Development of a Remote Field Operation Support System with Functions of Free Viewpoint Observation and Free Hand Gesture Instruction via Metaverse

Takashi Numata[1], Yuya Ogi[1], Keiichi Mitani[1], Kazuyuki Tajima[1], Yusuke Nakamura[1], Naohito Ikeda[1], Kenichi Shimada[1]

takashi.numata.rf@hitachi.com
[1]Research & Development Group, Hitachi, Ltd. 1-280 Higashi-Koigakubo, Kokubunji, Tokyo, Japan.
Keywords: Field operation support, free viewpoint observation, hand gesture instruction, augmented reality, virtual reality.

ABSTRACT

We developed a remote field operation support system with functions for free viewpoint observation and hand gesture instruction via metaverse. This novel system enabled users to switch between observation and instruction seamlessly without switching devices. The system improved remote support time efficiency by 14.6 % compared with a conventional system.

1 Introduction

The world has been facing global problems including shrinking workforces mainly due to low birthrate and aging population, the need for physical distancing, and resource shortages. These problems seriously affect human life and society. Remote support technologies including virtual reality (VR) and augmented reality (AR) technologies have been considered as some of the key technologies that will help address these problems. These technologies will enable a few skilled workers to remotely support the physical tasks of many onsite unskilled workers. These technologies are useful in various field operations such as machine maintenance, medical surgery, and work in harsh environments [1][2]. By extension, these technologies help solve the problems of skilled worker shortages.

To develop remote support systems for physical tasks, previous studies have focused on two key functions: observation and instruction from remote locations. The observation function has two types: observation by dependent view and that by independent view. Observation by dependent view is where a remote skilled worker observes the onsite work from the same viewpoint as the onsite worker [3][4]. Observation by independent view is where a remote worker observes the onsite work by independently structured viewpoints [5][6]. For the instruction function, there are typically three types of instruction cues: hand gestures [2][3], sketches [4], and pointers [5][6]. Remote support systems in previous studies have used different types of observation and/or instruction functions.

The remote support systems in previous studies, however, have focused on tasks in small task spaces. Therefore, large-scale changes of viewpoint during observation and of positions during instruction have not been taken into consideration. In field operations, onsite workers usually need large task spaces with large machines, such as in maintenance of manufacturing equipment and large vehicles. While performing the task, they have to move around the machines. In such cases, remote skilled workers have to switch between observation and instruction many times to understand the onsite situation by observing around the large task spaces in three dimensions. They also have to instruct the onsite unskilled workers on the correct tasks by giving correct hand gestures at the right position. Nevertheless, no system for such complicated tasks has been developed so far. The conventional observation and instruction systems have different user interfaces (use of remote controller and hand gestures). Therefore, it would take considerable time to switch between observation and instruction.

In this study, we developed a remote field operation support system that combined the functions for free viewpoint observation including along the vertical direction and for free hand gesture instruction, for the first time. As a noteworthy feature, we developed a seamless and hands-free user interface for combining observation and instruction for remote skilled workers, which enabled better time efficiency than that of the conventional system.

2 Methods and Materials

2.1 Remote Field Operation Support System

2.1.1 System overview

Our remote field operation support system was developed by combining two key functions: free viewpoint observation and free hand gesture instruction as shown in Fig. 1. The two functions were carried out by switching modes (observation mode and instruction mode) seamlessly. The functions and their combination are explained below.

2.1.2 Hands-free viewpoint observation function

The observation function was developed in three steps: 3D measurement, merging of multiple 3D data, and development of a VR application for remote

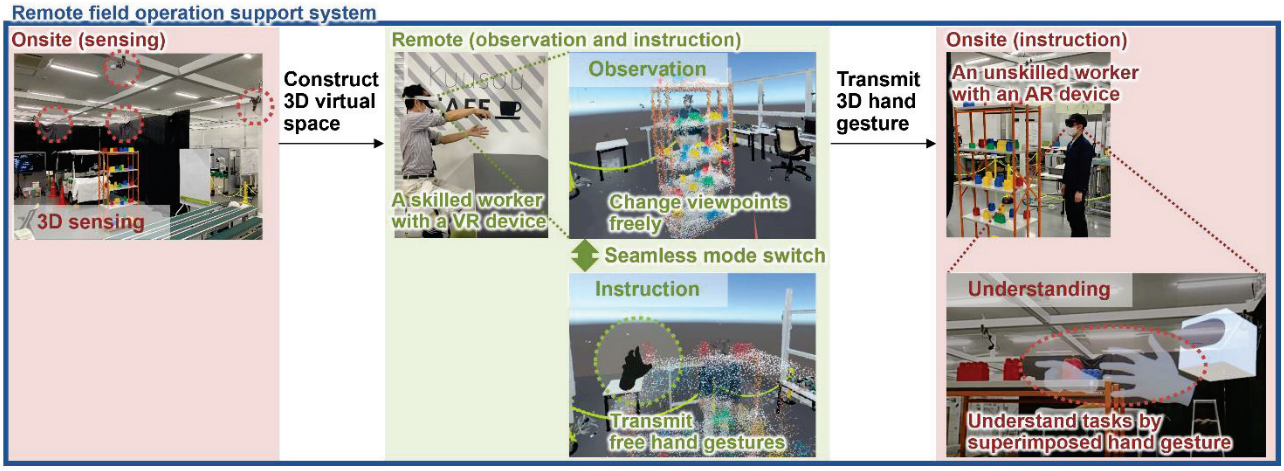

Fig. 1 Remote field operation support system developed in this study.

observation in VR space. First, multiple RGBD (Red, Green, Blue, and Depth) sensors were installed in a large task space, and 3D data were measured by each sensor. Four LiDAR Cameras were used as RGBD sensors. Second, multiple 3D data were merged, and a 3D space was reconstructed by using the merged 3D data. Finally, a virtual 3D space was constructed by using the 3D space data. The detail of novel technology for real-time 3D space reconstruction was introduced in another study [7]. Remote skilled workers were able to immerse into the virtual 3D space as a virtual avatar by using a VR device. This enabled them to observe the onsite situation and onsite unskilled worker's task at any viewpoint. During the observation mode, hand feature points were detected by using multiple cameras (Fig. 2(A)). As a novel method, remote skilled workers were able to change their viewpoints by hand gestures like making strokes in water

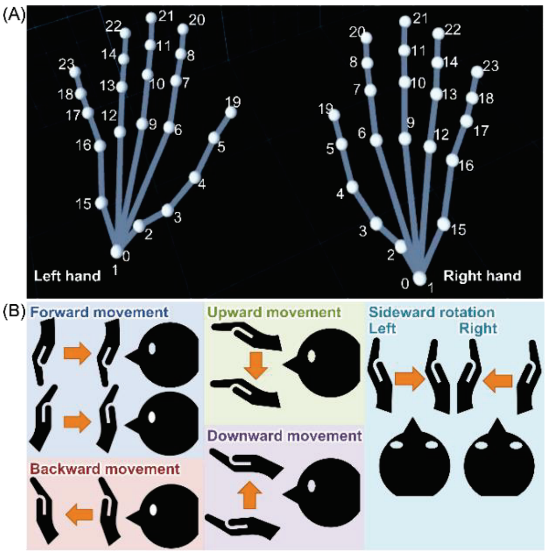

Fig. 2 Hand detection and gesture definition for viewpoint movements in the developed system. (A) Detection of hand feature points. (B) Gesture definitions for viewpoint movements.

(Fig. 2(B)). For example, the viewpoint of the remote skilled worker's avatar moves forward when the skilled remote worker moves his/her hand backward. Since most people are familiar with making strokes in water, even skilled workers who are not familiar with VR devices were able to change viewpoints naturally. A VR device with a software development kit was used as the VR device with cameras and hand feature detection software. Because 3D data were measured and the VR 3D space was updated in real-time, the remote skilled workers always observed the actual onsite situation.

2.1.3 Free hand gesture instruction function

The instruction function was also developed by detecting hand gestures. During the instruction mode, hand gestures of the remote skilled worker were continually measured. The hand gestures were then transmitted as AR contents to a head-mounted AR device attached to the onsite unskilled worker. The transmitted hand gestures were superimposed on the onsite environment by the AR device. Coordinate systems of the virtual space and onsite environment were matched by installing a marker onsite. The marker was then detected by the cameras of the AR device.

2.1.4 Seamless switch of observation and instruction

Detection of hand gesture for the mode switching was implemented in the developed system to carry out the switch between observation and instruction. The developed system enabled the remote skilled worker to switch modes by putting the tips of the thumb and forefinger of the left hand together. As a novel feature, both functions were implemented by detecting hand gestures, enabling the remote skilled worker to observe and give instructions seamlessly with free hands without having to switch devices. Therefore, the developed system saved time for switching between observation and instruction. Using the developed system, the remote skilled worker was able to observe the onsite unskilled worker via a virtual space (metaverse). In addition, the onsite unskilled worker was able to understand the appropriate physical task by hand gesture instruction via the superimposed AR hand gestures of the remote skilled worker's avatar. Thus, the remote and onsite workers were able to have smooth and effective communication via the metaverse in the developed system.

2.2 Experiment

2.2.1 Participants

We performed our experiment with 28 participants. The mean age ± standard deviation of the participants was 37.3 ± 9.1 years old. Data from the participants were obtained with their written informed consent.

2.2.2 Experimental environment and procedure

To evaluate the efficiency of the developed system, we constructed an original experimental environment and performed an experiment. A steel rack measuring 1805-mm x 915 mm x 450 mm with five shelves was set up in a task space as the experimental environment. Blocks of four colors were then placed on the first, fourth, and fifth shelves of the rack, as shown in Fig. 3(A).

Tasks combining observation and instruction were performed by the participants performing the role of remote skilled workers. They were asked to observe the onsite situation and give instructions on a physical task in accordance with the onsite situation from a remote location. The physical task was to place two blocks at correct positions and angles on the third shelf of the rack. Before performing the task, participants were asked to memorize two patterns (patterns A and B) of positions and angles for two blocks. They were then told that the correct positions and angles should follow pattern A when the total number of blocks in the rack was odd, and pattern B when it was even. Either 13 or 14 blocks were randomly placed in the rack. At the start of the task, participants counted the number of blocks in the rack to determine if the number was odd or even. They then gave instructions on the correct pattern of positions and angles of blocks by using hand gestures (Fig. 3(B)). As the features of the task, participants had to change viewpoints around the rack from top to bottom and front to back, and give instructions on the physical task by hand gestures on the shelf of the rack. The environment and the task, therefore, matched the features of tasks in large task spaces.

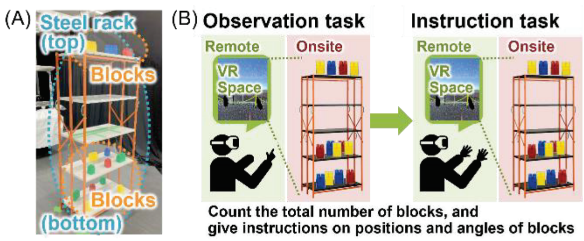

Fig. 3 (A) Experimental onsite environment and (B) explanation of experimental task. Participants performed the task from a remote location.

2.2.3 Comparison with conventional system

One of the advantageous features of the developed system is its seamless and hands-free user interface combining observation and instruction. We evaluated the time efficiency for remote support by comparing the system with a conventional user interface. Conventionally, viewpoints in virtual spaces are changed by using remote controllers of a VR device. Therefore, we also implemented the observation function by using controllers to compare the efficiency of remote support between the developed system and the conventional system. Using the developed system, participants made hand gestures for both observation and instruction to perform the task (Fig. 4(A)). In addition, using the conventional system, they used the controllers for observation, and then put the controllers on a desk when they made hand gestures for instruction (Fig. 4(B)). For the conventional system, we asked participants to be careful not to hit their hands on the desk when putting down the controllers and when making hand gestures while giving instructions. Participants were asked to practice performing observation and instruction using both systems before the experiment. They performed five trials for each system.

Time efficiency was evaluated using the time to complete the task (task time). The time was measured using a stopwatch from the beginning of the observation to the completion of the instruction. The shortest task

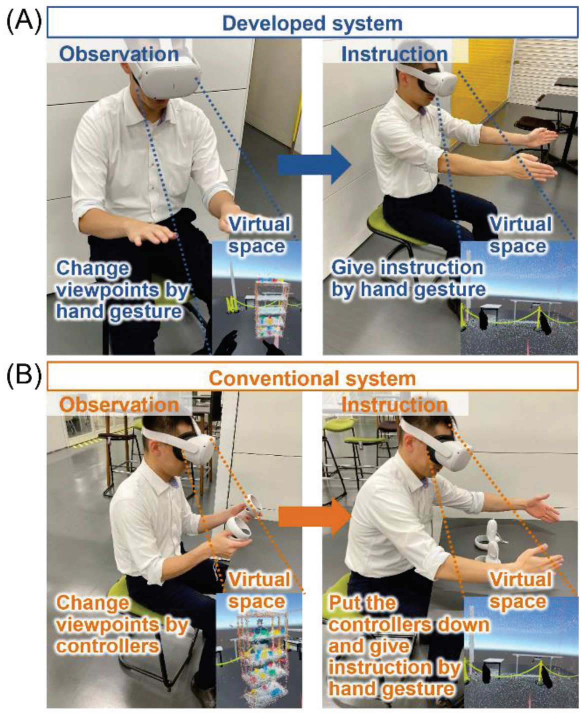

Fig. 4 Experimental conditions. (A) Observation and instruction using the developed system. (B) Observation and instruction using the conventional system.

time among the trials in each system by each participant was obtained. They were then statistically compared between the two systems by using the statistical paired t test. The level of significance was set to $p < 0.05$ according to common standads.

3 Results

The averaged time among participants for observation and instruction using the developed system was 12.33 ± 4.34 s, while that for the conventional system was 14.13 ± 3.88 s. The averaged time for the developed system was statistically significantly shorter than that for the conventional system ($p < 0.05$), as shown in Fig. 5.

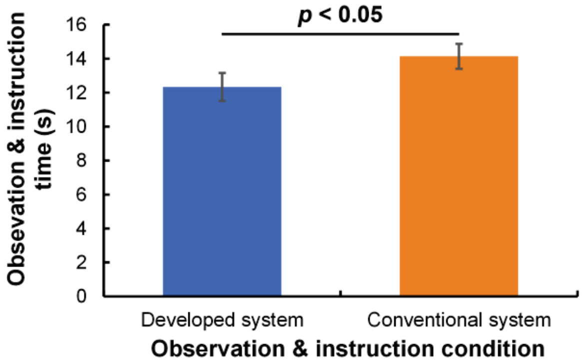

Fig. 5 Comparison of observation and instruction time between the developed system and the conventional system. Bars show average values ± standard error among participants.

4 Discussion

The developed system was more advantageous than the conventional system in terms of time efficiency for remote support. This advantage is attributed to the difference in user interface. In the developed system, participants were able to smoothly observe and give instructions by using only hand gestures. On the other hand, they had to put the controllers on the desk to switch from observation to instruction in the conventional system. In addition, the space for hand gesture instruction was limited in the conventional system since the desk had to be set near the participants. This space limitation reduced the usability of the conventional system. The developed system, therefore, was more advantageous in terms of time efficiency owing to its higher usability.

The time efficiency for remote support using the developed system was 14.6% higher than the conventional system. The experimental task was simple, and switching from observation to instruction was performed only once in the experiment. In actual tasks, however, remote support needs to be done through a trial and error process, and switching between observation and instruction needs to be done many times. Therefore, the difference in efficiency between the two systems would be magnified in actual conditions. The developed system, therefore, would make a significant contribution to remote support in large task spaces.

Going forward, we will use the developed system for remote support in actual field maintenance operations with large work spaces and large machines and demonstrate the efficiency of the developed system by comparing it with conventional remote support methods.

5 Conclusions

We developed a remote field operation support system with functions for free viewpoint observation and free hand gesture instruction via metaverse. This novel system enabled users to switch between observation and instruction seamlessly with free hands. From our experiment, the developed system improved remote support time efficiency by 14.6% compared with the conventional system. Thus, the developed system should enable remote skilled workers to understand the onsite situation even in large onsite task spaces. It should also enable them to instruct onsite unskilled workers on how to deal with problems in field operations by effectively giving instructions using correct hand gestures at the right position.

References

[1] R Palmarini, J.A. Erkoyuncu, and R. Roy, “An innovative process to select augmented reality (AR) technology for maintenance,” Procedia CIRP, Vol. 59, pp.23-28 (2017).

[2] E. Oyama, N. Shiroma, N. Watanabe, A.Agah, T.Omori, and N. Suzuki, “Behavior navigation system for harsh environments,” Advanced Robotics, Vol. 30, No. 3, pp.151-164 (2016).

[3] K. Robert, D. Zhu, W. Huang, and L. Alem, “MobileHelper: remote guiding using smart mobile devices, hand gestures and augmented reality,” Siggraph Asia 2013 Symposiumu on Mobile Graphics and Interactive Applications, Article. 38, pp. 1-5 (2013).

[4] S. Gauglitz, B. Nuernberger, M. Turk, T. Höllerer, “In Touch with the Remote World: Remote Collaboration with Augmented Reality Drawings and Virtual Navigation,” In Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology, pp. 197-205 (2014).

[5] T. Teo, L. Lawrence, G.A. Lee, M. Billinghurst, and M. Adcock, “Mixed reality remote collaboration combining 360 video and 3D reconstruction,” In Proceedings of the 2019 CHI conference on human factors in computing systems, article. 201 (2019).

[6] M.L. Chénéchal, T. Duval, V. Gouranton, J. Royan, and B. Arnaldi, “Help! I need a remote guide in my mixed reality collaborative environment,” frontiers in Robotics and AI, Vol. 6, Article. 106, (2019).

[7] K. Mitani, K. Tajima, and Y. Nakamura, “Real-time 3D space reconstruction for OT metaverse as interactive virtual site,” 29th International Display Workshops (IDW’22), PRJ1-2 (2022).