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-

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.

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.

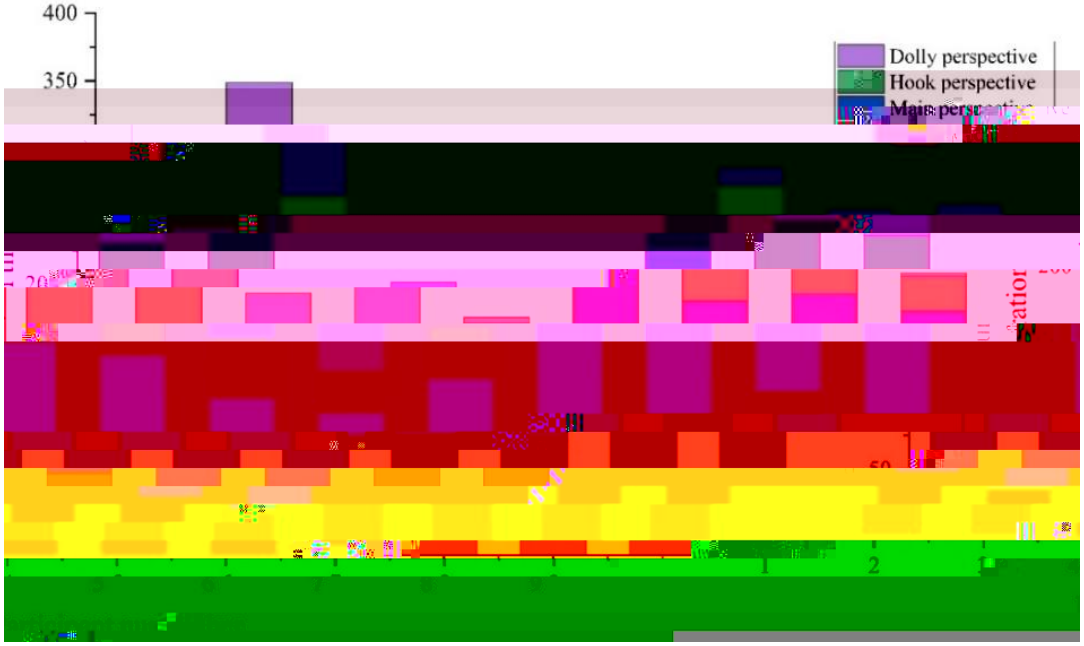

Figure 8: The resource allocation results of the participants in the four perspectives of the teleoperation interface.

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.

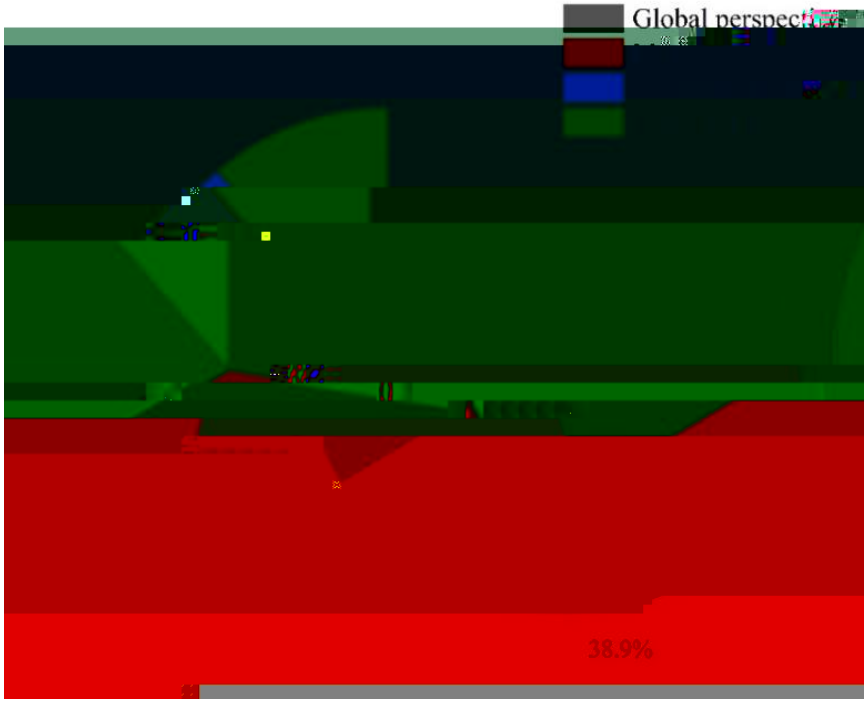

Figure 9: Statistical chart of participants' fixation duration in the four perspectives of the teleoperation interface.

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 BTF1 12 Tfd(l)382 Tfd(l)382 Tfd(l)3ingisito tyc6wl -(e)4(v)-9