2014 IEEE International Conference on Robotics & Automation (ICRA)
Hong Kong Convention and Exhibition Center
May 31 - June 7, 2014. Hong Kong, China

An Autonomous Multi-Camera Control System Using Situation-Based Role Assignment for Tele-Operated Work Machines

Mitsuhiro Kamezaki, *Member, IEEE*, Junjie Yang,
Hiroyasu Iwata, *Member, IEEE*, and Shigeki Sugano, *Fellow, IEEE*

***Abstract* – A method to autonomously control multiple environmental cameras, which are currently fixed, for providing more adaptive visual information suited to the work situation for advanced unmanned construction is proposed. Situations in which the yaw, pitch, and zoom of cameras should be controlled were analyzed and imaging objects including the machine, manipulator, and end-point and imaging modes including tracking, zoom, posture, and trajectory modes were defined. To control each camera simply and effectively, four practical camera roles combined with the imaging objects and modes were defined as the overview-machine, enlarge-end-point, posture-manipulator, and trajectory-manipulator. A role assignment system was then developed to assign the four camera roles to four out of six cameras suitable for the work situation, e.g., reaching, grasping, transport, and releasing, on the basis of the assignment priority rules, in the real time. Debris removal tasks were performed by using a VR simulator to compare fixed camera, manual control, and autonomous systems. Results showed that the autonomous system was the best of the three at decreasing the number of grasping misses and error contacts and increasing the subjective usability while improving the time efficiency.**

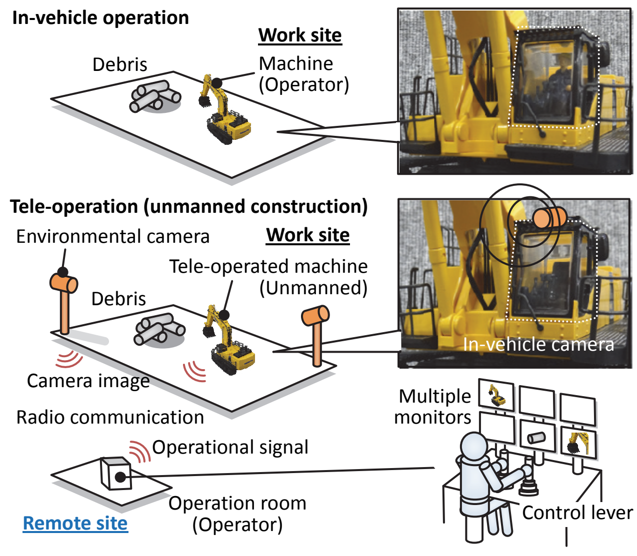

Fig. 1 Unmanned construction using tele-operation technologies

I. Introduction

DISASTER rescue and recovery operations are expected by the public to be performed safely and effectively in an expeditious way. Such operations are generally conducted using construction machinery, which has the advantage of being able to produce massive force. Construction machinery is maneuvered by an operator who is in a cockpit installed inside the machine (hereinafter called in-vehicle operation), as shown in the upper part of Fig. 1. However, in disaster response work, in-vehicle operation is often subjected to two serious problems: one, the endangerment of the lives of the operators, owing to secondary disasters such as collapsing debris, machines toppling over, and the falling and breakage of transported objects, and two, the limitations of the environment, owing to unusual or dangerous circumstances such as underwater condition or the presence of carbon monoxide or radiation, where human cannot live biologically.

Tele-operation technologies have been introduced into such works for addressing the above problems [1]–[3]. Note that they are called unmanned construction but do not mean autonomous control. In unmanned construction, an operator is in a safe operation room located at a remote place from the disaster site, as shown in the lower part of Fig. 1. Various data is exchanged between the work and remote sites by using radio communication techniques. Operational signals from the control levers are transmitted to tele-operated machines on the work site, such as hydraulic shovels, bull-dozers, and dump trucks. Sensor signals such as positional data from satellite positioning systems and images from in-vehicle and environmental cameras are transmitted to the operation room and they are projected onto multiple two-dimensional (2-D) monitors, as shown in the lower right of Fig. 1.

The safety of operators is physically ensured and the environmental limitations are partially resolved by using unmanned construction systems. However, these systems are simply an integration of existing tele-operation, monitoring, and communication technologies, and so they face various problems. The most critical problem is the decrease of time efficiency. In excavation work using a tele-operation system, the time efficiency is lowered by about 40% compared to in-vehicle operation [4]. The improvement of the time efficiency is a high-priority task because disaster response work should be performed as quickly as possible. Another problem

M. Kamezaki is with the Research Institute for Science and Engineering (RISE), Waseda University, 17 Kikui-cho, Shinjuku-ku, Tokyo 162-0044, Japan (corresponding author to provide phone and fax: +81-3-3203-4457; e-mail: kamezaki@ieee.org).

J. Yang is the Major in Modern Mechanical Engineering, Graduate School of Creative Science and Engineering, Waseda University, 17 Kikui-cho, Shinjuku-ku, Tokyo 162-0044, Japan.

H. Iwata is with the Department of Modern Mechanical Engineering, School of Creative Science and Engineering, Waseda University, Green Computing System Research and Development Center, 27 Waseda-machi, Shinjuku-ku, Tokyo 162-0042, Japan (e-mail: jubi@waseda.jp).

S. Sugano is with the Department of Modern Mechanical Engineering, School of Creative Science and Engineering, Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan (e-mail: sugano@waseda.jp).

URL: http://www.sugano.mech.waseda.ac.jp/

is that unmanned construction is recently required for use in more emergent and dangerous cases where it has not previously been applied such as nuclear-plant accidents [5], [6].

II. Requirements for Advanced Tele-Operation

The problems with the current unmanned construction systems are clarified and ways of addressing these problems to enable advanced tele-operation are then proposed.

A. Factors degrading work performance

1) Three dominant factors: To clarify exactly what causes degraded time efficiency, the differences between in-vehicle operation and tele-operation were analyzed and three dominant factors owing to primordially different operational conditions were found: incomplete visual information, radio communication delay of operational and camera image data, and lack of tactile and body sensory information, as shown in the upper part of Fig. 2. These factors seriously degrade the work performance, and various technologies have been developed in response, including an immersive interface using augmented reality [7], radio communication array vehicles [8], and force feedback systems [9]. Unfortunately, actual work performance has thus far shown little improvement.

2) Importance of visual information: As is well known, the visual sense is superior to other senses in terms of planning, judgment, and action in human behavior [10]. Thus, the first issue to be addressed of the above three factors for improving work performance should be the incomplete visual information. Moreover, more effective, adaptive, and intuitive visual information is inevitably required for advanced unmanned construction, which requires safer and more precise operations, as stated in section I. This study thus focuses on the enhancement of visual information.

B. Conventional studies on visual information

1) Current imaging system: As stated in section I, visual information in unmanned construction is provided by using multiple cameras. The in-vehicle camera is installed in the cockpit of a machine and provides images closer to the operator's view in in-vehicle operation. The environmental camera is mounted on a camera-car and provides panoramic images of the work environment. The position (x, y, and z), posture (roll, yaw, and pitch), and optical zoom of each camera are currently adjusted by hand before starting a task, and these parameters are not adjusted while this task is performed, as shown in the upper part of Fig. 2.

2) Enhancement using adaptive imaging: Questionnaires have indicated that the subjective usability of current imaging systems varies greatly depending on the operator [11]. This is because the required properties of visual information (e.g., viewpoint and imaging range) can change largely according to the work content, environmental conditions, and operational skill levels. If adaptive images suitable for work situations (desired images for operators) are provided, the operator can precisely control a machine with less stress and fatigue while avoiding dangerous and excessively careful operations, as shown in the lower part of Fig. 2. However, most conventional studies have focused on modifying the configuration of the in-vehicle camera for reproducing an operator's eyesight in the in-vehicle operation and reducing the number of environmental cameras and monitors on the operation room. The environmental cameras must be adjusted according the work situations, but there are no studies on adaptive imaging systems in current unmanned construction systems.

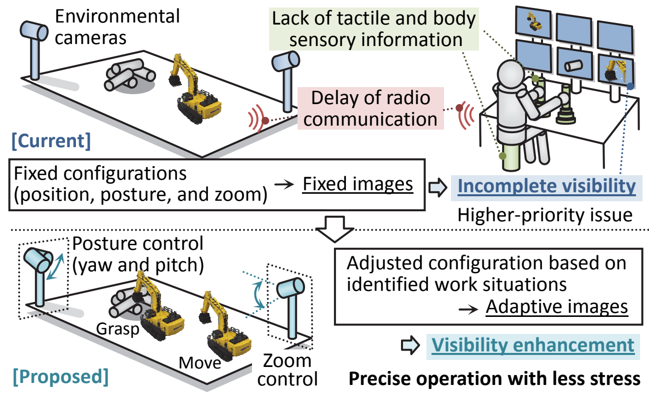

Fig. 2 Current problems and proposed visibility enhancement

C. Autonomous environmental-camera control

The purpose of this study is to develop a method to autonomously control multi-environmental cameras based on the work situation for enhancing the operator's visibility. The proposed system controls posture (yaw and pitch) and zoom, which are the three basic parameters for camera control. Installable positions greatly depend on the types of constraints in the actual environment, so controlling the position will be discussed in a future work. The following developments and experiments were thus conducted.

1) Camera role definition: To simply and effectively control cameras, camera roles R were defined by using imaging objects, where a camera captures within a specified imaging range, and imaging modes, which is a way to control the imaging object. Four camera roles were defined (section III).

2) Role assignment system: To adequately assign m camera roles to m out of n cameras ($m < n$), a role assignment system including checking play possibilities and selecting m cameras suitable for each R in the situation on the basis of the priority rule in real time was developed (section IV).

3) Comparison experiment using VR simulator: To evaluate the usefulness of the proposed autonomous imaging system, a debris transport experiment was performed and the proposed system was compared with a fixed camera system and manually controllable camera system (section V).

III. Camera Role

On the basis of analysis of the camera images that needs to be provided for operators in specific situations, the camera roles R_x needed for autonomous control were defined.

A. Analysis of required images depending on situation

Environmental cameras for compensating the in-vehicle camera image are not currently controlled and provide basic panoramic views, as stated in section II. Thus, adaptive im-

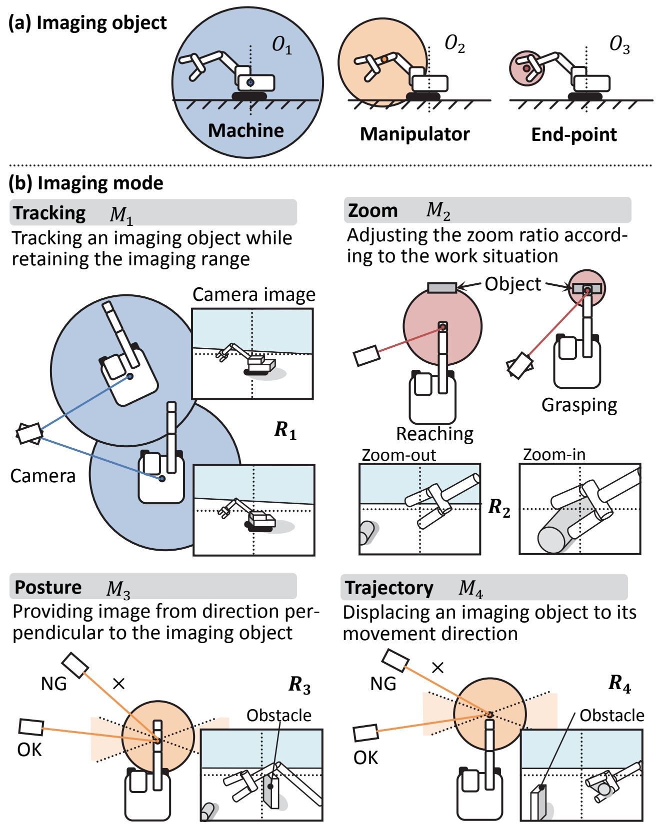

Fig. 3 Camera roles: Imaging objects and modes

ages needed for improving work performance were analyzed, on the basis of sequences common to disaster response work, such as base movement, reaching, grasping, transport, and releasing. In base movement, safe and efficient positioning is needed, so machine images involving the surroundings are required. In reaching, efficient operations without making contact with obstacles are needed, so images to indicate the manipulator's posture in relation to neighboring obstacles are required. In grasping, the position and posture of the grapple must be controlled more precisely, so enlarged end-point images to show the distance between the grapple and an object are required. This kind of image is also required in releasing. In transport, careful operation is required because the dropping of transported objects leads to extremely dangerous situations, so end-point images involving spaces displaced in the manipulator's moving direction are required.

B. *Imaging object and imaging mode*

The above analysis indicates that several camera roles must be defined. These roles consist of an imaging object, where a camera captures within a specified imaging range, and the imaging mode, which is a way to control the imaging object.

1) Imaging object: Three imaging objects O_x were derived: the machine including the surrounding environments (O_1), the manipulator (O_2), and the end-point (O_3), as shown in Fig. 3 (a). Target points capturing at the center of the imaging range are defined as the center of the pivot joint (O_1), the center of the arm link (O_2), and the center of the grapple (O_3).

2) Imaging mode: Four imaging modes M_x were derived: tracking (M_1), zoom (M_2), posture (M_3), and trajectory (M_4), as shown in Fig. 3 (b). M_1 keeps following an imaging object by controlling the yaw and pitch joints. M_2 adjusts the zoom ratio depending on the situation. M_3 provides images from the direction perpendicular to the imaging object for easiness of depth recognition. M_4 displaces an imaging object in its movement direction for effective planning and operation.

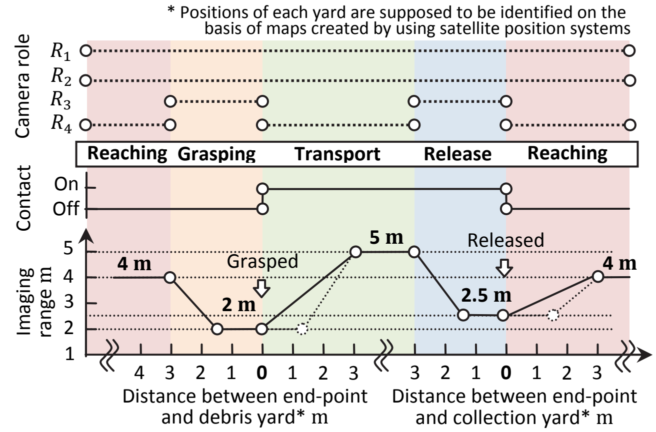

Fig. 4 Situation identification and adaptive image range (zoom) control

C. *Definition of camera roles*

Camera roles R_x can be defined by combining arbitrary imaging objects O_x and modes M_x. In this study, four camera roles were defined considering effectiveness, independence, and practicality, as shown in Fig. 3. The camera roles used for each situation, which is simply identified by using distances between the end-point and each yard and the on-off state of object contacts, are shown in Fig. 4.

1) Overview-machine R_1 (O_1, M_1): R_1 tracks the machine with a panoramic image. This role provides the most basic image, so it is assigned to two cameras whose relational angle is perpendicular to reduce blind areas (R_{1A} and R_{1B}). The angle of view was fixed to the maximum in this study.

2) Enlarged-end-point R_2 (O_3, M_1, M_2): R_2 adjusts the zoom ratio depending on the identified work situation while tracking the end-point, as shown in Fig. 4. The imaging range (zoom ratio) is linearly interpolated between each situation. These parameters were determined by pre-experiments considering sufficient visibility and comfortable operation.

3) Posture-manipulator R_3 (O_2, M_1, M_3): R_3 keeps the camera direction at a right angle to the manipulator while tracking the manipulator. The right angle is actually a strict condition, so the tolerance angle is defined as $\pm 20°$.

4) Trajectory-manipulator R_4 (O_2, M_1, M_3, M_4): R_4 tracks the target point in O_2 displaced in the movement direction of the end-point while satisfying R_3. The offset size is proportionally changed by the end-point's velocity. The offset at the maximum velocity was set to 2 m in this study. This role stands in for R_3 during the reaching and transport.

IV. Role Assignment System

To assign the four defined roles R_x to four cameras, a role assignment system to check role possibility, select suitable

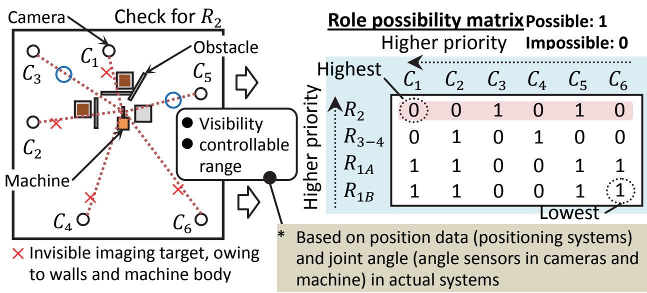

Fig. 5 Check of role possibility and role possibility matrix

cameras, and assign roles in real time was developed.

A. Role Possibility Matrix

1) Check of role possibility: The system first confirms if the cameras can play each camera role in real time. Play is possible if there are no obstacles in the camera direction toward the imaging object and no limitations of camera parameters and is confirmed by using the position among cameras, imaging objects, and obstacles, as shown in the left part of Fig. 5. This is then used to create the $m \times n$ matrix of the role possibility (m: the number of cameras and n: the number of roles), as shown in the right part of Fig. 5. In this study, four monitors were prepared for environmental cameras referring to current systems. There should be more environmental cameras than monitors because the system then has more options when selecting the most suitable cameras for the situations. In this study, m and n are thus set to six (C_1–C_6) and four (R_{1A}, R_{1B}, R_2, and R_{3-4}, as stated in section III C).

2) Definition of priority: Conditions for R_2 and R_{3-4} are stricter than R_1, and moreover, the grasping is highly difficult operation, so R_2 must have the highest priority and R_1 should have lower priority. To facilitate the role assignment, the index of row is arranged in order of R_2, R_{3-4}, R_{1A}, and R_{1B}, as shown in the right part of Fig. 5. Priority for camera is set in the numerical number in this study. The uppermost row and leftmost column have thus the highest assignment priority.

B. Role assignment system

An assignment system independently assigns camera roles R to cameras C on the basis of the defined priority.

1) Basic assignment rule: The possible cameras for R_2 are first checked in order of the camera number C_1–C_6, and R_2 is tentatively assigned to the first-appeared camera (C_3), as shown in Fig. 6 (a). In the same way, R_{3-4}, R_{1A}, and R_{1B} are tentatively assigned to cameras (C_2, C_1, and C_1). Then, overlapping, meaning that multiple roles are assigned to one camera, is checked. The tentative roles assigned to C_1 are first checked, and overlapping R_{1A} and R_{1B} is confirmed, as shown in Fig. 6 (b). R_{1B}, which has lower priority than R_{1A}, is tried to be transferred to a camera with no role (first) and cameras with role (if no candidates are found) for reducing frequent image-switching, as shown in Fig. 6 (b). In the same way, remaining tentative roles are also checked. Through these processes, (R_2, R_{3-4}, R_{1A}, and R_{1B}) were assigned to (C_3, C_2, C_1, and C_5), respectively, as shown in Fig. 6 (c).

2) Advanced reassignment rule: As work progresses, the

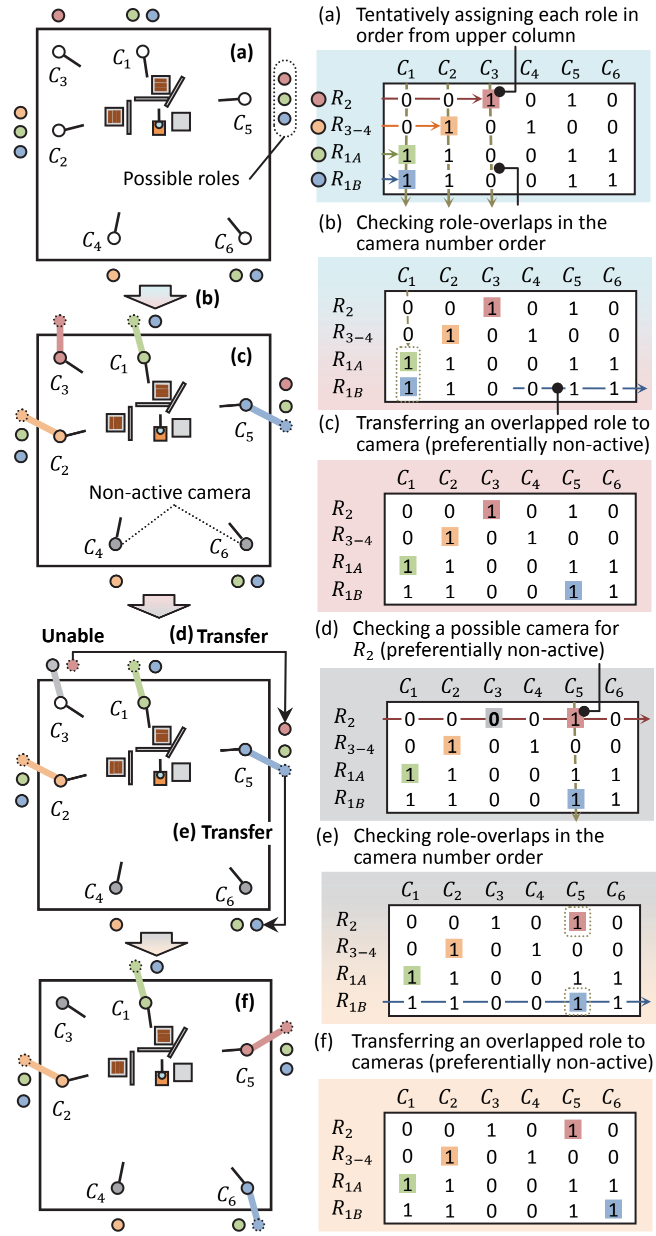

Fig. 6 Camera role assignment rules

positional relationship between cameras, machine, and obstacles change. If a camera cannot play an assigned role, the role must be reassigned to another camera while considering the overall balance. As shown in Fig. 6 (d), C_3 cannot play R_2. Possible cameras for R_2 are first checked in cameras with no role (C_4 and C_6), but they cannot play C_3. Next to be checked are cameras with a role. C_5 can play R_2 but is already play R_{1B}, as shown in Fig. 6 (e). Thus, possible cameras for R_{1B} are re-checked in cameras with no role, and C_6 is found. As a result, R_2 is transferred to C_5 and R_{1B} is transferred to C_6, as shown in Fig. 6 (f). If there are no cameras that can play the role, the camera assigned the role before is adopted.

The developed assignment system is easily implemented using linear programming [12] and can be applied to arbitrary configurations. This simple but effective system is important for practical applications such as unmanned construction.

V. Experiments

Debris transport experiments using a virtual reality (VR) simulator, as shown in Fig. 7, were conducted to evaluate the effectiveness of the developed autonomous control system.

A. Experimental conditions

1) VR simulator: A tele-operation simulator using a VR environment was developed. The VR environment was built by using OpenGL and Open Dynamics Engine. Control levers for a demolition machine with a grapple and environmental cameras with yaw, pitch, and zoom were developed (Fig. 7). Each configuration is shown in the upper part of Fig. 8. One in-vehicle camera image, four environmental camera images, and one machine status are displayed on a 2-D monitor.

2) Task: The evaluation task was a debris transport. The six objects to be transported were set in two debris yards, as shown in Fig. 8. To obscure the operator's vision, three walls were placed in front of the debris yards. The operator was meant to grasp the middle part of each object without making contact with anything else and transport it as quickly as possible without dropping it. The proposed autonomous system is compared with fixed camera and manual control systems.

3) Conditions: Camera parameters for the fixed and manual (initial state) systems were set as follows: the used cameras (C_3, C_2, C_5, and C_4) projected in the order of monitor number, pitch (-4, -3, -16, and -0.8°), angle of view (44, 30, 38, and 31°), and yaw (line directions as shown in Fig. 8) and were defined through pre-experiments for enhancing the in-vehicle camera image. For the autonomous system, R_2, R_{3-4}, R_{1A}, and R_{1B} were assigned to the monitor 1, 2, 3, and 4, respectively (Fig. 7). Four novice operators who were familiar with operating a machine in our VR simulator performed the task in four sets with each system. The completion time, number of grasping misses, dropped objects, and overload contacts, and the grasped position of each object were measured. The operators filled out questionnaires about subjective usability.

B. Experimental results

1) Time efficiency: The average task times for the three systems are listed in Table 1 (a). The task time for the autonomous system was the shortest and that for the manual system was the longest. In the manual system which inevitably needs time to control the cameras, machine operations often stopped when the cameras were controlled. In contrast, in the autonomous system, the cameras were autonomously controlled, so operators could continue to control the machine. The task time for the autonomous system is also shorter than that for the fixed system, meaning that it can provide effective images for smoothly controlling the machine and reducing waste stoppages and operations, as shown in Fig. 9.

2) Safety: The number of grasping misses, dropped objects, and overload contacts per set are listed in Table 1 (b). All items for the autonomous system were two times fewer than the other systems. Dropped objects in the fixed system were caused by the grapple grasping the end of the object while in the manual system they were caused by the pivot joint rotating at an excessive speed, although the object was initially

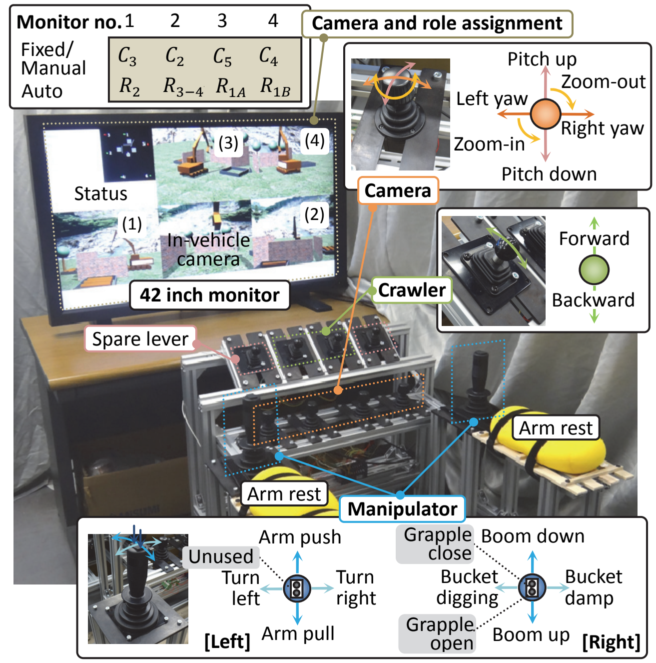

Fig. 7 Developed virtual reality simulator

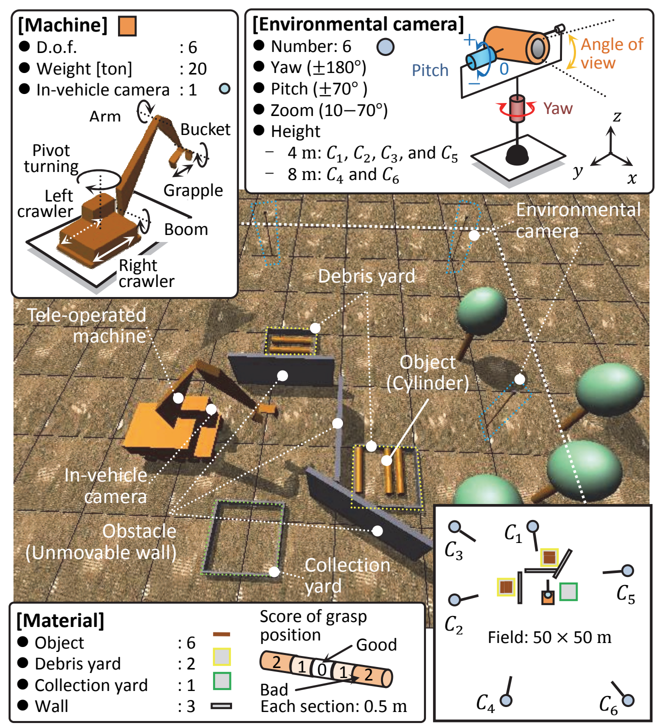

Fig. 8 Experimental environment

stable. This is because the operator could not ensure the stability of the grasped object by referring the fixed images since the images could not be controlled while the machine was being controlled. In contrast, the autonomous system can track an imaging object in accordance with the work situation, so the operators could always refer to the images to check careful spots peculiar to the situation.

3) Quality: To evaluate the quality of object grasping, the grasping position was measured. The relationship between the position and the score is shown in the lower part of Fig. 8.

TABLE I
EXPERIMENTAL RESULTS

Average for four operators	Auto	Manual	Fixed
(a) Time efficiency			
− Completion time s	395.1	427.7	414.3
(b) Safety (frequency per set)			
− Grasping miss	0.5	0.94	1.17
− Dropped object	0.06	0.25	0.19
− Overload contact	0	0.06	1.19
(c) Quality (good: 0↔bad: 2)			
− Score of grasping position	0.55	0.54	0.95
(d) Subjective usability (good: 1↔bad: 4)			
− Difficulty of overall task	1.25	1.75	2.75
− Frequency of blind areas	1.75	2.25	2.5
− Sensing difficulty of depth feeling	2	2	2.75

■: Best ■: Worst

Score 0 is best and score 2 is worst. The result is listed in Table 1 (c). In the manual and autonomous systems, the scores are two times lower than that in the fixed system. This result confirms that adaptive imaging to provide the desired view enables operators to perform more precise operations.

4) Subjective usability: Questionnaires were administered to the operators at the end of the experiments with three questions relating to the difficulty of the overall tasks, the frequency of blind areas, and the sensing difficulty of depth perception. The results are listed in Table 1 (d). The lower score represents better usability. The averages of all items are 1.67 (autonomous), 2.00 (manual), and 2.67 (fixed). The manual system is better than the fixed system, meaning that only the use of the controllable cameras is effective to improve subjective usability. In addition, by autonomously controlling cameras, both the difficulty of the task and the frequency of blind areas were more effectively reduced.

The analysis results indicate that the autonomous system adequately improved the time efficiency, safety, work quality, and subjective usability for all operators by providing adaptive images relevant to the situation.

VI. CONCLUSION AND FUTURE WORK

An autonomous system to control multiple environmental cameras was proposed for providing adaptive visual information suitable to the situation for unmanned construction. Four practical camera roles including the overview (machine), enlarge (end-point), posture (manipulator), and trajectory (manipulator) were defined by combining three imaging objects and four imaging modes. The role assignment system was developed for effectively assigning four roles to suitable cameras depending on the situation. To evaluate the developed autonomous system, debris removal tasks were conducted using a VR simulator and the results compared with fixed camera and manual control systems. The autonomous system decreased the number of grasp misses and error contacts and increased the subjective usability while improving time efficiency. As future works, the camera role parameters, including position adjustment, will be evaluated in other environmental conditions. More adaptive assignment system considering the cognitive ability of human operators and data acquisition methods in actual systems must be also discussed.

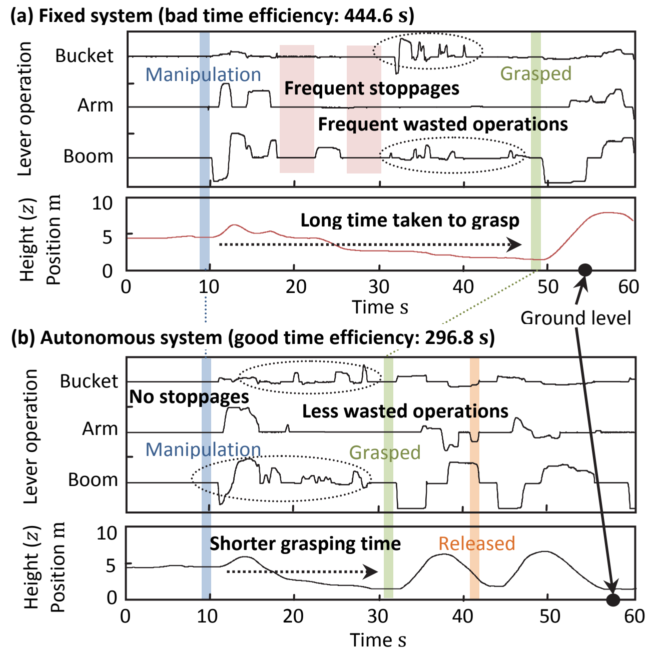

Fig. 9 Analysis of time spent for grasping an object

ACKNOWLEDGMENTS

This research was supported in part by the research grant FY 2011 from the Japan Construction Machinery and Construction Association and in part by the Research Institute for Science and Engineering, Waseda Univ.

REFERENCES

[1] S. Nishigaki, K. Saibara, S. Kitahara, H. Iwasaki, K. Yamada, and H. Satoh, "ICT-based work management to support unmanned construction for post-disaster restoration," in *Proc. Int. Symp. Automation and Robotics in Construction*, pp. 508–513, 2011.

[2] K. Yokoi, K. Nakashima, M. Kobayashi, H. Mihune, H. Hasunuma, Y. Yanagihara, T. Ueno, T. Gokyuu, and K. Endou, "A tele-operated humanoid robot drives a backhoe in the open air," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems,* pp. 1117–1122, 2003.

[3] Y. Ban, "Unmanned construction system: present status and challenges," in *Proc. Int. Symp. Automation and Robotics in Const.*, pp. 241–246, 2002.

[4] T. Xinxing, Z. Dingxuan, H. Yamada, and N. Tao, "Haptic interaction in tele-operation control system of construction robot based on virtual reality," in *Proc. Int. Conf. Mechatronics and Automation*, pp. 78–83, 2009.

[5] T. Hirabayashi, T. Yamamoto, J. Akizono, M. Iwasaki, and H. Yano, "Experimental land model of tele-operated underwater backhoe with AR technology," in *Proc. Int. Symp. Underwater Tech.*, pp. 339–344, 2004.

[6] D. Ryu, S. Kang, M. Kim, and J.B. Song, "Multi-modal user interface for teleoperation of ROBHAZ-DT2 field robot system," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, pp. 168–173, 2004.

[7] M. Jinki, S. Youngwan, P. Shinsuk, and K. Jinwook, "Development of immersive augmented reality interface for construction robotic system," in *Proc. Int. Conf. Control, Automation and Systems*, pp. 1192–1197, 2007.

[8] S. Kitahara and Y. Takashi, "Deployment of construction robots applying the information technology and network system," in *Proc. Int. Symp. Automation and Robotics in Construction*, pp. 19–23, 2006.

[9] D.B. Kaber, Y. Lee, M. Clamann, and Y.S. Lee, "Investigating human performance in a virtual reality haptic simulator as influenced by fidelity and system latency," *IEEE Trans. Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 42, no. 6, pp. 1562–1566, 2012.

[10] D.E. Crundall and G. Underwood, "Effects of experience and processing demands on visual information acquisition in drivers," *Ergonomics*, vol. 41, no. 4, pp. 448–458, 1998.

[11] M. Honma, N. Ikura, and K. Takai, "A study on video technologies for unmanned construction: back-hoe operations," in *Proc. Symp. Japan Const. Machinery and Const. Association*, pp. 61–64, 2007 (in Japanese).

[12] T. Ibaraki and S. Muroga, "Adaptive linear classifier by linear programming," *IEEE Trans. Syst. Sci. and Cybernetics*, vol. 6, no. 1, pp. 53–62, 1970.