REMOCOP: Remote Collaboration Platform for a Next Generation Remote Collaboration Support System

Daisuke Shirai, Yasuhiro Mochida, and Tatsuya Fujii

Abstract

There is increasing demand for work environments that enable collaborative work by people at remote sites. However, existing systems such as video conferencing systems and interactive whiteboards are designed for video and teleconferences, and they do not satisfy requirements for capabilities that include wide-ranging actual workflows that use specific devices and applications for work. We propose the Remote Collaboration Open Platform (REMOCOP), an application platform for a real-time remote collaboration environment. REMOCOP provides functions for transmitting and presenting work support information via audio/video communication.

Keywords: remote collaboration, application platform, computer supported collaborative work

1. Introduction

With business activities being carried out around the globe, the demand for remote collaboration technologies is increasing. As a result, video conferencing systems and web-based teleconferencing systems designed for information sharing and decision making are widely used. However, these systems are not optimized for remote collaboration of work processes such as product design and content creation because they have not been designed for remote collaboration systems other than basic video and teleconferencing. Remote product design requires sharing design applications such as computer-aided design (CAD) software. Remote content creation requires sharing media assets such as videos and pictures and carrying out real-time previews. These shared applications and media assets might also need to be operated or edited by multiple remote users at the same time. The existing widely used video and teleconferencing systems, which include interactive remote whiteboards, are single-function systems, so it is not possible to create

the remote collaboration environment described above or to implement additional customized functions.

A future collaborative and creative environment for users in remote locations will need a new framework that enables users to share tools and applications for work in real-time with audio/video communication, and that also enables transmitting and presenting work support information (**Fig. 1**).

2. Proposed application platform

To satisfy the previously mentioned requirements, NTT Network Innovation Laboratories is proposing the Remote Collaboration Open Platform (REMOCOP), which has been developed as a core framework of a future remote collaboration environment. The main elements and functions of the remote collaboration environment are expected to be:

- Audio/video transmission for real-time communication between users in remote locations
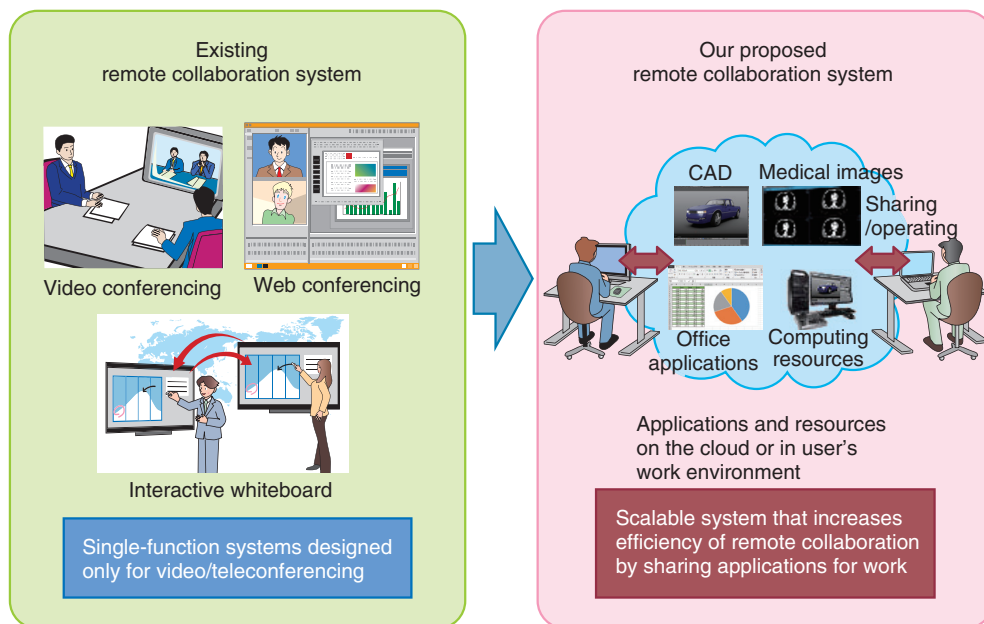- Sharing, editing, and operation of working objects and applications

Fig. 1.   Proposed remote collaboration system.

- Displaying work instructions and augmenting information to improve work efficiency
- Controlling input devices needed for work such as a mouse and keyboard, and transmitting their input data

However, the actual requirements for the remote collaboration environment depend on the type of work and workflow. How much audio/video quality is needed for sharing working objects or collaborative work? What kind of input devices and editing functions are needed? In addition, there is a need for network infrastructure control including a way to achieve a dynamic secure connection between various users and locations, as well as cloud resource provisioning. Thus, having to clarify every required function for every workflow or to develop a huge monolithic system that supports all those functions is not a realistic approach. The creation of a truly practical remote collaboration environment requires an application platform as illustrated in **Fig. 2**, which can replace, customize, or extend modularized components for remote collaboration so as to optimize the system for each workflow.
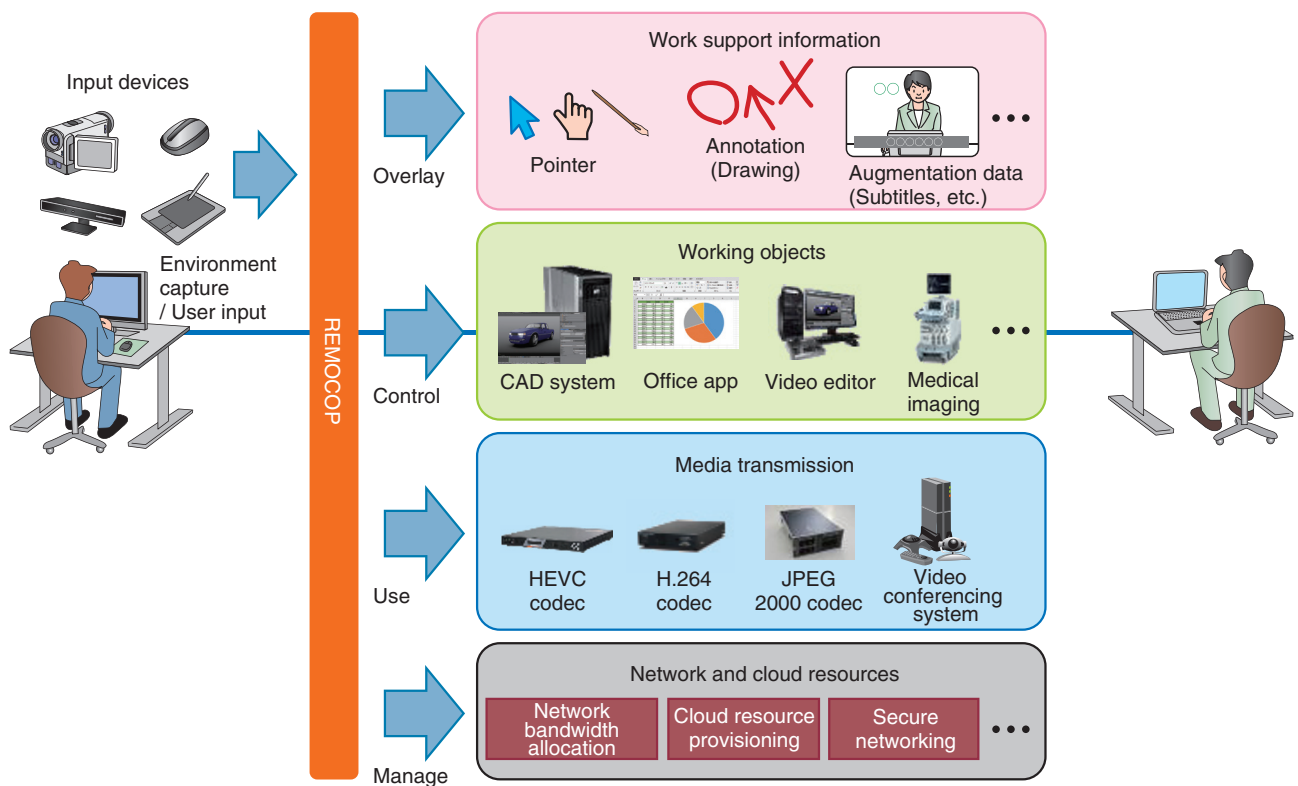
3.   Components of REMOCOP

A block diagram of REMOCOP is shown in **Fig. 3**. REMOCOP is composed of three core modules: an Augmentation Layer, Media-Time Manager, and Collaboration User Interface (UI). These modules are described in more detail below.

3.1   Augmentation Layer for overlaying work support information

In remote collaboration, a real-time audio and video transmission function is necessary for real-time communication and for sharing working objects between users in remote locations. However, the required media quality and latency depend on the workflow. Current teleconferencing systems and remote collaboration systems are able to use only the embedded codec on the system; therefore, it is difficult to satisfy such a wide range of requirements. The Augmentation Layer provides the function that overlays work support information and other additional information on real-time audio/video. Because it works separately from the audio/video codec, any type of audio/video codec can be used regardless of whether it is hardware-based or software-based, according to quality requirements. An example in which the 4K H.264 codec is used for real-time communication, and the HD (high definition) JPEG 2000[*1] codec is used for sharing CAD software as a

---

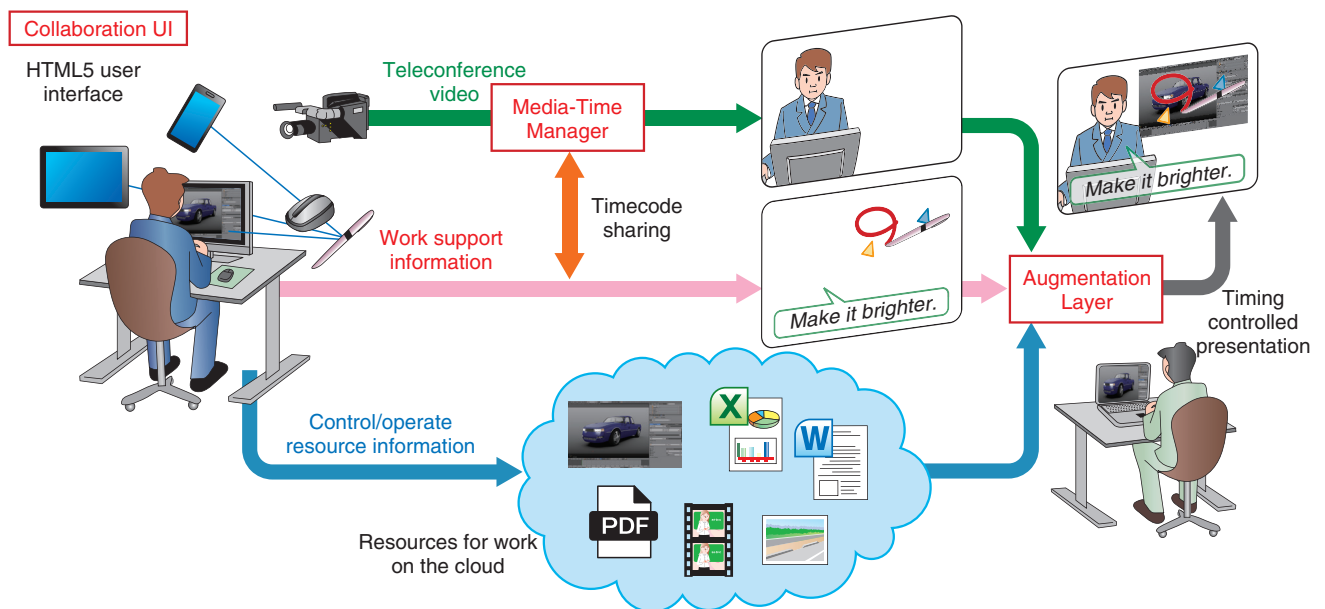Fig. 2.   Components of remote collaboration system.

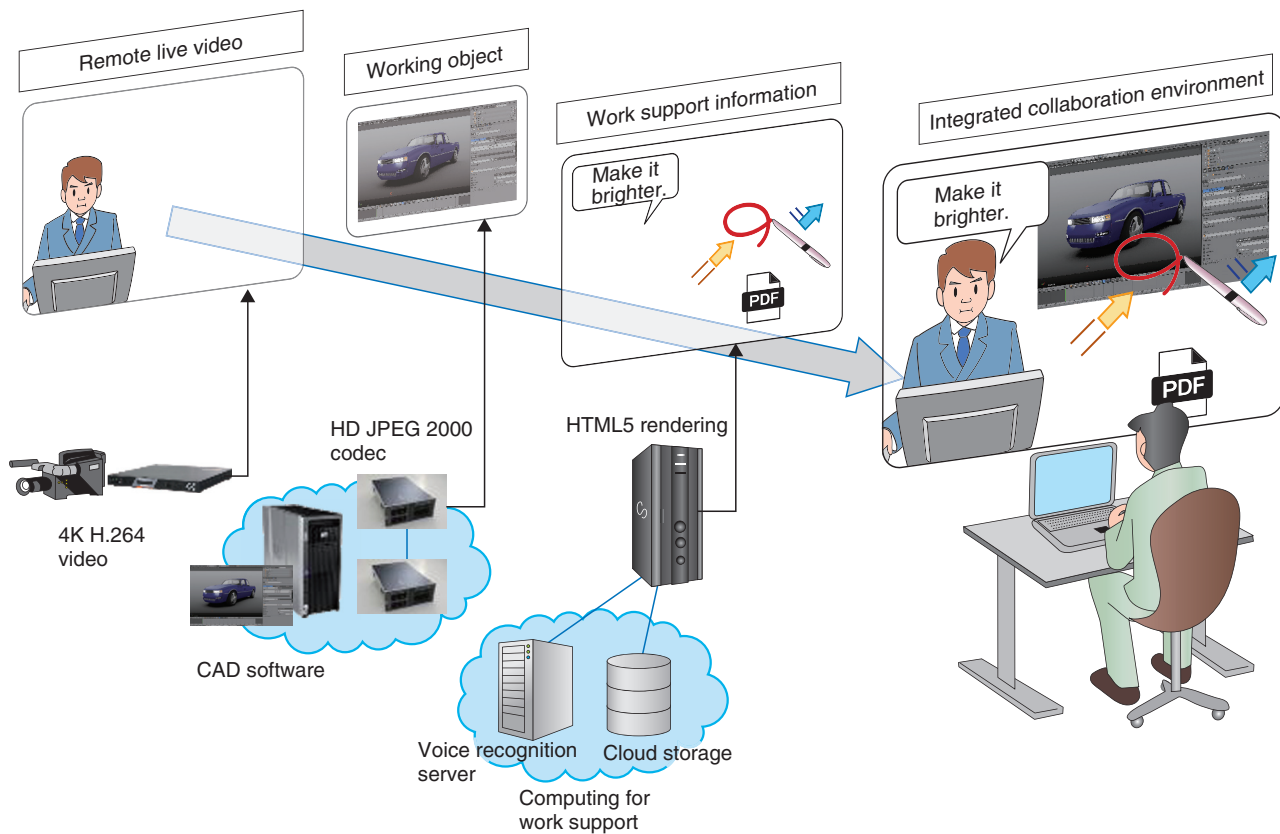

Fig. 3.   Block diagram of REMOCOP.

Fig. 4.   Overview and application example of Augmentation Layer.

working object is shown in **Fig. 4**. The Augmentation Layer overlays work support information such as pointers, annotations, instruction text by voice recognition, and PDF (portable document format) files on the cloud storage. Additionally, the Augmentation Layer displays work support information with HTML5[*2] rendering, which works in synchronization with audio/video for communication. Hence, the separate components shown in Fig. 2 can work as an integrated remote collaboration system.

3.2   Media-Time Manager for controlling temporal synchronization between multiple types of media and information

To enable the Augmentation Layer to display work support information in synchronization with audio/video, a framework is needed that manages and controls temporal information on all types of media as well as information for remote collaboration. The Media-Time Manager extracts timecodes from packets in the audio/video stream for real-time communication. Then it transmits them to the Collaboration UI

that manages user input interfaces; in doing so, it embeds timecodes in user input data packets—for example, the times of mouse movements—and sends the timecodes to remote locations. On the receiver side, the Augmentation Layer controls the presentation timing and overlays the data, receiving and referring to the embedded timecodes in the audio/video packets from the Media-Time Manager. It is also possible to present the data before or after the actual synchronized timing depending on the characteristics and requirements of the work support information.

3.3   Collaboration UI to build user interface optimized for each workflow

A function for sharing the pointer and a function for drawing and annotation are assumed to be common required functions for remote collaboration. Moreover, many other specific functions are required

---

*2   HTML5: The fifth revision of hypertext markup language (HTML), the standard programming language for creating web pages.

depending on the workflow such as those enabling operation with special input devices or those utilizing cloud storage. Demands for capabilities enabling users to join the collaboration environment from anywhere using devices such as tablet computers and smartphones are also increasing. The Collaboration UI makes it possible to implement these required functions easily with multi-device and multi-platform support by enabling us to build user interfaces using HTML5 and to run required functions on a web browser. The recent rapid advances in browser functions mean that various input devices including a keyboard, mouse, stylus, and gesture-based interfaces can be utilized by HTML5. Web APIs (application programming interfaces) also help us leverage external resources such as voice recognition and cloud storage. Therefore, the Collaboration UI makes the development of work support functions easy by utilizing various input devices and external resources with multi-platform support.

4. Application examples

Here, we discuss application examples of remote collaboration using REMOCOP.

4.1 Remote collaborative industrial design with CAD system

The first example involves remote collaborative industrial design with a CAD system running on virtualized GPU (graphics processing unit) resources, where the system is controlled by multiple remote users. In this workflow, by sharing the CAD software screen using a high quality video codec, remote users can discuss points to be fixed while drawing elements by hand. A system enabling two remote users to operate the CAD application running on the cloud and to communicate smoothly with multiple-channel 60p live video and multi-channel audio is shown in **Fig. 5**. In addition, pictures, movies, and documents needed for work as reference data can be overlaid. To instruct a remotely located person on design using hand drawing, the users only need to draw with a stylus pen and a tablet. To point out specific reference data, the users only need to use their finger to point at the object because the gesture interface senses hand movements. These intuitive uses of input devices in accordance with specific intentions dramatically accelerate the speed of collaboration work. Additionally, the HTML5 based implementation enables users to join the collaboration work even in a mobile environment simply by accessing the session URL (uniform

Fig. 5. Application example of remote collaborative industrial design.

resource locator), although there are limitations in media quality or spatial resolution of the work space.

4.2 Remote collaboration in the medical field
Another example is remote collaboration in the medical field, which requires sharing of medical images with remote users, searching for similar images from a database, and operating medical equipment remotely. In this case, the system has to support a display environment for sharing high resolution images. Functions that enable annotation on images and transmission of data between networked storage systems are also needed. REMOCOP enables us to build a very high resolution display environment using multiple computers and display devices in sync, and also to easily implement similar image searches with external resources using web APIs. Even in an operating room where use of physical interfaces such as a mouse and keyboard is prohibited, gesture-based interfaces enable users to control the system in order to browse or point at shared information.

5. Future prospects

As described in this article, REMOCOP allows us to create remote collaboration services that satisfy a wide variety of user needs. It is also expected that REMOCOP will be the leading platform for applying emerging technologies for real working environments such as new interface devices, machine-learning-based human thought process support, and cloud

functions. To achieve this, the business ecosystem consisting of users, application developers, device developers, and service providers should be created by making the platform open. Finally, we believe that the future remote collaboration environment using REMOCOP will provide higher work efficiency than a face-to-face environment will.

Daisuke Shirai
Senior Research Engineer, Media Processing Systems Research Group, NTT Network Innovation Laboratories.
He received his B.E. in electronic engineering, M.E. in computer science, and Ph.D. in media design from Keio University, Kanagawa, in 1999, 2001, and 2014. He pioneered the world's first 4K JPEG 2000 codec system, which enables low latency 4K60p video transmission on a gigabit network. He is a veteran in the design and implementation of parallel processing architecture, hardware codec boards, codec control software, and high-performance forward error correction (FEC) technology. He also developed FireFort, a high-performance FEC technology, which is part of the FEC standard in MMT (MPEG Media Transport). He has applied his expertise across multiple domains through his study of practical applications in digital audio and video broadcasting technology, image coding, information theory, networking, human-computer interaction, and software architecture. He is currently researching the design and implementation of Network Supported Collaborative Work (NSCW), which centers on the unification of high-quality robust media transmission technologies, multimodal interaction design, and emerging web-based technologies.

Yasuhiro Mochida
Researcher, Media Processing Systems Research Group, NTT Network Innovation Laboratories.
He received a B.E. and M.E. from the University of Tokyo in 2009 and 2011. He is currently researching remote collaboration elements, including system architecture, network protocol, interface design, and web technologies. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).

Tatsuya Fujii
Group Leader, Media Processing Systems Research Group, NTT Network Innovation Laboratories.
He received his B.S., M.S., and Ph.D. in electrical engineering from the University of Tokyo in 1986, 1988, and 1991. He joined NTT in 1991. Throughout much of his career, Dr. Fujii has been researching parallel image processing and super-high-definition image communication networks. In 1996, he was a visiting researcher at Washington University in St. Louis, USA, where he directed a digital cinema development project. He is currently studying advanced applications based on a real-time video communication system architecture, as well as reliable content distribution systems. He is a member of IEICE, the Institute of Image Information and Television Engineers of Japan, and the Institute of Electrical and Electronics Engineers.