was also a JSPS Fellow from 2002 to 2005 and a visiting researcher at Yerkes National Primate Research Center, Emory University, USA, from 2002 to 2003.

Shinji Watanabe
Research Scientist, Signal Processing Research Group, Media Information Laboratory, NTT Communication Science Laboratories.
He received the B.S., M.S., and Dr.Eng. degrees from Waseda University, Tokyo, in 1999, 2001, and 2006, respectively. Since joining NTT in 2001, he has been working at NTT Communication Science Laboratories. From January to March in 2009, he was a visiting scholar in Georgia Institute of Technology at Fred Juang's Laboratory. His research interests include Bayesian learning, pattern recognition, and speech and spoken language processing.

NTT Singapore
Partnership for Growth—The Preferred Choice for Trusted and Reliable IT Solutions

Karen Kang†

Marketing Executive, Sales and Marketing,

NTT Singapore Pte Ltd

Abstract

NTT Singapore was established in 1997. It started out as a systems integrator providing information technology (IT) services mainly to Japanese companies in Singapore. Over the years, it has evolved to provide a whole suite of information and communications technology (ICT) solutions to enterprises worldwide.

1. Introduction

Singapore is renowned as one of the most business-friendly environments. It is also rapidly emerging as an ideal location for the centralisation of services such as information and communications technology (ICT), finance, and logistics. The Singapore Government is making proactive efforts to facilitate the adoption of ICT technologies by various industries such as manufacturing, financial, logistics, and healthcare to enhance Singapore's competitiveness and position it as a regional hub [1]. Singapore is one of the key nodes in the global networks of many of the top technology companies: a strong testament to Singapore's strategic position as a global ICT hub (**Fig. 1**).

NTT Singapore [2], which established in 1997, is a subsidiary of NTT Communications (NTT Com) [3], the long-distance and international telecommunications provider in the NTT Group. It started out as a systems integrator providing information technology (IT) services mainly to Japanese companies in Singapore. Over the years, it has evolved to provide a whole suite of ICT solutions to enterprises worldwide (**Fig. 2**).

† NTT Singapore Pte Ltd
20 Cecil Street Equity Plaza #11-01/08 Singapore 049705

2. Growth and development

2.1 Partnerships and collaborations

In order to offer high-quality end-to-end solutions, NTT Singapore has been actively expanding to include more services to better serve the needs of its customers. Its parent company, NTT Com has acquired Integralis [4], which specializes in security services, to enhance its managed services. NTT Com has also acquired GlobeSoft Pte Ltd (Emerio) [5], a global IT outsourcing service provider, to reinforce

Fig. 1. Marina Bay Sands—Currently the No. 1 Singapore landmark.

Fig. 2. Equity Plaza Building, in which the NTT Singapore office is located.

its IT outsourcing capabilities around the world. With Dimension Data [6] having joined the NTT Group and with working partnerships having been established with StarHub and Orange Business, NTT Singapore is well positioned to serve the needs of multinational companies.

2.2 Investments

In April 2010, NTT Com's Arcstar™* Global IP-VPN became the first international IP-VPN service provided by a Japanese telecommunications company to support IPv4/IPv6 dual stack (IP: Internet protocol, VPN: virtual private network, v4: version 4, v6: version 6).

In response to the growing demand for cloud services, NTT Com offers Global Virtualization Services (GVS) [7] in several countries, including Singapore. NTT Singapore's GVS offers a suite of virtual hosting services for the Singapore market. It is a dedicated virtualisation infrastructure and virtualised network with improved server availability and enhanced network security. It enables end-users to manage their virtualised server clusters, which can be located in Singapore or elsewhere.

With the positioning of Singapore as the global ICT hub, NTT Com has invested in the Serangoon Data Center, its first greenfield datacentre outside Japan. We expect the new datacentre to significantly upgrade our world-class ICT solutions and quality of service for enterprise customers, especially for multinationals with multi-hub operations in Asia. Our datacentre will also boast a large-capacity backbone for high-speed, seamless rack-to-rack connectivity between core datacentres in Singapore, Hong Kong, Japan, and the USA.

The Asia Submarine-cable Express (ASE), constructed by a consortium of NTT Com, Malaysia-based Telekom Malaysia, Philippines-based PLDT, and Singapore-based StarHub, will serve Japan, Malaysia, Singapore, and the Philippines. A separate route to Hong Kong will be added in December 2012. ASE will be equipped with the latest 40-Gbit/s optical technology, which will boost the capacity and strengthen the redundancy of NTT Com's Asian cable networks. It can also incorporate 100-Gbit/s optical technology in the future. ASE is specially designed to withstand earthquakes and typhoons, particularly in areas such as the Bashi Channel south of Taiwan. Routes between Japan, Hong Kong, and Singapore will take the shortest possible distances to maximise reliability and minimise latency.

3. Services

3.1 ICT solutions portfolio

The main ICT solutions provided by NTT Singapore are listed in **Table 1** below.

3.2 One-stop solutions

NTT Singapore is rapidly expanding to cater to the

* Arcstar is a registered trademark of NTT Communications.

Table 1. ICT solutions portfolio.

Outline	Description	Services
Global networks	High-capacity, high-quality seamless connections worldwide	IP-VPN, Ethernet, Leased Line, eVLAN, Global IP Network (IP Transit)
Datacentre services	World-class facilities with hosting and managed services	Co-location, Connectivity, Global Virtualization Service, Hosted Firewall Service, Data Backup, and Storage Service, etc.
Global managed IT services	One-stop management for all IT assets	TeamWorks (Mail & Groupware), PC Management Service, IT Monitoring Services
Internet services	Back by Asia's Tier 1 IP backbone to ensure fast connectivity	NTT i-net (ISP)
Communications	Integrated communications that enhance collaboration and boost productivity	Unified Communications as a Service, ClearCall 1517 (IP Telephony)
Solutions integration	One-stop shop that provides hassle-free management of multi-vendor solutions	All the above and third-party products

ISP: Internet service provider
PC: personal computer
VLAN: virtual local area network

e-Kamo
Our NTT i-net Mascot
e-Kamo (from the Japanese *ii kamo*, meaning not bad).

needs of enterprises. The ability to offer a one-stop shop for customized solutions, coupled with NTT Com's strengths in global connectivity, makes NTT Singapore an ideal ICT partner for many enterprises. With new partnerships and with the Serangoon Data Center coming online soon, NTT Singapore is set to take centre stage in Singapore's ICT arena.

3.3 High-quality consumer Internet services

NTT Singapore has built a reputation as a leading provider of end-to-end network and IT solutions for multinational corporations. We currently provide the same high-quality solutions and exemplary customer service to residential customers as well. If you are planning to stay in Singapore, please sign up for our Internet plans [8].

References

[1] IDA Singapore (Infocomm Development Authority of Singapore, under the Government of Singapore). http://www.ida.gov.sg
[2] NTT Singapore. http://www.sg.ntt.com/en/about-us.html
[3] NTT Communications. http://www.ntt.com/
[4] NTT Com press release. http://www.ntt.com/aboutus_e/news/data/20091019.html
[5] NTT Com press release. http://www.ntt.com/aboutus_e/news/data/20100422.html
[6] Dimension Data. http://www.dimensiondata.com/
[7] GVS. http://www.sg.ntt.com/en/business/data-center-services/hosting-and-managed-services/global-virtualization-services.html
[8] http://www.sg.ntt.com/en/residential

NTT Singapore—short column

Singapore—a food paradise

Singapore is a multiracial country and its ethnic diversity is reflected in the food it offers. The food is influenced by the native Malay, Chinese, Indonesian, India, as well as Western traditions since the founding of Singapore by the British in the 19th century. Traces of cuisines such as Thai and Middle Eastern also exist in local food culture.

Lau Pa Sat Festival Market, the largest Hogarths (food stall village) market that opens 24 hours, 365 days.

Inside the Lau Pa Sat Festival Market.

What to eat

If you visit Singapore, the following are dishes you should try.

1) Hainanese Chicken Rice (though not available on Hainan island)
2) Singapore Laksa
3) Black Pepper Crab (a favourite among many Japanese expatriates)
4) Bak Kut The (pork ribs in spiced soup)
5) Rojak (a mixture of fried dough, pineapple, dried bean curd, fruits, and vegetables in prawn bean paste)

Chicken rice is a famous type of local food, as seen in a restaurant at Lau Pa Sat.

Menus at a food court at the Lau Pa Sat Festival Market.

Strange but true

Did you know that in Singapore's Central Business District you can reserve a seat simply by putting a packet of tissues on it? Owing to the limited time for lunch, people have started putting tissue packets on seats to reserve them while they queue at the food counter. This trend has led to some indigenous suppliers starting to print "Seat Reserved" on their tissue packets!

Other common items used to reserve seats include newspapers, umbrellas, and even company ID badges. If someone forgets to remove the tissues, the seat may be left empty all day. When you visit Singapore, do keep a packet of tissues with you. It will come in handy when you need to find a seat!

The packets of tissues mean the seats here are reserved.

External Awards

DICOMO 2011 Paper Awards

Winners: Takamitsu Narumi, Yasuhiko Yoshimura, Takeshi Miyasaka, and Kouichi Genda, NTT Network Service Systems Laboratories
Date: Aug. 28, 2011
Organization: Multimedia, Distributed, Cooperative, and Mobile Symposium

For "Adapting Dynamic Load Balancing Method for Video Streaming Services".

Although various load balancing methods have been proposed, they cannot deal with the rapid spread of information and the unanticipated surge in request rates. In this paper, we propose a method of dynamically balancing the load of video streaming services depending on the resources in the network and then analyze the effects of this method by simulation.

Published as: T. Narumi, Y. Yoshimura, T. Miyasaka, and K. Genda, "Adapting Dynamic Load Balancing Method for Video Streaming Services," Proc. of Multimedia, Distributed, Cooperative, and Mobile Symposium 2011 (DICOMO 2011), p. 509 (3F-1), Kyoto, Japan (in Japanese).

APNOMS Best Paper Award

Winner: Hiroshi Matsuura, NTT Service Integration Laboratories
Date: Sep. 23, 2011
Organization: Asia-Pacific Network Operations and Management Symposium 2011

For "Multipath Creation Algorithm Optimizing Traffic Dispersion on Networks".

A new heuristic k shortest simple path algorithm called k-shortest paths first (k-SPF) is proposed. It creates k shortest simple paths faster than the conventional Yen's algorithm does. In addition, its parameters can be changed effectively to achieve appropriate traffic dispersion without making any bottleneck links.

Published as: H. Matsuura, "Multipath Creation Algorithm Optimizing Traffic Dispersion on Networks," Proc. of the 13th Asia-Pacific Network Operations and Management Symposium (APNOMS2011), Taipei, Taiwan, 2011.

Papers Published in Technical Journals and Conference Proceedings

An SoC Demonstration of ONU Discovery and Dynamic Bandwidth Allocation for 10G/1G Dual-rate 10G-EPON

M. Nakanishi, K. Kawai, J. Kato, N. Miura, A. Miyazaki, H. Kamitsuna, H. Katsurai, N. Tanaka, Y. Ohtomo, M. Urano, and T. Shibata

Proc. of the National Fiber Optic Engineers Conference (NFOEC), Los Angeles, CA, USA, 2011.

Newly developed dual-rate optical line terminal (OLT) and optical network unit (ONU) system-on-a-chip devices (SoCs) combined with our latest transceivers demonstrate highly efficient 10G/1G (G: Gbit/s) ONU simultaneous discovery processing and hardware-accelerated dynamic bandwidth allocation of the transmission time according to the requirements.

A Heuristic Algorithm for Reducing System-level Test Vectors with High Branch Coverage

K. Yamazaki, Y. Sekihara, T. Aoki, E. Hosoya, and A. Onozawa

Proc. of the IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1475–1478, Rio de Janeiro, Brazil, 2011.

We introduce a heuristic that generates as few test vectors as possible with high branch coverage for the functional verification of digital design. The challenge is how to save time and effort while achieving sufficient verification at the system level. We focus on generating test vectors from a circuit specification written in C. We reuse them in a SystemC description by removing their redundancies while maintaining the branch coverage as is. Experimental results for our practical design show that over 90% on average of the redundant test vectors were reduced with 100% branch coverage maintained. The reused test vectors for SystemC Bus Cycle Accurate models scored 80% on average for branch coverage. These results are significant for saving verification cost and beneficial for simplifying debugging work.

Stack Queue Mixed Layout of Bipartite Graph Subdivisions

M. Miyauchi and H. Enomoto

Proc. of the 14th Korea-Japan Joint Workshop on Algorithms and Computation (WAAC2011), pp. 153–158, Busan, Korea, 2011.

This paper studies the problem of stack queue mixed layout of bipartite graph subdivisions. Dujmovic and Wood showed that for all integers s, $q > 0$, every graph G has an s-stack q-queue mixed subdivision layout with at most either $4[\log_{(s+q)q} sn(G)]$ division vertices per edge, where $sn(G)$ is the number of stacks of G, or $2 + 4[\log_{(s+q)q} qn(G)]$)

division vertices per edge, where $qn(G)$ is the number of queues of G. This paper improves upon their result for complete bipartite graphs $K_{m,n}$ ($m \geq n$) with m and n partite sets and shows an algorithm for constructing an s-stack q-queue mixed subdivision layout with at most $2[\log_{(s+q)q}n]$-1 division vertices per edge.

S^3: Smart Shadow System for Real World Service and Its Evaluation with Users

K. Umakoshi, T. Kambayashi, M. Yoshida, M. Takemoto, and M. Matsuo

Proc. of the 2011 IEEE/IPSJ International Symposium on Applications and the Internet, Vol. 1, No. 1, pp. 394–401, Munich, Bavaria, Germany.

Many studies have been focusing on building smart environments, which provide useful real world services (RWSs). With RWSs, however, feature interactions (FIs), device conflicts, and service conflicts occur more often because many devices run automatically. For this FI problem, we have been developing a Smart Shadow System (S3), which provides RWSs and can dynamically detect and resolve FIs. We describe an experiment in which general users tried RWSs provided by S3 in two scenarios. The purpose of one of the scenarios was to clarify the convenience of RWSs and that of the other was to determine the acceptability of FI resolution approaches. We obtained results from three viewpoints and analyzed them according to the two purposes mentioned above. We conclude that RWSs are convenient for general users and that an approach for resolving FIs through user interaction is the most acceptable for general users. Finally, we discuss issues that need to be addressed.

Efficient Data Selection for Speech Recognition Based on Prior Confidence Estimation

S. Kobashikawa, T. Asami, Y. Yamaguchi, H. Masataki, and S. Takahashi

Acoustical Science and Technology, Vol. 32, No. 4, pp. 151–153, 2011.

This paper proposes a technique that selects data to be recognized before speech recognition on the basis of rapid prior confidence estimation by using context independent and speech models. Experimental results show that this technique achieves the equivalent data selection performance at 30 times the speed of the conventional posterior confidence measure after speech recognition.

Cheat-sensitive Commitment of a Classical Bit Coded in a block of $m \times n$ Round-trip Qubits

K. Shimizu, H. Fukasaka, K. Tamaki, and N. Imoto

Phys. Rev. A, Vol. 84, No. 2, p. 022308, 2011.

This paper proposes a quantum protocol for a cheat-sensitive commitment of a classical bit. The receiver of the bit (Alice) can determine that the bit sender (Bob) is being dishonest if he changes or postpones his choice. Moreover, Bob can identify a dishonest Alice who violates concealment. For each round-trip case, Alice sends one of two spin states $|S\pm\rangle$ by choosing basis S at random from two conjugate bases X and Y. Bob chooses basis $C \in \{X,Y\}$ to perform a measurement and returns the resultant state $|C\pm\rangle$. Alice then performs a measurement with the other basis R ($\neq$S) and obtains the outcome $|R\pm\rangle$. In the opening phase, she can discover dishonest Bob who unveils a wrong basis with a faked spin state, or Bob can discover a dishonest Alice who infers basis C but destroys $|C\pm\rangle$ by setting R to be identical to S in the commitment phase. If a classical bit is coded in a block of m × n qubit particles, impartial examinations and probabilistic security criteria can be achieved.

Ultrasonic Earphone System Combined with Bone-conduction Vibration for Low Sound Leakage

M. Okamoto

Journal of INCE, Institute of Noise Control Engineering of Japan (INCE/J), Vol. 35, No. 4, pp. 323–325, 2011 (in Japanese).

This paper describes an open-air type earphone system with flat frequency characteristics and little sound leakage. This earphone is a hybrid ultrasonic earphone system combined with an audible-band bone-conduction earphone.
