View Systems for High Efficiency Teleoperations for Unmanned Construction based on Human Cognition Characteristics

無人化施工の高効率遠隔操作を目指した ヒトの認知特性に基づく視覚情報提示手法

February 2020

Ryuya SATO
佐藤　隆哉

View Systems for High Efficiency Teleoperations for Unmanned Construction based on Human Cognition Characteristics

無人化施工の高効率遠隔操作を目指したヒトの認知特性に基づく視覚情報提示手法

February 2020

Waseda University
Graduate School of Creative Science and Engineering,
Department of Modern Mechanical Engineering,
Research on Neuro-Robotics

Ryuya SATO
佐藤　隆哉

Abstract

Unmanned construction, which involves the teleoperation of construction machinery, has been introduced to areas affected by disasters such as earthquakes and volcanos. Such areas may be too dangerous for humans to enter. The unmanned construction machinery is controlled remotely by operators watching the views from cameras installed at the disaster sites. The crucial problem with unmanned construction is low efficiency; specifically, the work efficiency of unmanned construction is less than half that of ordinary on-board operation. Therefore, improving the efficiency of teleoperating heavy machinery under unmanned construction is crucially important. This thesis focuses on visual information for three cognitive reasons. First, humans acquire 70% of their information through their vision. Second, problems related to visual information are the most important in unmanned construction. Third, teleoperators mainly attend to visual information, ignoring other information for up to 30% of their teleoperation time.

Several researchers have developed visual support systems that provide information other than the simple images captured by machinery cameras. For instance, third-person views can be provided by drones or image processing, external cameras can be controlled to follow work states (e.g., grasping and releasing), and 3D and wide cab views can be obtained. These studies have provided various information to teleoperators.

However, as most of these studies do not consider human cognition characteristics, the systems impose excessive cognitive load on teleoperators. In previous studies, all information was provided to the operator during operations, requiring the teleoperators to simultaneously control the remote machinery and plan the moving paths and trajectories of the machinery arms. View systems to provide environmental information in advance are required to help the operators to plan their moving or grasping actions, reducing their cognitive load by removing the need to plan while working. Moreover, the techniques of previous studies provide no intuitive views (e.g., camera placement), although arbitrary third-person views are available. Furthermore, providing excessive information can cause cognitive tunneling, which focuses the teleoperators' attention on specific views while ignoring other views. Teleoperators are required to change their views depending on the work states. Therefore, a visual interface that avoids cognitive tunneling and attracts the operator's gaze to views appropriate for the work states is important to improve work efficiency.

In this thesis, the author develops a view system based on human cognition characteristics. In particular, the author addresses the following three technical challenges: (i) developing a view system that provides environmental information in advance, (ii) investigating the optimum and allowable camera placements, and (iii) developing a visual interface that avoids cognitive

tunneling. The thesis is divided into five chapters.

Chapter 1 summarizes the unmanned construction system, the problems of unmanned construction, and the causes of low efficiency (the crucial problem of unmanned construction). The importance of visual information in enhancing the work efficiency is also explained. Related studies on visual information, the limitations of these studies, and the purpose of the present study are highlighted.

Chapter 2 develops a prior view system for inputting environmental information based on the characteristics of a cognitive map, defined as a mental representation of the area. Cognitive maps can be roughly divided into two perspectives: survey and route perspectives. In the prior view system, the survey perspective is obtained through the third-person view of an arbitrary viewpoint, and the route perspective is obtained by a subjective view that can be changed by the teleoperator. Experimental results proved that the proposed prior view system can improve the quality and quantity of cognitive maps of important landmarks, including the target objects. Therefore, plans can be easily implemented in the proposed system. The acquisition of the survey perspective enables total planning, while the acquisition of the route perspective enables partial planning and improves the work efficiency. However, as some operators can forget their planned paths and trajectory, the author developed an augmented reality reminder which improves the work efficiency and eases the cognitive load.

Chapter 3 proposes an optimum and allowable camera placement for manipulation tasks. External views are essential even when teleoperators can watch wide 3D cab views. The author hypothesized an optimum and allowable area based on canonical views, which provide the highest performance in object recognition. Canonical views are characterized by minimal occlusion and an allowable rotation range of $\pm 30°$, and are almost unaffected by object size. Thus, the optimal pan and tilt angles were expected as 90° because this angle gives the canonical view. Meanwhile, the allowable pan and tilt angles were hypothesized as $\pm 30°$ to match the allowable rotation angles of the canonical views. The optimal and possible positions of the camera placements in manipulation tasks were experimentally investigated in a scale model and an actual machine with novice and skilled teleoperators as subjects. The experimental results are discussed and summarized. The results are applicable to camera-placement optimization in actual unmanned construction.

Chapter 4 develops a visual interface that avoids cognitive tunneling during teleoperation. Cognitive tunneling is caused by (i) focusing on views with high visual saliency, and (ii) low visual momentum. Visual saliency defines the ease of attracting a human's attention to an area, and visual momentum indexes the ease of integrating information through view transitions. The developed visual interface increases the visual momentum and attracts the teleoperator's eyes to views with low visual saliency. The visual momentum can be enhanced by including the same landmarks in the views of each work state. Moreover, human attention tends to focus on objects that vibrate at a specific frequency (5 Hz) in the effective field of view (±30°). Thus, whenever the work-state changes, the proposed interface displays a different external view within the teleoperator's effective viewing field, and vibrates it at 5 Hz for 0.5 s to capture the teleoperator's attention. The experimental results indicated that the proposed view system can decrease cognitive tunneling and improve the work efficiency in tasks requiring precise operations, such as grasping.

Chapter 5 summarizes the thesis and discusses the practical implementations of the proposed systems.

The developed view system is based on human cognition characteristics. A prior view system that inputs environmental information based on the characteristics of cognitive maps was first proposed. Next, an optimum and allowable camera placement based on the characteristics of canonical views was proposed, and was investigated in a scale model and on actual machinery. Finally, a view interface that avoids cognitive tunneling by increasing the visual momentum and lowering the visual saliency of views. The effectiveness of the proposed view system was evaluated in experiments using a simulator, a scale model, and an actual machinery.

Acknowledgements

This thesis was conducted at the Graduate School of Creative Science and Engineering, Waseda University. The research was supported by the Tateisi Science and Technology Foundation, Japan Institute of Country-ology and Engineering, the Casio Science Promotion Foundation, the Research Institute for Science and Engineering at Waseda University, the Leading Graduate Program in Embodiment Informatics, Japan, and a Research Fellowship for Young Scientists awarded by the Japan Society for the Promotion of Science.

First, I would like to gratefully thank my supervisor Prof. Hiroyasu Iwata, and Dr. Mitsuhiro Kamezaki. Their suggestions and comments on my research have helped me through seven years of laboratory life.

I also thank Dr. Takeshi Hashimoto and Mr. Mitsuru Yamada for providing us with the opportunity to conduct experiments on actual heavy machinery. I further thank my advisor during my intern at Fuji Xerox Palo Alto Laboratory (FXPAL), Dr. Don Kimber, and other researchers and interns. I could have a great time in Silicon Valley with all of you.

I greatly appreciate the research assistance of Dr. Junjie Yang during my bachelor's degree. I extend appreciation to the graduate and undergraduate students of the "Construction Machinery" group in Iwata Laboratory: Mr. Satoshi Niuchi, Ms. Saki Nakamura, Mr. Shinji Yamamura, Mr. Takahiro Eto, Mr. Yuki Yamashita, Mr. Yuichi Mizukoshi, Mr. Saiya Mizushima and Mr. Takumi Moteki. I would like to thank the other Iwata Laboratory members, the Leading Graduate Program members, and the Global Robot Academia members. I also thank the secretaries of Iwata Laboratory and all staff related to the Leading Graduate Program in Embodiment Informatics for their support.

Especially, I thank my friends and family for their support. Thanks to them, I have enjoyed a rich and fruitful student life.

Tokyo, February 2020

Ryuya SATO

Table of contents

List of figures

List of tables

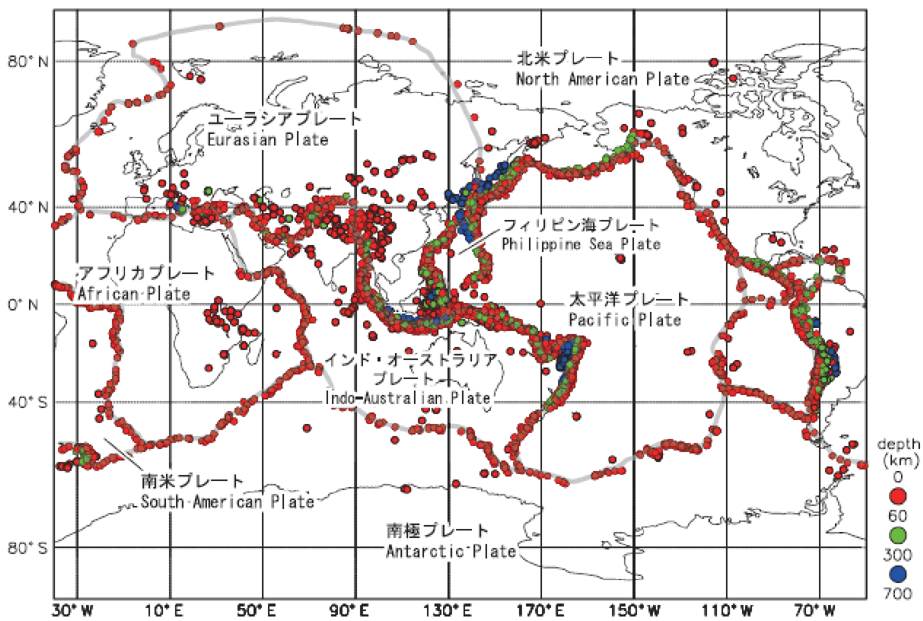
Chapter 1: Introduction

This chapter provides the background of this study, including a brief explanation of unmanned construction systems. It describes the teleoperation of heavy machinery, problems of unmanned construction systems, the importance of visual information, and previous researches on visual information in unmanned construction and teleoperation. The study purpose is also clarified.

1.1 Background of Unmanned Construction

1.1.1 Disaster-prone Japan

Japan has experienced many devastating natural calamities, including the Great East Japan earthquake, the heavy rains of July in Heisei 30, and the eruption of the Unzen volcano. Japan is known to be a disaster-prone country [1.1]. Figure 1.1 shows the plates and epicenter distributions of earthquakes with magnitudes of 5.0 or more worldwide from 2004 to 2013. Fig. 1.2 shows the



（注）2000～2009年，マグニチュード５以上。

Fig. 1.1 Worldwide distribution of earthquake epicenters [1.2]

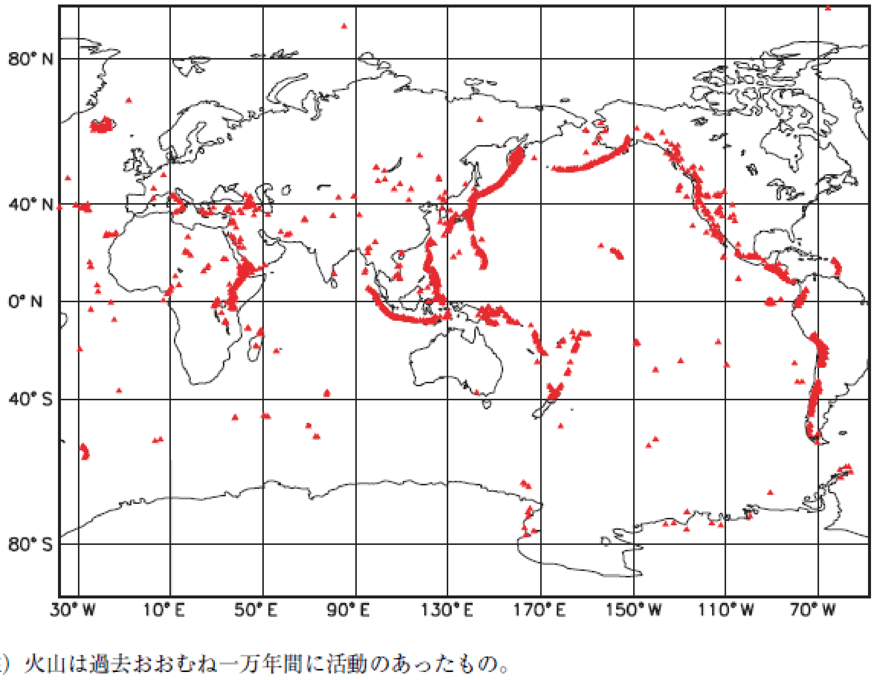


Fig. 1.2 Worldwide distribution of major volcanoes [1.2]

worldwide distribution of major volcanoes [1.2]. Fig. 1.3 shows the percentages of earthquakes in Japan and worldwide with Richter-scale magnitudes of 6 or more, and Fig. 1.4 shows the percentages of volcanoes in Japan and worldwide [1.2]. Japan has extremely high rates of earthquakes and volcano eruptions, considering that Japan's land area is only 0.25% of the world

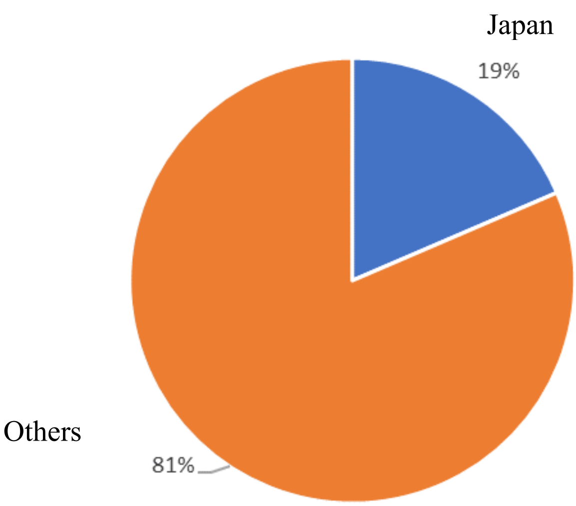


Fig. 1.3 Percentages of earthquakes with magnitudes of 6 or more in Japan and worldwide [1.2]

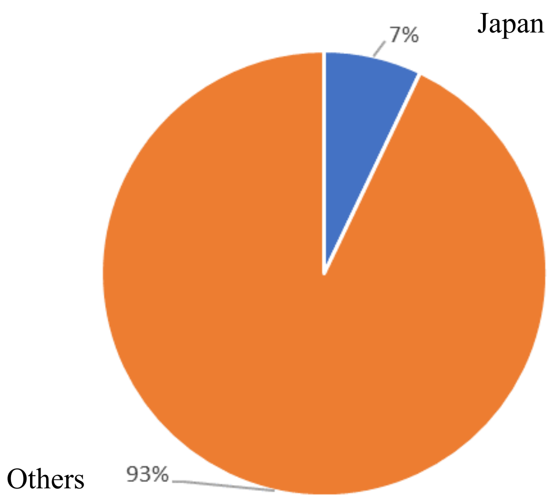


Fig. 1.4 Percentages of volcanoes in Japan and worldwide [1.2]

land area. Furthermore, 27,368 people have been killed or reported missing as a result of natural disasters from 1994 to 2013 [1.2].

1.1.2 Unmanned construction

Figure 1.5 shows the disaster site of the 2016 Kumamoto earthquake. In this case, there were risks of secondary disasters such as landslides and aftershocks. Because it was dangerous for humans to enter the disaster sites, an unmanned construction system, which included the teleoperation of construction machinery, was introduced [1.4–1.6]. Unmanned construction systems carry the construction machineries to a disaster site, and are teleoperated by humans from a distance (Fig. 1.6). Operators acquire information of the disaster sites by watching through cameras located at the disaster sites.

Unmanned construction technology consists of seven components: (a) construction machineries, (b) cab cameras, (c) environmental cameras, (d) communication equipment, (e) teleoperation room, (f) teleoperators of construction machineries, and (g) camera operators [1.7].

(a) Several construction machineries have been introduced at disaster sites. Examples are common backhoes, rough terrain cranes, and bulldozers that can run both on land and water [1.8]–[1.10]. Several attachments such as grapples, buckets, and breakers have also been used.

(b) Cab cameras are installed in the cabins. Cab cameras are important because they can provide similar views to operators on the construction machineries [1.7].

(c) Environmental cameras are installed for various purposes, such as viewing the entire

Fig. 1.5 Landslides at Kumamoto (taken at Kumamoto on 7/3/2016)

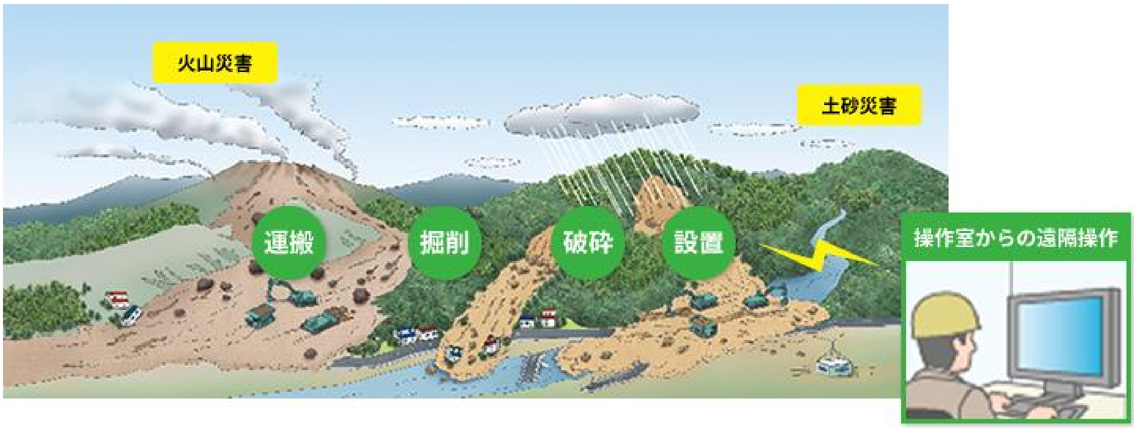


Fig. 1.6 Conceptual diagram of unmanned construction [1.3]

operation site and the areas being dug. Environmental cameras can be fixed or movable. The positions of fixed cameras cannot be changed. Movable cameras are installed on construction machineries, which can change the camera position. Both types of camera can zoom in and rotate.

(d) Communication equipment is used for several purposes including teleoperation signaling of the construction machinery and the viewing of relayed images. Relay cars (see Fig. 1.7) are

Fig. 1.7 Picture of environmental relay cars used for unmanned construction work ordered by Unzen Restoration Work Office, Ministry of Land, Infrastructure and Transport, and constructed by Asunaro Aoki Construction Co., Ltd.

used if communication is rendered difficult by long distance or obstacles. Relay cars are sometimes equipped with environmental cameras, as shown in Fig. 1.7.

(e) The teleoperation room houses several equipment, including monitors for the cab cameras and environmental cameras (see Fig. 1.8). One teleoperator of a construction machinery usually watches approximately four views, including one cab view.

(f) The teleoperators of construction machinery must have higher spatial awareness than usual boarding operators, because they must recognize the work sites and the construction machinery without entering the construction machinery. This spatial awareness is gained through regular training or experience on unmanned construction systems. Skilled teleoperators mainly work in Unzen, where unmanned construction was first introduced. Japan now has a deficit of skilled operators (approximately 20 operators nationwide), because of the lack of training fields and few regular works.

(g) Teleoperators for camera operators (usually called camera switchers) are required to control the environmental cameras and provide important information (along with the teleoperators) for the construction machinery. Camera switchers also require regular training to control the cameras adequately.

In Unzen, a serious pyroclastic flow occurred after the volcanic eruption in 1994 [1.11]. After the Unzen eruption, unmanned construction was introduced at the sites of the Mount Usu eruption, the 2004 Chuetsu earthquake, and the 2008 Iwate–Miyagi Nairiku earthquake. Until 2018, unmanned construction in Japan has been performed in 197 cases [1.3]. The major construction

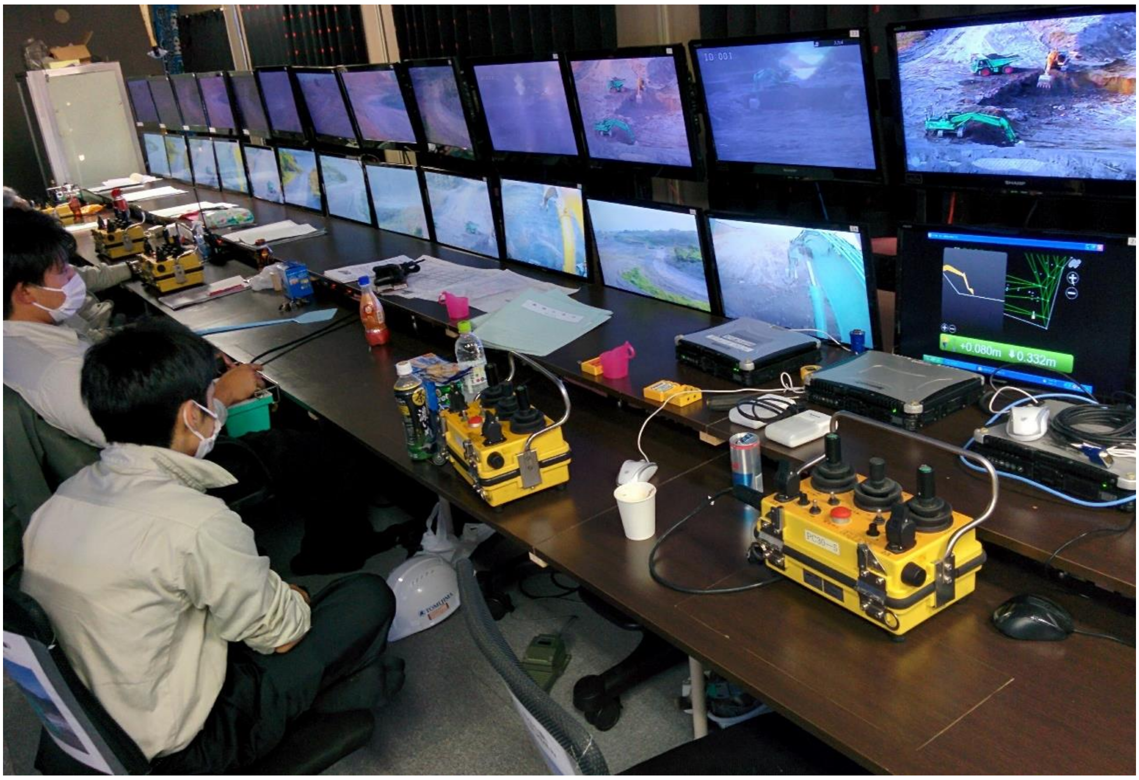


Fig. 1.8 Teleoperation room of an unmanned construction work ordered by Unzen Restoration Work Office, Ministry of Land, Infrastructure and Transport and constructed by Asunaro Aoki Construction Co., Ltd.

technology and works undertaken by unmanned construction are listed below [1.3].

- Compacted Concrete

  Fig. 1.9 shows a compacted concrete made by several construction machineries such as backhoes and bulldozers.
- Steel slit

  Fig. 1.10 shows a steel slit usually used in dams. During construction, the steel slit is grasped by attachments, including grapples, installed on backhoes.
- Concrete block masonry for dams

  Fig. 1.11 shows a concrete block masonry for dams. This work also requires backhoes with attachments for grasping the block.
- Earthwork

  Fig. 1.12 is an example of earthwork performed by combinations of backhoes, bulldozers, crawler dumps, and dump trucks.
- Placing sandbags

  Fig. 1.13 shows the placement of sandbags by backhoes with a hanging allowance.
- Box culvert

  Fig. 1.14 shows a box culvert made of concrete, which is usually used for waterways and

Fig. 1.9 Compacted concrete [1.3]

Fig. 1.10 Steel slit [1.3]

Fig. 1.11 Concrete block masonry for dams [1.3]

communication lines. Box culverts are grasped by backhoes with attachments and are transported by dump trucks.

Fig. 1.12 Earthwork [1.3]

Fig. 1.13 Placing sandbags [1.3]

Fig. 1.14 Box culvert [1.3]

1.1.3 Problems of unmanned construction

The main problem with unmanned construction systems is low efficiency. Moteki et al. showed that the efficiency of an unmanned construction system was less than half that of on-boarding operation [1.12]. Moteki et al. modeled an unmanned construction task as model tasks [1.13], which require operators to move in a curve, hook the target up by a bucket, and release it to the designated area (see Fig. 1.15). Moteki et al. compared the task completion time between the unmanned construction operation and the on-board operation. Subjects teleoperated the construction machinery by watching three views (one cab view and two environmental views), as shown in Fig. 1.16. As shown in Fig. 1.17, the task completion time of the unmanned construction was more than double that of the on-board operation.

The Japanese government requires rapid disaster response and recovery [1.14]. Humans can survive for approximately 72 h without food and water in the event of a disaster, the so-called "Golden 72 h" [1.15–1.17]. Unmanned construction was used to rescue missing people in the 2016 Kumamoto earthquakes [1.18]. Roads closed by landslides or fallen trees can prevent the transportation of food, water, and other essential supplies, and hence degrade the recovery speed. One of the most important roles of unmanned construction is securing the road for transportation by removing debris and gravel [1.14]. Therefore, the low efficiency of unmanned construction systems is a critical problem.

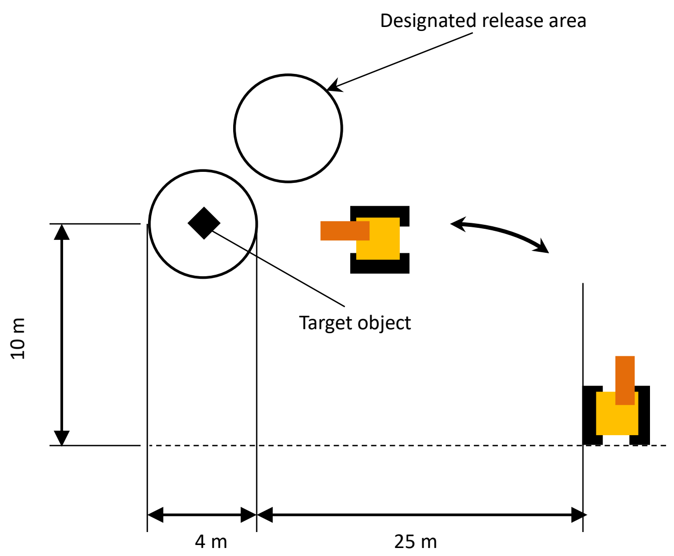


Fig. 1.15 Diagram of the model task (created with reference to [1.13])

1.2 Requirements to Improve Work Efficiency

1.2.1 Causes of low efficiency

The main cause of low efficiency is the difficulty involved in creating a mental map of the work sites or maintaining a situational awareness; that is, teleoperators can hardly recognize the environment and the machine situation [1.14, 1.19]. The author analyzed the causes of these difficulties based on previous studies on unmanned construction systems and teleoperations. The

Fig. 1.16 Teleoperation interface [1.12]

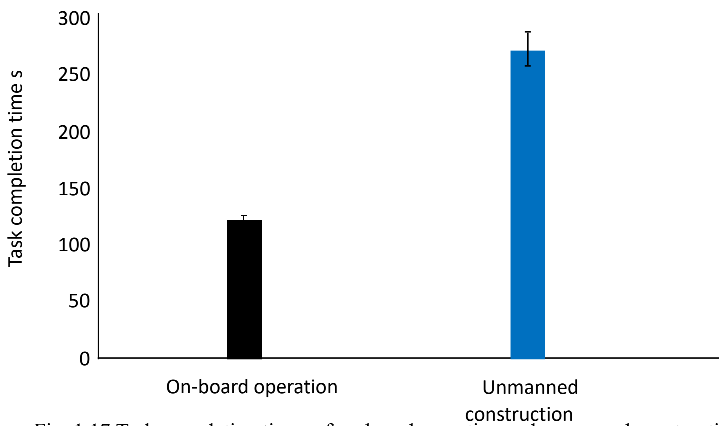


Fig. 1.17 Task completion times of on-board operation and unmanned construction (created by the author with reference to [1.12])

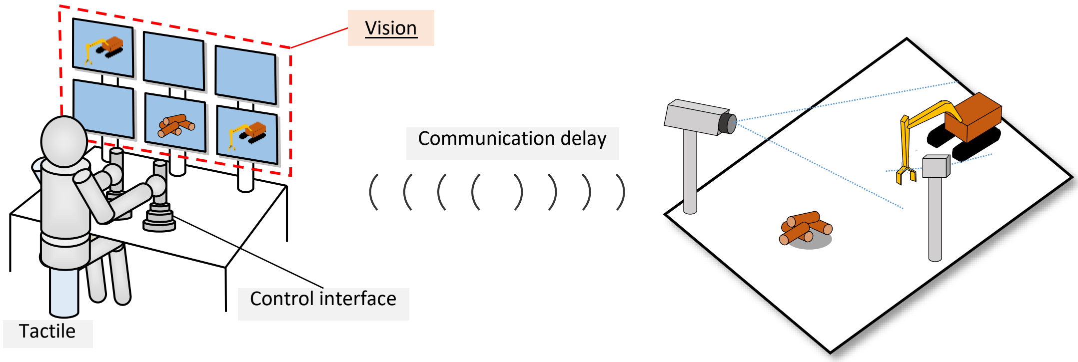


Fig. 1.18 Four major causes of low work efficiency

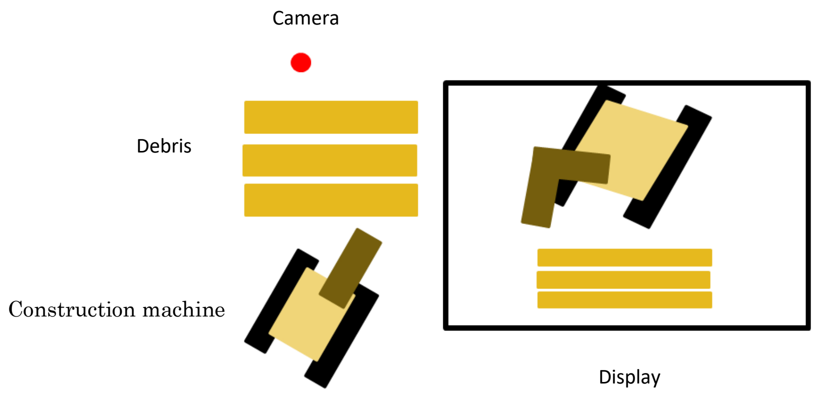


Fig. 1.19 Differences between the right and left sides of the construction machineries and camera views

analysis revealed that four main causes degrade the work efficiency: lack of visual information, lack of haptic information, communication delays, and inadequate interface for controlling the construction machinery (see Fig. 1.18). These causes lead to several problems including difficulties in recognizing the posture of the construction machinery, slips of the construction machinery, and recognizing the environment [1.14].

1.2.1.1 Lack of visual information

Lack of visual information can degrade the work efficiency [1.20–1.22]. Yamaguchi et al. [1.20] analyzed the causes of low efficiency based on previous research on unmanned construction systems. They revealed that teleoperators suffer from low depth perception (44% of visual problems), difficulty of differentiating the right and left sides of the construction machineries and camera views (Fig. 1.19) (17% of visual problems), and fatigue after three-dimensional (3D) viewing (11% of visual problems). Moteki et al. [1.21] conducted experiments of two cab views: a wide field of view and a conventional narrow field of view. They revealed that the wide field of view enabled higher work performance than the narrow field of view. Thus, the lack of visual information caused by the narrow field of view reduced the work efficiency. Woods et al. [1.22] revealed that teleoperators must imagine the work sites and machine situation when the work sites or machinery are partially missing in the views. This situation, called the keyhole effect [1.22], requires additional mental effort that can further reduce the efficiency.

1.2.1.2 Lack of haptic information

Lack of haptic information also degrades the work efficiency [1.21, 1.23, 1.24]. This fact was noted by Moteki et al. [1.21] after interviewing ten experienced teleoperators of unmanned construction. They revealed the necessity of haptic information, especially during contacts between attachments and objects. Moreover, haptic information provides teleoperators with material information about the object, including its hardness or softness. Several haptic devices have been developed for unmanned construction systems [1.25, 1.26]. The authors of [1.23] and [1.24] explained the necessity of obtaining haptic feedback, and developed a haptic feedback device.

1.2.1.3 Communication delay

Communication delay is another reducer of work efficiency [1.20, 1.27]. In a questionnaire study, Yamaguchi et al. [1.20] determined the cause of low efficiency among previous studies and teleoperators of two unmanned construction works. They showed that communication delay was the major cause of low efficiency; all communication was sent via relay cars and antenna base stations. Nitta et al. [1.27] investigated the effects of communication delay. Their experimental results proved that communication delays of less than 1.5 s are adequate for rough movements, but delays less than 1.0 s should be allowed for precise movements. Several systems that account for communication delays have been developed in teleoperation fields [1.28–1.30].

1.2.1.4 Control interface

Inappropriate control interfaces also degrade the work efficiency [1.31]. The Public Works Research Institute conducted experiments with two control interfaces: a conventional controller and a joystick similar to the interface of on-board operations. The operators of the joystick worked faster than those using the conventional controller. Moreover, several control interfaces with master–slave systems have been developed for unmanned construction [1.32, 1.33], and control interfaces using voices and gestures have been developed in the teleoperations field [1.34–1.36].

1.2.2 Importance of visual information

Among the four causes of reduced work efficiency given in Section 1.2.1, problems related to visual information are the most serious for the following five reasons. First, teleoperators of construction machinery complain primarily about poor visual information [1.20]. Second, humans acquire approximately 70% of their information from vision [1.37]. Third, human operators can hardly maintain their situational awareness when visual information is lacking [1.38]. Fourth, visual information accounts for approximately 30% of the attention span of teleoperations; haptic and other information commands minimal attention [1.39]. Fifth, teleoperators mainly judge and plan based on visual information, especially in navigation [1.38, 1.40].

The main cause of low efficiency is maintaining situational awareness [1.14, 1.19], which is required for planning and judging, and which is mainly acquired from visual information [1.38, 1.40]. Thus, teleoperators tend to focus on the information obtained from their camera views while ignoring other information [1.39]; this behavior can reduce the effectiveness of visual information [1.20].

Therefore, in this thesis, the author first addresses the problem of poor visual information.

1.3 Previous Researches

This section describes some of the previous researches on visual information in unmanned construction and other teleoperation fields.

1.3.1 Researches on unmanned construction

1.3.1.1 Next-generation remote-controlled machinery system

Furuya et al. developed a “next-generation remote-controlled machinery system” that aims to make a teleoperation room similar to a cabin in a construction machinery [1.41]. This system provides 3D views, a 360° view, and operator cabin chairs that rotate correspondingly to the rotation of the construction machinery (see Fig. 1.20). Experimental results indicated that this system increased the work efficiency.

Fig. 1.20 Next-generation remote-controlled machinery system [1.41]

(a) Cameras installed on the construction machinery

(b) Created overlooking view

Fig. 1.21 Overlooking views created by image processing [1.42]

1.3.1.2 Overlooking views by image processing

Sato et al. developed a bird's-eye view system based on image processing [1.42]. Using four fish-eyed cameras installed on the construction machinery, this system creates an overlooking view from above the construction machinery (see Fig. 1.21). In experimental tests, this system helped operators to avoid obstacles during the movement of the machinery, thereby improving the accuracy of the stop positions.

1.3.1.3 Third-person view

Nagatani et al. developed a third-person view system using a tethered drone [1.43]. The tethered drone reduces the weight of the battery and extends the flight time (see Fig. 1.22). Fig. 1.23 shows the views acquired from the drone. Fuchida et al. developed an arbitrary viewpoint

Fig. 1.22 Views from a drone [1.43]

Fig. 1.23 A tethered drone [1.43]

visualization system [1.44] using four fish cameras, which allows the movements of viewpoints by overlaying (see Fig. 1.24). Hojo et al. also developed a multi viewpoint visualization system [1.45, 1.46]. This system provides an arbitrary third-person view from omnidirectional past

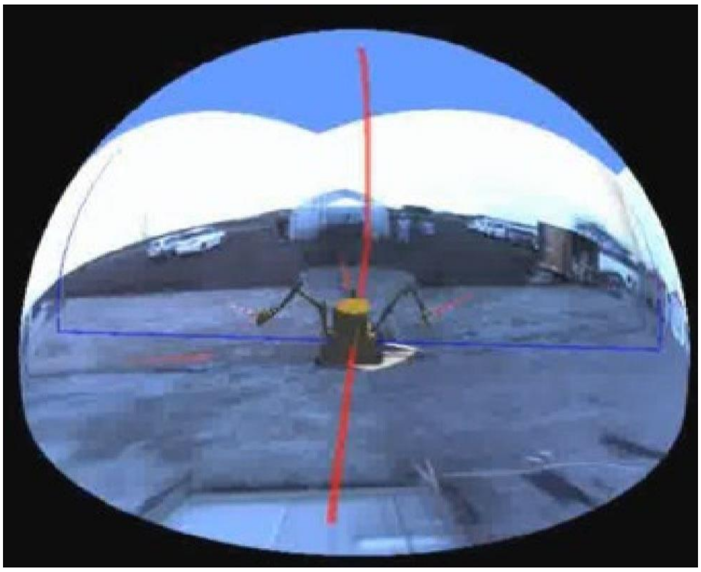


Fig. 1.24 Arbitrary viewpoint system by image processing from fish-eyed cameras [1.44]

Fig. 1.25 Arbitrary viewpoint system by past images [1.46]

images and by overlaying the computer graphics of the construction machinery (see Fig. 1.25) [1.46].

1.3.1.4 Autonomous multi-camera control system

Kamezaki et al. developed an autonomous multi-camera control system [1.47]. This system provides two overlooking views, two enlarged views, one overlooking view from above the environment, and a cab view based on the camera roles (see Fig. 1.26). The pan angles, tilt angles, and zoom of all environmental cameras were controlled to follow different work states, such as reaching and transporting an object. In experimental evaluations, this system decreased the blind spots and improved the work efficiency. Yamada et al. [1.48] introduced a viewpoint movement

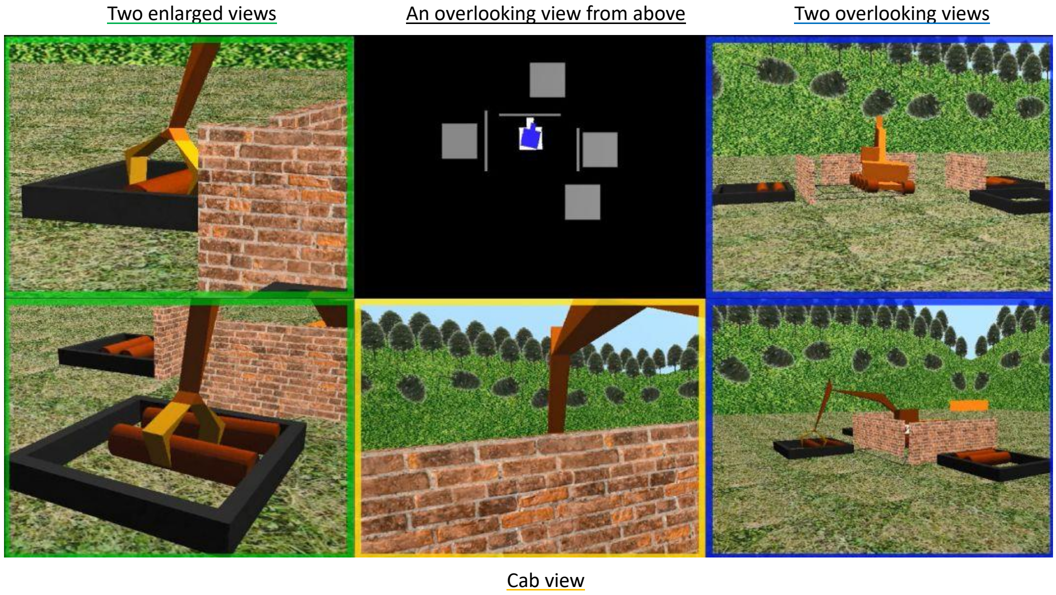


Fig. 1.26 Autonomous camera control system [1.47]

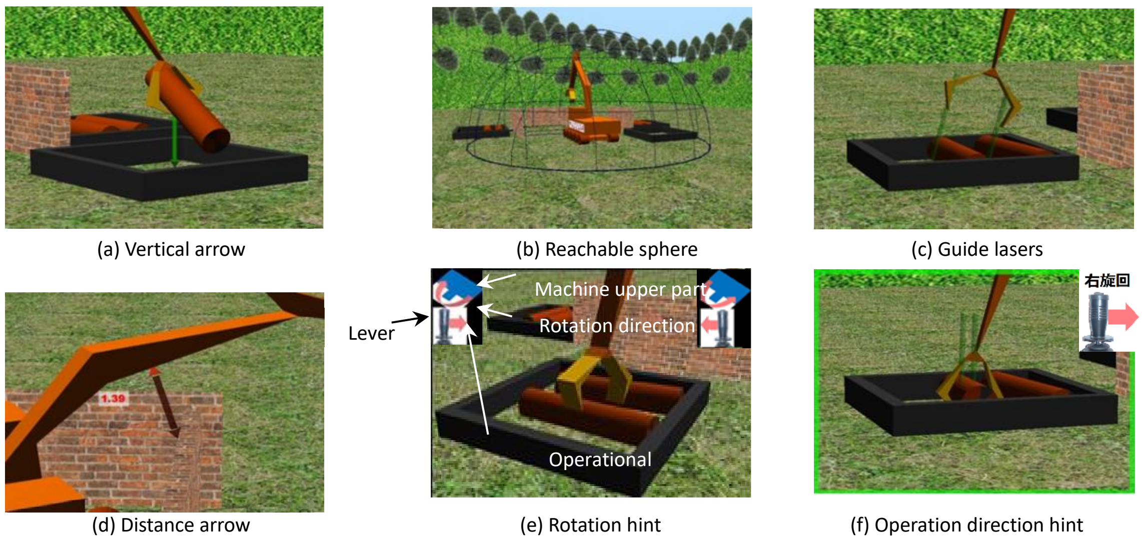


Fig. 1.27 AR for unmanned construction [1.49]

system using drones for visual support, which compares the effectiveness of several drone movements based on the work efficiency. They reported that the work efficiency can be improved if the drones are moved after the machine has grabbed the object.

1.3.1.5 Augmented reality

Yang et al. [1.49] introduced augmented reality (AR) for unmanned construction. Their system displays six ARs (Fig. 1.27): vertical arrows, guided lasers, and reachable spheres (for depth comprehension), distance arrows (for obstacle avoidance), rotation, and direction hints (for

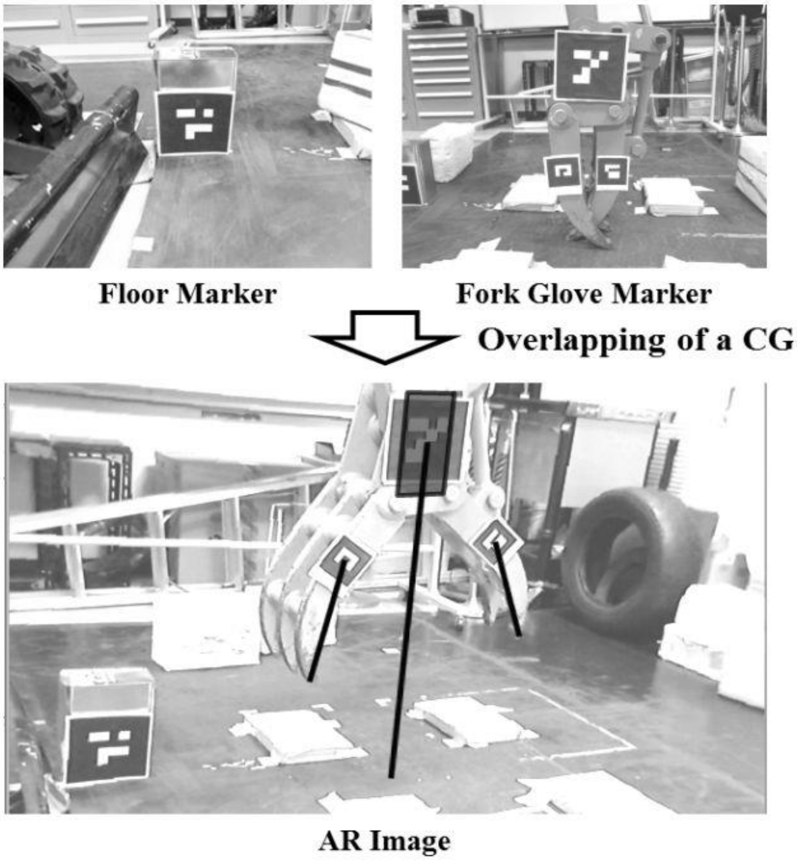


Fig. 1.28 Augmented reality using markers [1.50]

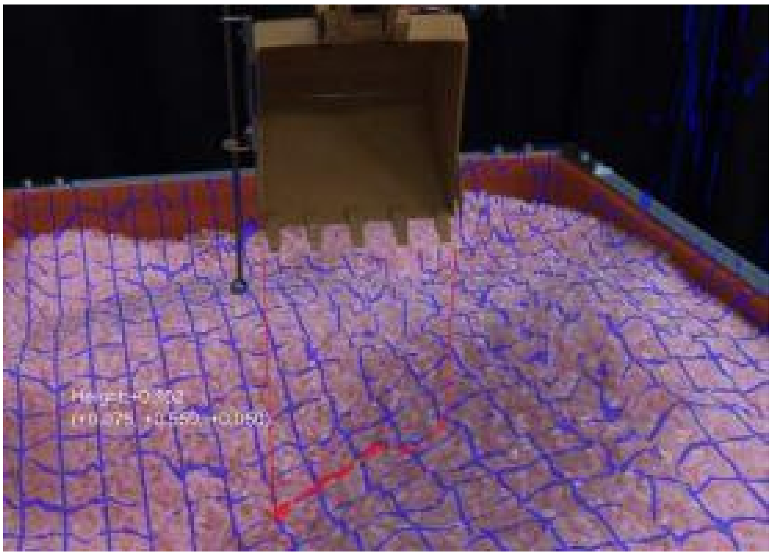


Fig. 1.29 Grid and shadow in augmented reality [1.51]

operation support). Experimental results indicated that this system can enhance the operators' comprehension of depth and improve their work efficiency. Yamada et al. [1.50] developed AR systems using markers in the laboratory (see Fig. 1.28). In experiments, this system improved the mental workload and work efficiency. The AR system of Tanimoto et al. [1.51] uses cameras and depth sensors [1.51] to display grids and shadows for depth perception (Fig. 1.29).

1.3.2 Studies on teleoperations other than unmanned construction

This section introduces visual-information studies in teleoperations other than unmanned construction systems.

1.3.2.1 Increasing field of view

A restricted field of view degrades the depth perception and reduces the number of essential distance cues [1.52]. To mitigate this problem, Scribner et al. introduced a wide field of view for driving [1.53], which proved useful under unfamiliar terrain conditions [1.53]. A wide field of view has also improved the performance of other teleoperations [1.54–1.56].

1.3.2.2 Three-dimensional views

Three-dimensional views have also been considered [1.57–1.59]. Drascis et al. [1.57] showed stereoscopic and monoscopic videos to trained operators, and confirmed that stereoscopic videos are more comfortable than monoscopic videos. Scribner et al. [1.58] revealed that stereoscopic videos reduced the error contacts. Draper et al. [59] showed that monoscopic videos improved work efficiency for difficult tasks. However, these types of 3D views can induce motion sickness [1.60]. In unmanned construction work, teleoperators need to work about 8 hours a day, so 3D views are inappropriate.

1.3.2.3 Immersive displays

Several immersive displays, including head-mounted displays (HMD), also have been developed [1.61–1.63]. Tachi et al. [1.61] introduced surrounding displays to enhance telexistence. Kot et al. [1.62] used HMD for a comfortable interface. Martins et al. [1.63] showed that HMD enhanced depth perception and situational awareness for search-and-rescue tasks. However, because these immersive displays increase motion sickness [1.64], they are unsuitable for unmanned construction work, which requires 8 hours per day of teleoperation.

1.3.2.4 Sensor fusion

Other studies have developed sensor fusion interfaces for teleoperations [1.65–1.67]. The visual interfaces of Meier et al. and Fong et al. fuse the information acquired from several sensors, including stereo vision and sonar [1.65, 1.66]. Livatino et al. developed an AR-based visual interface that displays videos and laser information [1.67].

1.4 Purpose of this thesis

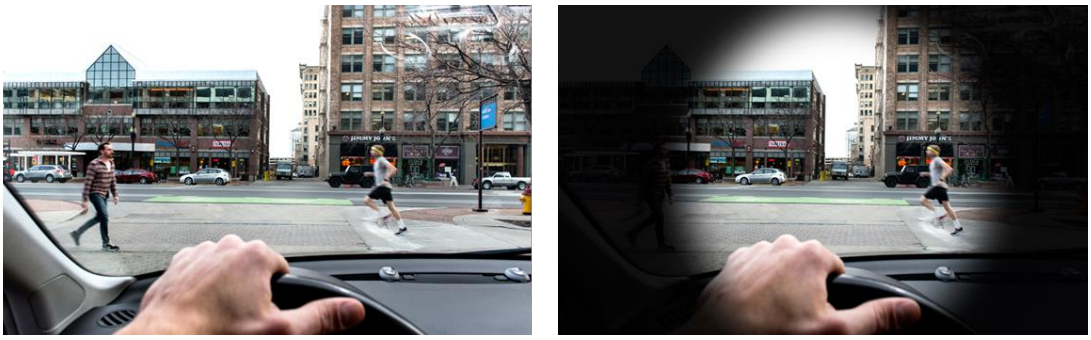
This section describes the purpose of this thesis.

1.4.1 Theme

Visual support systems for teleoperators have mainly focused on providing additional information. Because they do not consider the operator's cognitive factors, these systems impose excessive cognitive load on teleoperators. The problems in the previous studies can be divided into three categories: (1) provision of all information during the operations, (2) lack of intuitive views, and (3) provision of information from multiple (4−6) views. Problem (1) forces teleoperators to simultaneously work and plan; for example, they must simultaneously plan the trajectory and control the arm movements. Regarding problem (2), third-person views have been investigated but without considering which placement of cameras provide easily interpretable views. Problem (3) causes cognitive tunneling, meaning that operators focus on some views while ignoring others [1.68]. Fig. 1.30 shows an example of cognitive tunneling in a driving case [1.69]. When drivers are under high cognitive load, they will likely focus on a specific area as shown in the figure. To mitigate these problems, the author proposes a view system based on human cognition characteristics.

1.4.2 Target teleoperator

In this thesis, the target teleoperators are skilled at controlling construction machinery but novice at teleoperation. In Japan, because there are only 20 teleoperators skilled at both controlling construction machinery and at teleoperation, it usually is difficult to hire skilled teleoperators when local construction companies suddenly have to deal with disaster response [1.70]. This may mean that operators who have no experience in teleoperation but are skilled in controlling construction machinery are used in such situations.



(a) Driver not affected by cognitive tunneling (a) Driver in a cognitive tunneling scenario

Fig. 1.30 Example of cognitive tunneling in a driving scenario [1.69]

1.4.3 Objective

The objective of this study is to develop a view system based on human cognition characteristics. The problem is approached from three perspectives: (1) inputting environmental information in advance, (2) investigating an optimum and allowable camera placement, and (3) developing a visual interface that avoids cognitive tunneling.

(1) Current teleoperation systems require teleoperators to plan and work simultaneously (as mentioned in the previous section). Inputting environmental information in advance can help teleoperators to plan the required movements and grasping, reducing the cognitive load of planning during operations.

(2) Previous methods acquire an arbitrary view. An optimum and allowable camera placement can enhance the efficiency of teleoperations.

(3) Current teleoperation systems use four to six views, which can cause the cognitive tunneling effect. Teleoperators are required to change views depending on their work states [1.70]. A visual interface that avoids cognitive tunneling will enable the operators' gaze to move to the appropriate views depending on the state of the work in progress, thereby improving the work efficiency.

1.4.4 Originality

The original contributions of this study are listed below.

- Inputting information in advance
  - Most of the studies provide the information during operations.
- Investigating an optimum and allowable camera placement
  - Most of the studies acquire only third-person views.

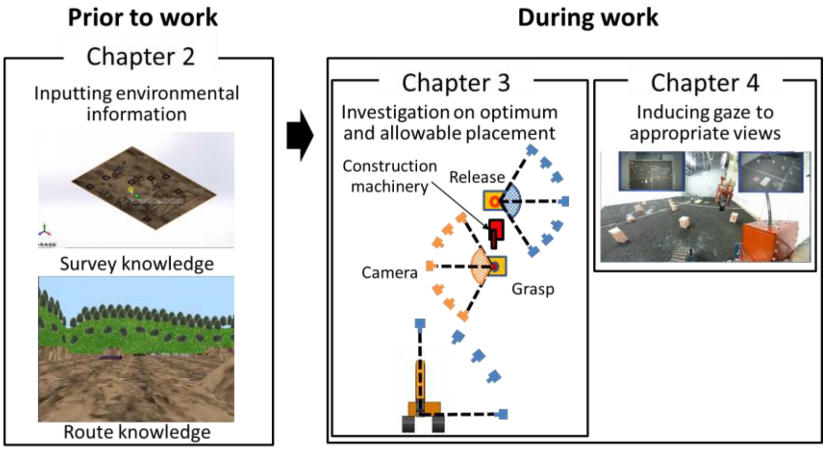


Fig. 1.31 Overview of this thesis

- Developing a visual interface to avoid cognitive tunneling
  - Most of the studies simply add information.

1.5 Overview

This section overviews the five chapters of the present study. Fig. 1.31 shows the overview of this thesis.

- Chapter 2

This chapter describes the developed prior view system that inputs environmental information based on the characteristics of the cognitive map within the teleoperator's mind. Cognitive maps can be roughly divided into two perspectives: survey and route perspectives. Survey perspectives are acquired from third-person views, whereas route perspectives are acquired form subjective views. Thus, the prior view system takes the characteristics of cognitive maps, and a third-person view from an arbitrary viewpoint provides the survey perspective. The route perspective can then be acquired by the changeable subjective view. Experimental results show that the proposed prior view system improves the quality and quantity of important landmarks in the cognitive maps, including the target objects to be grasped. Therefore, the proposed system enables easy planning; that is, the survey perspective enables total planning, while the route perspective enables partial planning and improved work efficiency. However, as operators can forget their planned paths and trajectory, the author develops an AR reminder that improves the work efficiency by reducing the cognitive load.

- Chapter 3

This chapter investigates an optimum and allowable camera placement for manipulation tasks. External views are essential even when the teleoperators are watching wide and 3D cab views.

Thus, we hypothesize an optimum and allowable area based on the characteristics of canonical views (views with the highest performance for object recognition). Canonical views are characterized by minimal occlusion, an allowable rotation range of ±30,, and little effect on object sizes. We hypothesize the optimum pan and tilt angle as 90, , where the canonical views have least occlusion, and the allowable pan and tilt angle as ±30, , equaling the allowable rotation angle of the canonical view. To investigate the optimum and allowable camera placement for manipulation tasks, experiments are conducted on a scale model and an actual machine. The subjects are novice and skilled teleoperators. The experimental results are summarized and discussed in each case, and are applicable to camera placement in actual unmanned construction.

- Chapter 4

This chapter presents the visual interface that avoids cognitive tunneling; that is, the propensity of teleoperators to focus on specific views and ignore other views. Cognitive tunneling occurs when the observed views have high visual saliency and low visual momentum. Visual saliency describes the easiness of focusing on the view, and visual momentum indexes the easiness of integrating information through a view transition. For example, the moon in the night sky has high visual saliency because it easily attracts humans' attention, and two views displaying completely different objects have low visual momentum. The developed visual interface increases the visual momentum and draws teleoperators' attention to views with low visual saliency. To raise the visual momentum, the displayed views in each work state include the same landmarks. Humans are attracted to objects vibrating at a specific frequency (5 Hz) in the effective field of view (±30,). Thus, the external views vibrating at 5 Hz in the effective field of view are displayed only when the work states are changed. The experimental results indicated that the proposed view system decreases cognitive tunneling and improves the work efficiency in tasks requiring precise operation such as grasping.

- Chapter 5

This chapter summarizes the study and explains real-life implementations of the proposed systems.

Chapter 2: View System for Inputting Environmental Information in Advance

This chapter describes the developed view system that inputs environmental information in advance. First, the importance of providing the advance environmental information is discussed. Next, a view system based on the characteristics of cognitive maps (maps in the human mind) is developed. Finally, the proposed view system is evaluated in experiments on a simulator.

The sentences, figures and tables in this chapter refer to author's work throughout the course of this study [2.1−2.6].

2.1 Importance of obtaining environmental information in advance

This section describes the importance of obtaining environmental information in advance.

When planning the paths that move, reach, grasp, and release objects, teleoperators benefit from knowing the prior environmental information of the work sites. This problem is explained in a sample environment with an obstacle and three objects that must be grasped and transported (see Fig. 2.1.) Without knowing the environmental information of the four transparent objects on the right side of Fig. 2.1(a), operators may take a longer path (e.g. the black dotted line in Fig. 2.1(a)) than if this information were given. Moreover, operators may stop moving to search for objects of unknown positional information. Meanwhile, without the environmental distance information between obstacles (black two-way arrow in Fig. 2.1(b)), operators may choose an oblique path to the target (e.g. the black dotted line in Fig. 2.1(b)).

Environmental information can be obtained either before introducing the unmanned construction or during the unmanned construction operations. The typical introduction period of unmanned construction is one week. Several researchers have developed view systems that provide environmental information to teleoperators during the unmanned construction work. Among such viewing systems are multiple-display systems [2.7], drones that display arbitrary views [2.8], and systems that automatically control the rotation and viewing angles of an environmental camera for multiple views [2.9]. Researchers outside the unmanned construction

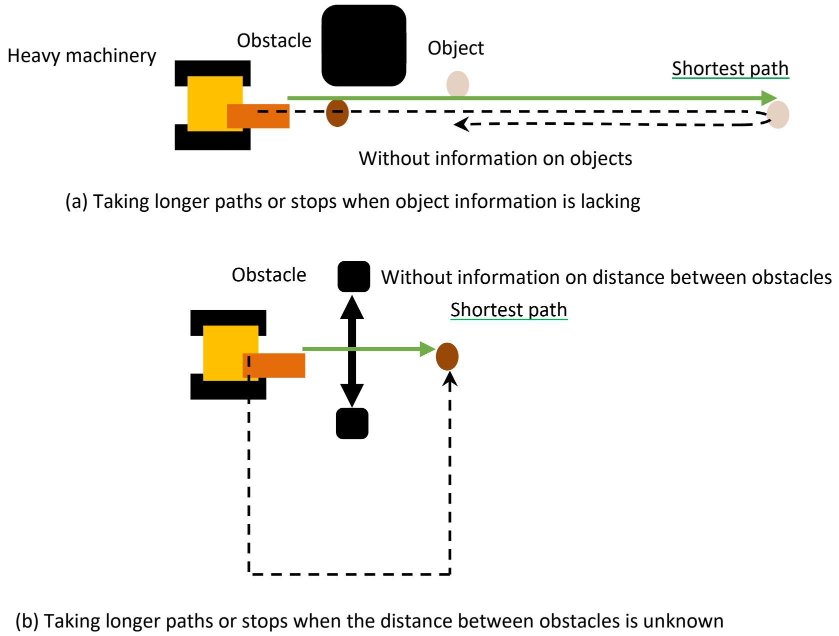


Fig. 2.1 Degradation of work efficiency when environmental information is lacking

field have also developed view systems for use during operations, including moving map displays for robotic orientation [2.10], gravity-referenced view displays showing a robot's attitude [2.11], and stereoscopic displays providing depth perception [2.12].

However, none of these studies provide the information before starting operations on the unmanned construction. When the advance environmental information is insufficient, the planning of movement paths and work strategies is suboptimal, because it must rely on the information obtained during the work [2.13]. A number of navigation systems display the correct path at the current moment. Examples are car navigation systems, head-up displays [2.14], and AR navigation systems [2.15]. Applying these navigation systems to unmanned construction is difficult, because they require the answer paths. At disaster sites, the answer paths are determined from diverse information such as the hardness of the ground, the ground water content, and 3D maps at disaster sites. Current systems can barely acquire all the necessary information at disaster sites, although they usually acquire the 3D map [2.16]. Furthermore, when the planning uses only the information acquired during work, the operators are forced to work and plan simultaneously.

As operators acquire 70% of their information from visual cues [2.17], vision is the most important of the five senses for judging and planning [2.18]. Thus, the author aims to develop a prior view system that provides environmental information, then investigate the work performance and operator-mental representation effects of this system. Furthermore, the author assumes that skilled teleoperators can forget their plans because teleoperators work only by watching several views for 8 hours per day. Thus, the author aims to develop an AR planning reminder system, which allows teleoperators to remember their planning right before each work commences.

2.2 View system for inputting environmental information in advance

This section develops the view system that inputs environmental information prior to starting work. The system is based on the characteristics of cognitive maps.

2.2.1 Human space cognition

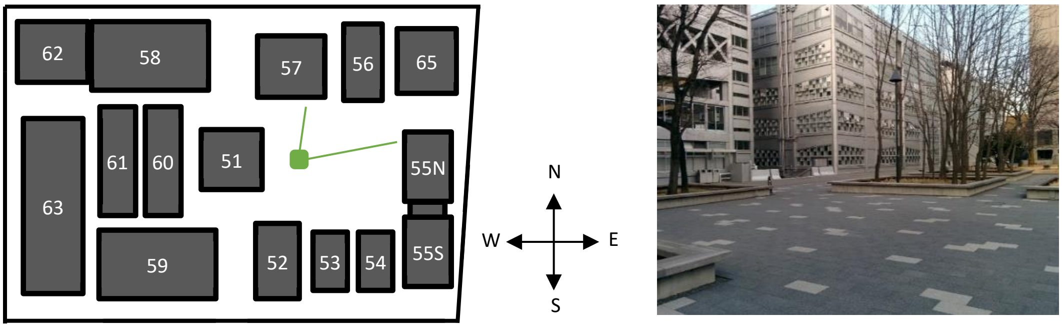
Humans store mental representations of space known as cognitive maps in their minds [2.19]. These cognitive maps are built by acquiring knowledge from survey and route perspectives [2.20]. Survey knowledge can be acquired from external viewpoints, whereas route knowledge can be acquired from personal or internal viewpoints. Therefore, survey knowledge is ordinally expressed in absolute coordinates such as east and west, whereas route knowledge is ordinally expressed in terms of relative coordinates such as front and back. Furthermore, the survey and route knowledge are ordinally depicted in maps and images, respectively, where the latter are taken from human viewpoints. For example, the map illustrated in Fig. 2.2 (a) provides survey knowledge and is expressed in terms such as south and north, whereas the picture shown in Fig. 2.2 (b) depicts the route knowledge and is expressed in terms of right and left.

2.2.2 Effectiveness of cognitive maps for teleoperators

The author investigated the effects of cognitive maps upon teleoperation-work performance. Operators can plan general paths with the survey knowledge, and work strategies in a particular area with the route knowledge.

2.2.2.1 Survey perspective

The acquisition of survey knowledge is assumed at a work site, as shown in Fig. 2.1 (a). Teleoperators can recognize the positions of objects and obstacles from external viewpoints.



(a) Survey perspective
(Map of Nishi-Waseda campus of Waseda University)

(b) Route perspective
(Picture taken from the green point in (a))

Fig. 2.2 Representation and difference of two perspectives (survey and route) of a cognitive map

Therefore, operators can plan the shortest movement path (green line in Fig. 2.1(a)), and thus avoid unscheduled stops when searching for objects.

2.2.2.2 Route perspective

The acquisition of route knowledge is also assumed in a work site (see Fig. 2.1 (b)). Teleoperators can recognize the distance between objects from the route knowledge because the views of the route perspective are similar to the views watched by operators during usual on-board operation. Therefore, teleoperators can take the shortest path to the target (e.g., green line in Fig. 2.1 (b)) because they can recognize the distances between various obstacles (black arrow in Fig. 2.1 (b)). In the work-site example of Fig. 2.3, the teleoperators are required to simultaneously work and plan their work strategies to grasp the left and right fallen trees without any prior route knowledge. Thus, the teleoperators may stop their operations to plan their work strategies for grasping the fallen trees, or might reduce their grasping speed. If the route knowledge is available, the teleoperators can plan their work strategies prior to starting the work

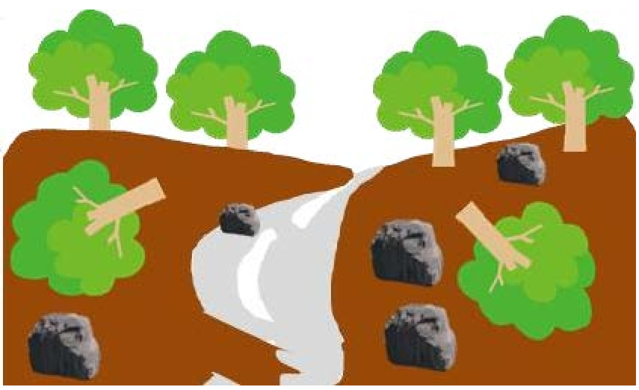


Fig. 2.3 Example of route perspective at a work site

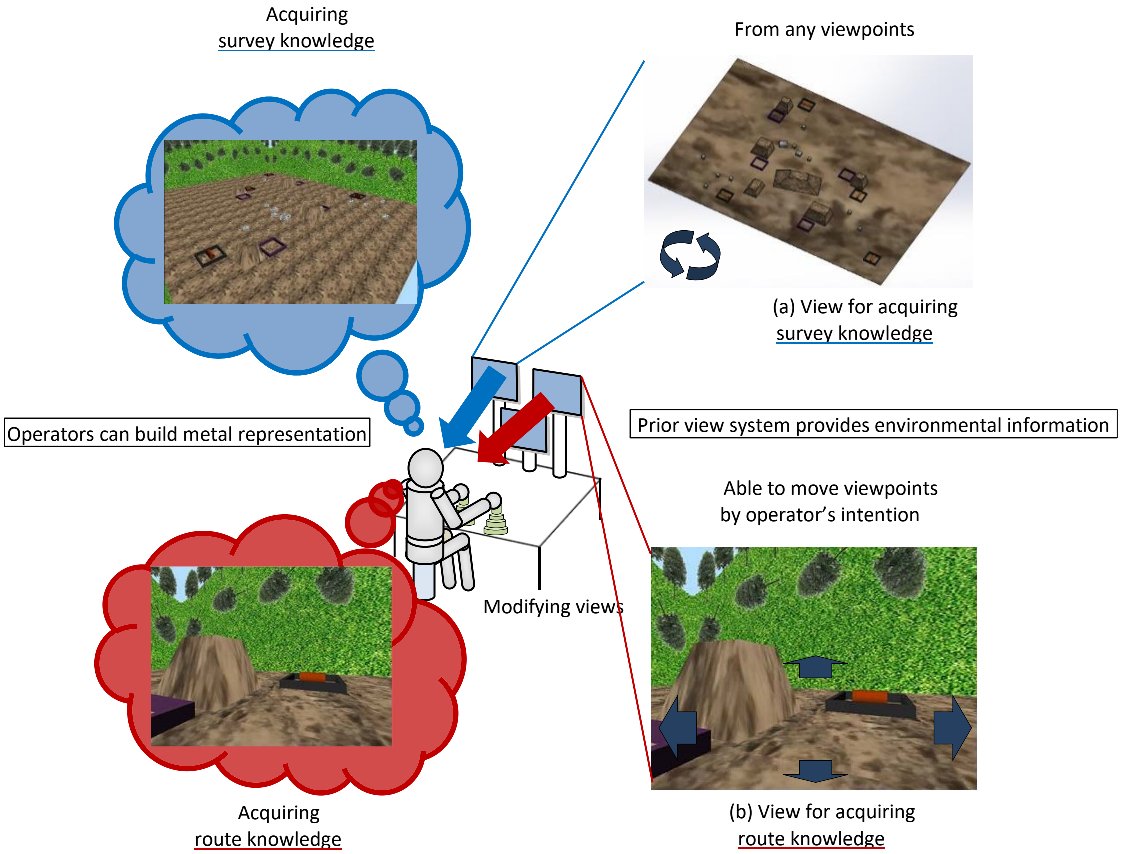


Fig. 2.4 Proposed prior view system for acquiring cognitive maps

because they can recognize the environment from familiar viewpoints. Therefore, the teleoperators can work without stopping and can maintain high-speed work during grasping.

2.2.3 Development of a view system that acquires cognitive maps

The work performance can be increased by building cognitive maps, as described in Section 2.2.2. Accordingly, the author developed a prior work-view system that acquires cognitive maps for teleoperators (see Fig. 2.4).

2.2.3.1 Survey perspective

Survey knowledge can be acquired from third-person views. Differences between the viewpoints received prior to and during the work must be corrected by mental rotation, which is not always easy [2.21]. Moreover, the proper viewpoints differ among operators [2.22] and tasks. Therefore, in the proposed system, teleoperators are presented with a third-person view that can be individually changed, as shown in Fig. 2.4(a).

2.2.3.2 Route perspective

Route knowledge can be acquired from first-person views. Route knowledge is more efficiently acquired by active movements, which modify the views according to intentions, than predetermined views [2.23]. Therefore, a first-person view that can be changed to fit the operators' intentions is displayed to teleoperators (see Fig. 2.4 (b)). Cognitive distance (the distance recognized by a human), is more important in decision-making than the actual distance [2.24]. Thus, the author proposes a first-person view that can be changed by operators at the speed of the heavy machinery.

2.2.4 Effectiveness of proposed system for teleoperators

The author assumes that introducing the proposed prior view system will benefit teleoperators.

2.2.4.1 Survey perspective

The proposed prior view system for survey perspectives enables teleoperators to accurately remember many substances from third-person viewpoints. However, it is not humanly possible to remember the positions of all substances at a disaster site [2.25]. Therefore, when following planned paths, teleoperators may choose to remember only the essential substances, including the debris to be transported and the release area of the debris. Moreover, teleoperators can plan paths in general and work without stopping, as described in Section 2.2.2.1.

2.2.4.2 Route perspective

Teleoperators usually control heavy machinery on-board at work sites, with little risk of secondary disasters. They also tend to watch cab views rather than third-person views [2.26]. Thus, teleoperators can plan their work strategies for grasping objects using the proposed prior route-knowledge view, rather than by remembering the positions of substances. The proposed views are similar to the views that operators usually watch during their on-board operations.

Furthermore, they can work at an increased grasping speed without stopping. The movement distance is decreased as described in Section 2.2.2.2.

2.3 Experiments

The effectiveness of the proposed prior view systems was evaluated in simulator experiments [2.27] (see Fig. 2.5), which are easier to perform experiments with large environments than physical experiments. The experimental procedures were approved by the ethics committee for human research at Waseda University.

2.3.1 Experimental settings

The experimental subjects were 16 novice participants with no experience in controlling heavy machinery. Skilled operators were not recruited because their number in Japan is very small (20) [2.28]. However, after completing the pre-experimental training tasks, the participants obtained sufficient skills to teleoperate the heavy machinery in this simulator. All participants (22–25 years) were male students enrolled at Waseda University.

The experimental tasks involved grasping four cylindrical target objects and transporting them one by one to the designated release area (boxes) in three environments (Environment 1, Environment 2, and Environment 3) (see Fig. 2.6). All environments contained target objects, designated release boxes, clods, stones, and slopes. The participants were asked to avoid any contacts with obstacles such as clods and stones.

Fig. 2.7 shows the experimental procedure and the views prior to and during the work. First, the author trained each participant in the required teleoperation skills. Second, the 16 participants were divided into two groups: an 8-participant control group and an 8-participant knowledge group. Equal division of the two groups ensured approximately equal average work times of the training task in each group. Third, the 16 participants tried three sets of the experimental tasks with differently displayed views prior to commencing the operations. In one experimental task, the participants watched the views to gain the environmental information before starting the work and subsequently teleoperating the heavy machinery. Two fixed third-person views of all three environments (sets) were displayed to the eight participants in the control group prior to the work. In addition to these fixed third-person views, a survey-knowledge view in the 1st set, a route-knowledge view in the 2nd set, and survey- and route-knowledge views in the 3rd set were displayed to the eight participants in the knowledge group prior to the work. A cab view and two

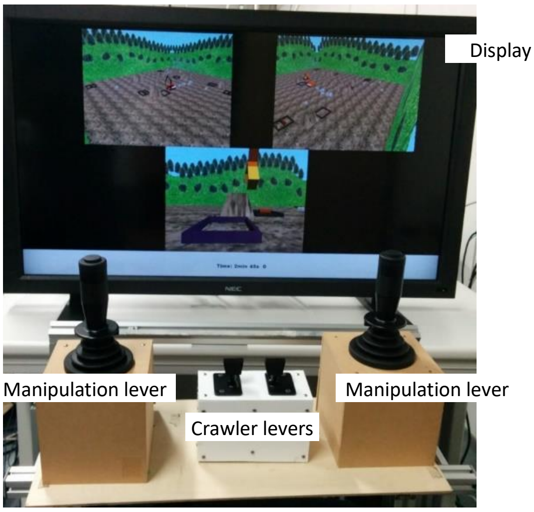
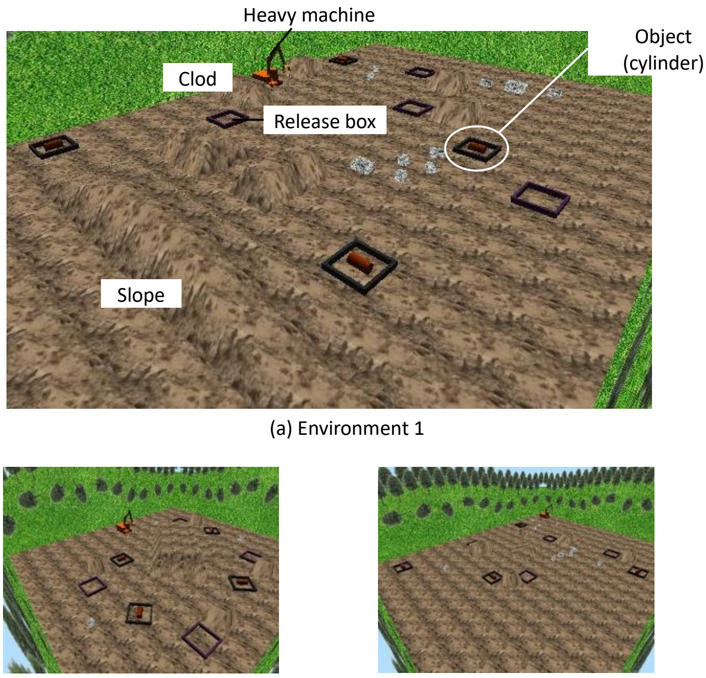


Fig. 2.5 Interface of the developed simulator



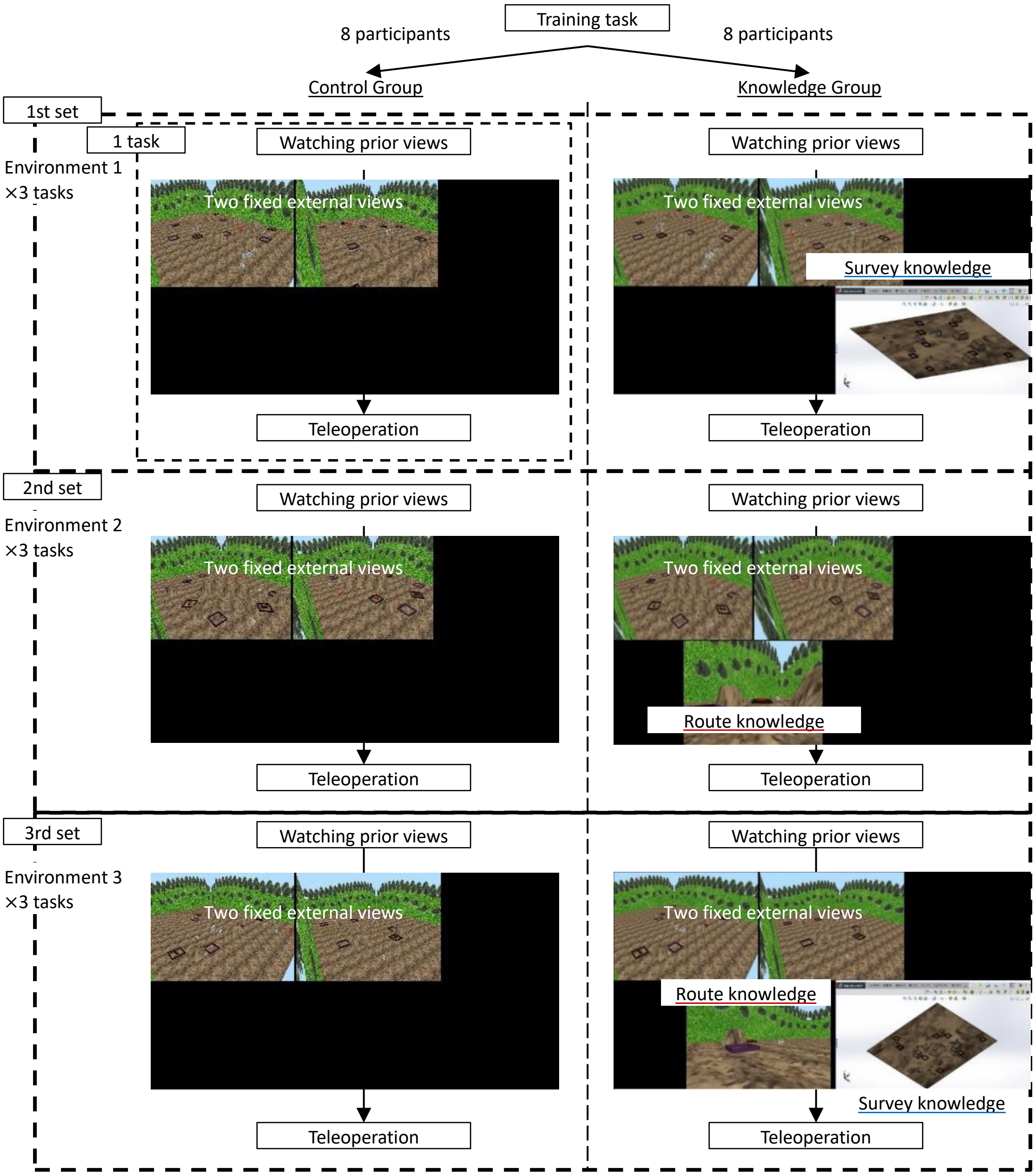
(a) Environment 1

(b) Environment 2

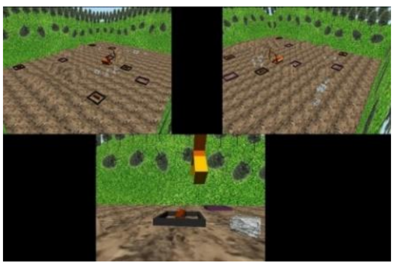
(c) Environment 3

Fig. 2.6 The three experimental environments

fixed external views of all three sets were displayed to the 16 participants during operations. Each set included three tasks, and all participants tried three sets under three conditions in the three environments. Environment 1 was used for the 1st set, and Environment 2 was used for the 2nd set. Environment 3 was used for the 3rd set. All participants were asked to prepare for the work by watching the displayed views for up to 10 minutes beforehand. Four parameters were measured during the work: the total task time, the number of stops, the movement distance, and the speed during grasping; the participants were asked to respond to questionnaires. Cognitive maps were also measured before work commencement, as explained in the next section.



(a) Procedure of the experiments and the prior displayed views



(b) Displayed views for teleoperation work

Fig. 2.7 Procedure of the experiments and the displayed views before and during work

2.3.2 Analysis method of cognitive maps

The cognitive maps were measured using sketch maps, as widely done in geography and cognitive psychology because of their high reliability [2.29]. Immediately after watching views and commencing operations, all participants were asked to sketch maps from their memory in PowerPoint 2013. Templates of the landmarks, which included the target objects, designated releasing boxes, clods, stones, and slopes, and the frames of the three environments, were already prepared as shown in Fig. 2.8(a) (the same template was prepared for the clods and slopes). Fig. 2.8(b) is an example of a sketch map made by a participant, and Fig. 2.8(c) is the corresponding actual map of Environment 1.

The two essential features of cognitive maps are the quantity (number of recognized landmarks among the sketched substances) and the quality (distance between the recognized landmarks and the landmarks on an actual map). These two features are important for building mental representations of work sites, including the positions of substances. To analyze cognitive maps, one must identify which sketched landmarks are recognized, and their correspondence to landmarks in actual maps. For instance, whether Object (i) in Fig. 2.8(b) can be recognized or not

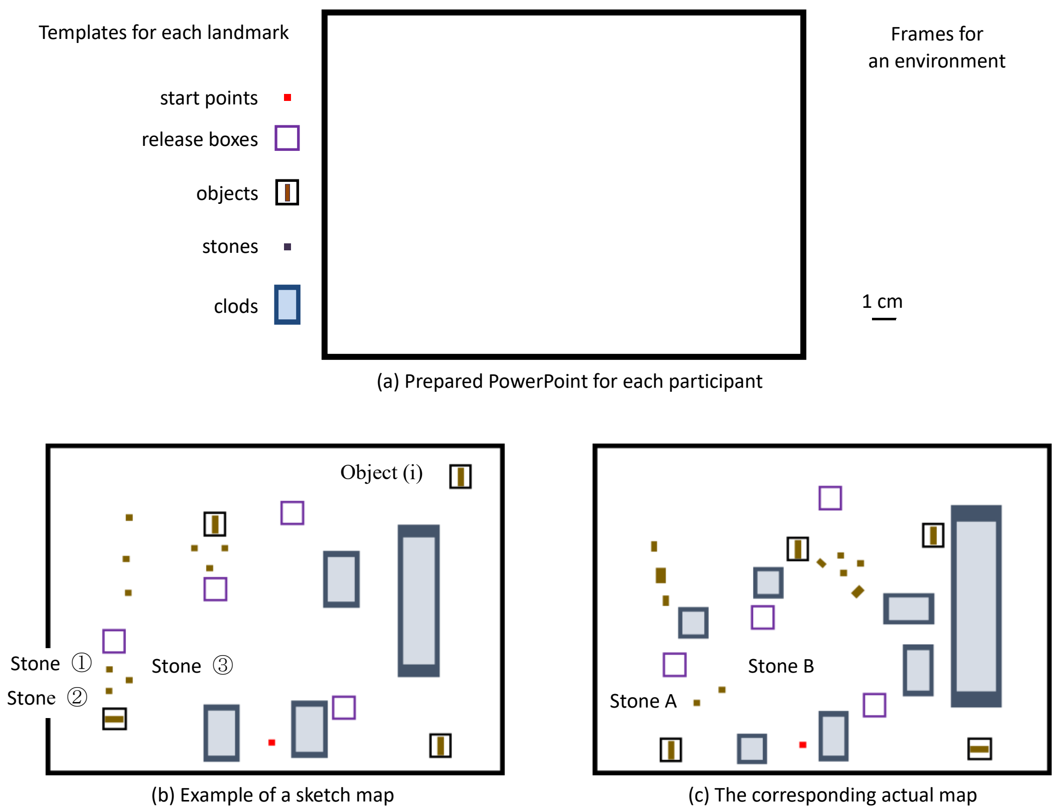


Fig. 2.8 Measuring cognitive maps using sketch maps

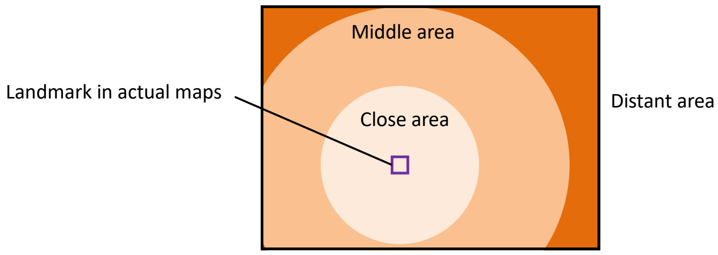


Fig. 2.9 Probability of landmarks being drawn in each area of the sketch maps

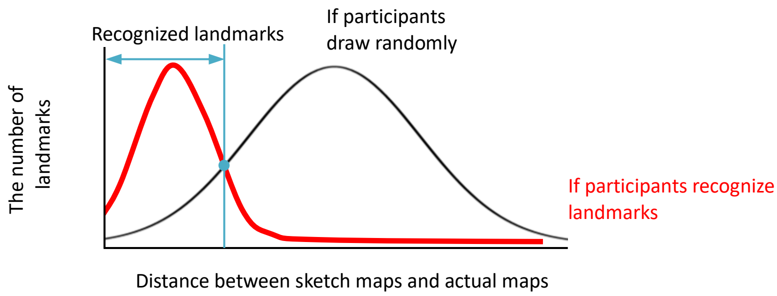


Fig. 2.10 Hypothesized histogram of distances between sketch and actual maps

is important, and determining its correspondence to an object in Fig. 2.8(c) is also necessary. The landmarks in the present cognitive maps differed from those in geography and cognitive psychology [2.30–2.32], which have specific names such as Tokyo Disneyland, and which are usually located in cities. In sketch maps with the same landmarks, no analytical methods can identify which landmarks in the sketch maps are recognized and correspond to which landmarks in the actual maps. Thus, the author developed an analytical method to identify the landmarks that were recognized in the sketch maps and their correspondences with the landmarks on the actual maps.

If the landmarks were randomly sketched, they were probably sketched in a middle area of the actual maps, and were unlikely to be sketched close to or distant from the landmarks in the actual maps, (see Fig. 2.9). Therefore, the histogram of the distances between the landmarks in the sketch maps and the actual maps should be Gaussian, as shown by the black curve in Fig. 2.10. On the contrary, if recognized landmarks are sketched, the histogram will follow the red curve in Fig. 2.10 because the positions of the sketched landmarks will approximate those in the actual maps. Therefore, the landmarks in the cognitive maps are recognized as actual landmarks at distances less than the distance from an intersection in a histogram. The width of each column of the histogram was calculated by the Freedman–Diaconis rule, which uses a quartile basis. The distance between the mid-points of the sketched and actual landmarks was calculated, and duplicate landmarks were eliminated. To illustrate this process, consider the three stones (① to ③) in the bottom left of Fig. 2.8 (b)). The distances between Stone ① and all ten stones in the actual maps were calculated. The distance calculation was repeated for stones (② and ③), and the results are given in Table 1. If a threshold distance were 3.6, stones ①, ②, and ③ were

Table 2.1. Example results of distance measurements

	Stone A	Stone B
Stone ①	2.6	3.5
Stone ②	1.4	3.0
Stone ③	2.4	3.2

recognized as duplicate landmarks (Stones A and B)., which must be eliminated. The author eliminated them by minimizing the total distance under the condition that Stone ② was recognized as Stone A and Stone ③ was recognized as Stone B. Therefore, Stone ② was recognized as Stone A, and Stone ③ was recognized as Stone B; the other stones were eliminated.

2.3.3 Experimental results and discussion on work efficiency

2.3.3.1 Survey perspective (1st set)

Fig. 2.11 shows the results of the first set. Presented are the task time (Fig. 2.11(a)), number of stops (Fig. 2.11(b)), movement distance (Fig. 2.11(c)), and speed during grasping (Fig. 2.11(d)). The task time, number of stops, and movement distance were significantly lower in the knowledge group than in the control group (Welch's t-test, task time: $t(40) = 3.22$, $p = 0.003$, number of stops: $t(33) = 2.26$, $p = 0.03$, movement distance: $t(43) = 2.10$, $p = 0.04$). Also, the speed during grasping was significantly higher in the knowledge group than in the control group (Welch's t-test, $t(44) = 2.23$, $p = 0.03$). These results prove that watching the proposed prior survey-knowledge view can decrease the task time, the movement distance, and number of stops, and increase the grasping speed.

2.3.3.2 Route perspective (2nd set)

Fig. 2.12 shows the results of the second set. Panels (a), (b), (c) and (d) of this figure show the

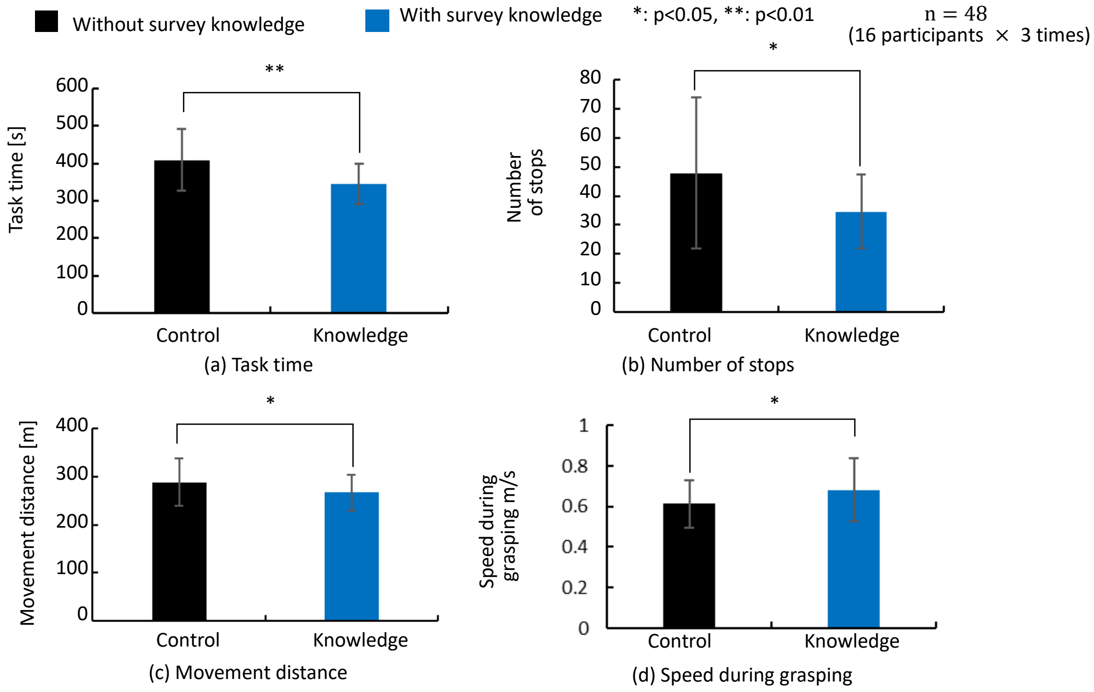


Fig. 2.11 Results of the 1st set (survey knowledge)

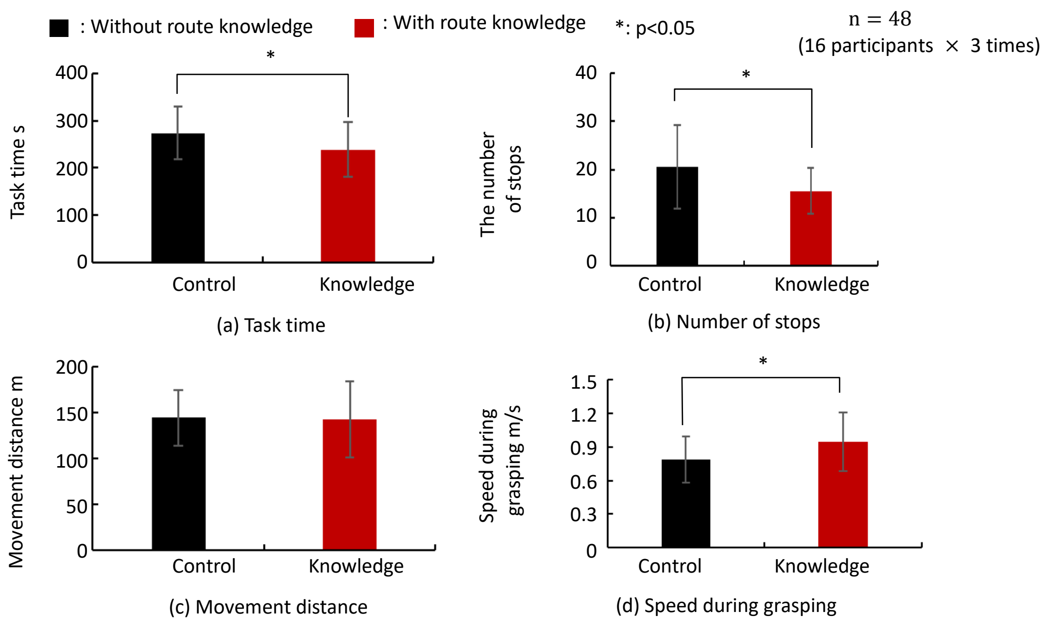


Fig. 2.12 Results of the 2nd set (route knowledge)

task time, number of stops, movement distance, and speed during grasping, respectively. The task time and number of stops were significantly lower in the knowledge group than in the control group (Welch's t-test, task time: $t(46) = 2.12$, $p = 0.04$, number of stops: $t(36) = 2.50$, $p = 0.02$). Meanwhile, the speed during grasping was significantly faster in the knowledge group than in the control group (Welch's t-test, $t(44) = 2.37$, $p = 0.02$). Those results show that watching the

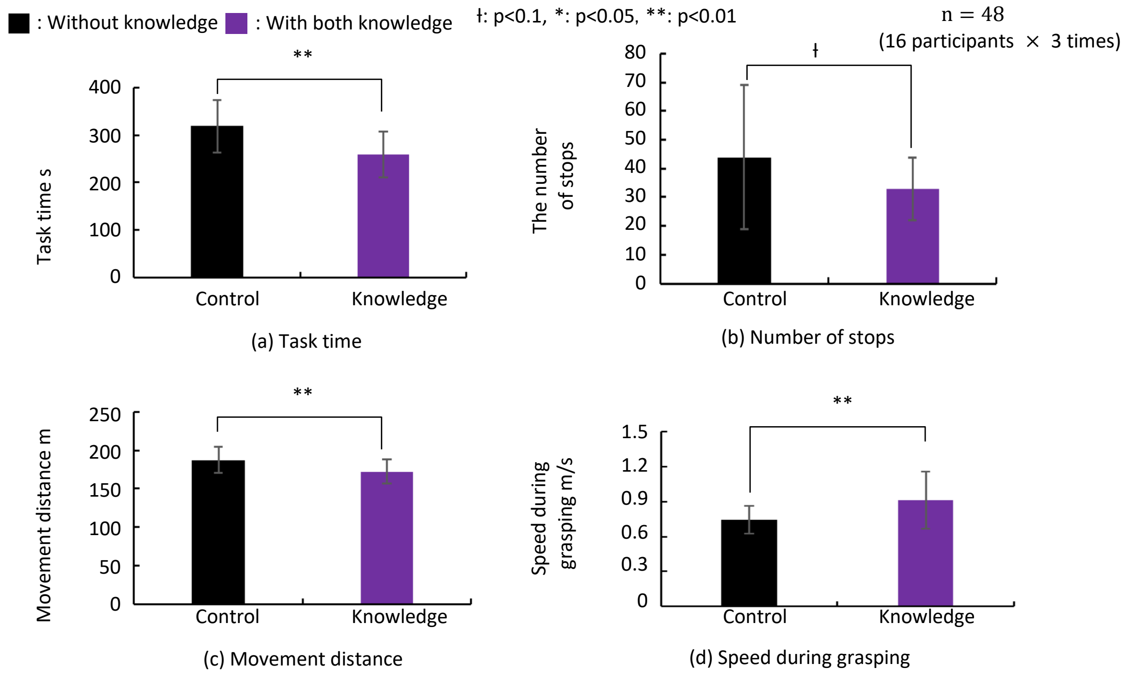


Fig. 2.13 Results of the 3rd set (both forms of knowledge)

proposed prior route-knowledge view can decrease the task time and number of stops, and increase the speed during grasping.

2.3.3.3 Survey and route perspective (3rd set)

Fig. 2.13 shows the results of the 3rd set. Again, panels (a), (b), (c) and (d) present the task time, number of stops, movement distance, and speed during grasping, respectively. The task time and movement distance were significantly lower in the knowledge group than in the control group (Welch's t-test, task time: $t(45) = 4.04$, $p < 0.001$, movement distance: $t(46) = 3.25$, $p = 0.002$), and the number of stops was marginally significantly lower in the knowledge group than in the control group (Welch's t-test, $t(31) = 1.98$, $p = 0.06$). Furthermore, the speed during grasping in was significantly higher in the knowledge group than in the control group (Welch's t-test, $t(34) = 3.00$, $p = 0.005$). These results prove that watching the proposed prior survey- and route-knowledge view decreases the task time, the movement distance, and the number of stops, and increases the speed during grasping.

2.3.3.4 Discussion

Table 2.2 summarizes the effectiveness of the proposed prior view system. The survey-

Table 2.2. Summary of the results from each perspective

✓: $p < 0.1$, ×: $p \geq 0.1$

	Task time	The number of stops	Movement distance	Speed during grasping
Survey knowledge	✓	✓	✓	✓
Route knowledge	✓	✓	×	✓
Both knowledge	✓	✓	✓	✓

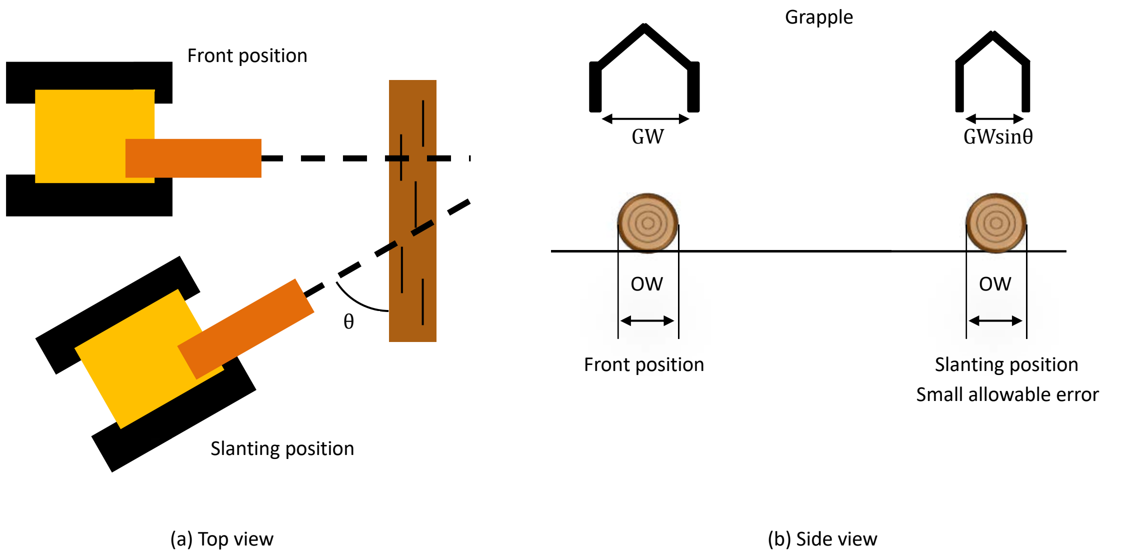


Fig. 2.14 Difficulties of grasping from slanted positions

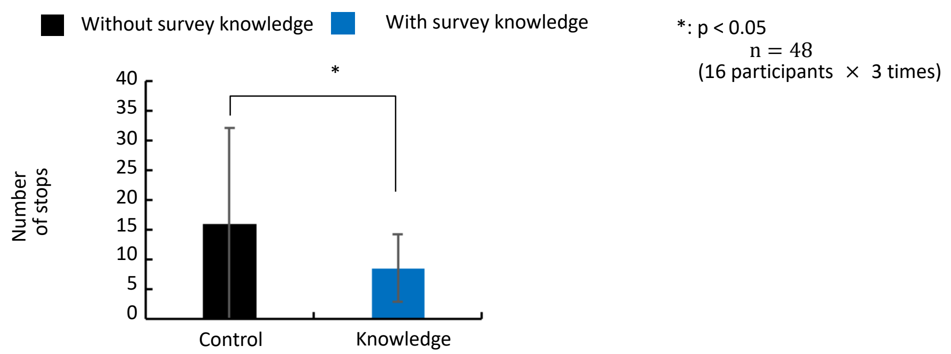


Fig. 2.15 Number of stops when grasping the circled object in Fig. 2.7 (a)

knowledge view improved the task time, the number of stops, the movement distance, and the speed during grasping. Meanwhile, the route-knowledge view improved the task time, the number of stops, and the speed during grasping. Two of these results differ from the assumptions described in Section 2.2.4: First, the survey-knowledge view increased the speed during grasping, and second, the route-knowledge view did not decrease the movement distance. These findings are discussed below.

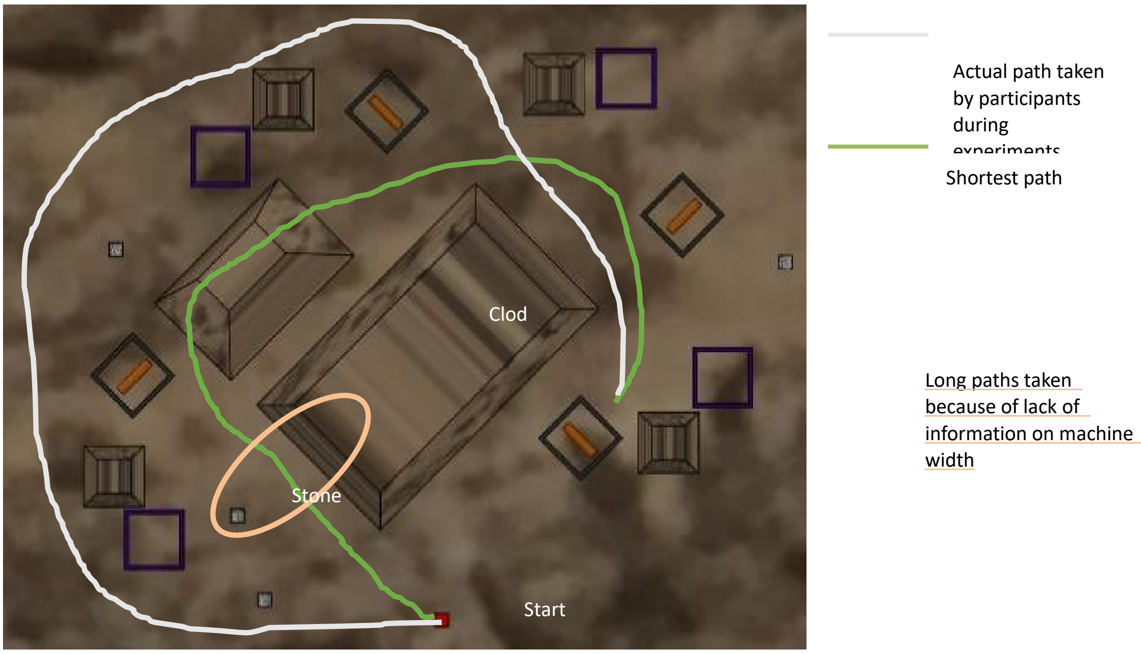


Fig. 2.16 Shortest path (green) and the actual path taken by the participants during the experiments

1) Improvement of speed during grasping with survey-perspective knowledge: The speed of grasping the target object might depend on the grasp position. Teleoperators will need to stop more often if they attempt to grasp target objects from slanting positions. When the angle θ between the heavy machine and the object is small, the grasping can be difficult (see Fig. 2.14). The allowable error in grasping the target object is given by $GW\sin\theta - OW$, where GW is the grapple width, and OW is the object width. Therefore, the object grasping increases in difficulty as θ decreases, causing an increase in the number of stops. Fig. 2.15 shows the number of stops when the participants tried to grasp the circled object in Fig. 2.7 (a). The number of stops was significantly fewer in the knowledge group than in the control group (Welch's t-test, $t(29) = 2.17$, $p = 0.04$). Furthermore, the participants in the control group tried to grasp the circled object in Fig. 2.7(a) five times under the condition $\theta < 60°$, but those in the knowledge group required 0 tries under the same conditions. These results proved that watching the proposed prior survey-knowledge view enhanced the operators' grasp of target objects from positions of large θ and avoided all stops, thus increasing the speed during grasping.

2) Reduction of movement distance with route perspective: Difficulties in recognizing the relationship between the heavy machines and the obstacles degraded the distance judgment. Fig. 2.16 shows the shortest path (green line) and a roundabout path. Participants in the knowledge group followed the white path in the 2nd set. Although the distance between the encircled obstacles in Fig. 2.16 was recognized from the questionnaire, the participants took the roundabout (white) path because they could not easily recognize the distance relationship between the

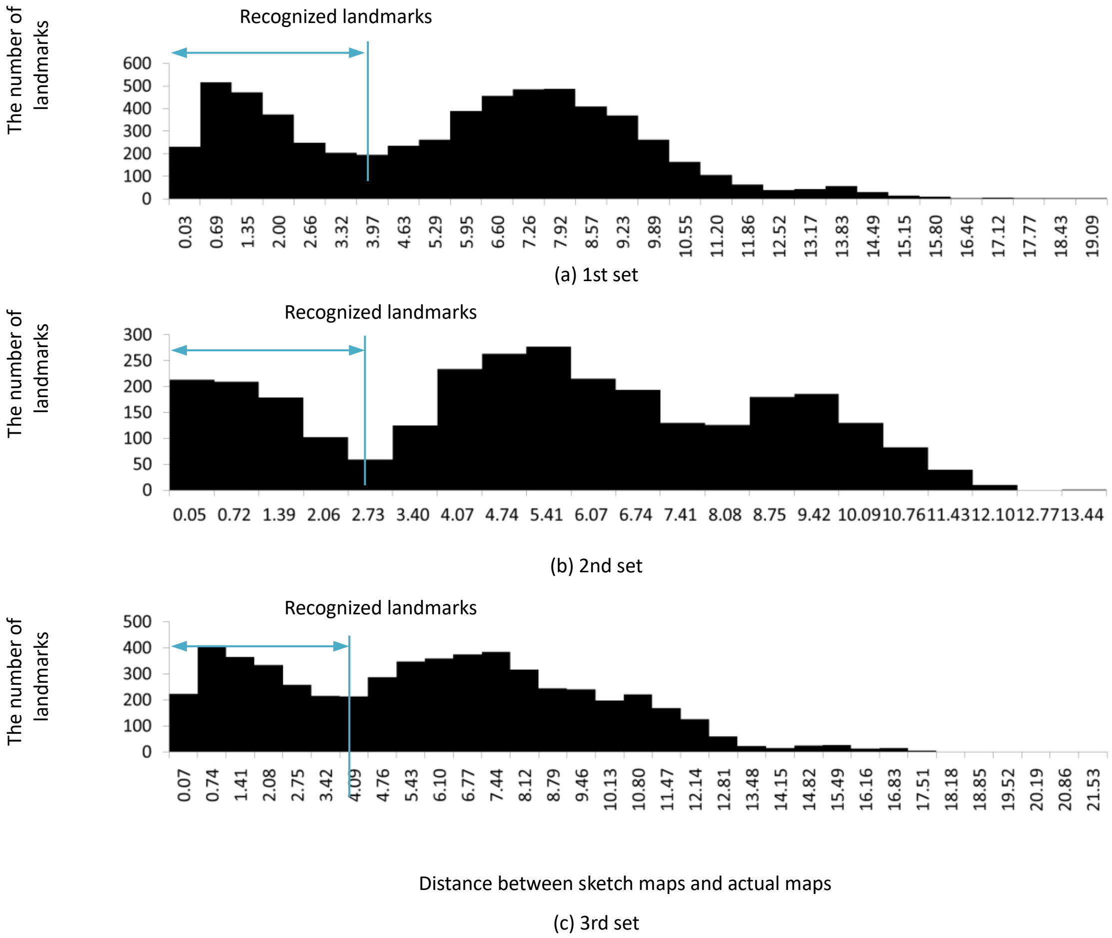


Fig. 2.17 Results of histograms of the distances between sketched and actual maps

obstacles and the heavy machines, which minimizes the movement distance.

2.3.4 Experimental results and discussion on cognitive maps

This section analyzes the results of the cognitive maps and relates them to the work performance.

2.3.4.1 Results of recognized landmarks

Fig. 2.17 shows the histograms of the distances in all sketched landmarks and the same landmarks in the actual maps of the three sets. All three histograms include at least two peaks,

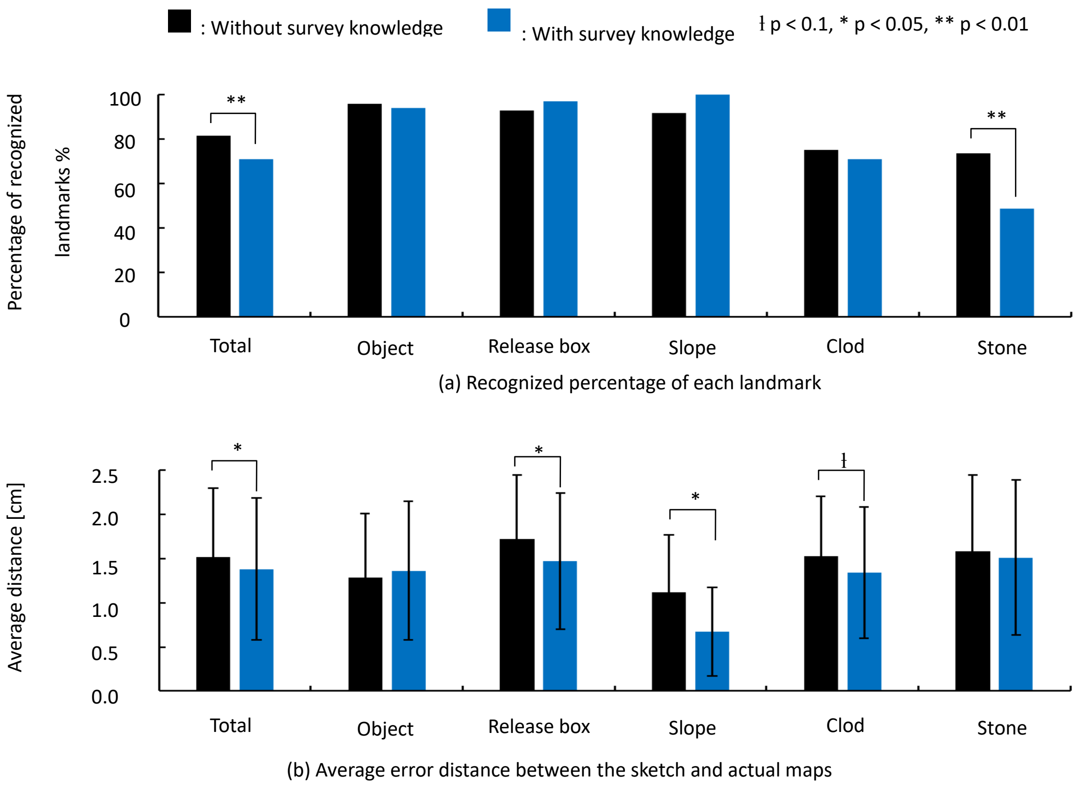


Fig. 2.18 Results of sketch maps for the 1st set (survey perspective)

and their shapes are similar to that of the hypothesized histogram in Fig. 2.10. Therefore, the intersections were set as the midpoints of the columns of local minima, and the thresholds were determined as the distances from these intersections.

2.3.4.2 Results and discussion of the survey-knowledge view (1st set)

Fig. 2.18 shows the percentages of recognized sketched landmarks in set 1 (i.e., the number of recognized sketched landmarks divided by the number of landmarks in the actual map). The average distance error between the recognized sketched landmarks and the actual map of each landmark, and the total of the first set, are also shown. The chi-squared test shows that participants in the control group recognized significantly more stones ($\chi^2(1) = 31.59$, $p < 0.001$) and total ($\chi^2(1) = 17.64$, $p < 0.001$) than participants in the knowledge group. However, the participants in the knowledge group recognized the release boxes, the slopes, and the total significantly more accurately than those in the control group (Welch's t-test, release boxes: $t(180) = 2.26$, $p = 0.02$, slopes: $t(39) = 2.55$, $p = 0.01$, total: $t(889) = 2.55$, $p = 0.01$). Moreover, the average distances between clods were marginally significantly different between the knowledge and control groups (Welch's t-test, $t(204) = 1.89$, $p = 0.06$). These results indicate that watching the proposed prior

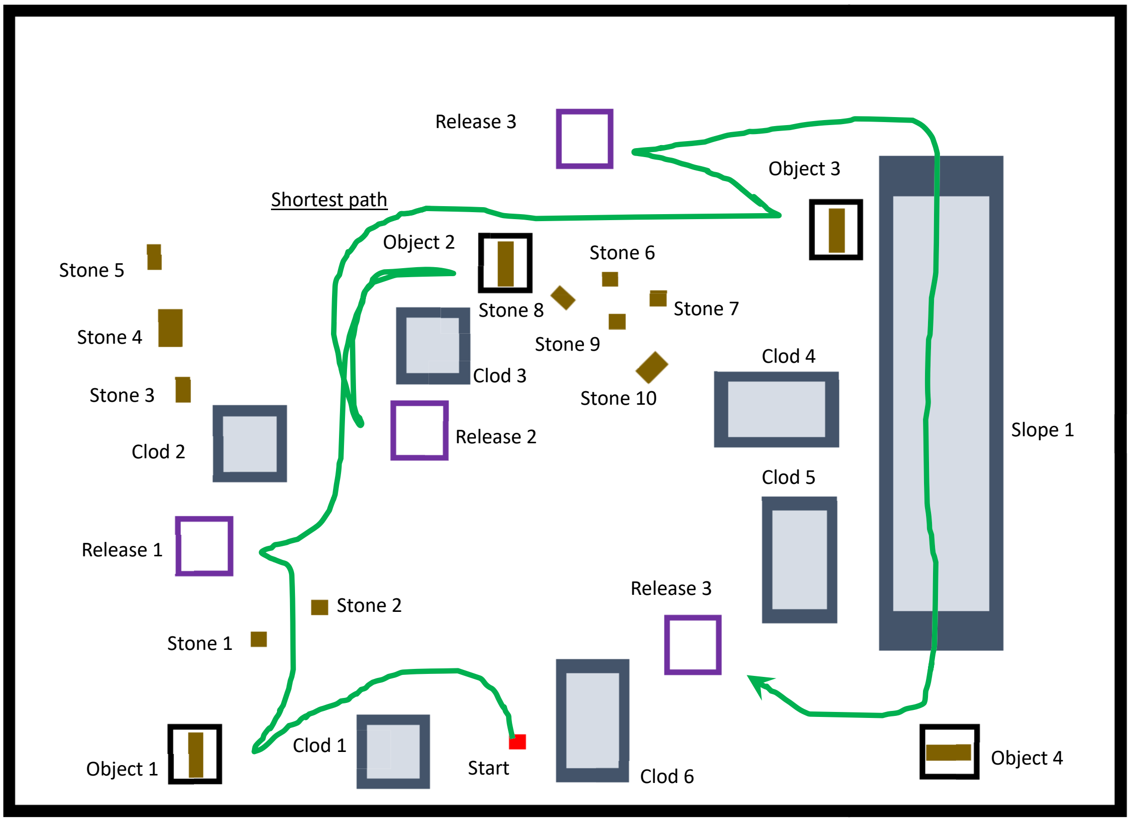


Fig. 2.19 Shortest path among all trials and the actual map of environment 1

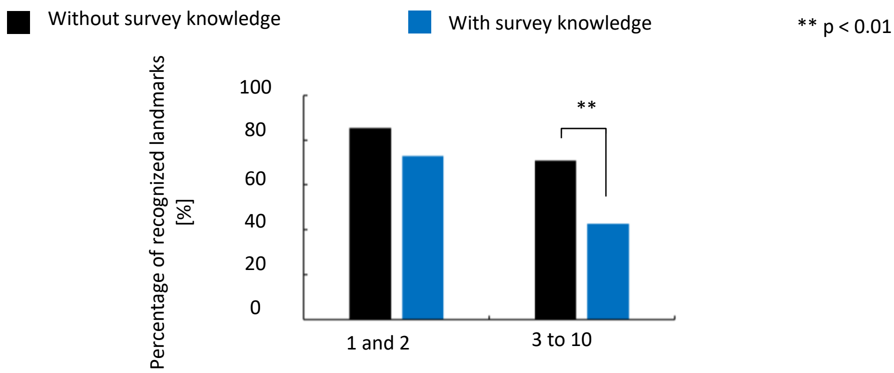


Fig. 2.20 Recognized percentages of stones (1 and 2) and (3 to 10)

survey-knowledge view can improve the accuracy of recall by the operators, but cannot improve the quantity of remembered objects.

The author now discusses: (1) the high percentage of objects recognized by the control group, and (2) the lack of improvement in the average error distance between objects and stones.

(1) The high recognition percentage of the control group would be explained if participants in the control group attempted to remember every landmark, whereas those in the knowledge group focused only on the important landmarks, such as the debris to be transported and the release boxes. Stones are less essential in path planning than release boxes or target objects because

they are small obstacles. Fig. 2.19 shows the shortest path (green line) on the actual map of Environment 1. Stones 1 and 2 are adjacent to the shortest path, but the other stones are distant from this path. Fig. 2.20 shows the recognized percentages of Stones 1 and 2 and Stones 3 to 10. The chi-squared test showed no significant difference between the knowledge group and the control group in recognizing Stones 1 and 2 ($\chi^2(1) = 2.27$, $p = 0.13$), but a significant difference in recognizing Stones 3 to 10 ($\chi^2(1) = 30.94$, $p < 0.001$). Furthermore, the number of Clod 4 recognized at a distance from the shortest path (Fig. 2.19) was 20 for the control group and, but 14 in the knowledge group. This difference was marginally significant by the chi-squared test ($\chi^2(1) = 3.65$, $p = 0.06$). Moreover, the results indicated that the participants in the knowledge group recognized most of the landmarks adjacent to the shortest path, such as the objects, the release boxes, and the slopes (see Fig. 2.19). Furthermore, the movement distance was significantly shorter in the knowledge group than in the control group (Section 2.3.2.1). These results prove that after watching the proposed prior survey-knowledge view, operators can plan short paths and focus on remembering the important landmarks adjacent to those short paths.

2) Lack of improvement in the average error distance of objects and stones: This result can be explained by the perceived importance of the landmarks. The objects were the most important landmarks in the experiment because the task involved grasping the objects. The results of the questionnaire proved that the participants in both groups attempted to remember objects. On the contrary, stones are less essential than the other substances, including the above-described objects. Therefore, their positions were largely ignored by the participants in both groups. Therefore, the lack of improvement in the average error distance of objects and stones was probably caused by the relative importances of the landmarks.

The results of this subsection prove that watching the proposed prior survey-knowledge view helps operators to accurately remember important landmarks, thereby decreasing the task time, number of stops, and movement distance, and increasing the speed during grasping.

2.3.4.3 Results and discussion of route-knowledge view (2nd set)

Fig. 2.21 shows the recognized percentages of the sketched landmarks and the average error distance between the sketched recognized landmarks and landmarks in the actual map in the 2nd set. A chi-squared test shows that the participants in the control group recognized significantly more clods ($\chi^2(1) = 6.43$, $p = 0.01$), stones ($\chi^2(1) = 6.07$, $p = 0.01$) and total landmarks ($\chi^2(1) = 10.93$, $p < 0.001$) than those in the knowledge group. A marginally significant difference was seen for release boxes ($\chi^2(1) = 3.57$, $p = 0.06$; Welch's t-test, $t(161) = 1.68$, $p = 0.10$). These results prove that watching the proposed prior route-knowledge view cannot effectively input either the quality or quantity of environmental information. However, the proposed route-knowledge view

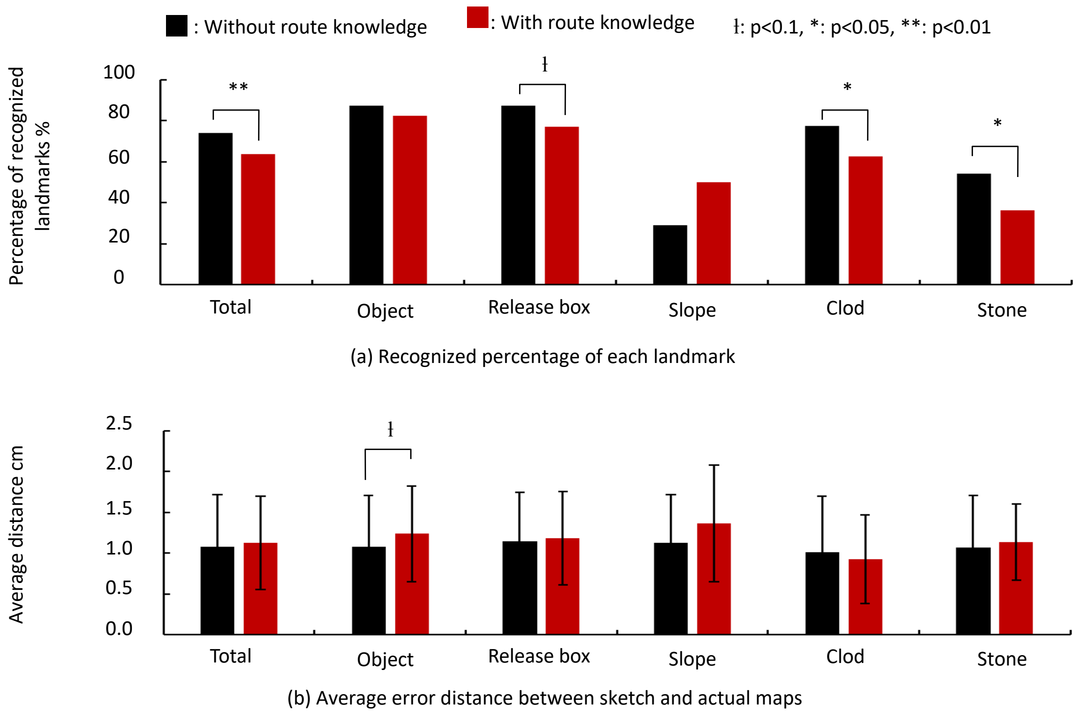


Fig. 2.21 Results of sketch maps in the 2nd set (route perspective)

significantly improved the task time, number of stops, and the grasping speed, as stated in Section 2.3.3.2.

It is surmised that participants in the knowledge group attempted to mentally build their work strategies (such as the way to grasp objects) rather than remembering the locations of various objects in the environment. Fig. 2.16 shows the shortest path in Environment 2, which proceeds clockwise. Thus, Environment 2 had limited complexity. The questionnaire results confirmed that the participants in the knowledge group tried to plan their work strategies, and were focused on planning rather than on remembering where the objects were located, because this environment was relatively simple. Therefore, watching the proposed prior route-knowledge view could not effectively input the environmental information. These results proved that watching the proposed prior route-knowledge view helped teleoperators plan their working strategies, thus improving the task time, number of stops, and the speed during grasping.

2.3.4.4 Results and discussion of survey- and route-knowledge view (3rd set)

Fig. 2.22 shows the recognized percentages of the sketched landmarks and the average error distance between the sketched recognized landmarks and landmarks in the actual map in the 3rd set. The chi-squared test indicates that participants in the knowledge group recognized

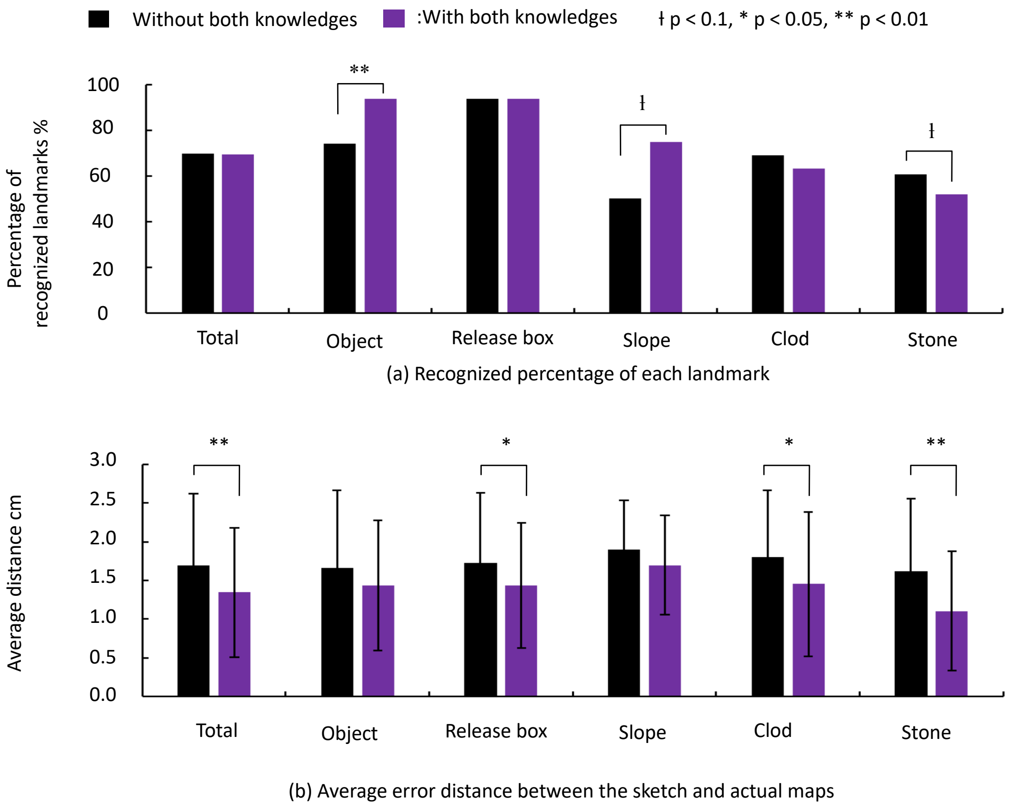


Fig. 2.22 Sketch map results of the 3rd set (both perspectives)

significantly more objects ($\chi^2(1) = 13.89$, $p < 0.001$), and marginally more slopes ($\chi^2(1) = 3.2$, $p = 0.07$), than those in the control group. However, the chi-squared test shows that participants in the control group recognized more stones ($\chi^2(1) = 3.74$, $p = 0.05$) than those in the knowledge group, and this result was marginally significant. Significant differences between the two groups were observed in the distances of the release boxes, the clods, the stones, and total landmarks (Welch's t-test, release boxes $t(176) = 2.24$, $p = 0.03$; clods $t(153) = 2.41$, $p = 0.02$; stones $t(269) = 4.99$, $p < 0.001$; and total $t(793) = 5.65$, $p < 0.001$). These results prove that watching the proposed prior survey- and route-knowledge view can effectively input both high-quality and high-quantity environmental information. These results are further discussed because they differ from the results of Sections 2.3.4.2 and 2.3.4.3.

These results might be explained by the greater difficulty of remembering Environment 3 than remembering Environment 1 by participants in the control group. Environment 1 included four objects, four release boxes, one slope, six clods, and ten stones. Environment 3 included four objects, four release boxes, one slope, five clods, and ten stones. The two environments differ only in their number of clods (one more clod in Environment 1 than in Environment 3). However, the percentage of recognized landmarks was 81% in Environment 1 (1st set, evaluated by participants in the control group) and ~70% in Environment 3 (3rd set). The chi-squared test shows that this difference was statistically significant ($\chi^2(1) = 21.27$, $p < 0.001$). Furthermore, the most important landmarks were objects because the participants were tasked with grasping the objects. However, the recognized percentage of objects in the 3rd set was 74% in the control group and 94% in the knowledge group, indicating a very significant difference between the two groups. These results suggest that the proposed prior survey-knowledge view can help teleoperators recognize important landmarks even in different environments, shortening their chosen paths and reducing the number of stops (as explained in Section 2.3.3.3).

2.3.4.5 Originality of the proposed prior view system

The proposed prior view system provides the environmental information of both survey and route perspective in advance. Although several systems provide the information during working [2.7–2.12], advance provision has not been reported. At actual unmanned construction sites, teleoperators walk the disaster sites and acquire route knowledge in advance, but this is difficult when the disaster sites are too dangerous for humans to enter. Teleoperators can watch survey views of the 2016 Kumamoto Earthquake because the Government released a 3D model of the disaster sites on the internet within 3 days after the earthquake. However, the advance provision of environmental information has not been investigated from a route perspective. This thesis highlights the importance of the route perspective, especially in manipulation.

2.3.4.6 Practical usability

The proposed survey and route views with 3D environmental information can be displayed in common computer graphics software such as Unity and Blender. 3D environmental information has been obtained in previous studies [2.33, 2.34]. Common computer graphics software can provide views from any viewpoint if the 3D environmental information is available. Therefore, the proposed survey and route views are available for practical use.

Although the experimental environments were quite different from actual unmanned construction sites, they confirmed the usefulness of providing the survey and route views to teleoperators. Using this information, teleoperators can plan the movement paths and their work strategies in tasks requiring long-distance movements such as laying compacted concrete, and in manipulation tasks such as earthworks and box culverts.

2.3.4.7 Effects on skilled operators

The targeted teleoperators in this thesis are skilled at controlling construction machinery, but the participants were all novices, so the author discusses it. Because the Japan Ministry of Land, Infrastructure, Transport, and Tourism promotes i-Construction, which includes modeling (measuring and visualizing) the worksites in 3D [2.35], skilled operators have more opportunity than novice operators to watch third-person views. This may mean that skilled operators can obtain environmental information more effectively than can novice operators from the proposed prior survey view, which also means that work efficiency can be improved more than in novice-operators cases. Skilled operators usually control construction machines on board. Thus, they can plan more effectively than novice operators from route perspective, which also means that skilled workers' advance review of routing can improve work efficiency more than novice operators.

2.4 Reminder system

This section develops an AR reminder system that helps teleoperators to remember their movement paths and trajectories right before each work commences.

2.4.1 Problems of forgetting the plan

When the plan is forgotten, the operators tend to choose oblique paths, stop the operation, and make erroneous contacts. For example, consider the environments shown in Fig. 2.23(a) and (b). If teleoperators forget their planned paths (black line in Fig. 2.23(a)), they may select a roundabout path (blue line in Fig. 2.23(a)). If the planned work strategies are forgotten during the teleoperation, the operators must also re-plan their work strategies to grasp the object over the tree (see Fig. 2.23(b)). Stops and error contacts with obstacles may result from the excess cognitive load of planning and working simultaneously. Therefore, a reminder system can increase the work efficiency and lower the cognitive load.

Previous studies handled the above-mentioned problems by displaying additional information such as paths and trajectories. For instance, some AR-based navigation systems have been devolved for showing the movement paths of cars and pedestrians [2.11, 2.36]. These AR navigation systems can also display vessels and internal organs to surgeons [2.37, 2.38]. Other AR navigation systems show animations and ghosts mimicking real-time movements for manipulation tasks [2.39, 2.40].

However, these previous systems do not consider the work-state changes such as moving, grasping, and cognitive loads. Unmanned construction work requires changes in work states. For example, the removing task, one of the most important tasks in disaster response, requires movement toward the target object. The object needs to be reached, grasped, transported (to the

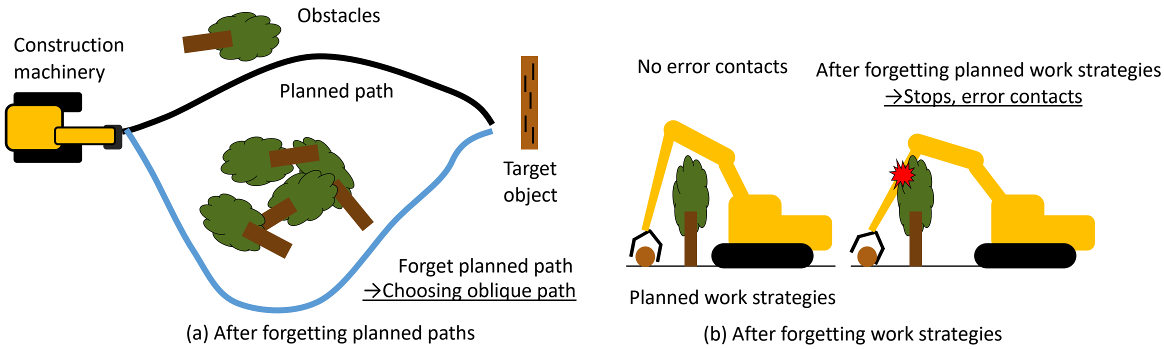


Fig. 2.23 Decrease of work efficiency by forgetting the plan

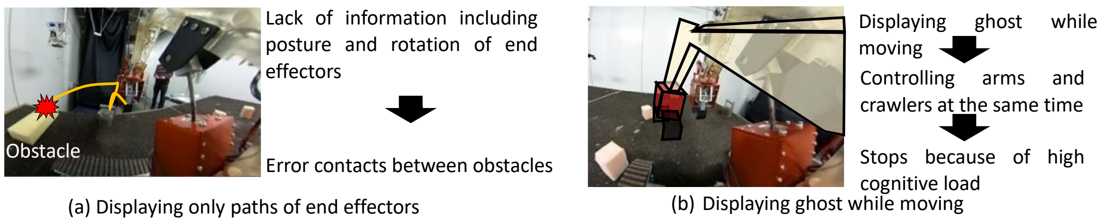
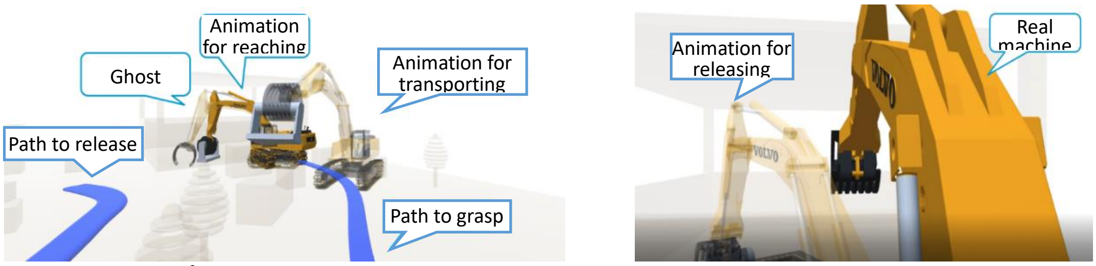


Fig. 2.24 Problems in recognizing the same AR in different work states

designated area), and released. Moreover, some unmanned construction tasks such as earthworks require changes in work states. Thus, a reminder system should adapt to changes in the work state by altering its AR. Displaying only the suitable AR paths for moving during grasping can hardly provide the important information of grasping, such as rotation of the end effectors and arm postures. The incomplete information can lead to error contacts with obstacles during grasping, as shown in Fig. 2.24(a). Moreover, displaying a ghost, which is suitable AR for manipulation tasks, can provide excess information during rotational movements of end effectors and arm postures. This problem increases the cognitive load of teleoperators, causing them to pause when controlling the arms and end effectors during movement (see Fig. 2.24(b)).

Excessive cognitive load can lead to several serious problems, such as focusing only on specific views or areas while ignoring other information [2.41], and mental fatigue [2.42]. Therefore, cognitive load is very important for teleoperators who must select suitable views depending on the work states [2.22], and who maneuver machines approximately eight hours every day. When presented with all information all the time, teleoperators can miss important information for specific work states, as shown in Fig. 2.25 (a). Displaying suitable information immediately before the task can also cause excess cognitive load (Fig. 2.25 (b)), because teleoperators sometimes pay little attention to additional information due to excessive cognitive loads by some precise teleoperation tasks including manipulation.

Thus, this study aims to develop a planning reminder system that adapts to changes in the work states. This reminder system involves two functions: a reminder function and a function that



Excess information
・ Missing suitable information

(a) All additional information displayed at the same time

Excess cognitive load caused by both working and watching the additional displayed information
・ Stops to watch the additional information
・ Missing additional information when focusing on work

(b) Displaying additional information right before the work.

Fig. 2.25 Problems caused by high cognitive load

identifies work states. This paper focuses on the reminder function because work states have been already distinguished in previous studies [2.43, 2.44].

2.4.2 Development of a reminder system

This section develops a plan reminder system that adapts to changes in the work states and maintains a low cognitive load, as explained in Section 2.4.1.

2.4.2.1 Adapting to changes in work states

Appropriate information, which is different from the work states, must be provided to teleoperators, as explained in Section 2.4.1. This appropriate information is displayed by AR, which allows teleoperators to watch both the additional information and the real environments [2.45].

1) Moving toward the target: This movement approaches an area close to the target object, and teleoperators often control only crawlers. As movement paths include the essential information for movement, they have been used to increase user interactions in various fields, such as pedestrian and car navigations [2.15, 2.36]. Thus, fig. 2.26(a) displays a machine movement path in the proposed reminder system.

2) Reaching and transportation: Reaching brings the end-effectors close to the target objects, and transportation brings the end-effectors close to the designated release area. The only difference between these two work states is the grasping of the target object during transportation; accordingly, similar information can be displayed in both states. The motion of the arms plays an important role in reaching and transportation because teleoperators control only the arms. Arm motions are commonly displayed in sports and instrument teaching [2.46, 2.47]. Therefore, the

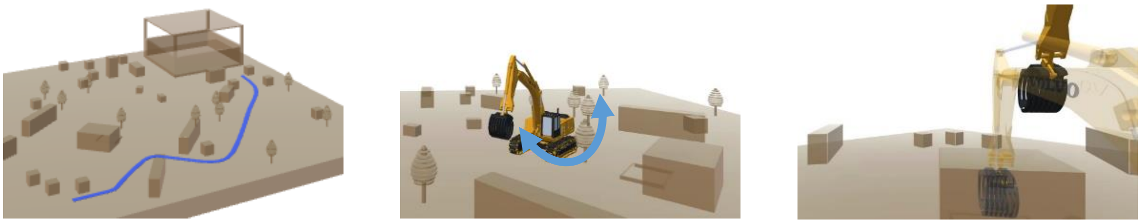


Fig. 2.26 AR for each work state

proposed reminder system also displays arm motions, as shown in Fig. 2.26(b).

3) Grasping and releasing: Grasping holds the target objects, and releasing detaches the grasped target objects. These two work states differ only by the end-effector motions (closing in grasping, opening in releasing), which implies similar appropriate information in both states. Comprehending the positional relationship between the current and goal situations, including the postures and positions of the end-effectors, is essential for these two work states. In situations requiring the assembly of instruction systems, the goal postures and positions are displayed [2.40, 2.48]. Accordingly, the proposed reminder system also displays the goal postures and positions of the end effectors, as shown in Fig. 2.26(c).

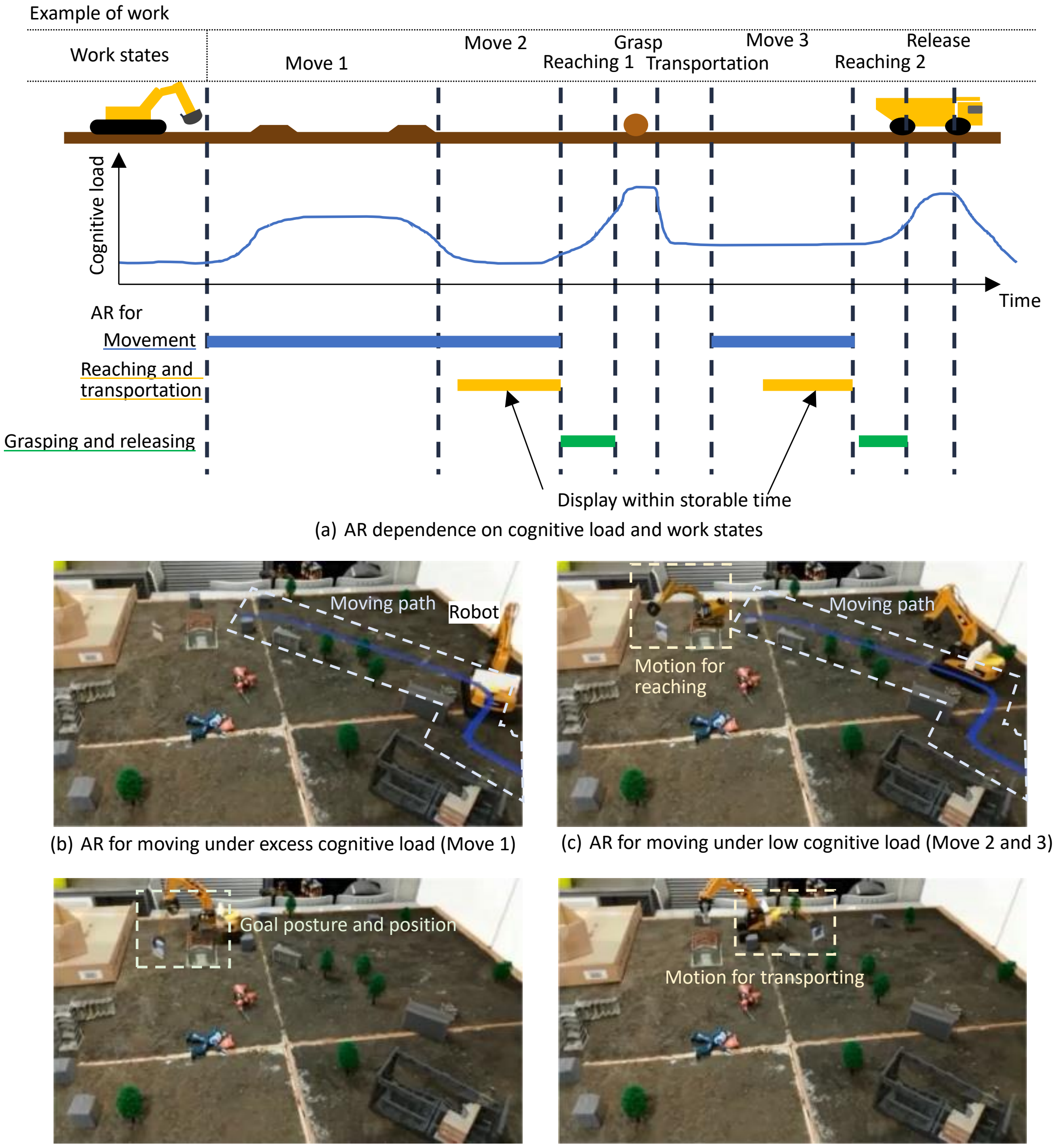
2.4.2.2 Low cognitive load

The displayed information should consider the cognitive load because teleoperators can suffer serious problems under excess cognitive loads, as described in Section 2.4.1.

1) Amount of displayed information: If all information is displayed all the time, the appropriate information might be buried in the information overload, as described in Section 2.4.1. The proposed reminder system determines the amount of displayed information, relieving teleoperators from remembering much information during the teleoperation work. This is effective because human memory is limited [2.25]. Furthermore, the cognitive load increases with the complexity of the task, demanding more working memory [2.47]. Therefore, the amount of displayed information should be decreased when the teleoperators perform complex tasks, including grasping and releasing.

2) Time to display: Teleoperators hardly remember information that is displayed too early for retention in their short-term memory [2.50]. Furthermore, teleoperators' ability to remember information is degraded under excessive cognitive load [2.51]. Therefore, the proposed reminder system displays the AR information within a memory-storable time, maintaining the teleoperators under low cognitive loads.

3) Example of the proposed system: An example of the proposed reminder system is shown in



(a) AR dependence on cognitive load and work states

(b) AR for moving under excess cognitive load (Move 1)

(c) AR for moving under low cognitive load (Move 2 and 3)

(d) AR for reaching

(e) AR for transportation

Fig. 2.27 Example of AR adapting to work states and cognitive load.

Fig. 2.27. The machine movement paths (Fig. 2.27(b)) are always displayed during the movement, reminding the teleoperator of the paths to follow. Reaching (see Fig. 2.27(c)) is also displayed within a storable time when teleoperators are under low cognitive loads, such as during moves 2 and 3 in Fig. 2.27(a). The goal postures and positions of end-effectors (e.g., Fig. 2.27(d)) are displayed during reaching for two reasons: first, to decrease the excessive cognitive load during manipulation tasks, which require control of four degrees of freedom (whereas moving tasks require only two degrees of freedom); second, teleoperators might have already watched the reaching action during the movement. Nothing is displayed during grasping and releasing because those manipulation tasks require precision and a high cognitive load. Transportation motions are

displayed at the beginning of the transportation task (Fig. 2.27(e)), as the cognitive load of transportation is lower than that of reaching.

2.4.3 Experiment

The author conducted experiments to verify the proposed reminder system in a scale model environment because a scale model enables physical experiments in a feasible setting (see Fig. 2.28). The participants were asked to teleoperate a robot by controlling two levers for crawlers and eight buttons for the arm. Three views—a cab view and two external views—were displayed

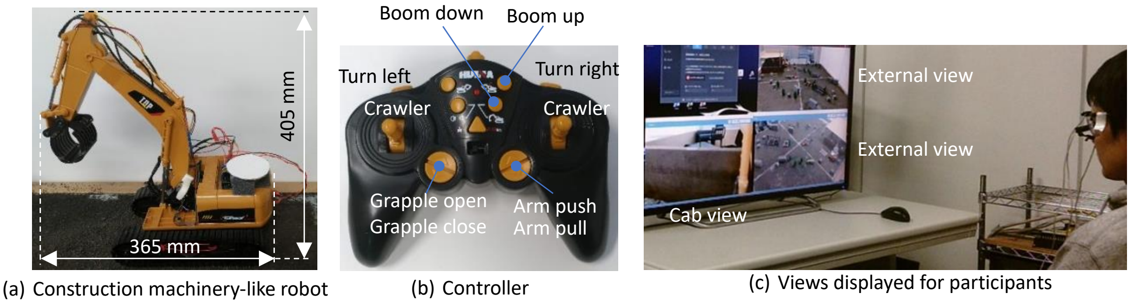


Fig. 2.28 Scale model and the interface

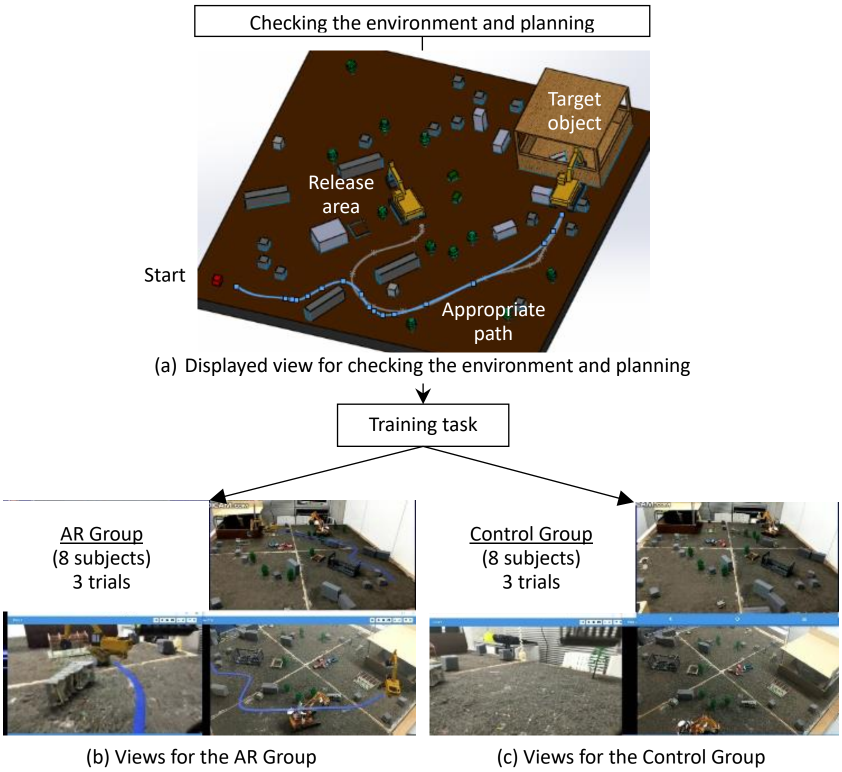


Fig. 2.29 Experimental procedure

Fig. 2.30 Experimental environment

to all participants, and the AR was displayed in ARCore software.

Fig. 2.29 shows the experimental procedure. Sixteen novice participants (students) with no experience in the teleoperation of construction machineries were invited to participate because only 20 skilled teleoperators reside in Japan (20) [2.28]. Fig. 2.30 shows the environment. The participants were prompted to move to the target object without contacting any obstacle. They were asked to grasp the object, transport it, and release it to the designated release box. For up to 10 min, all participants preliminarily watched the appropriate movement paths and the stop positions for the grasping and releasing of objects (see Fig. 2.29(a)); here, the paths and positions were determined by the author. Next, all participants were trained until they had acquired enough skills to teleoperate the robot. The 16 participants were then divided into two groups (the AR group and the control group). The groups were evenly divided (with eight participants each) to ensure that their average task times in the training sessions were sufficiently close to verify the proposed reminder system. The eight participants in the AR Group tried three tasks with the AR displayed in all three views, as shown in Fig. 2.29(b). The eight participants in the control group tried the same tasks without AR, as shown in Fig. 2.29(c). The displayed AR, which replicated the movement paths and the stop positions for grasping and releasing, was also determined by the author and was changed for each work state. As the AR switcher, the author identified each work state and changed the displayed AR by clicking the buttons for the eight participants in the AR Group. The task time, mental workload (measured by NASA-TLX [2.51]), and the error contacts between the machinery parts and obstacles were measured throughout the tasks. The experimental procedures were approved by the ethics committee for human research at Waseda University.

2.4.4 Results and discussion

2.4.4.1 Task time

Fig. 2.31 shows the task times of the total work and each work state in the first trial. Significant differences between the groups were observed for the total work and the moving, reaching, and transportation tasks (Mann–Whitney U test, total: $U = 0$, $p < .001$, $r = .84$; moving $U = 3$, $p = 0.002$, $r = 0.76$; reaching and transportation time $U = 3$, $p = 0.002$, $r = 0.76$). These results indicate that the proposed reminder system can increase the total work efficiency by improving the efficiency of the machine movement and the way in which the machine reaches and transports the target. However, there were no significant differences in the grasping and releasing times (Mann–Whitney U test, $U = 27$, $p = 0.60$, $r = 0.13$), possibly because the stop positions differed in the AR and the real robot. The goal situations (including the stop positions for grasping) were displayed to the AR Group by the proposed reminder system, but the AR did not exactly replicate the real situations. Fig. 2.32 shows a typical cab view during the grasping process. The differences between the AR and real situations could negatively affect precise manipulation tasks, such as grasping and releasing, and might explain the lack of any improvement in the grasping and

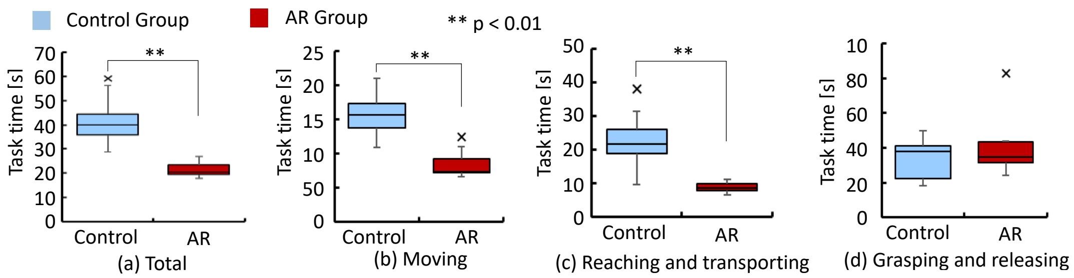


Fig. 2.31 Results of task times in the first trial

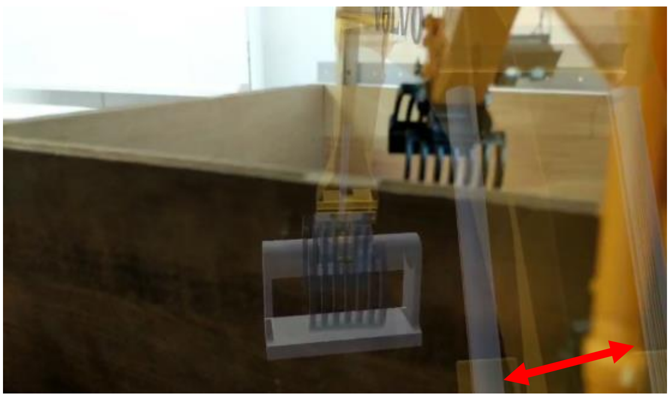


Fig. 2.32 Error between AR and the real robot

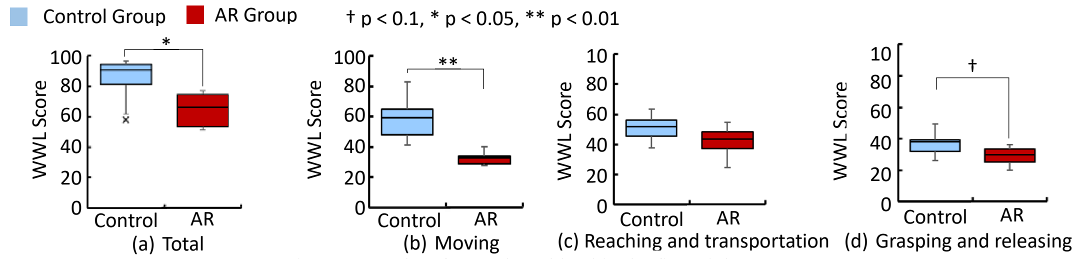


Fig.2.33 Results of mental workload in the first trial

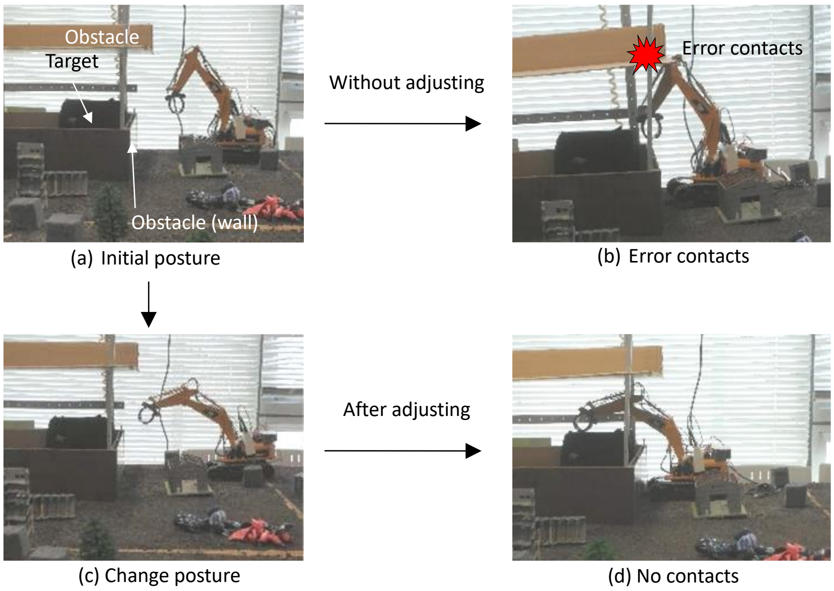


Fig. 2.34 Error contacts between the machinery parts and walls

releasing times. Therefore, the AR for grasping and releasing needed to be adapted to real situations based on the machine positions.

2.4.4.2 Cognitive load

Fig. 2.33 shows the mental workloads of the total work and each work state in the first trial. Significant differences between the two groups were observed in the total and moving work states, and a marginal significance was observed in the grasping and releasing states (Mann–Whitney U test, total: $U = 10, p = 0.02, r = 0.58$; moving: $U = 0, p < 0.001, r = 0.84$; grasping and releasing: $U = 14, p = 0.06, r = 0.47$). These results prove that the proposed reminder system can downgrade the workload of the total operation and the workloads of moving, grasping and releasing. No significant differences in the machine's reaching and transportation actions were observed between the groups (Mann–Whitney U test, $U = 16.5, p = 0.10, r = 0.41$).

The lack of improvement in the AR group might be explained by participants missing the AR for reaching and transportation. To avoid erroneous wall contacts, the teleoperators needed to adjust the arm height before approaching the target-object area (see Fig. 2.34). However, only two

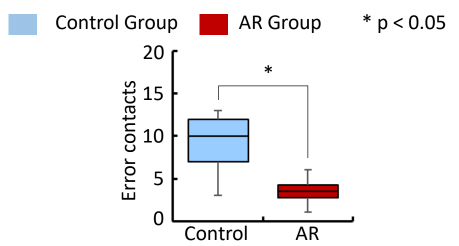


Fig. 2.35 Results of error contacts in the first trial

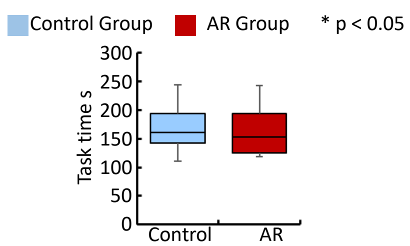


Fig. 2.36 Results of task time in the third trial

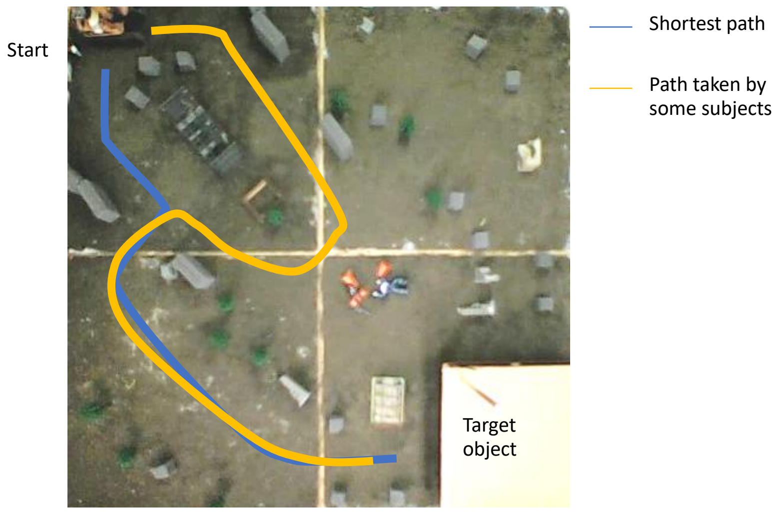


Fig. 2.37 Results of movement paths in the third trial

out of eight participants adjusted the arm height in the AR Group. The results from the interviews showed that the participants in the AR Group did not recognize the AR for reaching and transportation because the AR motion occurred at high speed and made a sudden appearance in the AR. The error contacts were made by the participants who missed the AR could explain the lack of improvement in the mental workload.

2.4.4.3 Error contacts

Fig. 2.35 shows the number of error contacts in the first trial. A significant difference was observed between the two groups (Mann–Whitney U test, $U = 8$, $p = 0.01$, $r = 0.63$). This result suggests that the proposed reminder system can reduce the number of error contacts.

2.4.4.4 Discussion 1 ~Effects of the reminder system on repetitious tasks~

Here, the author discusses the effects of the proposed reminder system on repetitious tasks. Although the proposed reminder system increased the work efficiency of the first trial, unmanned construction work includes some repetitious tasks whose effects need to be discussed.

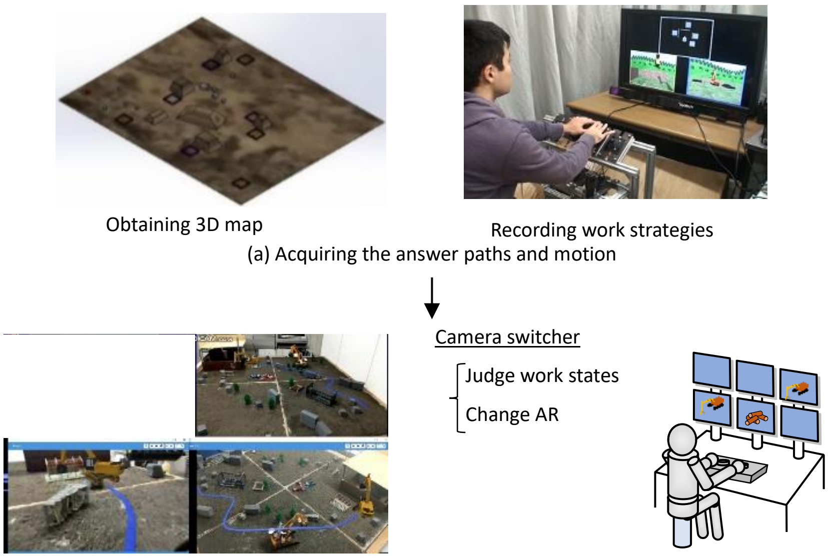


Fig. 2.38 Flow of practical use of the proposed reminder system

Fig. 2.36 shows the results of the total task times in the third trial. No significant differences between the control and AR groups were observed in the Mann–Whitney U test ($U = 28$, $p = 0.67$, $r = 0.11$), possibly because the participants in the control group could plan the movement paths and work strategies of two repetitious tasks. However, some participants in the control group planned roundabout paths during repetitious tasks, as shown in Fig. 2.37. These results suggest that the reminder system will not assist teleoperators who can plan easily during repetitious tasks, but can be helpful for teleoperators who cannot plan during operations.

2.4.4.5 Discussion 2 ~Practical use of an AR reminder system~

The AR reminder system can be used in two ways: to obtain the appropriate paths and motions, and to change the AR depending on the work state. Experiments were conducted using the answer movement paths and motions determined by the author, who identified each work state and changed the AR accordingly. These roles are expected to typify practical applications (see Fig. 2.38).

1) Ways of acquiring appropriate movement paths and motions: Two kinds of information are necessary for obtaining the appropriate movement paths and motions: 3D environmental information and the work strategies recorded by a simulator (see Fig. 2.38(a)). The 3D environmental information at disaster sites was taken from the literature [2.33, 2.34]. The teleoperators can record their appropriate movement paths and motions via a simulator such as [2.27], as automatically calculating the appropriate movement paths and motions is technically difficult.

2) Ways of changing the AR in each work state: Some unmanned construction operations have camera switchers who modify the pan, tilt, and zoom of the external views for high-efficiency operations. Here, the AR could be modified at work sites by employing camera switchers who identify each work state and change the AR accordingly, as shown in Fig. 2.38 (b). However, as this solution could increase the number of tasks performed by camera switchers, an identification system for work states is required [2.43, 2.43].

2.5 Summary

The author proposed a prior view system that provides environmental information based on human spatial cognition characteristics. The system displays two views: an external view from any viewpoint to acquire the survey perspective, and a view from the operator's viewpoint that can be changed by the teleoperator to acquire a proper route perspective. The proposed prior view system was verified in simulator experiments. The experimental results indicated that the proposed prior view system improves the work time and enhances the teleoperator's planning process. Watching the survey-knowledge view reduced the task time, the movement distance, and the number of stops. Moreover, watching the route-knowledge view improved the task time, the number of stops, and the grasping speed. After analyzing the cognitive maps, it was found that watching the survey-knowledge view helped teleoperators to memorize important landmarks such as the target objects and release boxes. This recognition may have improved the task time, the movement distance, and the number of stops. The analysis results of cognitive maps further suggested that watching the route-knowledge view helped teleoperators to plan the manipulation of tasks, thereby reducing the number of stops and enhancing the grasping speed.

The author also proposed a reminder system that adjusts to the changing work states and reduces the cognitive load of teleoperators. In an AR display, the proposed reminder system provides the appropriate information depending on each work state. In particular, it displays the movement paths of the machine, the arm motion for reaching and transportation, and the goal postures and positions of the end-effectors for grasping and releasing the target object. The amount of displayed information was downgraded during complex manipulation tasks such as grasping and releasing to avoid cognitive overload of the teleoperator's working memory. Moreover, the proposed reminder system displays AR within a memory-storable time while teleoperators are working under low cognitive loads. The experimental results of a scale model indicated that the task time, the mental workload, and number of error contacts were improved by the proposed system.

Chapter 3: Optimal and Allowable Position for Camera Placement

This chapter investigates the optimal and allowable position for camera placement. First, the author describes the necessity of having external views and the importance of deriving optimal and allowable locations for camera placement. Next, the cameral placement is investigated in scale-model experiments performed by novice participants. An actual machine is also used by novices and skilled operators. Finally, the author discusses applications of the obtained results.

Some of the sentences, figures and tables in this chapter are borrowed from the author's works [3.1-3.6].

3.1 Necessity of external views and optimal location for camera placement

This section describes the necessity of external views and the importance of deriving an optimal and allowable location for camera placement.

External views captured from cameras introduced at disaster sites are essential, especially during manipulation tasks requiring depth perception, such as digging and releasing; such views are required even when a 3D cab view with a wide field of view is provided to the teleoperators [3.7]. External views are usually captured by camera dollies at disaster sites such as Unzen–Fugendake. Furthermore, external views can be acquired by image processing or by drones [3.8, 3.9]. However, no research has derived an optimal position for camera placement in unmanned construction work. Teleoperators can work efficiently by switching their gaze views to correspond with their work states [3.10]. An optimal camera placement system corresponding to each work state can improve the work efficiency. However, the optimal placement of cameras is difficult in the extreme environments of disaster sites, such as heavy rains and steep slopes. Therefore, determining an allowable location for camera placement is essential for proper investigations.

3.2 Parameters of camera placement and human object recognition

Camera placement is defined by four parameters. After describing these parameters, the author hypothesizes an optimal and allowable camera placement based on human object-recognition characteristics.

3.2.1 Parameters of camera placement

This subsection defines the four parameters of camera placement: the view targets, the pan angle (φ), the tilt angle (θ), and the distance (r), as illustrated in Fig. 3.1.

3.2.2 Human object recognition

The highest performance in object recognition is achieved by watching the objects from the canonical view with the fewest occlusions [3.5]. For instance, the object's shape is easily recognized from the canonical view (Fig. 3.2 (a)). In contrast, humans can barely discern the object from a noncanonical view (see Fig. 3.2 (b)). The characteristics of canonical views are as follows:

A) They have the fewest occlusions [3.11].

B) In the canonical view, humans can recognize an object with the same performance when the view angle changes by less than 30° [3.12].

C) The object sizes in canonical views have little effect on object recognition [3.12].

3.2.3 Hypothesis of optimal and allowable camera placement

Precise manipulation tasks, such as digging and releasing, require teleoperators to discern the 3D positional relationship between the end-effectors and the target object or the designated release area. In a canonical view, teleoperators can easily discern this relationship.

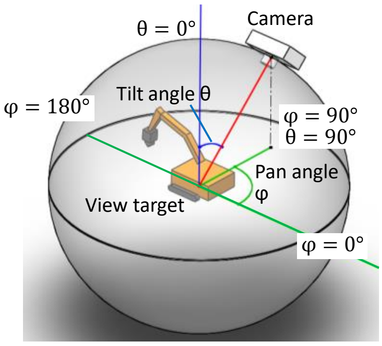
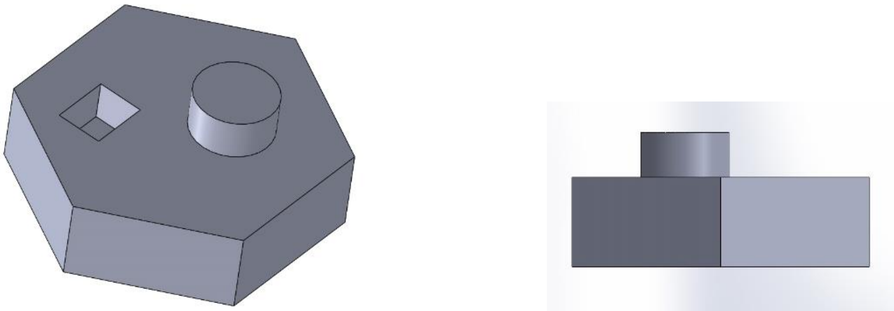


Fig. 3.1 Four Camera placement parameters: view target, pan angle, tilt angle, and distance



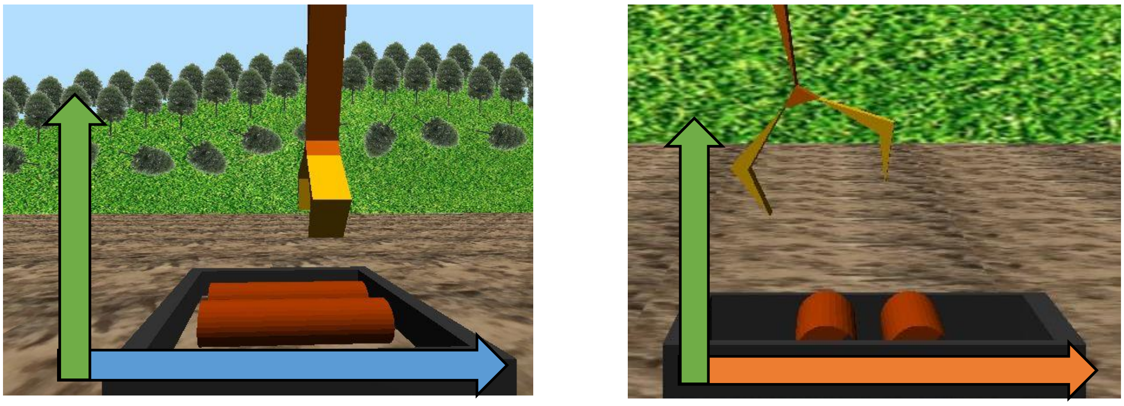
(a) Canonical view

(b) Non-canonical view

Fig. 3.2 Example of a (a) canonical and (b) non-canonical view

Manipulation tasks, such as digging and releasing in unmanned constructions, require precision. For instance, teleoperators are required to grasp 100-mm-size debris and place the debris approximately 9 m ahead of a cockpit using an ordinary 0.8m^3 construction machine. For this purpose, they must discern the 3D (X, Y, and Z axes) positional relationships between the end-effectors and the target object. Teleoperators are always provided with a cab view, which is similar to the view that operators watch during boarding operations, so is more heavily relied on than external views [3.13]. A cab view helps teleoperators to discern the positional relationship in the X and Z axes (as shown in Fig. 3.3(a)) because this view is the canonical view with the fewest occlusions along the X and Z axes. However, a cab view is not suitable for discerning the positional relationship along the Y axis. Thus, an external view is required to help teleoperators to discern the Y axis. Therefore, the canonical-view parameters required for recognizing the Y axis are described next.

1) Viewpoint targets: The views displayed during the manipulation tasks should include the target objects (or the designated releasing area) and the end-effectors, because teleoperators are required to discern the positional relationship between these two object types. Views from the moving cameras can cause difficulties during teleoperation [3.14]. The target objects and the designated release areas are close to the end-effectors during the manipulation tasks. Therefore, the author determined that the target objects or the designated release area will be viewed during



(a) Cab view (b) External view along the Y axis

Fig. 3.3 Analysis of optimal external view

these tasks.

2) Pan angle (φ) and tilt angle (θ): As mentioned above, a canonical view is the view with the fewest occlusions. Thus, a side external view with pan and tilt angles of 90° (see Fig. 3.3(b)), can be a canonical view in manipulation tasks. The target objects can be discerned as those observed on the canonical views if the absolute value of the rotation changes by less than 30° [3.12]. Thus, when the camera is optimally placed, the work efficiency of the canonical views is maintained if the pan angle is 90° ± 30° and the tilt angle range is 90°–30°. The author defines the allowable camera placement as the conditions in which statistical comparisons with the fastest task time yield p values within 0.1 in this study.

3) *Distance (r)*: The distance (r) increases with increasing view target and vice versa. Therefore, r can be the same as the zoom level. The discerning of the target object hardly depends on the object size, and is mostly done by rotation angles [3.6]. Therefore, object recognition in the present study is mediated by rotation angles and the zoom level is ignored.

3.3 Experiments by novice operators using a scale model

Two experiments (Experiment 1 and Experiment 2) were performed on a scale model to derive the optimal and allowable pan and tilt angles in the manipulation tasks. The scale model was chosen because tasks are easily repeated and the parameters are easily changed in a real physical environment. Another consideration is that distances between objects are underestimated in simulators [3.15, 3.16], and therefore, results may differ between simulations and scale models. Also, manipulation tasks require depth perception, which means the participants need to discern

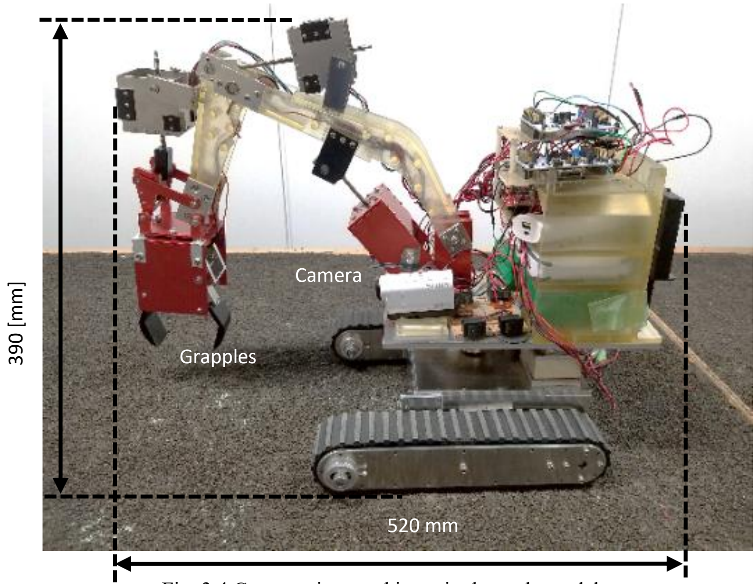


Fig. 3.4 Construction machinery in the scale model

Fig. 3.5 Control interface of the scale model

distance between the end-effectors and the target objects. Moreover, from previous experiments, the author found that participants using simulators tended toward rapid control because no physical elements that could be broken were involved. In contrast, they tended to control scale models slowly because physical models can be broken. The procedures of all experiments in this chapter were approved by the ethics committee for human research at Waseda University.

3.3.1 Experiments

Fig. 3.4 illustrates the 1/20 scale model, which simulates unmanned construction including the teleoperation of construction machinery. Fig. 3.5 shows the interface of this scale model. The scale model was developed through the law of similarity [3.17]. Thus, the speed of the scale model v' was computed as

$$v' = \sqrt{\frac{l}{l'}}v, \tag{1}$$

where v is the speed of the actual construction machinery, l is the length of the actual construction machine, and l' is the length of the scale model.

In the scale model, the participants were asked to teleoperate the construction machine in the scale model by using two manipulators for the arm and two levers for the crawlers, while watching a cab view and two fixed external views displayed on a 42-inch monitor. The participants were prompted to perform the tasks as quickly as possible.

The experimental tasks were determined by the model tasks of unmanned construction [3.17], and were identical in Experiment 1 and Experiment 2. Specifically, they involved moving to the target object, grasping and transporting the object, and releasing it at the designated release area (see Fig. 3.6). On the interface screen (Fig. 3.5), the external views at the top left and top right guided the grasping and releasing tasks, respectively. The author measured the grasping time

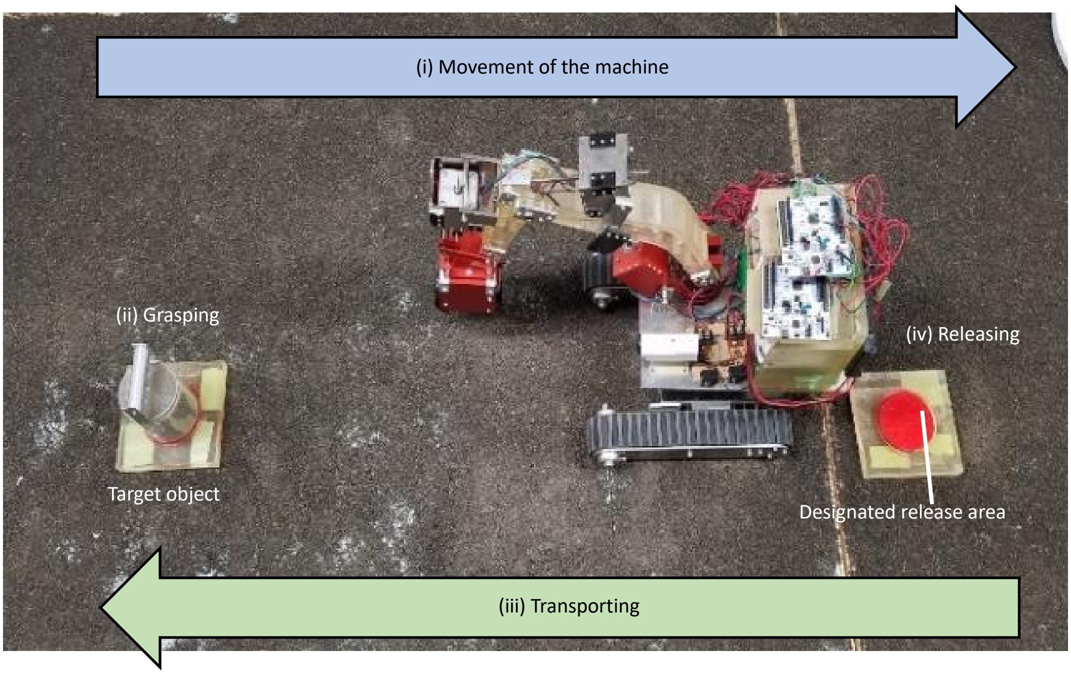


Fig. 3.6 Experimental task

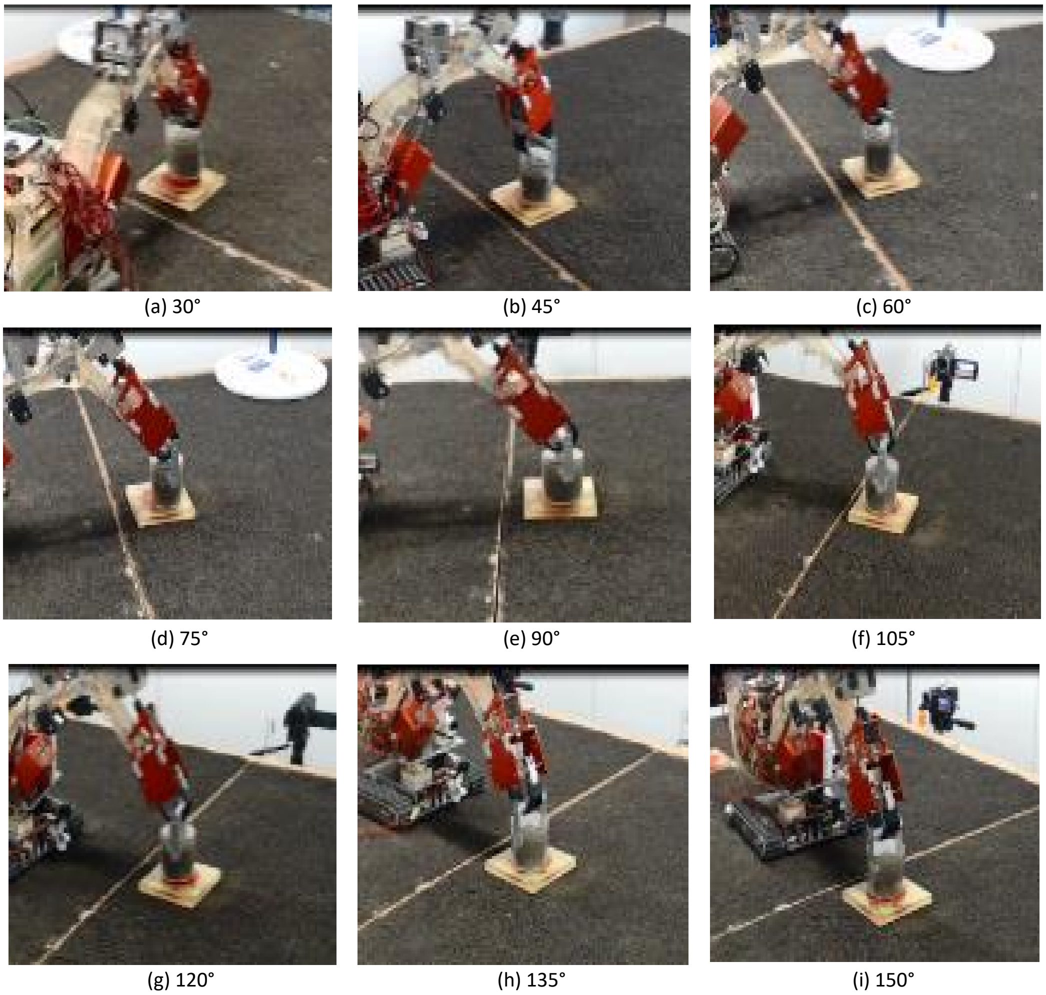


Fig. 3.7 External views at different pan angles

when the target object and the end-effectors were separated by less than 40 mm, and the releasing time when the designated release area and the end-effectors were separated by less than 40 mm. The error distances between the designated and actual release areas were also measured.

3.3.1.1 Experiment 1: pan angle experiment

1) Experimental settings: In Experiment 1, the author derived the optimal and allowable pan angle. Eight novice participants (students) with no experience in operating construction machinery were invited because only 20 skilled teleoperators reside in Japan (20) [3.17]. The participants acquired sufficient skills to teleoperate the construction machine of the scale model through training tasks prior to Experiment 1.

The pan angle was ranged from 30° to 150° in 15°-increments (a total of nine experimental conditions). Fig. 3.7 shows the nine external views captured at the nine pan angles. First, the participants performed the experimental tasks thrice without external views for normalization, as

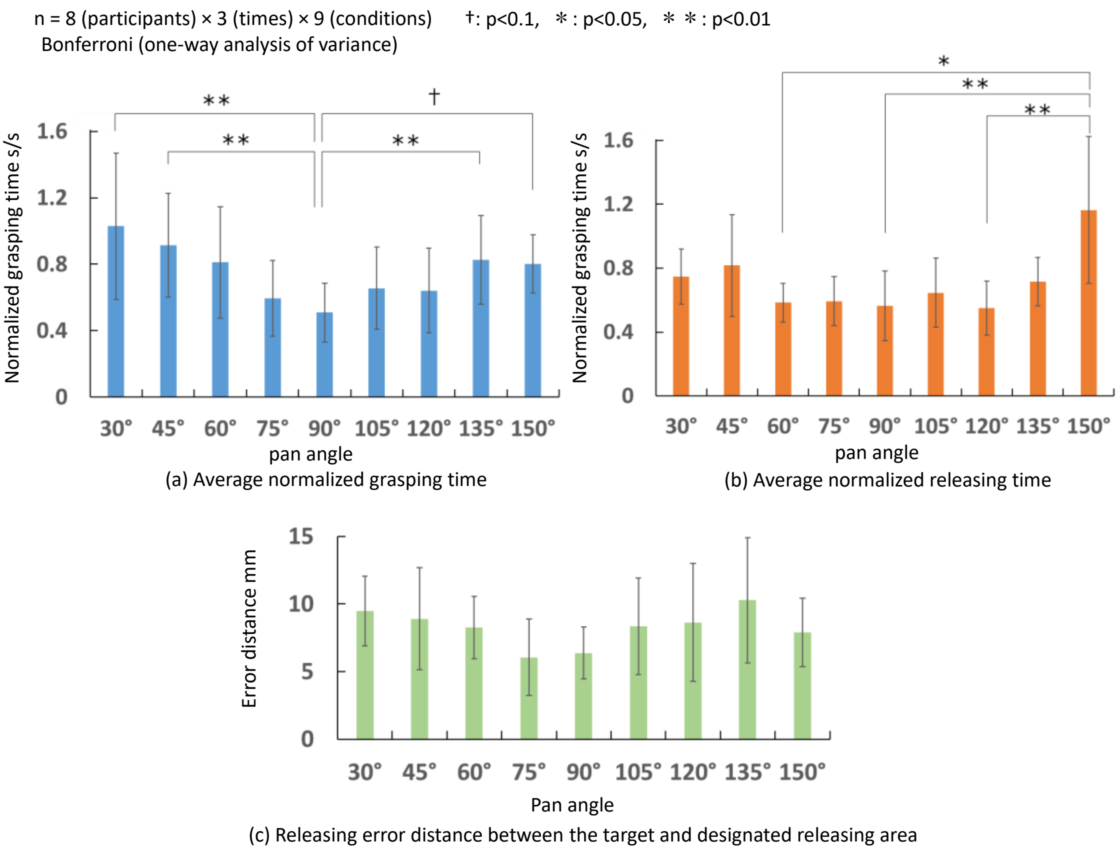


Fig. 3.8 Results of Experiment 1

explained in the Results section. Next, the participants tried the tasks thrice under the nine conditions. As shown in Fig. 3.7, the conditions were randomly ordered with one cab view and two external views.

The other experimental conditions were constant under all conditions and were set as follows: tilt angle = 60°, view angle = 20°, frame rate = 30 fps. The communication delay could be ignored because all communications were local.

2) Results: Fig. 3.8 shows the results of the average grasping time, average releasing time, and average releasing error distance. The grasping and releasing times were normalized by the condition without external views to eliminate the influence of the participants' skills on the work. The normalization was calculated as

$$\text{Normalized time} = \frac{\text{Work time of each conditon [s]}}{\text{Work time of the condition without external views [s]}} \quad . (2)$$

Panels (a), (b) and (c) of Fig. 3.8 show the results of the normalized grasping times, normalized releasing times, and releasing distance errors, respectively, at each pan angle. The results under each condition were compared by one-way analysis of variance.

As indicated in Fig. 3.8 (a), the normalized grasping time decreased as the pan angle tended to

Fig. 3.9 Image when releasing error distance was approximately 6.5 mm

90°. The grasping time was significantly lower at the 90° pan angle than at the 30°, 45°, and 135° pan angles, and marginally changed between pan angles of 90° and 150° (Bonferroni method). Moreover, there was no significant difference between the 90° pan angle and the 60°, 75°, 105°, and 120° pan angles (Bonferroni method). Therefore, the results showed that the optimal pan angle for grasping was determined as 90°, and the allowable range was 90° ± 30°.

As shown in Fig. 3.8 (b), the fastest release was observed at a pan angle of 90°. A pan angle of 150° significantly increased the normalized releasing time from those at 60°, 90°, and 120° (Bonferroni method), but there were no significant release-time differences between the 90° pan angle and the 30°, 45°, 60°, 75°, 105°, 120°, and 135° pan angles (Bonferroni method). Thus, the optimal pan angle for releasing was determined as 90°, and the allowable pan angle ranged from 30° to 135°.

Finally, the releasing error distance decreased as the pan angle approached 90° (Fig. 3.8(c)), but the differences among the pan-angle conditions were not significant (Bonferroni method).

3) Discussion: The optimal range was hypothesized as 90° ± 30°, but the results suggested an optimal range from 30° to 135°. It was surmised that the participants did not release the target object precisely under any condition. The smallest releasing error distance among all conditions was approximately 6.5 mm. Fig. 3.9 shows the image captured from above when the releasing error distance was approximately 6.5 mm. As shown in the figure, the participants easily discerned that the target object was released with some error. Therefore, the imprecise release of the target object by the participants changed the allowable pan angle of the releasing action from its expected 90° ± 30° value.

3.3.1.2 Experiment 2: tilt angle experiment

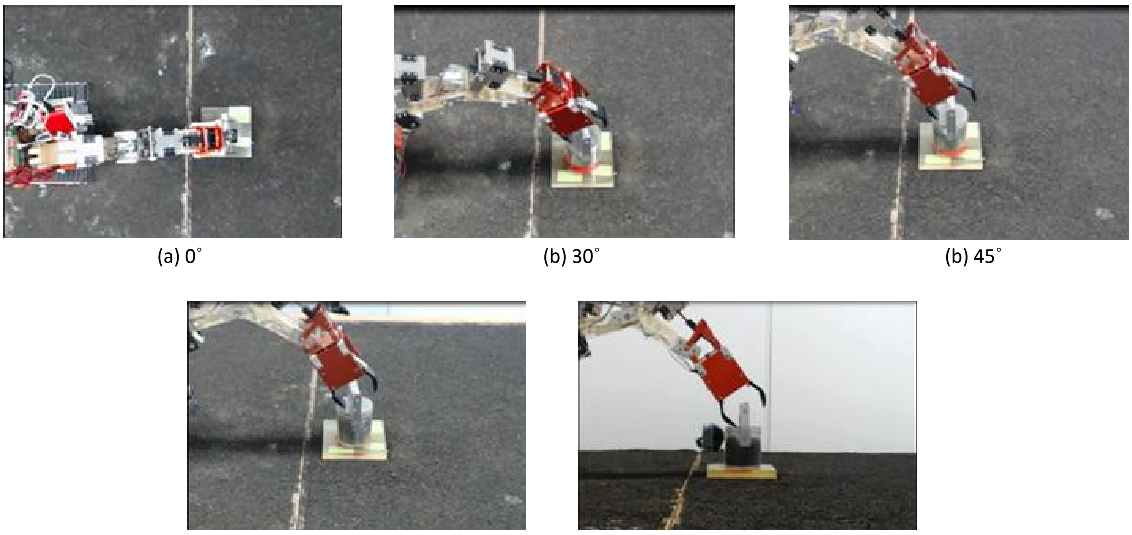
1) Experimental settings: In Experiment 2, the author derived the optimal and allowable pan angles. The participants were six novice students with no prior experience in operating construction machines. The participants acquired sufficient skills to teleoperate the construction machine of the scale model in training tasks administered prior to Experiment 2.

The five tilt angles were 0°, 30°, 45°, 60°, and 90°. Fig. 3.10 shows the five external views at each pan angle. First, the participants performed the experimental tasks thrice with no external views for normalization (i.e., to remove the effects of any acquired skills). Next, the participants tried the tasks thrice under five conditions in a randomly selected order (see Fig. 3.10). During these tasks, the participants were presented with a cab view and two external views.

The pan angle, view angle, and frame rate were fixed at 90°, 20°, and 30 fps, respectively. Communication delay was ignored.

2) Results: Fig. 3.11 shows the results of Experiment 2. The grasping and releasing times were normalized by Eq. (2), as done in Experiment 1. Panels (a), (b) and (c) of this figure present the normalized grasping times, normalized releasing times, and releasing error distances, respectively, at the different tilt angles.

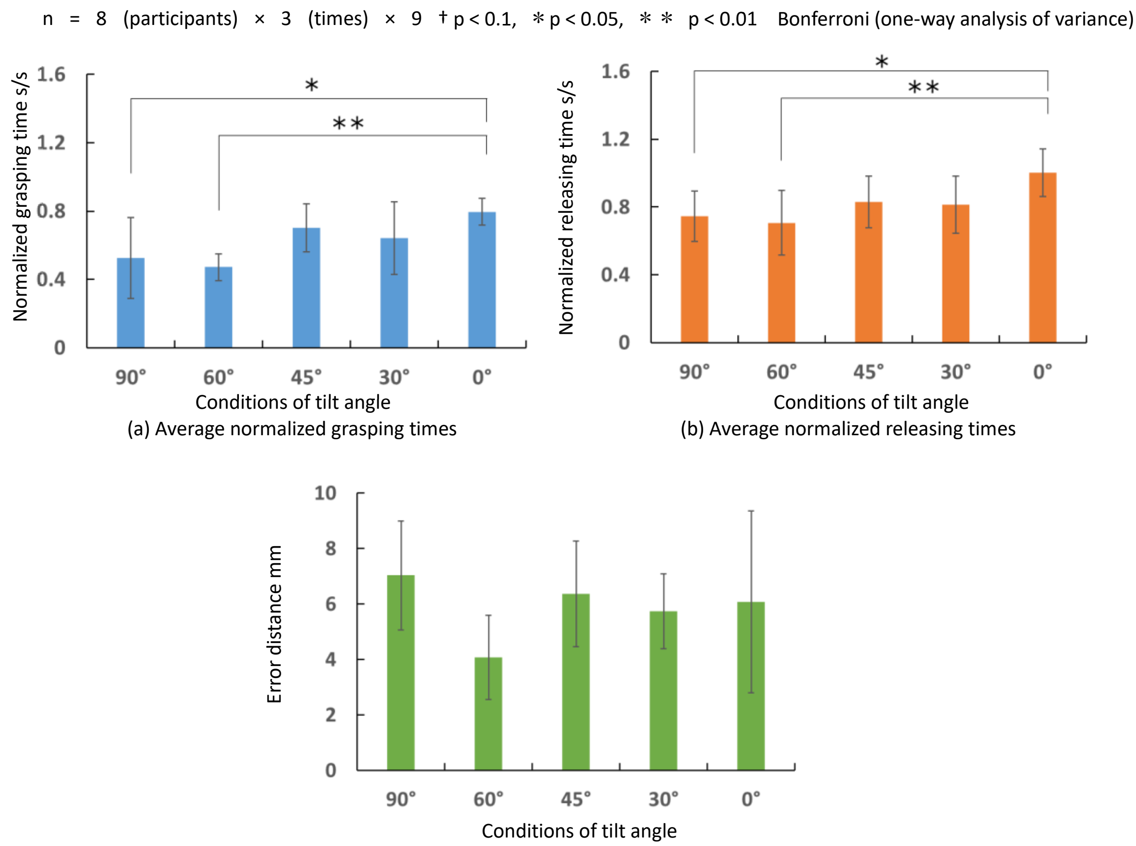
As shown in Fig. 3.11 (a), the fastest tilt angle for grasping is 60°. A tilt angle of 0° incurred more grasping-time costs than tilt angles of 60° and 90° (Bonferroni method). No significant differences were observed between the 60° tilt angle and the 30°, 45°, and 90° tilt angles. Therefore, the optimal tilt angle for grasping was determined as 60°, and the allowable tilt angle was 60° ± 30°.



(a) 0° (b) 30° (b) 45°

(b) 60° (b) 90°

Fig. 3.10 External views of each tilt angle



(a) Average normalized grasping times

(b) Average normalized releasing times

(c) Error distances between a target object and designated place area

Fig. 3.11 Results of Experiment 2

The releasing time was also fastest at a tilt angle of 60° (Fig. 3.11(b)), and was slower at 0° tilt angle than at 60° and 90° tilt angles. No significant differences were observed between tilt angles of 60° and tilt angles of 30°, 45°, and 90°. Therefore, the optimal tilt angle for releasing was determined as 60°, and the allowable tilt angle was 60° ± 30°.

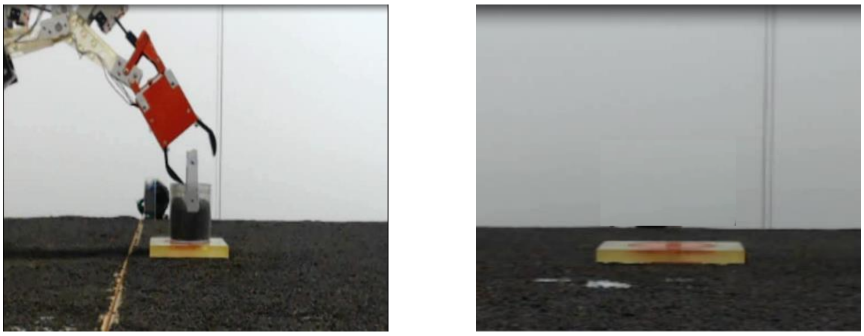
Although the error distance was minimized at the 60° tilt angle (Fig. 3.11(c)), this parameter was not significantly influenced by tilt angle (Bonferroni method).

3) Discussion: The original tilt angle was hypothesized as 90° but the obtained optimal angle was 60°. Possibly, the participants were able to discern the target objects in 3D only when the object was tilted by 60° in the external view (see Fig. 3.12(a)). In contrast, when the tilt angle was 90°, the participants were unlikely to discern the target objects in 3D (Fig. 3.12 (b)).

3.3.2 Summary

From the results of a scale-model experiment, the author derived the optimal and allowable ranges of the pan and tilt angles during grasping and releasing tasks. The author hypothesized that canonical views with the fewest occlusions were optimal camera placements, and that rotations

(a) Grasping and releasing an objected tilted at 60°



(b) Grasping and releasing an object tilted at 90°

Fig. 3.12 Grasping and releasing at different tilt angles

of less than 30° from the canonical view were allowable camera placements. Within the allowable camera placement, humans recognize images to the same extent as canonical views.

The hypotheses were tested in scale-model experiments. The optimal and allowable pan angles for grasping and releasing were estimated as 90° and 90° ± 30°, respectively, but the measured allowable pan angles in releasing tasks ranged from 30° to 135°. Furthermore, the optimal tilt angle in grasping and releasing tasks was determined as 60°, and the allowable tilt angles in grasping and releasing tasks were 60° ± 30°.

However, the participants in these experiments were all novice operators, and the experiments were conducted on a scale model. The author aims to conduct additional experiments using an actual construction machine run by professional teleoperators. Furthermore, the experiments were conducted under limited conditions; that is, the tilt angle was fixed in Experiment 1, and the pan angle was fixed in Experiment 2. To remove this limitation, the author conducted additional experiments with different combinations of pan and tilt angles, as described in the next section.

3.4 Experiments on an actual machine with novice operators

This subsection investigates the optimal and allowable pan and tilt angles in manipulation tasks on an actual machine operated by novice teleoperators. The results confirm that the results obtained via the scale model can be applied to the actual construction machine.

3.4.1 Limitations of experiments using actual machines

Cameras at disaster sites should always provide views during unmanned construction work. If the cameras fall down, their views are no longer available to teleoperators, and the fallen cameras themselves become obstacles. Furthermore, falling cameras can lead to secondary disasters such as mudslides. To ensure their safety in earthquakes and gales, cameras should be set low (i.e., with tilt angles near 90°). As executing numerous experiments on actual construction machinery is infeasible in terms of safety, time, and cost, the present experimental investigations aim to prove that tilt angles from 45° to 90° have little impact on the work efficiency.

3.4.2 Hypothesis

Based on the characteristics of the canonical views explained in Section 3.2, the experimental results of the scale model described in Section 3.3, and the limitations of experiments on actual machine described in Section 3.4.1, the author proposed the following four hypotheses:

1) The optimal pan angle is 90°.
2) The optimal tilt angle is 60°.
3) The allowable range of the pan angle is 90° ± 30°.
4) Tilt angles between 45° and 90° have little impact on the work efficiency.

3.4.3 Experimental settings

The experiments were conducted on an actual construction machine (Hitachi ZAXIS 35 U; see Fig. 3.13). The interface of the construction machine is shown in Fig. 3.14. The participants

Fig. 3.13 Actual construction machinery (ZAXIS 35U)

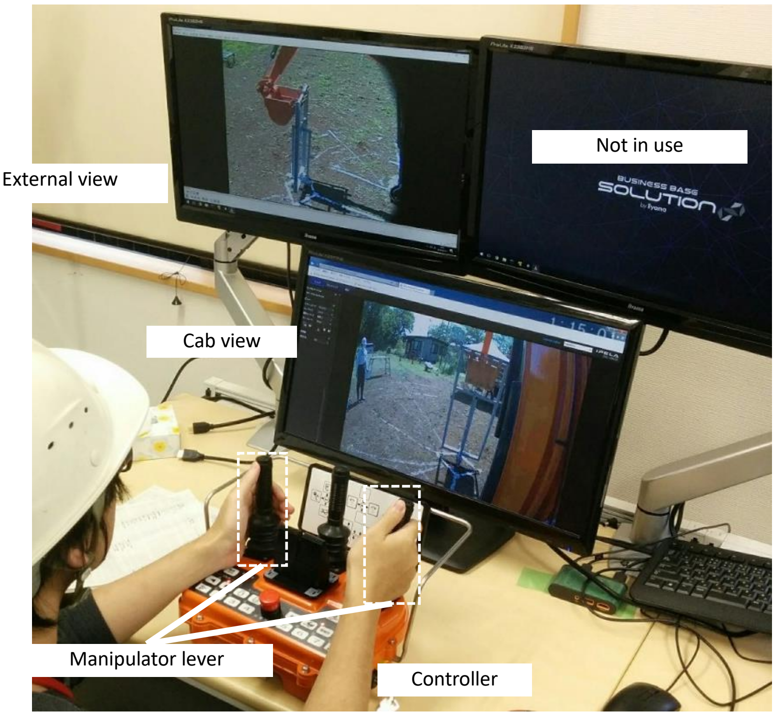
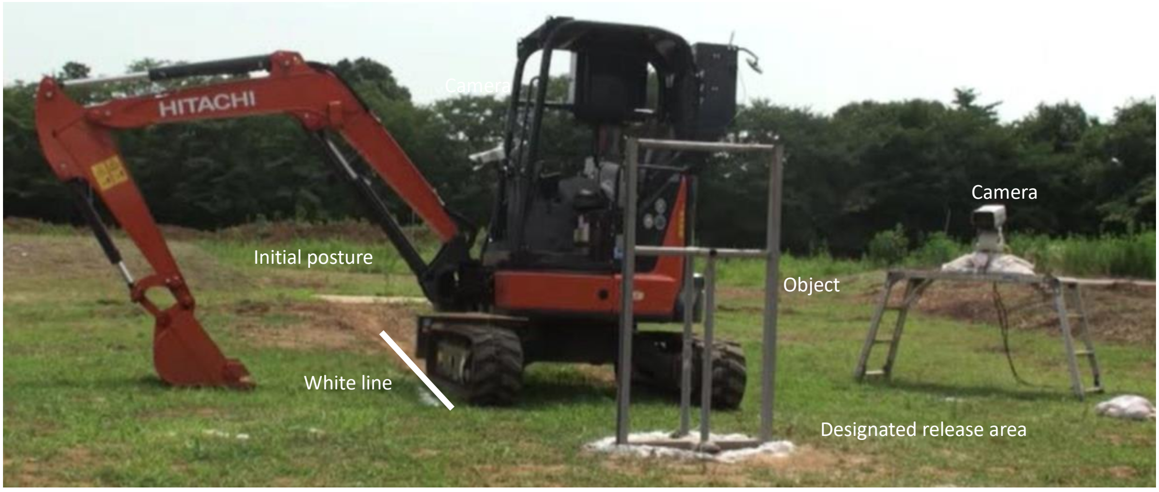


Fig. 3.14 Interface for the actual construction machinery

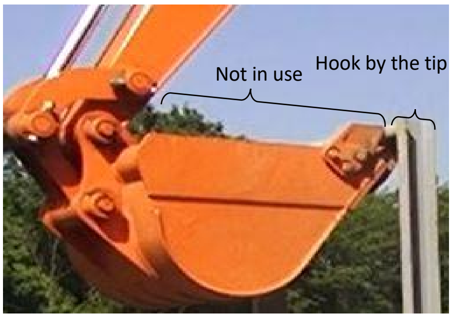
teleoperated the construction machine via two manipulator levers while watching a cab view and an external view on the monitors.

Five student participants with no prior experience in teleoperation of any construction machineries were invited to the experiment, because the number of skilled operators in Japan (~20) is very small [3.18]. However, after pre-training on the experimental tasks, the participants obtained sufficient skills to teleoperate the construction machine in the actual experiments.

The experimental tasks were referenced to a model task of unmanned construction [3.17]. Digging tasks were modeled as hooking an object because reproducing a given volume of gravel



(a) Initial posture and experimental environment



(b) Hooking an object

(c) Turning right until the machine passed over the white line

(d) Experimenter re-hooking

(e) Releasing

Fig. 3.15 Experimental procedure

or sand in a bucket is infeasibly difficult. The experimental tasks consisted of digging (i.e., hooking an object) and releasing it to a designated release area, as shown in Fig. 3.15 (a). An experimenter judged whether the hooked object was released inside or outside the designated area. If the object was released outside the designated area, the experimenter prompted the participants to re-hook and release the object. The steps of the experimental procedure were as follows:

1) Turn 90° to the left from the initial position (see Fig. 3.15 (a))
2) As a proxy of digging, hook an object by the tip of the bucket (see Fig. 3.15 (b))
3) Turn to the right until the bucket containing the object passes over the white line (see Fig. 3.15(c))

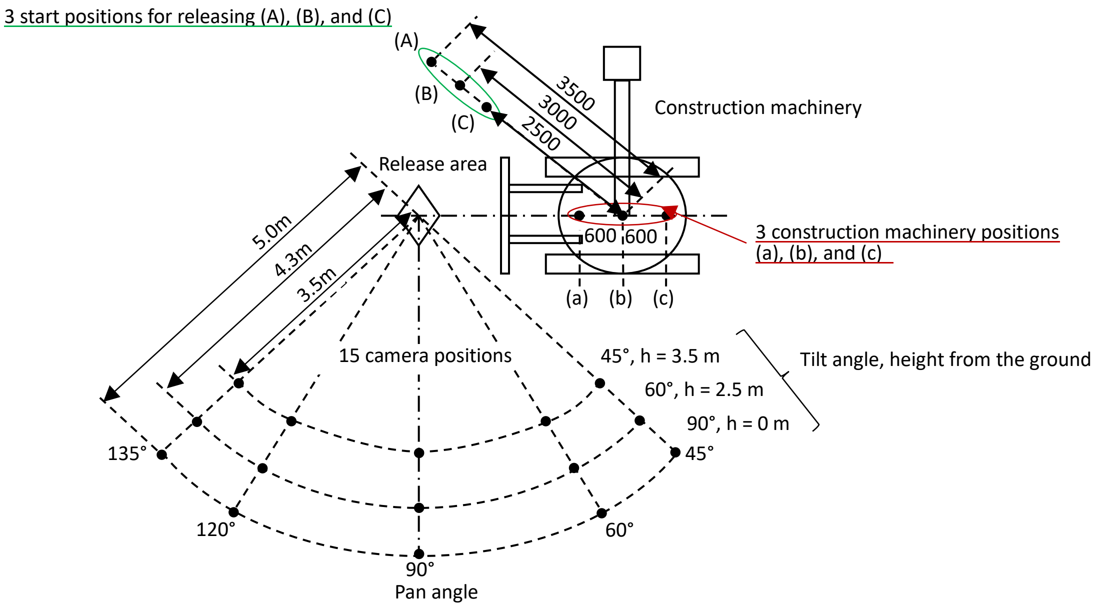


Fig. 3.16 Layout of the environment and the conditions of the experiment

Fig. 3.17 External views under each condition, where (x°, y°) represents (pan angle, tilt angle)

4) An experimenter hooks an object again during the boarding operation (see Fig. 3.15 (d)) to eliminate the effects of digging (or hooking) skills of the participants
5) Releasing the dug object in the designated release area (see Fig. 3.15 (e)) after the experimenter completes hooking and disembarks from the construction machine
6) Turn to the right until the bucket passes over the white line (see Fig. 3.15 (a))

The pan angle was varied as 45°, 60°, 90°, 120°, and 135°, and the tilt angle was varied as 45°, 60°, and 90°. The experiments were conducted under 15 conditions (combinations of the five pan angle conditions and three tilt angle conditions), as shown in Fig. 3.16. Fig. 3.17 shows the views under all conditions. To exclude the zoom effects, an external camera was installed to maintain a

5-m distance between the object and the external camera under all conditions. The participants tried the experimental tasks thrice under each condition. The positions of the construction machine and the start positions for release were varied among the three trials to exclude the effects of mastering the manipulations at specific positions (see Fig. 3.16). In the first trial under each condition, the construction machine started from position (b) in Fig. 3.16, and the release was started at position B. In the second and third trials under each condition, the construction machine was started from positions (a) and (c), respectively, and the release was started at positions A and C, respectively. Three participants performed the experiments in order A, defined as (pan angle, tilt angle) = (135, 45), (120, 60), (90, 90), (60, 60), (45, 45), (45, 60), (60, 90), (90, 60), (120, 45), (135, 60), (135, 90), (120, 90), (90, 45), (60, 45), and (45, 90). Two participants performed the experiments in order B, which was the reverse of order A; that is, (pan angle, tilt angle) = (45, 90), (60, 45), (90, 45), (120, 90), (135, 90), (135, 60), (120, 45), (90, 60), (60, 90), (45, 60), (45, 45), (60, 60), (90, 90), (120, 60), and (135,45). The order was changed to minimize the order effect and possible mastery caused by task repetition under the same conditions.

The digging and releasing times, defined when the bucket was inside the white line (see Fig. 3.15(a)), were measured. The digging time began when the bucket without the object passed over the white line, and ended when the bucket containing the hooked object passed over the white line. The release time began when the bucket with the hooked object passed over the white line, and ended when the bucket without the object passed over the white line.

3.4.4 Results

1) Order effect: The order effect must be considered because Order A and Order B were performed by different numbers of participants. If the order effect is significant, the first half of the trials should take longer than the latter half. The average work time, that is, the average summed digging and releasing times, was 86.0 s (SD = 34.6 s) in the first half and 90.5 s (SD = 36.8 s) in the latter half. Therefore, the order effect can be ignored because the first half did not consume more time than the second half.

2) Digging: Fig. 3.18 shows the digging results. The digging time was normalized by Eq. (2), based on the participants' work time in the training tasks.

Fig. 3.18(a) shows the average normalized digging time under each condition. The digging time was minimized at a pan angle of 60°. However, the Bonferroni method revealed no significant differences between any of the (pan angle, tilt angle) pairs, possibly because the number of participants was very small.

Fig. 3.18(b) shows the average normalized digging time at each pan angle, which was again

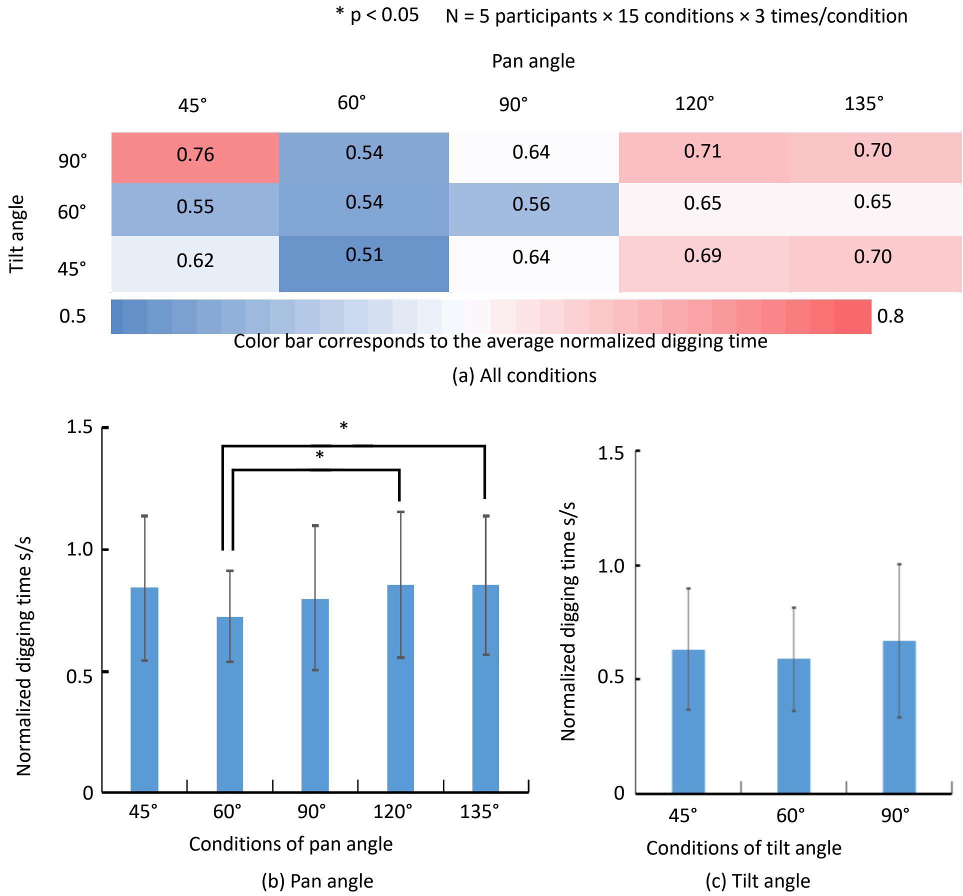


Fig. 3.18 Results of digging

minimized at a pan angle of 60°. The Bonferroni method revealed significant differences between pan angles of 60° and 120° and between 60° and 135°, but no significant differences between pan angles of 60° and 45° and between 60° and 90°. Thus, the results indicate that the optimal pan angle for digging was determined as 60°, and the allowable range was 45–90°.

Fig. 3.18(c) shows the average normalized digging time at each tilt angle. The digging time was minimized at a tilt angle of 60°. The Bonferroni method confirmed no significant differences between any pairs of conditions, thus proving that the optimal tilt angle and allowable range of digging were 60° and 45–90°, respectively.

3) Release: Fig. 3.19 shows the results of releasing the object. The releasing time was normalized by Eq. (2), based on the participants' work time in the training tasks. Fig. 3.19 (a) shows the average normalized release times under the various conditions. When both the pan and tilt angles were 45°, the releasing times were minimized and almost independent of the conditions (see Fig. 3.19(a)). No significant differences between any pairs of conditions were found by the Bonferroni method.

Fig. 3.19(b) shows the average normalized release times at all pan angles. The release time was lowest at a pan angle of 45°, and was not significantly different among the conditions (Bonferroni

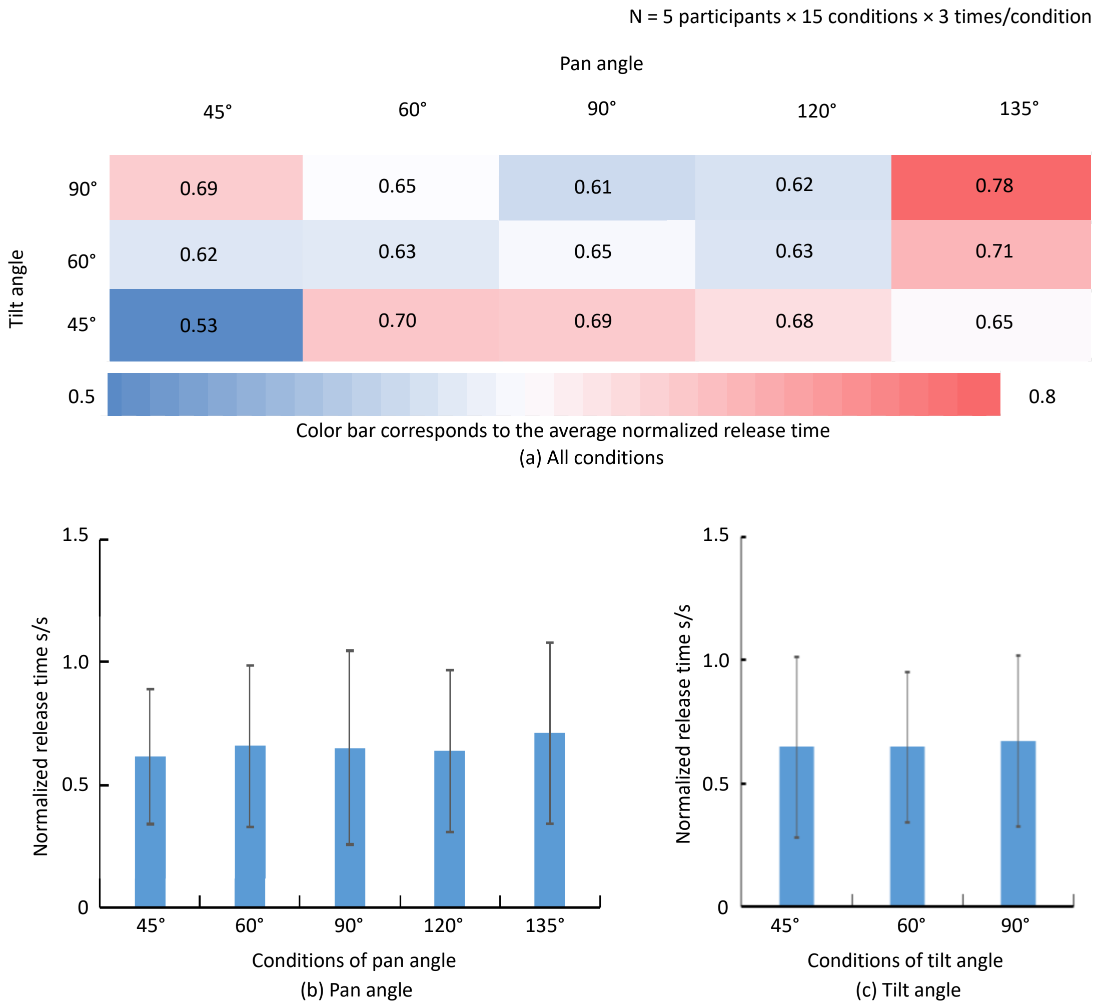


Fig. 3.19 Results of release

method). Therefore, the optimal pan angle for releasing was determined as 45°, and the allowable range was 45–135°. However, the release times were almost independent of pan angle (see Fig. 3.19 (b)).

Fig. 3.19(c) shows the average normalized release time at each tilt angle. The release time was minimized at a tilt angle of 60°, but the differences among the tilt-angle conditions were not significant (Bonferroni method). Therefore, the optimal tilt angle for releasing was determined as 60°, and the allowable range was 45°–90°.

3.4.5 Discussion

The author discusses two findings in which the assumptions and the acquired experimental results did not match. First was the choice of 60° as the optimal pan angle for digging, and the second is the minimal effect of the pan angle on the releasing time.

1) Optimal pan angle for digging: The 3D positional relationship between the object to be hooked and the bucket was difficult to discern from the external view with a 90° pan angle. Fig. 3.20 shows the views with pan angles of 60° and 90°. When the pan angle was 90°, the teleoperators could not easily discern the x-axis of the external views (see Fig. 3.20(b)), but the pan angle of 60° revealed the x-axis more clearly than 90° (see Fig. 3.20 (a)). Therefore, the participants could discern the 3D positional relationship from the external views with a pan angle of 60°, which might explain why this pan angle accelerated the work time.

Participants could also discern the 3D positional relationship from external views with a 120° pan angle. However, as the digging time was significantly lower at the 60° pan angle than at the 120° angle, the 60° pan angle was assumed in further discussion. The highly efficient digging at the 60° pan angle can be attributed to few occlusions and effective mental rotation.

Fig. 3.21 shows the views at pan angles of 60° and 120°. At the 60° pan angle, the participants discerned the tip of the bucket because the view was almost unobstructed by occlusions (see Fig. 3.21 (a)). In contrast, at the 120° pan angle, the tip of the bucket was obscured; especially, the right side of the bucket was hidden by occlusions (see Fig. 3.21 (b)). For this reason, participants might not fully discern the tip of the bucket at pan angles above 90°. Fig. 3.22 compares the average normalized digging times at pan angles below 90° (combination of 45° and 60°) and pan angles above 90° (combination of 120° and 135°). The digging time was significantly faster at pan angles < 90° than at pan angles > 90° (t-test). Furthermore, the questionnaire survey showed that teleoperations were much easier at pan angles below 90° than at pan angles above 90°. Thus, the participants could discern the bucket efficiently by watching views with pan angles < 90°, which might explain why 60° was the optimal pan angle for grasping.

Next, the influence of mental rotation is discussed [3.18]. If a camera is installed ahead of the machine, the effective pan angle is > 90°, and the left and right parts of the construction machine differ between the internal and external views, as described in 1.17. Because the participants were required to mentally rotate the views, the optimal pan angle for grasping could be 60° rather than 120°. If mental rotation retards the task time, it should influence both the digging and releasing tasks. Fig. 3.23 compares the average normalized releasing times at pan angles less than 90° (combination of 45° and 60°) and pan angles above 90° (combination of 120° and 135°). No significant difference was found between these two conditions (t-test). During release, the positional relationship between the designated release area and the bottom of the hooked object can be more important than the tip of the bucket, because the participants must release the hooked object inside the designated release area. Thus, occlusions of the bucket tip have little influence on releasing. Therefore, occlusions probably exert a bigger impact than mental rotation because there was no significant difference between less than 90° and more than 90° pan angles during releasing.

(a) 60° (b) 90°

Fig. 3.20 External views with pan angles of 60° and 90°

(a) 60° (b) 120°

Fig. 3.21 External views with pan angles of 60° and 120°

N = 5 participants × (3 trials × 3 tilt conditions) × 2 × 2 pan conditions ** $p < 0.05$

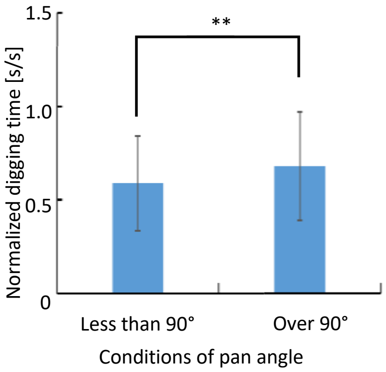


Fig. 3.22 Normalized digging times at pan angles < 90° and > 90°

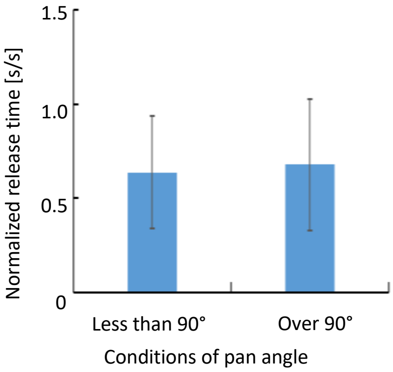


Fig. 3.23 Normalized releasing times at pan angles < 90° and > 90°

2) Effects of the pan angle on releasing: The pan angle only minimally affected the releasing time. When the bottom of the hooked object was close to the designated release area, the designated release area was occluded at all pan angles, as shown in Fig. 3.24. Thus, the participants could not discern the positional relationship between the designated release area and the bottom of the hooked object. This effect might explain why the pan angle exerted little influence on the releasing time.

3.4.6 Summary

The optimal and allowable pan and tilt angles in the external views during digging and releasing tasks were determined in experiments on an actual construction machine. In digging, the optimal pan angle was 60°, and the allowable range was 45−90°. The optimal tilt angle was also 60°, with an allowable range of 45−90°. The following experiments were conducted by skilled operators.

3.5 Experiments using an actual machine with skilled operators

This section examines the optimal and allowable pan and tilt angles for digging and releasing in experiments on an actual machine operated by skilled participants. A teleoperator's gaze and the required field of view differ between unmanned construction [3.10] and other teleoperation fields, such as surgery [3.20] and robotic operations [3.21]. Skilled operators of unmanned construction usually work at disaster sites. The above experimental results proved that the optimal pan angle for digging by novice teleoperators was 60° (see Fig. 3.1). In this view, the novice teleoperators could discern the 3D positional relationship from the external view alone (as explained in Section 3.4.6). However, skilled operators can discern the crosswise direction of the construction machine by watching a cab view, because a similar view is used in on-board

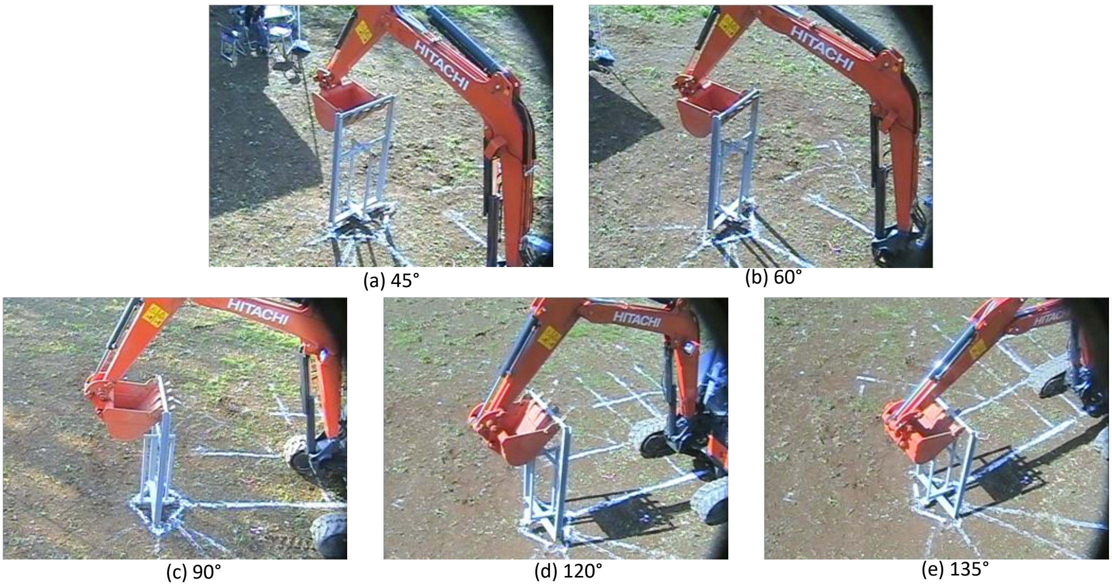


Fig. 3.24 External views during releasing

operations. Therefore, the optimal pan angle for skilled operators can be 90°, which is more effective for depth perception than 60°. The differences in the optimal and allowable pan and tilt angles between the novice and skilled operators can then be determined. Installing cameras for novice teleoperators may degrade the work efficiency when the skilled operators are at work. Therefore, this section aims to derive the optimal and allowable pan and tilt angles for skilled operators during digging and releasing. Moreover, revealing the differences between novice and skilled operators can help to develop future support systems for novice teleoperators. The experimental settings were those of the experiments on the actual machine operated by novice operators.

3.5.1 Results

3.5.1.1 Digging results

Fig. 3.25(a) shows the average digging time under all conditions. The digging time significantly depended on the camera placement (one-way analysis of variance, $F = 2.40$, $p = 0.005$). The task time was minimized under the condition (pan angle φ, tilt angle θ) = (90°, 60°), and was significantly slower under (φ, θ) = (120°, 60°) ($p = 0.02$) than under the fastest condition. A significant trend was observed between (φ, θ) = (45°, 90°) and (φ, θ) = (90°, 60°) (Holm–Sidak method, $p = 0.07$).

Fig. 3.25(b) shows the average digging time at each pan angle. The digging time significantly depended on the pan angle (one-way analysis of variance, $F = 5.28$, $p < 0.001$). The task time was minimized at a pan angle of 90°. Significant differences were observed between 45° and 90° ($p = 0.01$), 120° and 90° ($p = 0.001$), and 135° and 90° ($p = 0.01$). No significant differences were observed between 60° and 45° (Holm–Sidak method). These results suggested that the optimal pan angle was 90°, and the allowable pan angles ranged from 60° to 90°.

Fig. 3.25(c) compares the average digging times at different tilt angles. The digging task was finished earliest at a tilt angle of 45°. The digging time did not significantly depend on tilt angle (one-way analysis of variance, $F = 0.03$, $p = 0.97$). The results suggested that the optimal tilt angle was 45°, and the allowable tilt angles ranged from 45° to 90°.

3.5.1.2 Releasing results

Fig. 3.26(a) shows the average releasing times under all conditions. The releasing time significantly depended on the camera placement (one-way analysis of variance, $F = 2.09$, $p =$

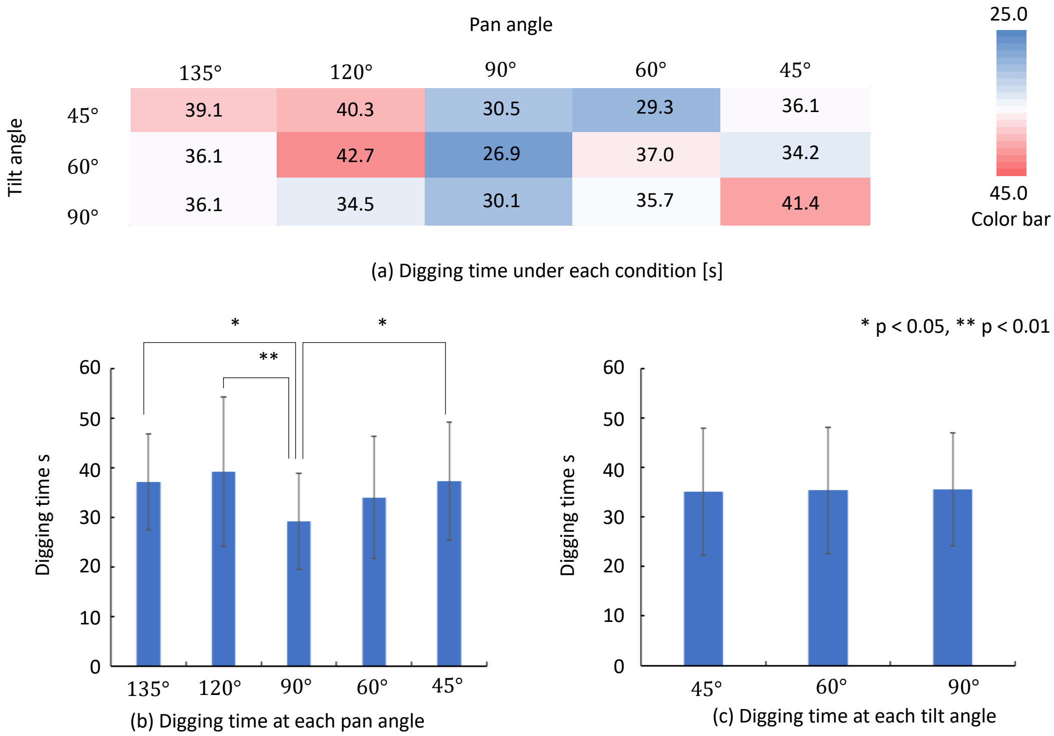


Fig. 3.25 Results of digging time

0.01), and was completed fastest under the condition (pan angle φ, tilt angle θ) = (90°, 90°). Under the condition (φ, θ) = (90°, 45°), the task was significantly slower than under the fastest condition (Holm–Sidak method).

Fig. 3.26(b) shows the average releasing time at each pan angle. The releasing time did not significantly depend on pan angle (one-way analysis of variance, $F = 0.79$, $p = 0.53$), but was fastest at 45°. Thus, the optimal pan angle was 45°, and the allowable pan angles ranged from 45° to 135°.

Fig. 3.26(c) shows the average releasing time at each tilt angle. The releasing time significantly depended on tilt angle, although the significance was marginal (one-way analysis of variance, $F = 2.97$, $p = 0.06$). The task time was minimized at a tilt angle of 90°, and a marginal significance was observed between the 45° and 90° tilt angles (Holm–Sidak method, $p = 0.06$). Therefore, the optimal tilt angle was determined as 90°, and the allowable tilt angles ranged from 60° to 90°.

3.5.1.3 Summary of results

Table 3.1 summarizes the optimal and allowable pan and tilt angles for both novices and skilled operators during the digging and releasing work. The differences between the novice and skilled operators will be discussed for the following factors: (i) optimal pan angle for digging, (ii) allowable pan angle for digging, (iii) optimal tilt angle for releasing, (iv) allowable tilt angle for

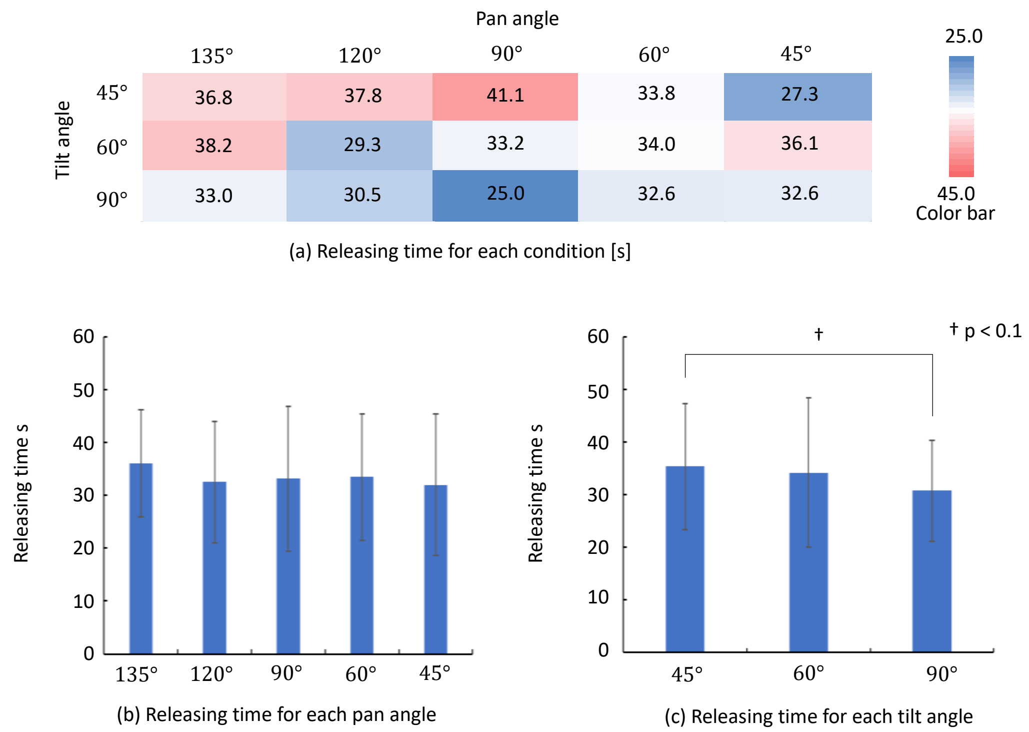


Fig. 3.26 Results of the releasing time

Table 3.1. Comparison of experimental outcomes of novices and skilled operators

			Novice	Skilled
Digging	Pan	Optimum	60°	90°
		Allowable	45°–90°	60°–90°
	Tilt	Optimum	60°	45°
		Allowable	45°–90°	45°–90°
Releasing	Pan	Optimum	45°	45°
		Allowable	45°–135°	45°–135°
	Tilt	Optimum	60°	90°
		Allowable	45°–90°	60°–90°

releasing, and (v) optimal tilt angle for digging. The different optimal tilt angles in digging might have arisen because the tilt angle little affected the digging time of both novices and skilled operators.

3.5.2 Discussion

This section discusses four reasons for the differences between the skilled and novice operators, as described in Section 3.5.1.3.

3.5.2.1 Optimal pan angle of digging

First, the author discusses why the optimal pan angle of digging differs between the skilled and novice operators. The difference could have been caused by the better adjustability of the skilled operators to the X axis (i.e., the crosswise direction of the construction machine) as shown in Fig. 3.27. The optimal pan angle for the novice teleoperators during digging was 60°, because they could discern the 3D positional relationship between the bucket and the object from this view (as seen in Fig. 3.27(a) and described in the previous section). However, they could not easily discern the X-axis information from a pan angle of 90° (see Fig. 3.27 (b)).

All of the participants (skilled operators and novices) were asked to first adjust the X-axis direction by watching the cab view, because the task first required turning to the left (see Section 3.4.3). This task could be difficult for the novice teleoperators because they lacked experience in teleoperating a construction machine, but was easily accomplished by the skilled operators, as the cab view was similar to the views they watched during on-board operations. Fig. 3.28 shows the number of rotations (number of lever inputs) required by the participants to adjust the X-axis

Fig. 3.27 External views from pan angles of 60° and 90°

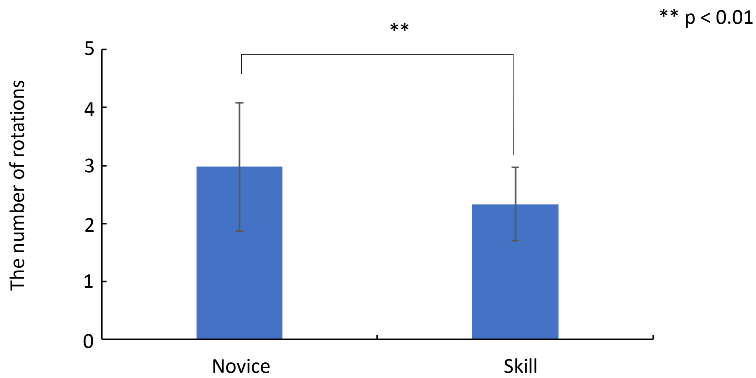


Fig. 3.28 Comparison of number of rotations when the pan angle in the external view is 90°

direction. The number of rotations significantly differed between the skilled and novice participants (Welch's t-test, $t(73) = 3.21$, $p = 0.002$). The participants best perceived the depth from the views with a pan angle of 90°. Therefore, the optimal pan angle for the skilled operators was 90° because from this angle, they could adjust the X-axis direction using the cab view alone.

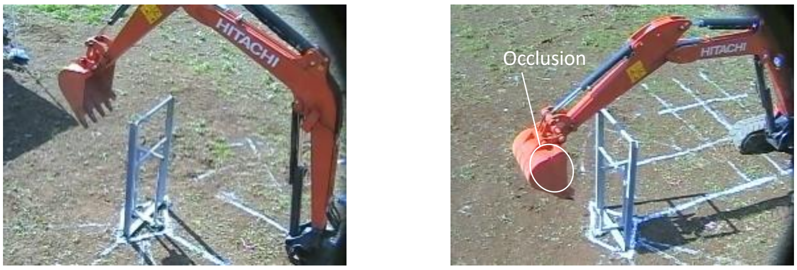
3.5.2.2 Allowable pan angle of digging

Next, the author discusses why the allowable pan angle for digging differs between the skilled and novice operators. As explained in Section 3.5.2.1, the optimal pan angle was 60° for novices and 90° for skilled operators. Owing to the characteristics of canonical views, humans can deliver the same performance if the canonical view rotates by less than 30° [3.12]. This characteristic suggests that a pan angle of 45° is allowable for novices but not for skilled operators. Specifically, the object was only 15° from the optimal pan angle of the novices, but was 45° from that of the skilled operators.

The largest allowable pan angle for the skilled participants was not 120°, but 90°. This limit might be imposed by (i) occlusion and (ii) mental rotation, as observed in novice participants.

i) Occlusion: Fig. 3.29 shows the external views from pan angles of 60° and 120°. As explained in Section 3.4.6, the participants might not clearly discern the tip of the bucket, especially the right tip, at a pan angle of 120° because the views are occluded. However, this phenomenon is less likely when the pan angle of the views is 60°. Fig. 3.30 compares the average digging times at pan angles below and above 90°. The digging time was longer at pan angles below 90° than at pan angles above 90°, but the difference was not statistically significant (t-test, $t(71) = 1.40$, $p = 0.17$).

For a more detailed analysis, the digging time was calculated for each participant because the above insignificant differences were significant for the novice participants. A participant required more digging time when viewing from pan angles below 90° than when viewing from pan angles above 90°. This result might be explained by a single participant who extended the arm of the construction machine after (rather than before) crossing the white line (see Fig. 3.31). In this case, the digging time involved the time of arm extension because it defines the time from the bucket crossing the white line to the bucket re-crossing the white line after hooking the object. The arm-extension time was included in 8 out of 18 digging trials at pan angles below 90°, and in 3 out of 18 digging trials at pan angles above 90°. This difference in the number of digging time with the lengthened arm-extension time might be responsible for the longer digging time of the participant viewing pan angles below 90°. This anomalous participant may have extended the arm insufficiently at pan angles under 90° because the external views excluded the bucket during the reaching process. Fig. 3.32 shows the external views when the bucket appeared at pan angles of



(a) 60°　　(b) 120°

Fig. 3.29 Environmental views with pan angle of (a) 60° and (b) 120°

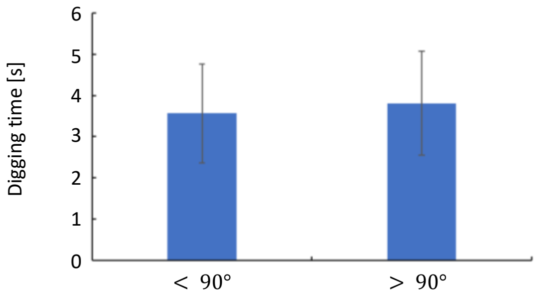


Fig. 3.30 Comparison of digging time at pan angles of < 90° and > 90°

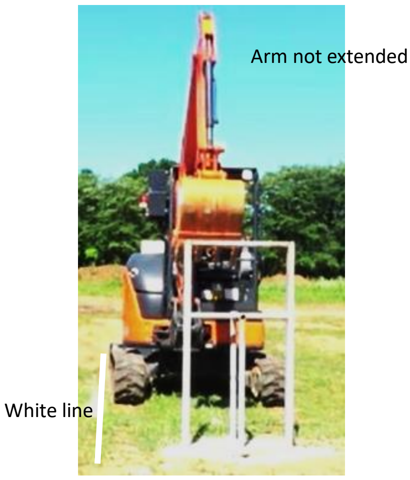


Fig. 3.31 Arm not extended during rotation

60° and 120°. In the image viewed from 60°, the bucket was invisible until it approached the object (Fig. 3.32(a)), but in the image viewed from 120°, the bucket was seen at its initial position. Therefore, this participant could not extend the arm during the rotation period because the external view taken at 60° excluded the bucket during this period.

Fig. 3.33 compares the digging times at pan angles below and above 90° for the three participants who extended the arm during the rotation period. The difference between the digging times under the two conditions was marginally significant (t-test, $t(53) = 1.84$, $p = 0.07$), confirming that the pan angle of 120° was outside the allowable range.

ii) Mental rotation: Mental rotation, which includes the differences between the left and right

(a) 60° (b) 120°

Fig. 3.32 External views of the bucket appearing at pan angles of (a) 60˚ and (b) 120˚

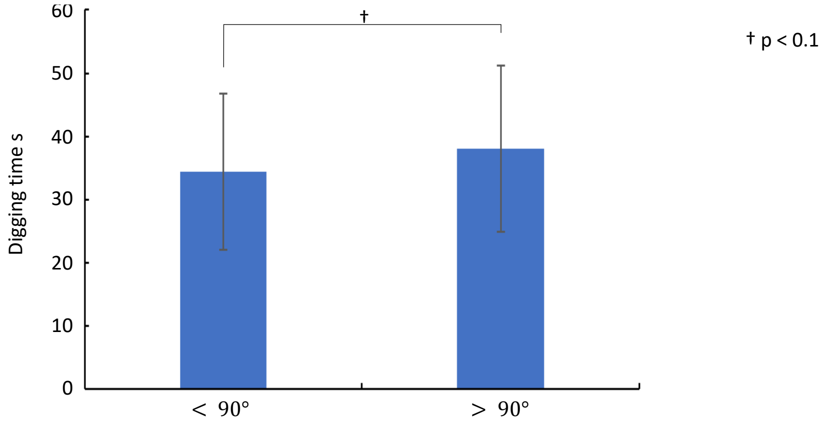


Fig. 3.33 Comparison of digging times at pan angles < 90˚ and > 90˚, excluding the participant who did not extend the arm before digging

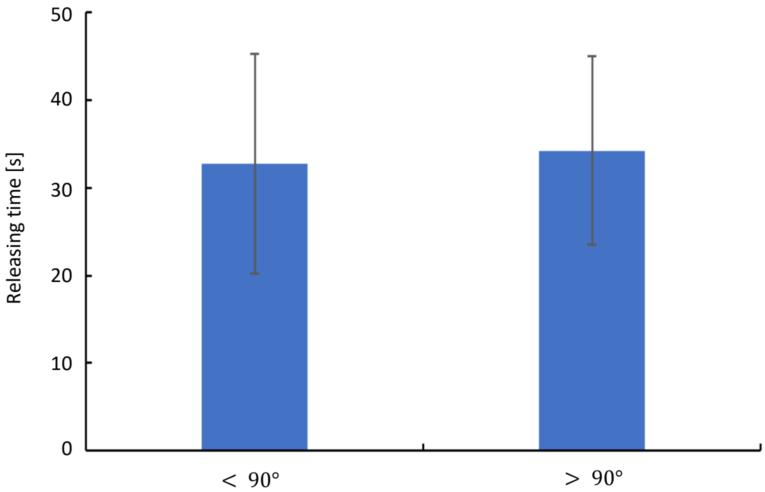


Fig. 3.34 Comparison of releasing times at pan angles of 90˚ and 90˚

sides in the views and in the construction machines, can degrade the work efficiency [3.18], as explained in Section 3.4.6. When comparing releasing times, the effects of mental rotation are important because the participants must mentally rotate the image during both digging and releasing tasks, whereas occlusions are important only in digging tasks. Fig. 3.34 shows the releasing time at pan angles of above and below 90°. The releasing time was reduced when viewing from pan angles below 90°, but the difference was not significant (t-test, $t(71) = 0.88$, $p = 0.39$). This result might be attributable to the less frequent rotation control in the experimental tasks.

The skilled participants rotated the levers 2.3 times on average (see Fig. 3.28). The experimental tasks of both digging and releasing required at least two controls of the rotation levers: one for turning left and the other for turning right. Therefore, the skilled participants gave 0.3 additional inputs to the rotation levers on average. Furthermore, the participants could learn the required rotation controls because these controls were required under all camera conditions. The participants tried the digging and releasing tasks 90 times. These results proved that occlusions could be more important than mental rotation ability.

3.5.2.3 Optimal tilt angle of releasing

The author now discusses the difference in the optimal tilt angle for releasing by novice and skilled participants. Large tilt angles provide more explicit views for discerning the distance between the object and the ground (see Fig. 3.35). Therefore, a tilt angle of 90° was optimal for the skilled participants, as they could easily discern the height information.

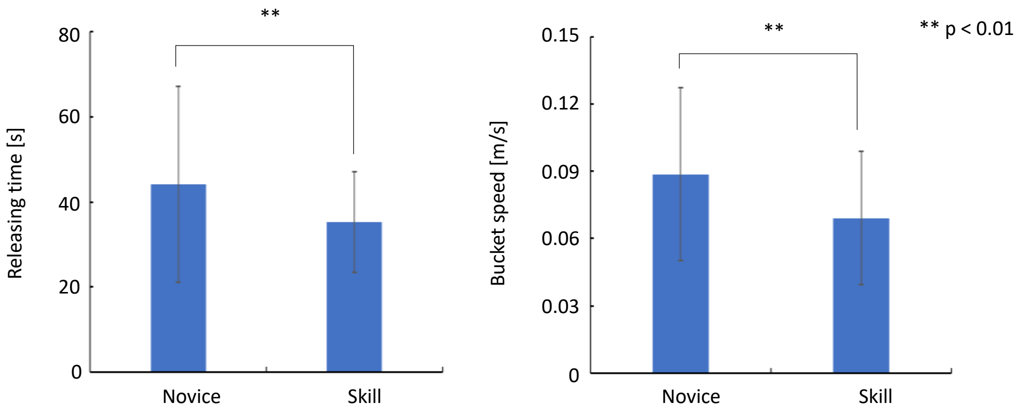
For novice participants, the optimal tilt angle for releasing was not 90°, possibly because these participants released the hooked object without sufficiently discerning the height information (Z-axis). Fig. 3.36 compares the releasing times and bucket speeds of the novice and skilled participants at a tilt angle of 45°. The skilled participants were significantly faster than the novice participants (Welch's t-test, $t(115) = 2.83$, $p = 0.005$), but the novice participants moved the bucket significantly faster than the skilled participants (Welch's t-test, $t(108) = 3.01$, $p = 0.003$). The participants might not easily discern the height information from views with smaller tilt angles. Moreover, in their interview responses, the novice participants expressed tentativeness in releasing the object because of their difficulties in discerning the height information. Nevertheless, the novice participants successfully released the object because the releasing task has a large allowable error (approximately ±8 cm). Accordingly, the tilt angle little affected the novices' releasing time because the object was released without sufficient height information.

3.5.2.4 Allowable tilt angle of releasing

Finally, the author discusses the difference in the allowable tilt angle of releasing by novices and skilled operators. This difference might have been caused by the different optimal tilt angles

(a) 45°　　(b) 60°　　(c) 90°

Fig. 3.35 Environmental views at tilt angles of (a) 45°, (b) 60°, and (c) 90°



(a) Releasing time　　(b) Bucket speed

Fig. 3.36 Comparison of releasing time and hand speed at a tilt angle of 45°

for the novice and skilled participants (60° versus 90°). As discussed in Section 3.5.2.2, an object can be viewed with the same performance when its fully canonical view is rotated by less than 30° [3.12]. Therefore, a tilt angle of 45° (which differed by 15° from the optimum for novice participants and by 45° from the optimum for skilled participants) was allowable for novice participants but unacceptable for skilled participants.

3.5.2.5 Summary of differences between the novice and skilled operators

1) Digging: The optimal pan angle was 60° for novice participants and 90° for skilled participants, possibly because the skilled participants could easily adjust the left and right directions of the construction machine. The allowable pan angle for both novice and skilled participants was −30° from the optimal pan angle. The + direction was prohibited by occlusions.

2) Releasing: A tilt angle of 45−90° was allowable for novice participants but skilled participants (with an optimal tilt angle of 90°) were limited to tilt angles of 60−90°. This discrepancy might be explained by the novice operators releasing the object without sufficient height information.

3.5.2.6 Practical applicability of the results

This research experimentally investigated optimal and allowable camera placements for digging and releasing tasks, which require depth perception. Therefore, the results are relevant to the camera placement systems [3.8, 3.9] used in real digging and releasing tasks. Depth perception in unmanned construction is demanded in tasks such as steel slit placements, concrete block masonry for dams, and sandbag placements. Furthermore, construction companies working on unmanned construction sites can refer to the results for deciding camera placement. For instance, the cameras for digging can be installed at a pan angle of 90° (the optimal pan angle for skilled operators) when the optimal area and its neighborhood are free of obstacles. When obstacles lie close to the optimal area, the cameras for digging can be installed at pan angles of 60°–90°, the allowable range for skilled operators. However, additional experiments are necessary for other work states such as transporting and reaching. Individual differences among the participants were observed even in the skilled operators; for example, the digging times at pan and tilt angles of 60° and 90° respectively were 17.0 s for one participant and 47.6 s for another participant. Therefore, additional experiments are required to determine the optimal and allowable camera placements for each individual.

3.5.3 Summary

The author derived the optimal and the allowable pan and tilt angles for skilled operators during digging and releasing tasks. During digging, the optimal pan angle was 90°, the allowable pan-angle range was 60–90°, and the allowable tilt angle was > 45°. During releasing, the allowable pan-angle range was 45°–135°, the optimal tilt angle was 90°, and the allowable tilt-angle range was 60–90°. These results are relevant to manipulation tasks requiring depth perception. Construction companies working on unmanned construction can refer to these results for optimizing their camera placements.

3.6 Summary

In this chapter, the author investigated the optimal and allowable camera placements in experiments using a scale model and an actual machine. The author’s hypotheses were based on the characteristics of object recognition by humans, which is optimized in the canonical view.

Table 3.2 Summary of all results in this chapter

			Scale (Novice)	Actual (Novice)	Actual (Skilled)
Grasp	Pan	Optimum	90°	60°	90°
		Allowable	90°±30°	45–90°	60–90°
	Tilt	Optimum	60°	60°	45°
		Allowable	60°±30°	45–90°	45–90°
Release	Pan	Optimum	60°	45°	45°
		Allowable	30–135°	45–135°	45–135°
	Tilt	Optimum	60°	60°	90°
		Allowable	60°±30°	45–90°	60–90°

First, the experiments were conducted on a scale model with novice operators. In the second set of experiments, the novice participants operated an actual machine. Finally, the experiments were conducted on an actual machine with skilled operators. The results are summarized in Table 3.2.

Chapter 4: Visual Interface to Decrease Cognitive Tunneling

This chapter describes the visual interface developed to decrease cognitive tunneling. First, the author describes the problems of cognitive tunneling in teleoperations and the necessity of reducing these problems by a visual interface. Next, the visual interface is developed based on the characteristics of visual momentum and visual saliency, which are the causes of cognitive tunneling. Finally, the proposed interface is evaluated in experiments.

Some of the sentences, figures and tables in this chapter are borrowed from the author's previously published works [4.1, 4.2].

4.1 Problems of cognitive tunneling and necessity of a visual interface

This section describes the problems of cognitive tunneling in teleoperations and the necessity of reducing these problems by a visual interface.

Excess information, such as too many views on a display, can cause "cognitive tunneling" in teleoperators, meaning that the teleoperators focus on a limited area and ignore other areas [4.3]. Teleoperators mainly watch the cab view; they rarely watch other external views [4.4]. To work at high efficiency, teleoperators should select different views depending on the work states because the advisable information differs among the work states [4.5]. For example, teleoperators need to observe a bird's-eye view for moving the construction machine because this view gives the machine's position. In contrast, grasping requires detailed views to satisfy the high precision demands of this task. Cognitive tunneling may cause teleoperators to select incorrect views, especially when their attention is already fixed on a specific view. For instance, if teleoperators fix their attention on a cab view, they can hardly discern the depth of grasping. In this case, the work efficiency may be degraded by problems such as overshoots of the manipulator controls. Teleoperators may stop controlling the manipulators while attempting to discern the depth by watching a cab view (see Fig. 4.1). Therefore, the work time can be increased by cognitive tunneling.

Several studies have investigated the cognitive tunneling problem [4.6–4.8]. Cognitive tunneling can be reduced by employing cooperators for unmanned aerial vehicles [4.6], or by

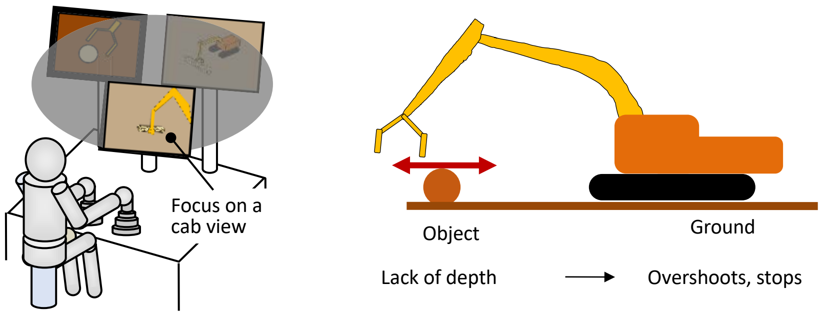


Fig. 4.1 Decrease in work efficiency caused by cognitive tunneling

introducing a wider field of view in robotic telepresence systems [4.7]. Cognitive tunneling is averted by installing only one display for unmanned ground vehicles [4.8]. However, the tasks of these researches mainly involve movement by imprecise operations that do not require external views. Solving the cognitive tunneling problem using multiple displays with a robot teleoperator has not been attempted.

In summary, the view systems should select appropriate visual information from multiple views while avoiding cognitive tunneling, which includes the fixation of the teleoperator's gaze. In this study, the author has developed a visual interface that decreases cognitive tunneling and enhances the ability of teleoperators to select views relevant to their work states. Using this interface, teleoperators can increase the efficiency of their unmanned construction operations.

4.2 Development of a visual interface to decrease cognitive tunneling

This section first describes the causes of cognitive tunneling, then describes the visual interface that alleviates cognitive tunneling.

4.2.1 Causes of cognitive tunneling

The two main causes of cognitive tunneling are low visual momentum and high visual saliency [4.9].

4.2.1.1 Low visual momentum

Visual momentum defines the degree to which the eye can transit among various views, and it involves the integration of information acquired through eye transitions [4.10]. When humans switch gazing views, they expend additional mental effort in acquiring information from the novel view and adding this information to their existing mental representation. The mental effort of switching views decreases on interfaces with high visual momentum. Fig. 4.2(b) shows a visual interface with high visual momentum displaying the environment shown in Fig. 4.2(a). For comparison, Fig. 4.2(c) shows a visual interface with low visual momentum, which will likely cause cognitive tunneling. Teleoperators are required to switch views depending on the work states and to integrate the information obtained from several views to discern the work sites [4.5]. Therefore, low visual momentum can lead to low-efficiency work because teleoperators have difficulty switching their views and integrating the information acquired from the switched views. In systems with high visual momentum, they can easily switch views and integrate the new information, and thus improve their work efficiency. Accordingly, view systems should be designed to increase the visual momentum.

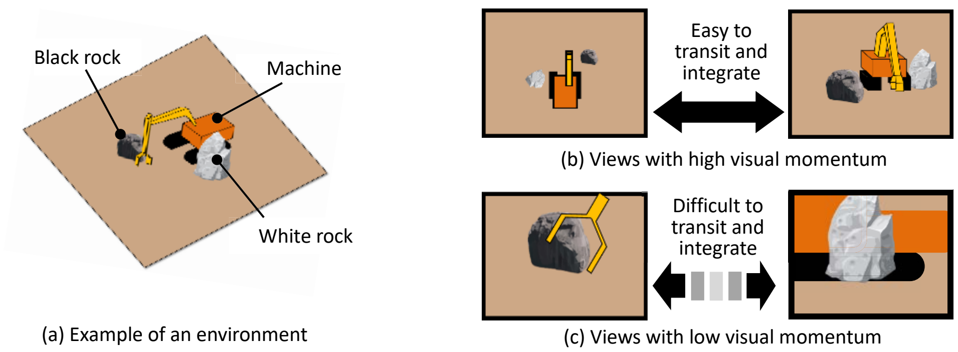


Fig. 4.2 Examples of view systems with high and low visual momentum

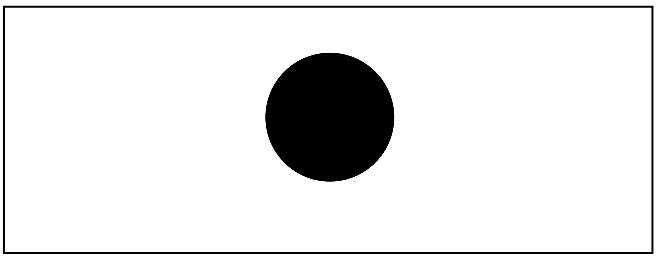


Fig. 4.3 Example of high visual saliency

4.2.1.2 Focusing on views with high visual saliency

Visual saliency describes the ease with which humans can focus on a specific area in a scene. Humans tend to focus on areas with relatively high visual saliency, and ignore other areas and views [4.11]. For instance, in Fig. 4.3, a person would likely focus on the black circle, which has high visual saliency. In this research, visual saliency is defined as a measure of the additional attention dedicated to one area relative over other areas [4.11]. The view with high visual saliency in unmanned construction is the cab view, the main focus of teleoperators [4.4]. Watching only a cab view can increase the work time, as described in Section 4.1. Therefore, view systems that lead teleoperators to switch their attention from the cab view (with relatively high visual saliency) to external views (with relatively low visual saliency) can decrease the work time [4.12]. According to the above analysis, view systems should be designed to release the teleoperators' attention from the cab view, which has relatively high visual saliency.

4.2.2 Visual interface with low cognitive tunneling

4.2.2.1 Increasing visual momentum

Low visual momentum can cause cognitive tunneling and increase the work time of teleoperators, as explained in Section 4.2.1.1. To mitigate this problem, the author develops a high visual momentum view system.

Fig. 4.4 shows the dimensions of visual momentum (the author created by selecting elements with relevance to teleoperation cases from [4.10]). If the views are changed completely, the view system lacks visual momentum (total replacement in Fig. 4.4). The autonomous camera control system for unmanned construction is always displayed in fixed format (enlarged views at the left of the interface, a cab view in the bottom center, and overlooking views at the right), as shown in Fig. 1.24 [4.13]. Such a fixed format can have visual momentum. A long shot provides an overview of the format of the view systems. Therefore, teleoperators do not need to remember the format, such as the enlarged views to the left in the case of [4.13]. Perceptual landmarks are to include the same landmarks in the views. Teleoperators can immediately discern the relationship between two views containing the same landmarks [4.14]. An overlap occurs when one view is part of another view, meaning that both views include the same scenes.

High visual momentum promotes work efficiency. However, overlapping can increase the number of views, increasing the cognitive load of determining which views to watch. Thus, the author chose the perceptual landmarks displaying the same landmarks. Table 4.1 shows the

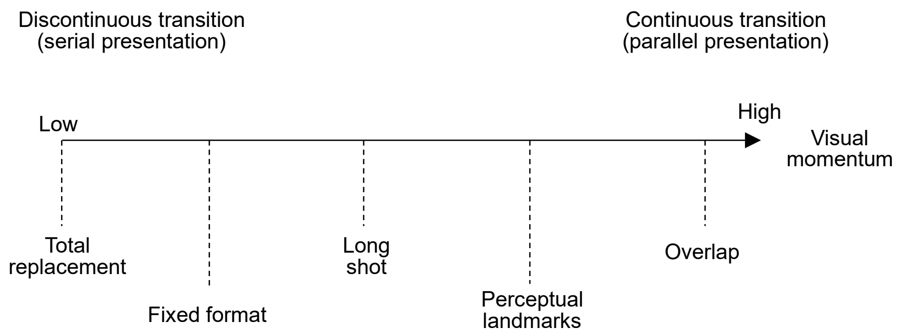


Fig. 4.4 Dimensions of visual momentum (created by the author with reference to [4.10])

landmarks in unmanned construction, which differ in each work state [4.15]. The visual interface should display the same landmarks that are suitably selected for each work state.

Next, the author considered two view types, namely, the bird's-eye view and detailed view, for displaying the landmarks. During movement, the teleoperators tend to obtain information from the bird's-eye views captured from above the object and diagonal to the environment [4.5]. In contrast, during manipulation tasks such as grasping and releasing, they tend to acquire information from the detailed views captured from the side of the target objects [4.5]. Therefore, the view systems were designed to display two bird's-eye views; one including the target objects, the construction machinery, and the obstacles above and diagonal to the environment during movement and transportation (see Fig. 4.5(a) and (c)), and a detailed sideways view during grasping and releasing (Fig. 4.5(b) and (d)).

4.2.2.2 Attracting attention to views with low visual saliency

Fixing attention on a specific view with high visual saliency (such as a cab view) without watching other views can reduce the work efficiency, as discussed in Section 4.2.1.2. Therefore, during unmanned construction, view systems should allow teleoperators to switch their attention from a cab view to other views [4.4].

Visual information inside the effective field of view (ranging by 30° to the right and left) can be acquired by eye movements alone, without any head movement (see Fig. 4.6) [4.16]. Objects vibrating at a specific frequency (5 Hz) inside the effective field of view also tend to attract a person's attention [4.17]. Therefore, the external views were displayed in the effective field of view, and were vibrated at 5 Hz for 0.5 s to attract the viewer's attention (see Fig. 4.6). The vibration time was kept short (0.5 s) because information is not easily extracted from vibrating views.

Table 4.1 Landmarks in each work state

(a) Move	(b) Grasp	(c) Transport	(d) Release
Machine	End-effectors	Machine	End-effectors
Obstacles	Target object	Obstacles	Grasped object
Target object		Grasped object	Release area
		Release area	

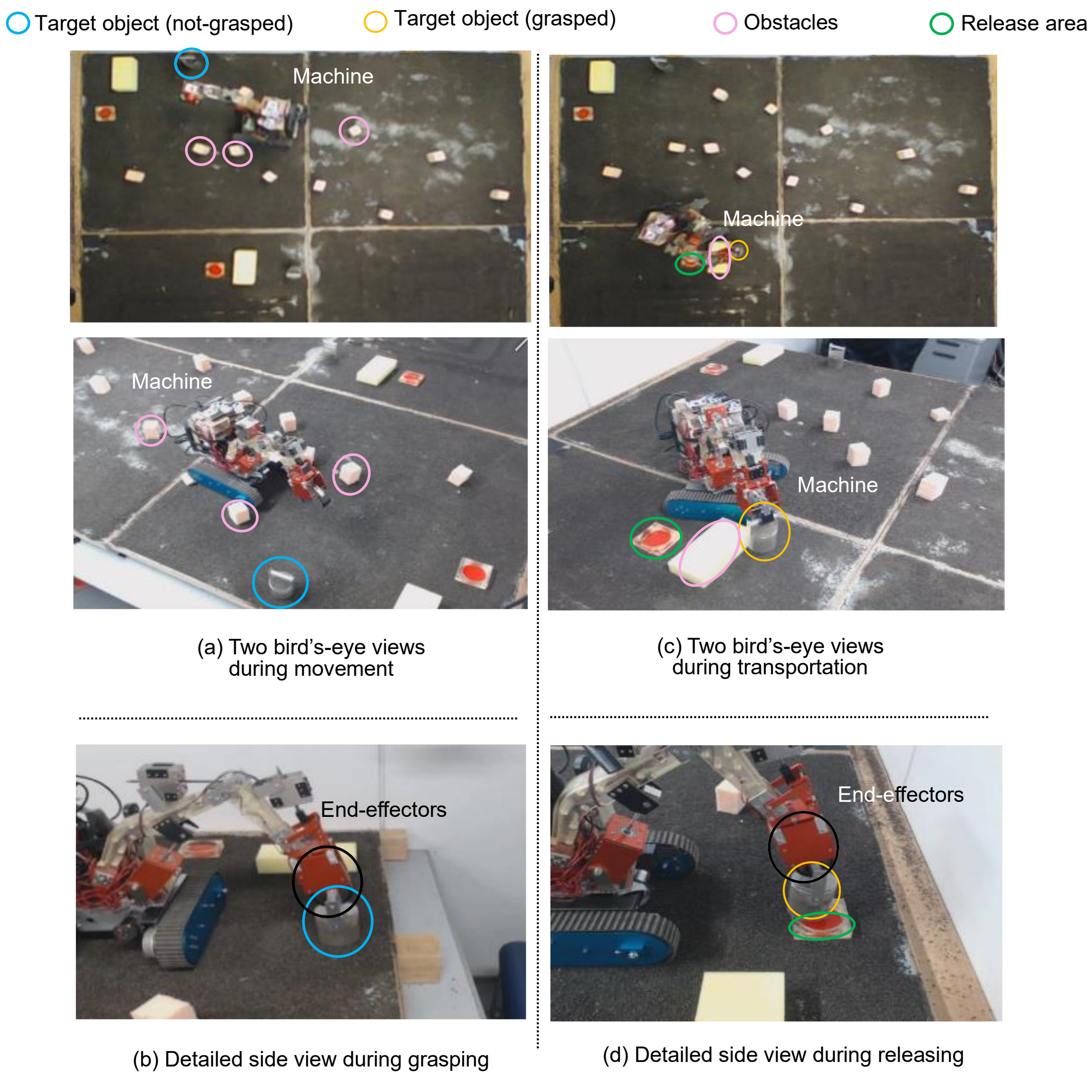


Fig. 4.5 Multiple views of each work state

4.2.2.3 Development of a visual interface

Fig. 4.7 shows the proposed visual interface. Skilled construction machinery operators watch at 107° horizontally and 56° vertically when controlling a construction machinery for on-board operation [4.18]. Thus, the proposed visual interface always allows teleoperators to watch a wide

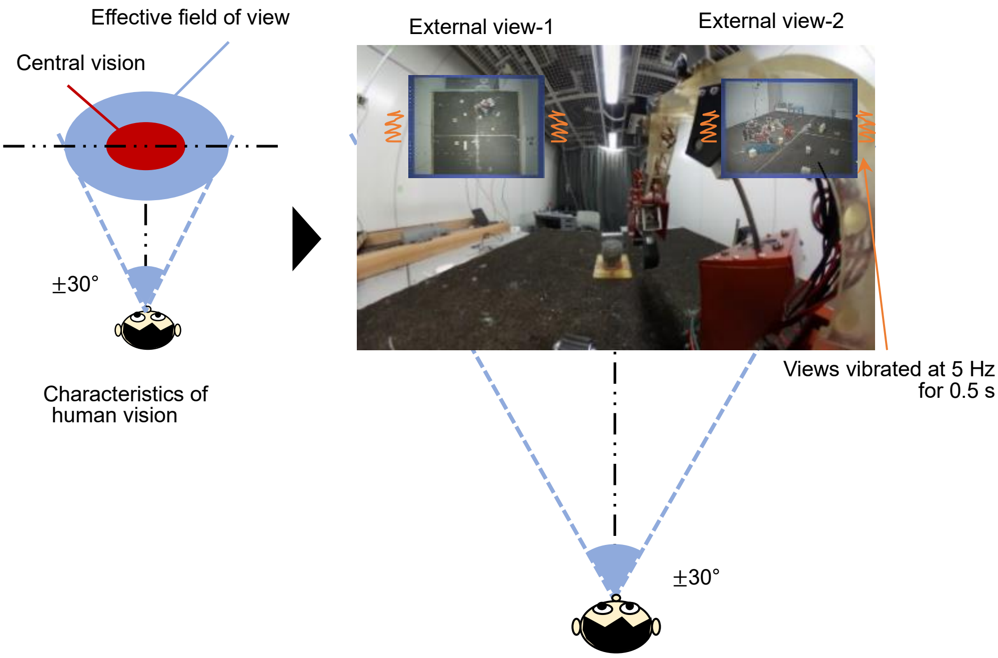


Fig. 4.6 Displaying external views based on the characteristics of human vision

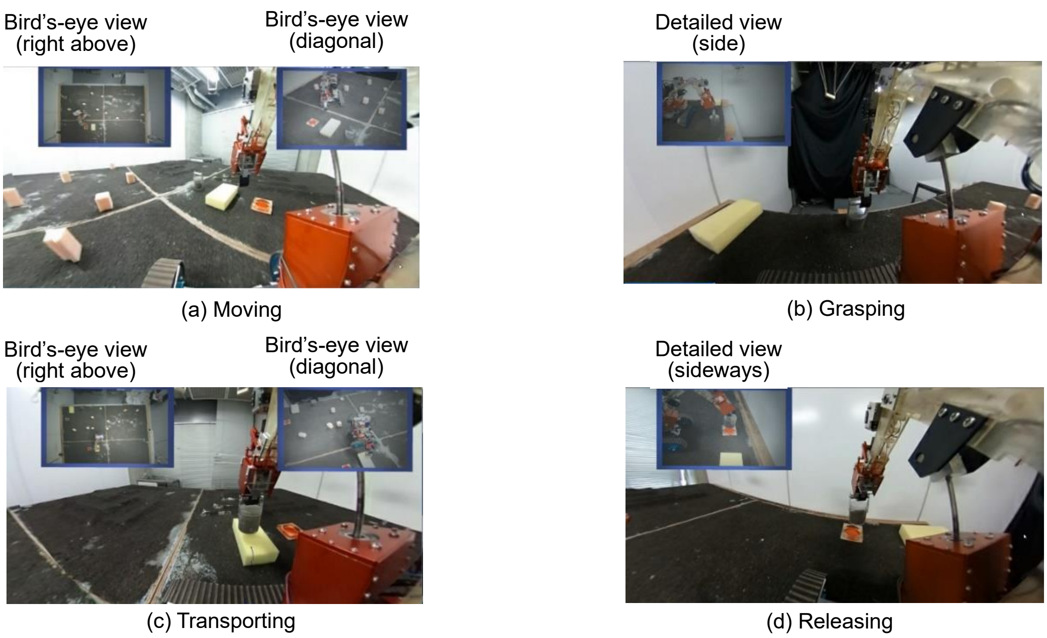
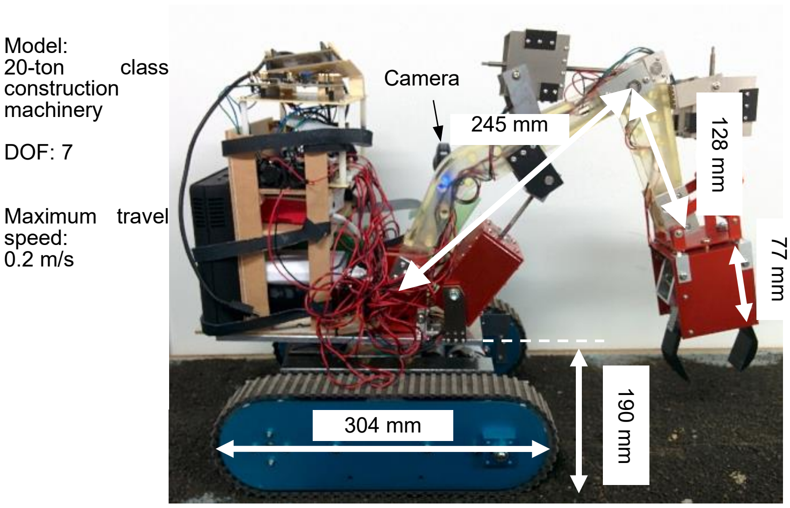


Fig. 4.7 Views displayed in each work state by the proposed visual interface

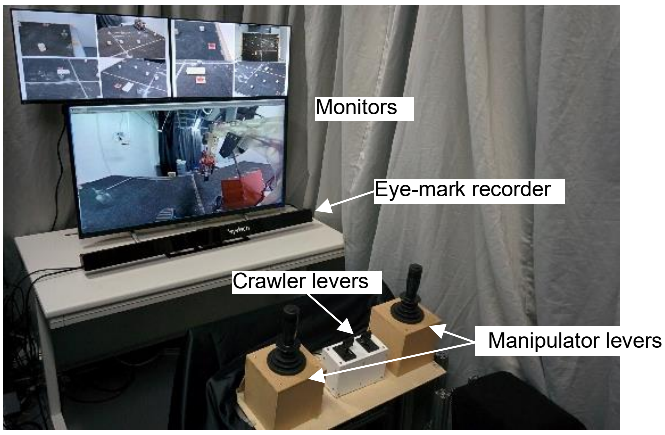
cab view with a horizontal view angle exceeding 107° and a vertical view angle exceeding 56°, which fully covers on-boarding operations. Furthermore, the proposed visual interface presents external views in the effective field of view corresponding to each work state and vibrates them at 5 Hz for 0.5 s. Fig. 4.6 shows the displayed views in each work state including the movements, the grasps, the transportation, and the release.

4.3 Experiment

The proposed visual interface was verified in scale-model experiments (Fig. 4.8 (a)). In a scale model, experiments are easily conducted in real environments. The participants teleoperated a construction machinery using two manipulator levers and crawler levers in this scale model, while



(a) Construction machinery in the scale model



(b) Interface of the scale model

Fig. 4.8 Experimental setup: 1/20 scale model watching the views on monitors (see Fig. 4.8 (b)).

4.3.1 Experimental setup

The crawlers were driven by two motors (Maxon Motor RE-max 29). The boom, arm, and bucket were driven by stepper motors (Nippon Pulse Motor, PJPL4233D4). The grapples were also driven by a stepper motor (Nippon Pulse Motor, PFCL25-48Q4C). A Raspberry Pi 3 Model B was used for signal communication. Two microcomputer boards (Nucleo F401RE) controlled the stepper motors and other motors. In the construction machinery of the scale model, visual communication was accomplished by a stick PC (Intel® Compute Stick STK2mv64CC, max turbo frequency: 2.80 GHz, memory: 4 GB). Wide cab views were provided by a camera (Theta S, resolution: full HD, frame rate: 29.97 fps, field of view: 360°). The external cameras were eight web cameras (Logicool C922 PRO STREAM WEBCAM, resolution: 1080 p, frame rate: 30 fps, field of view: 78°). These cameras were connected to eight laptop PCs (LIFEBOOK A500/B, OS: Ubuntu 14.04, max turbo frequency: 2.66 GHz, memory: 4 GB) through USB cables. The proposed visual interface was developed by Unity. The Unity interface and display were implemented on a laptop PC (NEXTGEAR-NOTE i5910GA2, OS: Windows, max turbo frequency: 3.50 GHz, memory: 32 GB). As all communication was local, the latency could be ignored. A desktop PC (OptiPlex 9020, OS: Ubuntu 14.04, max turbo frequency: 3.90 GHz, memory: 8 GB) was used for the control levers. The scale model was sized 2400 mm × 1600 mm. The experimental field contained two target objects to be grasped (diameter: 71 mm, height: 120 mm), two designated release areas (red circles of diameter 75 mm), and twelve obstacles made of sponge (see Fig. 4.9). Three monitors were provided; two monitors for external views (23.8 inch), and one monitor for a cab view (40 inch), as shown in Fig. 4.10.

4.3.2 Experimental procedure

As only 20 skilled operators of construction machinery reside in Japan, eight novice participants (students) with no experience in controlling construction machinery were recruited for this study [4.19]. However, after pre-training, the participants obtained sufficient skills to teleoperate the construction machinery in the scale model. The procedures of this experiment were approved by the ethics committee for human research at Waseda University. The experimental tasks were determined by referring to model tasks in unmanned construction [4.20]. During the experiment, the participants were required to move along the designated routes (Movement 1 with

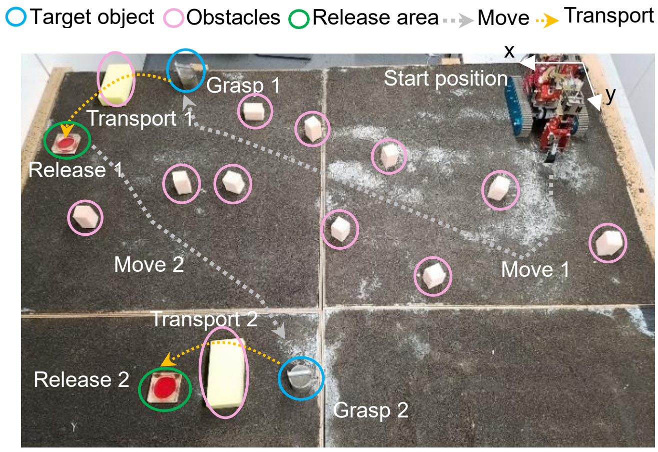


Fig. 4.9 Experimental environment and task

Fig. 4.10 Conventional viewing system for comparison

a bend and Movement 2 with a narrow (< 20 mm) tolerance range). The two target objects (Grasps 1 and 2) were to be grasped with precise teleoperations. The grasped objects were then expected to be transported (Transports 1 and 2) without contacting any obstacles. Finally, the objects were to be released in their designated areas (Release areas 1 and 2), also by precise teleoperations (see Fig. 4.8). The participants were asked to try the tasks as quickly as possible without making any contact with the obstacles in the environments.

The eight participants were divided into two groups: the control group (CG) and intervention group (IG); by assigning four participants to each group, we ensured almost the same average task times in the training tasks. The average time of the training task was 126.9 s (SD = 14.5 s) in the CG, and 130.0 s (SD = 28.8 s) in the IG. The Welch's t-test revealed no significant differences between the two groups. Therefore, the teleoperation skill levels of the two groups were quite similar. The four participants in the CG tried the experimental tasks with a conventional view system, which provided a wide cab view and eight external views (see Fig. 4.9). In contrast, the

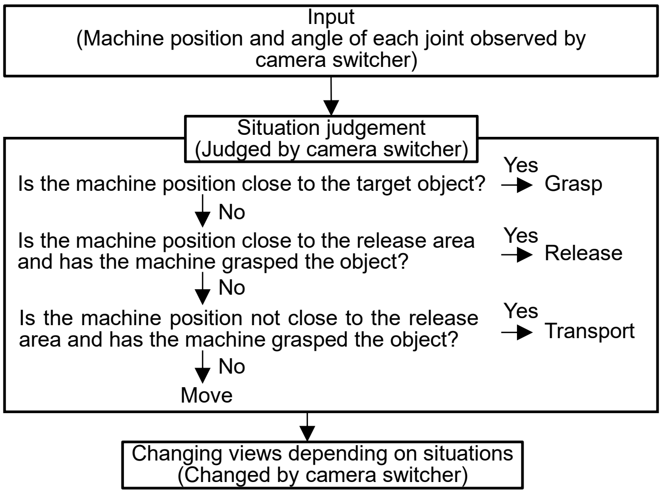


Fig. 4.11 System flow of the camera switcher

four participants in the IG tried the experimental tasks with the proposed visual interface shown in Fig. 4.6. All participants tried the tasks thrice, and their work times, stoppage times, and eye-marks were recorded.

The author played the role of the camera switcher and identified the work states based on the positions and joint angles of the construction machinery. The external views corresponding to each work state were switched via the mouse wheel of a wireless mouse. Fig. 4.11 shows the procedures for changing the external views, as performed by the camera switcher. The camera placements of all eight cameras were identical in the CG and IG, ensuring that the participants in both groups received the same external views.

4.3.3 Results

4.3.3.1 Work time

Fig. 4.12 compares the results of the required average total task times and the task times in each work state between the CG and IG. The total task time differed between the two groups (Welch's t-test, $t(19) = 2.02$, $p = 0.06$), and the significance was marginal. Significant differences were observed in the grasping times (Welch's t-test, $t(18) = 2.75$, $p = 0.01$) and releasing times (Welch's t-test, $t(22) = 2.08$, $p = 0.05$). These results indicate that the proposed visual interface reduced the total task time, the grasping time, and the release time. However, the Welch's t-test revealed no significant differences in the movement ($t(17) = 1.25$, $p = 0.23$) and transportation ($t(16) = 0.27$, $p = 0.79$) times. This finding is discussed in detail below.

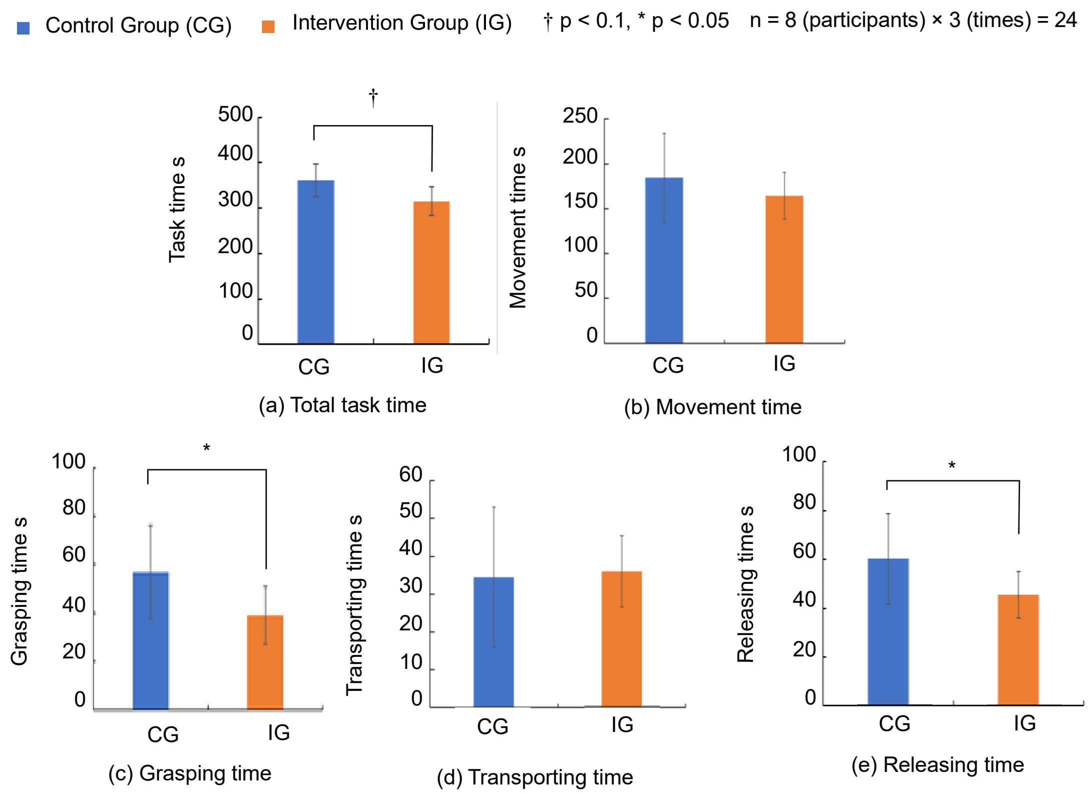


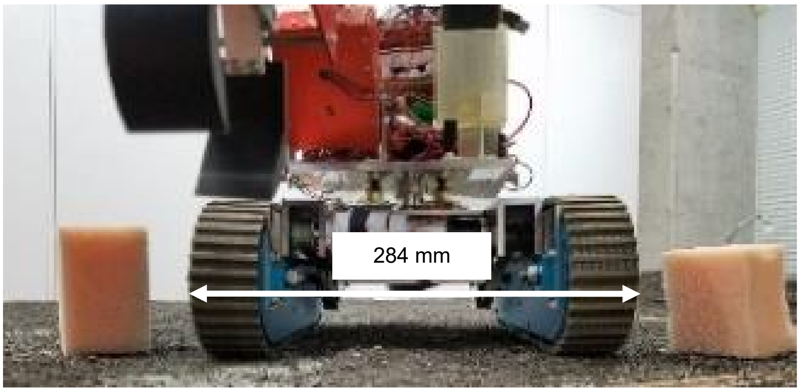
Fig. 4.12 Comparison of total task times and task times of each work state in the control and intervention groups

1) Movement times: Two moves (Movement 1 and Movement 2) were assigned in the experimental tasks. Fig. 4.13 (a) shows the tolerance range of Move 1, which was wider than 100 mm, and Fig. 4.13 (b) shows the tolerance range of Move 2, which was narrower than 20 mm. Fig. 4.14 shows the movement times of Moves 1 and 2. A marginally significant difference between the two groups was observed in Move 2 (Welch's t-test, $t(19) = 1.85$, $p = 0.08$), but not in Move 1 ($t(21) = 0.29$, $p = 0.77$). These results indicate that the proposed visual interface can improve the movement time along a route with a narrow tolerance range, which requires precise teleoperation.

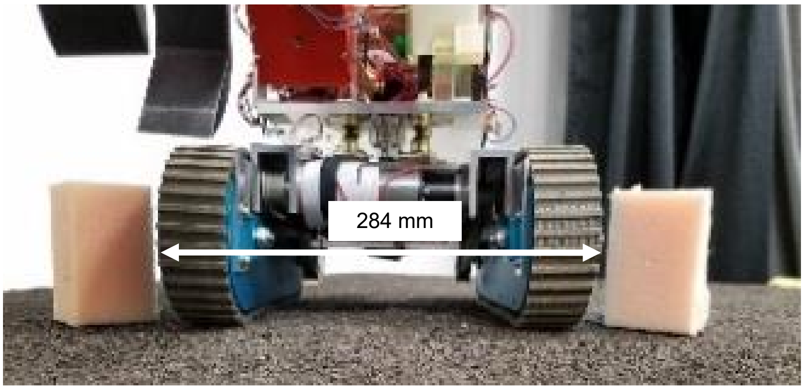
2) Transportation time: In the experimental tasks, the transportation needed to avoid only one obstacle placed around the transportation area (Fig. 4.9). As this task required few precise teleoperations, the proposed visual interface little improved the transportation time.

4.3.3.2 Stoppage time

Fig. 4.15 compares the results of the total required average total stoppage times and the stoppage time of each work state in the CG and IG. Significant differences were observed in the total stopping time (Welch's t-test, $t(22) = 2.51$, $p = 0.02$) and in the grasping time (Welch's t-test,

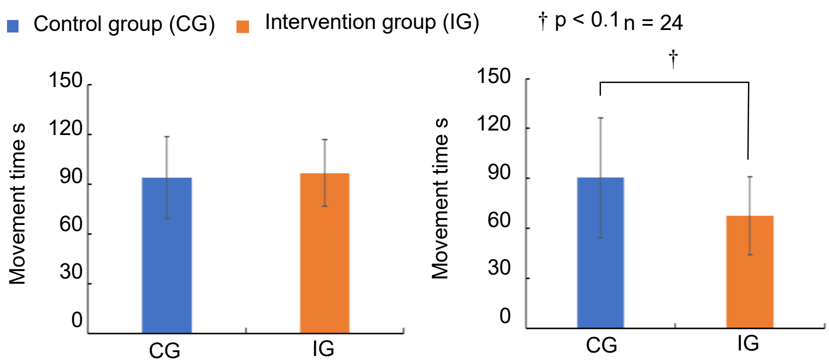


(a) Tolerance range wider than 100 mm (Move 1)



(b) Tolerance range narrower than 20 mm (Move 2)

Fig. 4.13 Tolerance ranges of Moves 1 and 2



(a) Move 1 (b) Move 2

Fig. 4.14 Movement times of Moves 1 and 2 in the control and intervention groups

$t(19)= 3.16$, $p = 0.01$), proving that the proposed visual interface can improve the total-stoppage and grasping times. However, the Welch's t-test indicated no significant differences in the times of movement ($t(22) = 1.63$, $p = 0.11$), transportation ($t(15) = 1.64$, $p = 0.12$), and release ($t(22) = 1.49$, $p = 0.15$). Reasons for these observations are discussed below. The lack of improvement in the stoppage time of transportation is probably explained by the non-necessity of precise teleoperations, as discussed for the task time of transportation.

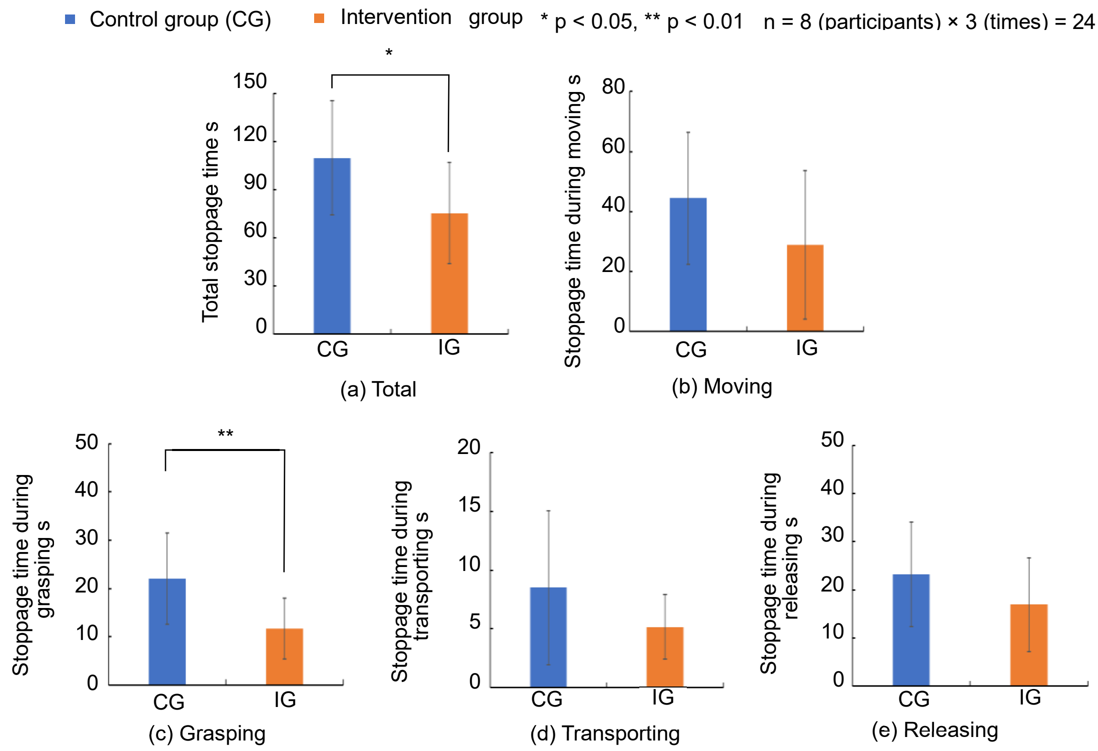


Fig. 4.15 Results of stoppage time

1) Movement: The experimental tasks included two moves (Movement 1 with a wide tolerance range and Movement 2 with a narrow tolerance range), as shown in Fig. 4.9. Fig. 4.16 compares the stoppage times of the CG and IG in Movement 1 and Movement 2. A marginally significant difference was observed in Move 2 (Welch's t-test, $t(22) = 1.86$, $p = 0.08$), but no significant difference appeared in Move 1 (Welch's t-test, $t(21) = 1.08$, $p = 0.29$). Those results indicate that the proposed visual interface can improve the stoppage time during movement along a route with a narrow tolerance range, which requires precise teleoperations.

2) Releasing: The stoppage time of release did not significantly differ between the groups, possibly because of occlusions. Fig. 4.17 shows an external view displayed during the release task. The occlusions largely prevented the participants from recognizing the entire designated release area. Therefore, they may have released the grasped target objects in an incorrect location because they could not discern whether the objects would fit into the designated release area. The imprecise deposition of the released objects would obscure any significant differences between the groups.

4.3.3.3 Eye marks

Fig. 4.18 compares the eye-marks between the CG and IG. The data of four participants (two

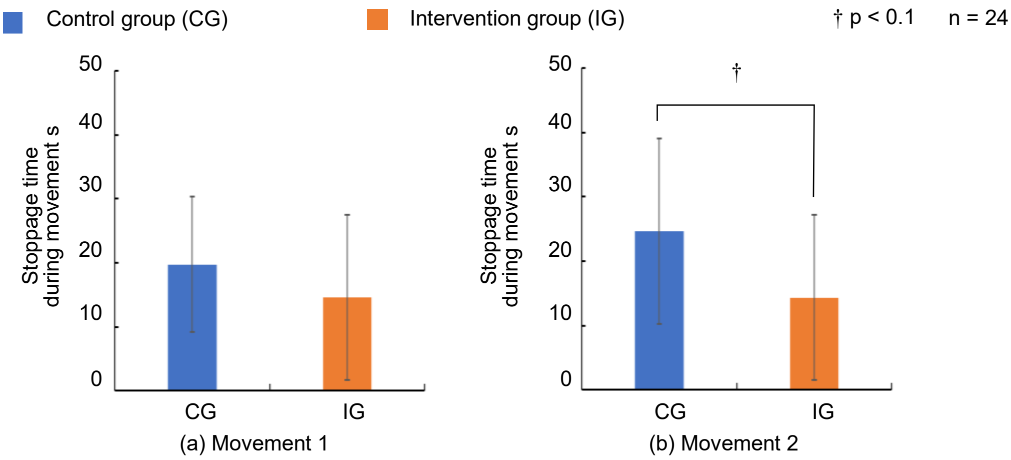


Fig. 4.16 Stoppage times during Moves 1 and 2 in the control and intervention groups


Fig. 4.17 Occlusions in the releasing task

from each group) were analyzed because their error percentage exceeded 50% of the total work time. The visual momentum of the proposed visual interface was verified by two measures: (i) the percentage of fixing times exceeding 1.5 s in one view, and (ii) the number of view switches within 1 s. The results of measures (i) and (ii) are displayed in Fig. 4.18(a) and Fig. 4.18(b), respectively. Significant differences in both measures were observed between the two groups (Welch's t-test, fixing time: $t(7) = 21.9$, $p < 0.01$; number of view switches: $t(5) = 20.4$, $p < 0.01$). These results prove that the participants in the IG shared their attention among various views and often switched their views. Therefore, the proposed visual interface had a high visual momentum. Fig. 4.18(c) shows the number of times the views were switched from the cab view in 1 s. The number of view switches from the cab view significantly differed between the CG and IG (Welch's t-test, $t(4) = 15.5$, $p < 0.01$). These results prove that the proposed visual interface released the teleoperators' attention from the cab view with high visual saliency.

All of the eye-mark results proved that the proposed visual interface reduces cognitive

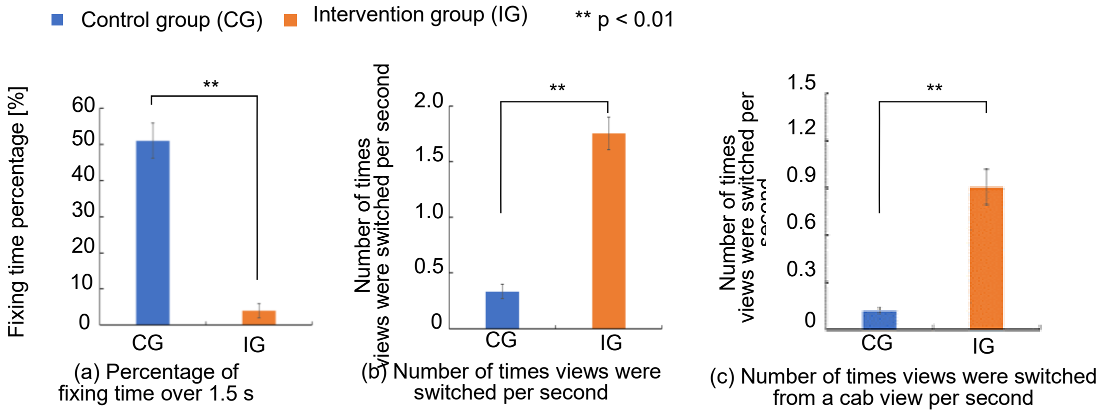


Fig. 4.18 Comparison of eye-marks in the control and intervention groups

tunneling; in particular, the interface has a high visual momentum and can release the teleoperator’s attention from the cab view with high visual saliency.

4.3.4 Discussion

4.3.4.1 Differences between the conventional and proposed view interfaces

In this subsection, the author first reviews the results obtained for various task times and eye-marks, then analyzes the differences between the proposed and conventional visual interfaces. The experimental results proved that the proposed visual interface can improve the teleoperators’ ability to switch views frequently, and can release their attention from a cab view with high visual saliency. Furthermore, the proposed visual interface improved the completion time of tasks requiring precise teleoperations, such as grasping, releasing, and moving (Move 2).

The conventional visual interface can trap teleoperators in cognitive tunneling, fixing their attention on specific views. Cognitive tunneling can reduce the work efficiency, especially in tasks requiring precise teleoperations. To remove cognitive tunneling, the proposed visual interface displays appropriate external views of each work state within the teleoperators’ effective field of view, and vibrates the external views for a short period when they first appear. The proposed visual interface enables higher work efficiency than the conventional visual interface, especially in tasks requiring precise teleoperations.

4.3.4.2 Habituation problems

Repeating the same stimulus can cause habituation, meaning a reduced response to the stimulus

[4.21]. The author vibrates the external views to attract the teleoperators' attention. Habituation to the vibration was not observed during the experiments, but could occur during eight hours of unmanned construction work. Thus, measures against habituation, such as changing the stimulus by adding red frames or sound, should be considered in the future.

4.3.4.3 Practical use of the proposed visual interface

1) Practically switching the external views corresponding to each work state: As the camera switcher, the author identified the work states and switched the views during the experiments. In unmanned construction work, camera switchers are usually required to modify views. Therefore, the proposed visual interface is accessible to real-life situations. However, unmanned construction cannot be introduced if skilled people are lacking [4.19]. Thus, the author should develop some systems to identify work states, such as those in [4.22].

2) Ways of obtaining external views in real life: In unmanned constructions, external views are provided by camera dollies. Tether-powered drones [4.23] or image processing techniques [4.24] also provide external views. Furthermore, image processing provides teleoperators with clear views even in low-visibility situations, such as foggy weather conditions [4.25]. Therefore, external views can be obtained in real-life situations.

4.3.4.4 Effects on skilled operators

Skilled operators tend to fix their attention on the cab view and ignore other views [4.4]. Thus, the results of the experiments using skilled operators might be similar to those using novice operators. The proposed visual interface will allow skilled operators to avoid cognitive tunneling, thereby increasing their work efficiency. However, interviews with the skilled teleoperators in Unzen-Fugendake revealed that they can work without external views. Thus, the proposed visual interface may not remarkably assist skilled teleoperators who can work with cab views alone.

4.4 Summary

The author developed a visual interface that discourages cognitive tunneling in teleoperation. The causes of cognitive tunneling are (i) low visual momentum and (ii) excessive focus on a view with high visual saliency. To improve the visual momentum, the author displayed external views with the same landmarks relevant to each work state. The author also displayed external views within the effective field of view and vibrated them at the specific frequency that attracts human operator’s attention (5 Hz), thus releasing the teleoperator’s attention from the cab view with high visual saliency. Experiments were conducted on a scale model to verify the proposed visual interface. The proposed visual interface was found to improve the task time and stoppage time, and helped the teleoperators to release their attention from the cab view to other views at high frequency.

Chapter 5: Conclusions

5.1 Summary

This study addressed three technical challenges: (i) developing a view system that inputs environmental information in advance, (ii) investigating an optimum and allowable camera placement for manipulation tasks, and (iii) developing a visual interface that avoids cognitive tunneling. The effectiveness of the developed view systems was evaluated in a simulator, a scale model, and an actual machine. The novelty of this study is that (i) it has a view system to input information by displaying survey and route views before work in advance, (ii) it proposes an optimal and allowable possible position for camera placement, (iii) it develops a visual interface displays external views in the cab views, thereby increasing the visual momentum and attracting the teleoperators to views with low visual saliency. The five chapters of this study are overviewed below.

Chapter 1 summarized the unmanned construction system, the problems associated with unmanned construction, and the causes of low efficiency (a crucial problem in unmanned construction operations). The importance of visual information to high work efficiency was also discussed, along with previous studies on visual information and their limitations. The purpose of the study was stated in this chapter.

Chapter 2 developed the view system that can be seen by teleoperators in advance. The environmental information input into the system in advance is based on the characteristics of cognitive maps, which are broadly composed of survey and route knowledges. Survey knowledge can be acquired from external viewpoints, whereas route knowledge can be acquired from personal or internal viewpoints. The survey perspective is obtained as a third-person view from an arbitrary viewpoint, because appropriate viewpoints differ among teleoperators. Meanwhile, the route perspective is acquired by a subjective view that can be changed by the teleoperators, because this knowledge is generally gained through active movements. The experimental results indicated that the proposed prior view system can improve the quality and quantity of the cognitive maps of important landmarks, including the target objects to be grasped. Therefore, the proposed system can improve the efficiency of planning. A view that provides the survey perspective enables general planning, as it improves the perception of the movement distance. Meanwhile, a view that acquires the route perspective enables partial planning, because the route view improves the speed of grasping. As operators sometimes forget their planned paths and trajectories, the author further developed an AR reminder that improves the work efficiency and eases the cognitive load.

Chapter 3 investigated the optimal and allowable camera placement for manipulation tasks. External views are essential even when teleoperators can watch wide and 3D cab views. Thus, an optimal and allowable camera placement was investigated based on the characteristics of the canonical view (which gives the best performance in object recognition. Canonical views are characterized by minimal occlusion, an allowable rotation range of $\pm30°$, and minimal sensitivity on object sizes. The optimal pan and tilt angles were thus hypothesized as 90°, which provides the canonical view, and the allowable pan and tilt angles were hypothesized as $\pm30°$ to match the rotation angles that give the same viewing acuity as the canonical view. To determine the optimal and allowable camera placements for task manipulation, experiments were conducted on a scale model and an actual machine using novice and skilled teleoperators as participants. The experimental results were summarized, and deemed suitable for optimizing camera placement in actual unmanned construction operations.

Chapter 4 developed a visual interface that avoids cognitive tunneling, defined as a focus on specific views with a tendency to ignore other views. Cognitive tunneling occurs (i) when teleoperators focus on views with high visual saliency, and (ii) when there is low visual momentum (that is, when information is not easily integrated through view transitions). The author developed visual interface to increase the visual momentum and attract the teleoperators' eyes to views with lower visual saliency. The visual momentum can be increased by displaying the same landmarks in different views. To this end, views with same landmarks were displayed in each work state. Moreover, because humans tend to respond to objects vibrating at a specific frequency (5 Hz) in their effective field of view ($\pm30°$), the external views were briefly vibrated at 5 Hz in the effective field of view of the teleoperator at every change of the work state. The experimental results proved that the proposed view system can decrease cognitive tunneling and improve the work efficiency in tasks requiring precise operations, such as grasping.

The second part of this chapter discusses the applicability of the proposed systems to real-life situations, and the third section proposes future directions of the proposed systems.

5.2 Practical use of the proposed system

This section discusses the practical applicability of the proposed system. Fig. 5.1 shows the use of the proposed system in real-life situations. First, the system requires environmental data including 3D maps of the disaster sites and simulators for the view system. The 3D maps are available in the literature [5.1, 5.2], and simulators for unmanned construction have been developed [5.3]. The proposed prior view systems are compatible with these simulators in real-

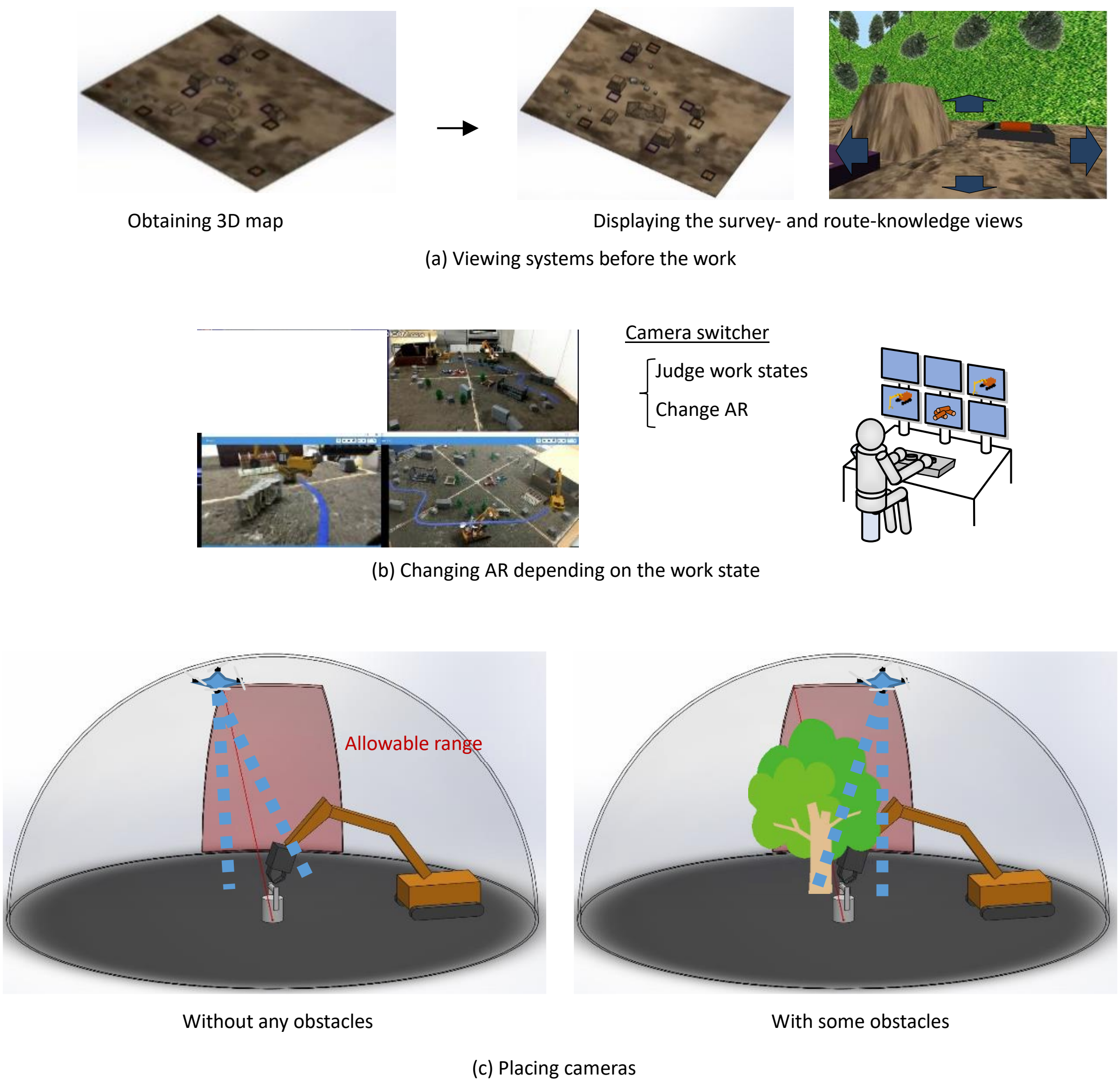


Fig. 5.1 Real-life applications of the proposed view system

life situations. Thus, teleoperators of works requiring long-distance movements such as compacted concrete can watch the survey view and improve their general planning. Moreover, teleoperators of works including manipulation tasks such as earthworks and box culverts, can watch the route view and improve their local planning.

The AR reminder system can use the data of current point clouds at the disaster sites and work states. Point clouds can be obtained from existing sensors, such as LiDAR sensors, and work states can be obtained from previous studies [5.4–5.6]. Thus, this AR reminder system can be applied at unmanned construction work sites involving compacted concrete, earthworks, and box culverts.

The cameras can be installed at the investigated optimal and allowable positions in tasks requiring manipulation, such as concrete block masonry for dams, placing sandbags, and placing box culverts. For example, an external camera can be installed at the optimal position that averts all obstacles, or in an allowable place when obstacles occlude the target in the optimal place (see Fig. 5.1 (c)).

Finally, the visual interface that avoids cognitive tunneling is available in works requiring precise operations such as steel slit and earthwork. This visual interface can be automated in work-state identification systems [5.4–5.6].

5.3 Future directions

The proposed low-cognitive view systems for teleoperators are based on human cognition characteristics. They are applicable to unmanned construction, as discussed in Section 5.2. This section describes the limitations of the proposed systems and their future research directions.

The view system enables teleoperators to acquire spatial knowledge prior to work commencement. In this study, the impact of the proposed prior view system was verified in simulator experiments. This system requires further verification in a scale model and an actual machine operated by skilled subjects. Furthermore, the interface should modify the views in an intuitive way.

The optimal and allowable camera placement was derived for task manipulation. In future work, a similar derivation is required for other work states, such as movement and transportation. Path planning methods for drones, including determining the required number of cameras, should also be considered in future research.

In the visual interface that avoids cognitive tunneling, the work states were identified by a camera switcher. By integrating existing work-state identification systems [5.4–5.6], the work states could be judged automatically. Finally, additional experiments using actual machinery with skilled operators are important for assessing the real-life applicability of this interface.

The proposed view systems can be applied to other teleoperation fields, such as in space and in the deep sea. Some teleoperated robots in these environments are required to move and grasp objects such as stones. The proposed prior view systems can be useful for such places because the survey and route knowledge can increase work efficiency of teleoperation. Also, the optimal and allowable camera placement can enhance work efficiency. Furthermore, the proposed view interface can help teleoperators maintain watching appropriate views on each work state.

References

Chapter 1

[1.1] Cabinet Office, Government of Japan, "Disaster management in Japan," 2018. URL: http://www.bousai.go.jp/1info/pdf/saigaipamphlet_je.pdf
Accessed on 8/27/2019

[1.2] Cabinet Office, Government of Japan, "White paper on disaster management", 2010. URL: http://www.bousai.go.jp/kaigirep/hakusho/h22/index.htm
Accessed on 8/27/2019

[1.3] 無人化施工協会，URL: http://www.kenmukyou.gr.jp/about/
Accessed on 8/27/2019

[1.4] Y. Ban, "Unmanned construction system", *Proceedings of International Symposium on Automation and Robotics in Construction*, pp. 241–246, 2002.

[1.5] K. Chayama, A. Fujioka, K. Kawashima, H. Yamamoto, Y. Nitta, C. Ueki, A. Yamashita, and H. Asama, "Technology of unmanned construction system in Japan," *Journal of Robotics and Mechatronics*, vol. 26, no. 4, pp. 403–417, 2012.

[1.6] 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," *Proceedings of International Symposium on Automation and Robotics in Construction*, pp. 508–513, 2011.

[1.7] M. Moteki and H. Yamamoto, "Restoration activities which is quick and safety by unmanned construction," *Journal of the society of instrument and control engineers*, Vol. 55, No. 6, pp. 495–500, 2016.

[1.8] S. Suzuki, Y. Miura, T. Sunakawa, Y. Kihara, H. Ueno, H. Murakami, and Y. Hayase, "Development of a shallow water construction carrier by altering a common carrier," *The proceedings of JSME Annual Conference on Robotics and Mechatronics*, 2A2-07a, 2016.

[1.9] M. Umakakeba, K. Nishimura, S. Sakamoto, and S. Tano, "Applicability of amphibious bulldozers in sediment dredging at the end of dam reservoirs," *Journal of Japan Society of Dam Engineers*, Vol 24, No. 1, pp. 17–24, 2014.

[1.10] 三村充，猪原幸司，"無線遠隔操縦式水陸両用ブルドーザの活用状況"，建設の施工企画，Vol. 702，pp. 71–76，2008.

[1.11] Y. Hiramatsu, T. Aono, and M. Nishio, "Disaster restoration work for the eruption of Mt Usuzan using an unmanned construction system," *Advanced Robotics*, Vol. 16, No. 6, pp. 505–508, 2002.

[1.12] M. Moteki, A. Nishiyama, K. Fujino, and S. Yuta, "Working hour measurement of the hydraulic excavator by the unmanned construction", *Civil Engineering Journal*, Vol.57, No. 4, pp. 16–19, 2015.

[1.13] M. Moteki, S. Yuta, K. Fujino, "The proposal of the model task for efficiency evaluation of the construction work by remote control of a hydraulic excavator," *Journal of JCMA*, Vol. 66, No. 8, pp. 71–79, 2014.

[1.14] H. Asama, et al., "Proposal for unmanned construction," *Council on Competitiveness-Nippon*, 2013.

[1.15] A. Macintyre, J. Barbera, and E. Smith, "Surviving collapsed structure entrapment after earthquakes: a "time to rescue" analysis," *Prehospital and Disaster Medicine*, Vol. 21, No. 1, pp. 4–17, 2006.

[1.16] K. Gurusamy, K. Samraj, C. Gluud, E. Wilson, and BR. Davidson, "Meta-analysis of randomized controlled trials on the safety and effectiveness of early versus delayed laparoscopic cholecystectomy for acute cholecystitis," *British Journal of Surgery*, Vol. 97, No. 2, pp. 141–150, 2009.

[1.17] T. Mayumi, T. Takada, Y. Kawarada, Y. Nimura, M. Yoshida, et al., "Results of the Tokyo consensus meeting Tokyo guidelines," *Journal of Hepato-Biliary-Pancreatic Surgery*, Vol. 14, No. 1, pp. 114–121, 2007.

[1.18] 産経 WEST ，"阿蘇大橋崩落現場で無人の重機が土砂を撤去",
https://www.sankei.com/west/news/160423/wst1604230033-n1.html
Accessed on 9/14/2019

[1.19] T. Fong, C. Thorpe, and C. Baur, "Multi-robot remote driving with collaborative control," *IEEE Transactions on Industrial Electronics*, Vol. 50, No. 4, pp. 699–704, 2003.

[1.20] 山口崇，吉田正，石松豊．"遠隔操作におけるマンマシンインターフェイスに関する実態調査"，土木学会第 59 回年次学術講演会，pp. 373–384，2004.

[1.21] M. Moteki, N. Akihiko, S. Yuta, H. Mishima, and K. Fujino, Research on improving work efficiency of unmanned construction, *Proceedings of International Symposium on Automation and Robotics in Construction*, pp. 478–486, 2016.

[1.22] D. D. Woods, J. Tittle, M. Feil, and A. Roesler, "Envisioning human-robot coordination in future operations," *IEEE Transactions on System, Man, Cybernetics, Part C (Applications and Reviews)*, Vol. 34, No. 2, pp. 210–218, 2004.

[1.23] A. M. Okamura, “Methods for haptic feedback in teleoperated robot‐assisted surgery,” Industrial Robot, Vol 31, No. 6, pp. 499–508, 2004.

[1.24] T. Fong, F. Conti, S. Grange, and C. Baur, “Novel interfaces for remote driving: gesture, haptic and pda,” *SPIE Telemanipulator and Telepresence VII*, 2000

[1.25] T. Hirabayashi, T. Yamamoto, H. Yano, and H. Iwata, “Experiment on teleoperation of underwater backhoe with haptic information,” *Proceedings of International Symposium on Automation and Robotics in Construction*, pp. 36–41, 2006.

[1.26] K. Hayashi, and T. Tamura, “Teleoperation performance using excavator with tactile feedback,” *IEEE International Conference on Mechatronics and Automation*, pp. 2759–2764, 2009.

[1.27] Y. Nitta, O. Matsuo, S. Kitahara, N. Kuroda, K. Tamura, and T. Shimoda, “A study of the application of ultra-long-range remote-controlled construction technology - A summary of ultra-long-range remote control experiment in Mt. Fugen in Unzen -," *The 13h Symposium on Construction Robotics in Japan*, pp. 41–50, 2012.

[1.28] Y. Li, Y. Yin, and D. Zhang, “Adaptive task-space synchronization control of bilateral teleoperation systems with uncertain parameters and communication delays,” *IEEE Access*, Vol. 6, pp. 5740–5748, 2018.

[1.29] Z. Cehn, Y. Pan, and J. Gu, “Integrated adaptive robust control for multilateral teleoperation systems under arbitrary time delays,” *International Journal of Robust and Nonlinear Control*, Vol. 26, No. 12, pp. 2708–2728, 2018.

[1.30] B. Jafari, and M. Spong, “Passivity-based switching control in teleoperation systems with time-varying communication delay,” *American Control Conference*, pp. 5469–5475, 2018.

[1.31] 土木研究所, “災害発生時における無人化施工技術の最新動向―土木研究所での無人化施工の作業効率向上に関する実験見学及び最新技術の紹介―”, 2016.

[1.32] T. Hirabayashi, “Human machine interface for underwater excavator,” *Journal of the Robotics Society of Japan*, Vol. 33, No. 6, pp. 412–415, 2015.

[1.33] H. Yamada, “Human interface of a tele-operated construction machine,” *Journal of the Robotics Society of Japan*, Vol. 33, No. 6, pp. 400–403, 2015.

[1.34] M. Draper, G. Calhoun, H. Ruff, D. Williamson, and T. Barry, “Manual versus speech input for unmanned aerial vehicle control station operations,” Proceedings of Human Factors Ergonomics Society 47th Annual Meeting, pp. 109–113, 2003.

[1.35] D. Perzanowski, D. Brock, W. Adams, M. Bugajska, A. Schultz, J. Trafton, S. Blisard, and M. Skubic, “Finding the FOO: A pilot study for a multimodal interface," 2003 IEEE Conference on Systems, Man, Cybernetics, pp. 3218–3223.

[1.36] C. Hu, M. Q. Meng, P. X. Liu, and X. Wang, "Visual gesture recognition for human–machine interface of robot teleoperation," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1560–1565, 2003.

[1.37] M. Heilig, "The cinema of the future," *Presence*, Vol. 1, No. 3, pp. 279–294, 1992.

[1.38] R. Murphy, "Human-robot interaction in rescue robotics," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 34, No. 2, pp. 138–153, 2004.

[1.39] H. Yanco, and J. Drury, "Where am I? acquiring situational awareness using a remote robot platform," *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, pp. 2835–2840, 2004.

[1.40] D. 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.

[1.41] H. Furuya, N. Kuriu, and C. Shimizu, "Development of next-generation remote-controlled machinery system," *Report of Obayashi Corporation technical Research Institute*, Vol. 76, pp. 1–10, 2012.

[1.42] T. Sato, H. Fujii, A. Moro, K. Sugimoto, A. Nozue, et al., "Development of bird's-eye view system in unmanned construction," *Transactions of the JSME*, Vol. 81, No. 823, pp. 1–13, 2015.

[1.43] K. Nagatani, K. Yakushigawa, S. Kiribayashi, and A. Watanabe, "Status display of unmanned construction machine for initial response of sediment disasters -1st report: Initial experiments of information acquisition system by multi-rotor UAV with a wired power feeding system mounted on an unmanned construction machine," *The proceedings of JSME Annual Conference on Robotics and Mechatronics*, 1P1-10a6, 2016.

[1.44] M. Fuchida, S. Chikushi, A. Moro, A. Yamashita, and H. Asama, "Arbitrary viewpoint visualization for teleoperation of disaster response robots," Journal of Advanced Simulation in Science and Engineering, Vol. 6, No. 1, pp. 249-259, 2019.

[1.45] R. Hojo, S. Nakata, Y. Hada, T. Asuma, and S. Kitahara, "Proposal and evaluation of multi-viewpoint generating method using omnidirectional past images," in *Proceedings of the 19th symposium on construction robotics in Japan*, paper no. O1-5, 2019.

[1.46] 阪田顕，中田真颯，北條怜，棟本真弘，羽田靖史，竹下嘉人，飛鳥馬翼，北原成郎，"無人クローラーダンプの後方から追従する映像の生成実験", 第３７回日本ロボット学会学術講演会，2A1-06，2019.

[1.47] M. Kamezaki, J. Yang, H. Iwata, and S. Sugano, "Visibility enhancement using autonomous multicamera controls with situational role assignment for teleoperated work machines," *Journal of Field Robotics*, Vol. 33, No. 6, pp. 802-824, 2016.

[1.48] H. Yamada, N. Bando, K. Ootsubo, and Y. Hattori, "Teleoperated construction robot using visual support with drones," *Journal of Robotics and Mechatronics*, Vol. 30, No. 3, pp. 406–415, 2018.

[1.49] J. Yang, M. Kamezaki, R. Sato, H. Iwata, and S. Sugano, "Inducement of visual attention using augmented reality for multi-display systems in advanced tele-operation," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5364–5369, 2015.

[1.50] K. Ootsubo, T. Kawamura, and H. Yamada, "Construction tele-robotics system with AR presentation," *Journal of Physics*, Vol. 433, pp. 1–8, 2013.

[1.51] T. Tanimoto, R. Fukano, K. Shinohara, H. Yoshinada, K. Kurashiki, and D. Kondo, "Research on effective teleoperation of construction machinery with 3D environmental map and superimposed image," *The proceedings of JSME Annual Conference on Robotics and Mechatronic*s, 2P1-O06, 2015.

[1.52] B. G. Witmer and W. J. Sadowski, "Nonvisually guided locomotion to a previously viewed target in real and virtual environments," Human Factors, Vol. 40, pp. 478–488, 1998.

[1.53] D. R. Scribner and J. W. Gombash, "The effect of stereoscopic and widefield of view conditions on teleoperator performance," U.S. Army Res. Lab., Aberdeen Proving Ground, MD, Tech. Rep. ARL-TR-1598, 1998.

[1.54] A. Oving and J. van Erp, "Driving with a head-slaved camera system," Proceedings of the Human Factors Ergonomics Society Annual Meeting, pp. 1372–1376, 2001.

[1.55] H. Nagahara, Y. Yagi, H. Kitamura, and M. Yachida, "Super wide view tele-operation system," *Proceedings of IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pp. 149–154, 2003.

[1.56] S. Pazuchanics, "The effects of camera perspective and field of view on performance in teleoperated navigation," *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*, pp. 1528–1532, 2006.

[1.57] D. Drascic, and J. Grodski, "Using stereoscopic video for defense teleoperation," *Stereoscopic Displays and Applications IV*, Vol. 1915, pp. 58–69, 1993.

[1.58] D. Scribner, J. W. Gombash, "The effect of stereoscopic and wide field of view conditions on teleoperator performance," *U.S. Army Research Laboratory, Aberdeen Proving Ground*, MD, Technical Report ARL-TR-1598, 1998.

[1.59] J. V. Draper, S. Handel, and C. C. Hood, "Three experiments with stereo-scopic television: When it works and why," in *Proceedings of 1991 IEEE International Conference on System, Man, Cybernetics*, pp. 1047–1052, 1991.

[1.60] B. Keshavarz and H. Hecht, "Stereoscopic viewing enhances visually induced motion sickness but sound does not," PRESENCE: Teleoperators and Virtual Environments, Vol. 21, No. 2, pp. 213–228, 2012.

[1.61] S. Tachi, K. Komoriya, K. Sawada, T. Itoko, and K. Inoue, "Telexistence control cockpit system for HRP," *Journal of the Robotics Society of Japan*, Vol. 19, No. 1, pp. 16–27, 2001.

[1.62] T. Kot, and P. Novak, "Utilization of the Oculus Rift HMD in mobile robot teleoperation," *Applied Mechanics and Materials*, Vol. 555, pp. 199–208, 2014.

[1.63] H. Martins, and R. Ventura, "Immersive 3-D teleoperation of a search and rescue robot using a head-mounted display," in *Proceedings of 2009 IEEE Conference on Emerging Technologies & Factory Automation*, pp. 1–8, 2009.

[1.64] S. Sharples, S. Cobb, A. Moody, and J. R. Wilson, "Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems," Displays, Vol. 29, No. 2, pp. 58–69, 2008.

[1.65] R. Meier, T. Fong, C. Thorpe, and C. Baur, "A sensor fusion based user interface for vehicle teleoperation," in *Proceedings of International Conference on Field Service Robot*, 1999.

[1.66] T. Fong, C. Thorpe, and C. Baur, "Advanced interfaces for vehicle teleoperation: collaborative control, sensor fusion displays, and remote driving tools," *Autonomous Robotics*, Vol. 11, pp. 77–85, 2001.

[1.67] S. Livatino, F. Banno, G. Muscato, "3-D integration of robot vision and laser data with semiautomatic calibration in augmented reality stereoscopic visual interface," IEEE Transactions on Industrial Informatics, Vol. 8, No. 1, pp. 69–77, 2012.

[1.68] G. Dirkin, "Cognitive tunneling: Use of visual information under stress," *Perceptual and Motor Skills*, Vol. 56, pp. 191–198, 1983.

[1.69] Learn.Genetics, URL: https://learn.genetics.utah.edu/content/memory/distraction/, accessed on 1/12/20

[1.70] Y. Nitta, "Evolution of unmanned construction system playing important role in the disaster recovery," *Civil Engineering*, Vol. 67, No. 4, pp. 16–23, 2012 (in Japanese)

[1.71] R. Sato, M. Kamezaki, S. Sugano, and H. Iwata, "Gaze pattern analysis in multi-display systems for teleoperated disaster response robots," *Proceedings of IEEE*

International Conference on Systems, Man and Cybernetics, pp. 3534–3539, 2016.

Chapter 2

[2.1] Ryuya Sato, Mitsuhiro Kamezaki, Satoshi Niuchi, Shigeki Sugano, and Hiroyasu Iwata, “A Pre-offering View System for Teleoperators of Heavy Machines to Acquire Cognitive Maps,” *The 15th IEEE International Symposium on Safety, Security and Rescue Robotics,* pp. 61–66, 2017.

[2.2] Ryuya Sato, Mitsuhiro Kamezaki, Satoshi Niuchi, Shigeki Sugano, and Hiroyasu Iwata, “Pre-offering Work-Site Views to Acquire Spatial Knowledge from Survey and Route Perspective for Advanced Teleoperation of Construction Machines,” *The 34th International Symposium on Automation and Robotics in Construction,* paper no. P05, 2017.

[2.3] Ryuya Sato, Mitsuhiro Kamezaki, Shigeki Sugano, and Hiroyasu Iwata, “A Basic Framework of View Systems Allowing Teleoperators to Pre-Acquire Spatial Knowledge from Survey and Route Perspectives,” *Presence: Virtual and Augmented Reality*. (under review)

[2.4] Ryuya Sato, Mitsuhiro Kamezaki, Satoshi Niuchi, Shigeki Sugano, and Hiroyasu Iwata, “A view system based on human’s cognition to provide spatial knowledge for teleoperators of disaster response heavy machines,” *Proceedings of 17th Symposium on Construction Robotics in Japan*, paper no. O3-2, 2017.

[2.5] Ryuya Sato, Mitsuhiro Kamezaki, Satoshi Niuchi, Shigeki Sugano, and Hiroyasu Iwata, “Investigation of effects on pre-offering images from operator’s viewpoint to acquire spatial knowledge for teleoperated heavy-duty machines,” *Proceedings of JSME Annual Conference on Robotics and Mechatronics*, paper no. 1P2-A02, 2017.

[2.6] Ryuya Sato, Mitsuhiro Kamezaki, Satoshi Niuchi, Shigeki Sugano, and Hiroyasu Iwata, “Cognitive map analysis of operators using sketch maps by a view system before introducing teleoperation,” *Proceedings of SICE System Integration Division Annual Conference*, paper no. 3C4-03, 2017.

[2.7] M. Moteki, N. Akihiko, S. Yuta, H. Mishima, and K. Fujino, Research on improving work efficiency of unmanned construction, *Proceedings of International Symposium on Automation and Robotics in Construction*, pp. 478–486, 2016.

[2.8] S. Kiribayashi, K. Yakushigawa, and K. Nagatani, “Design and development of tether-powered multirotor micro unmanned aerial vehicle system for remote-controlled construction machine,” *Field and Service Robotics*, Vol. 5, pp. 637–648, 2018.

[2.9] M. Kamezaki, J. Yang, H. Iwata, and S. Sugano, “Visibility enhancement using autonomous multicamera controls with situational role assignment for teleoperated work machines," *Journal of Field Robotics*, Vol. 33, No. 6, pp. 802-824, 2016.

[2.10] S. Casner, "The effect of GPS and moving map displays on navigational awareness while flying under VFR," *Journal of Applied Aviation Studied*, Vol. 5, No. 1, pp. 153-165 2005.

[2.11] J. Wang, M. Lewis, and S. Hughes, "Gravity-referenced attitude display for teleoperation of mobile robots," *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting*, pp. 2662-2666, 2004.

[2.12] J. Draper, S. Handel, and C. Hood, "Three experiments with stereoscopic television: When it works and why," *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, pp. 1047-1052, 1991.

[2.13] R. Passini, "Spatial representations, a wayfinding perspective," *Journal of Environmental Psychology*, Vol. 4, No. 2, 153–164, 1984.

[2.14] G. Burnett, "A road-based evaluation of a head-up display for presenting navigation information," *Proceedings of International Conference on Human-Computer Interaction*, Vol. 3, pp. 180–184, 2003.

[2.15] W. Narzt, G. Pomberger, A. Ferscha, D. Kolb, R. Muller, et al, "Augmented reality navigation systems," *Universal Access in the Information Society*, Vol. 4, No. 3, pp. 177–187, 2005.

[2.16] F. Nex, and F. Remondino, "UAV for 3D mapping applications: a review," *Applied Geomatics*, Vol. 6, No. 1, 1–15, 2013.

[2.17] M. Heiling, and R. Futuro, "The cinema of the future," *Presence*, Vol. 1, No. 3, pp. 279–294, 1992.

[2.18] D. Crundall, and G. Underwood, "Effects of experience and processing demands on visual information acquisition in drivers," *Ergonomics*, Vol. 41, No. 4, 448–458, 1998.

[2.19] T. Edward, "Cognitive Maps in Rats and Men. Psychological Review," Vol. 55, No. 4, pp. 189-208, 1948.

[2.20] R. Golledge, and R. Stimson, R "Spatial behavior: A geographic perspective," *New York: Guilford Press*, 1997.

[2.21] M. Levin, I. Jankovic, and M. Palij, "Principles of spatial problem solving," *Journal of Experimental Psychology: General*, Vol. 111, No. 2, pp. 157-175, 1982.

[2.22] R. Sato, M. Kamezaki, S. Sugano, and H. Iwata, "Gaze pattern analysis in multi-display systems for teleoperated disaster response robots," in *Proceedings of 2016 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 3534–3539, 2016.

[2.23] M. Cadwallader, "A behavioral model of consumer spatial decision making," *Journal of Economic Geography*, Vol. 51, No. 4, pp. 339-349, 1975.

[2.24] T. Garling, and R. Golledge, "Behavior and environment: Psychological and geographical approaches," *North Holland*, 1993.

[2.25] G. Miller, "The magical number seven, plus or minus two: some limits on our capacity for processing information," *Psychological Review*, Vol. 63, No. 2, pp. 81-97, 1956.

[2.26] M. Moteki, K. Fujino, T. Ohtsuki, and T. Hashimoto, "Research on visual point of operator in remote control of construction machinery," *Proceedings of the International Symposium Automation and Robotics in Construction*, pp. 532-537, 2011.

[2.27] M. Kamezaki, J. Yang, H. Iwata, and S. Sugano, "A basic framework of virtual reality simulator for advancing disaster response work using tele-operated work machines," *Journal of Robotics and Mechatronics*, Vol. 26, No. 4, pp. 486-495, 2014.

[2.28] Y. Nitta, "Evolution of unmanned construction system playing important role in the disaster recovery," *Civil Engineering*, Vol. 67, No. 4, pp. 16–23, 2012 (in Japanese)

[2.29] P. Curseu, R. Schalk, and S. Schruijer, "The use of cognitive mapping in eliciting and evaluation group cognitions," *Journal of Applied Social Psychology*, Vol. 40, No. 5, pp. 1258-1291, 2010.

[2.30] N. Huynh, and S. Doherty, "Digital sketch-map drawing as an instrument to collect data about spatial cognition," *Cartographica*, Vol. 42, No. 4, pp. 285-296, 2007.

[2.31] Y. Wakabayashi, and S. Itoh, "Reexamination of the components of distortions in cognitive maps," *Journal of Geography*, Vol. 103, No. 3, pp. 221-232, 1994.

[2.32] M. Blades, "The reliability of data collected from sketch maps," *Journal of Environmental Psychology*, Vol. 10, No. 4, pp. 327-339, 1990.

[2.33] M. Spranger, F. Heinke, S. Becker, and D. Labudde, "Towards drone-assisted large-scale disaster response and recovery," *Int. Conf. Advances in Computation, Communication and Services*, pp. 5–10, 2016.

[2.34] F. Nex, F. Remondino, "UAV for 3D mapping applications: a review," *Applied Geomatics*, vol. 6, no. 1, pp. 1–15, 2014.

[2.35] S. Tsuneyama, S. Takeshita, H. Tsutsumi, and M. Shirosawa, "The practical usage of i-Construction for enhancing the safety of the construction site," *Journal of Japan Society of Civil Engineers*, Ser. F6, Vol. 73, No. 2, pp. I_1–I_6, 2017.

[2.36] B. Walther-Franks, R. Malaka, "Evaluation of an augmented photograph-based pedestrian navigation system," *International Symposium on Smart Graphics*, pp. 94-105, 2008.

[2.37] M. Baumhauer, M. Feuerstein, H. Meinzer, et al, "Navigation in endoscopic soft tissue surgery: Perspectives and limitations," *Journal of Endourology*; Vol. 22, No. 4, pp. 751-766, 2008.

[2.38] C. Hansen, J. Wieferich, F. Ritter, et al, "Illustrative visualization of 3D planning models for augmented reality in liver surgery," *Computer Assisted Radiology and Surgery*, Vol. 5, No. 2, pp. 133-141, 2010.

[2.39] U. Yang, G. Kim, "Implementation and evaluation of "Just Follow Me": An immersive, VR-based, motion-training system," *Presence: Teleoperators & Virtual Environments*, Vol. 11, No. 3, pp. 304-323, 2002.

[2.40] J. Zauner, M. Haller, A. Brandl, et al, "Authoring of a mixed reality assembly instructor for hierarchical structures," *IEEE/ACM International Symposium on Mixed and Augmented Reality,* pp. 237, 2003.

[2.41] G. Dirkin, "Cognitive tunneling: Use of visual information under stress," *Perceptual and Motor Skills*, Vol. 56, pp. 191-198, 1983.

[2.42] K. Mizuno, M. Tanaka, K. Yamaguti, et al, "Mental fatigue caused by prolonged cognitive load associated with sympathetic hyperactivity," *Behavioural and Brain Functions,* Vol. 8, pp. 1-7, 2011.

[2.43] M. Kamezaki, H. Iwata, S. Sugano, "Operator support system based on primitive static states in intelligent operated work machines," *Advanced Robotics*, Vol. 23, No. 10, pp. 1281-1297, 2009.

[2.44] M. Kamezaki, H. Iwata, S. Sugano, "A symbolic construction work flow based on state transition analysis using simplified primitive static states," *Journal of Robotics and Mechatronics*, Vol. 24, No. 6, pp. 939-948, 2012.

[2.45] R. Azuma, "A survey of augmented reality," *Presence: Teleoperators and Virtual Environments*, Vol. 6, No. 4, pp. 355-385, 2002.

[2.46] Z. Shengli, Y. Ming, Z. Yubin, "Application of 3D animation in wushu routine teaching and training," *Journal of Physical Education*, Vol. 19, No. 1, pp. 90-92, 2012.

[2.47] K. Hayakawa, D. Hasegawa, H. Sakuta, "Drum training support by using first-person 3D animations and the effect of animations using experts' eye movements on training," *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics*, Vol. 28, No. 1, pp. 511-521, 2016.

[2.48] A. Tang, C. Owen, F. Biocca, et al, "Comparative effectiveness of augmented reality in object assembly," *SIGGHI Conference on Human Factors in Computing Systems,* pp. 73-80, 2003.

[2.49] R. Backs, K. Seljos, "Metabolic and cardiorespiratory measures of mental effort: the effects of level of difficulty in a working memory task," *International Journal of Psychophysiology*, Vol. 16, No. 1, pp. 57-68, 1994.

[2.50] L. Peterson and J. Peterson, "Short-term retention of individual verbal items," *J. Experimental Psychology*, vol. 58, no. 3, pp. 193–198, 1959.

[2.51] J. Sweller, "Cognitive load during problem solving: Effects on learning," Cognitive Science, vol. 12, pp. 257–285, 1988.

[2.52] S. Hart and L. Staveland, "Development of NASA-TLX (Task Load Index): Results of

empirical and theoretical research," *Advances in Psychology*, vol. 52, pp. 139–183, 1988.
[2.53] M. Kamezaki, H. Iwata, and S. Sugano, "Operator support system based on primitive static states in intelligent operated work machines," *Advanced Robotics*, vol. 23, no. 10, pp. 1281–1297, 2009.

Chapter 3

[3.1] Ryuya Sato, Mitsuhiro Kamezaki, Satoshi Niuchi, Shigeki Sugano, and Hiroyasu Iwata, "Derivation of an Optimum and Allowable Range of Pan and Tilt Angles in External Sideway Views for Grasping and Placing Tasks in Unmanned Construction Based on Human Object Recognition," *2019 IEEE/SICE International Symposium on System Integrations*, pp. 776–781, 2019.
[3.2] Ryuya Sato, Mitsuhiro Kamezaki, Mitsuru Yamada, Takeshi Hashimoto, Shigeki Sugano, and Hiroyasu Iwata, "Experimental Investigation of Optimum and Allowable Range of Side Views for Teleoperated Digging and Release Works by Using Actual Construction Machinery," *2019 IEEE/SICE International Symposium on System Integrations*, pp. 788–793, 2019.
[3.3] Ryuya Sato, Mitsuhiro Kamezaki, Mitsuru Yamada, Takeshi Hashimoto, Shigeki Sugano, and Hiroyasu Iwata, "Experimental Derivation of the Optimum and Allowable Environmental Camera Placement for Skilled Operators during Digging and Releasing in Unmanned Construction," *Automation in Construction*. (under review)
[3.4] Satoshi Niuchi, Ryuya Sato, Mitsuhiro Kamezaki, Shigeki Sugano, and Hiroyasu Iwata, " Research on Suitable Positions for a Multi-Camera System for Advanced Teleoperation of Construction Machineries -- Analysis on Effect for Work Efficiency by Pan and Tilt Angles of Cameras --," *Proceedings of Robotics Symposia*, pp. 267–272, 2018.
[3.5] Ryuya Sato, Mitsuhiro Kamezaki, Mitsuru Yamada, Takeshi Hashimoto, Shigeki Sugano, and Hiroyasu Iwata, "Research on Suitable Positions for a Multi-Camera System for Advanced Teleoperation of Construction Machineries -- Experimental Analysis on Work Efficiency by Pan and Tilt Angle in Digging and Releasing Work –," *Proceedings of Robotics Symposia*, pp. 14–17, 2018.
[3.6] Ryuya Sato, Mitsuhiro Kamezaki, Mitsuru Yamada, Takeshi Hashimoto, Shigeki Sugano, and Hiroyasu Iwata, "Derivation of the optimum and suitable external camera placement for manipulation tasks requiring depth sense in unmanned construction," *Proceedings of 19th Symposium on Construction Robotics in Japan*, paper no. O3-01, 2019.

[3.7] S. Ito, Y. Sakano, K. Fujino, and H. Ando, “Remote controlled construction equipment by using high-resolution stereoscopic 3D images,” *Japan Society of Civil Engineers*, Vol. 73, No. 1, pp. 15–24, 2017.

[3.8] S. Iwataki, H. Fujii, A. Moro, A. Yamashita, H. Asama, H. Yoshinada, “Visualization of the surrounding environment and operational part in a 3DCG model for the teleoperation of construction machines,” in *Proceedings of IEEE/SICE International Symposium on System Integration*, pp. 81–87 2015.

[3.9] S. Kiribayashi, K. Yakushigawa, and K. Nagatani, “Design and development of tether-powered multirotor micro unmanned aerial vehicle system for remote-controlled construction machine,” *Field and Service Robotics*, Vol. 5, pp. 637–648, 2017.

[3.10] R. Sato, M. Kamezaki, S. Sugano, and H. Iwata, “Gaze pattern analysis in multi-display systems for teleoperated disaster response robots,” in *Proceedings of 2016 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 3534–3539, 2016.

[3.11] S. Palmer, E. Rosch, and P. Chase, “Canonical perspective and the perception of objects”, *Attention and Performance IX*, pp. 135–151, 1981.

[3.12] N. Logothetis, J. Pauls, and T. Poggio, “Shape representation in the inferior temporal cortex of monkeys,” *Current.Biology*, Vol.5, No.5 pp. 552–563, 1995.

[3.13] A. Nishiyama, M. Moteki, and K. Fujino, “Research on the comparison of operator viewpoints between manned and remote control operation in unmanned construction of the hydraulic excavator,” *Proc. The 17th Symp. Construction Rob. in Japan*, pp. 63–70, 2014.

[3.14] K. Nagatani, K. Yakushigawa, S. Kiribayashi, and A. Watanabe, “Status display of unmanned construction machine for initial response of sediment disasters,” *JSME Conference on Robotics and Mechatronics*, paper no. 1P1–10a6, 2016.

[3.15] D. R. Lampton, M. J. Singer, D. McDonald, and J. P. Bliss, “Distance estimation in virtual environments,” in *Proceedings of Human Factors and Ergonomics Society 39th Annual Meeting*, 1995, pp. 1268–1272.

[3.16] B. G. Witmer and P. Kline, “Judging perceived and traversed distance in virtual environments,” *Presence: Teleoperators Virtual Environment*, vol. 7, no. 2, pp. 144–167, 1998.

[3.17] M. Moteki, S. Yuta, K. Fujino, “The proposal of the model task for efficiency evaluation of the construction work by remote control of a hydraulic excavator,” *J. JCMA*, vol. 66, no. 8, pp. 71–79, 2014.

[3.18] Y. Nitta, “Evolution of unmanned construction system playing important role in the disaster recovery,” *Civil Engineering*, vol. 67, no. 4, pp. 16–23, 2012 (in Japanese).

[3.19] M. Levin, I. Jankovic, and M. Palij, “Principles of spatial problem solving,” *J.*

Experimental Psychology: General, Vol. 111, No. 2, pp.157–175, 1982.

[3.20] N. Ahmidi, M. Ishii, G. Fichtinger, G. Gallia, G. hager, “An objective and automated method for assessing surgical skill in endoscopic sinus surgery using eye-tracking and tool-motion data,” *International Forum of Allergy & Rhinology*, Vol. 2, No. 6, pp. 507–515, 2012.

[3.21] Y. Maeda, S. Suzuki, H. Igarashi, K. Hidaka, “Evaluation of human skill in teleoperation system,” in *Proceedings of 2006 SICE-ICASE International Joint Conference*, pp. 5339–5344, 2006,

Chapter 4

[4.1] Ryuya Sato, Mitsuhiro Kamezaki, Satoshi Niuchi, Shigeki Sugano, and Hiroyasu Iwata, “Development of a View System to Induce Operators to Watch Suitable Views Depending on Work States in Unmanned Construction,” *Proceedings of SICE System Integration Division Annual Conference*, paper no. 2C1-03, 2018.

[4.2] Ryuya Sato, Mitsuhiro Kamezaki, Satoshi Niuchi, Shigeki Sugano, and Hiroyasu Iwata, “Development of a Cognitive Untunneling Multi-view System Based on Visual Momentum and Saliency for Teleoperators of Heavy Machines,” *Automation in Construction*. (under review)

[4.3] G. Dirkin, “Cognitive tunneling: Use of visual information under stress,” *Perceptual and Motor Skills*, Vol. 56, pp. 191–198, 1983.

[4.4] A. Nishiyama, M. Moteki, and K. Fujino, “Research on the comparison of operator viewpoints between manned and remote control operation in unmanned construction of the hydraulic excavator,” *Proceedings of International Symposium on Automation and Robotics in Construction*, pp. 677–684, 2013.

[4.5] R. Sato, M. Kamezaki, S. Sugano, and H. Iwata, “Gaze pattern analysis in multi-display systems for teleoperated disaster response robots”, *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, pp. 3534–3539, 2016.

[4.6] A. Burkle, F. Segor, M. Kollmann, “Towards autonomous micro uav swarms,” *Journal of Intelligent and Robotic Systems*, Vol. 61, No. 1-4, pp. 339–353, 2011.

[4.7] S. Johnson, I. Rae, B. Mutlu, L. Takayama, “Can you see me now?: how field of view affects collaboration in robotic telepresence,” *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 2397–2406, 2015.

[4.8] G. Doisy, A. Ronen, Y. Edan, “Comparison of three different techniques for camera and motion control of a teleoperated robot,” *Applied Ergonomics*, Vol. 58, pp. 527–534, 2017.

[4.9] J. Chen, E. Haas, and M. Barnes, "Human performance issues and user interface design for teleoperated robots," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 37, No. 6, pp. 1231–1245, 2007.

[4.10] D. Woods, "Visual momentum: a concept to improve the cognitive coupling of person and computer," *Man-Machine Studies*, Vol. 21, pp. 229–244, 1984.

[4.11] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," *Proceedings of the 19th International Conference on Neural Information Processing Systems*, pp. 545–552, 2007.

[4.12] S. Ito, Y. Sakano, K. Fujino, and H. Ando, "Remote controlled construction equipment by using high-resolution stereoscopic 3D images," *Japan Society of Civil Engineers*, Vol. 73, No. 1, pp. 15-24, 2017.

[4.13] M. Kamezaki, J. Yang, H. Iwata, and S. Sugano, "Visibility enhancement using autonomous multicamera controls with situational role assignment for teleoperated work machines," *Journal of Field Robotics*, Vol. 33, No. 6, pp. 802-824, 2016.

[4.14] I. Biederman, "On the semantics of a glance at a scene," *Perceptual Organization*, Chapter 8, 1981.

[4.15] J. Yang, M. Kamezaki, H. Iwata, and S. Sugano, "Analysis of effective environmental-camera images using virtual environment for advanced unmanned construction," *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pp. 664–669, 2014.

[4.16] B. Wandell, "Foundation of vision," *Sunderland: Sinauer*, 1995.

[4.17] M. Shimura, H. Suzuki, Y. Shimomura, and T. Katsuura, "Perception of visual motion stimulation in peripheral vision during eye fixation," *Japan Society of Physiological Anthropology*, Vol. 20, No. 2, pp. 95–102, 2015.

[4.18] M. Yamada, K. Fujino, H. Kajita, and T. Hashimoto, "Survey of the line of sight characteristics of construction machine operators to improve the efficiency of unmanned construction," *Proceedings of International Symposium on Automation and Robotics in Construction*, pp. 588–593, 2017.

[4.19] Y. Nitta, Evolution of unmanned construction system playing important role in the disaster recovery, *Civil Engineering*, Vol. 67, No. 4, pp. 16–23, 2012.

[4.20] M. Moteki, S. Yuta, and K. Fujino, "The proposal of the model task for efficiency evaluation of the construction work by remote control of a hydraulic excavator," *Journal of Japan Construction Machinery and Construction Association*, Vol. 66, No. 8, pp. 71–79, 2014.

[4.21] P. M. Groves, and R. F. Thompson, "Habituation: A dual-process theory," Psychological Review, Vol. 77, No. 5, p. 419–450, 1970.

[4.22] M. Kamezaki, H. Iwata, and S. Sugano, "Work state identification using primitive static states – Implementation to demolition work in double-front work machines," *Proceedings of International Symposium on Automation and Robotics in Construction*, pp. 278–287, 2009.

[4.23] S. Iwataki, H. Fujii, A. Moro, A. Yamashita, H. Asama, H. Yoshinada, "Visualization of the surrounding environment and operational part in a 3DCG model for the teleoperation of construction machines," in *Proceedings of IEEE/SICE International Symposium on System Integration*, pp. 81–87 2015.

[4.24] S. Kiribayashi, K. Yakushigawa, and K. Nagatani, "Design and development of tether-powered multirotor micro unmanned aerial vehicle system for remote-controlled construction machine," *Field and Service Robotics*, Vol. 5, pp. 637–648, 2017.

[4.25] Y. Ogino, T. Shibata, M. Tanaka and M. Okutomi, "Coaxial visible and FIR camera system with accurate geometric calibration," *Proceedings of SPIE 10214, Thermosense: Thermal Infrared Applications XXXIX*, paper no. 1021415, 2017.

Chapter 5

[5.1] M. Spranger, F. Heinke, S. Becker, and D. Labudde, "Towards drone-assisted large-scale disaster response and recovery," *Int. Conf. Advances in Computation, Communication and Services*, pp. 5–10, 2016.

[5.2] F. Nex, F. Remondino, "UAV for 3D mapping applications: a review," *Applied Geomatics*, vol. 6, no. 1, pp. 1–15, 2014.

[5.3] M. Kamezaki, J. Yang, H. Iwata, and S. Sugano, "A basic framework of virtual reality simulator for advancing disaster response work using tele-operated work machines," *Journal of Robotics and Mechatronics*, Vol. 26, No. 4, pp. 486-495, 2014.

[5.4] M. Kamezaki, H. Iwata, and S. Sugano, "Work state identification using primitive static states – Implementation to demolition work in double-front work machines," *Proceedings of International Symposium on Automation and Robotics in Construction*, pp. 278–287, 2009.

[5.5] M. Kamezaki, H. Iwata, S. Sugano, "Operator support system based on primitive static states in intelligent operated work machines," *Advanced Robotics*, Vol. 23, No. 10, pp. 1281-1297, 2009.

[5.6] M. Kamezaki, H. Iwata, S. Sugano, "A symbolic construction work flow based on state transition analysis using simplified primitive static states," *Journal of Robotics and Mechatronics*, Vol. 24, No. 6, pp. 939-948, 2012.

Research achievements

種 類 別 (By Type)	題名、　発表・発行掲載誌名、　発表・発行年月、　連名者（申請者含む） (theme, journal name, date & year of publication, name of authors inc. yourself)
論文	
○	無人化施工の掘削・リリース作業における側面カメラの最適および好適配置の実験的導出，日本機械学会論文集，2019.8，Vol. 85，No. 876，pp. 1–12，佐藤隆哉，亀﨑允啓，山田充，橋本毅，菅野重樹，岩田浩康
○	Development of a Cognitive Untunneling Multi-view System Based on Visual Momentum and Saliency for Teleoperators of Heavy Machines, Automation in Construction, 2019.12, Vol. 110, pp. 1–9, Ryuya Sato, Mitsuhiro Kamezaki, Satoshi Niuchi, Shigeki Sugano, and Hiroyasu Iwata
○	A Basic Framework of View Systems Allowing Teleoperators to Pre-Acquire Spatial Knowledge from Survey and Route Perspectives, Presence, Ryuya Sato, Mitsuhiro Kamezaki, Shigeki Sugano, and Hiroyasu Iwata (in press)
講演 (査読有)	
	A Low Cognitive Load and Reduced Motion Sickness Inducing Zoom Method Based on Typical Gaze Movement for Master-Slave Teleoperation Systems with HMD, 2020 IEEE/SICE International Symposium on System Integration (SII), pp. 28–33, 2020.1, Yuichi Mizukoshi, Ryuya Sato, Takahiro Eto, Mitsuhiro Kamezaki, Ayaka MATSUZAKA, Yang Liu, Akio Namiki, Asaki Imai, Takashi Matsuzawa, Kenji Hashimoto, Atsuo Takanishi, Hiroyasu Iwata
○	Experimental Investigation of Optimum and Allowable Range of Side Views for Teleoperated Digging and Release Works by Using Actual Construction Machinery, 2019 IEEE/SICE International Symposium on System Integrations (SII 2019), pp. 788–793, 2019.1, Ryuya Sato, Mitsuhiro Kamezaki, Mitsuru Yamada, Takeshi Hashimoto, Shigeki Sugano, and Hiroyasu Iwata
○	Derivation of an Optimum and Allowable Range of Pan and Tilt Angles in External Sideway Views for Grasping and Placing Tasks in Unmanned Construction Based on Human Object Recognition, 2019 IEEE/SICE International Symposium on System Integrations (SII 2019), pp. 788–793, 2019.1, Ryuya Sato, Mitsuhiro Kamezaki, Satoshi Niuchi, Shigeki Sugano, and Hiroyasu Iwata
○	A Pre-offering View System for Teleoperators of Heavy Machines to Acquire Cognitive Maps, The 15th IEEE International Symposium on Safety, Security and Rescue Robotics (SSRR2017), pp. 61–66, 2017.10, Ryuya Sato, Mitsuhiro Kamezaki, Shigeki Sugano, and Hiroyasu Iwata
	Investigation on Effect of Angular Deviation from Perpendicular Line to Lateral View Camera on Work Efficiency for Teleoperated Construction Machines, The 34th International Symposium on Automation and Robotics in Construction (ISARC2017), paper no. P06, 2017.6, Satoshi Niuchi, Mitsuhiro Kamezaki, Ryuya Sato, Shigeki Sugano, and Hiroyasu Iwata
	Pre-offering Work-Site Views to Acquire Spatial Knowledge from Survey and Route Perspective for Advanced Teleoperation of Construction Machines, The 34th International Symposium on Automation and Robotics in Construction (ISARC2017), paper no. P05, 2017.6, Ryuya Sato, Mitsuhiro Kamezaki, Satoshi Niuchi, Shigeki Sugano, and Hiroyasu Iwata

種 類 別 By Type	題名、　発表・発行掲載誌名、　発表・発行年月、　連名者（申請者含む） (theme, journal name, date & year of publication, name of authors inc. yourself)
講演 (査読有) ○	Gaze Pattern Analysis in Multi-Display Systems for Teleoperated Disaster Response Robots, A Satellite Conference of 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC JUNIOR 2016), paper no. 2017, 2016.10, Ryuya Sato, Mitsuhiro Kamezaki, Shigeki Sugano, and Hiroyasu Iwata Visual Attention to Appropriate Monitors and Parts Using Augmented Reality for Decreasing Cognitive Load in Unmanned Construction, 6th International Conference on Advanced Mechatronics 2015 (ICAM2015), page 45 (1A1-23), 2015.12, Ryuya Sato, Mitsuhiro Kamezaki, Junjie Yang, Shigeki Sugano, and Hiroyasu Iwata Performance Analysis in Advanced Tele-operation System Based on Introduction of Danger-Avoidance View, 8th International Conference on Intelligent Robotics and Applications (ICIRA2015), pp. 401-412, 2015.8, Junjie Yang, Mitsuhiro Kamezaki, Ryuya Sato, Hiroyasu Iwata, and Shigeki Sugano Inducement of Visual Attention Using Augmented Reality for Multi-Display Systems in Advanced Tele-operation, 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2015), pp. 5364-5369, 2015.9, Junjie Yang, Mitsuhiro Kamezaki, Ryuya Sato, Hiroyasu Iwata, and Shigeki Sugano 無人化施工における奥行感の把握が必要な手先作業時における外部カメラの最適・好適配置の実験的導出，第 19 回建設ロボットシンポジウム論文集(SCR2019），paper no. O3-1，2019.10，佐藤隆哉，亀崎允啓，山田充，橋本毅，菅野重樹，岩田浩康 重機の遠隔操作性向上のためのマルチカメラ最適配置に関する研究－ 第二報　掘削・配置作業におけるパン・チルト角が及ぼす作業効率への影響の実験的検証 －，第 24 回ロボティクスシンポジア講演論文集，pp. 14–17，2019.3，佐藤隆哉，亀崎允啓，山田充，橋本毅，菅野重樹，岩田浩康 ダンプトラックの燃費向上・長寿命化に向けた状態識別に関する基礎検討，第 18 回建設ロボットシンポジウム論文集(SCR2018)，paper no. O1-5，2018.9，山村真司，亀崎允啓，佐藤隆哉，菅野重樹，岩田浩康 重機の遠隔操作性向上のためのマルチカメラ最適配置に関する研究－第一報　パン・チルト角による作業性への影響の検証－，第 23 回ロボティクスシンポジア講演論文集，pp. 267–272，2018.3，仁内智志，亀崎允啓，佐藤隆哉，菅野重樹，岩田浩康 ヒトの認知構造に基づいた事前の環境把握映像提示による災害対応遠隔重機作業の効率化に関する研究，第 17 回建設ロボットシンポジウム論文集(SCR2017)，paper no. O3-2，2017.8，佐藤隆哉，亀崎允啓，仁内智志，菅野重樹，岩田浩康

種 類 別 By Type	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む） (theme, journal name, date & year of publication, name of authors inc. yourself)
講演 (査読有)	無人化施工における視覚情報の強化に関する研究－作業状況に応じた環境カメラの自動制御と拡張現実技術を用いた注視支援－，第 15 回建設ロボットシンポジウム論文集(SCR2015)，paper no. O-24，2015.9，亀﨑允啓，佐藤隆哉，楊俊傑，岩田浩康，菅野重樹
講演 (査読無)	重機の遠隔操作性向上のためのマルチカメラ最適配置に関する研究－ 第三報 搭乗操作熟練者における掘削・配置作業でのパン・チルト角が及ぼす作業効率への影響の実験的検証ー，第 37 回日本ロボット学会学術講演会，paper no. 2A1-05，2019.9，佐藤隆哉，亀﨑允啓，山田充，橋本毅，菅野重樹，岩田浩康 災害対応人型ロボット遠隔マニピュレーション作業における操作者の疲労軽減および精度向上可能なスケール・ゲイン調整手法の実機適応性検証，第 37 回日本ロボット学会学術講演会，paper no. 2A1-02，2019.9，佐藤隆哉，亀﨑允啓，江藤孝紘，水越勇一，劉楊，並木明夫，今井朝輝，松澤貴司，孫瀟，橋本健二，高西淳夫，岩田浩康 低疲労空間内で人型ロボットの精密遠隔作業を可能とするマスタスレーブ相対手先位置調整手法の提案，第 26 回バイオメカニズム・シンポジウム，pp. 72–79，2019.7，佐藤隆哉，江藤孝紘，亀﨑允啓，並木明夫，岩田浩康 HMD を用いた脚ロボット遠隔操作時における酔いの低減が可能な注視の典型動作を用いた低認知負荷ズーム手法の効果検証，日本機械学会ロボティクス･メカトロニクス講演会 2019 論文集(Robomec’19) ，paper no. 2P1-D05，2019.6，水越勇一，佐藤隆哉，江藤孝紘，亀﨑允啓，松坂彩香，並木明夫，今井朝輝，松澤貴司，孫瀟，橋本健二，高西淳夫，岩田浩康 重機の遠隔操作における作業状態の変化に適応可能な拡張現実技術を用いた低認知負荷リマインド手法の構築，日本機械学会ロボティクス･メカトロニクス講演会 2019 論文集(Robomec’19) ，paper no. 1P2-D07，2019.6，佐藤隆哉，亀﨑允啓，山下雄輝，菅野重樹，岩田浩康 重機遠隔操作者の視線を作業状態に応じた映像に誘導可能な映像提示手法の構築，第 19 回計測自動制御学会システムインテグレーション部門講演会論文集(SI2018)，paper no. 2C1-03，佐藤隆哉，亀﨑允啓，仁内智志，菅野重樹，岩田浩康 マスタ・スレーブシステムにおける提示映像スケーリングの有効性検証，日本機械学会ロボティクス･メカトロニクス講演会 2018 論文集(Robomec’18) ，paper no.2A2-M04，2018.6，江藤孝紘，亀﨑允啓，佐藤隆哉，岩田浩康

種 類 別 By Type	題名、　　発表・発行掲載誌名、　　発表・発行年月、　　連名者（申請者含む） (theme, journal name, date & year of publication, name of authors inc. yourself)
講演 (査読無)	災害対応作業の複雑・連続・時限性を考慮したマスタ・スレーブシステムのスケール・ゲイン調整手法の開発，日本機械学会ロボティクス･メカトロニクス講演会 2018 論文集 (Robomec'18) ，paper no.2A2-M02，2018.6，亀﨑允啓，江藤孝紘，佐藤隆哉，岩田浩康 遠隔操作導入前の映像提示システムによる直接描画法を用いた操作者の認知地図への影響分析，第 18 回計測自動制御学会システムインテグレーション部門講演会論文集 (SI2017)，paper no. 3C4-03，2017.12，佐藤隆哉，亀﨑允啓，仁内智志，菅野重樹，岩田浩康 側面カメラ映像における撮影対象との垂直度が遠隔操作者の奥行き感把握に与える影響の調査，第 17 回建設ロボットシンポジウム論文集(SCR2017)，paper no. P2-01，2017.8，仁内智志，亀﨑允啓，佐藤隆哉，菅野重樹，岩田浩康 脚ロボットにおける連続的な移動・手先動作遷移に対応可能な遠隔操作インタフェースの開発，日本機械学会ロボティクス･メカトロニクス講演会 2017 論文集(Robomec'17)，paper no. 1P1-Q04，2017.5，江藤孝紘，佐藤隆哉，仁内智志，中村早紀，亀﨑允啓，岩田浩康 災害対応重機の遠隔操作における操作者視点映像の事前提供による環境把握性効果の検証，日本機械学会ロボティクス･メカトロニクス講演会 2017 論文集(Robomec'17)，paper no. 1P2-A02，2017.5，佐藤隆哉，亀﨑允啓，仁内智志，菅野重樹，岩田浩康 ウォード法を用いた複数映像提供型遠隔操作における視線パターンのクラスタリング，第 17 回計測自動制御学会システムインテグレーション部門講演会論文集(SI2016)，pp. 2592-2595（3H1-2），2016.12，佐藤隆哉，亀﨑允啓，菅野重樹，岩田浩康 複数画面を使った遠隔操作における注視映像と作業パフォーマンスの関連性分析，第 16 回計測自動制御学会システムインテグレーション部門講演会論文集(SI2015)，pp. 2354-2355 (3G2-6)，2015.12，亀﨑允啓，佐藤隆哉，楊俊傑，岩田浩康，菅野重樹 拡張現実感を利用した複数の視線誘導手法による遠隔操作者の認知負荷軽減に関する研究，第 16 回計測自動制御学会システムインテグレーション部門講演会論文集(SI2015)，pp. 2320-2323 (3G1-2)，2015.12，佐藤隆哉，亀﨑允啓，楊俊傑，菅野重樹，岩田浩康 遠隔重機作業の高度化に関する研究―拡張現実感を利用した映像の注視・解釈支援手法の開発―，日本機械学会ロボティクス･メカトロニクス講演会 2015 論文集(Robomec'15)，paper no. 2P1-P04，2015.5，亀﨑允啓，佐藤隆哉，楊俊傑，岩田浩康，菅野重樹

種 類 別 By Type	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む） (theme, journal name, date & year of publication, name of authors inc. yourself)
その他 （講演） （査読有）	Investigating the Relationship between Connection, Agency and Autonomy for Controlling a Robot Arm for Remote Social Physical Interaction, 21st International Conference on Human-Computer Interaction (HCI International 2019), pp. 277–290, 2019. 7, Ryuya Sato, Don Kimber, Yanxia Zhang Forest Roads Design and 3D Visualization for Steep Mountains, The IUFRO 125th Anniversary Congress, paper no. 1155, 2017.9, Naoki Nozawa, Shoma Fuji, Ryuya Sato, Ryo Suzuki, and Yuko Shirai Inverse Innovation; Ripple Railway Model to Acquire Local Industries Based on User's Viewpoint in Thailand, The 13th International Symposium on Autonomous Decentralized Systems (ISADS 2017), pp. 281-286, 2017.3, Satoshi Funabashi, Ryuya Sato, Tamon Miyake, Ryosuke Tsumura, and Kinji Mori
その他 （講演） （査読無）	Proposition on Thai Railway System Based on Field Work, IEEE International Symposium on Railway System (ISRS2016), 2016.6, Ryuya Sato, Tamon Miyake, Satoshi Funabashi, and Kinji Mori Innovation of Railway System, IEEE International Symposium on Railway System (ISRS2015), 2015.11, Ryuya Sato, Tamon Miyake, Satoshi Funabashi, and Kinji Mori
その他 （特許）	路網ルート設計装置及びそのプログラム、並びに路網ルート生成表示システム，特開2017-201936，2017.11，白井裕子，野澤直樹，藤井祥万，佐藤隆哉，加藤卓哉
その他 （研究費）	立石科学技術振興財団　研究助成（C）採択　2018.4 – 現在 日本学術振興会特別研究員 DC2 研究奨励費 2019.04 - 現在 博士課程教育リーディングプログラム 「実体情報学博士プログラム」奨励金 2015.04 - 2019.03
その他 （受賞）	日本機械学会三浦賞，2017.3 Best Fitness Hack, 1st FXPAL International Hack Day Expo Extravaganza, 2018.7

種類別 By Type	題名、　発表・発行掲載誌名、　発表・発行年月、　連名者（申請者含む） (theme, journal name, date & year of publication, name of authors inc. yourself)
その他 （招待講演, シンポジウム等での発表, 展示等）	BYD にて招待講演，2019.6 Japan Robot Week 2018 にて機器展示，2018.10 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2018), Fr-WS3: ImPACT Tough Robotics Challenge: A National Project of Disaster Robotics Aiming at Social Innovation in Safety and Security, 2018.10 ImPACT　タフ・ロボティクス・チャレンジシンポジウムにてポスター発表，2018.7 ImPACT　タフ・ロボティクス・チャレンジ第 5 回フィールド評価会にて非公開デモおよびポスター発表，2017.11 リーディングフォーラム（Program for Leading Graduate Schools Forum）にてポスター発表，2017.10 ImPACT　タフ・ロボティクス・チャレンジ第 4 回フィールド評価会にて非公開デモおよびポスター発表，2017.6 ImPACT　タフ・ロボティクス・チャレンジ第 3 回フィールド評価会にてポスター発表，2016.11 ImPACT　タフ・ロボティクス・チャレンジ第 2 回フィールド評価会にてポスター発表，2016.6 早稲田大学博士課程教育リーディングプログラム　リーディング理工学博士プログラム×実体情報学プログラム　1st 合同シンポジウムにて登壇発表，2016.3 Waseda-Tsinghua workshop にて口頭発表，2016.1

Appendix：私の考える実体情報学

リーディングプログラムで経験したことに基づく筆者の実体情報学の定義は，“creating values for users by integrating knowledge”である. 前半の“creating values for users”はタイにおける鉄道プロジェクト，後半の“by integrating knowledge”は山プロジェクト，における経験から学んだことである．また，筆者はこの実体情報学で学んだことを半年間のインターンで実践し，実体情報学を筆者なりに定義した．そこで，各プロジェクトにおける経験を述べる．

（1）タイにおける鉄道プロジェクト（L1~L3，先見力）

このプロジェクトでは, 現地調査に基づく課題設定の重要性を学んだ. この授業では,”タイ”と”鉄道”という 2 つのお題のみ与えられ，そこからタイの人々のためになる鉄道プランを提案するという内容であった.

現地調査の前は，日本において文献調査を行い，その結果より東北部の人々の収入は首都の半分程度であるため，東北部を対象として提案を行った．具体的には，東北部の土地の 70%は米用の農地であり，また，タイの人口の 50%は東北部に住んでおり，2015 年においてタイの米の輸出量は世界一であった．しかし，収入が首都のバンコクの半分程度であり, この主要な原因は米という付加価値の低いものをそのまま輸出しているためだと考え，1 次産業（生産）から 3 次産業（サービス）までを行う 6 次産業（1 次×2 次×3 次＝6 次産業）を導入すれば，東北部の人々の収入が増えると仮説を立てた．そこで，生産したコメ（1 次産業）を東北部の人たちで酒やせんべいのようなものに加工し（2 次産業），その後鉄道を用いて加工された大量の食品を運送する（3 次産業）という提案を行った.

しかし実際に現地に行き，チュラロンコン大学やタイの運輸省，バンコク・メトロの人々など計 17 の機関との議論や現地の鉄道工場見学から，本質的な課題は中進国から先進国への成長が困難となる中進国の罠に陥りつつあるということを抽出した. その原因として，30%以上の企業が 50%以上外資企業に依存していることがあげられる．鉄道の場合では，ヨーロッパの会社に全てのシステムを外注しているため，メンテナンスの知識等がタイに根付かず，また，メンテナンスのための費用を払うことができずに，閉鎖する路線もある状況であった．そこで，中進国の罠から抜け出すには自国産業が必要となることから, 従来人と物の輸送としてのみ使用されていた鉄道を技術の集合体ととらえ，鉄道を自国産業化し，中進国の罠を抜けるという提案を国際会議で行った．具体的には，日本が 100 年後には人口が半分になるため，日本の鉄道会社は海外進出を行う必要がある．しかし，提案時点ではホーチミンにしか進出できていないため，日本企業

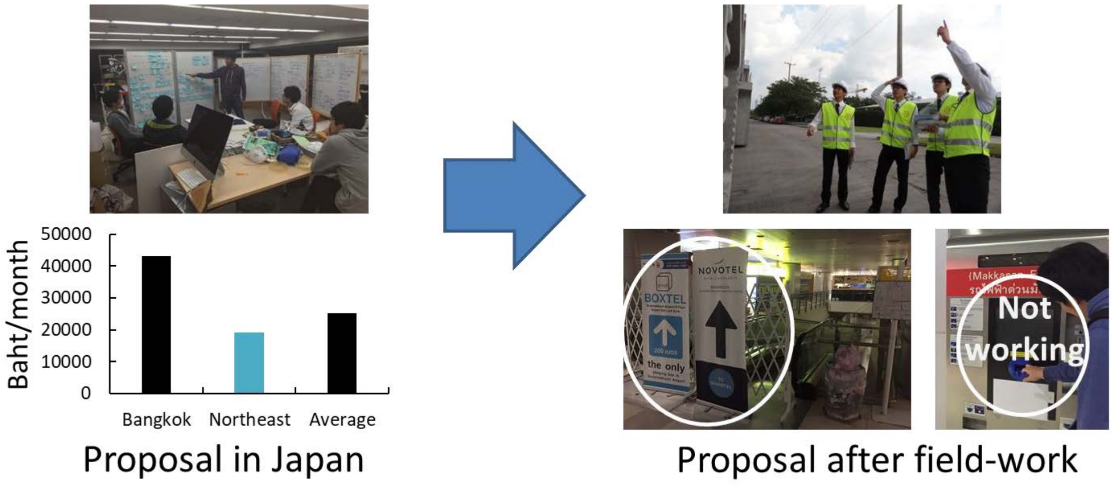


Fig. 1 Thai railway project

が技術提供を行うことで，日本企業は海外進出のメリットが，タイには鉄道の技術を学ぶことができるというメリットがある，という提案を行った．この成果を IEEE の国際会議において 3 件報告した．

（2）山プロジェクト（L1~L3，構想力）

このプロジェクトでは，林業における建設機械の通り道である作業道の自動生成手法の構築を自分の専門ではないコンピュータグラフィックスの技術を用いて行った．ここでは，自分の専門外の分野の知識を統合し，システム構築を行う重要性を学んだ．

まず，タイにおける鉄道プランの提案で学んだ現地調査を行い，建設機械が通る道である作業道に問題があることを発見した．林業には，建設機械を使用して木を伐採し，運搬する作業がある．その際，作業道を作設する必要があるが，職人の経験と勘のみで作業道を計画することから，効率が低下することや依頼主がどのような作業道が作設されるかを想像できずに作業道の作設自体を拒否することがある．そこで，数値的根拠に基づく作業道の生成を行い，その作業道をコンピュータグラフィックスの技術を用いて 3D で表示するシステムを異分野の学生と議論しながら構築した．その成果で特許を取得し，また，林業分野におけるトップカンファレンスである IUFRO にて発表を行い，高評価を受けた．

（3）インターン（L4，突破力）

このプロジェクトでは，自分が学んできた現地調査に基づく課題設定の重要性と異分野の知識を統合してシステムを構築する重要性を応用した．

まず，タイにおける鉄道プランの提案で学んだ現地調査を行った．インターンにおけるテーマは，Connecting space であったため，ディズニーランドを繋げる場所として選定し，ディズニーランドにおいて現地調査を行った．そこで，トイ・ストーリー・マニ

(a) Fieldwork in Tenryu

(a) Developed road design system

Fig. 2 Mountain project

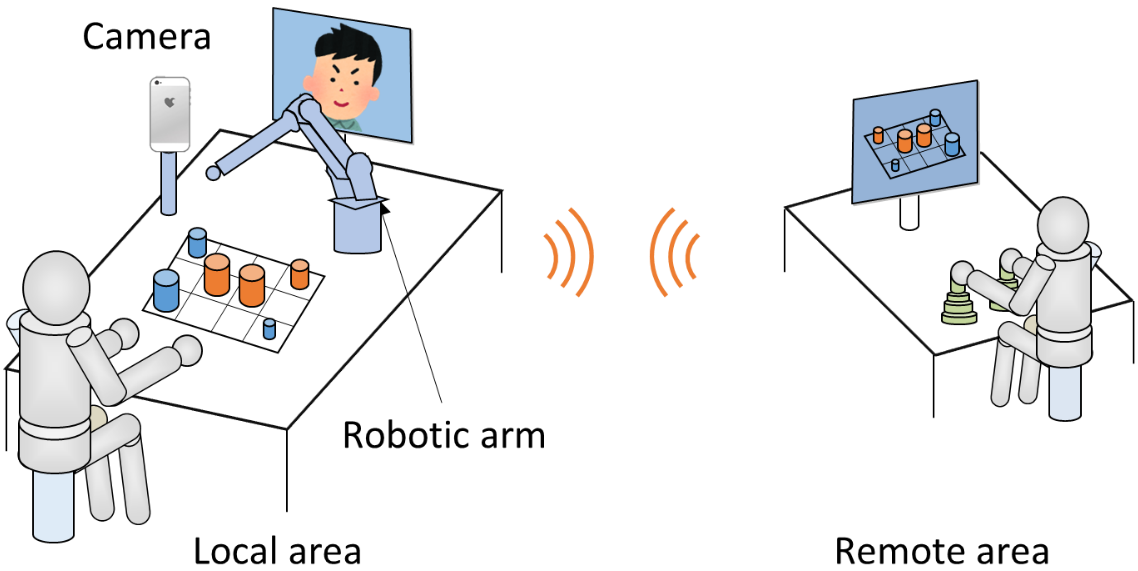


Fig. 3 Intern

アやバズ・ライトイヤー・アストロブラスター等のアトラクションでは，物理的なインタラクションが必要であることを発見した．例えば，バズ・ライトイヤー・アストロブラスターでは，光線銃のトリガーを引かなければアトラクションを楽しむことができない．しかし，従来の LINE や Skype 等の遠隔コミュニケーションツールではこういった物理的なインタラクションを実現することができない．そこで，ロボットアームがあり，遠隔操作を行えば物理的なインタラクションが可能になるという提案を行った．

次に，人間工学という異分野の知識を用いて研究を行った．ロボットアームの操作インタフェースが適切でなければ，遠隔操作者が行いたい事を行えなくなる可能性があるため，疎外感や孤独感を覚える可能性がある．したがって，遠隔操作インタフェースに関する研究を行った．具体的には，ヒトには Sense of Agency（主体感）というものがあり，これは今

回のケースでは自分が遠隔地のロボットアームを動かしているという感覚となる．この主体感が低下すると遠隔地とのつながりが希薄になる．そこで，Sense of Agency は人が意識していない Feeling of Agency と人が意識している Judgement of Agency に分けられることから，Feeling of Agency による影響を調査した．

以上が筆者が実体情報学で学んだことであり，これらの経験から，実体情報学を“creating values for users by integrating knowledge”と定義した．自身の研究においても，無人化施工が導入されている雲仙普賢岳に見学に行くことや, 熊本地震の現場に足を運ぶことをし，現地調査より課題を設定した．また，博士論文の研究対象は遠隔操作者であることから，人間工学の知識を組み込むことにより，ヒトの認知特性に基づいた映像提示手法を構築した．