A Multi-view Interface for Tower Crane Teleoperation

Ming ZHANG, Xiaowei LUO
Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong, China; Email: xiaowluo@cityu.edu.hk

Abstract. Implementing teleoperation technology on construction sites remains challenging. The reconstruction of a traditional cockpit environment based on a limited view is one of the critical issues. Therefore, this study proposed a multi-view interface as a teleoperation terminal for tower crane teleoperation. A virtual multi-view interface for tower crane teleoperation was developed using the game engine. Four typical operation views were designed to reconstruct the visual environment of a traditional tower crane cockpit. Moreover, virtual annotations are designed and integrated into the virtual interface to enhance information display efficiency. Human-subject experiments were adopted to verify the effectiveness of the proposed teleoperation interface. User teleoperation performance and eye-tracking data were collected and analyzed during the experiment. The experimental results report the effectiveness of the proposed teleoperation system and provide empirical evidence for the future teleoperation interface design and development.

1. Introduction

The construction industry is facing various challenges, such as low productivity, labour shortages, and poor safety performance(Turner et al., 2021). Teleoperation (i.e., operating a machine from a distance) is an emerging technology that has many potential benefits, such as increasing workplace safety and engaging a more diverse workforce(Lee et al., 2022; Zhou et al., 2020; Zhu et al., 2021). Since human operators can stay away from dangerous workplaces, teleoperated systems are expected to be effective in reducing safety accidents on construction sites. For instance, remote-controlled semi-supervised construction robots can carry out operations (e.g., blasting (Melenbrink et al., 2020) and demolition (Jacinto-Villegas et al., 2017)) in hazardous environments, and perform collaborative construction tasks in complex construction site environments (Nagano et al., 2020; Okishiba et al., 2019).

The application of teleoperation still retains great value in the construction industry. Modular integrated construction (MiC), a new policy initiative to promote construction innovation in Hong Kong, offers new opportunities for teleoperation (Goh et al., 2019). The safe and efficient movement and assembly of the modules by the tower crane are critical to the success of MiC projects (Zhang & Pan, 2021). Despite the promising advantages, implementing teleoperation technology on construction sites remains challenging for two reasons. Firstly, it is challenging to reconstruct the cockpit visual environment based on the limited teleoperated interface. The teleoperation interface is presented by the camera's built-in cockpit monitors, which have the constraints of camera number and field of view (FoV). Recreating spherical field-of-view cockpit environments with limited monitor views is challenging for teleoperation interface design. Secondly, there is a lack of fundamental research on tower crane teleoperation interface design and deployment in construction sites. Teleoperator relies on the created sensations via the human-computer interface (HCI) interface to obtain field and operational information. It is still unclear how the teleoperation interface layout and information display affect the teleoperators' visual cognition and operation performance, which is critical for designing and optimizing teleoperation interfaces.

The multi-view HMI interface has been demonstrated as a potential solution to tackle the problem of operator information transmission (Kamezaki et al., 2021). Previous research suggested that virtual annotation(Hong et al., 2020) and multi-monitor systems (Kamezaki et al., 2021) may improve the performance of excavator operators. The multi-monitor HCI interface can improve the supply of visual resources to ensure sufficient information for teleoperation. However, the design of the multi-view of the tower crane teleoperation interface and the effectiveness of information transmission still need to be further explored. Moreover, the extensive and complex information feedback via multi-view can also cause the operator information overload problem and affect the operational performance (Edmunds & Morris, 2000; Maes, 1995). Therefore, it is critical to investigate the operator's information perception and teleoperation performance based on the multi-view teleoperation interface to validate the effectiveness of the proposed teleoperation system.

Motivated by this gap, this study proposed a multi-view teleoperation interface for future tower crane teleoperation. The contribution of this research is twofold. Firstly, a four-view teleoperation interface is proposed to reconstruct the visual environment of the tower crane cockpit. Secondly, a human-subject eye-tracking experiment was conducted to validate the effectiveness of the teleoperation system and indicate the visual distribution of the teleoperation interface. The results of this study provide empirical evidence and theoretical application for future teleoperation systems and interface research and development (R&D).

2. Method

The design and optimization of teleoperation interface is critical to realize the safe and stable teleoperation in construction site. This section proposed a four-view teleoperation interface for tower crane teleoperation for future on-site construction. Firstly, a virtual construction environment for tower crane teleoperation simulation is proposed. Then, the development pipeline of the four-view teleoperation interface is developed for tower crane teleoperation. Finally, the human-subject experiment design, apparatus, and data analysis method are introduced in this section.

2.1 Environment and system design

A typical construction project is developed in Unity software to simulate the hoisting tasks in real-world scenarios, as shown in Figure 1. A construction tower crane teleoperation system was developed in Unity software for human-subject experiments. The virtual scene includes the construction site under construction, the tower crane system and the basic construction houses, and a wealth of virtual assets are used to ensure the ecological validity of the virtual scene. A typical tower crane teleoperated hoisting task is designed through the HCI interface in the game engine to simulate the on-site hoisting process.

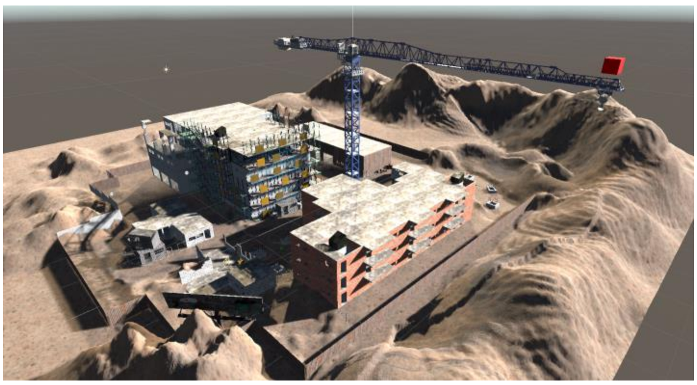

Figure 1: Schematic diagram of construction site and tower crane scene in Unity.

2.1 Multi-view teleoperation interface

The traditional tower crane cockpit has a near-circular FoV to assist the operator in obtaining a wide range of environmental information. The teleoperation system of the tower crane needs to reconstruct the cockpit environment based on the limited viewing angle of the camera and transmit it through the monitor, as shown in Figure 2. Therefore, previous research has explored teleoperation interface design from multiple views, including side and global view (Chi et al., 2012; Hong et al., 2020; Kamezaki et al., 2021; Yu et al., 2021). These studies focus on the perspective of the on-site environment, that is, a comprehensive display of the task structure. However, there is a lack of teleoperation interface design for the tower crane hoisting process, especially involving multiple information resources such as moving arms, dolly, and hook view. The reconstruction of the perspective of the cockpit and the comprehensive presentation of the information sources required in the teleoperation process are one of the core issues in the design of the teleoperation system for tower cranes.

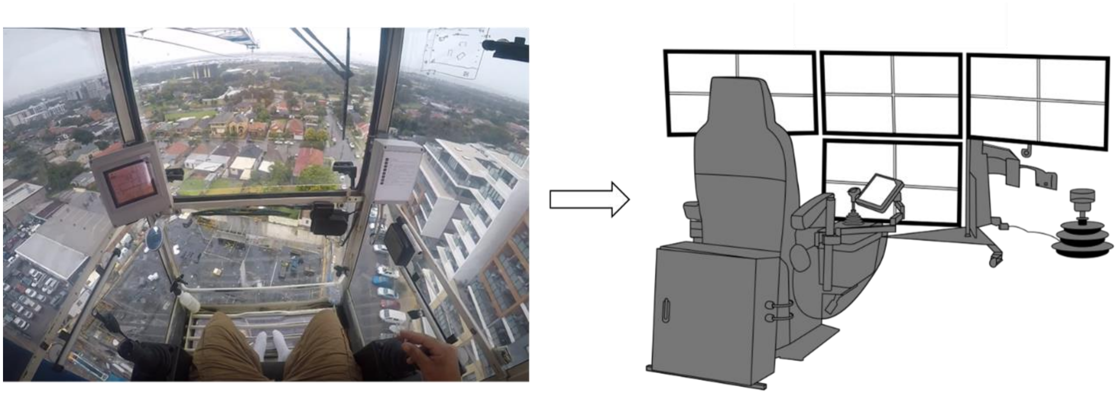

Figure 2: Schematic diagram of construction site and tower crane scene in Unity.

This research proposed a multi-view tower crane teleoperation interface was developed to assist the teleoperator in tower crane hoisting tasks. As shown in Figure 3, a four views monitor was designed and integrated into the teleoperated HCI interface, which integrates traditional perspective and enhanced perspective. Traditional perspective provides basic information on the state of the tower crane, including the main perspective and the global perspective. The former is the perspective of the traditional tower crane cockpit, and the latter show the status of the tower crane through the global positioning system (GPS) or wireless sensor network. The global perspective is a top-down perspective used to supplement the environmental information of the construction site. The main perspective is to simulate the first perspective of the tower crane driver inside the cockpit. The enhanced perspective is the additional cameras on the tower crane dolly and hook to provide closer and accurate information on the hook and the hanging objects. The dolly and hook are both cameras facing the ground and attached to the dolly and hook of the tower crane. A significant difference is that both cameras will move with the dolly movement, while only the camera on the hook will fall and rise with the hook during hoisting. In addition, an additional crosshair alignment system and auxiliary operating system are added to simplify the complexity of teleoperation.

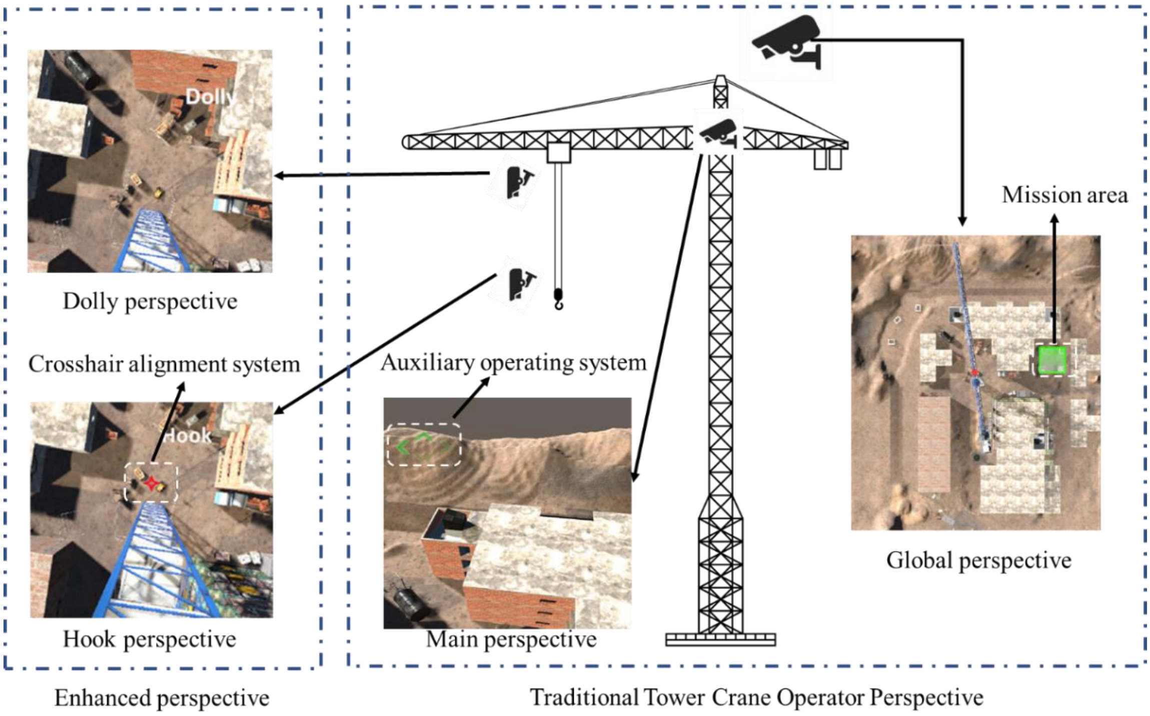

Figure 3: Muti-views tower crane teleoperation interface development framework diagram.

Based on the development framework, a four perspectives tower crane teleoperation HCI interfaces are designed for experimental testing. A desktop VR-based HCI teleoperation interfaces with four perspectives were developed for experimental testing, as shown in Figures 4.

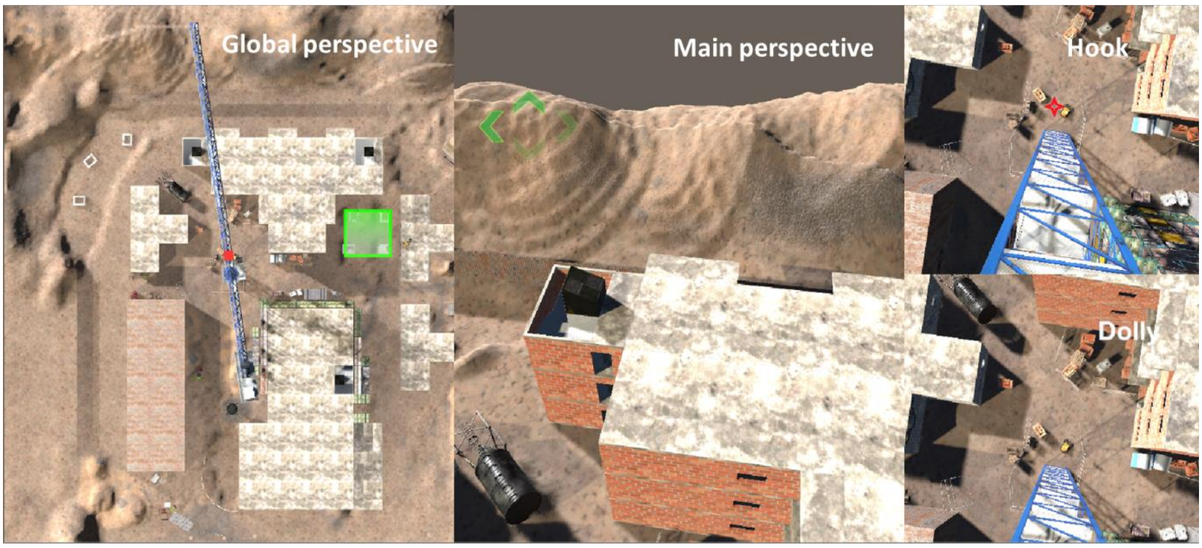

Figure 4: Four-perspectives desktop VR tower crane remote operation interface.

2.2 Experiment design

Three hoisting tasks were designed for the experiment. Participants need to lift items from the material area and need to transfer and place the items in the fluorescent task area as shown in Figure 4. In the three experimental tasks, the position of the material area is the same, but three different hoisting task areas are set in the scene, which appear sequentially with the completion of the previous hoisting task. The participants control the rotation of the tower crane arm, the forward and backward movement of the dolly and the up and down movement of the hook through the external game joystick (as shown in Figure 5), to realize the control of the virtual tower crane in the Unity scene. Before the experiment, participants need to participate in operation training for about half an hour to ensure that the experimental results will not be affected by individual operating skills.

The proposed HCI teleoperation interface was used in a human-subject experiment and the participants' visual cognition and task performance data were recorded. This article uses social media to post job advertisements for volunteer recruitment. Finally, 9 participants participated in the experiment and was included in analyzing and discussion in this study. Participants can terminate the experiment at any time if they feel unwell during the experiment, and each participant will be paid 60 yuan after the experiment. This experiment was approved by the Ethical Research Committee at the City University of Hong Kong.

Figure 5: Schematic diagram of the participants controlling the HCI system using joysticks.

2.3 Apparatus and data analysis

To further verify the effectiveness of the proposed four-view teleoperation interface, eye-tracking technology was used in this study to analyze the gaze behavior of the operator. Eye movement data collection was performed with Tobii Pro Fusion with sampling frequencies of up to 250 Hz. The data analysis in this research is divided into two categories, including experimental performance data and participant eye movement data analysis. The operating data of the participating personnel is recorded and analyzed locally through custom code integrated on the tower crane and operating handle. The eye movement data of the participants were analyzed by Tobii Pro Lab.

3. Results

3.1 Teleoperation performance

Finally, all operators completed three teleoperation lifting tasks, suggesting the effective of the proposed tower crane teleoperation system. To be specifically, the average time for 9 participants to complete the teleoperation task was 370.07s with standard error (S.D.) of 110.50. The operator who completed the task the fastest took 261.76s, while the operator who completed all tasks the slowest took 600.33s. In the experiments, 3 operators had a slight collision with the building during the task completion. The experimental results show that the tower crane teleoperation system proposed in this study can meet the requirements of tower crane teleoperation tasks.

3.2 Effectiveness of each supportive view

The visual search heat map of participant 3 after completing the hoisting operation of the tower crane is shown in Figure 6. According to the heat map, the participants still pay more attention to the main perspective. The main perspective is used to monitor the relative position of the suspended object to the building environment, providing location relationship information during the lifting movement phase. In addition, the participants also allocated many gaze points in the global perspective for observing the task position and the current space state of the tower crane. Some fixation points were also observed from the hook and dolly perspectives, indicating that the participants gazed at these two perspectives during the hoisting process. Overall, the four perspectives of the visual-enhanced teleoperation interface provide feedback information to the participants to complete the teleoperation task.

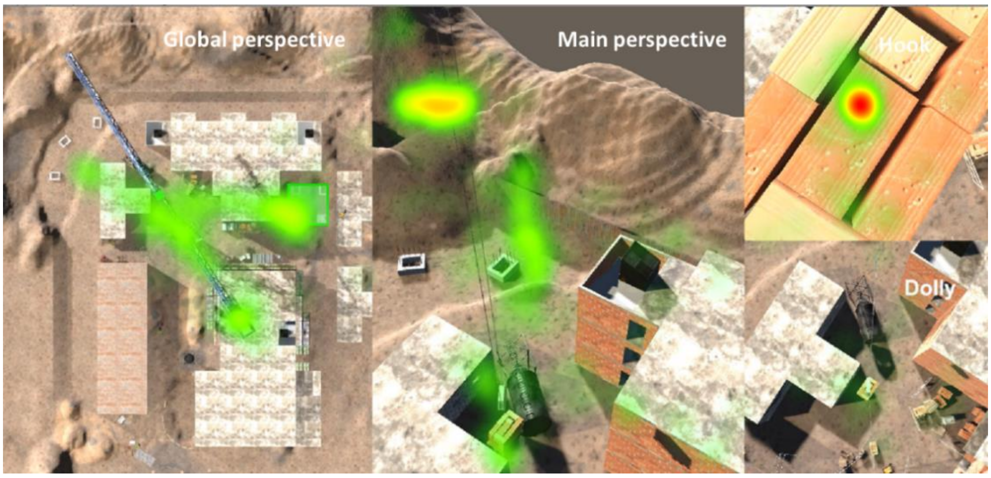

Figure 6: Visual heat map of the teleoperation interface.

To further analyze the visual distribution of the participants on the teleoperation interface, this study conducts a partition analysis on the four perspectives on the teleoperation interface. As shown in Figure 7, the four perspectives of the teleoperation interface are partitioned to count the fixation duration of each partition separately. The statistical analysis of fixation persistence results based on partitions was performed by Tobii Pro Lab.

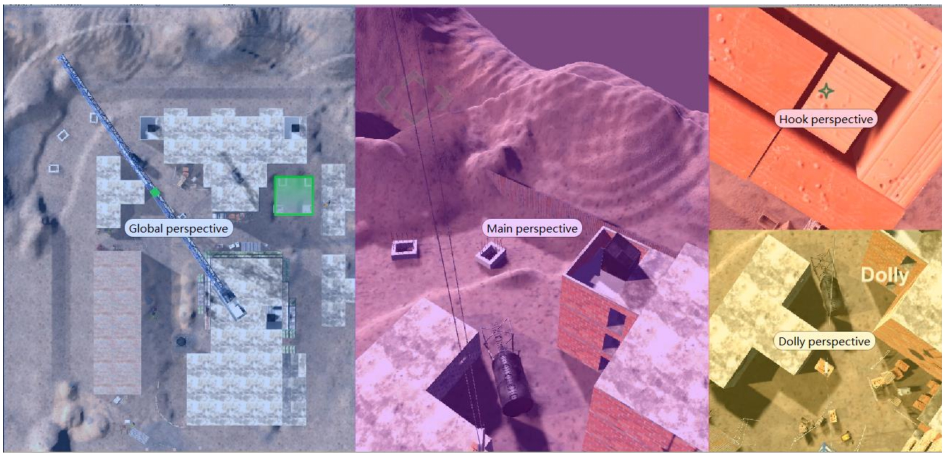

Figure 7: A screenshot of the gaze behaviour zonal analysis of the teleoperation interface.

Figure 8 shows the statistical results of the gaze time of the nine participants on the four partitions in the teleoperation experiment. Most participants allocated more fixation time in traditional perspectives (i.e., global perspective and main perspective), indicating that traditional perspectives can provide the useful feedback for the teleoperation process. It should be noted that all participants have invested certain visual resources in the hook perspective. In experimental observations, the hook perspective can provide effective feedback for specific tasks such as lifting and lowering loads. The dolly perspective does not receive many visual resources, because the participants usually only need to move the dolly of the tower crane to observe its position, which accounts for a small proportion in the teleoperation process.

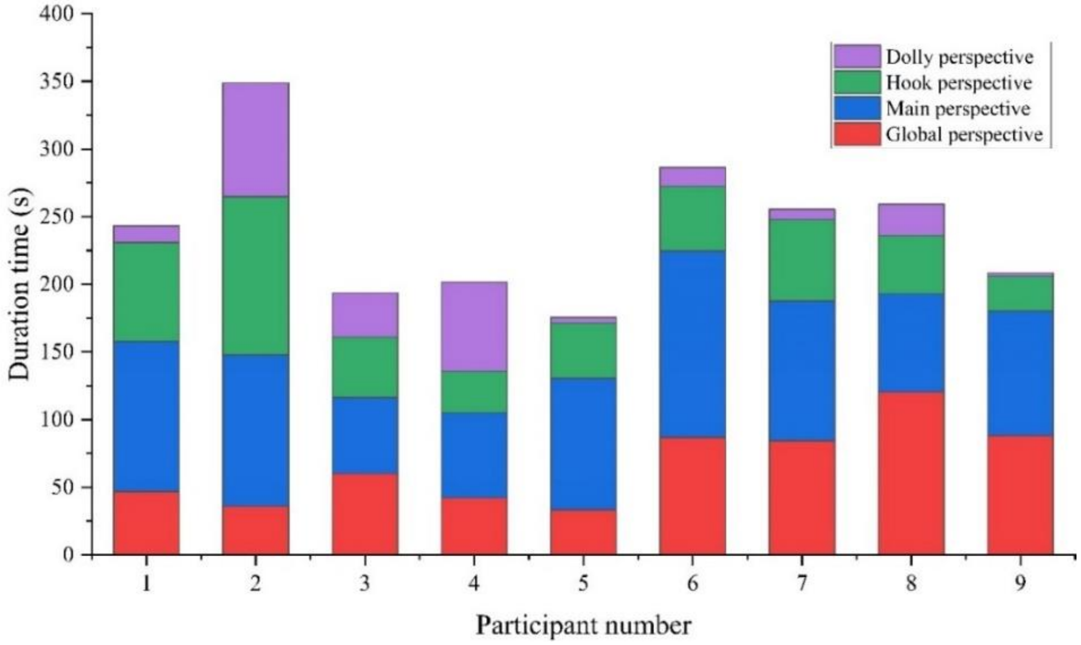

Figure 8: The resource allocation results of the participants in the four perspectives of the teleoperation interface.

Finally, a statistical analysis was carried out on the gaze duration of the four viewing angles of the nine participants on the tower crane teleoperation interface, as shown in Figure 9. The results show that the main perspective is the most gazed perspective, accounting for 38.9% of the total fixation time. In addition, the global perspective and the hook perspective, as functional perspectives, get a lot of fixation time, accounting for 27.5% and 22.2% respectively. The dolly perspective is only applicable to the tower crane dolly movement process, and the gaze time accounts for about 11.3%. It should be noted that the sum of the decimals is not 1 is caused by the decimal rounding problem.

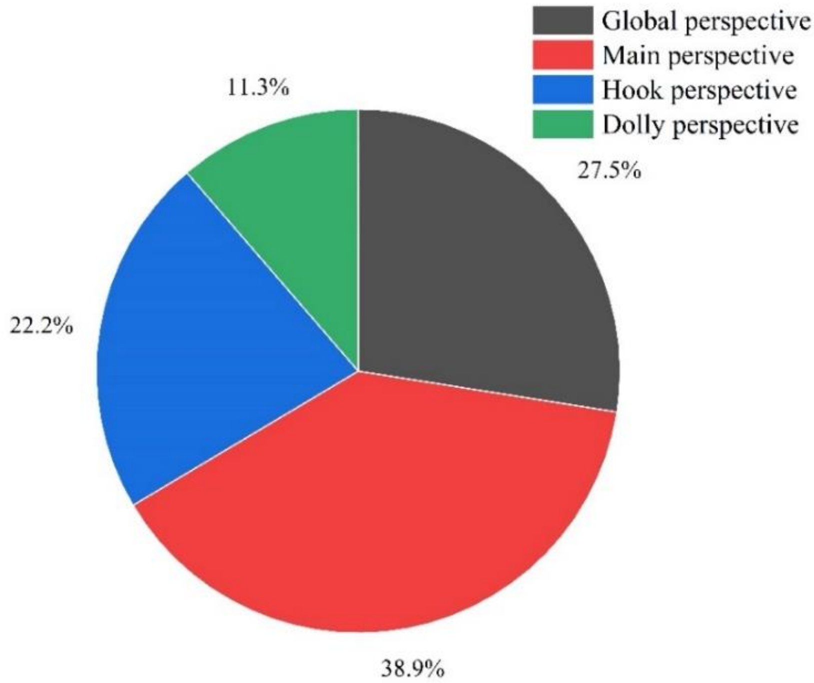

Figure 9: Statistical chart of participants' fixation duration in the four perspectives of the teleoperation interface.

4. Discussion

This research explores a novel visual-enhanced teleoperation interface and provides the user testing results. Overall, the results based on eye-tracking behaviour prove the effectiveness of the proposed vision-enhanced teleoperation interface design, which users can obtain teleoperation site information through multiple monitor perspectives. Thus, this study discusses focus on the interface design and visual-enhanced interface in this section.

4.1 Teleoperation interface design

In teleoperation, the teleoperator relies on the created sensations via the HCI interface to obtain construction site environmental information. The operator in the traditional tower crane cockpit has a broad perspective, which can usually be understood as a spherical viewing range but is limited by the operator's vision and sight. Therefore, effectively conveying the environmental and operation information of the construction site through the limited monitor interface of the teleoperation interface still needs to be solved. This research proposes a visual-enhanced HCI operation interface, which includes four perspective systems: main perspective, global perspective, hook perspective, and dolly perspective. Compared with the traditional tower crane operator, the proposed perspectives of the HCI interface (e.g., hook and dolly viewing angle) can provide more accurate operation information without being limited by the operator's line of sight distance. The experiment results show that all participants can complete three hoisting tasks through the teleoperation interface, which validates the effectiveness of the proposed teleoperation system.

4.2 Operator Visual Distribution

It is challenging to reconstruct the traditional tower crane cockpit perspective based on the limited monitor perspectives. Although increasing the number of monitors can increase the supply of teleoperation environmental information, it may still cause problems such as information overload(Edmunds & Morris, 2000). Therefore, optimizing HCI information supply and interface design and exploring the operator's cognitive behavior signify great importance to teleoperation. This study developed a visual-enhanced HCI interface design that aims to enhance the provision of environmental and operational information by providing additional perspectives beyond the primary perspective. According to the results of the eye-tracking analysis, about 60.11% of the visual resources of the participants were allocated to perspectives other than the main perspective. The enhanced perspectives (hook and dolly perspectives) accounted for 22.2% and 11.3% of the total fixation time, respectively, which verified the auxiliary effect and effectiveness of the enhanced viewing angles on the teleoperation process. It should be noted that in the traditional tower crane operation process, the operator needs to be equipped with a ground commander to assist, because it is usually difficult for the operator to see the ground and the hanging object at a far distance. The traditional transportation mode requires operators and ground managers to communicate through walkie-talkies, which increases the risk of human error. The proposed vision-enhanced tower crane teleoperation interface provides monitoring of the tower crane dolly and hook through the enhanced perspective, without additional manual assistance to complete the teleoperation hoisting task. Therefore, the proposed teleoperation system may reduce the influence of human factors in future teleoperation, which will be more reliable and safer than traditional tower crane operations.

5. Conclusion

This research proposed a novel visual-enhanced tower crane teleoperation interface for future teleoperation applications in construction sites. Four perspectives are designed and integrated into the HCI interface to provide the system's teleoperation environment and operational information, including main, global, hook, and dolly perspectives. The effectiveness of the proposed teleoperation interface is tested based on human-subject experiments, and the participants' experimental data and eye movement behaviour data are analyzed. The results show that the teleoperation system proposed in this study can effectively support the participants in completing the teleoperation lifting task. In addition, the visual-enhanced HCI interface proposed in this research can effectively provide feedback on teleoperation information to assist the teleoperator in information perception and decision-making in the teleoperation process. Several limitations are presented in this section. This study only selected a small sample (9 volunteers) for pre-experimental testing, and the experimental results are not universal. In addition, according to the pre-test results, the teleoperation interface and system should be optimized and designed for further research and application.

6. Acknowledgments

The National Science Foundation of China Grant # 51778553 and City University of Hong Kong Strategic Research Grant # 6000781 jointly supported this work. The conclusions herein are those of the authors and do not necessarily reflect the views of the sponsoring agencies.

7. References

Chi, H.-L., Chen, Y.-C., Kang, S.-C., & Hsieh, S.-H. (2012). Development of user interface for tele-operated cranes. Advanced Engineering Informatics, 26(3), 641-652.

Edmunds, A., & Morris, A. (2000). The problem of information overload in business organisations: a review of the literature. International Journal of Information Management, 20(1), 17-28.

Goh, J., Hu, S., & Fang, Y. (2019). Human-in-the-Loop Simulation for Crane Lift Planning in Modular Construction On-Site Assembly. In Computing in Civil Engineering 2019 (pp. 71-78).

Hong, Z., Zhang, Q., Su, X., & Zhang, H. (2020). Effect of virtual annotation on performance of construction equipment teleoperation under adverse visual conditions. Automation in Construction, 118, 103296.

Jacinto-Villegas, J. M., Satler, M., Filippeschi, A., Bergamasco, M., Ragaglia, M., Argiolas, A., Niccolini, M., & Avizzano, C. A. (2017). A Novel Wearable Haptic Controller for Teleoperating Robotic Platforms. IEEE Robotics and Automation Letters, 2(4), 2072-2079.

Kamezaki, M., Yang, J., Sato, R., Iwata, H., & Sugano, S. (2021). A situational understanding enhancer based on augmented visual prompts for teleoperation using a multi-monitor system. Automation in Construction, 131, 103893.

Lee, J. S., Ham, Y., Park, H., & Kim, J. (2022). Challenges, tasks, and opportunities in teleoperation of excavator toward human-in-the-loop construction automation. Automation in Construction, 135, 104119.

Maes, P. (1995). Agents that Reduce work and information Overload. In R. M. Baecker, J. Grudin, W. A. S. Buxton, & S. Greenberg (Eds.), Readings in Human–Computer Interaction (pp. 811-821). Morgan Kaufmann.

Melenbrink, N., Werfel, J., & Menges, A. (2020). On-site autonomous construction robots: Towards unsupervised building. Automation in Construction, 119, 103312.

Nagano, H., Takenouchi, H., Cao, N., Konyo, M., & Tadokoro, S. (2020). Tactile feedback system of high-frequency vibration signals for supporting delicate teleoperation of construction robots. Advanced Robotics, 34(11), 730-743.

Okishiba, S., Fukui, R., Takagi, M., Azumi, H., Warisawa, S. i., Togashi, R., Kitaoka, H., & Ooi, T. (2019). Tablet interface for direct vision teleoperation of an excavator for urban construction work. Automation in Construction, 102, 17-26.

Turner, C. J., Oyekan, J., Stergioulas, L., & Griffin, D. (2021). Utilizing Industry 4.0 on the Construction Site: Challenges and Opportunities. IEEE Transactions on Industrial Informatics, 17(2), 746-756.

Yu, Z., Luo, J., Zhang, H., Onchi, E., & Lee, S.-H. (2021). Approaches for Motion Control Interface and Tele-Operated Overhead Crane Handling Tasks. Processes, 9(12), 2148.

Zhang, Z., & Pan, W. (2021). Virtual reality supported interactive tower crane layout planning for high-rise modular integrated construction. Automation in Construction, 130, 103854.

Zhou, T., Zhu, Q., & Du, J. (2020). Intuitive robot teleoperation for civil engineering operations with virtual reality and deep learning scene reconstruction. Advanced Engineering Informatics, 46, 101170.

Zhu, Q., Du, J., Shi, Y., & Wei, P. (2021). Neurobehavioral assessment of force feedback simulation in industrial robotic teleoperation. Automation in Construction, 126, 103674.