Remote Robot Control with Haptic Feedback Enabled by Low-latency Transport and Precision Bilateral Control Technologies

Junki Ichikawa, Takuro Yamaguchi, Yasuhiro Mochida, Hitoshi Masutani, Yoshihide Tonomura, and Hirokazu Takahashi

Abstract

The anticipation for advancements in remote robot control technology has been growing. NTT and Sony Group Corporation have collaborated to develop and demonstrate a new remote robot control technology that includes the sensation of touch, such as the feeling of pressure and weight, using low-latency transport and precision bilateral control technologies. This technology provides detailed feedback to operators located remotely, making it feel as though they are physically interacting with objects from afar. This article discusses the potential applications of this technology in various fields, including healthcare and manufacturing.

Keywords: remote robot control, RDMA communication, precision bilateral control

1. Expectations for remote collaborative robots

Demand for robots that cooperate with humans, known as collaborative robots, or cobots, is increasing. These robots are capable of complex cognition and decision-making tasks, as well as precise manipulations that can be complemented by human operators, leading to their anticipated utilization across a wide range of industries. In the construction industry, for example, construction robots can perform tasks that are challenging for human workers, while operators can focus on quality control and machine management, enhancing productivity and improving working conditions. In the medical field, surgical robots provide precise operational assistance, while doctors can make clinical decisions during procedures. Providing operators with the robot's sense of force (the sensation of pressure and weight when touching objects) also allows for even more delicate manipulations. There are various methods for enabling force feedback, but bilateral control, which coordinates the robot with the operator's movements to provide force feedback, is well-known. Thus, in this article, we refer to these robots as "bilateral control robots."

As the demand for collaborative robots continues to rise, remote robot control has also garnered increasing attention as a solution to mitigate the impact of the pandemic and labor shortages. For instance, doctors in Tokyo are able to carry out surgical procedures remotely on patients in rural areas (see **Fig. 1**). Although surgeries can be conducted despite the physical distances separating the doctor and patient, a high-quality network is essential to connect the locations.

Requirements for networks and applications are expected to differ significantly on the basis of the operations they support. The specific requirements
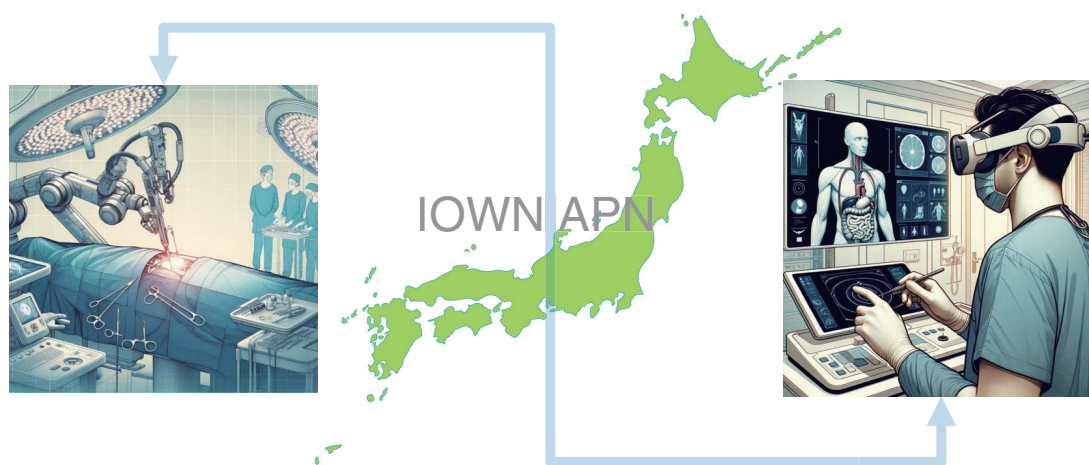
Fig. 1.   Remote surgery by teleoperated robot.

Table 1.   Network requirements for remote robots.

		Construction robots	Remote surgery robots	Bilateral control robots
Requirements	Position accuracy	Order of centimeters	Less than order of millimeters	Less than order of millimeters
	Latency	Order of several hundred milliseconds	Less than 100 milliseconds	Order of several dozen milliseconds
	Jitter	–	Order of several hundred microseconds	Order of 100 microseconds

for construction robots, remote surgery robots, and bilateral control robots are detailed in **Table 1**. The requirements for remote surgery robots, which necessitates precise control, are more stringent than those for construction robots. The requirements for bilateral control robots, which provide haptic feedback through synchronization, are even more demanding. This is due to the need for these robots to maintain synchronization within microseconds.

2.   Technical overview

To enable remote robot control, the integration of robotic technology and communication technology is essential. Therefore, NTT and Sony Group Corporation have combined their respective technologies to undertake a demonstration of remote robot control that incorporates a new tactile experience. This section provides an overview of NTT's low-latency transport technology and Sony's precision bilateral control technology.

2.1   Low-latency transport technology

NTT is advancing the development of the All-Photonics Network (APN) as part of its commitment to implementing the Innovative Optical and Wireless Network (IOWN)[*1], a next-generation communication technology. In APN IOWN1.0, the network can provide a stable, low-latency physical infrastructure by exclusively dedicating optical wavelengths from end-to-end, crucial for mission-critical services. However, it is essential to extend this low latency and stability to robot-side applications. To fully leverage the capabilities of the IOWN APN, our research and development (R&D) efforts are focused on developing low-latency transport technologies that ensure high-quality data delivery directly to the application layer.

2.1.1   Uncompressed video transmission technology [1]

Uncompressed video transmission technology is a

---

*1   IOWN: An innovative network based on photonics technology, openly developed through collaborative architecture formulation at the IOWN Global Forum.
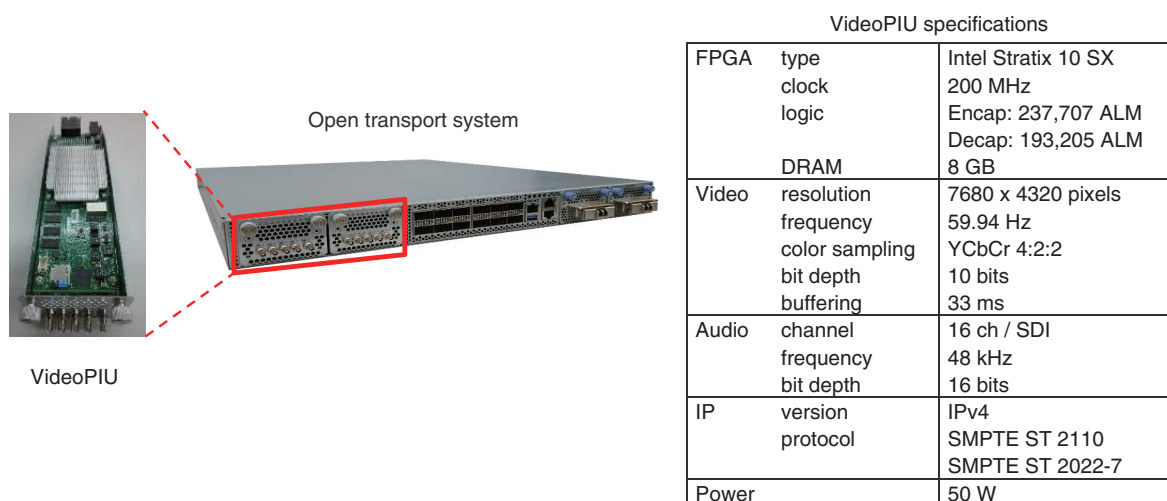
VideoPIU specifications		
FPGA	type	Intel Stratix 10 SX
	clock	200 MHz
	logic	Encap: 237,707 ALM
		Decap: 193,205 ALM
	DRAM	8 GB
Video	resolution	7680 x 4320 pixels
	frequency	59.94 Hz
	color sampling	YCbCr 4:2:2
	bit depth	10 bits
	buffering	33 ms
Audio	channel	16 ch / SDI
	frequency	48 kHz
	bit depth	16 bits
IP	version	IPv4
	protocol	SMPTE ST 2110
		SMPTE ST 2022-7
Power		50 W

Fig. 2.   Uncompressed video transmission plug-in unit (VideoPIU) and specifications.

method that enables low-latency feedback of visual information from remote robots to operators. With the recent increase in network capacities, although the required transmission bandwidth becomes larger compared with buffering and compressing video frames, the ability to transmit video with low latency has become more advantageous in real scenarios.

Traditional optical transmission devices were offered as integrated solutions by individual manufacturers, including both hardware and software functions, which made it difficult for service providers to flexibly add additional functionalities for video transmission. In contrast, new open optical transmission devices have started to emerge with a disaggregated configuration. These devices allow for the separation of various functions and the addition of flexible configurations and enhancements through standardized interface control. We have seized this trend and developed an uncompressed video transmission plug-in unit (VideoPIU) that can be used with disaggregated open optical transmission devices (**Fig. 2**).

VideoPIU converts serial digital interface (SDI) signals[*2] directly into SMPTE ST2110 streams [2], the standard for Internet protocol (IP)-based video transmission, allowing direct video-signal transmission through open optical transmission devices. VideoPIU is also implemented in hardware, reducing the delay from video input on the transmitting side to video output on the receiving side to within 1 ms. It can process 8K60p video[*3] per image and, by linking two images, can achieve 8K120p transmission. It also

supports seamless protection (SMPTE ST2022-7 [3]), ensuring continuous delivery without interruption even in the event of network disruptions or temporary failures by using two different optical paths.

2.1.2 Remote direct memory access acceleration technology

Remote direct memory access (RDMA) communication is widely used in datacenters and high-performance computing domains as a method for achieving high-bandwidth and low-latency data access between computers. RDMA allows for the direct transfer of data from memory to the network without central processing unit (CPU) intervention, making it a highly promising technology for fast, low-latency data transfers essential for time-sensitive and real-time services.

However, RDMA reliable connection, which facilitates reliable data transfers, is primarily designed for short-distance communications within datacenters. This design limitation has historically led to performance issues when adapted for medium to long-distance communications. The latency induced by longer distances typically results in increased waiting times at the requester due to slower acknowledgement (ACK) reception from the responder, subsequently lowering throughput.

To address this issue, we devised a method for

---

*2 SDI signal: A video transmission method using coaxial cables, widely used in broadcasting equipment.

*3 8K60P video: A video format that transmits ultra-high-resolution video at a resolution of 7680 x 4320 pixels, with a refresh rate of 60 frames per second, delivering clear and detailed visuals.

(a) Configuration of the pseudo-ACK packet method

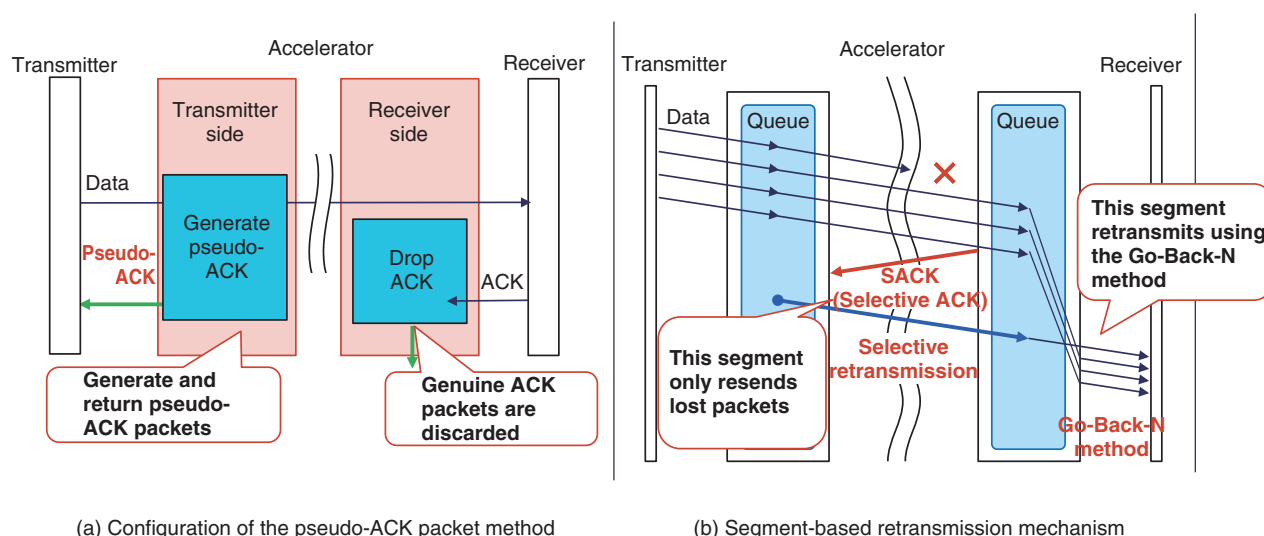(b) Segment-based retransmission mechanism

Fig. 3. Technical highlights of RDMA accelerator.

generating pseudo-ACKs near the requester, allowing for the early release of the requester's work queue element and enabling the acceptance of subsequent work requests. This method helps prevent performance degradation by reducing wait times and maintaining throughput (**Fig. 3(a)**).

This innovative method places accelerators near both the requester and responder, bridging long-distance networks. These accelerators create pseudo-ACKs using connection establishment data and request packets, effectively minimizing the waiting period for transmissions. However, this initially resulted in disabled packet-loss detection and retransmission mechanisms, posing significant reliability concerns. To resolve these, we implemented a suitable retransmission mechanism for the communication segments (**Fig. 3(b)**). This mechanism uses the Go-Back-N and selective retransmission strategies by segment, enabling quick recovery from packet loss while maintaining high-speed communication over long distances.

We conducted simulation evaluations of the above method. In the evaluations, we simulated propagation delays typical of long-distance networks and compared the throughput with traditional RDMA communications. The results indicate that under conditions simulating a 1000-km delay, with and without packet loss, there was an average improvement in throughput of 10 and 40 times, respectively, for a message size of 4 KiB (**Fig. 4**).

2.2 Precision bilateral control technology

We have sought collaboration with Sony, which is developing precision bilateral control technology. Sony possesses advanced bilateral control technology capable of flexibly adapting to changes in the external environment in response to human operations and controlling very minute forces with extreme accuracy. This section outlines the overview of the precision bilateral control technology.

Precision bilateral control technology is achieved by synchronizing the actions of two robots: a leader and follower. The follower robot moves in response to the leader robot's actions, and when the follower robot touches an object, the reactive force is transmitted back to the leader robot [4] (**Fig. 5(a)**).

2.2.1 High-sensitivity force sensing technology [5]

To feedback the sensation of touching human soft tissues to the operator, it is necessary to detect minute force changes as small as 1 gf (gram-force, equivalent to 0.0098 N) at the tool tip. However, incorporating force sensors into the system resulted in inertial forces being observed as noise, which made it challenging to detect slight changes in tip force. To address this issue, Sony has applied an optical strain sensor called a fiber Bragg grating (FBG) sensor, which features a diffraction grating etched into part of an optical fiber. Despite its ultra-fine fiber shape, the FBG sensor can measure the strain in the sensor part with high sensitivity. By mounting an ultra-sensitive FBG sensor at the tool tip and improving the algorithm that estimates the three-dimensional force
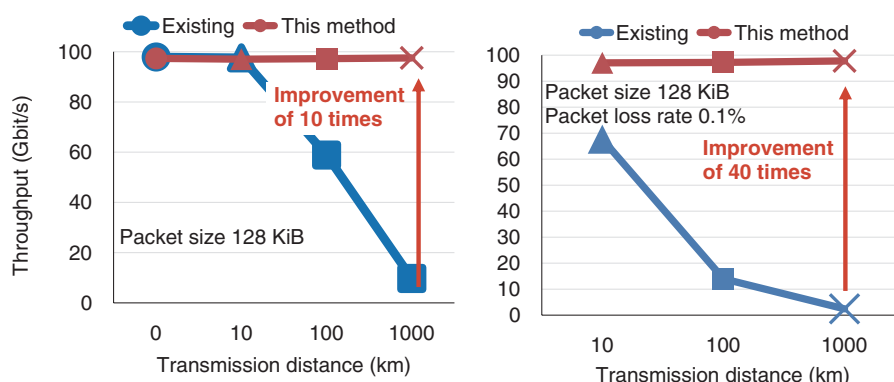
Fig. 4. Effect of accelerators in RDMA communications.

(a) Overview of precision bilateral control
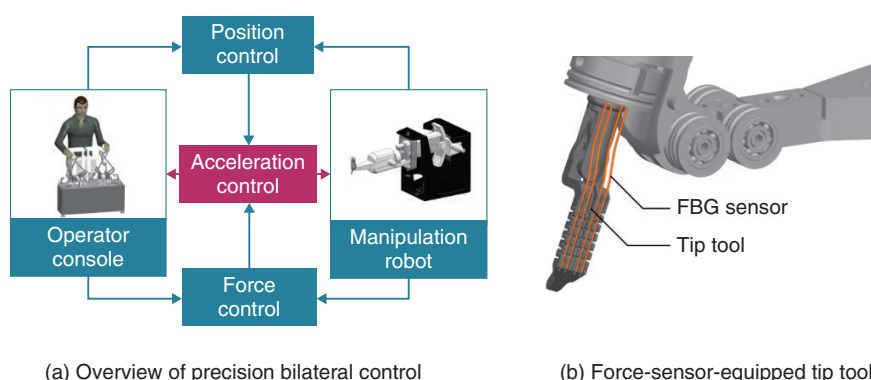
(b) Force-sensor-equipped tip tool

Fig. 5. Precision bilateral control technology.

applied to the tool tip from the sensor's strain amount, Sony has significantly reduced dynamic noise. This advancement allows for the precise detection of even minute force changes as small as 1 gf (**Fig. 5(b)**).

2.2.2 Precision position and force control technology

To achieve precise bilateral control, precise management of both position and force is essential. Traditional algorithms are affected by modeling errors due to variations in the robot's posture, leading to degraded position-tracking performance and vibrations when interacting with external environments. To address this issue, Sony has incorporated its proprietary technology, the generalized inverse dynamics (GID) library[*4], into their systems. This integration minimizes the impact of modeling errors caused by changes in robot posture, significantly enhancing responsiveness and operational stability, achieving position accuracy of less than 1 mm.

High-speed real-time processing is crucial for the signal processing involved in precise bilateral control. These signal processing tasks have traditionally been implemented on generic CPUs within a host personal computer (PC), which could not meet control system specifications due to communication delays and interrupt processing delays. To overcome these challenges, Sony has implemented part of the electrical processing on a field programmable gate array (FPGA), which is suitable for high-speed parallel processing. A proprietary protocol has also been developed for communication between the host PC and FPGA, resulting in a system that is 50 times faster than previous systems. By cascading multiple FPGAs using optical fibers, it is now possible to

---

*4 GID library: A model-based control algorithm that executes optimization calculations considering various constraints to accurately determine the necessary actuation in robots for specific tasks, along with a library that implements this algorithm.
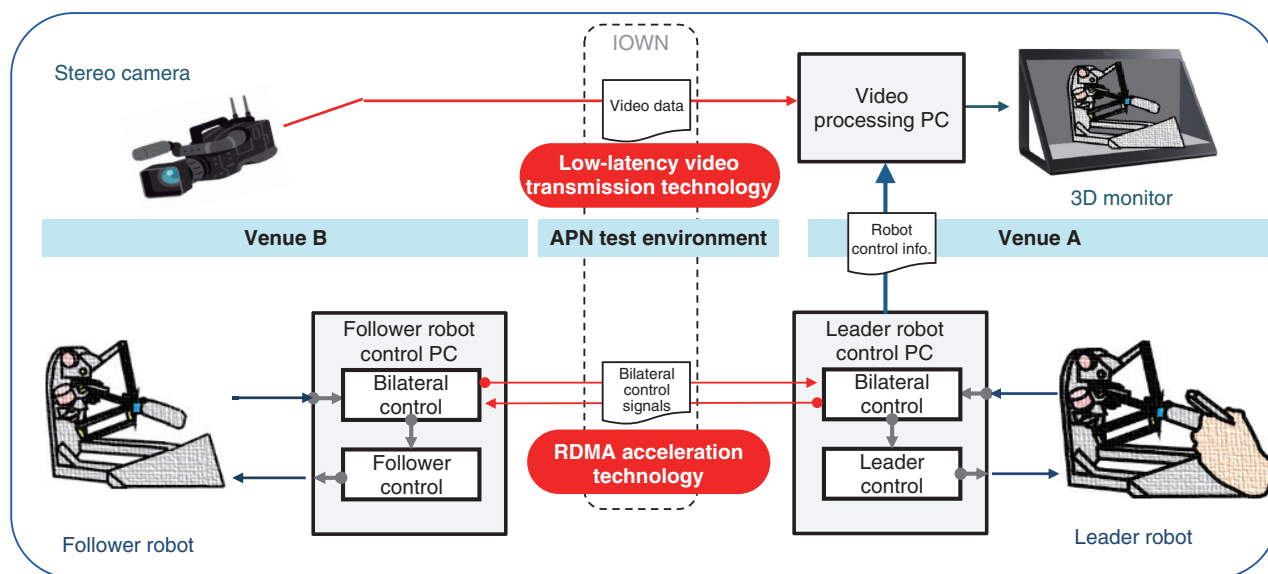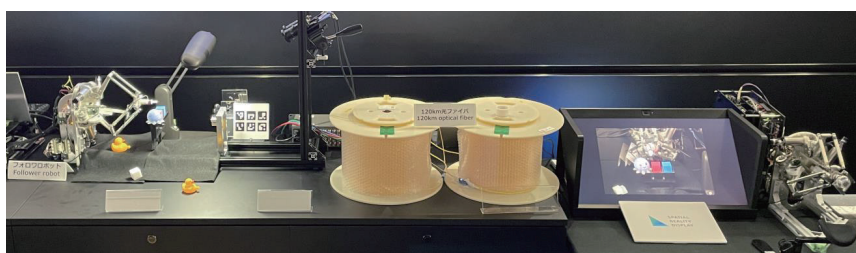
Fig. 6. Configuration and setup of the remote robot control demonstration experiment.

transmit signals with minimal delay in multi-degree-of-freedom systems such as the precision bilateral control system composed of leader and follower robots.

3. Results of the joint demonstration experiment

We conducted a demonstration experiment to verify whether remote robot control with haptic feedback is functional over distances exceeding 100 km. The experiment was carried out in an APN testing environment established at the NTT Musashino R&D Center. This testing environment included configurations such as APN-G (Gateway) and APN-I (Interchange) [6], as discussed in the IOWN Global Forum, representing a projection of future service environments. For this experiment, approximately 120 km of optical fiber was used to connect the setup. The configuration and setup of the experiment are illustrated in **Fig. 6**.

In the demonstration, the leader and follower robots were connected using Sony's precision bilateral

robots. The robot control PCs on both the leader and follower sides were equipped with RDMA transfer capabilities, facilitating the transmission and reception of bilateral control signals via an RDMA accelerator. Regarding the visual feedback, side-by-side stereo video was transmitted using VideoPIU and displayed on Sony's ELF-SR2 spatial reproduction display, which enables unaided stereoscopic viewing. It is important to note that the FBG sensor was not used in this experiment as additional tests beyond force feedback evaluation were required.

3.1 Communication evaluation results

Throughout the long-distance demonstration experiment spanning 120 km via the APN, video was transmitted with a remarkably low total latency of 1.6 ms, broken down as 1 ms of video-processing delay and 0.6 ms of APN-transmission delay for the 120 km. For the bilateral control signal communication via RDMA, the total latency was maintained at 1 ms, with an accelerator-processing delay of 0.4 ms, APN-transmission delay of 0.6 ms, and jitter of only 10 μs.

This setup allowed for exceptionally stable and low-jitter communication due to the elimination of CPU intervention, confirming the feasibility of achieving the high stability required by the bilateral control robots, as outlined in Table 1.

3.2 Operational evaluation of bilateral control robots

The operational evaluation revealed that the operator could not perceive any significant distance between themselves and the follower robot, enhancing the sense of immediacy and responsiveness. In this experiment, interactions with objects of varying hardness were performed remotely through the bilateral control robot, enabling the operator to discern even subtle differences in the object's surface texture. The ELF-SR2 spatial-reproduction display also provided natural stereoscopic vision enabled by head tracking, allowing the operator to carry out natural manipulations with a clear sense of spatial depth.

4. Future initiatives

In this initiative, a joint demonstration was conducted in an APN test environment, merging NTT's low-latency transport technology with Sony's precision bilateral control technology. This was aimed at achieving remote control operations that do not let users feel the distance, even when performed over distances exceeding 120 km. The results confirmed that it is possible to meet the stringent requirements demanded of bilateral control robots, allowing operators to feel as though they are directly touching objects in front of them.

Going forward, we plan to expand the scope of precise remote operations beyond geographical limitations by conducting demonstration experiments tailored to specific use cases. Through these efforts, we aim to contribute to the creation of a richer society.

References

[1] Y. Mochida, D. Shirai, and K. Takasugi, "Ultra-low-latency 8K-video-transmission System Utilizing Whitebox Transponder with Disaggregation Configuration," IEICE Transactions on Electronics, Vol. E106.C, No. 6, pp. 321–330, 2023. https://doi.org/10.1587/transele.2022LHP0003
[2] SMPTE ST 2110-10: "Professional Media Over Managed IP Networks: System Timing and Definitions," 2017.
[3] SMPTE ST 2022-7: "Seamless Protection Switching of RTP Datagrams," 2019.
[4] K. Ohnishi, S. Katsura, and T. Shimono, "Motion Control for Real-world Haptics," IEEE IEM, Vol. 4, No. 2, pp. 16–19, June 2010. https://doi.org/10.1109/MIE.2010.936761
[5] H. Suzuki, H. Masuda, K. Hongo, R. Horie, S. Yajima, Y. Itotani, M. Fujita, and K. Nagasaka, "Development and Testing of Force-sensing Forceps Using FBG for Bilateral Micro-operation System," IEEE Robot. Autom. Lett., Vol. 3, No. 4, pp. 4281–4288, Oct. 2018. https://doi.org/10.1109/LRA.2018.2864771
[6] IOWN Global Forum, "Open All-Photonic Network Functional Architecture," Ver. 2.0, Oct. 2023. https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-Open_APN_Functional_Architecture-2.0.pdf

Junki Ichikawa

Research Engineer, Advanced Networking Research Group, Frontier Communication Laboratory, NTT Network Innovation Laboratories.

He received a B.E. and M.E. from Chiba University in 2012 and 2014. He joined NTT Network Innovation Laboratories in 2014 and studied software-defined networking (SDN) and network functions virtualization (NFV) technologies. From 2017 to 2019, he was a member of the Lagopus project, which is an open source project for developing a software openflow switch and software multi-function router called "Lagopus." He deployed the Lagopus switch and router to ShowNet, which is a project for building a network within a venue held at Interop Tokyo in 2017, 2018, and 2019. He is currently studying RDMA acceleration technologies for long-distance optical networks such as the IOWN APN. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan.

Takuro Yamaguchi

Research Engineer, NTT Network Innovation Laboratories.

He received a B.E., M.E., and Ph.D. in electrical engineering from Keio University, Kanagawa, in 2014, 2016, and 2018. He joined NTT Network Innovation Laboratories in 2020 and engaged in research on video transmission over IP networks. His research interest is in high-resolution uncompressed video transmission over IP networks. He is a member of IEICE.

Yasuhiro Mochida

Senior Research Engineer, NTT Network Innovation Laboratories.

He received a B.E. and M.E. from the University of Tokyo in 2009 and 2011. Since he joined NTT in 2011, he has been engaged in research on video transmission over IP networks including transport protocol, presentation synchronization, and video conferencing. His current research interest is in low-latency video transmission over high-speed optical networks. He is a member of IEICE.

Hitoshi Masutani

Senior Research Engineer, Advanced Networking Research Group, Frontier Communication Laboratory, NTT Network Innovation Laboratories.

He received a B.E. in communication engineering, and M.E in electrical, electronic and information engineering from Osaka University in 1999 and 2001. After joining NTT Network Innovation Laboratories in 2001, he studied multicast networking and SIP (session initiation protocol)-based home networking. In 2005, He moved to the Visual Communication Division of NTTBizlink, where he was responsible for developing and introducing visual communication services, including an IP-based high-quality large-scale video-conferencing system, and real-time content-delivery system on Ipv6 multicast. He also worked on developing their service order management system and network management system for video-conferencing services. When he moved back to NTT Network Innovation Laboratories in 2012, he was engaged in R&D of programmable network nodes, including SDN and NFV, e.g., high-performance software openflow switch "Lagopus." When he moved to NTT Network Technology Laboratories in 2019, he was engaged in R&D of deterministic communication services' technologies, including time-sensitive networking. He has been developing high-performance and low-latency network technologies based on RDMA since 2023.

Yoshihide Tonomura

Senior Research Engineer, Supervisor, NTT Network Innovation Laboratories.

He received a B.E. and M.E. in electronics engineering from Nagaoka University of Technology, Niigata, and a Ph.D. from Tokyo Metropolitan University in 2002, 2004, and 2010. He joined NTT Network Innovation Laboratories in 2004. From 2012 to 2013, he was a visiting scientist at MIT Media Lab, USA. He received the 26th TELECOM System Technology Award from the Telecommunication Advancement Foundation in 2011. He was honored with the SUEMATSU-Yasuharu Award from IEICE in 2015. In 2018, he was awarded the Contribution Award for Information and Communication Technology from the Telecommunication Technology Committee. His research has consistently focused on signal processing and information theories and their applications for cutting-edge networks.

Hirokazu Takahashi

Senior Research Engineer, Supervisor, NTT Network Innovation Laboratories.

He received a B.E. and M.E. in electrical engineering from Nagaoka University of Technology, Niigata, in 2000 and 2002. He joined NTT Network Innovation Laboratories in 2002, where he studied multicast technologies and high-performance software packet processing technologies. He then moved to NTT Communications in 2010, where he developed network services such as IPv6 access and distributed denial-of-service protection. He returned to NTT Network Innovation Laboratories in 2013. His current research is high-performance and low-latency communication technologies.