Ryuya Sato*
Department of Modern Mechanical Engineering
Graduate School of Creative Science and Engineering
Waseda University, Tokyo, Japan

Mitsuhiro Kamezaki
Research Institute for Science and Engineering
Waseda University, Tokyo, Japan

Shigeki Sugano
Department of Modern Mechanical Engineering
School of Creative Science and Engineering
Waseda University, Tokyo, Japan

Hiroyasu Iwata
Department of Modern Mechanical Engineering
School of Creative Science and Engineering
Waseda University, Tokyo, Japan

Presence, *Vol. 27, No. 3, Summer 2018*, 309–332
https://doi.org/10.1162/PRES_a_00333

A Basic Framework of View Systems Allowing Teleoperators to Pre-Acquire Spatial Knowledge from Survey and Route Perspectives

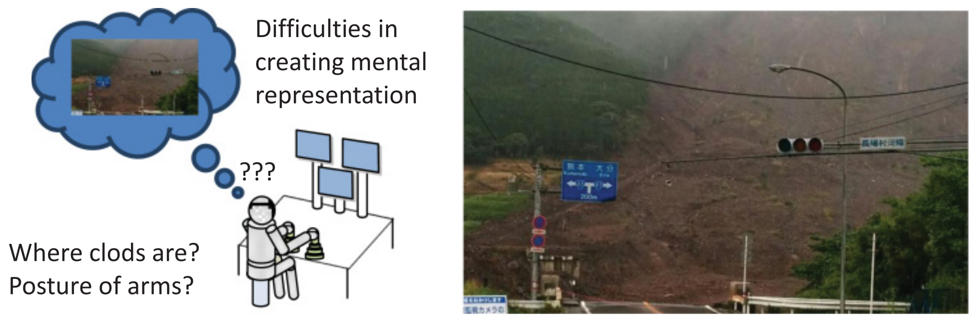
Abstract

One of the most important problems in teleoperation of heavy machinery is that the work efficiency of teleoperation is lower than half that of a typical boarding operation. This difference is primarily caused by operators' difficulty in creating mental representations (i.e., cognitive maps) of work sites. Operators have two opportunities to acquire information, namely before work and during work, because the introduction of teleoperation requires about one week. Therefore, we have developed a view system to be used before work to provide environmental information concerning work sites on the basis of human spatial cognition. Cognitive maps can be built by acquiring knowledge from two perspectives—the survey perspective and the route perspective. We display an external view from any viewpoint to acquire knowledge from a survey perspective and a view from an operator's viewpoint, which can be modified by the operator's intention to acquire knowledge from the route perspective. Experimental results using a simulator suggested that a proposed view system could help operators acquire cognitive maps, which may lead to a decrease in task time, the number of stops, and the moving distance and an increase in speed during grasping.

1 Introduction

Japan has experienced many natural disasters, such as the eruption of Unzen-Fugendake in 1991 (Nakada & Fujii, 1993) and the Great East-Japan Earthquake in 2011 (Lay & Kanamori, 2011). After such disasters, teleoperation of heavy machinery such as construction equipment has been introduced since secondary disasters such as landslides may occur (Hiramatsu, Aono, & Nishino, 2002; Kawatsuma, Fukushima, & Okada, 2012; Chayama et al., 2014). Operators can maneuver heavy machinery in a safe and distant area by watching several views from cameras in disaster sites. One of the most crucial problems affecting teleoperation is degradation of work performance. The work efficiency of teleoperation is lower than half that of a typical boarding operation (Moteki, Akihiko, Yuta, Mishima, & Fujino, 2016). This degradation primarily arises because operators have difficulty creating mental representations of work sites in their mind because of the current view system, as shown in Figure 1 (Asama & Ueki, 2013; Fong, Thorpe, & Baur, 2003). Operators

*Correspondence to ryuya-sato@iwata.mech.waseda.ac.jp.



(a) Teleoperation room (b) Example of an actual work site

Figure 1. *Difficulties in creating mental representations of work sites.*

need two types of information to create mental representations of work sites (Asama & Ueki, 2013). The first one is environmental, for example, the locations of things such as debris or clods (see Figure 1(b), a picture of a disaster site in the 2016 Kumamoto earthquake). The second one is the heavy machinery itself, for example, the posture of the arms and the orientation of the equipment during work. In this article, we focus on mental representations of environment because environmental information is essential for teleoperation.

Operators can hardly plan paths and working strategies without mental representations of work sites because humans judge based on what they recognize (Endsley, 1995). For example, in the work sites with seven objects to transport and four obstacles, as illustrated in Figure 2(a), operators can choose an oblique path (black dotted line) without the information concerning the four diaphanous objects on the right of the figure. Moreover, operators may stop to search for objects without their positional information. Furthermore, in the work site as illustrated Figure 2(b), operators can choose a roundabout path (black dotted line) without the information concerning the distance between the obstacles (black two-way arrow).

Operators have two opportunities to obtain such information, namely before work commences (since the introduction of heavy-machinery teleoperation requires about a week; Nitta, 2012) and during work. Several researchers have developed systems to provide such information during work periods. These include, for example, adding displays to increase fields of view (Moteki, Akihiko, Yuta, Mishima, & Fujino, 2016), providing an arbitrary view using a drone (Kiribayashi, Yakushigawa, & Nagatani, 2018), and providing views from various perspectives by automatically changing the rotation and view angle of the cameras (Kamezaki, Yang, Iwata, & Sugano, 2016). Furthermore, moving map displays designed to give a sense of robot orientation (Casner, 2005), gravity-referenced view displays to show robot attitude (Wang, Lewis, & Hughes, 2004), and stereoscopic displays to offer depth perception (Draper, Handel, & Hood, 1991) have been developed outside of the construction-research field.

However, no studies have focused on providing information before work commences. Lack of environmental information before work causes difficulties in planning paths and working strategies because operators tend to do so before work. People can plan only with information acquired during work (Passini, 1984). Thus, many navigation systems for displaying which way to go at that moment, including car navigations, head-up displays (Burnett, 2003), and augmented reality navigation systems (Narzt et al., 2005), have been developed. However, these systems require the answer paths. Various information including a 3D map at disaster sites, the hardness of the ground, and water content of the ground is necessary for finding the answer paths at the disaster sites. However, current systems can hardly acquire all the information, even though some information such as a 3D map at disaster sites can be

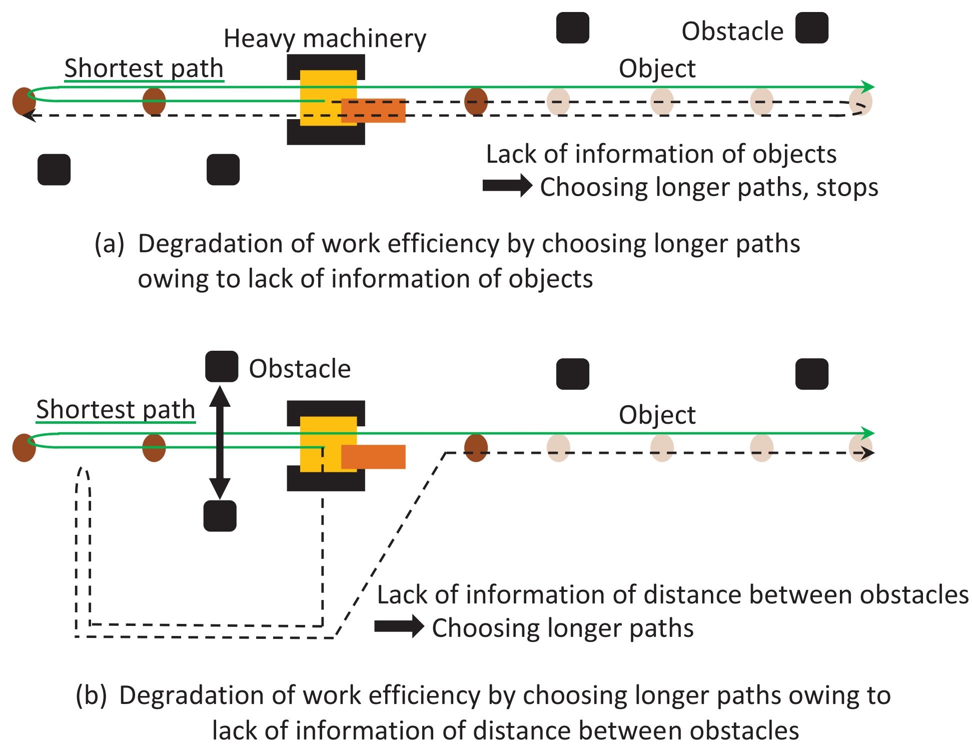


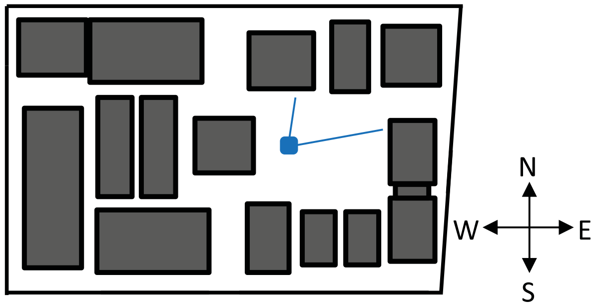
Figure 2. *Degradation of work efficiency owing to lack of environmental information.*

acquired (Nex & Remondino, 2014). In summary, previous research including navigation systems can barely adapt for teleoperation of heavy machinery owing to difficulties to acquire necessary information for planning answer paths. Moreover, planning only with information acquired during work forces them to plan and work simultaneously. Humans acquire 70% of their information from sight (Heilig & Futuro, 1992), and visual information is the most important among the five senses for judging and planning (Crundall & Underwood, 1998). Therefore, in this study, we develop a view system to be used in advance of work so as to provide environmental information; we also analyze the work-performance and operator-mental-representation effects of this system. The contribution of the study is to help operators plan before work by providing prior environmental information, though previous studies have focused on providing information during work. This can increase work efficiency because operators can focus on operation with enough prior environmental information.

In Section 2, we first introduce human spatial cognition, and then develop a view system based on it and hypothesize the effects of the proposed system. In Section 3, experiments to evaluate the proposed view system are conducted, and the results are analyzed in terms of work efficiency and operator-mental-representation. In Section 4, we conclude the study and explain future work.

2 A View System for Providing Environmental Information Prior to Work

In this section, we develop a view system to provide environmental information to teleoperators before work on the basis of human spatial cognition. First, we explain human spatial cognition according to psychological knowledge. Next, we assume the effects of acquiring environmental information for teleoperation. Finally, we develop a view system based on human spatial cognition and presume the effects of this view system.



(a) Survey knowledge
(map of the campus of Waseda University)

(b) Route knowledge
(captured from the blue point in (a))

Figure 3. *Comparison of two perspectives of a cognitive map.*

2.1 Human Spatial Cognition

Humans store mental representations known as cognitive maps (Edward, 1948). These maps can be built by acquiring knowledge from survey and route perspectives (Golledge & Stimson, 1997). Humans acquire survey knowledge from external viewpoints, whereas route knowledge can be acquired from internal or personal viewpoints. Thus, survey knowledge is usually described in absolute coordinates such as east and west, whereas route knowledge is described in terms of relative coordinates such as front and back. Moreover, maps are usually used to express survey knowledge, whereas pictures taken from a human viewpoint are usually used to express route knowledge. For example, the maps illustrated in Figure 3(a) are used for the survey knowledge and described in terms such as south and north, whereas the pictures shown in Figure 3(b) come from the route knowledge and are described in terms of right and left (Figures 3(a) and (b) represent the campus of Waseda University).

2.2 Assumption of Effects of Acquiring Cognitive Maps

We assume the effects of acquisition of cognitive maps upon teleoperation work performance. Operators can plan general paths with knowledge from the survey perspective and work strategies in a particular area with knowledge from the route perspective.

2.2.1 Survey Knowledge. We hypothesize the effects of acquiring knowledge from the survey perspective in a work site as illustrated in Figure 2(a). Operators can recognize where objects and obstacles are from external viewpoints if they acquire knowledge from a survey perspective. Thus, operators can plan and choose the shortest path such as the green line in Figure 2(a) and thus avoid stopping to search.

2.2.2 Route Knowledge. We assume the effects of acquiring knowledge from route perspective in a work site as illustrated in Figure 2(b). Operators recognize the working environment, including distance between obstacles, as illustrated in Figure 2(b) inside the cabins during a usual boarding operation. Therefore, operators can recognize the distance between obstacles because route perspective is close to the views operators watch during a usual boarding operation. Thus, operators can choose the shortest path, such as the green line in Figure 2(b). Moreover, we hypothesize the effects in a work site, as illustrated in Figure 4. Operators must simultaneously work and plan how to grasp the right and left fallen trees illustrated in Figure 4 without prior route knowledge. Therefore, the speed of grasping may be degraded or operators may need to stop to plan how to grasp the fallen trees. However, operators can finish planning prior to the work if they acquire route knowledge before work because they recognize environments from viewpoints where operators usually are.

Figure 4. *Hypothesized effects of acquiring cognitive maps.*

Thus, operators can maintain a high speed during grasping and keep working without stopping.

2.3 Development of a View System for Acquiring Cognitive Maps and Hypothesis on Effects of a View System

Acquiring cognitive maps can enhance work performance as described in Subsection 2.2. Therefore, we develop a view system before work to enable operators to acquire cognitive maps, as illustrated in Figure 5. Moreover, we hypothesize the effects of this proposed view system.

2.3.1 Survey Knowledge. This knowledge can be acquired from external viewpoints. Different viewpoints that occur before and during work may cause difficulties in recognizing environments, since mental rotation is required to match different viewpoints (Levin, Jankovic, & Palij, 1982). Furthermore, suitable viewpoints differ between operators (Sato, Kamezaki, Sugano, & Iwata, 2016) and tasks. Thus, we display a zoomable view from an external viewpoint that can be modified by operators as shown in Figure 5(a). We display this view in 2D because teleoperators maneuver by watching 2D views.

This survey-knowledge view can let operators plan general paths, which allows them to keep working without stops as explained in Subsection 2.2.1. Operators can remember more objects more accurately from external viewpoints, but remembering all positions of objects in disaster sites is almost impossible because humans have a limited memory (Miller, 1956). Thus, operators may prefer to remember important objects, such as the debris that needs to be transported and the area to which the debris should be released, to follow planned paths.

2.3.2 Route Knowledge. This knowledge can be acquired from personal viewpoints. Operators can acquire it more effectively from active movements which include modifying views according to their intention, rather than watching predetermined views (Cadwallader, 1975). Thus, we display a view from an operator's perspective that can be modified according to the operators' intention as shown in Figure 5(b). Cognitive distance, which is a distance humans recognize, works more effectively in decision making than actual distance (Garling & Golledge, 1993). Cognitive distance is formed by various things including human experience (Ankomah, Crompton, & Baker, 1996). Therefore, we display a view so that operators can modify this view at the same speed as heavy machinery moves.

This route-knowledge view can let operators plan how to work by means of grasping objects. This is because these views are similar to those that operators watch when they operate heavy machinery and cab views during teleoperation. Moreover, operators prefer watching cab views from cameras in heavy equipment, rather than views from cameras around the work site (Moteki, Fujino, Ohtsuki, & Hashimoto, 2011). This planning can help operators keep working without stops, speed of the attachment can be faster during grasping, and moving distance can be decreased as explained in Subsection 2.2.2.

3 Experiment

We conducted experiments, including grasping objects and transporting them using a simulator (Kamezaki, Yang, Iwata, & Sugano, 2014), to evaluate the effects of the proposed view system. The ethics committee for human research at Waseda University approved the procedures of the experiments in the study.

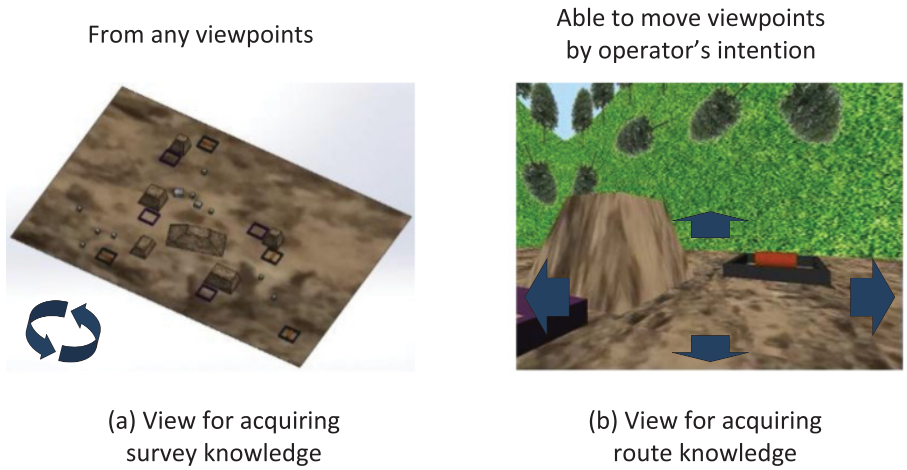


Figure 5. *A view system to acquire cognitive maps before work commences.*

3.1 Method

3.1.1 Participants. Sixteen novice operators with no experience in operating heavy machinery were involved as participants because there are only about 20 skilled operators in Japan (Nitta, 2012). However, they acquired the skills required to teleoperate heavy machinery in a simulator through training before the start of the experiments. The participants were all male students, aged 22 to 25, majoring in mechanical engineering at Waseda University. The number of participants was determined by using G∗Power (Faul, Erdfelder, Lang, & Buchner, 2007; Faul, Erdfelder, Buchner, & Lang, 2009). We conducted the pre-experiments, which were the same procedure as the experiments in this study, with eight participants who tried the tasks three times. Four parameters, including the effect size d, α error prob, power (1-β error prob), and allocation ration N2/N1, are required to calculate the required sample size for t-test (statistical test; means: difference between two independent means (two groups), type of power analysis; a priori: compute required sample size). Mean and SD values are required to calculate the effect size d. The results of mean work time and SD obtained by the pre-experiments with 24 samples were 282.0 ± 33.2 for participants watching the proposed prior view systems and 331.9 ± 65.7 for those watching the conventional view systems. Thus, the effect size d was 0.96. Moreover, we determined α error prob as 0.05, power as 0.8 (Cohen, 1992), and allocation ration N2/N1 as 1 because of the same sample size. The required total sample size was 38 based on the calculation. Thus, we set the number of sample size as 48, which means 16 participants with three trials, so that it should be over the required sample size.

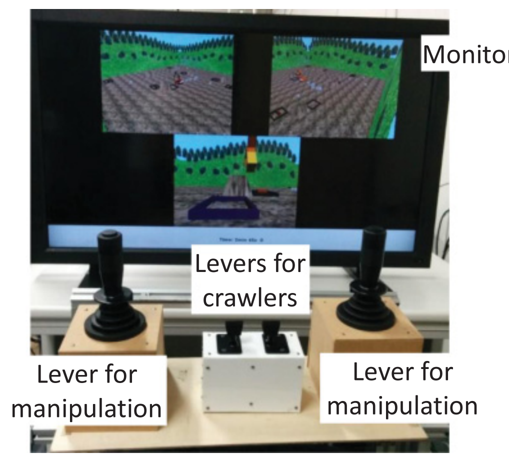
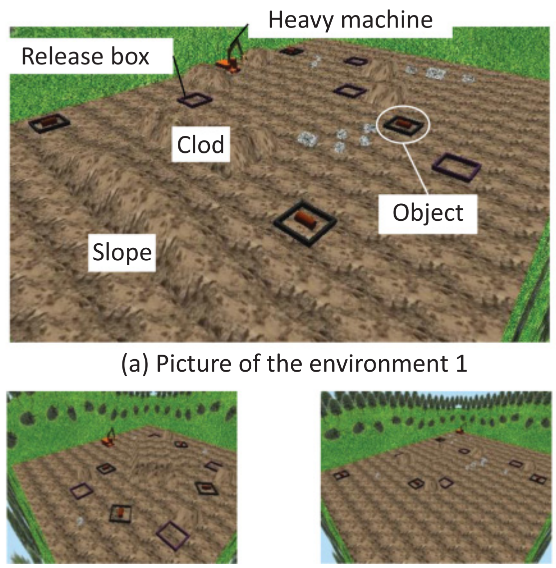


Figure 6. *Interface of the developed simulator.*

3.1.2 Material. Figure 6 shows the interface to control the heavy machine in the simulator (Kamezaki, Yang, Iwata, & Sugano, 2014). This simulator



(b) Picture of the environment 2 (c) Picture of the environment 3

Figure 7. *Three experimental environments.*

calculates physics by ODE and renders by OpenGL. Details of the simulator are described in a previous paper (Kamezaki, Yang, Iwata, & Sugano, 2014). We used a desktop computer (processor: Interl®Core^{TM}i7-4770 CPU @3.40-GHz, memory: 8.00-GB, OS: Windows 10) for the simulator. We used two joysticks for manipulation (type: S90JBM-YO-21R2G, made by Sakae), and two levers for crawlers (type: 30JLK-X1-10R1H, made by Sakae). Participants controlled heavy machinery by watching a 42-inch monitor. The experimental tasks included grasping four cylindrical objects and transporting them to release boxes one by one in the three environments (environment 1, environment 2, and environment 3) shown in Figure 7. Figure 8 shows the drawing of the environment. All environments included objects, release boxes, clods, stones, and slopes. The participants were asked to avoid contacts with clods and stones.

We displayed the survey-knowledge view by SolidWorks 2014 because all the participants learned how to modify the view, including changing the viewpoint and zoom levels, by using the mouse wheel in the classes. Participants modified the route-knowledge view by joysticks for crawlers the same as moving the heavy machine. Thus, participants could navigate the environment as they maneuvered the heavy machine.

We measured cognitive maps using sketch maps in PowerPoint 2013, which are widely used in the field of cognitive psychology and geography because they offer high reliability (Blades, 1990).

3.1.3 Procedure. Figure 9 shows the experimental procedure and views displayed before and during teleoperation work. First, all of the participants tried the training tasks, which includes grasping one object and transporting it to the designated box, to acquire enough skills to teleoperate in the environment, as shown in Figure 10. We used the same conditions as used in previous research to judge whether participants acquired skills (Yang, Kamezaki, Sato, Iwata, & Sugano, 2015; Sato et al., 2019). At first, all the participants tried five tasks. Then, we decided whether to finish the training tasks. We first selected the three trials of the last five trials, except for the fastest and the slowest ones. Next, we calculated the percentage of the time difference between the average work time of the three trials and each work time. If all the three percentages of time difference were less

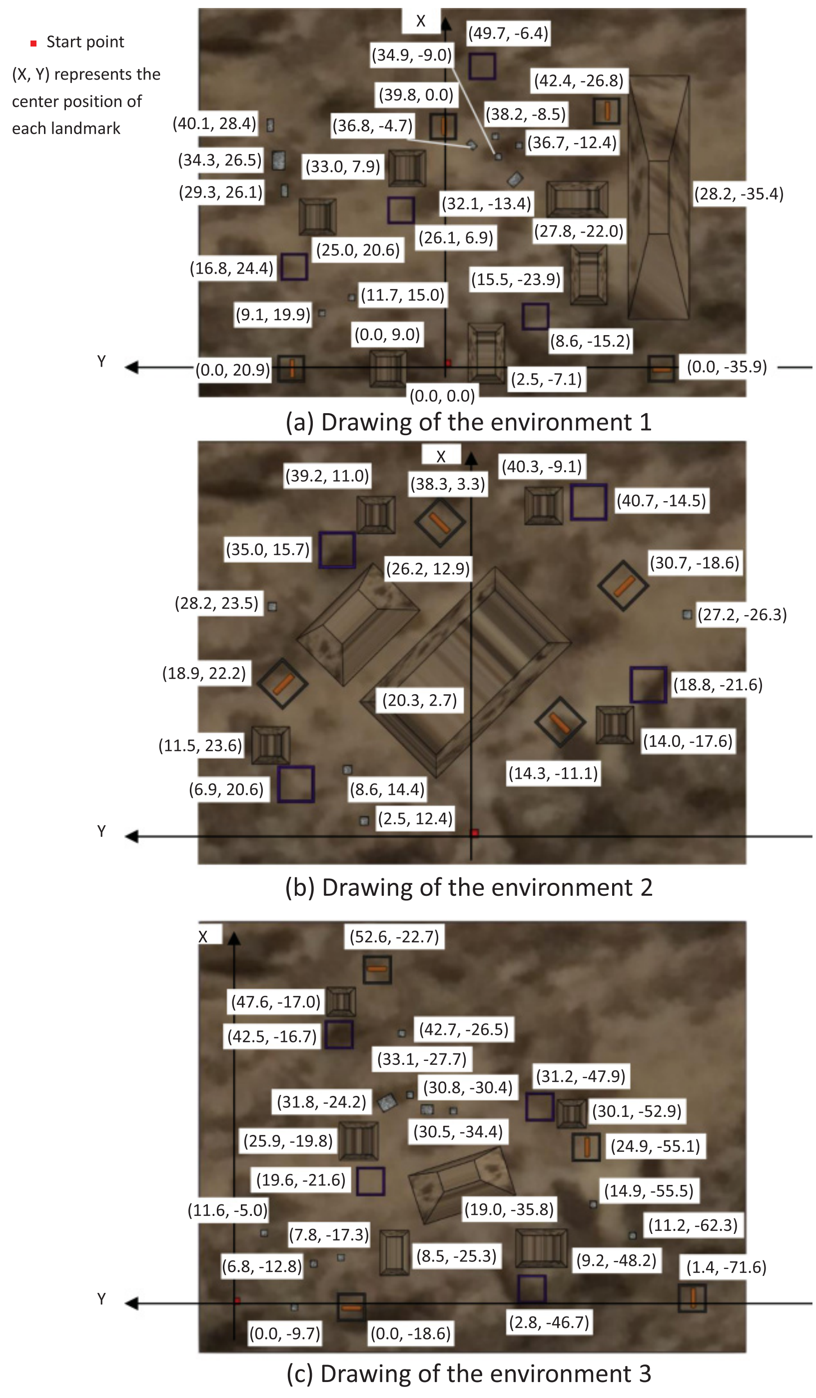


Figure 8. *Drawings of three experimental environments.*

than 5%, we finished the training tasks. If not, we continued the training tasks until all the three percentages were less than 5%.

Next, we divided the sixteen participants into two groups (Control Group and Knowledge Group) of eight participants, such that the average time of training

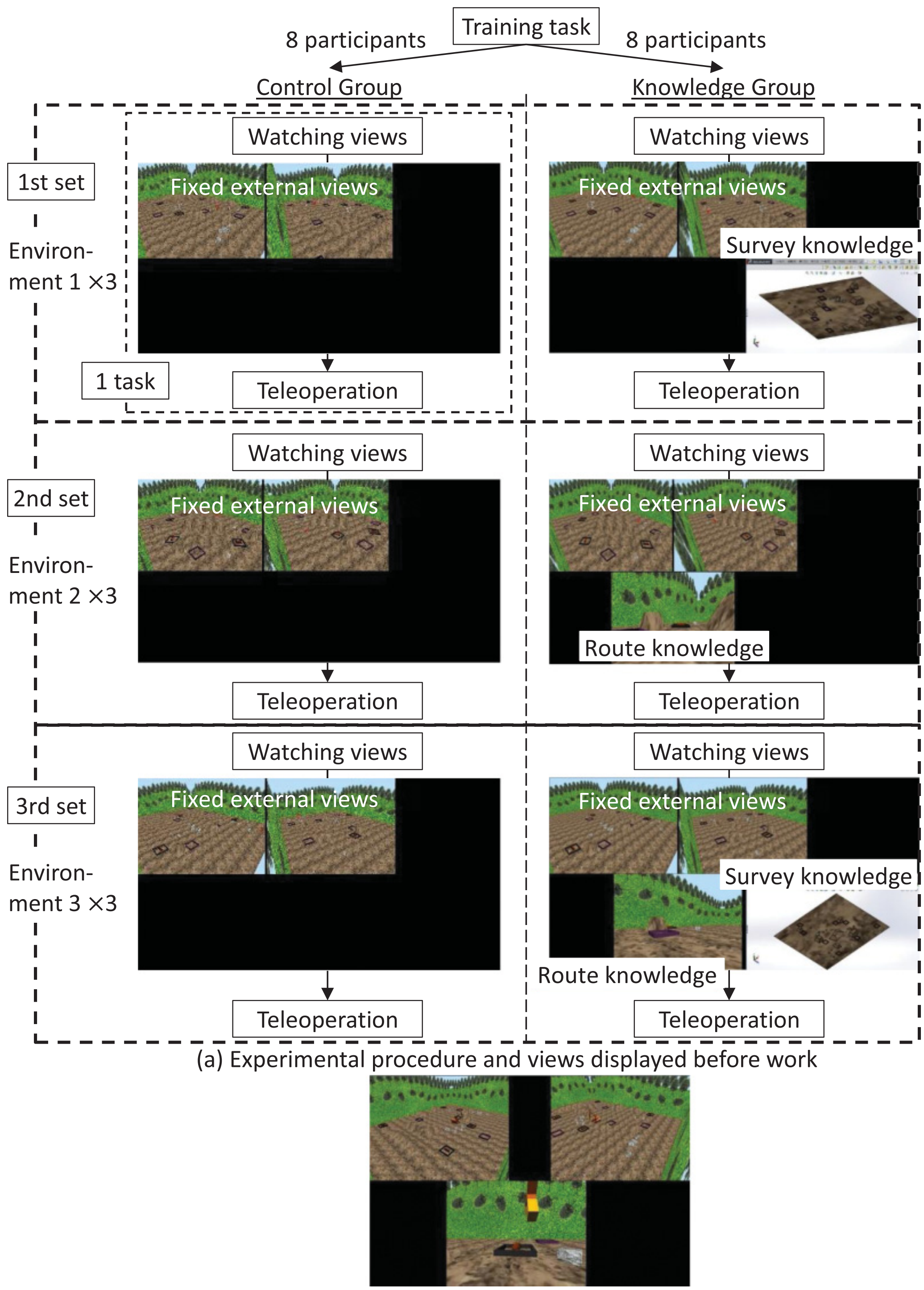
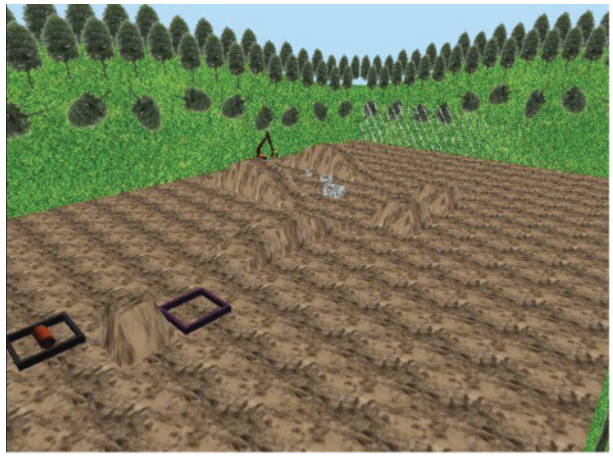
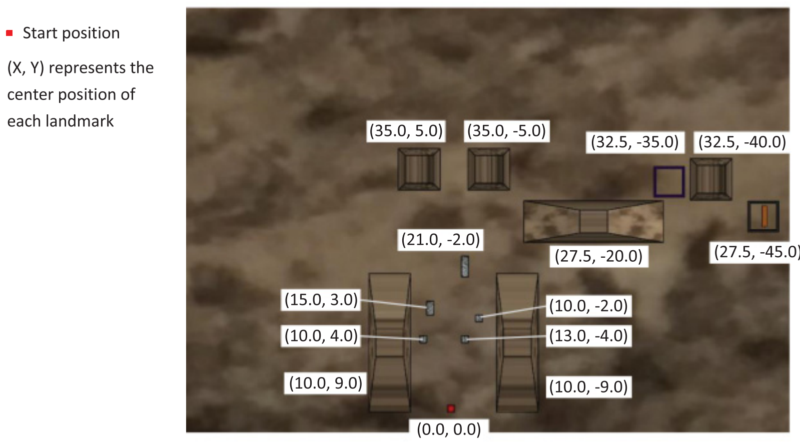


Figure 9. *Experimental procedure and views displayed before and during work.*



(a) Picture of the environment used for training tasks



(b) Drawing of the environment used for training tasks

Figure 10. *Training task environment.*

tasks in each group would be almost the same to evaluate the proposed view system. The mean work time and SD of the selected three trials of Control Group were 85.7 ± 5.9 s, and those of Knowledge Group were 82.2 ± 14.3 s. No significant differences were observed between the Control Group and Knowledge Group by Welch's t-test $t(31) = 1.08$, $p = .28$.

Then, the participants performed three sets of the experimental tasks with the different view systems before work. Each experimental task involved watching views to input environmental information before the start of teleoperation and then teleoperating heavy machinery. The participants in the Control Group watched two fixed external views of all sets before work commenced. In addition to these external views, the participants in the Knowledge Group watched a survey-knowledge view in the first set, a route-knowledge view in the second set, and two views to acquire both the survey and route knowledge in the third set before work. All participants teleoperated heavy machinery by watching a cab view and two fixed external views of all sets during work. Each set involved three experimental tasks for which the participants tried three sets under three conditions in three different environments. Operators were asked to prepare for teleoperation by watching views up to 10 min beforehand. We measured four parameters during work, that is, total task time, number of stops, moving distance, and speed during grasping,

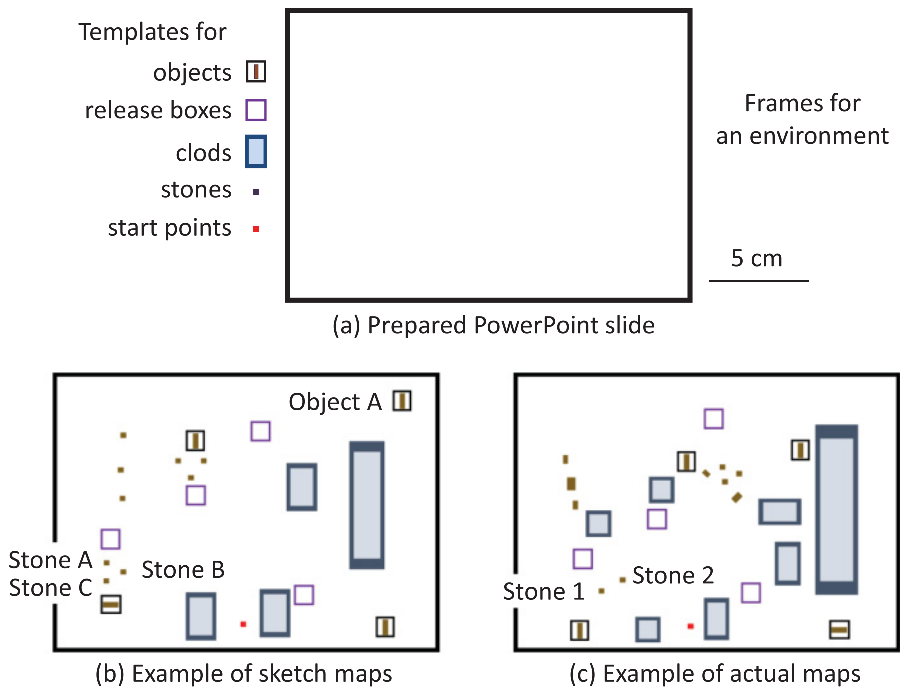


Figure 11. *Measuring cognitive maps by sketch maps.*

and asked the participants to answer interviews as done in previous research (Lum, Rosen, Lendvay, Sinanan, & Hannaford, 2009; Korte et al., 2014; Bidwell, Holloway, & Davidoff, 2014.). We define the grasping as the distance between the root of the end-effector and the ungrasped target object less than 3 m the same as was done in previous research (Yang et al., 2015; Kamezaki et al., 2016).

The participants were asked to sketch maps based on those in their minds immediately after watching views before work. The templates for each landmark (objects, releasing boxes, clods, stones, and slopes) and the frames of each environment were prepared as shown in Figure 11(a) (the templates for clods and slopes were the same). Figure 11(b) shows exemplary sketch maps acquired by the experiment, and Figure 11(c) shows an example of an actual map.

We now explain how to analyze the sketch maps. The two important aspects of cognitive maps are quantity, that is, the number of recognized landmarks, and quality, that is, the accuracy of recognized landmarks, because operators must create mental representations of work sites, including where things are located. For analytical purposes, we must identify which landmarks in cognitive maps are recognized and how they correspond to landmarks in actual maps. For example, we must identify whether object A in Figure 11(b) is recognized and which object it corresponds to in Figure 11(c). Some researchers have analyzed cognitive maps in the fields of cognitive psychology and geography (Curseu, Schalk, & Schruijer, 2010; Huynh & Doherty, 2007; Wakabayashi & Itoh, 1994). However, the cognitive maps in these studies did not include the same landmarks as in this study because these studies are usually performed in cities, which have landmarks with specific names, such as Tokyo Station. Thus, no analytical methods can identify which landmarks in cognitive maps are recognized and correspond to landmarks in actual maps if maps include the same landmarks. Therefore, we developed an analytical method to identify which landmarks are recognized in cognitive maps and their correspondence with actual maps.

If the participants draw landmarks randomly, we assume that the probability that they will be drawn close

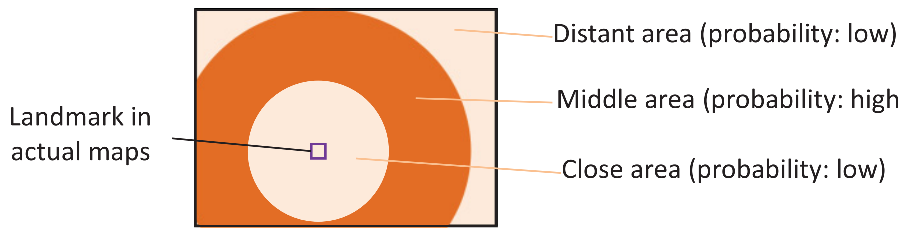


Figure 12. *Probability that landmarks are drawn in each area in sketch maps.*

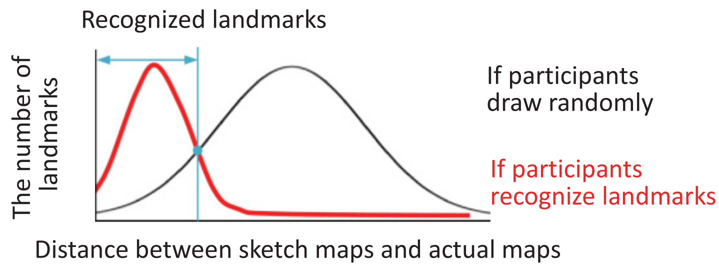


Figure 13. *Histogram of distance between sketch maps and actual maps.*

or distant from landmarks on actual maps can be low and that of being drawn at a middle distance from actual landmarks can be high because middle area has the most surface area as illustrated in Figure 12. Thus, we assume that the histogram of distance can be Gaussian, like the black curve shown in Figure 13. On the other hand, if the participants recognize landmarks, we assume that this histogram can obey the red curve shown in Figure 13 because they can draw landmarks close to those in actual maps. Thus, we recognize landmarks in cognitive maps as actual landmarks if the distance is less than that of an intersection in a histogram. We calculate the width of each column of a histogram using the Freedman–Diaconis rule because this rule is calculated on a quartile basis. We calculate the distance between the center point of actual and drawn landmarks. If there are some duplicate landmarks, we need to eliminate them. For example, in the case of three stones in the bottom left of Figure 11(b), we calculate the distances between Stone A and all the ten stones in the actual maps, and this calculation is done for Stones B and C as well. If the threshold distance is 3.5, and the distances between stones in the drawn maps and those in actual maps are as in Table 1, then Stones A, B, and C are recognized as both Stones 1 and 2, which are duplicate landmarks. Therefore, we need to eliminate those duplicate landmarks to minimize the total distance. In this case, the condition with Stone B recognized as Stone 1 and Stone C recognized as Stone 2 has the minimum distance. Thus, we recognize Stone B as Stone 1 and Stone C as Stone 2, and eliminate the others.

Table 1. *Example of the Results of Distance*

	Stone 1	Stone 2
Stone A	2.5	3.4
Stone B	1.3	2.9
Stone C	2.3	3.1

3.2 Results and Discussion

In this section, first we explain the results and discussion for cognitive maps, and after that we explain them for work efficiency. We verified normal distribution by the Shapiro–Wilk test for all the analysis using Welch's t-test.

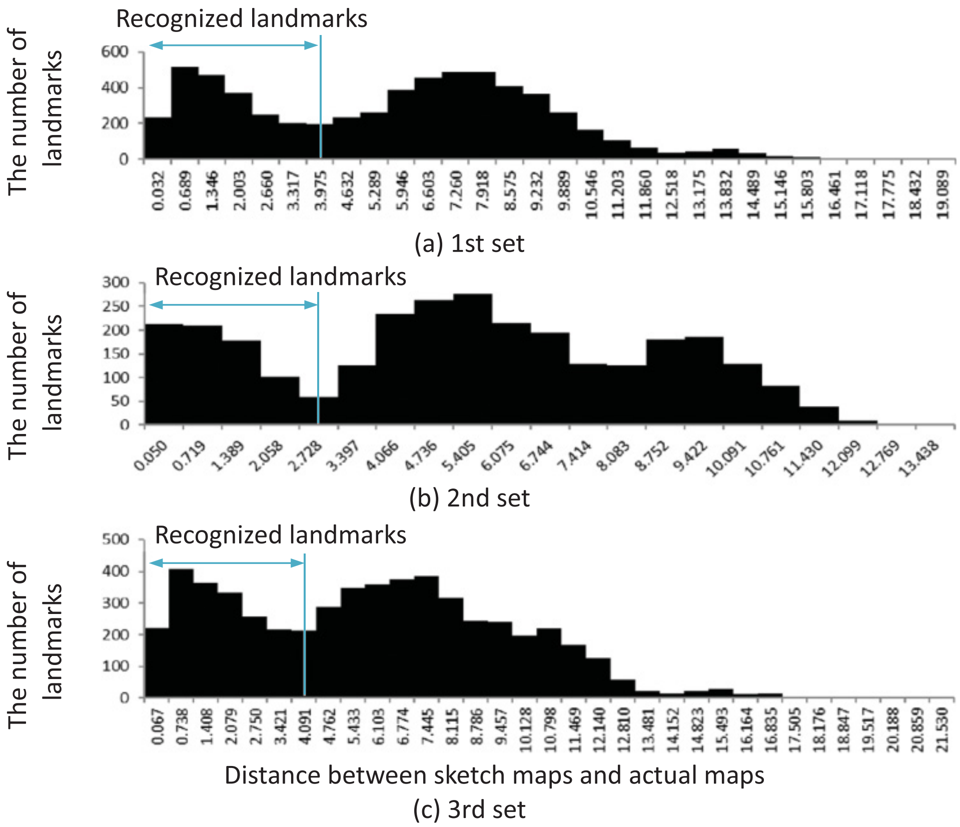


Figure 14. *Results of histogram of the distance between sketch maps and actual maps.*

3.2.1 Results on Recognized Landmarks. Figure 14 shows the histogram of distance between all landmarks in the sketch maps and all the same landmarks in the actual maps in each set. These histograms have at least two peaks, and the shape of them is similar to the one illustrated in Figure 13 that was our hypothesized histogram. Thus, we determined the intersection as the midpoint of the column of local minimum and set the distance of the intersection as the threshold.

3.2.2 Effects for Cognitive Maps by Survey Knowledge (First Set). Figure 15 shows the percentage of recognized landmarks (recognized landmarks/the number of landmarks in the actual map) and the average error distance between recognized landmarks and those in the sketch maps for all same landmarks in the first set. The Mann–Whitney U test indicates that participants in the Control Group recognized more stones $U = 177$, $p = .02$ than those in the Knowledge Group significantly, and a marginally significant difference is observed between total of each group $U = 203.5$, $p = .08$. Moreover, Welch's t-test indicates that participants in the Knowledge Group recognized release box $t(180) = 2.26$, $p = .02$, slopes $t(39) = 2.55$, $p = .01$, and total $t(889) = 2.55$, $p = .01$ more accurately than those in the Control Group, and a marginally significant difference is observed between the average distances in the clods $t(204) = 1.89$, $p = .06$ of the Control and Knowledge Groups. These results suggest that watching a survey-knowledge view before work can help operators remember environments accurately, but it cannot input environmental information in terms of quantity.

We now discuss two points, which are (1) the reason why percentage of recognized landmarks in the Control Group was higher than that in the Knowledge Group, (2) the reason why the average error distance of objects and stones was not improved significantly.

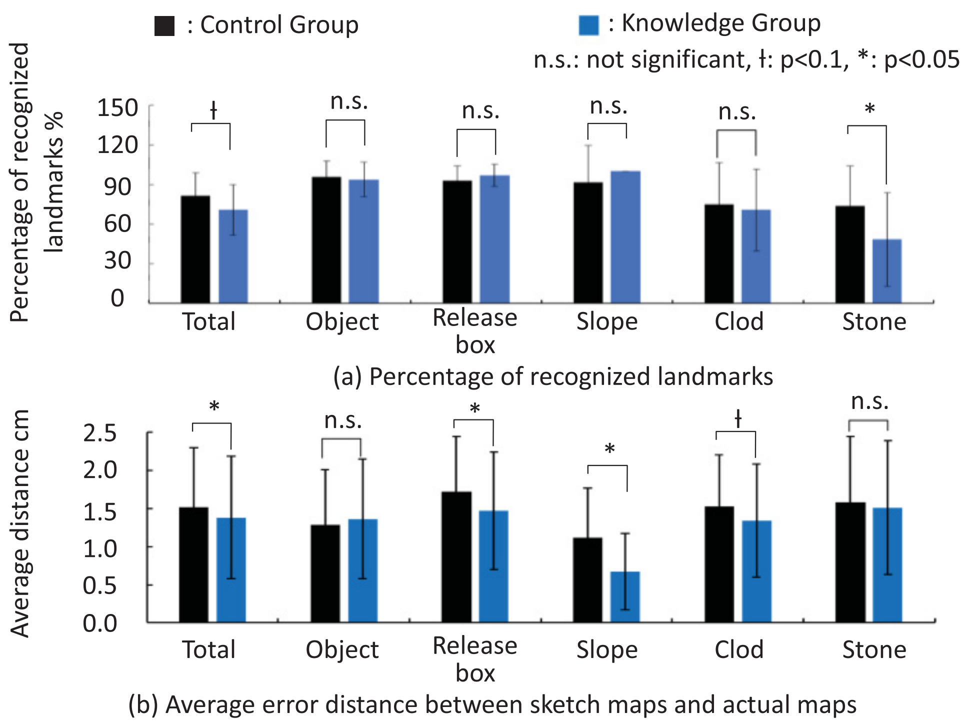


Figure 15. *Results of sketch maps for the first set (survey knowledge).*

Percentage of recognized landmarks in Control Group was higher than that in Knowledge Group: This could be caused because participants in the Control Group tried to remember all the landmarks; even though those in the Knowledge Group tried to remember important landmarks. Stones are not as important as objects or releasing boxes for path planning because they are small obstacles. Figure 16 shows the actual map and the shortest path of all the trials (black line) in the environment 1. Stones 1 and 2 are close to the shortest path and Stones 3 to 10 are far from the shortest path. Figure 17 shows the percentage of recognized Stones 1 and 2, and 3 to 10. The Mann–Whitney U test indicates that no significant differences are observed between the Knowledge Group and the Control Group in Stones 1 and 2, $U = 263$, $p = .47$, but a significant difference is observed in Stones 3 to 10, $U = 172.5$, $p = .01$. Moreover, "clod 4" is far from the shortest path as shown in Figure 16, and the Mann–Whitney U test indicates that a significant difference between the recognized percentage of each group, $U = 192$, $p = .02$. Furthermore, participants in the Knowledge Group recognized almost all of landmarks close to the shortest path such as objects, release boxes, and slopes, as shown in Figure 16, though those in the Control Group recognized unimportant landmarks including Stones 3 to 10, as shown in Figure 11(a) and 17. This difference could lead the Knowledge Group to recognize more accurately, as shown in Figure 11(b).

Average error distance of objects and stones were not improved significantly: This could be caused by the importance of the landmarks. Objects could be the most important landmark in this experiment because the task included grasping objects. All of the participants tried to remember objects from the interview. On the other hand, stones are not so important, as explained in this section. All the participants tried to remember stones with little effort from the interview. Those results suggest that the average error distance of objects and stones were not improved significantly because of the importance of landmarks.

These results suggest that watching a survey-knowledge view can help operators remember important

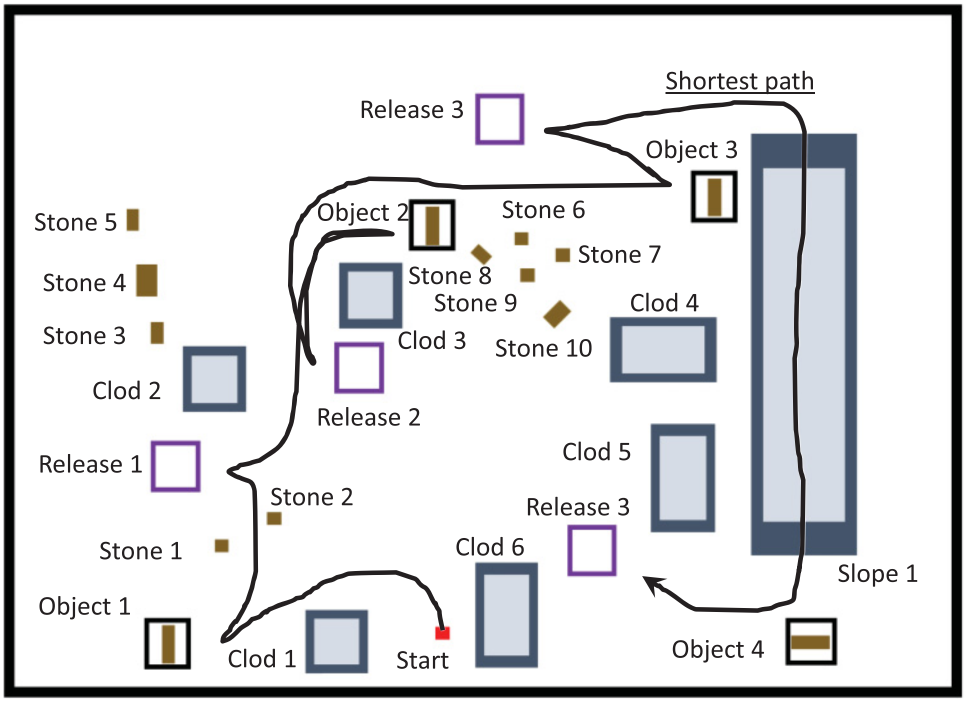


Figure 16. *The actual map and the shortest path of all trials in the environment 1.*

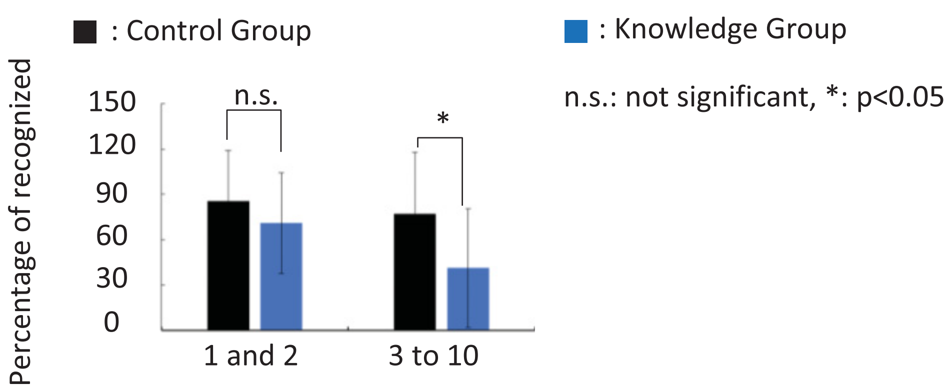


Figure 17. *Percentage of recognized Stones 1 and 2, and 3 to 10.*

landmarks accurately because of improvement in accuracy, as shown in Figure 11(b) and recognizing only important landmarks, as shown in Figures 11(a) and 17.

3.2.3 Effects for Cognitive Maps by Route Knowledge (Second Set). Figure 18 shows the percentage of recognized landmarks and the average error distance between recognized landmarks and those in the sketch maps for all same landmarks in the second set. The Mann–Whitney U test indicates that participants in the Control Group recognized more stones, $U = 202$, $p = .07$, and total, $U = 206$, $p = .09$, than those in the Knowledge Group, by a marginally significant number. Welch's t-test indicates that a marginally significant difference was observed in objects, $t(161) = 1.68$, $p = .10$. These results suggest that watching a

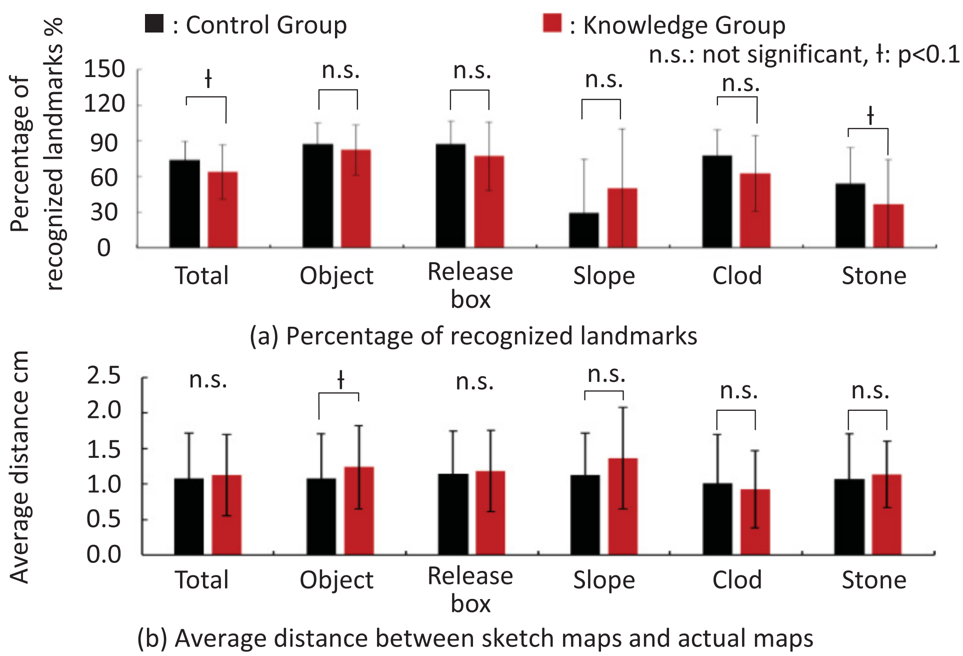


Figure 18. *Results of sketch maps for the second set (route knowledge).*

route-knowledge view before work cannot input environmental information in terms of both quality and quantity. We will discuss this with results on work efficiency in Subsection 3.2.6.

3.2.4 Effects for Cognitive Maps by Survey and Route Knowledge (Third Set). Figure 19 shows the percentage of recognized landmarks and the average distance between recognized and actual landmarks in the third set. The Mann–Whitney U test indicates that participants in the Knowledge Group recognized significantly more objects, $U = 201$, $p = .04$, than those in the Control Group, and a marginally significant difference was observed in slopes, $U = 216$, $p = .08$. Welch's t-test indicates that participants in the Knowledge Group recognized release boxes, $t(176) = 2.24$, $p = .03$, clods, $t(153) = 2.41$, $p = .02$, stones, $t(269) = 4.99$, $p < .001$, and total, $t(793) = 5.65$, $p < .001$, more accurately than those in the Control Group. These results suggest that watching a survey-knowledge view and a route-knowledge view before work can input environmental information in terms of both quality and quantity. These results are different from results explained in Subsections 3.3.2 and 3.3.3, and we will address this further. This difference in results could be caused because the participants in the Control Group could not remember environment 3 as well as environment 1. Environment 1 has 4 objects, 4 release boxes, 1 slope, 6 clods, and 10 stones, and environment 3 has 4 objects, 4 release boxes, 1 slope, 5 clods, and 10 stones. Thus, the only difference is that environment 3 has 1 less clod than environment 1. However, the participants in the Control Group recognized about 81% in the first set (environment 1), but they recognized about 70% in the third set (environment 3). The Mann–Whitney U test indicates that this difference is significant, $U = 174.5$, $p = .02$. Moreover, objects are important landmarks because the task requires participants to grasp the objects. However, the participants in the Control Group remembered only 74%. These results indicate that a view to acquire knowledge from the survey perspective can help operators recognize important landmarks stably, even for different environments.

3.2.5 Effects for Work Efficiency by Survey Knowledge (First Set). Figure 20 shows (a) the task time,

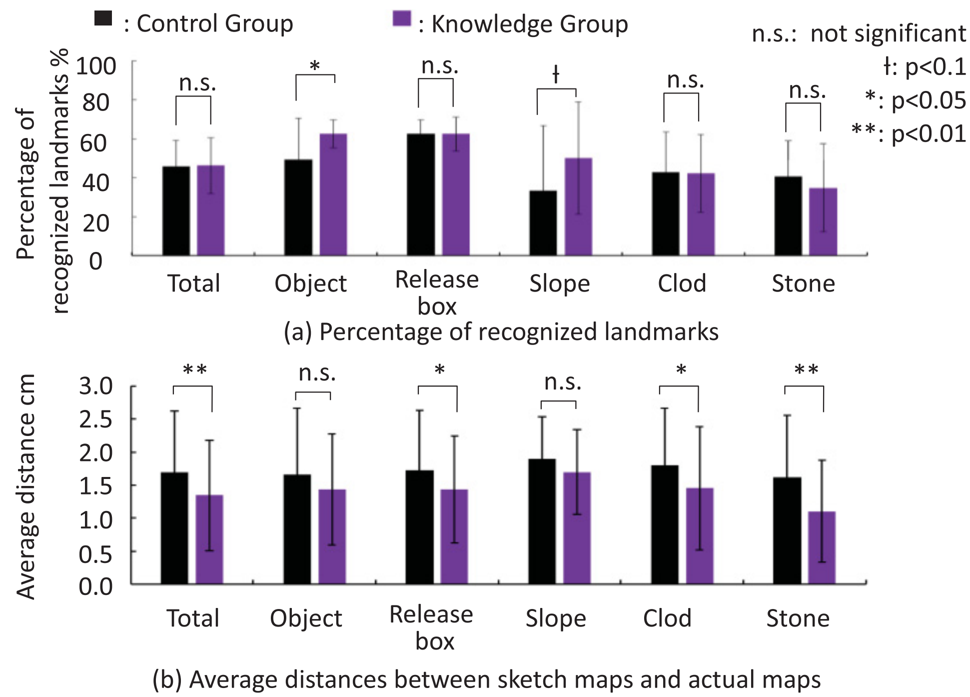


Figure 19. *Results of sketch maps for the third set (both knowledge).*

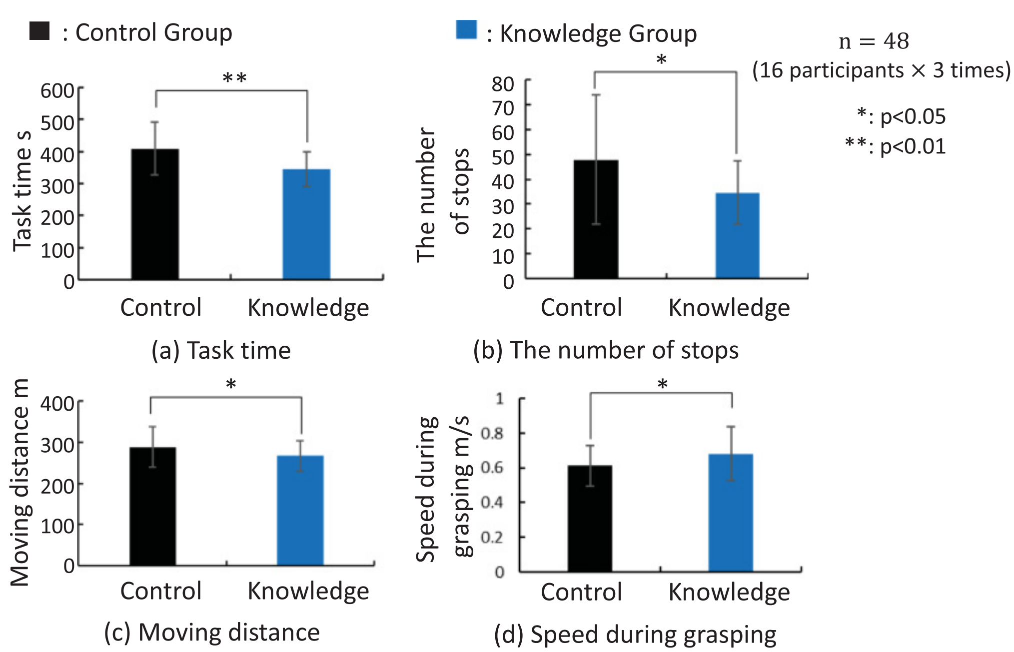


Figure 20. *Results of the first set (survey knowledge).*

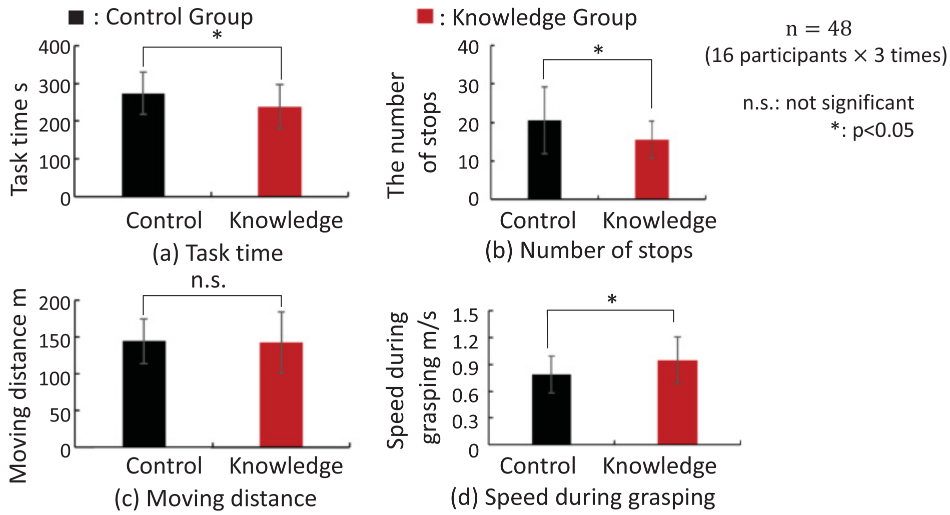


Figure 21. *Results of the second set (route knowledge).*

(b) the number of stops, (c) the moving distance, and (d) the speed during grasping for the first set. Welch's t-test indicates that the task time $t(40) = 3.22$, $p = .003$, the number of stops $t(33) = 2.26$, $p = .03$ and moving distance $t(43) = 210$, $p = .04$ in the Knowledge Group are significantly less than those in the Control Group and that the speed during grasping $t(44) = 2.23$, $p = .03$ in the Knowledge Group is significantly faster than that in the Control Group. These results suggest that watching a survey-knowledge view before work can decrease the task time the moving distance, and the number of stops. Moreover, the Knowledge Group remembered more accurately than in the Control Group significantly, and the Knowledge Group focused on remembering only important landmarks, as explained in Subsection 3.2.1. Thus, those results suggest that watching a survey-knowledge view before work can help operators plan short paths and focus on remembering landmarks close to the short paths.

3.2.6 Effects for Work Efficiency by Route Knowledge (Second Set). Figure 21 shows (a) the task time, (b) the number of stops, (c) the moving distance, and (d) the speed during grasping for the second set. Welch's t-test indicates that the task time $t(46) = 2.12$, $p = .04$, and the number of stops $t(36) = 2.50$, $p = .02$, for the Knowledge Group were significantly less than that in the Control Group and that the speed during grasping $t(44) = 2.37$, $p = .02$, in the Knowledge Group is significantly faster than that in the Control Group These results suggest that watching a route-knowledge view can decrease task time and the number of stops and increase speed during grasping.

3.2.7 Effects for Work Efficiency by Survey and Route Knowledge (Third Set). Figure 22 shows (a) the task time, (b) the number of stops, (c) the moving distance, and (d) the speed during grasping for the third set. Welch's t-test indicates that the task time $t(45) = 4.04$, $p < .001$, and the moving distance $t(46) = 3.25$, $p = .002$, in the Knowledge Group are significantly less than those in the Control Group and a marginally significant difference is observed between the number of stops $t(31) = 1.98$, $p = .06$, in the Knowledge Group and that in the Control Group. Moreover, Welch's t-test indicates that speed during grasping $t(34) = 3.00$, $p = .005$, in the Knowledge Group is significantly faster than that in the Control Group These results suggest that watching a survey-knowledge view and route-knowledge view before work can decrease task

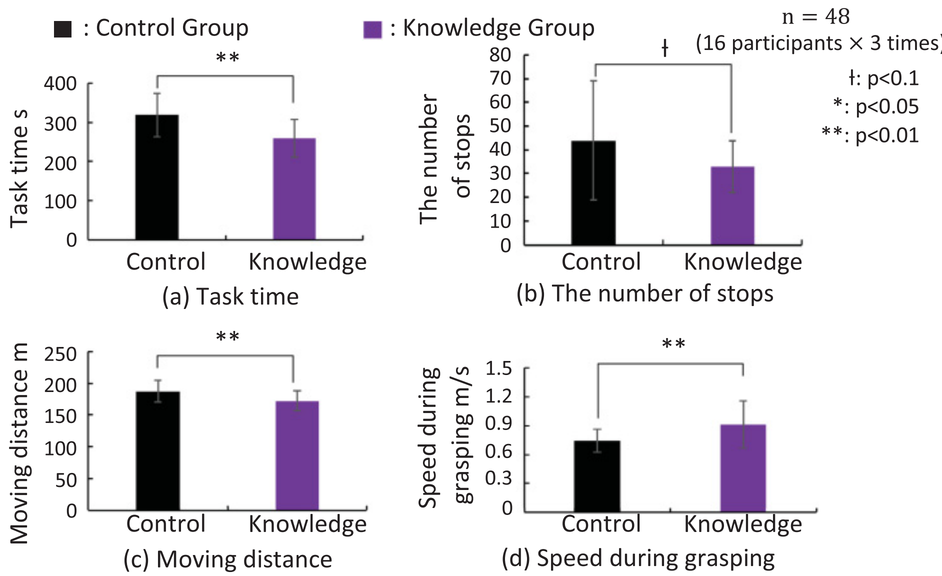


Figure 22. *Results of the third set (both knowledge).*

Table 2. *Summary of the Results from each Knowledge (√: $p < 0.05$, ✗: $p \geq 0.05$)*

	Task time	The number of stops	Moving distance	Speed during grasping
Survey knowledge	√	√	√	√
Route knowledge	√	√	✗	√
Both knowledge	√	✗	√	√

time, moving distance, and the number of stops, and increase speed during grasping.

3.2.8 Discussion on Work Efficiency. We summarize the effects of the proposed view system in Table 2 and Figure 23. Table 2 shows whether differences between Contol Group and Knowledge Group are significant. Figure 23 shows the normalized results of each set. The standard deviation bars of the Control Group were also normalized, so each result has three standard deviation bars because each result and standard deviation ratio is different from each set. For example, the blue standard deviation bars of the task time is normalized by dividing standard deviation of the first set by the average task time of that set. The task time, the number of stops, the moving distance, and the speed during grasping can be improved with survey knowledge Moreover, the task time, the number of stops, and the speed during grasping can be improved with route knowledge. The task time, the moving distance, and the speed during grasping can be improved with survey and route knowledge. Thus, we discuss the following four points.

Improvement of speed during grasping with survey perspective: This could be caused by the positions during grasping. Grasping objects from slanting positions, which means the angle between the object and the heavy machine θ is small, can increase the number of stops owing to difficulties in grasping, as illustrated in Figure 24. Allowable error to grasp the object is $(\mathrm{WG}\sin\theta - \mathrm{WO})$, when the width of grapples is WG and the width of the object is WO. Thus, the more parallel to the object, the more difficult grasping the objects is, which can increase the number of stops. Figure 25 shows the

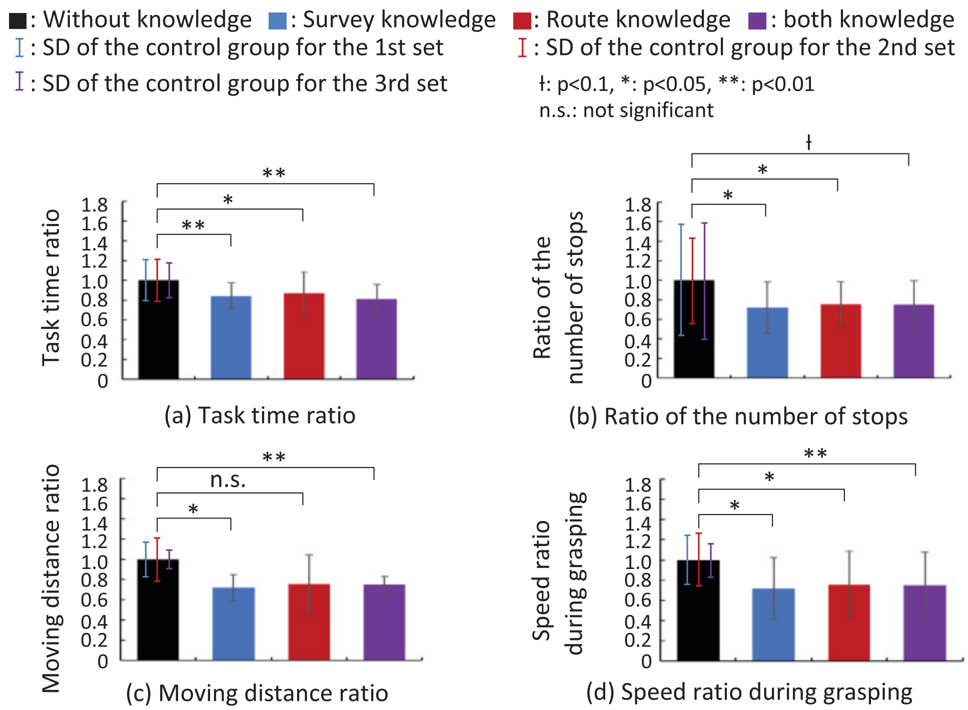


Figure 23. *Summary of the normalized results.*

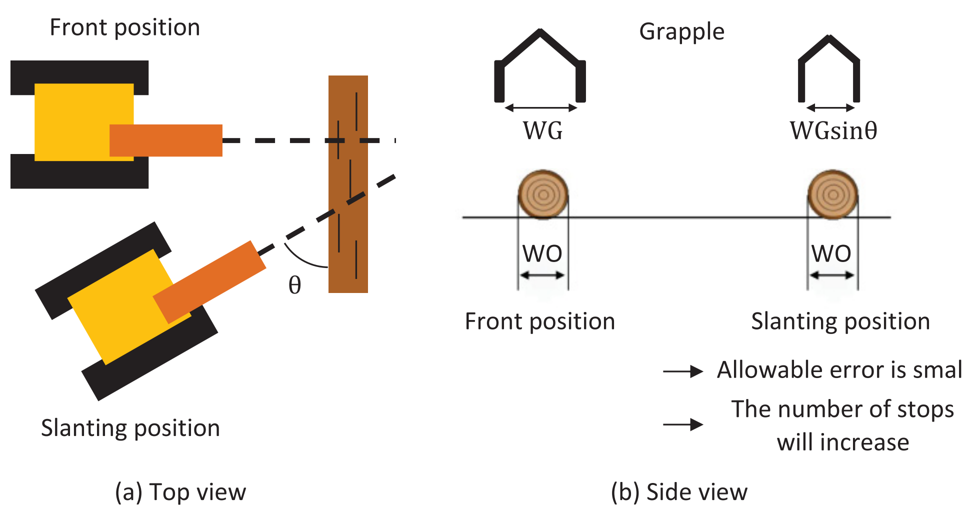


Figure 24. *Difficulties in grasping from slanting positions.*

results of the number of stops during grasping the circled object in Figure 7(a). Welch's t-test indicates that the number of stops, $t(29) = 2.17$, $p = .04$, for the Knowledge Group is significantly less than that in the Control Group. Moreover, the number of trials with $\theta < 60°$ in the Control Group was 5, but the one in the Knowledge Group was 0. These results suggest that watching a survey-knowledge view before work can help

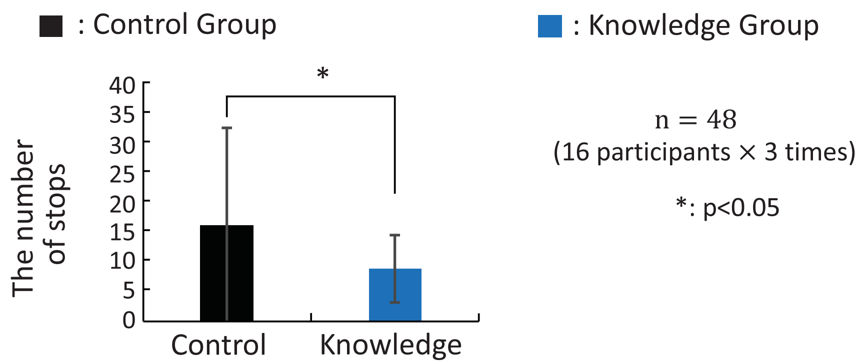


Figure 25. *Results of the number of stops during grasping the circled object in 7a.*

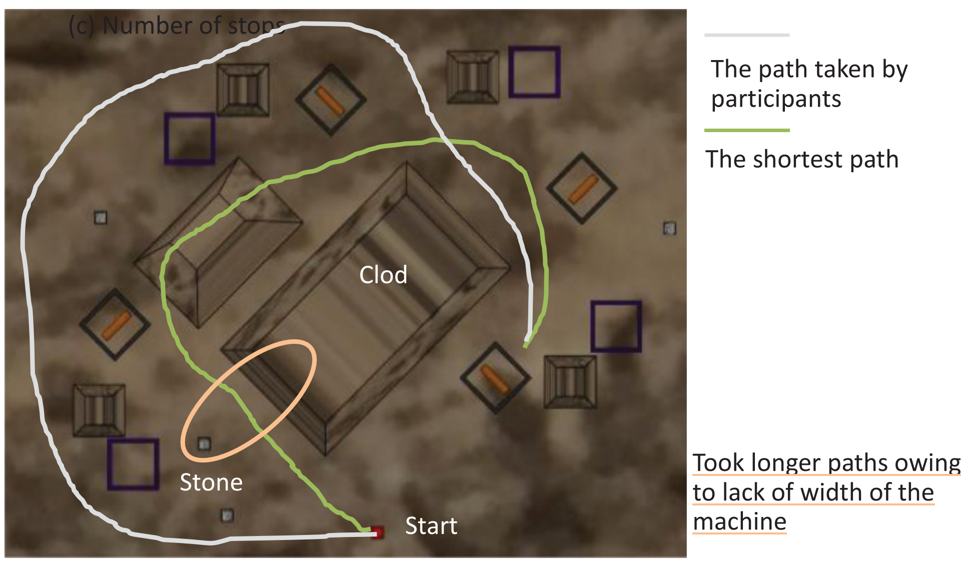


Figure 26. *Comparison of the shortest path and the path taken by participants.*

operators grasp objects from places where θ is large and prevent workers from stopping, which may cause the improvement of speed during grasping.

Unimprovement of moving distance with route perspective: This could be caused because participants could hardly recognize the relationship between the heavy machines and distance between obstacles. Figure 26 shows the shortest path (the green line) and the path participants in the Knowledge Group took (the white line) in the second set. Participants recognized the distance between obstacles circled in Figure 26 from the interview, even though they took the longer path (the white line). This was because they could not recognize the relationship between the heavy machines and the distance between obstacles. That could be the reason why the moving distance was not decreased significantly.

Improvement of the task time, the number of stops, and speed during grasping without improvement of cognitive maps by watching a route-knowledge view: This could be caused because participants in the Knowledge Group tried to remember the working strategies, including how to grasp objects, rather than remembering the environment. The shortest path in environment 2 was just clockwise and there were a few obstacles, so environment 2 was not so complicated, as shown in Figure 26.

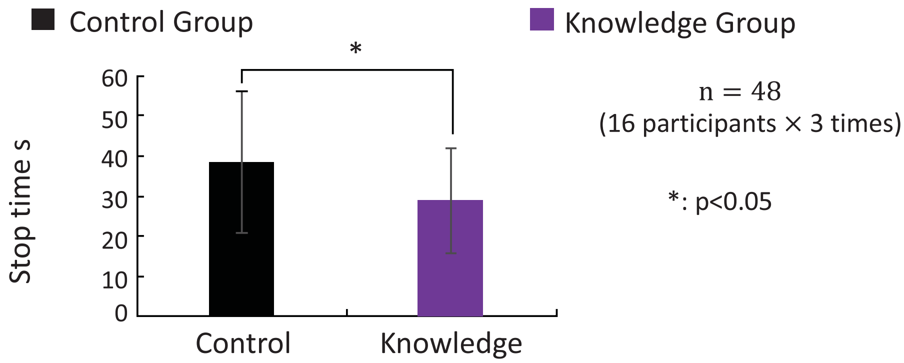


Figure 27. *Results of the stop time in the third set.*

Participants in the Knowledge Group focused on planning the working strategies from the interview. These results suggest that participants in the Knowledge Group could try to focus on planning rather than remembering the environment because the environment was not so complicated. This could be why watching a route-knowledge view before work could not input environmental information in terms of both quality and quantity. These results suggest that watching a route-knowledge view can help operators plan working strategies, which may lead to reducing the task time, the number of stops, and increasing speed during grasping.

Here, we discuss the effects of a route-knowledge view for complicated environments. Remembering the environment from a route-knowledge view can be more difficult than a survey-knowledge view, especially for complicated environments. Thus, we assume that operators can focus on planning rather than remembering the complicated environment.

Unimprovement of the number of stops with survey and route perspective: We discuss the point that the number of stops was improved with each knowledge, but not improved with knowledge of both. Figure 27 shows the results of the stop time in the third set. Welch's t-test indicates that the stop time in the Knowledge Group is significantly smaller than that in Control Group, $t(42) = 2.10$, $p = .04$. Thus, the proposed view system could improve the stop time, though it might not be able to improve the number of stops. Therefore, these results suggested that the proposed view system could improve the stop time.

3.2.9 Discussion on Practical Use. We now discuss how to apply the proposed view system for practical use because the system was developed in the simulator. First, we need a 3D map of a disaster site for this system. This 3D map can be acquired by using drones (Nex & Remondino, 2014; Spranger, Heinke, Becker, & Labudde, 2016). Then, both a survey-knowledge view and a route-knowledge view can be displayed by using common 3D computer graphic software, including Blender and Unity.

4 Conclusion

We developed a view system based on human spatial cognition to provide teleoperators with environmental information prior to commencement of work. We displayed two views, an external view from any viewpoint to show survey knowledge and another from an operator's viewpoint, representing the route knowledge, which could be modified based on the operator's intent. We conducted experiments using a simulator to evaluate a proposed view system. The results suggest that this system can increase work efficiency and help operators plan paths. The results showed that the task time, moving distance, and number of stops could be reduced by watching the survey-knowledge view. Furthermore, the task time, stops, and speed degradation during grasping could be prevented by watching a route-knowledge view. From the analysis of cognitive maps, the survey-knowledge view was found to help operators remember important landmarks stably, which might lead to reducing task time, preventing stops, and choosing shorter paths. Moreover, the route-knowledge view could help operators plan their work, possibly allowing them to prevent stops and speed degradation during grasping.

We will consider providing the proposed view system in 3D because views during work can be 3D in the future, with an interface to modify the viewpoint. In future work, we hope to analyze cognitive maps in terms of other factors, such as the dimensions and rotation of landmarks. Moreover, we will examine the effects of a route-knowledge view for complicated environments. Furthermore, we will develop systems to provide environmental information in case the environment changes, including head-mounted and other immersive displays.

Acknowledgments

This research was supported by JSPS KAKENHI; under grant numbers 26870656 and 16K06196, Casio Science Promotion Foundation, and the Research Institute for Science and Engineering, Waseda University.

References

Ankomah, P., Crompton, J.& Baker, D. (1996). Influence of cognitive distance in vacation choice. *Annals of Tourism Research*, *23*(1), 138–150.

Asama, H., & Ueki, C. (2013). A proposal for unmanned construction. *Council on Competitiveness-Nippon*. Japanese. URL Retrieved from http://www.cocn.jp/report/thema50-L2.pdf, accessed on 10/23/2019

Bidwell, J., Holloway, A., & Davidoff, S. (2014). Measuring operator anticipatory inputs in response to time-delay for teleoperated human-robot interfaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1467–1470.

Blades, M. (1990). The reliability of data collected from sketch maps. *Journal of Environmental Psychology*, *10*(4), 327–339.

Burnett, G. (2003). A road-based evaluation of a head-up display for presenting navigation information. *Proceedings of International Conference on Human–Computer Interaction*, *3*, 180–184.

Cadwallader, M. (1975). A behavioral model of consumer spatial decision making. *Journal of Economic Geography*, *51*(4), 339–349.

Casner, S. (2005). The effect of GPS and moving map displays on navigational awareness while flying under VFR. *Journal of Applied Aviation Studied*, *5*(1), 153–165.

Chayama, K., Fujioka, A., Kawashima, K., Yamamoto, H., Nitta, Y., et al. (2014). Technology of unmanned construction system in Japan. *Journal of Robotics and Mechatronics*, *26*(4), 403–417.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159.

Crundall, D., & Underwood, G. (1998). Effects of experience and processing demands on visual information acquisition in drivers. *Ergonomics*, *41*(4), 448–458.

Curseu, P., Schalk, R., & Schruijer, S. (2010). The use of cognitive mapping in eliciting and evaluation group cognitions. *Journal of Applied Social Psychology*, *40*(5), 1258–1291.

Draper, J., Handel, S., & Hood, C. (1991). Three experiments with stereoscopic television: When it works and why. *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, 1047–1052.

Edward, T. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*(4), 189–208.

Endsley, M. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, *37*(1), 32–64.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G∗Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G∗Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.

Fong, T., Thorpe, C., & Baur, C. (2003). Multi-robot remote driving with collaborative control. *IEEE Transactions on Industrial Electronics*, *50*(4), 699–704.

Garling, T., & Golledge, R. (1993). *Behavior and environment: Psychological and geographical approaches*. Amsterdam: North Holland.

Golledge, R., & Stimson, R. (1997). *Spatial behavior: A geographic perspective*. New York: Guilford Press.

Heiling, M., & Futuro, R. (1992) The cinema of the future. *Presence: Teleoperators and Virtual Environments 1*(3), 279–294.

Hiramatsu, Y., Aono, T., & Nishino, M. (2002). Disaster restoration work for the eruption of Mt Usuzan using an unmanned construction system. *Advanced Robotics*, *16*(6), 505–508.

Huynh, N., & Doherty, S. (2007). Digital sketch-map drawing as an instrument to collect data about spatial cognition.*Cartographica*, *42*(4), 285–296.

Kamezaki, M., Yang, J., Iwata, H., & Sugano, S. (2014). A basic framework of virtual reality simulator for advancing disaster response work using tele-operated work machines. *Journal of Robotics and Mechatronics*, *26*(4), 486–495.

Kamezaki, M., Yang, J., Iwata, H., & Sugano, S. (2016). Visibility enhancement using autonomous multicamera controls with situational role assignment for teleoperated work machines. *Journal of Field Robotics*, *33*(6), 802–824.

Kawatsuma, S., Fukushima, M., & Okada, T. (2012). Emergency response by robots to Fukushima-Daiichi accident: Summary and lessons learned. *Industrial Robot*, *39*(5), 428–435.

Kiribayashi, S., Yakushigawa, K., & Nagatani, K. (2018). Design and development of tether-powered multirotor micro unmanned aerial vehicle system for remote-controlled construction machine. *Field and Service Robotics*, *5*, 637–648.

Korte, C., Nair, SS., Nistor, V., Low, TP., Doarn, CR., & Schaffner, G. (2014). Determining the threshold of time-delay for teleoperation accuracy and efficiency in relation to telesurgery.*Telemedicine and e-Health*, *20*(12), 1076–1086.

Lay, T., & Kanamori, H. (2011). Insights from the great 2011 Japan earthquake. *Physics Today*, *64*(12), 33–39.

Levin, M., Jankovic, I., & Palij, M. (1982). Principles of spatial problem solving. *Journal of Experimental Psychology: General*, *111*(2), 157–175.

Lum, L., Rosen, J., Lendvay, T., Sinanan, M., & Hannaford, B. (2009). Effect of time delay on telesurgical performance. IEEE International Conference on Robotics and Automation, 4246–4252.

Miller, G. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information.*Psychological Review*, *63*(2), 81–97.

Moteki, M., Akihiko, N., Yuta, S., Mishima, H., & Fujino, K. (2016). Research on improving work efficiency of unmanned construction. *Proceedings of the 33rd International Symposium on Automation and Robotics in Construction*, 478–486.

Moteki, M., Fujino, K., Ohtsuki, T., & Hashimoto, T. (2011). Research on visual point of operator in remote control of construction machinery. *Proceedings of the International Symposium Automation and Robotics in Construction*, 532–537.

Nakada, S., & Fujii, T. (1993). Preliminary report on the activity at Unzen Volcano (Japan), November 1990–November 1991: Dacite lava domes and pyroclastic flows. *Journal of Volcanology and Geothermal Research*, *54*(3–4), 319–333.

Narzt, W., Pomberger, G., Ferscha, A., Kolb, D., Muller, R., et al. (2005). *Universal Access in the Information Society*, *4*(3), 177–187.

Nex, F., & Remondino, F. (2014). UAV for 3D mapping applications: A review. *Applied Geomatics*, *6*(1), 1–15.

Nitta, Y. (2012). Evolution of unmanned construction system playing important role in the disaster recovery. *Civil Engineering*, *67*(4), 16–23. Japanese.

Passini, R. (1984). Spatial representations, a wayfinding perspective. *Journal of Environmental Psychology*, *4*(2), 153–164.

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

Sato, R., Kamezaki, M., Yamada, M., Hashimoto, T., Sugano, S., & Iwata, H. (2019). Experimental investigation of optimum and allowable range of side views for teleoperated digging and release works by using actual construction machinery. *Proceedings of the IEEE/SICE International Symposium on System Integration*, 788–793.

Spranger, M., Heinke, F., Becker, S., & Labudde, D. (2016). Towards drone-assisted large-scale disaster response and recovery. *Proceedings of International Conference on Advances in Computation, Communication and Services*, 5–10.

Wakabayashi, Y., & Itoh, S. (1994). Reexamination of the components of distortions in cognitive maps. *Journal of Geography*, *103*(3), 221–232.

Wang, J., Lewis, M., & Hughes, S. (2004). Gravity-referenced attitude display for teleoperation of mobile robots. *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting*, 2662–2666.

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