Design and Evaluation of Human-Centered Visualization Interfaces in Construction Teleoperation

Yeon Chae[1]; Samraat Gupta[2]; and Youngjib Ham, Ph.D., A.M.ASCE[3]

[1]Ph.D. Student, Dept. of Construction Science, Texas A&M Univ., College Station, TX. Email: yeonchae62@tamu.edu
[2]Master's Student, Dept. of Computer Science and Engineering, Texas A&M Univ., College Station, TX. Email: samraatg@tamu.edu
[3]History Maker Homes Endowed Associate Professor, Dept. of Computer Science and Engineering and Dept. of Construction Science, Texas A&M Univ., College Station, TX. Email: yham@tamu.edu

ABSTRACT

Teleoperation is widely used in hazardous and uncertain site settings, allowing scheduled procedures to be carried out across long distances while workers are away from the sites. Teleoperators in off-sites collect both the site information and feedback from the interfaces which provide synthesized information that a robot collects. This interface mainly conveys visionary information for the operator's intuitiveness such as the spatial awareness of objects and surroundings. To achieve a rich visual understanding of the site, the interface should fully contain and intuitively convey the associated contextual information. Excessive or unintuitive information not only makes it difficult for operators to exert their full potential but also increases their cognitive load. This study explores how different visual interface configurations affect operators' work performance and their cognitive load during the teleoperation task. The findings from the experimental studies are expected to help develop human-centered interfaces for teleoperation in the context of construction tasks and provide the cornerstone for not only an intuitive but fruitfully informative interface in a provided task setting.

INTRODUCTION

The introduction of unmanned construction technology has enabled the remote operation of construction machinery from a secure location, facilitating challenging work tasks (e.g., restoration of disaster sites) (Hiramatsu et al. 2002). As such, teleoperation is beneficial for the manipulation of excavation in an unfamiliar scene (Lee et al. 2022), as well as for problem-solving in such situations through trial-and-error techniques according to Rasmussen's SRK framework (Rasmussen 1983). Especially in extremely unstable areas, accidental operation of an excavator can result in tipping over due to changes in the center of gravity and the effects of inertial forces (Shigematsu et al. 2021). Therefore, careful operation is necessary based on spatial understanding provided by the interface (Lunghi et al. 2016). Successful teleoperation depends heavily on the design of the visual interface, which must provide operators with necessary information and feedback from the robot (Naceri et al. 2019). Operations are often challenged by environmental factors, and therefore require spatial understanding based on these factors, which can be enhanced through the information conveyed by the visualization interface (Lee et al. 2022; Wang and Dunston 2012).

Prior studies such as (Hedayati et al. 2018; Nielsen et al. 2007) have attempted to evaluate the effectiveness of visualization interfaces, which provide multiple 2D viewpoints or 3D wearable displays to teleoperators for enhancing the visual understanding of the site. Despite the benefits, there is still a gap in understanding their effects on teleoperator in the sloping terrain, which could cause a significantly degraded spatial awareness and high risks in manipulation. To address this gap, this study explores the effects of visualization interfaces on task performance and teleoperators' side in challenging conditions, such as a sloping terrain and physical restrictions on excavator manipulation due to closely located obstacles. This research investigates how displays can enhance spatial awareness despite misaligned viewpoints caused by challenging terrain, and how this may lead to data overload for teleoperators.

BACKGROUND

Enhancing spatial awareness though the visualization interface in teleoperation. Improving visualization and situational awareness is crucial in teleoperation (Naceri et al. 2019), where limitations in range of motion and visibility require clear conveyance of spatial awareness while minimizing cognitive load (Lee and Ham 2022). Prior studies in the virtual reality domain (Nielsen et al. 2007; Yanco et al. 2004) suggest minimizing the number of windows to avoid distractions. However, as excavator operators require a comprehensive understanding including occluded areas, recent studies (Motohashi et al. 2023; Tanimoto et al. 2017) pointed out the importance of incorporating multiple viewpoints into visualization interfaces for teleoperation in construction. This conflict shows the gap between studies in the virtual reality domain and the practical needs. Despite the teleoperation challenges posed by unstructured work environments, it has been rarely explored how the sloping terrain challenges teleoperators and how it could be alleviated from the benefits of visualization interfaces.

Factors causing jobsites as a challenging working environment in a sloping terrain. Construction sites are inherently challenging working environments that can be stressful and demanding for workers (Lee et al. 2022). Especially in the case of destroyed buildings (e.g., demolition), workers face even greater challenges as the topography itself can be treacherous and hazardous (Shigematsu et al. 2021). A slope terrain in such challenging sites poses a significant challenge for excavator workers due to its unstable and unpredictable nature, leading to an increased risk of accidents and injuries and reducing productivity (Seraji and Howard 2002). Additionally, obstacles can increase the task difficulty, forcing operators to adjust their strategies continuously (Li et al. 2016). As the task is deeply related to the physical environment surrounding the equipment (Carayon et al. 2006), the evaluation of the visualization interface should be conducted taking account of environmental factors.

Degraded spatial awareness in challenging work environments. Limited spatial awareness is a major challenge in teleoperation of construction tasks (Lee et al. 2022), and such negative effects are further amplified as the teleoperators are provided with the visual information through the interface (Verner et al. 2018). Moreover, the distorted view in uneven and sloping environments can reduce the intuitiveness of the teleoperator and lead to a degradation of spatial awareness. To overcome this, multiple viewpoints and 3D user-centered wearable display can be provided, which enable enhanced depth perception (Su et al. 2022). Nevertheless, according to the Yerkes-Dodson Law (Yerkes and Dodson 1908), providing more information to the operator is not always better. There is a need for investigating how added multiple viewpoints as multiple screens and wearable 3D display could enhance those teleoperators' spatial awareness in challenging environments.

METHODOLOGY

This study investigated the impact of visual interface configurations on the teleoperation performance and cognitive load of operators during excavation tasks in demolition sites (Fig.1). Objective measures, including work done, collisions, and completion time, were collected. Data analysis examined the effects of visual interfaces on task performance and workload perception using NASA-TLX and Presence Questionnaire (Witmer and Singer 1998) quantified the sense of presence in the virtual environment.

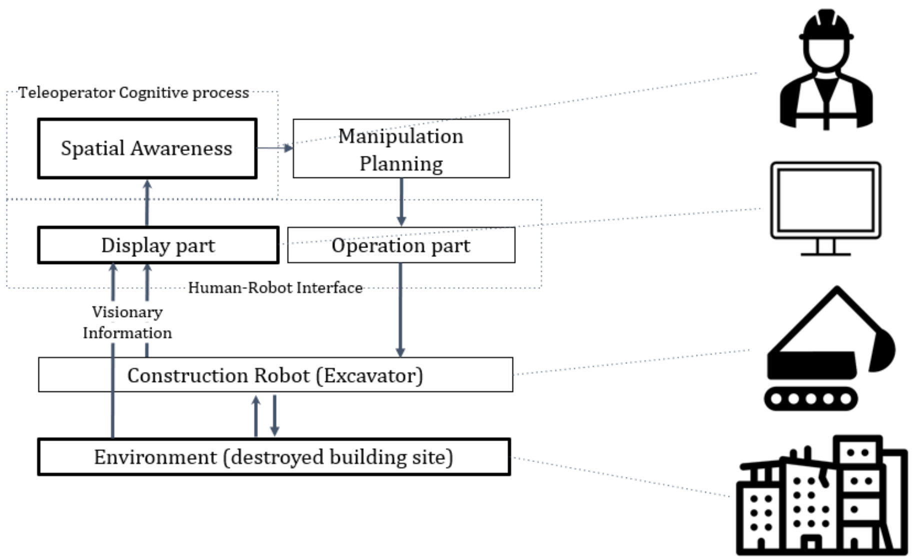

Figure 1. Visualization interface in human-robot interaction

Scenario Design and Virtual Environment Modeling. Moving the debris around destroyed buildings after disaster was selected for the experimental task. Two scenes, both of which included obstacles, debris, and the dumping area, were created. One was designated as the baseline, while the other was designated as the challenging scenario. Two key factors selected for modeling the challenging scenario are a hazardous terrain with a slope and closer obstacles requiring delicate manipulation, which imposes greater physical restrictions during work (Fig.2). Bricks were used as debris in the VR model, where participants were required to move debris in both hazardous and unhazardous terrains by completing four sequential tasks of picking up debris, avoiding obstacles, and dumping the debris at the destination. Each participant in the study experienced all the three visualization interfaces randomly, including single screen display, multiple screen display, and Head-Mounted (HMD) display. 1st person view is given to participants as a default screen. In a multiple screen display, 3rd person viewpoints are added including top-view (Kamezaki et al. 2016) and side-view (Ito et al. 2017) (Fig.3). In the model, these additional viewpoints were given with the adjustment of camera location to ensure that participants received the scene information without any occluded areas, providing relative distance information for debris, destination, and obstacles. In HMD display, viewpoints could be automatically adjusted as the participants rotated their head. The study built upon the Unity game engine to create an immersive VR environment, and to prevent motion sickness, each set of trials was limited to a maximum of 10 minutes.

Data Collection on Performance Measures and Eye Movement. During the experiment, eye movements of the participants were recorded. Both objective measures and subjective ratings were collected from participants. We employed the One-way repeated measures analysis of variance (ANOVA) with a Tuckey-adjusted post hoc paired t-test for the comparison. The association between variables is assumed as linear, and linear regression analyses were used to evaluate the relationship between variables and analyze the effects of factors on task performance.

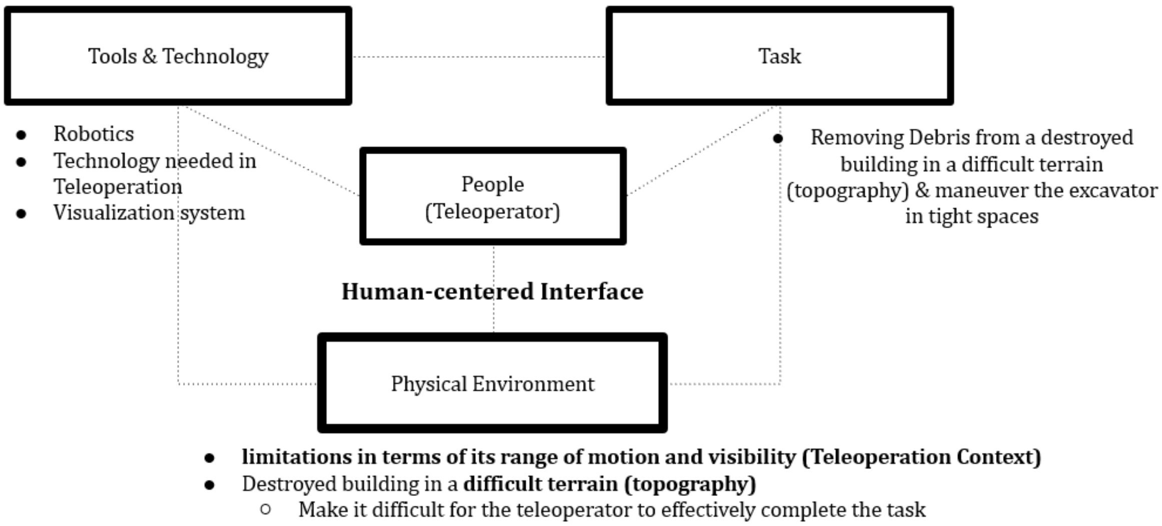

Figure 2. Elements for designing human-centered interface in teleoperation

Conditions and Procedures of the Pilot Experiment. A pilot experiment was conducted at Texas A&M University, involving 10 graduate students comprising eight males and two females (M_{age} = 24, SD_{age} = 1.95). The experiment consisted of two sections: baseline section and hazardous terrain section (Fig.4). The participants were randomly assigned to different scene conditions and display types. Before the experiment, participants watched a demonstration video and had a practice session to become familiar with the joystick.

Figure 3. Example images of distorted axis in 1st person view (left) compared to the 3rd person view from an uneven and sloping jobsite (right)

During the practice session, they were presented with only picking up and dumping debris. Before the experiments, 6 minutes limit is notified to the participants. The task was considered as completed after four sequential sets of picking up the debris, avoiding obstacles, dumping the debris are done or after 10 minutes. As a performance related to human-robot interaction in a challenging environment, completion time, completion amount, and the number of collisions are measured. In each section, after each trial ended in given display types, the participants were asked to rate their cognitive load with respect to mental demand, physical demand, temporal demand, self-rated performance, effort, and frustration level based on NASA-TLX (Hart and Staveland 1988). To examine the effect of scene conditions and display types to such objective measures, participants completed the Presence Questionnaire (Witmer and Singer 1998) at the last of each section. Presence Questionnaire consists of a series of questions that assess the strength of participants' sense of being present and engaged in the virtual environment while using the visualization interface.

Figure 4. Visualization display setting for normal scene (up) and sloping environment (bottom); single screen display (left), multiple screen display (middle), head-mounted display (right).

EXPERIMENT RESULTS

Impact of visualization interfaces on task performance. The results indicate the performance were significantly different in two scenes ($p < 0.01$). In a flat scene (normal), the mean amount of the work done (both picking up the debris and dumping the debris) does not show significant difference in both multiple screens and HMD compared to the single screen. However, in the sloping scene, the mean amount of work done in HMD were significantly higher than that in single screen display. Still, the mean amount of the work done does not show significant difference in multiple screens compared to single screen display even in the sloping scene (Fig.5).

Impact of visualization interfaces on collision occurrence and completion time. The results indicate that in the baseline scene, the number of collision occurrences doesn't show significant differences among the three types of visualization interfaces. However, in the hazardous terrain, the use of HMD display had a significant impact on collision occurrences (p =

0.02), while the use of multiple screens still did not show a significant effect compared to single screen use (Fig.6 [a]). Regarding the completion time, the impact of HMD display and multiple screens is found to be insignificant when compared to the single screen display. However, the median completion time is higher in the multiple screens case than in both the HMD display and single screen display cases (Fig.6 [b]).

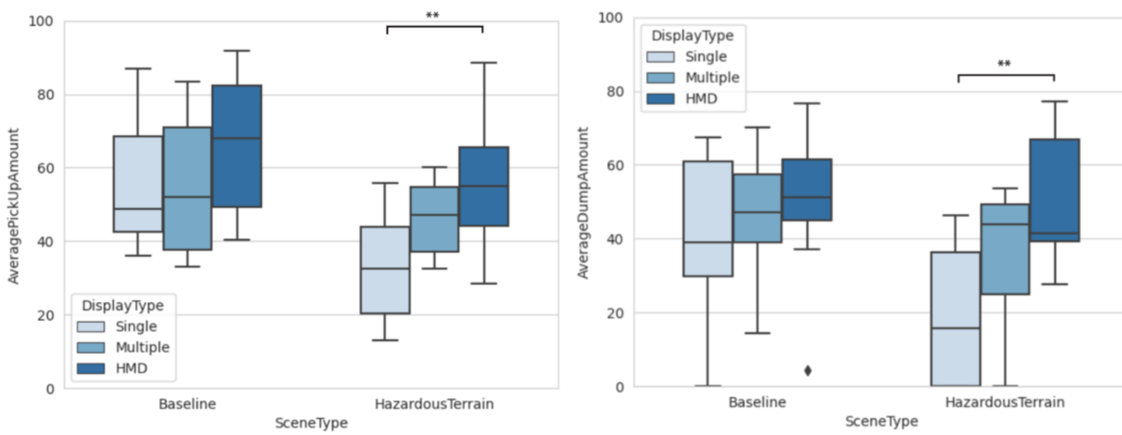

Figure 5. Results of amount of work done in picking up task (left) and dumping task (right).

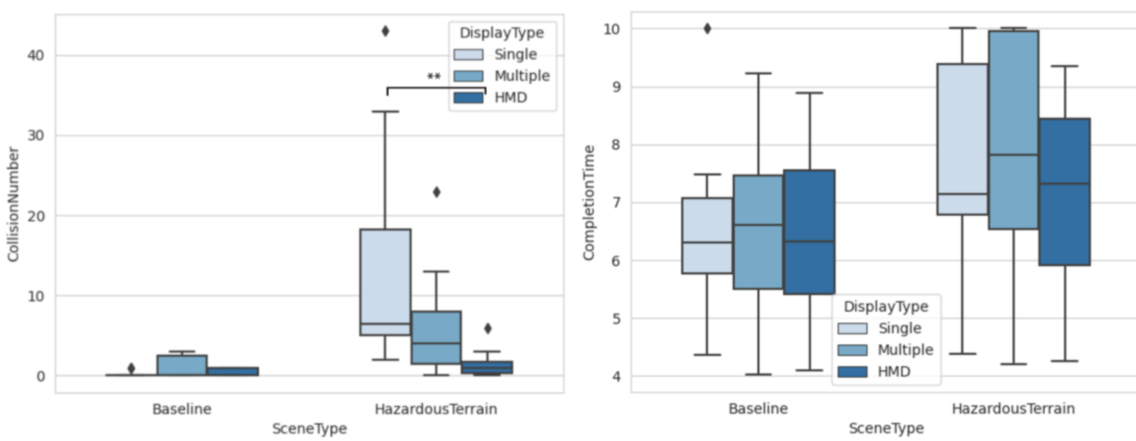

Figure 6. Results of the number of collision (a) and the completion time (b).

Impact of visualization interfaces on presence perception and task workload. The result showed that NASA-TLX score is significantly higher when participants were experiencing the hazardous terrain ($p < 0.01$). In both scenes, there's a significant difference in NASA-TLX score in single screen display and HMD display; $p = 0.02$ (in the baseline) and $p = 0.01$ (in the hazardous terrain), but there's no significant difference between both pairs of multiple display with HMD display and single screen display. In the hazardous terrain, NASA-TLX averaged scores were significantly different among three types of interfaces ($p = 0.01$). In both scenes, the mean differences of presence perception among the visualization interfaces are significant ($p < 0.01$). This indicates regardless of how challenging the scene is, the interface affects the teleoperator a lot in terms of presence perception. HMD provided significantly higher presence perception to teleoperators compared to multiple (p <0.01) and single screen (p <0.01). However, there was no significant difference between single screen and multiple screens (Fig.7 [a]).

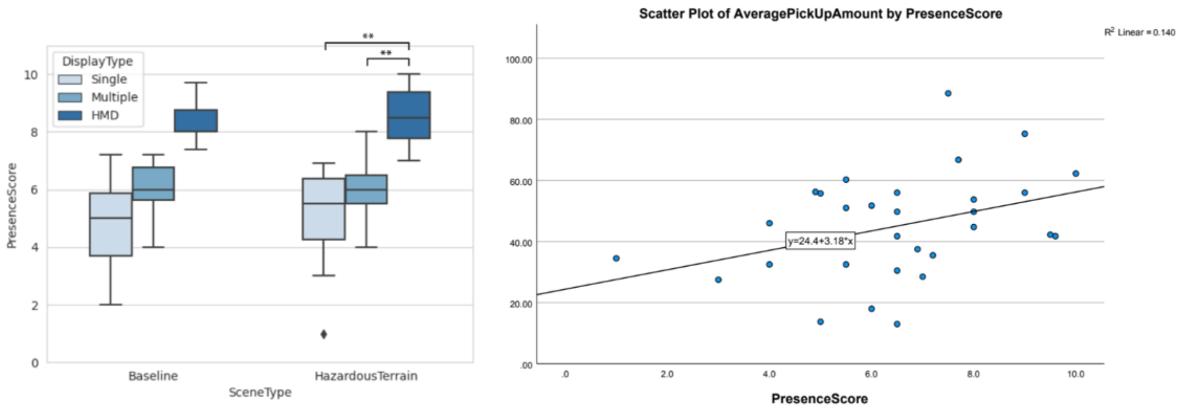

Figure 7. Averaged presence score (a) and linear relationship between presence score and the amount of work done (b).

The results of linear regression predicting the average amount of picking up bricks from participants' presence perception shows presence could not significantly affect the amount of work done in the baseline scene. However, in the hazardous terrain, participants experienced a significant positive relationship between presence perception and the average amount of picking up bricks, with R^2= 0.14, F (1,28) = 4.55, p =0.04 (Fig.7 [b]). This indicates when the participants show higher presence perception in virtual environments, they do their allocated work better done in virtual environments and this affects go significant when the participants experienced hazardous environments.

DISCUSSION

The experimental results demonstrated that the visualization interface significantly impacted work performance, particularly in hazardous and demanding site conditions. This finding suggests that the design of worker-centered visualization interfaces is increasingly necessary, particularly in severe physical conditions. The experiment employed two methods to enhance the information displayed on the visualization interface: adding viewpoints and adjusting the viewpoints to fit the user's head with a wearable visual display device. Interestingly, the addition of multiple screens (i.e., viewpoints) to the display interface did not lead to a significant improvement in performance compared to a single screen display in terms of the amount of work completed. This suggests that although additional useful information is provided, it does not necessarily lead to a significant improvement in performance. This lack of effect could be attributed to the fact that comparing information simultaneously did not have a significant impact on performance in the experimental setting; instead, it could distract the participants by providing too much information at once. Objective measures, including median completion time and collision rates, along with eye movements during collisions (Fig.8), support this interpretation. As depicted in the figure 8, the size of the circle corresponds to the fixation duration of the participants, and the line represents the saccades during that period. It was observed that during the collision occurrence, participants shifted their gaze to different viewpoints. Moreover, the median completion time for the multiple screens was higher compared to the single screen display, and there was no significant difference in the number of collisions between the two displays. These findings suggest that providing additional information to operators may not necessarily improve their task performance and abilities.

Figure 8. Eye Movements observed during the collision occurrence in the experiments.

The results could provide insights to consider when designing the visualization interface. Firstly, the interface should be designed to convey information intuitively and immersively, with a rigorous evaluation of its impact on teleoperators. Secondly, how the visualization interface could provide presence perception to workers should be considered in designing the visualization interface. Lastly, designing the interface to encourage high presence can improve workers' ability. While this study has provided valuable insights into the impact of different visualization interfaces on teleoperation performance, there is a limitation of the small sample size. Future studies should aim to address this limitation by recruiting larger participants, and this would allow for a more comprehensive evaluation of the impact of visualization interfaces on teleoperation performance and provide a more robust basis for the design of human-centered interfaces in construction tasks.

CONCLUSION AND FUTURE STUDY

To address the limited viewpoints for teleoperation in demolition sites with sloping terrain, the human-robot interface design should incorporate information from multiple viewpoints, particularly for the visualization interface, to support teleoperators. However, this approach could potentially result in data overload, necessitating an exploration of the effects of added and adjusted viewpoints with enhanced dimensions on teleoperators during the human-robot interface design phase. The impact of different visualization interfaces on teleoperation performance during excavation tasks was investigated in the experimental study, providing valuable insights into the development of intuitive and informative interfaces tailored to teleoperator needs. Furthermore, a VR environment was proposed to expose operators to hazardous terrain and enhance the crucial factor of presence, which affects work performance. The findings of the study suggest when designing the interface, it is critical to consider how to convey information intuitively and immersively to teleoperators and rigorously evaluate its impact on their performance.

ACKNOWLEDGEMENTS

This material is in part based upon work supported by the National Science Foundation (NSF) under Award #2026574. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

REFERENCES

Carayon, P., A. S. Hundt, B. Karsh, A. P. Gurses, C. J. Alvarado, M. Smith, and P. F. Brennan. 2006. "Work system design for patient safety: the SEIPS model." *Qual. Saf. Health Care*, 15 (Suppl 1): i50–i58.

Hart, S. G., and L. E. Staveland. 1988. "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research." *Adv. Psychol., Human Mental Workload*, P. A. Hancock and N. Meshkati, eds., 139–183. North-Holland.

Hedayati, H., M. Walker, and D. Szafir. 2018. "Improving Collocated Robot Teleoperation with Augmented Reality." *Proc. 2018 ACMIEEE Int. Conf. Hum.-Robot Interact., HRI '18*, 78–86. New York, NY, USA: Association for Computing Machinery.

Hiramatsu, Y., T. Aono, and M. Nishio. 2002. "Disaster restoration work for the eruption of Mt Usuzan using an unmanned construction system." *Adv. Robot.*, 16 (6): 505–508.

Ito, S., Y. Sakano, K. Fujino, and H. Ando. 2017. "Remote Controlled Construction Equipment by Using High-Resolution Stereoscopic 3d Images." *J. Jpn. Soc. Civ. Eng. Ser F3 Civ. Eng. Inform.*, 73 (1): 15–24.

Kamezaki, M., J. Yang, H. Iwata, and S. Sugano. 2016. "Visibility Enhancement using Autonomous Multicamera Controls with Situational Role Assignment for Teleoperated Work Machines." *J. Field Robot.*, 33 (6): 802–824.

Lee, J. S., and Y. Ham. 2022. "The effect of challenging work environment on human-robot interaction and cognitive load during teleoperation: a case study of teleoperated excavator in a virtual experiment." *Proc. 1st Future Constr. Workshop Int. Conf. Robot. Autom. ICRA 2022*. International Association for Automation and Robotics in Construction (IAARC).

Lee, J. S., Y. Ham, H. Park, and J. Kim. 2022. "Challenges, tasks, and opportunities in teleoperation of excavator toward human-in-the-loop construction automation." *Autom. Constr.*, 135: 104119.

Li, W., H. Gao, L. Ding, and M. Tavakoli. 2016. "Trilateral Predictor-Mediated Teleoperation of a Wheeled Mobile Robot With Slippage." *IEEE Robot. Autom. Lett.*, 1 (2): 738–745.

Lunghi, G., R. M. Prades, and M. D. Castro. 2016. "An Advanced, Adaptive and Multimodal Graphical User Interface for Human-robot Teleoperation in Radioactive Scenarios:" *Proc. 13th Int. Conf. Inform. Control Autom. Robot.*, 224–231. Lisbon, Portugal: SCITEPRESS - Science and and Technology Publications.

Motohashi, S., Z. Qiao, T. Moteki, and H. Iwata. 2023. "Analysis of the Effect of the Operator's Spatial Cognition on Their Planning Skills in Unmanned Construction." *2023 IEEESICE Int. Symp. Syst. Integr. SII*, 1–7.

Naceri, A., D. Mazzanti, J. Bimbo, D. Prattichizzo, D. G. Caldwell, L. S. Mattos, and N. Deshpande. 2019. "Towards a Virtual Reality Interface for Remote Robotic Teleoperation." *2019 19th Int. Conf. Adv. Robot. ICAR*, 284–289.

Nielsen, C. W., M. A. Goodrich, and R. W. Ricks. 2007. "Ecological Interfaces for Improving Mobile Robot Teleoperation." *IEEE Trans. Robot.*, 23 (5): 927–941.

Rasmussen, J. 1983. "Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models." *IEEE Trans. Syst. Man Cybern.*, SMC-13 (3): 257–266.

Seraji, H., and A. Howard. 2002. "Behavior-based robot navigation on challenging terrain: A fuzzy logic approach." *IEEE Trans. Robot. Autom.*, 18 (3): 308–321.
Shigematsu, K., T. Tsubouchi, and S. Sarata. 2021. "Tip-over prevention system based on motion prediction for teleoperated excavator." *Adv. Robot.*, 35 (23): 1438–1449. Taylor & Francis.
Su, Y., X. Chen, T. Zhou, C. Pretty, and G. Chase. 2022. "Mixed reality-integrated 3D/2D vision mapping for intuitive teleoperation of mobile manipulator." *Robot. Comput.-Integr. Manuf.*, 77: 102332.
Tanimoto, T., K. Shinohara, and H. Yoshinada. 2017. "Research on effective teleoperation of construction machinery fusing manual and automatic operation." *ROBOMECH J.*, 4 (1): 14.
Verner, I. M., S. Gamer, and A. Polishuk. 2018. "Development of spatial awareness and operation skills in a remote robot laboratory." *2018 IEEE Glob. Eng. Educ. Conf. EDUCON*, 389–393.
Wang, X., and P. S. Dunston. 2012. "Mixed Reality — Enhanced Operator Interface for Teleoperation Systems in Unstructured Environment." 1–8. American Society of Civil Engineers.
Witmer, B. G., and M. J. Singer. 1998. "Measuring Presence in Virtual Environments: A Presence Questionnaire." *Presence Teleoperators Virtual Environ.*, 7 (3): 225–240.
Yanco, H. A., J. L. Drury, and J. Scholtz. 2004. "Beyond Usability Evaluation: Analysis of Human-Robot Interaction at a Major Robotics Competition." *Human–Computer Interact.*, 19 (1–2): 117–149. Taylor & Francis.
Yerkes, R. M., and J. D. Dodson. 1908. "The relation of strength of stimulus to rapidity of habit-formation." *J. Comp. Neurol. Psychol.*, 18 (5): 459–482.