Construction Tele-Robotics System with AR Presentation

K Ootsubo*, T Kawamura and H Yamada

1Department of Human and Information Systems, Gifu University, 1-1 Yanagido, Gifu, 501-1193, Japan

*E-mail: otsubo@gifu-u.ac.jp

Abstract. Tele-Robotics system using bilateral control is an effective tool for task in disaster scenes, and also in extreme environments. The conventional systems are equipped with a few color video cameras captures view of the task field, and their video images are sent to the operator via some network. Usually, the images are captured only from some fixed angles. So the operator cannot obtain intuitively 3D-sense of the task field. In our previous study, we proposed a construction tele-robotics system based on VR presentation. The operator intuits the geometrical states of the robot presented by CG, but the information of the surrounding environment is not included like a video image. So we thought that the task efficiency could be improved by appending the CG image to the video image. In this study, we developed a new presentation system based on augmented reality (AR). In this system, the CG image, which represents 3D geometric information for the task, is overlaid on the video image. In this study, we confirmed the effectiveness of the system experimentally. Additionally, we verified its usefulness to reduction of the communication delay associated with a tele-robotics system.

1. Introduction
Recently, there is a need for system to perform task in areas where it is difficult for human such as disaster scenes and extreme environments; space, seabed, deep underground, nuclear plant and so on. Especially in Japan, a new system which is more effective compared with the conventional systems and can perform advanced task is strongly required after the Great East Japan Earthquake.

Tele-Robotics system using bilateral control is an effective tool for such task. In Japan, some conventional systems were used at real disaster scenes such as volcano eruption and earthquake. The development of the technique was started in order to avoid a secondary disaster in the recovery from the eruption of Mount Unzen in 1991. In this case, construction robot was tele-operated by feedforward control, and the operator performed tasks with a video image obtained from a stereovision camera. However, the image contains little spatial information of the task field, and the physical load was not negligible.

In the past, we investigated a tele-robotics system for a construction machine. The system consists of a servo-controlled construction robot, two joysticks for operating the robot from a remote place, and a 3-degree-of-freedom motion base. In order to improve the controllability of the system, we have examined the master and slave control method between joysticks and robot arms [1,2], the presentation method for the motion base [3,4], and the visual presentation of the task field to the operator [5]. Because the visual presentation is the information most essential to the operator, in this study we continued to focus on the presentation method for a remote operation field.

In our previous paper [6], we proposed a presentation method that used a mixture of stereographic video and the CG image of the robot, and clarified that the task efficiency was improved. In this
system, the task object and the robot are presented by CG. These are simulated in real time in accordance with the operation, and the operator can obtain a view from a desired angle. So the operator intuits the 3D geometries of the object and the robot. However, most of the environmental elements were not represented. Since the visual information about the task field was poor compared with a video image, we think this drawback has a negative effect on the task efficiency. Additionally, the task viewing the multiple images putted a big load on the operator.

So we thought that the task efficiency and the load could be improved by appending the CG image to the normal color video image. In this study, we developed a new presentation system based on augmented reality (AR). In this system, the CG image, which represents 3D geometric information assists to perform a task, is overlaid on the video image. The operator can intuits spatial information about the task while obtaining the environmental information in the task field.

In this study, we confirmed the effectiveness of the system experimentally. Additionally, we verified it can deter decline of the task efficiency resulting from the communication delay associated with a tele-robotics system. In the experiments, the operators were asked to repeat a designed task, and the results demonstrated the effectiveness by using the task efficiency and their load.

2. Construction tele-robotics system

Figure 1 shows an illustration diagram of our tele-operation system of construction robotics. This system is composed of joy stick (JS) for operation, camera for capturing the task field, construction robot, PC for controlling the robot (PC1) and PC for processing data from input devices (PC2). The control system is based on master-slave method, and displacement of the JS (master’s side) corresponds to that of the piston (slave’s side) which is fed back to master’s system. An operator performs task looking the monitor connected to PC2.

![Diagram](image)

**Figure 1.** Construction Tele-Robotics System.

3. AR based presentation system

In this system, AR is achieved by overlaying the 3D-CG image, which supports the operation, on the video image.

The image contains 2-types of objects drawn as line segments, and overlaid on the video image by using the marker method. One represents the claw, and another represents a virtual gauge which
indicates vertical distance to the other objects. In this method, the marker is a pattern image board with a square outline, and the system identifies it by image processing of the captured video image. Then the system calculates a rigid body transformation from the coordinate system of the marker to that of the video camera, and the CG objects are drawn onto the corresponding real objects by the transformation. In our system, the programming module of AR is based on the “ARToolKit” [7] which is a license-free library written in C for AR development.

In this study, we use two types of markers (Figure 2). One is a “floor marker” which is used to recognize the task field, and another is a “claw marker” which is used to recognize the claw’s position and posture. As shown in figure 3, line segments are drawn from markers (a, b, c) to floor or claw’s ends on the video image. This system allows the operator to intuit the opening degree of the claw and the distance between the robot and the other objects by looking at the image of the task field. In this study, communication delay is simulated by stopping the video image, and the CG image is updating at the time. Since the operator can continue the operation by relying on the image, we expected that improve the task efficiency as a result.

![AR markers](image1)

**Figure 2. AR markers**

![AR presentation](image2)

**Figure 3. AR presentation**

**4. Experimentations and the results**
In this study, we performed experiments described later in order to verify the usefulness of our system. In the experiments, the subjects were asked to iterate a work set composed of designed tasks.

We performed the quantitative evaluation by using the task efficiency and the number of failures. Here, the task efficiency is calculated as a ratio of total tasking time to the criterion time.
Also the qualitative evaluation is performed by using NASA-TLX and the questionnaire of free-description type as the indexes. The former is used to evaluate the task load, and the latter is used to evaluate the usability of our system. In the evaluation based on the former, the task load is defined by 6-items (table 1). In the first procedure, a subject decides the weights based on his/her individual importance by using pairwise comparison between the items. And after the experiments, decides the scores of the items according to degree of the loads. If a score is high, it means the load of the item is high for the operator. Finally, a numerical value is calculated by averaging the weighted scores. It indicates the comprehensive task load, and is called Weighted Work Load (WWL).

| Name                | Abbr. | Description                                                                 |
|---------------------|-------|-----------------------------------------------------------------------------|
| Mental Demand       | MD    | How mentally demanding the task?                                            |
| Physical Demand     | PD    | How physically demanding the task?                                          |
| Temporal Demand     | TD    | How hurried or rushed was the pace of the task?                             |
| Performance         | OP    | How successful were you in accomplishing what you are asked to do?          |
| Effort              | EF    | How hard did you have to work to accomplish your level of performance?      |
| Frustration         | FR    | How insecure, discouraged, irritated, stressed, annoyed were you?           |

### 4.1. Experiment description

In the experiments, the operators were asked to grasp the center region of a block (figure 4), and carry it to a designated position (figure 5). A set of the tasks required transfers of blocks (1) from B1 to B2, (2) from B2 to B3, (3) from B3 to B4 and (4) from B4 to B1. The operators did 3-sets of tasks continuously, and we monitored the task time and the number of failures. Failure means the cases that the target block is not grasped at the center region, and the number is the total number in all the task time. The region is set near the center of gravity of the block. If the other region is grasped, the block falls easier during the transfer.

![Figure 4. Blocks](image1.png)

![Figure 5. Transfer position in the task field](image2.png)

In the experiments, we tried to evaluate the usefulness of our presentation system (AR) by comparing with the conventional presentation system (Conv.) in different state of communication delay. The delay is given randomly as 1-3 seconds in each 3-10 seconds. The experiments were performed in the cases shown in table 2.
4.2. Results
Here, the evaluation results based on quantitative index and qualitative index are shown in the following subsections.

4.2.1. Quantitative results
The task efficiencies given by 10-operators are shown in figure 6. Tasking times given by the operators in a condition are averaged, and the task efficiency is calculated as the ratio to that of the case I.

As shown in the figure, the task efficiency of the case III is lower than that of the case I, and also the same trend is seen between the case IV and the case II. Additionally, the difference between the latter cases is a little lower than that of the former cases. We think this is because that the operation was stopped more or less while the delay was occurred, and the time was shortened by using the CG image updated in real time.

The value of the case II is higher than that of the case I, and also the same trend is seen between the case IV and the case III. We think this is because that the recognition accuracy of the relative distance between the robot and the target object is improved by using AR. Additionally, the effect would be independent of the communication delay.

Next, the numbers of failure given by 5-operators are shown in figure 7.
As shown in the figure, the number of failure of the case II is less than half that of the case I. We think this is because that the recognition accuracy of the relative position of the claw and the center region is improved by using AR.

4.2.2. Qualitative results
The results based on the NASA-TLX are shown in figure 8. In the figure, (a) shows a result given by an operator experienced the operation using our system, and (b) shows that given by a non-experienced operator.

As shown in the figure, WWL of the case III and IV are higher than that of the case I and II. Especially, the differences of MD, EF and FR are larger than the other indexes, and these trends are
seen in both of the results (a) and (b). We think this is because that randomly stopping of the video image increased time to recognize the present situation in the task field. As a result, it would have increased difficulty of the task.

WWL of the case II and IV are lower than that of the case I and III, and the trend is seen in both of the results (a) and (b). We think this is because that AR allowed the operator to intuit the relative geometry between the robot and the target object easily.

In the result (a), WWL of the case IV is lower than that of the case I, but the relation is reversed in the result (b). We think this means that the reduction effect of the task load can be improved with the experience.

At last, we describe the result of the questionnaire. The contents are shown in table 3, and the averaged result given by 10-operators is shown in figure 9. Each item was scored by 5-point scale.

When the item A is observed, you can see the score falls by communication delay but is improved by using AR. Also, the same trend is seen in the item B and C. From this result, we think that ease of understanding the relative position of the claw and the other objects is directly linked to ease of the grasping of the block. This result means our consideration about the quantitative evaluation is confirmed.

Additionally, you can see the scores of the item D and E have the same trend. This means that the stopping of the video image increases the mental load, and AR can be effective answer for the problem.

| ID | Item |
|----|------|
| A  | Ease of the operation to grasp the block |
| B  | Ease of understanding the relative position of the claw and the block |
| C  | Ease of understanding the relative position of the claw, the floor and the transfer positions |
| D  | Safety of the task |
| E  | Comfort of the task |

Figure 8. NASA-TLX
5. Conclusions
In our previous study, we proposed a VR based presentation system using multiple images. However, the simulated environment was limited, and the task load to look at the multiple images was not negligible. So we thought that the task efficiency and the load could be improved by implementation of AR technique achieves overlay the CG image of the task on the video image.

In this study, we developed an AR based presentation system for tele-operation of construction robot. From the experimental results, we confirmed that the task efficiency and the load were improved by using the system. Also, the usefulness against the communication delay problem was confirmed.

In future works, we will try to further increase the task efficiency and the quality by implementing the following improvements to the system.

- Construction of accurate CG model and achievement of markerless AR by using a multiple-distance sensor with high resolution
- Development of new visual interface for supporting complex tasks

Acknowledgments
The authors would like to express their deep appreciation to Mr. Masashi HAMAGUCHI and Mr. Hidenori SATOH of the Graduate School at Gifu University for their extensive cooperation in performing the experiments.

6. References
[1] Yamada H, Muto T, Xia Y and Iwase T. 2000 Tele-operated Construction Robot Using Virtual Reality (CG Presentation of Virtual Robot for Increasing Working Efficiency): Proc. of ISMM2000 pp 652-658.
[2] Yamada H, Muto T and Ohashi G. 1999 Development of a Telerobotics System for Construction Robot Using Virtual Reality: Proc. of European Control Conference (ECC’99) F1000-6.
[3] Zhao D, Yamada H, Muto T, Huang H and Kato H. 2001 A New Method of Presenting Realistic Motions for a 3-Degree-of-Freedom Motion Base Applied in Virtual Reality System: Proc. of the 5th International Conference on Fluid Power Transmission and Control (ICFP’ 2001) pp254-258.
[4] Zhao D, Xia Y, Yamada H and Muto T. 2002 *Presentation of realistic motion to the operator in operating a tele-operated construction robot: Journal of Robotics and Mechatronics* 14-2 pp98-104.

[5] Yamada H, Muto T, Xia Y and Iwase T. 2000 Tele-operated Construction Robot Using Virtual Reality (CG Presentation of Virtual Robot for Increasing Working Efficiency). Proc. of VSMM2000: 652-658.

[6] Yamada H and Muto T. 2007 *Construction Tele-Robotic System with Virtual Reality(CG Presentation of Virtual Robot and Task Object using Stereo Vision System): International Journal of Control and Intelligent Systems* 35-3 pp195-201.

[7] Kato H, Billinghurst M. 1999 *Marker Tracking and HMD Calibration for a video-based Augmented Reality Conferencing System: Proc. of the 2nd International Workshop on Augmented Reality (IWAR 99)* pp85-94.