Developing Technologies for Services that Deliver the Excitement of Games Worldwide

Akihito Akutsu, Kota Hidaka, Masayuki Inoue, Naoki Ito, Tetsuya Yamaguchi, Shigeru Fujimura, and Atsushi Nakadaira

Abstract

NTT Service Evolution Laboratories has floated the concept of an ultra-high-presence service that will enable viewers to experience the feeling of actually being in a sporting venue or the like, wherever they are, and is therefore developing an immersive telepresence technology called *Kirari!*, towards its implementation. This article introduces the supporting technologies used with Kirari!, which involve ultra-high-presence media generation, spatial and environmental information synchronized delivery, and content license management.

Keywords: ultra-high-presence services, immersive telepresence, media synchronization technology

1. Introduction

NTT Service Evolution Laboratories floated the concept of delivering complete game spaces in real time—first within Japan and then to the world—with introduction intended for 2020, and is working on the immersive telepresence technology called *Kirari!* together with ultra-high-presence media generation technology, spatial and environmental information synchronized delivery technology, and content license management technology to support Kirari!. Our aim is to establish these technologies and implement a world where viewers can experience the feeling and excitement of being at a sporting venue, wherever they are.

2. Techniques behind ultra-high-presence services

2.1 Kirari! immersive telepresence technology

We have been working on the research and development (R&D) of the immersive telepresence technology called Kirari!, which is a technique for directly transmitting not just the images and sounds of players but also the environment in which the players exist. It also links real objects that exist in the environment, such as lighting and robots, at the transmission destination and reproduces them together with the sound by projection mapping. Kirari! is a combination of the next-generation video compression technology H.265/HEVC (High Efficiency Video Coding) [1], the lossless speech coding technology MPEG-4 ALS (Moving Picture Experts Group-4 Audio Lossless Coding) [2] that can compress audio information without distortion, and the highly realistic media synchronization technology Advanced MPEG Media Transport (MMT) that we have recently started researching and developing. The objective with Advanced MMT is to make it seem as if a game is being played in front of the viewers' eyes, in a form that far exceeds the reality enjoyed from a flat image. It does this by using synchronized reproduction that transmits a life-size image of the subject, for example, a player on a team, while transmitting all

A world enabling users to feel as if they were experiencing
the atmosphere of the sporting venue, wherever they are

The experience of physical
sensations such as "So fast!!"
and "So high!!" lights up the
spectators' eyes (*Kirari!*).

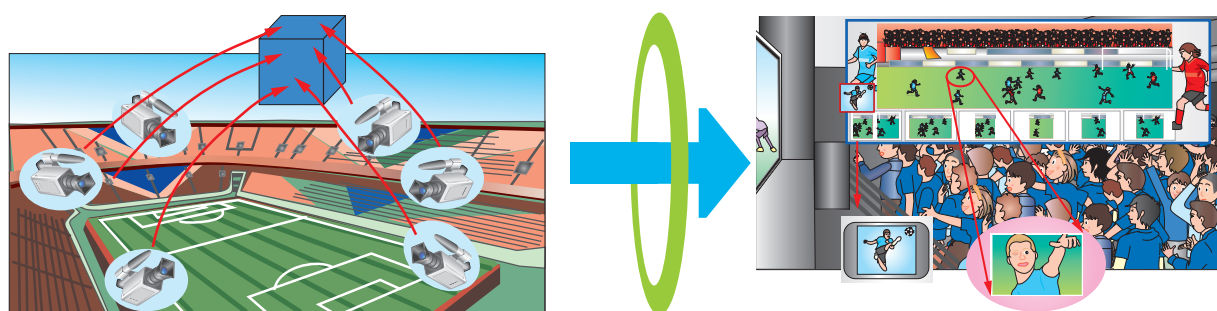Fig. 1.   The world of immersive telepresence technology *Kirari!*.

Fig. 2.   Images of the highly realistic service suited to the reproduction environment.

information other than the subject to fit into the dimensions of the transmission destination, such as a television (TV) screen in a bar or restaurant, even when the transmission destination facilities have different dimensions, lighting, and acoustics.

It is predicted that in 2020, many viewers will be able to share the excitement of sports from sites around the world on TV or by public viewing, due to the spread of 4K and 8K broadcasting, although there are limits to the realism that can be portrayed on a flat screen. That is why our intention is to apply Kirari! technology to provide viewers at multiple points around the world with the physical sensations of being at games, as if they were at the sporting venue at the same time (**Fig. 1**).

2.2   Ultra-high-presence media generation technology

Ultra-high-definition video is not limited to 4K

video. We intend to establish media processing technologies that can transmit ultra-high-definition video that surpasses even 8K video. More specifically, we aim to implement ultra-high-presence media generation technology that combines and integrates high-definition video and sound data that has been captured at a number of locations by a number of cameras, for reproduction using multiple display devices. This technology adjusts flexibly to the reproduction environment; it enables us to display content that is tailored to the user on a portable terminal such as a smartphone, or to display ultra-high-definition video that has been put together on a large screen at a public viewing venue, for example. It also reproduces a high level of realism that has not been seen before, by combining different kinds of media data such as video and sound (**Fig. 2**).

We previously implemented trials of 4K video transmission over a commercial network, taking

- Professional baseball relay broadcasting from Hanshin Koshien Stadium to Abeno Q's Mall
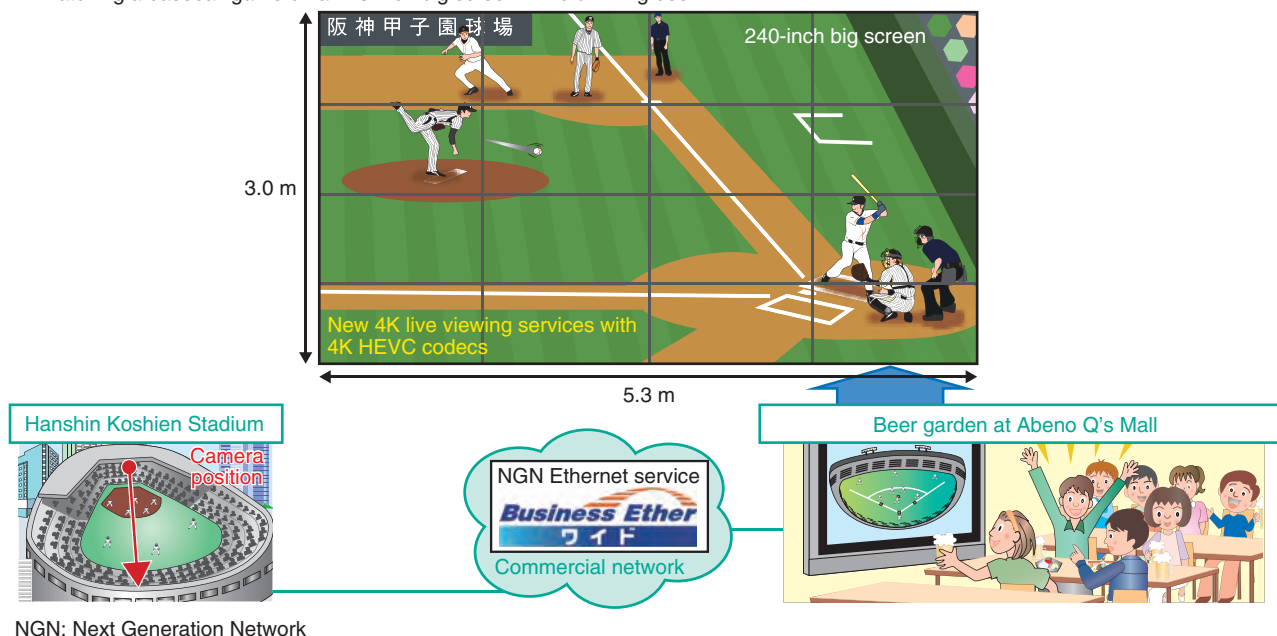- Watching a baseball game on a 240-inch big screen while drinking beer!

Fig. 3. Trial conducted at Hanshin Koshien Stadium.

advantage of the HEVC encoding technologies that NTT has developed. In a trial of 4K live viewing carried out at Hanshin Koshien Stadium [3], video of a professional baseball game captured by 4K cameras at the stadium was compressed by an H.265/HEVC encoder, and 4K video was transmitted in real time to a live-viewing venue through the Business Ether Wide business-oriented network [4] of NTT WEST (**Fig. 3**). In a trial carried out with public viewing at the All-Japan Phone-Answering Competition [5], we demonstrated 4K video delivery to multiple sites by best-effort IPv6 multicast delivery, using FLET'S Cast service for content providers who have the FLET'S Hikari Next service users. In a trial of live viewing at the 2014 Saga International Balloon Fiesta [6], we demonstrated 4K video delivery to multiple points from video captured by 4K cameras at the Saga Balloon Fiesta site, using the best-effort service FLET'S Hikari Next Hayabusa [7].

Through these trials, we were able to confirm that the HEVC encoding technology and commercial network were effective as a means of transmitting high-definition 4K video to remote sites. We plan to continue working on the R&D of ultra-high-presence media generation technologies intended for the transmission of ultra-high-definition video of 8K and

higher.

2.3 Spatial and environmental information synchronized delivery technology

To achieve natural communication services, it is necessary to concentrate video, sound, and other spatial and environmental information, synchronize it, and deliver it efficiently with little lag. For example, when presenting ultra-high-resolution video that has been extracted and concentrated from various spaces and environments and that is to be shared between a number of playback devices, or when carrying out switching control of environmental information such as lighting and temperature in addition to video and sound to match timing, it is necessary to present various different bits of information in an integrated form to the user rather than individually for each playback device, which makes it important to have a mechanism for synchronized delivery and presentation. To exchange spatial and environmental information bidirectionally in the future, information will need to be transferred with little lag, so techniques for reducing the processing time for information transfer and the buffering time at the receiving terminals will be important. Images of a system that delivers synchronized ultra-high-resolution video, sound, and
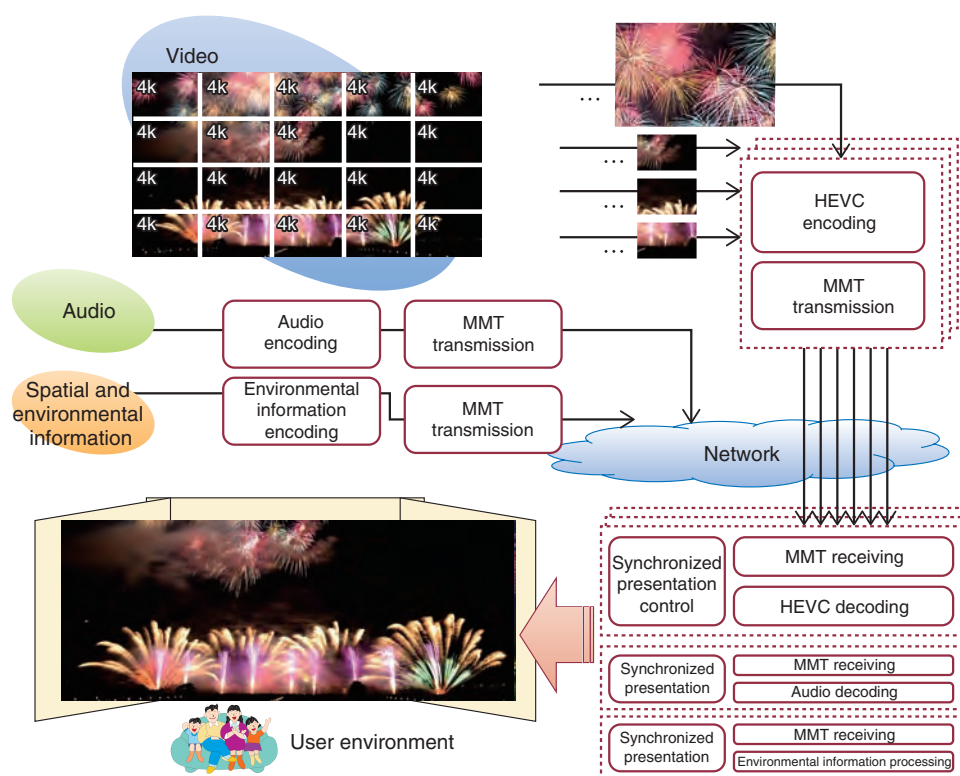
Fig. 4.　Images of system that reproduces space and environment in the users' environment.

environmental information to reproduce the space and environment in the user's environment are shown in **Fig. 4**.

NTT Service Evolution Laboratories is first implementing synchronized delivery and presentation that focus on high-definition movies and sound, and is also working on the R&D of a multi-content synchronized delivery and presentation system, as part of efforts to establish spatial and environmental information synchronized delivery technology. By adopting the MMT method that is specified as the transmission format in ISO/IEC[*1] 23008-1 and ARIB[*2] STD-B60, this multi-content synchronized delivery and presentation system delivers presentation timestamp information (UTC: Coordinated Universal Time) together with the video and sound content, in order to implement synchronized presentation at the receiving terminal.

It is also possible to freely integrate and present a number of different sources of video and sound content at the receiving terminal to suit the user's intentions, by transferring and receiving MPEG Composition Information (CI) that consists of XML (extensible markup language) information for presentation

control as defined by ISO/IEC 23008-11. The multi-content synchronized delivery and presentation system implements a content synchronized delivery function and a presentation control information delivery function in the configuration shown in **Fig. 5**, making it possible to deliver a number of high-definition video images and partial images generated by extracting them from high-definition video, and to present these images in synchronization to a number of terminals to suit the user's intentions. This means that users can freely enjoy the content they want to play, without any constraints on the size of the display or terminal, or of processing power.

2.4　Content license management technology

In ultra-high realistic services, because smartphones capable of shooting 4K video are becoming available in the market, it is assumed that utilization of high-quality user-generated content such as photos or videos will increase. To ensure that this content can

---

*1　ISO/IEC: International Organization for Standardization, International Electrotechnical Commission

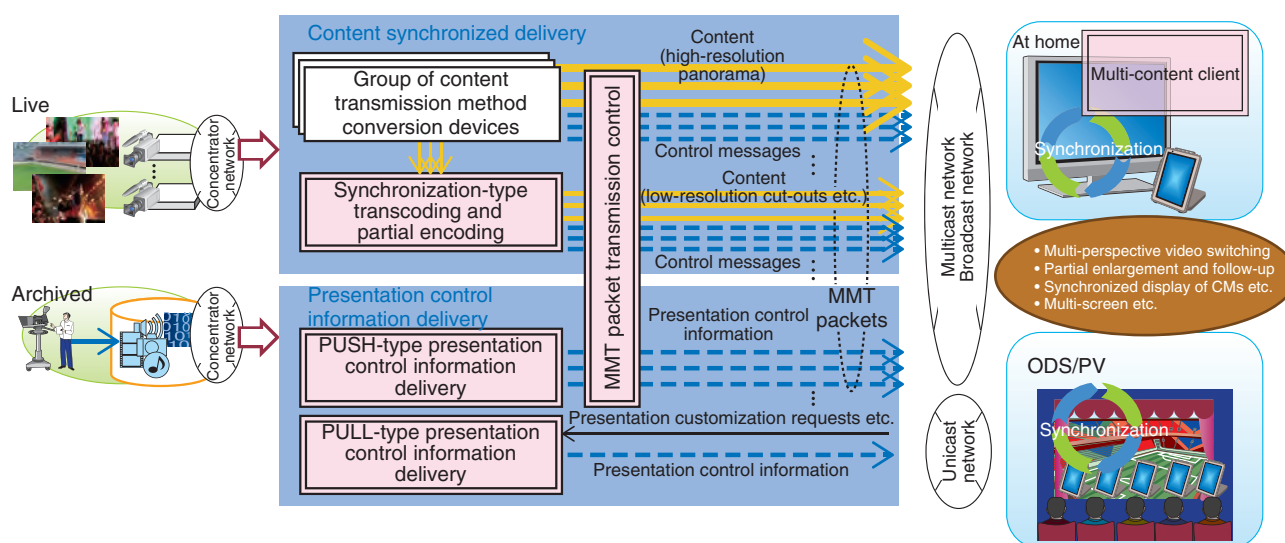*2　ARIB: Association of Radio Industries and Businesses

Fig. 5.   Multi-content synchronized delivery and presentation system.

be utilized securely, a new highly convenient license management technology is needed; therefore, we are currently researching this kind of technology to ensure that such content is never used in unintended ways, as high-quality digital content can easily be digitally copied.

It is generally assumed that centralized control is needed in conventional license management technology. However, there have been several problems with this style. For example, if the security of the centralized management is breached, the whole system will entirely collapse, and it is difficult to provide flexible management with licenses set by the content creators themselves. In consideration of these problems, NTT is developing completely new license management technology through collaborative research with Muroran Institute of Technology, and our proposed technology can ensure sufficient reliability without centralized control by gathering small amounts of computing resources from each of the participants in the network.

Based on a so-called blockchain[*3] that records the entire history related to license information and that is shared by all network participants, our technology achieves highly reliable license management by combining blockchain and cryptographic technologies such as electronic signatures created by a public-key encryption method. In this case, the data structure of a blockchain is literally like chains; that is, each piece of data is recorded in a form that links the data before and after it.

The license issuing mechanism using a blockchain is shown in **Fig. 6**.
(1) First of all, the content creator uses a private key that is held only by the content owner, and broadcasts license information that records who has been given the license, with an electronic signature attached, to the entire network. At this point, it is guaranteed by the electronic signature that no one other than the content owner can issue a license. Accessibility and durability are ensured by recording the broadcast license information in a blockchain.
(2) When license information is added to the blockchain—in other words, when a new block is added to the blockchain—the computationally expensive problem, which involves searching for additional information (nonce) to ensure that the leading n digits are zero within the hash value calculation relating to the added block, is calculated competitively by all the participants. To make it difficult to fake the blockchain, a computationally expensive problem is set deliberately.
(3) When someone eventually finds the appropriate nonce, a new block containing license information is added to the blockchain, and the license information is approved with its

---

*3 Blockchain: Used in applications such as digital (virtual) currency, a blockchain is a public ledger in which the exchanges of data are recorded in a chain-like form. The blockchain is shared with network users participating in the service.

(1) License information on the content creator is issued.

Content ID:
    125AEAD...
Licensed User ID:
    A700EB7...
License type:
    Playback License

Creator's e-signature

License information is broadcast to the entire network.

(2) All participants competitively try to solve computationally expensive problems.

$(N+1)^{th}$ new block

+ nonce

Hash value of Nth block: 5F48B9...

A new block that gathers all the license information is made at an approximately fixed time.

Additional information for ensuring that the leading n digits of the hash calculation are zero

participants

Finding appropriate nonce is computationally very expensive.

(3) If someone finds the appropriate nonce, a new block is added and the license is approved.

Blockchain

Reason of secure reliability

• Since everyone holds the same unique blockchain, the attacker must change everyone's blockchain to a spoofed one.
• To add spoofed information to the blockchain, the attacker must forge all the blocks added after the spoofed block.

Addition of $(N+1)^{th}$ block

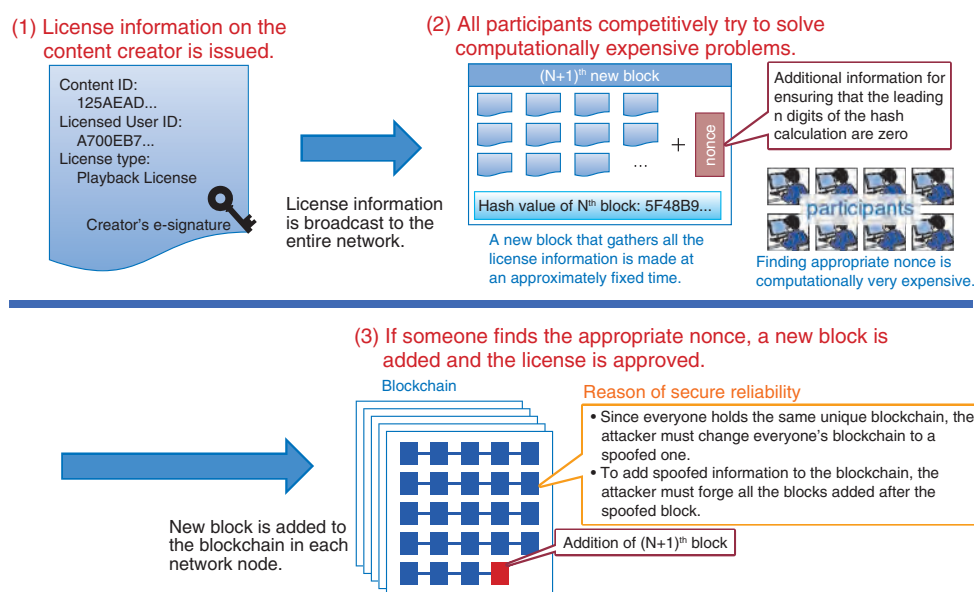New block is added to the blockchain in each network node.

Fig. 6. Blockchain mechanism.

reliability being ensured by the efforts of all the participants.

To be accepted as a standard, it is not only important for this new content license management technology to be secure and convenient, but efforts must also be made to achieve increased use of it through community activities. In the future, we intend to continue working on this approach with the cooperation of many stakeholders such as content creators and manufacturers.

3. Future expansion

At NTT, we are involved in R&D that will enable people around the world to physically experience ultra-high-presence technology that allows them to feel like they were spectators at sports and live events, through collaboration with various partners up to 2020. Kirari! enables spectators at remote sites to observe games as they happen, by transmitting and reproducing entire game spaces at many distant sports arenas and live venues, so they can share physical sensations such as those conveying speed and strength. We also plan to investigate applications to events that have been difficult for many people to appreciate in the past because of their remoteness, such as local festivals that serve as examples of our intangible cultural heritage.

References

[1] Feature Articles: "Video Technology for 4K/8K Services with Ultra-high Sense of Presence," NTT Technical Review, Vol. 12, No. 5, 2014.
     https://www.ntt-review.jp/archive/2014/201405.html
[2] Selected Papers: "Research Activities in Laboratories of New Fellows Part1," NTT Technical Review, Vol. 5, No. 12, 2007.
     https://www.ntt-review.jp/archive/2007/200712.html
[3] NTT WEST press release issued on June 27, 2014 (in Japanese).
     https://www.ntt-west.co.jp/news/1406/140627a.html
[4] NTT WEST, Overview of Business Ether Wide solution (in Japanese).
     https://www.ntt-west.co.jp/solution/solution/category/wide/
[5] NTT WEST press release issued on November 10, 2014 (in Japanese).
     https://www.ntt-west.co.jp/news/1411/141110a.html
[6] NTT WEST press release issued on October 24, 2014 (in Japanese).
     http://www.ntt-west.co.jp/saga/s_release/2014/20141024.pdf
[7] NTT WEST, FLET'S Hikari.
     http://flets-w.com/english/flets_hikari/next/

Akihito Akutsu

Executive Research Engineer, Project Director, Natural Communication Project, NTT Service Evolution Laboratories.

He received the M.E. in engineering from the Center for Environmental Remote Sensing, Chiba University. In 2001, he received the Ph.D. in natural science & technology from Kanazawa University. Since joining NTT in 1990, he has been engaged in R&D of video indexing technology based on image/video processing, and man-machine interface architecture design. From 2003–2006, he worked at NTT EAST, and engaged in managing a joint venture between NTT EAST and Japanese broadcasters. In 2008, he was appointed as Director of NTT, Cyber Solutions Laboratories. He was engaged in the R&D project of the broadband and broadcast services. In October 2013, he was appointed as Executive Producer of 4K/8K HEVC in NTT Media Intelligence Laboratories.

He received the Young Engineer Award and Best Paper Award from the Institute of Electronics, Information and Communication Engineers (IEICE) in 1993 and 2000. He is a member of IEICE.

Kota Hidaka

Senior Research Engineer, Planning Section (double as Natural Communication Project), NTT Service Evolution Laboratories.

He received the M.E. in applied physics from Kyushu University, Fukuoka, in 1998, and the Ph.D. in media and governance from Keio University, Kanagawa, in 2009. He joined NTT in 1998 and studied video skimming method based on emphasized speech. From 2009 to 2013, he worked at NTT EAST and engaged in commercial developments of Home Gate Way (HGW), Home-ICT based on HGW platform, and smart mater system. He has been studying the immersive telepresence technology called Kirari! since 2014.

Masayuki Inoue

Senior Research Engineer, Natural Communication Project, NTT Service Evolution Laboratories.

He received the B.E. and M.E. in electrical engineering from Tokyo University of Science in 1994 and 1996, respectively, and the Ph.D. in engineering from Tokyo Institute of Technology in 2008. He joined NTT in 1996 and has been engaged in research on 3D cyberspace and video communication systems. He is currently working on the business development of H.265/HEVC 4K LSI. His research interests also include user interaction and video system evaluation. Dr. Inoue is a member of IEICE and the Institute of Image Information and Television Engineers (ITE).

Naoki Ito

Senior Research Engineer, Natural Communication Project, NTT Service Evolution Laboratories.

He received the M.E. from Toyohashi University of Technology, Aichi, in 2001. He joined NTT in 2001 and engaged in research on character recognition. He moved to NTT EAST in 2004 and engaged in the development of security systems. He moved to NTT Cyber Space Laboratories (now, NTT Media Intelligence Laboratories) in 2008. Since 2014, he has been working on the development of high-definition visual communication service systems. He is a member of IEICE.

Tetsuya Yamaguchi

Senior Research Engineer, NTT Service Evolution Laboratories.

He received his B.E., M.E., and Ph.D. in information engineering from Osaka University in 1997, 1999, and 2008, respectively. Since joining NTT in 1999, he has been engaged in R&D of content distribution and navigation systems. His current interests are advanced media transport and ultra-realistic communications.

Shigeru Fujimura

Research Engineer, Natural Communication Project, NTT Service Evolution Laboratories.

He received the M.S. in information science and technology from the University of Tokyo in 2005 and joined NTT the same year. Since then, he has been engaged in research on web mining and web engineering, especially on effective implementation methods for web applications.

Atsushi Nakadaira

Senior Research Engineer, Natural Communication Project, NTT Service Evolution Laboratories.

He received the B.S. in physics from Kyoto University in 1992, the M.S. in physics from the University of Tokyo in 1994, and the Ph.D. in applied physics from the University of Tokyo in 2001. He joined NTT in 1994 and has been studying III-nitride semiconductors, display devices of polymer distributed liquid crystals, interactive systems of 3D displays, and visual communication systems. He is a member of the Japan Society of Applied Physics and ITE.