Towards a Collaborative Future in Construction Robotics: A Human-centered Study in a Multi-user Immersive Operation and Communication System for Excavation

Di Liu, Youngjib Ham, Jeonghee Kim, Hangue Park

Abstract — **When operating a construction robot, i.e., an excavator, the excavator operator's unsafe behavior directly affects the underground utility damage occurrence during excavation process. Operator's behavior is greatly affected by the environment and further the communication with other coworkers, i.e., spotter. In this paper, we propose a multi-user immersive operation and communication system for excavation. Further, we investigate how the different types of environments and operator-spotter communication channels affect operator's attention demand and performance during excavation.**

I. Introduction

The damage to utility lines in excavation is one of the most significant crises for contractors and creates great economic and societal loss especially in dense urban areas having the congested underground utility lines [1, 2, 3-5]. Unfortunately, current practices and damage prevention systems are insufficient to prevent these accidents [2,4-5,8-10, 11], and excavator operators heavily rely on their own judgement to avoid utility line damages [6,7]. In fact, in the safety guidelines recommended by CGA Best Practice and some state code [11, 12], once the excavation starts, working with a spotter is a key step to prevent damages and enhance the safety of excavation, and excavation tasks are performed as a teamwork most of the time. When an operator controls an excavator and interacts with a spotter at the same time, the operator often experiences cognitive overload, and unsafe behaviors and accidents are more likely to occur. In this regard, studying the operator-spotter interaction is crucial to ensure the safe human-excavator collaboration in construction tasks, especially with utility lines buried in a challenging environment (e.g., urban jobsite).

To better train operators for preventing the accidents in the real jobsite, excavation simulators are commonly used for task practicing and studying the human factors. The typical excavation simulator in the current market is composed by joysticks and pedals, monitor-based display, and available for a single user to practice the basic excavator operation. Despite the advanced development of technologies, majority of excavation simulator provide an acceptable but less immersive simulation environment. Furthermore, collaborative excavation which has multiple users involved has not been included in the simulator design.

In this study, we develop and evaluate an immersive multi-user simulation system for excavator operation and investigate the operator-spotter communication under different environments. Specifically, we focus on how the types of environments and the operator-spotter communication formats affect operator's attention demand and performance during excavation. The main contributions of the proposed multi-user immersive system are:

- to allow more than one construction workers, i.e., operator and spotter, collaborate with construction robot(s), i.e., excavator, in a high level of immersion
- to assess the human performance in a collaborative human-robot-interaction workplace as needed.

II. System Design

A multi-user immersive operation and communication system is composed of two parts, a virtual human excavator interaction platform (Fig. 1), and an immersive multi-user communication system (Fig. 2), which are supported by a set of hardware and software.

A. Virtual Human-Robot Interaction Platform

- Hardware system

The primary goal of the hardware design is to serve as the physical excavator simulator. As the main part of this VR-based platform, HTC Vive Pro Eye is the VR headset being functional as the user display with a resolution of 1440 x 1600 per eye. The embedded eye tracking feature enables eye data collection during the experiment. The realistic excavator joysticks that had a USB connection that could be plugged directly into the PC were selected for this simulator. Due to compatibility issues with the external controller and the Unity software, a software called JoyToKey is selected to emulate the joystick movements as keystrokes. By doing so, we were able to tie the joystick movements to key presses and set these key press inputs as the inputs in Unity model. One of the most common control patterns, ISO control pattern, is utilized in this system. The realistic excavator pedals interfaced with a fabricated printed circuit board (PCB) and Arduino UNO. The analog signals provided by the pedal were sent to two analog pins located on the Arduino UNO. These signals were able to be read by Unity by using a Unity Asset called Uduino. Uduino

*Research supported by the National Science Foundation.

Di Liu is the Ph.D. student, Department of Construction Science, Texas A&M University, College Station, TX 77840 USA (e-mail: catsquito@ tamu.edu).

Youngjib Ham, Associate Professor, Department of Construction Science, Texas A&M University, College Station, TX 77840 USA (corresponding author to provide phone: 979-458-0184; e-mail: yham@tamu.edu).

Jeonghee Kim, Assistant Professor, Department of Engineering Technology & Industrial Distribution and Electrical & Computer Engineering, Texas A&M University, College Station, TX 77840 USA (e-mail: jeonghee.kim@tamu.edu).

Hangue Park, Assistant Professor, Department of Electrical & Computer Engineering, Texas A&M University, College Station, TX 77840 USA (e-mail: hpark@ece.tamu.edu).

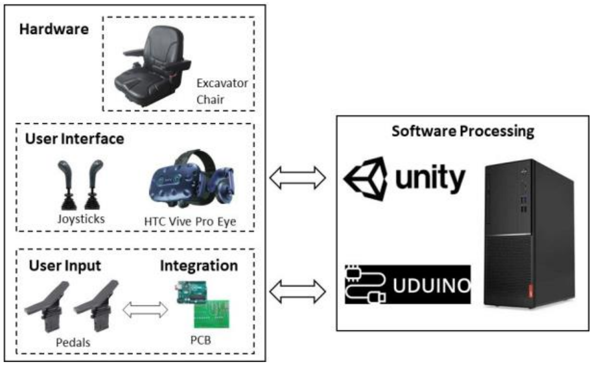

Figure 1. Virtual Human Robot Interaction Platform.

is an asset that helps simplify the communication between the Arduino UNO and Unity software. Uduino can read and write analog and digital signals though C# scripts written in Unity, rather than through the Arduino IDE. Besides the above components, a set of mechanical assembly made by plywood enclosure is built to place the excavator chair, joysticks, and pedals.

- Software system

The excavator simulation is created completely on the Unity3D Game Engine. In addition, to properly stream the simulation to the VR headset with Unity and other external controller inputs, SteamVR is used. A Lenovo ThinkStationP620 is the computer environment of the software system and connected to the VR headset, joysticks, and PCB.

B. *Immersive Multi-User Communication Environment*

The immersive multi-user communication environment allows a real-person operator (a participant), a real-person spotter, and the excavator simulator to work together. The real-person operator wearing an HTC Vive Pro Eye headset which has an eye tracker and a headphone embedded performs a set of excavation tasks. The real-person spotter wore a Logitech H390 microphone to communicate with the operator and held two Vive controllers to track hand gestures in real time. When the operator was performing the task, the spotter kept monitoring the virtual excavation process and virtual environment from multiple view directions displayed by two 24 inches monitors. The VR headset, microphone, and excavator simulator are connected to a Lenovo ThinkStationP620.

In terms of excavation task, Unity3D is the main platform for modeling and visualization. A roadwork scenario was simulated as the baseline environment, and a downtown scenario is simulated as the challenging environment. Both scenarios include visual and auditory urban elements. An excavation job site area were modeled and a virtual excavator was placed in the center. Underground utility lines were placed in front of the excavator and hidden from the operator's view direction which provides a close-to-real excavation experience. A Collision Detection script recorded the excavator-utility collision automatically. Moreover, a virtual operator and a virtual spotter were simulated in the same scenarios. The virtual operator was represented by a virtual camera located inside of the virtual excavator cabinet. The real-person operator could constantly see the scenario from a First-Person Point-of-View through VR headset display which showed the view of virtual camera. A virtual avatar representing the spotter stand in front of the virtual camera and can be seen by the operator. To mimic the real-life work, the avatar was model with a construction worker's appearance. The avatar's arm gestures were controlled by the real-person spotter's arm movements via two controllers. Two audio converting scripts were used to enable the real-time verbal communication between operator and spotter. The buried utility lines were set to be visible to the real-person spotter so that the spotter could guide the operator to avoid hitting the utility lines.

Multiple experiments were conducted. Excavation tasks with multiple view perspectives and different sound were recorded by OBS screen recording software. Operator's gaze information was collected through Vive eye tracking SDK and iMotions software. A log file recorded collision numbers and other system information.

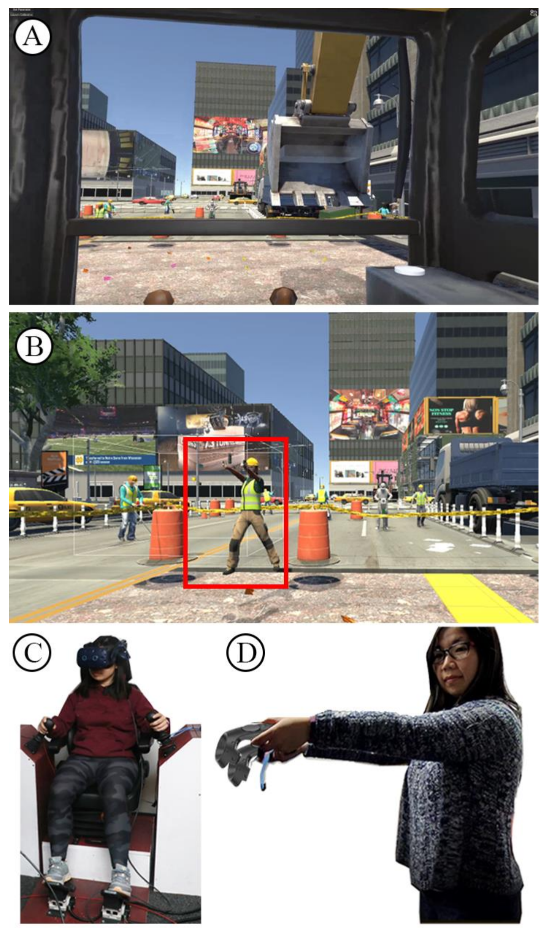

Figure 2. Immersive Multi-User Communication Environment: (A) Operator's view (B) Spotter's avatar (C) Real-person operator (D) Real-person Spotter

III. EXPERIMENT

We conducted a small group of user study by using the designed system prototype to examine three types of operator-spotter communication formats (hand signals, verbal signals, a mixture of hand-verbal signals), and two types of environmental conditions (baseline, challenging environment) on the tested subjects. A total of six participants were divided into three groups based on hand signals, verbal signals, and a mixture of hand-verbal signals. Each subject repeatedly performed four trials of excavation tasks (three loads of soil per trial) in both baseline and challenging environments. To ensure the job completion quality, the tasks were monitored by the spotter. Dependent variables are defined as task-oriented performance variables including collision numbers (COLLI) and missed-signal rates (SGMR), as well as several cognitive responses variables including attention demand.

A. *Experiment Procedures*

On the experiment day, a total of six experimental sessions were conducted by each participant upon the completion (Fig. 3). In Session 1, participants were acknowledged by the consent form and given an introduction of the research. Participants completed a background questionnaire. Session 2 provided basic knowledge of operating an excavator. In Session 3, a 20-min practice in VR is provided. In Sessions 4, 5, 7, 8, participants were asked to perform tasks of excavating three loads of soil by following spotter's signals, as well as avoiding the collision with buried utility lines. In Session 6, a 5-min break is provided after Session 5. To counterbalance the learning effect, Sessions 4 and 8 are conducted in the baseline environment and Sessions 5 and 7 are in the challenging environment. In Session 9, participants were asked to complete post-experiment questionnaires.

B. *Human Factor Measurements*

Four types of measurements including performance, attention demand, mental workload, and awareness are assessed in the experiments. The number of collisions is recorded in a log file. Missed signals are counted from screen recording after the experiments. Participants' Signal Missing Rate (SGMR) per trial is calculated as below:

SGMR = the number of signals / a total number of signals from spotter.

Operator's attention demand is assessed by dynamic attention intensity and attention spatial density generated from eye tracking data.

Besides the instrument-based measurements, a set of subjective evaluation are collected, including NASA Task Load Index (NASA-TLX) to assess mental workload, 10-D SART Scale to evaluate the situation awareness, a 5-item environmental distraction questionnaire to evaluate the perceived distraction by different environmental elements, and a W&S presence questionnaire to access the sense of presence of the VR environment.

IV. RESULTS AND DISCUSSION

A. *Performance Results (SGMR, COLLI)*

Fig. 4 shows the results of signal-missing-rate (SGMR) and collision number (COLLI) of all subjects of two environments (B – baseline, C – challenging environment). For each subject, SGMR in challenging environment shows a higher value than SGMR in baseline environment. The average SGMR in challenging environment is 14.37%, which is higher than the average SGMR in baseline environment (3.77%). The results of SGMR indicate that all participants missed more signals from spotter when they performed the task in the challenging environment than they did in the baseline environment. Among different signal formats, a higher SGMR is occurred in the verbal signal group. In terms of collision number, Fig. 4 shows that, among all participants, COLLI in the challenging environment, with an average of 2.5, is higher than in the baseline environment in which the average COLLIS is 1. This result indicates that the operator tends to make more collisions in the challenging environment than in the baseline environment. Regarding the completion time of tasks, it was observed that participants completed the tasks faster in the baseline environment with verbal communication with spotter.

B. *Attention Intensity and Attention Spatial Density*

Fig. 5 shows the distribution of attention intensity. Visualizing attention intensity (heatmap) is commonly based on static 2D display on which the user has a static field of view. In an immersive VR environment, as the operator changes the view direction constantly and intuitively, i.e., look around by moving head positions, rotate the virtual excavator cabinet, it is necessary to categorize the dynamic view fields for analyzing attention demands. Therefore, operator's views are categorized into three classes: Trenching (TR), Rotating (RO), Dumping (DU). When performing task of each load, a static

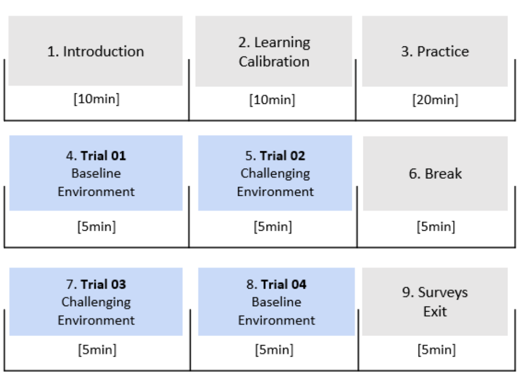

Figure 3. Experiment Procedure

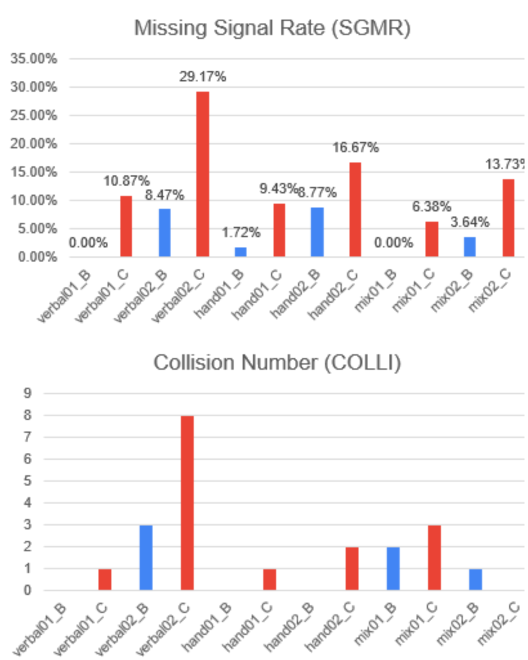

Figure 4. Performance Results

Figure 5. Dynamic Attention Intensity : (a) Attention heatmap in trial 1 (b) Attention heatmap in trial 2

scene with a short time interval for each view class is abstracted, the attention intensity of each static frame per load is rendered into a heatmap (Fig. 5). Fig.5 shows that in the challenging environment, a wider distributed attention is achieved than in the baseline environment. Fig. 6 shows that there is a higher fixation spatial density in challenging environments than in the baseline environment. These results indicate that in a challenging environment, the operator is more distracted than in the baseline environment, which strengthened the likelihood of accidents.

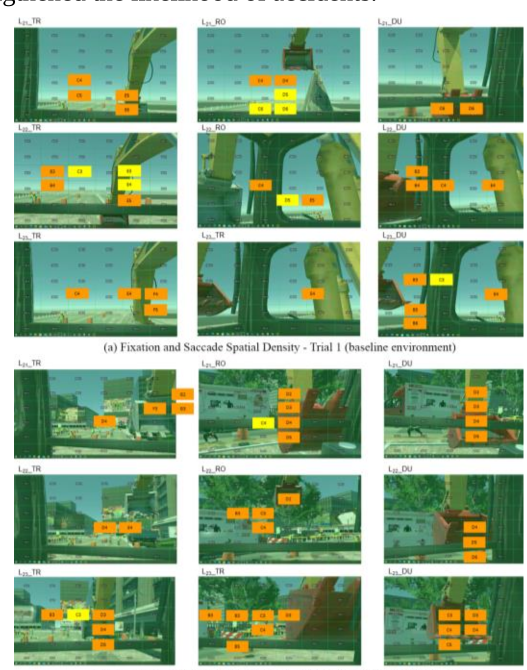

Figure 6. Attention spatial density : (a) Attention spatial density in trial 1 (b) Attention spatial density in trial 2

V. Conclusion

We proposed a multi-user immersive operation and communication system for excavation, and investigated how environment types and operator-spotter communication formats affect operator's performance and attention demand during human-excavator collaboration. In the user study, we found that in challenging environments, the operator tends to make more collisions and miss more signals from the spotter. Also, operator tends to have a wider distributed attention in the challenging environment. In the future work, we will conduct larger user studies and will analyze eye tracking and subjective evaluation results in a quantitative manner.

Acknowledgment

This work is supported by the National Science Foundation (NSF) under Grant No.2026574. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

[1] S.N. Naghshbandi, L. Varga, Y. Hu, Technologies for safe and resilient earthmoving operations: A systematic literature review, Automation in Construction. 125 (2021) 103632. https://doi.org/10.1016/j.autcon.2021.103632.

[2] S. Talmaki, V.R. Kamat, H. Cai, Geometric modeling of geospatial data for visualization-assisted excavation, Advanced Engineering Informatics. 27 (2013) 283–298. https://doi.org/10.1016/j.aei.2013.01.004.

[3] T. Çelik, S. Kamali, Y. Arayici, Social cost in construction projects, Environmental Impact Assessment Review. 64 (2017) 77–86. https://doi.org/10.1016/j.eiar.2017.03.001.

[4] H. Jeong, C. Arboleda, D. Abraham, Imaging and Locating Buried Utilities, Purdue University, West Lafayette, IN, 2003. https://doi.org/10.5703/1288284313237.

[5] S.A. Talmaki, S. Dong, V.R. Kamat, Geospatial Databases and Augmented Reality Visualization for Improving Safety in Urban Excavation Operations, (2012) 91–101. https://doi.org/10.1061/41109(373)10.

[6] A.J. Al-Bayati, L. Panzer, Reducing Damage to Underground Utilities: Lessons Learned from Damage Data and Excavators in North Carolina, Journal of Construction Engineering and Management. 145 (2019) 04019078. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001724.

[7] DIRT Report, (n.d.). https://commongroundalliance.com (accessed April 12, 2022).

[8] S. Li, H. Cai, V.R. Kamat, Uncertainty-aware geospatial system for mapping and visualizing underground utilities, Automation in Construction. 53 (2015) 105–119. https://doi.org/10.1016/j.autcon.2015.03.011.

[9] X. Su, S. Talmaki, H. Cai, V.R. Kamat, Uncertainty-aware visualization and proximity monitoring in urban excavation: a geospatial augmented reality approach, Vis. in Eng. 1 (2013) 2. https://doi.org/10.1186/2213-7459-1-2.

[10] S. Talmaki, V.R. Kamat, Real-Time Hybrid Virtuality for Prevention of Excavation Related Utility Strikes, J. Comput. Civ. Eng. 28 (2014) 04014001. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000269.

[11] CGA Best Practices Version 18.0, (n.d.). https://bestpractices.commongroundalliance.com/ (accessed April 12, 2022).

[12] Ohio Revised Code – OHIO811 | Call 811 Before You Dig | OHIO811, (n.d.). https://www.oups.org/law/ (accessed April 12, 2022).