Automation in Construction 167 (2024) 105683

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

Divergent effects of visual interfaces on teleoperation for challenging jobsite environments

Yeon Chae [a], Samraat Gupta [b], Youngjib Ham [a,*]

[a] Department of Construction Science, Texas A&M University, 3137 TAMU, College Station, TX 77843, USA.
[b] Department of Computer Science and Engineering, Texas A&M University, 3137 TAMU, College Station, TX 77843, USA.

A R T I C L E I N F O

Keywords:
Human-machine interface
Teleoperation
Spatial awareness

A B S T R A C T

In challenging work environments, teleoperation has been adopted for executing tasks over long distances, ensuring worker safety off-site. The physical distance from jobsites typically poses challenges to teleoperators, limiting their cognition and perception of sites. Particularly, these human-machine interface challenges vary depending on the work environment conditions. This paper analyzes the impact of different visual interfaces on teleoperators and their work performance in challenging jobsite environments, using interactive workplace modeling and simulation to execute teleoperation tasks. Based on experiments with 33 subjects in construction-related fields with industrial experience, the study observed data overload problems and distraction with effects varying by work environment. Additionally, the analysis includes how their eye gaze patterns could demonstrate these effects. The findings contribute valuable insights into human factors crucial for designing user-centered teleoperation interfaces in construction, forming the basis for an intuitive and informative interface design in task-specific settings.

1. Introduction

To support construction works in challenging environments like restoration of disaster sites, teleoperation has been adopted, enabling workers to remotely operate construction machinery from a secure location [1]. In these environments, various tasks such as clearing debris and conducting search operations are essential as an initial response. Teleoperation proves particularly advantageous in managing excavation tasks in unfamiliar environments [2]. This is because it requires operators to continuously engage in problem-solving [3], adapting to the unique challenges presented by each scene. Despite the benefits, teleoperation often suffers underlying problems with low productivity [2,4], along with heightened risks of tip-overs, especially in sloping terrain due to changes in the center of gravity and the effects of inertial forces [5]. Therefore, precise and safe operation is necessary for teleoperators based on their thorough spatial understanding provided by the human-machine interface [6]. During teleoperation with distance, by its nature of remote operation, teleoperators heavily rely on interfaces to interact with both machinery and their surrounding environment [7]. Unlike direct operation, visibility is often limited unless visual interfaces provide necessary information [8,9]. This underscores the critical role of visual interfaces in enhancing teleoperators' situational awareness and operational effectiveness. Moreover, teleoperators face additional challenges posed by surrounding work environments such as sloping terrain [5] and obscured areas [8], further complicating their tasks. These challenges accentuate the importance of spatial understanding during operations, as teleoperators should understand dynamic and potentially hazardous environments. Given that operators predominantly rely on visual interfaces for accessing crucial information, including spatial layouts and the real-time status of teleoperated machinery [10], the quality and effectiveness of the interfaces are paramount. Robust visual interfaces not only facilitate seamless interaction but also play a pivotal role in enhancing teleoperators' work performance and overall user experience [2,11]. Therefore, investing in the development and implementation of robust visual interfaces is essential for enhancing teleoperation outcomes.

Previous works [7,12] have endeavored to evaluate the effectiveness of visual interfaces, which provide multiple viewpoints in 2D or 3D wearable displays to teleoperators for enhancing their visual understanding of the site. To mitigate challenges of spatial understanding for remote operation, [8] has investigated how the autonomous control system of multiple cameras affects their work. Additionally, evaluations

* Corresponding author.
E-mail addresses: yeonchae62@tamu.edu (Y. Chae), samraatg@tamu.edu (S. Gupta), yham@tamu.edu (Y. Ham).

https://doi.org/10.1016/j.autcon.2024.105683
Received 11 February 2024; Received in revised form 4 August 2024; Accepted 5 August 2024
Available online 20 August 2024

have been conducted on the impact of visual prompts in multi-monitor systems to address challenges in multiple viewpoint settings [13]. Additionally, to provide immersive environment for teleoperation, 3D wearable displays has been leveraged, demonstrating a stronger sense of presence and higher level of immersion across multi-user collaboration [9] and tasks with crane training [14]. As such, prior works largely explored how visual interfaces affect teleoperators during work. Despite the efforts, a gap still exists in understanding how different visual interfaces affect teleoperators' operating capabilities and perceptions, specifically in challenging work environments where spatial awareness is substantially degraded, posing higher risks during tasks. In our study, we pose the research question: how the approaches to mitigate the challenges of degraded scene understanding in remote operation, such as additional viewpoints or immersive displays, effectively mitigate the challenges given from demanding job sites characterized by sloping terrain and limited visibility?

Figure 1 depicts the overview of this study. For systematic human-machine interface design, different factors for the system design (e.g., environments, tools and technology, construction robot, task, user) should be comprehensively considered [15]. This research evaluates the impact of visual interfaces, specifically multiple viewpoints with enhanced visibility and head-mounted displays, in challenging jobsite environments, focusing on task performance and teleoperators' human factors. The research objective is to investigate how these approaches alleviate challenges when teleoperation is conducted in challenging job sites. We analyze the effects of each interface compared to a single viewpoint setting, which serves as the baseline. During the experiments ($N = 33$), task performance (e.g., completion time, amount of work done, number of collision occurrences) are analyzed with the questionnaire such as NASA-TLX, situational awareness, and presence questionnaire. Human data including eye tracking data (e.g., fixation, saccades) is also collected and analyzed along with the following in-depth discussion regarding how the visual interfaces affect teleoperators in challenging environment settings.

This paper is structured as follows: Section 2 explains the visual interfaces and degraded spatial awareness in teleoperation. Section 3 describes the modeling and simulation along with the experimental design. Section 4 describes the experiment results and discussion on eye gaze patterns as visual attention with the analysis of the visual interface usage by work environments. Section 5 summarizes findings and discusses future works, followed by section 6, which outlines limitations.

2. Research background

2.1. *Visual interfaces enhancing visibility to mitigate challenges in teleoperation*

A primary factor contributing to limited spatial awareness of teleoperators is the scarcity of information regarding work environments and their surroundings, which poses a challenge in teleoperation tasks [2]. The constraints of the teleoperation interface compromise the operator's ability to accurately perceive work environment such as terrain slopes and obstacles with respect to the machine position [5]. Consequently, the operators often struggle to assess the excavator's orientation, its proximity to potential hazards, and the precise angles necessary for safe and efficient manipulation. This reduction in spatial awareness has exacerbated when the operator has not been fully provided with the required information (e.g., occlusion) [8]. To enhance teleoperator's visibility by providing additional viewpoint, understanding the characteristics of viewpoints is paramount [16]. A viewpoint can be categorized into two: first-person view (FPV) and third-person view (TPV). Within the realm of FPV, users' actions and engagements directly synchronize with their physical bodies, fostering a sense of immersion and embodiment within the given landscape, the perception of the environment is linked to the viewpoint offered [17]. Moreover, it could create an intuitive and natural sense of depth and spatial relationships within the visible area. On the other hand, the third-person view offers a viewpoint from an external position, allowing operators to observe the work environment as if they were viewing it through a camera or an observer's eyes [12,17]. This enables teleoperators to evaluate the spatial arrangement of the environment and the relative positions of objects without being limited by their personal physical actions.

Based on this exploration, to enhance visibility and spatial understanding, it is crucial to provide appropriate information tailored to task settings in an effective and intuitive manner [13]. In manipulation of heavy machinery, previous literatures have provided different viewpoints tailored to their task and work environment conditions. In [18], a direct vision teleoperation system was developed, and experiments were conducted by observing the excavator's diagonal front and side views, presuming a horizontal towing operation near underground utilities. Remote operation from a distance poses additional challenges compared to direct teleoperation, as teleoperators must rely solely on interfaces, limiting their exploration. To mitigate such challenges, [8] implemented an adaptive imaging system for the manipulator with end effector. Tailored to the specific task needs, including a third-eye viewpoint and zoom-in images to illustrate the distance between the end effector and the target object, [8] highlighted relationship between required

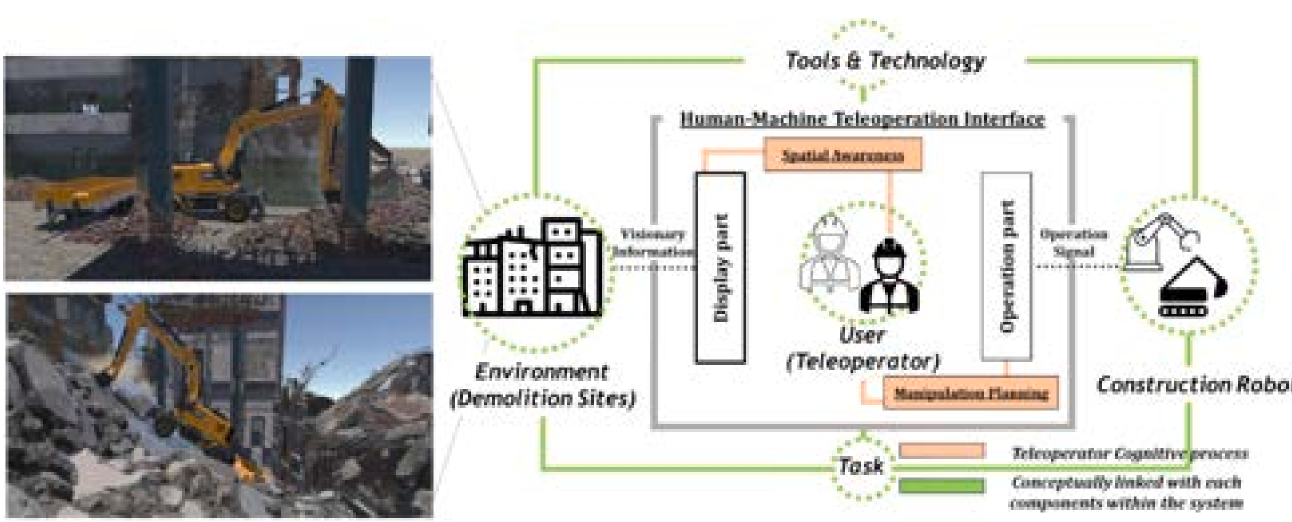

Fig. 1. User-centered and systematic study on visual interfaces of construction teleoperation.

operations and images in work situation during each of situation encompassing movement, manipulation with end effector attached to the manipulator. [19] investigated the allowable ranges of pan and tilt angles of viewpoint of additional side view when cabin view is provided in default focused on digging and releasing which especially requires external views for delicate manipulation. Furthermore, [20] has implemented a top view-based visual monitor for operators engaged in heavy machinery manipulation, aimed at mitigating collision-related safety hazards. This viewpoint has been used where the task requires spatial understanding, including navigation. Prior studies [7,21] selected top view based on its characteristics of showing spatial layout of teleoperated remote robot with their surrounding environments.

Based on the characteristics of viewpoints, in the realm of teleoperation, the choice between multiple and single view interface has emerged as a critical consideration. Still concerns about distraction arising from multiple viewpoints have indicated the importance of the single view interface [21,22]. Recent studies [21,23,24] have collectively demonstrated the effectiveness in enhancing the teleoperator's ego-centric spatial representation. Despite the benefits of ego-centric spatial representation, in teleoperation of heavy machinery on sites, additional viewpoints may be required to address the scarcity of information, such as occlusion [8], depending on the context. Still, it is essential to note that adding multiple viewpoints does not always improve their visual understanding; operators must judiciously select and switch the viewpoint that best suits their task every moment [13]. This implies the importance of not just supplying adequate data but also meticulously considering how to present it, demanding a thorough examination of display methods. Given that challenges during teleoperation are site-specific, work environment considerations are crucial in designing visual interfaces for challenging construction contexts.

Especially, when individuals are tasked with combining two different perspectives, such as their own viewpoint (egocentric) and that of another person (allocentric), they may face challenges [25]. [13] highlighted the potential issue of multiple viewpoints potentially distracting teleoperators and proposed the design of visual prompts to enhance their situational understanding. However, this approach could exaggerate the issue of cognitive tunneling [26] which could be originated from excess visual information [27]. The effectiveness of multiple viewpoint setting can be subject to work environment variations (i.e., terrains). Particularly when dealing with misaligned viewpoints, this issue becomes critical when the perspectives of teleoperators clash with the real-world terrain slope, potentially leading to information overload. How multiple viewpoint setting affect performance and this can be varied by work environment setting even through the same visual configuration (viewpoint) is provided yet to explored, posing limit to understand the multiple viewpoints' affect to manipulating heavy machinery remotely.

2.2. Approaches to providing immersive experiences to mitigate degraded spatial awareness

To mitigate the challenges of degraded spatial awareness, [9,14,28] have investigated the evaluation of HMD (Head-Mounted Display) as an immersive display, exploring how it impacts user experience and work performance. In teleoperation of moving vehicles, HMD-based display has been investigated to enhance teleoperator's telepresence [29,30] by enhancing distance perception [31]. To provide 3D layouts in 2D display has an inherent challenge on its depth ambiguity [32]. [33] demonstrated its effect to increase situational awareness during underwater operations with grabbing the target object. Comparing the characteristics of HMD over 2D displays, [28] measured the impact of HMD-based experience over 2D display setting on construction design review tasks. The task was to review the piping system via a virtual walkthrough, detecting design errors and planning for the installation sequence of the piping system. [14] thoroughly investigated the effectiveness of virtual reality (VR) crane training using head-mounted displays (HMDs). VR training has been provided with controlling cranes, demonstrating VR training's effectiveness in enhancing crane control skills as it provides realistic virtual experiences. [34] demonstrated effectiveness of VR training using HMD on improving construction workers' knowledge and safety behavior in teleoperating a demolition robot. These investigations into the impact of HMDs have demonstrated their potential to enhance teleoperator's situational awareness and improve task performance. However, the effectiveness of visual interfaces can vary depending on the work environment (e.g., different terrains). In previous literature, the effectiveness of HMD has primarily been investigated when the viewpoints between the user's physical position and the virtual environment are aligned. A human's perception can be different based on their standing posture and visual information's position conditions [35,36]. The misalignment of viewpoints between the user's physical position and the work environment can influence how well the HMD supports teleoperation tasks, making it essential to consider work environments in such scenarios.

Though the benefits of HMDs, simulation sickness is representative drawbacks in HMD [37]. It is caused by the discrepancy between the degree-of-freedom of the user's perception and their actual movement [38]. Additionally, multi-modal interfaces offer promising avenues for improving teleoperation systems as haptic feedback can provide assistive support during teleoperation tasks. Especially, force feedback [39], provide users with tactile sensations or force response based on their interactions with virtual or remote environment. In human-robot collaboration, for an intuitive and direct control in teleoperation tsks, various methods with brain-computer interfaces (BCI) [40], electromyography (EMG) [41] were investigated. These methods enhance hands-free control using human gestures reducing human's workload in collaborations with robots. These studies investigated the assistive technology to support their work within a certain set of visual interfaces. However, considering visual modality captures up to 70% of waking attention [38], a comprehensive understanding of visual interfaces should precede integration with assistive modalities in order to effectively design assistive technology to support their work.

2.3. Knowledge gap

Despite efforts to measure the impact of visual interfaces and understand the factors influencing the impact, challenges persist, particularly in work environments with demanding conditions such as sloping terrain and obstructed views. In such challenging jobsites, operators may face additional challenges that hinder their ability to perceive and interpret visual information accurately. For instance, limited visibility condition due to obstacles or misaligned viewpoints can lead to distorted spatial understanding, impeding the operator's performance. As there's a limitation in exploring their surroundings freely due to distance, a degree of teleoperators' awareness is limited due to the restriction to the camera view [42]. Moreover, during teleoperation tasks, obstacles and debris may continually emerge, further complicating the work [8,43]. However, how visual interfaces affect teleoperation in such scenarios is still unknown. To bridge this gap, there's a need to investigate the interaction between visual interfaces, challenges during teleoperation, and teleoperator performance in such scenarios. Our study focuses on how these visual interfaces can mitigate the challenges in teleoperation when additional challenges arise in work environments (e.g., sloping terrain, limited visibility) which can deteriorate teleoperator's spatial understanding.

3. Methodology

Figure 2 illustrates the overview of this study. Workplace modeling and simulation are conducted, allowing teleoperators to manipulate heavy machinery based on their interaction with the surroundings. Experiments were conducted to analyze the impact of different visual interfaces to teleoperators and their work performance under challenging

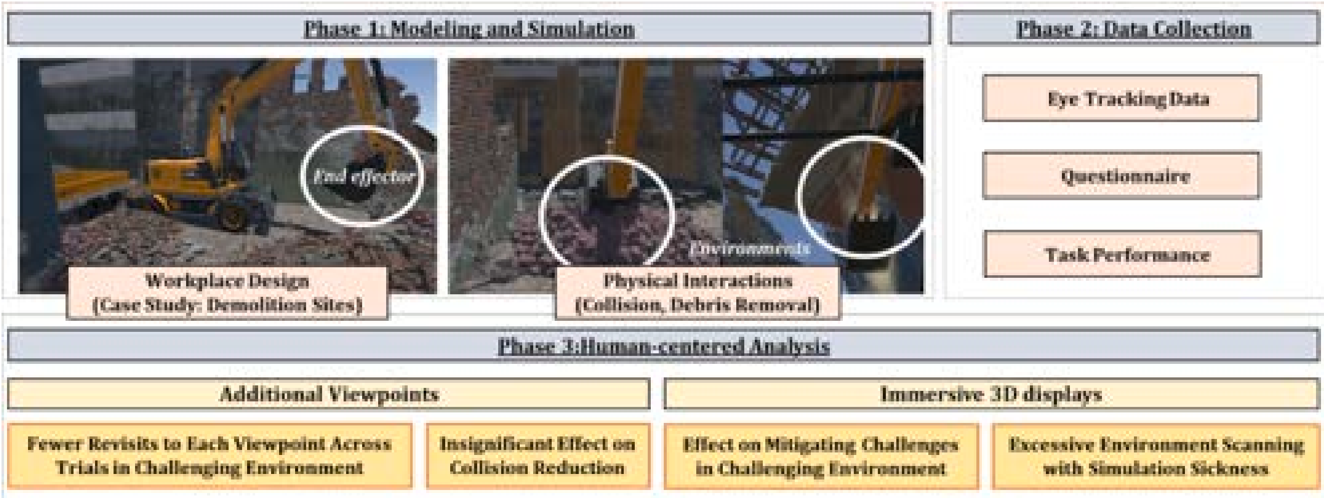

Fig. 2. Exploring divergent impacts of visual interfaces by challenging jobsite environments.

jobsite environments.

3.1. Modeling and simulation

3.1.1. Workplace modeling

To analyze the impact of visual interfaces in challenging work environments, this study includes designing a model to replicate the interactive work environments in demolition sites as a case study. The model was built in C#, providing features for handling collisions between rigid bodies in the virtual environment. Having the excavator as a physics body responding to collisions with rigid bodies can lead to undesirable situations such as inaccurate responses to collisions and the excavator continuing to move after colliding with its surroundings which is not occurring in real-world conditions. Thus, we developed a new system for handling collision responses of the excavator with objects of interests such as obstacles and debris, allowing for the simulation of complex interactions inherent in debris removal tasks across diverse terrains. These simulated environments not only serve as workspaces for teleoperation but also function as dynamic interactive platforms that respond to teleoperation commands.

The modeling specifically aims to replicate the interactive work environments in demolition sites as a case study. Our goal is to model jobsite characteristics that could be generally applicable, including diverse slope conditions and obstacles not limited to demolition sites. The scenario involves two work environments, baseline and challenging scenes, as illustrated in Fig. 3. In the baseline scene (Fig. 4 (b)), the site has a flat terrain, and there are obstacles that could work as a hindrance during teleoperation, and unsound building structures. The operators were supposed to attentively avoid such obstacles during their manipulation. Additionally, dust is modeled surrounding the debris pile located. On the other hand, in the challenging scene, the terrain has a slope of around 30° based on OSHA safety rule [44]. In terms of layout perspective, the obstacles are located nearer than the baseline, and the level of difficulty to avoid the obstacles is harder compared to the baseline scene. Excluding the slope of terrain and the proximity of the obstacles, the scene settings are consistent with the baseline. Both scenes have three primary obstacles consisting of two pillars with surrounding environments. In the baseline scene, the poles are 11.62 m and 9.62 m away from the center of the excavator respectively. The distances to the pillars are designed in such a way that the operators are forced to maneuver away from them and cannot simply control the excavator in the given default configuration. The wall prevents the operators from avoiding the pillars by rotating in the opposite direction. In the challenging scene, the distance to the pillars is 8.78 m and 8.41 m away from the center of the excavator respectively.

The physics part designed in our model is divided into the following:

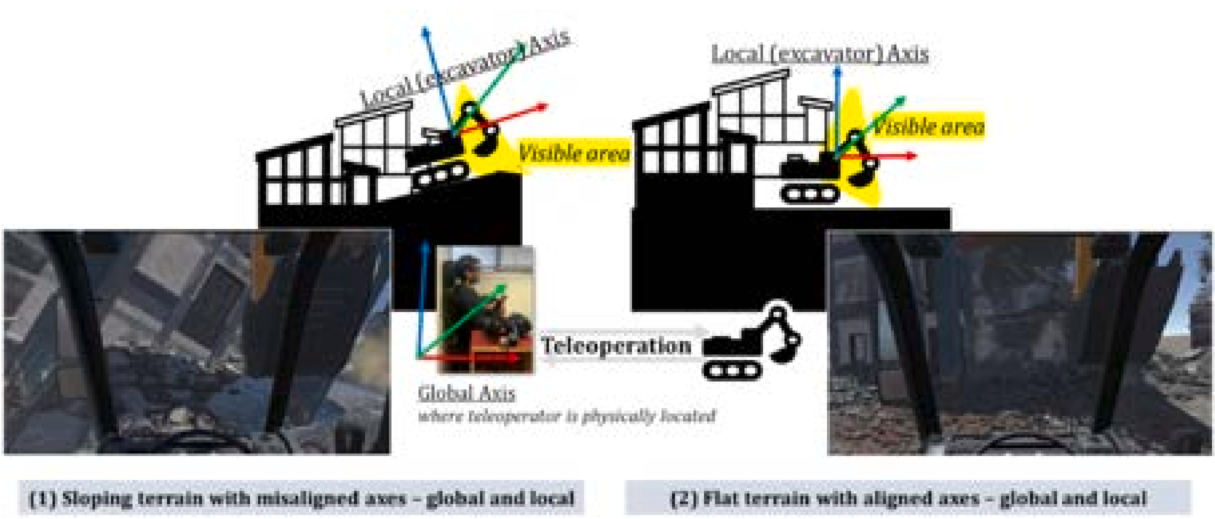

Fig. 3. Challenges in misaligned axes of different perspectives from a teleoperator and an excavator in a sloping terrain.

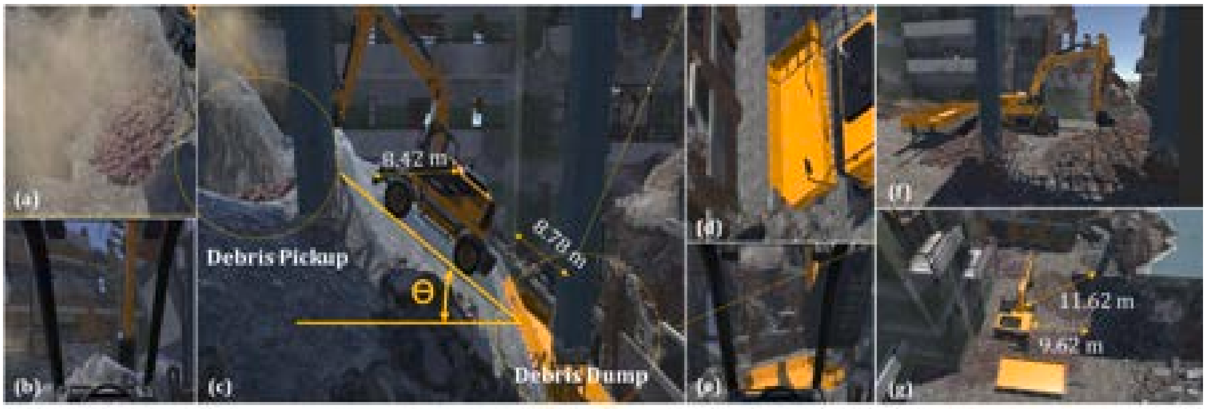

Fig. 4. Scene layout for demolition site modeling: (a-e) sloping terrain scene (a: enlarged image of the debris pickup area, b: cabin view for picking up debris, c: scene layout of sloping terrain, d: enlarged image of the dumpster area, e: cabin view for releasing debris), (f-g) flat terrain scene.

(1) scatter debris module, (2) count debris module, (3) debris control module, (4) terrain detection module, (5) collision detection module, as illustrated in Fig. 5.

As illustrated in Fig. 5. (a), scatter debris module, in order to make a pile of debris, we randomly sample unit vectors and apply translation and scaling transformations as required. In this module, we are generating multiple debris by creating clones of a single debris. During this process, we utilize GPU instancing to reduce the computational time required for rendering debris.

The task starts when the debris has fallen on the ground. After the excavator picks up the debris, to keep track of the amount of debris in the bucket, as illustrated in Fig. 5. (b), a global variable named "debris count" is set up. This variable can be called and controlled globally in the model. Each debris, as a game object, is assigned a tag to determine whether it is debris or not. The value of a "debris count" variable increases every time as a game object with a tag collides or is in collision with the bucket, and it decreases when the collision ends, as illustrated in Fig. 5. (b). The same method is used when counting the amount of debris in the trailer. The logic flow for counting collisions and debris during each task is illustrated in Fig. 6. C# scripts with physics engine is used to follow the logic during the manipulation.

After debris are picked up, operators should rotate the excavator to reach the dumpster area. For debris control, especially when debris are inside the bucket, to avoid false movement (e.g., falling out of debris), debris control module was custom designed in our model (as illustrated in in Fig. 5. (c)). To overcome the issues of friction not to allow debris to come out of the bucket after they have been picked up, collision-based parenting is conducted when once a debris was picked up and the angle between the stick and the bucket reaches the empirically tested angle δ as illustrated in Fig. 7. (a). In this status, the debris would be attached to the bucket and would no longer fall out. The debris would only be allowed to fall out from the bucket once the angle of the stick and the bucket reaches larger than δ, by the rotation of the bucket manipulated by the controller, which typically occurred during picking up and dumping tasks. In this study, δ was empirically set to around 48.70° based on multiple trials relative to the starting positions and the rotation of the bucket, which enables to control the release or retention of debris based on the teleoperators' intention. This value is not optimized yet, so further adjustments could be made in the future. The logic of this module is illustrated in Fig. 8.

By default, the unity game engine has support for handling collisions between rigid bodies. All objects in the virtual environment are designed

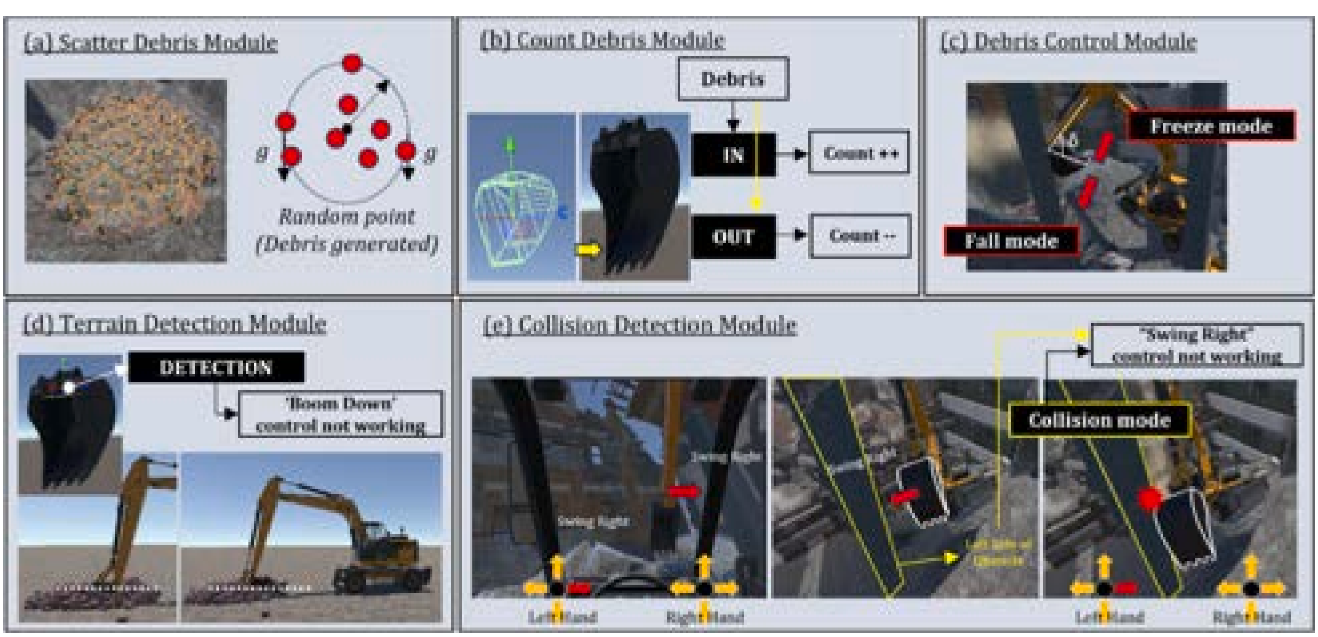

Fig. 5. Custom-designed modules for debris removal task in our model.

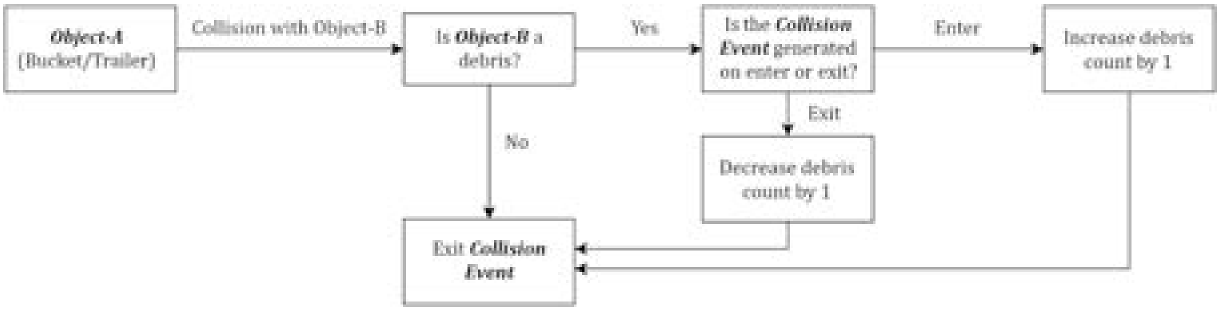

Fig. 6. Logic flow for counting debris for picking up and dumping task.

Fig. 7. Debris control module: (a) tracking δ during bucket rotation, (b) freeze mode after δ below than threshold, (c) release mode with "bucket curl out".

Algorithm 1 Debris Control Module

```
function DEBRIS CONTROL UNDER FREEZEMODE(FreezeMode: Debris-Movement)
    if FreezeMode is true: then
        Debris' individual movement is frozen, and it moves along with Bucket
    else
        Debris not frozen and can move independently
    end if
end function
function DEBRIS CONTROL(DebrisInTheBucket , δ : FreezeMode)
    if Debris are in the Bucket then
        DebrisInTheBucket ← Add the objects tagged as Debris in the Bucket
        if DebrisInTheBucket and δ > 0.85 then
            FreezeMode ← true
            Debris Control Under FreezeMode
        else if δ < 0.85 then
            for each debris in DebrisInTheBucket do
                FreezeMode ← false
                Debris Control Under FreezeMode
            end for
        end if
    end if
end function
```

Fig. 8. Pseudo code for debris control module.

as rigid bodies, ensuring that occlusions occur realistically when objects interact. Building upon this functionality, we customize the train detection module (as illustrated in Fig. 5. (d)), to detect terrain and control bucket movements accordingly. Additionally, we customize collision detection module, as illustrated in Fig. 5. (e) from the challenges of utilizing collisions between rigid bodies supported in the unity game engine. While this functionality makes excavator movements once it collides with the other objects (e.g., obstacles) in the virtual environment, excavations with collisions are infeasible in real-world environments. Furthermore, in our model, we made excavator's main body remains fixed while its movable parts (including digger, bucket, and arm) should be operable.

The collision detection module follows the logic depicted in Fig. 9. First, once a collision occurs with obstacles, as illustrated in Fig. 10. (c), our model is designed to track operator input signal and recognize collision's direction (e.g., left or right). Based on the collision's direction, we limit the operation so that they can recognize they are collided with obstacles, as their manipulation is not working, and they should manipulate with different input command. Additionally, as illustrated in Fig. 10. (c), sudden dust and hitting sound effect occur when a collision happens during work (e.g., hitting a structure). Regarding the interaction with the environment and heavy machinery, when the end effector (i.e., bucket) interacts with the work environment, "collider" physics engine was leveraged to detect the collision.

3.1.2. Teleoperation simulation: Hardware and visual interface setup

In the operation interface, the input from the joystick is converted to keyboard inputs. The movement of the excavator is controlled with these inputs, using scripts in C# which define how much a given component should translate or rotate based on a given input. The script in our model also includes the functionality to control the movement speed of the cab, the stick, and the bucket. This allows for the excavator to have eight degrees of freedom with the ability for multiple inputs to be entered simultaneously following the Society of Automotive Engineers (SAE) control pattern, as illustrated in Fig. 11. The operation involves rotation, pickup and dump, and obstacle avoidance tasks. To reduce simulation sickness, excavator's rotation speed is set to 15.88 (°/s). The movement speed for the stick and bucket is set following the hydraulic excavator model, and finetuned based on the trials. Additionally, bucket rotation speed is set as 10.91(°/s) and digger reach out speed is set as 12 (°/s). All these parameters are set based on our pilot testing [45]. The detailed setting and logic regarding the tasks are

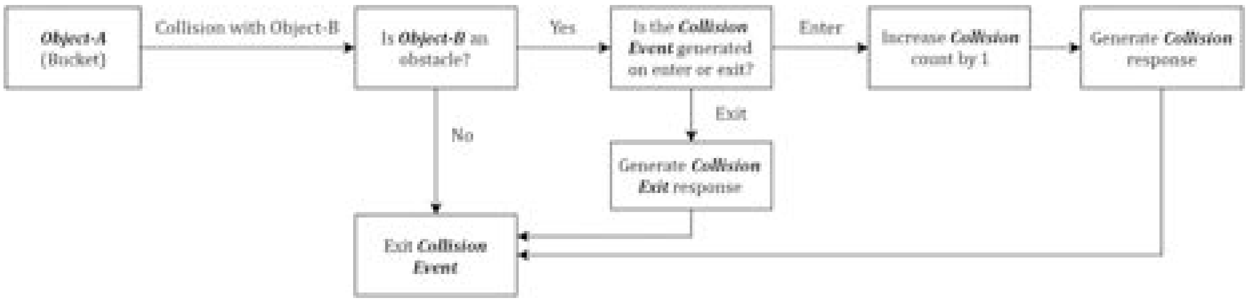

Fig. 9. Logic flow of collision detection module.

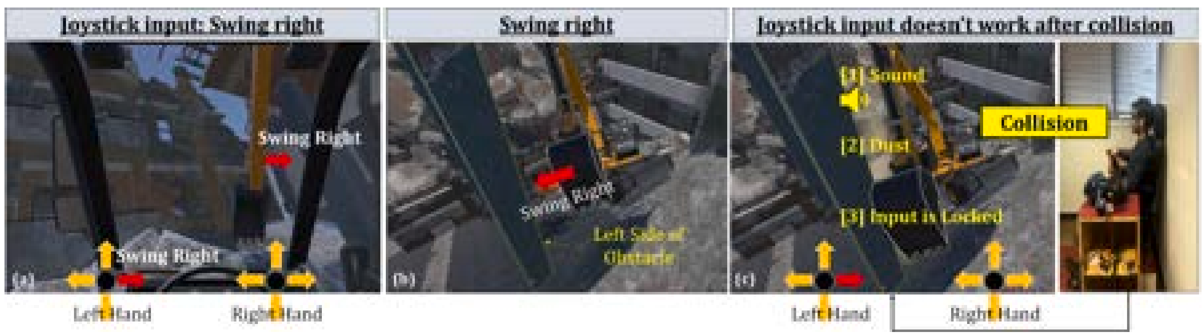

Fig. 10. Collision detection module: (a) cabin view during 'swing right' manipulation, (b) collision tracking, (c) collision-based control restriction.

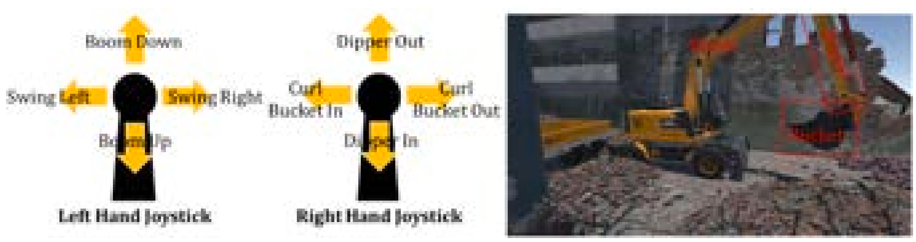

Fig. 11. SAE joystick control pattern for excavator controls.

described in Section 3.1.1, and the overall hardware settings are illustrated in Fig. 12.

In the experiments, we explored different visual interfaces, such as settings with single viewpoint, multiple viewpoints, and HMD during teleoperation. The HMD is limited to a single view, as many current approaches, as indicated in numerous studies [9,14], aim to enhance the immersive experience using a single view. As our study is to analyze the impact of current practices adopted to enhance spatial understanding of jobsite environments to workers, adding multiple viewpoints or providing HMD is within our scope. Both single viewpoint and HMD settings, the viewpoint is ego-centered from the camera located in the machine (i.e., cabin camera). With the view generated from the main camera located inside the cabin (where the operator sits in during on-site operation), teleoperators feel ego-centric perception as they sit inside the cabin. On the other hand, in the multiple viewpoint setting, additional viewpoints are provided, offering depth information that might be challenging to discern using the cabin-located camera alone. Based on previous studies [8,13] on the configuration and implementation of multiple viewpoint settings, multiple viewpoint setting in this experiment consists of the side view and the top view along with the cabin view (default setting). In our model, visibility is limited in the cabin view especially for picking up area. The purpose of additional viewpoint selection is to provide the essential information required. During manipulation of the bucket and rotation of the heavy machinery, the information of distance between the bucket and work area, along with a sense of depth, should be provided to teleoperator for the given

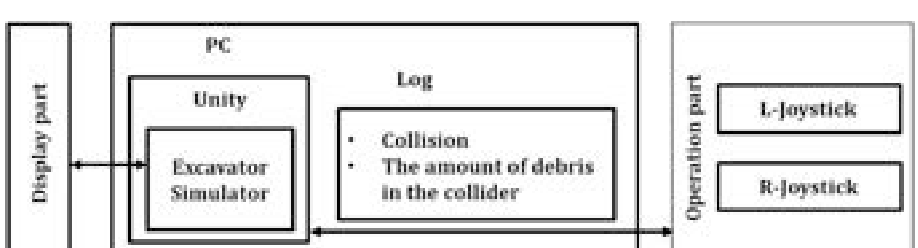

Fig. 12. Hardware settings for teleoperation simulation in this study.

task [8]. Following this principle, external side views are required [19], and construction machinery studies [18,19] have selected side view in experimental studies. In our experiment, illustrated in Fig. 13, side view is selected as additional viewpoint as this provides depth-related data from the bucket to the work area, which have limited visibility in the cabin view, based on its proven effectiveness from the pilot testing ($N =$ 10) [45].

On the other hand, for tasks that involve movement, such as rotating excavator's body, ensuring safe and efficient positioning is paramount. To avoid obstacles, even if they are visible from the cabin's view, a top view has a strength considering the ambiguity of depth perception in a 2D display [42]. This view provides spatial information [46], including obstacles typically found in proximity to heavy machinery, essential for successfully completing their tasks [8]. From this rationale, top viewpoint was adopted and configured as an additive viewpoint. In terms of the layout perspective, the cabin view is selected as the main viewpoint. This selection is based on setting the cabin view as the default viewpoint in the single viewpoint setting. This layout in multiple viewpoint setting offers teleoperators to selectively utilize additional viewpoints as needed.

3.2. Experimental setup

Table 1 presents the information of a total of 33 participants in the experiments approved by the Texas A&M University Institutional Review Boards (IRB). All participants were recruited via the bulk mail system (only for those who are in construction-related majors) and before starting the experiment, the questionnaire regarding the relevant industrial experience included was shared to ensure the eligibility of operating heavy machinery without visual and intellectual difficulties. Only those who were satisfied with these conditions participated in the study, assuming that they were aware of construction-related issues such as productivity and safety. Prior to the experiment, a pilot study has been conducted to test the prototype functionality with employing a small group of users ($N < 10$) [45]. A within-group experiment was conducted in a single day.

As illustrated in Fig. 14., the experiment started with a brief introduction of the overview of the experiment and the given tasks emphasizing both productivity and safety, which indicates they should complete their task within the given time while avoiding collision with obstacles and maximizing the work completed. The practice session allows the participants to be familiarized with controls.

The experiment consists of two sessions, one with sloping terrain scene (as illustrated in Fig. 15. (e)), and the other with flat terrain scene (as illustrated in Fig. 15. (c)). After the practice session is completed, all subjects participated in the two sessions, but in each session, the order of the scene setting was randomly given. Each session consists of three different interface settings, consisting of single viewpoint setting, multiple viewpoint setting, and HMD. All participants participated in all six settings but in a random order to avoid their order affecting the results.

The given task is to move debris, known as a main task especially in demolition sites [43]. The task is designed to assess the operators'

Table 1
Pre-survey results for the participants ($N = 33$).

Variables	Scale	#
Age	19–21	10
	22–24	11
	25–27	8
	>27	4
Gender	Male	30
	Female	3
Education Level	Undergraduate	16
	Graduate	17
Industry Experience	<1 yr	23
	1–3 yr	6
	>3 yr	4
Job Titles	Field Engineer	24
	Civil Engineer	3
	Project Manager	2
	Project Engineer	1
	Architect	3

spatial awareness and the depth perception of the target objects (e.g., debris for picking up, dumpster for unloading) and obstacles in the surrounding environments. One trial includes picking up the debris, avoiding the obstacles along the way to the dumpster, and finally unloading all the debris in the bucket to the dumpster and turn back to the starting point. The task consists of two sequential sets of picking up debris, avoiding obstacles, and dumping them, with the aim of keeping the time taken under 7 min to prevent simulation sickness. The limit of time is set based on the previous study [47] which was aligned with our pilot testing [45]. After the session, an open-ended interview gathers feedback on participants' experiences in the context of human factors. The metrics of the task follows [9]. The metrics to evaluate the task can be divided into task performance and safety. Regarding task performance, the amount of work done, and the completion time are measured. For evaluate their unsafe behavior, the number of collision occurrences is measured.

3.3. Evaluation metrics

The performance assessed encompasses the amount of work done, completion time and collision occurrences [9]. We also evaluate the effectiveness among different types of visual interface for teleoperation in aspects of presence in challenging environment settings and simulation sickness. The questionnaire encompassed a comprehensive assessment of participants' experiences, involving the NASA-TLX questionnaire to gauge the perceived task load, the presence questionnaire to evaluate the sense of presence within the work environment, and the situational awareness assessment to ascertain their awareness of the work context. The NASA-TLX evaluates the task load in terms of mental demand (MD), physical demand (PD), temporal demand (TD), performance (P), effort (E), frustration (F). Situational awareness [48] was measured from the questionnaire of Situational Awareness Rating Technique (SART), measuring with multiple dimensions as follows; familiarity of the situation, focusing of attention, information quality,

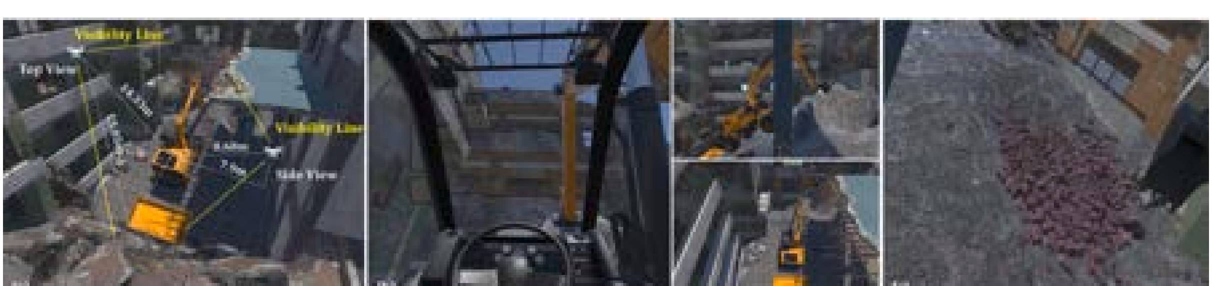

Fig. 13. Viewpoint setting and camera setup: (a) camera setup, (b) multiple viewpoint setting, (c) free field of viewpoint controlled by teleoperator using HMD.

Fig. 14. Experiment procedure.

Fig. 15. Experimental settings: (a-c) screen setting and (d-e) HMD setting.

information quantity, instability of the situation, concentration of attention, complexity of the situation, variability of the situation, arousal, and spare mental capacity. The given work environments are dynamic, with frequent unexpected events and changing conditions, demanding that teleoperators possess a thorough grasp of the scene settings during work. The simulation sickness is evaluated only during the HMD session to assess participants' level of simulation sickness levels and how it might be influenced by the use of HMD. The combination of these questionnaires and open-ended discussions provided a broaden understanding of participants' behaviors and perception during teleoperation, ensuring a comprehensive evaluation of the impact of different visual interfaces on various aspects of their engagement and work performance. Presence is a theoretical concept that describes the extent to which a medium represents real life in both physical and social environments [14]. Such perception is commonly understood as the feeling of "being there", a psychological state where individuals feel as if the virtual experience was real. Presence questionnaires in our experiment built upon [14], following 7 point Likert scale using the questions: 1) "The obstacle seemed to be real." 2) "The scenery felt real." 3) "I felt as if I really removed debris and avoided obstacles during the manipulation." 4) "I felt like I was sitting in the excavator cabin." Additionally, to assess the immersive display, simulation sickness is evaluated [49], specifically for HMD sessions. The given work environments are dynamic, with frequent unexpected events and changing conditions, demanding that teleoperators possess a thorough grasp of the scene settings during work.

4. Experimental results and discussion

4.1. Task performance varied by visual interfaces in different work environments

In this research, work performance was assessed based on the number of debris that have been dumped, instances of collision with obstacles, and the time taken to complete the assigned tasks. The results show significant differences in different work environment settings. To examine group differences, we built upon the one-way ANOVA followed by Tukey post-hoc *t*-tests. Fig. 16 (a) shows the amount of work done in both sloping and flat terrains. In the sloping terrain, the amount of work done of both multiple viewpoint setting (*M*:134.27) and HMD setting (*M*:132.53) show significant performance enhancement ($p < 0.001$) compared to the single viewpoint setting. Also, in terms of completion time, both multiple viewpoint setting (*M*:3.69 (unit: mins), $p < 0.01$) and HMD setting (*M*:3.25 (unit: mins), $p < 0.001$) show significant time reduction effect compared to the single viewpoint setting. However, the multiple viewpoint setting (*M*:0.27 (in a flat scene), *M*:1.18 (in a slope scene)) doesn't have a significant reduction of collision occurrences in both scenes ($p > 0.05$) compared to single viewpoint setting (*M*:0.16 (in a flat scene), *M*:1.91 in a slope scene)). HMD setting (*M*:0.58) significantly lower the collision occurrences in the sloping terrain ($p < 0.001$). However, in the flat terrain, there were no significant differences between any group of single viewpoint setting (*M*:0.15 (for collision), *M*:2.83 (for completion time, unit (mins)), *M*:136.58 (for the amount of work done), multiple viewpoint and (*M*:0.27 (for collision), *M*:3.11 (for completion time, unit (mins)), *M*:135.64 (for the amount of work done)

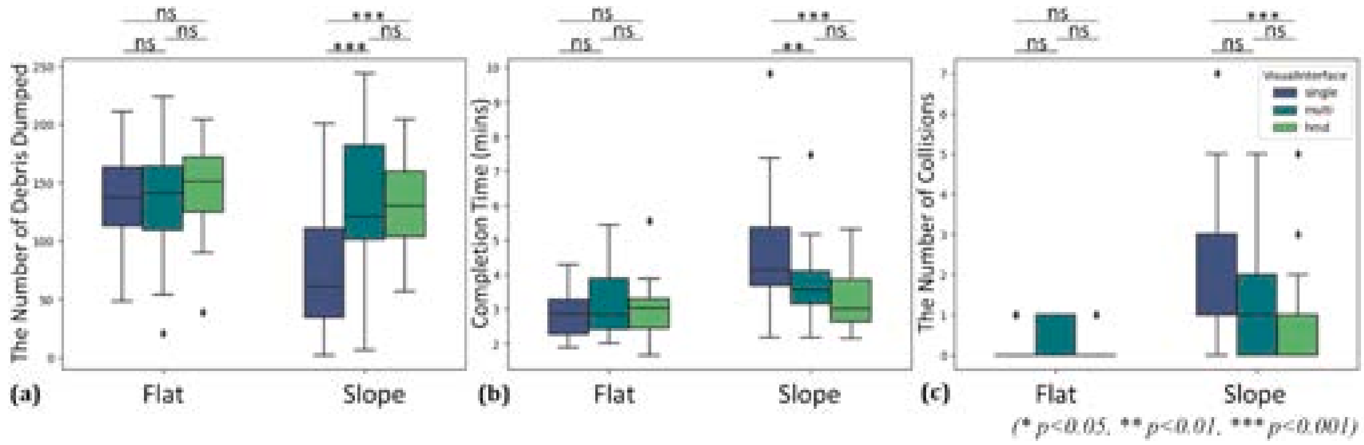

Fig. 16. Work performance across different visual interfaces and different work environments assessed by (a) work completed, (b) completion time, and (c) collision occurrences.

HMD setting (*M*:0.06 (for collision), *M*:2.92 (for completion time, unit (m)), *M*:145.72 (for the amount of work done) for all the measures in productivity ($p > 0.05$).

Based on their age, participants were divided into the following age groups: Group A (19–21), B (22–24), C (25–27), D (>27). In the sloping terrain, the average completion times for each group were: Group A (*M*:3.54), Group B (*M*:3.66), Group C (*M*:4.43), and Group D (*M*:4.23). Additionally, the average amount of work completed by each group was: Group A (*M*:120.7), Group B (*M*:140.31), Group C (*M*:86.67), and Group D (*M*:86.17). Collision occurrences with the following averages were Group A (*M*:1.47), Group B (*M*:1.06), Group C (*M*:1.29), and Group D (*M*:0.92). In terms of industrial experiences, participants can be divided into following groups: Group A (<1 yr experience), Group B(1–3 yr), Group C (>3 yr). As illustrated in Fig. 17., only in multiple viewpoint setting, Group B ($p < 0.01$) and Group C ($p < 0.001$) whose industrial experience is over 1 yr significantly show a higher performance in terms of the number of debris dumped compared to Group A whose industrial experience is <1 year.

4.2. Human factors

In the assessment of the perceived workload based on the NASA-TLX, the impact of different visual interfaces on task workload perception in both flat and sloping terrains was examined. The NASA-TLX evaluates the task load in terms of mental demand (MD), physical demand (PD), temporal demand (TD), performance (P), effort (E), frustration (F). The level of presence and simulation sickness are evaluated to quantify how immersive their experience is and the difficulties they might suffer with. Fig. 18 depicts the NASA-TLX across different visual interface settings and different work environments. The average score of the NASA TLX is significantly different ($p < 0.001$) between different work environments.

In the flat terrain, the average score of the NASA-TLX in the multiple viewpoint setting is not significantly different compared to the single viewpoint setting ($p > 0.05$). However, the HMD interface, in contrast to the single viewpoint setting, exhibited a statistically significant reduction in perceived task workload ($p < 0.05$). Multiple viewpoint setting shows insignificant differences in MD, PD, TD, P compared to single viewpoint settings ($p > 0.05$), while still demonstrating significantly lower E and F ($p < 0.05$) compared to the single viewpoint setting. Meanwhile, the HMD setting still exhibited a significant decrease in task load for MD, P, E ($p < 0.05$) and F ($p < 0.001$) compared to the single viewpoint setting.

In the sloping terrain, the average score in the multiple viewpoint setting is not significantly different compared to the viewpoint setting ($p > 0.05$). However, the HMD interface, in contrast to the single viewpoint setting, exhibited a statistically significant reduction in perceived task workload ($p < 0.01$). The multiple viewpoint setting shows significant reduction in all six sectors including Md, PD, TD, P, E and F ($p < 0.05$). The HMD setting displayed a significant decrease in task load compared to the single viewpoint setting ($p < 0.001$ for MD, P, and F; $p < 0.01$ for TD; $p < 0.05$ for E), except for PD ($p > 0.05$).

Table 2 presents the level of presence that the participants perceived in different visual interface settings and different work environments. In both environments, the participants mentioned the HMD setting has a

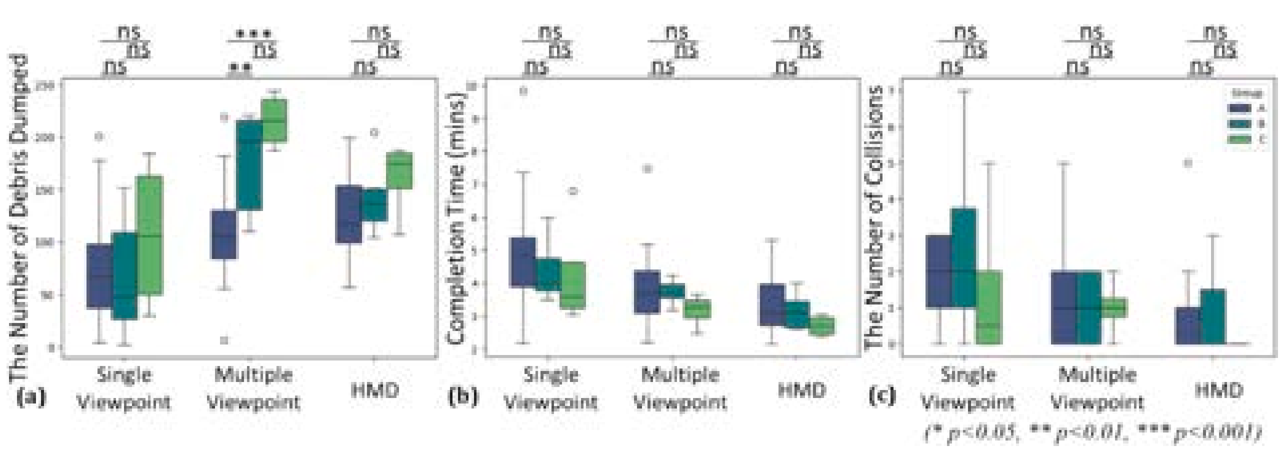

Fig. 17. Work performance across different visual interfaces in sloping terrain grouped by industrial experience, assessed by (a) work completed, (b) completion time, and (c) collision occurrences.

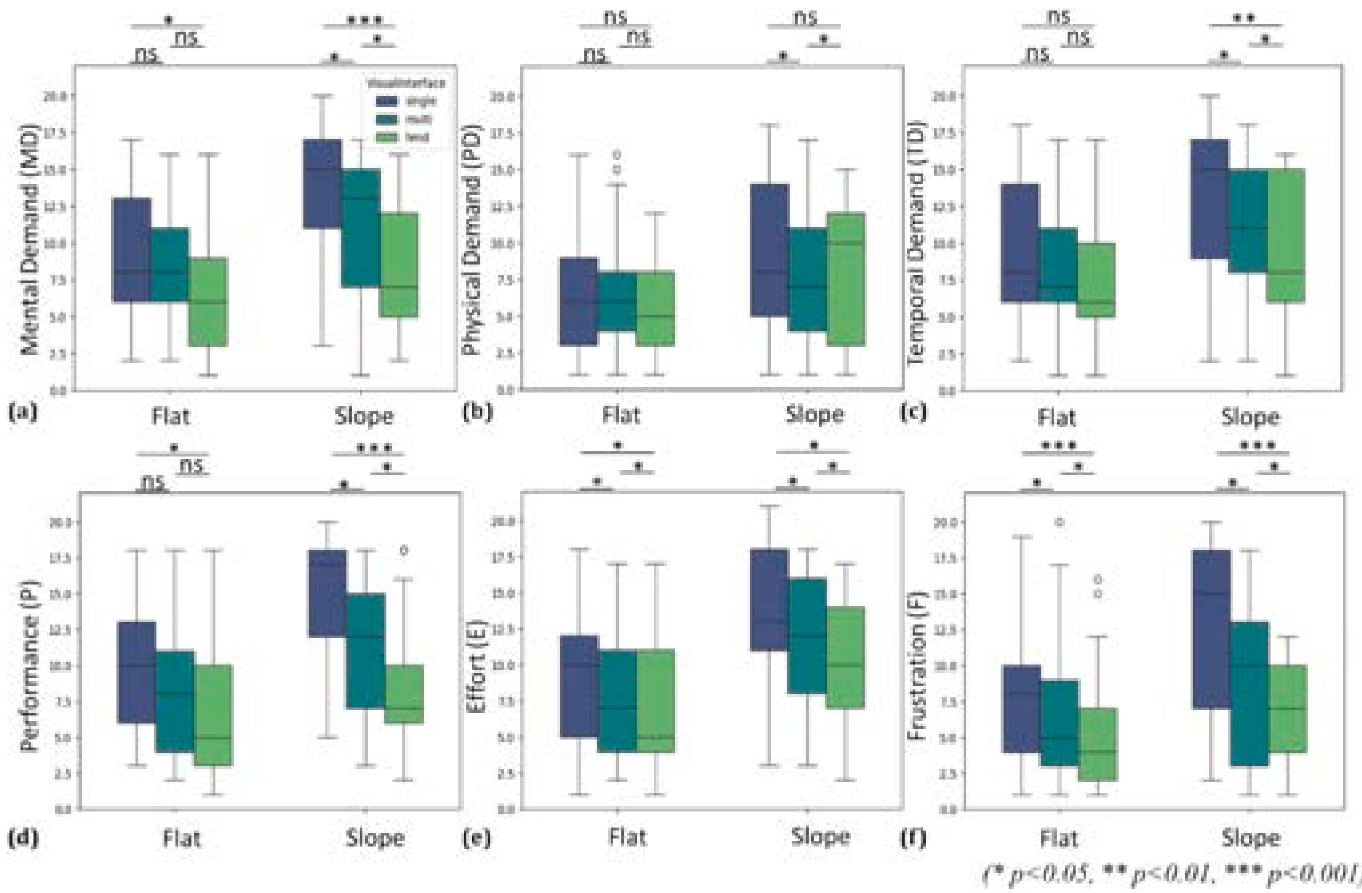

Fig. 18. NASA-TLX scores across different visual interfaces and different work environments for (a) mental demand, (b) physical demand, (c) temporal demand, (d) performance (e) effort (f) frustration.

Table 2
Presence level across different visual interface and work environments.

	Flat Terrain						Sloping Terrain					
	2D				3D		2D				3D	
	Single Viewpoint		Multiple Viewpoint		HMD		Single Viewpoint		Multiple Viewpoint		HMD	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Presence	3.56	1.15	4.04	1.19	6.22	0.61	3.15	1.04	3.95	1	5.86	0.65

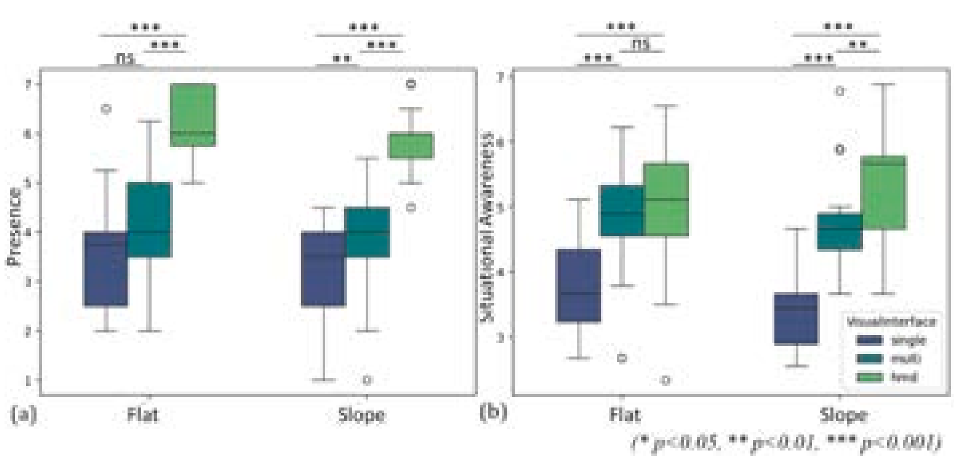

Fig. 19. Subjective experience across different environment settings: (a) presence, (b) situational awareness.

significant higher presence level compared to the single viewpoint setting ($p < 0.001$) (Fig. 19. (a)). This aligns with the findings from [50], highlighting the advantage of a 3D interface in delivering depth perception, a feature lacking in 2D interfaces. Regarding the presence level in 2D interfaces, employing multiple viewpoints would diminish the sense of presence, especially when adopting a third-person perspective [51]. Interestingly, this study reveals a contrasting trend within the context of excavator teleoperation that the level of presence appears to be elevated with the provision of multiple viewpoints in sloping terrain setting ($p < 0.01$) (Fig. 19(a)).

This study explores the role of different visual display configurations to mitigate degraded spatial awareness caused by hazardous terrain. The challenges stem from misaligned axes of different perspectives, which are heightened for teleoperators. As described in Section 4.1, the impact of visual interfaces is significant, especially in challenging work environments. According to the survey, all participants reported increased impact of visual interface differences, particularly in challenging work environments. This comprehensive demonstration emphasizes the heightened importance of how information is visually presented through the interface in challenging scenarios. Designing interfaces that adapt to varying complexities in work environments is crucial, facilitating the development of interfaces tailored to teleoperator needs to improve work efficiency and safety in challenging work contexts and guide the creation of user-centered interfaces that convey immersive information intuitively.

4.3. Perceptions regarding the use of multiple viewpoints

4.3.1. Distraction from multiple viewpoints

In the multiple viewpoint setting, participants were presented with multiple viewpoint display aimed at augmenting their performance and enhancing their work productivity by providing supplementary information. Surprisingly, contrary to its original intention, introducing additional viewpoint do not show significant effect of enhancing the collision prevention (Fig. 20 (a)). Integrating this result with completion time (Fig. 20 (b)) illustrates that in the flat terrain, where the environment doesn't present many obstacles or complexities, the benefits of reducing time or avoiding collisions by switching viewpoints are negligible. In all performance measures, the effect of provision of multiple viewpoints was insignificant in the flat terrain ($p > 0.05$).

In the sloping terrain, a challenging scenario, performance regarding time (Fig. 20 (b)) and the amount of work done (Fig. 20(c)) was enhanced when viewpoints are added, compared to the single viewpoint setting ($p < 0.01$ and $p < 0.001$ for each). However, in both scene conditions, collision reduction effect was not significant ($p > 0.05$). The participants reported attempting a continuous shift to explore different viewpoints as a mentally demanding task. This report aligns with previous study [25], which demonstrated that human can face challenges when individuals are tasked with combining two different perspectives, first-person viewpoint and third-person viewpoint. Operators often need to simultaneously monitor all viewpoints, especially when the excavator requires movement, such as rotating the excavator's body, to ensure safe and delicate operations. However, when individuals rapidly switch their gaze between two viewpoints, they may not spend enough time processing the visual information from each perspective thoroughly. From the findings of [13], higher duration time in danger-avoidance view shows a lower number of collisions. Inferring from this, rapidly switching gaze attention between these perspectives may cause individuals to overlook important signals or hazards, increasing the likelihood of collisions or errors.

Figure 21 illustrates the eye movements during the sequential work setting, especially during obstacle avoidance task. The size of circle and the line represent the fixation duration and the saccade trajectory (Fig. 21 (a&e)). In the sloping terrain, none of the metrics among fixation revisit count, saccade count, and peak velocity of average saccade shows significant differences between non-collision and collision cases. In the flat terrain, fixation revisit count of collision cases was significantly higher than that of non-collision cases ($p < 0.05$). This could demonstrate their attention distributed in multiple viewpoints could rather hinder their maneuvers, leading into the collision occurrences.

The participants reported challenges was intensified when the excavator is in motion (e.g., rotation). As the machine moves, the patterns and perspectives shift dynamically. This movement necessitates constant adjustment and recalibration of mental models to align with the changing viewpoints. Moreover, the cognitive demand becomes even more significant when considering the need to synchronize these viewpoints with the excavator movements. The operators must not only interpret the changing perspectives but also correlate them with the excavator actions to ensure accurate spatial awareness and collision avoidance. While multiple viewpoints hold the potential to enhance spatial understanding, the cognitive demands of managing such multiple viewpoints, especially in dynamic and complex work environments like operating heavy machinery on challenging jobsites, can be substantial. The challenge lies not just in acquiring diverse perspectives but also in effectively processing and synthesizing the work-related information in real-time to inform safe and effective control decision-making via human-machine interfaces.

4.3.2. Varied user preferences in viewpoint selection

Figure 22. illustrates participants' predominant use of a display among available view options (i.e., cabin view, side view, top view) during tasks. Despite having identical viewpoint settings, individuals tend to consistently favor a particular viewpoint throughout their tasks. This inclination toward a specific viewpoint differs among users, indicating distinct individual preferences or strategies in viewpoint selection. Interestingly, the selection of the preferred viewpoints is

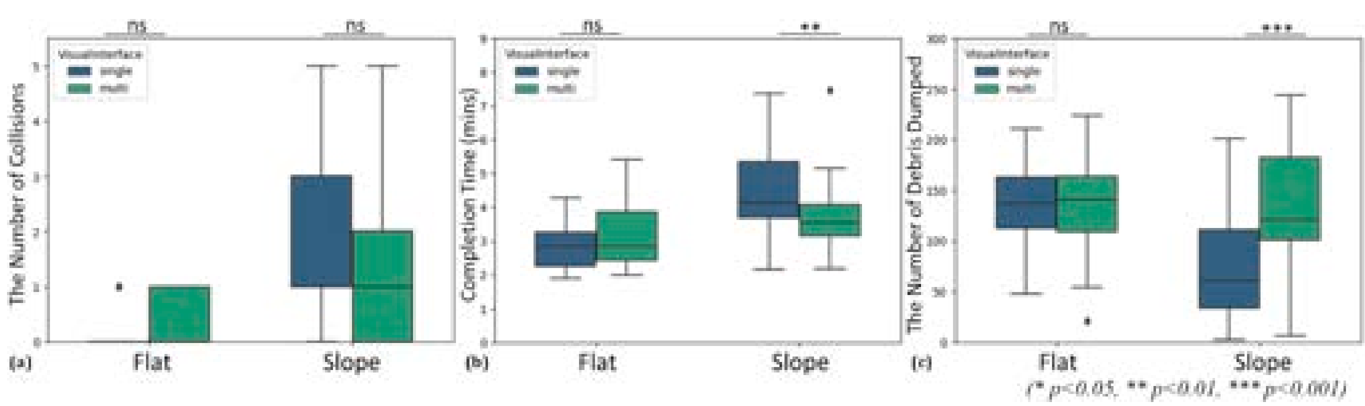

Fig. 20. Performance with the interface settings for single and multiple viewpoints in different work environments evaluated via (a) collisions, (b) completion time, and (c) work completed.

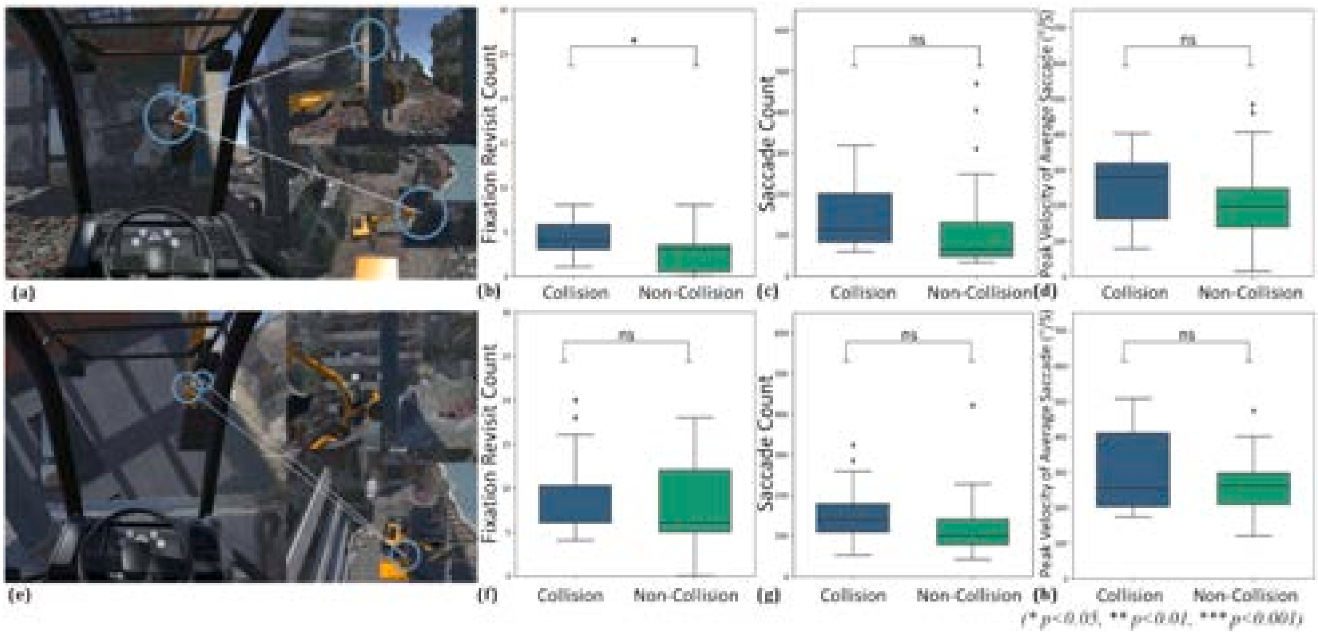

Fig. 21. Examples of eye gaze pattern in the multiple viewpoint setting in flat (a-d) and sloping terrain (e-h): (a & e) eye gaze pattern during obstacle avoidance task, (b & f) fixation revisit count (c & g) saccade count (d & h) peak velocity of the average saccade.

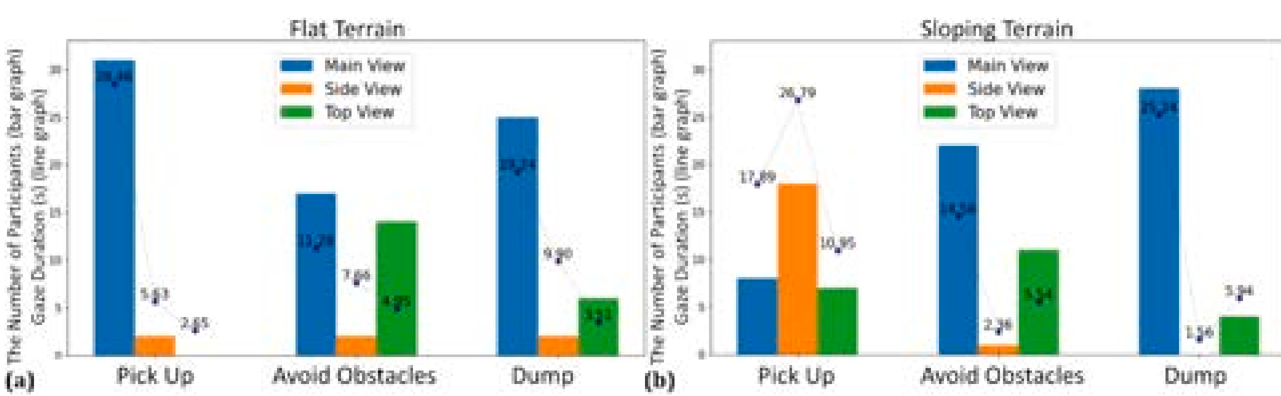

Fig. 22. The number of participants selecting each viewpoint as the main view and the average gaze duration time for each viewpoint: (a) flat terrain (b) sloping terrain.

influenced by their task currently working on. Adding on to the task, the site characteristics including visibility condition such as occlusion and terrain itself, could also potentially shape participants' choices, implying that different work environments might lead operators to favor specific viewpoints over others. This variance in preference could be attributed to work-related challenges or visual cues presented by distinct conditions of work environments. The observation highlights the potential impact of work environment-related factors, particularly terrain type, on users' primary viewpoint selections, hinting at how different terrains would prompt a more salient preference for specific viewpoints.

The variability in preferred viewpoints among users indicates distinct individual preferences or strategies in viewpoint selection. Interestingly, the participants tend to employ a wide array of viewpoints during their tasks, opting to continually compare information rather than relying solely on one viewpoint. This behavior is inferred from Fig. 22, which shows the average fixation duration at different viewpoints through a dot line. This diversity suggests that each operator develops their unique inclination or strategy for selecting a primary perspective. This could be originated from personal preferences based on their own comfort levels, past experiences, or individual strategies of cognitive processing. In challenging jobsite environments, where the target object is occluded from the direct view (i.e., from the cabin perspective), teleoperators lean more toward alternative viewpoints, like the side or top view. These viewpoints offer crucial cues regarding the location and depth of the target object. The average gaze duration per viewpoint during each task is illustrated in Fig. 22. (as a line graph). Interestingly, despite relying on alternative views due to the occluded visibility condition from the cabin view, the participants consistently maintain the use of the cabin view (Fig. 22 (b)). Through the survey interviews, the participants reported utilizing the cabin view because it is familiar to their on-site experiences, giving them a sense of physically sitting in that location. This suggests that the cabin view might serve as the primary or default reference point for teleoperators, even when the work area is somewhat occluded from the cabin view.

The results could demonstrate the potential impact of work environment-related factors, such as sloping terrain with restricted movements of machine, on operator's primary viewpoint choices in teleoperation. This insight would inform the interface system design, suggesting the importance of adaptability or customization based on diverse work environments or tasks to accommodate their preferred viewpoints. Such adaptability could significantly enhance efficiency and user experience.

4.3.3. Dynamic adaptation of teleoperators and viewing patterns across consecutive trials with multiple viewpoint settings

This experiment consists of two sets of sequential tasks including pickup, avoiding obstacles, and dumping. Fig. 23 illustrates the number of revisits to the main view per trial during each task. The main view is determined as the viewpoint where the participant spends the longest duration of fixation, investing the most time as described in Section 3.1.2. Depicted in Fig. 23, there is a reduction in the frequency of gaze revisits to alternative viewpoints observed during the second trial compared to the initial trial. This implies an adaptation process among teleoperators, wherein they progressively adjust their viewing patterns to the work environment while gaining experience with the given visual interface. Their adaptation indicated by reduced eye gaze revisit, seems different by the work environment condition. Its pattern is salient in a sloping terrain, in pickup task ($p < 0.01$) and obstacle avoidance task ($p < 0.01$) while there's no significant reduction of eye gaze revisits in dumping task ($p > 0.05$). Through the survey interview, it was reported that the most challenging part of the experiment was pickup and obstacle avoidance task in a sloping terrain. In such tasks, participants scan the multiple viewpoints to grasp their work at their first trial, and they are quite used to how to utilize the information provided by each viewpoint, leading to significant reduction in their revisits. Interestingly, in flat terrain, there is no significant reduction in eye gaze revisits ($p > 0.05$) in their second trial, indicating that their adaptation is not significantly salient in this work condition. The participants reported feeling considerably distracted by multiple viewpoints, especially in flat terrain. Comprehensively, reflecting on the findings in Section 1, it remains crucial to acknowledge that an abundance of available viewpoints can often cause distraction, possibly leading to counterproductive results like collisions. This highlights the importance of balancing viewpoint availability to ensure enhanced task performance.

4.4. Enhanced performance in high presence level group in challenging scenarios

As illustrated in Fig. 24., based on the presence level with the average of 4.46, groups can be divided into two: one with low-presence level (lower than 4.46) and the other with high-presence level (higher than 4.46). Comparing the performance metrics regarding the amount of work done and completion time, the group with low-presence perception shows particularly lower performance (higher completion time and lower amount of work done) in the sloping terrains ($p < 0.01$ for both) compared to the group with high-presence perception. This emphasizes the crucial role of perceived presence in mitigating collisions in challenging work environments. Interestingly, this correlation loses its significance in a flat terrain ($p > 0.05$). This disparity in findings suggests that presence can be one of the factors related to work performance, and this impact can be dependent on the work environments in which teleoperators operate for. Based on the finding from previous studies that presence could benefit or determine performance [50,52], this result could support that presence level can be one of the factors related to task performance, still this effect could be determined by the work environments. In challenging environments where teleoperators' immersive engagement is required, how great they perceive in physical space could affect their performance. Essentially, higher levels of presence perception engender a sense of immersion, fostering increased engagement with the task at hand, ultimately boosting teleoperators' productivity.

4.5. Simulation sickness with increased environment scanning in different work environment

After the completion of the task with the HMD setting, some participants suffered simulation sickness (Table 3). Overall, the participants reported they experienced higher likelihood of simulation sickness symptoms specifically when navigating challenging environments like a sloping terrain during collision avoidance tasks that require rotation. Among 16 symptoms outlined in [53], seven symptoms (Increased salivation, Sweating, Nausea, Dizzy (eyes closed), Vertigo, Stomach Awareness, Burping) were disregarded from this study since all participants never exhibit any of these. Regarding the remaining nine symptoms in Table 2, the average level of simulation sickness among 33 participants infers the impact of work environments on simulation sickness. Despite our experiments being shortened to <7 min to mitigate simulation sickness [47], some participants still exhibited heightened simulation sickness symptoms, especially in the sloping terrain.

Figure 25 (a) illustrates the average eye movement patterns observed during collision avoidance task in the sloping terrain. The scanning ratio of work environments represents the ratio of the participants' eye fixation duration on the work environment relative to their overall duration engaged in the collision avoidance tasks. In Fig. 25 (b), examples are presented, highlighting participants who have their eye fixation on the excavator (top), contrasted with the others exploring and scanning the work environment during the collision avoidance tasks (bottom).

The participants who experienced simulation sickness shows a significantly different environment scanning ratio compared to those who did not ($p < 0.001$). Interestingly, this phenomenon is only shown in the sloping terrain ($p > 0.05$ for the flat terrain). In such challenging scenarios, the discrepancy arises from the misalignment between the teleoperator's axis of perception, typically centered on the cabin view, and the actual configuration of the surrounding work environment (which has a slope), especially regarding obstacles. As illustrated in Fig. 26 (a), during rotation, the movement of the excavator is parallel to

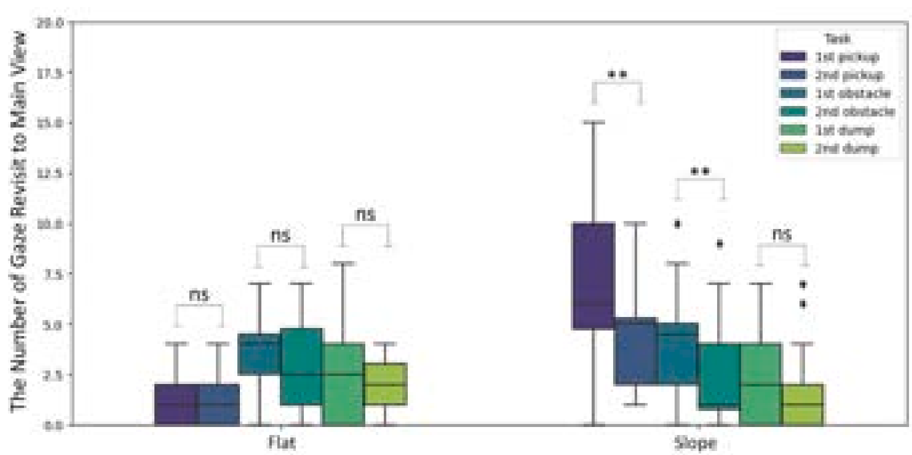

Fig. 23. The number of gaze revisit to main view in different work environments.

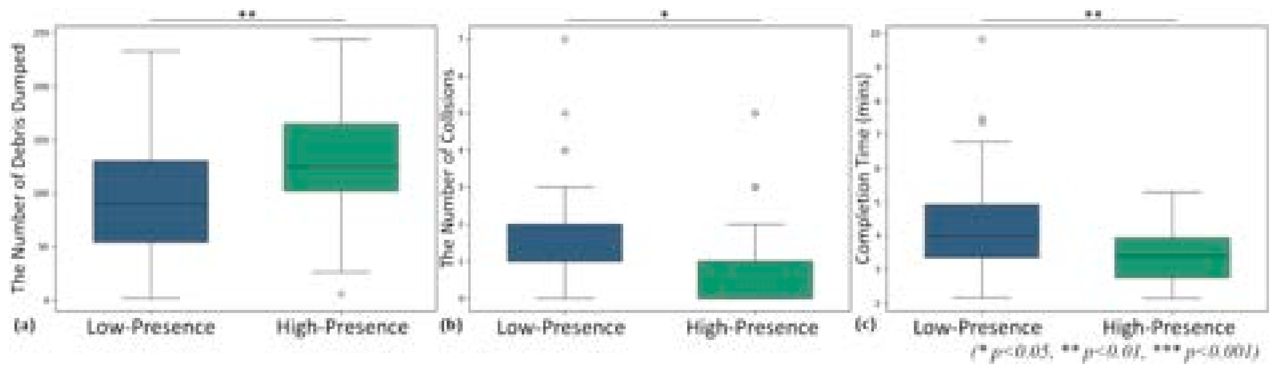

Fig. 24. Work performance across low-presence group and high-presence group for the experiment in sloping terrain: (a) work completed, (b) collision occurrences, and (c) completion time.

Table 3
Simulation sickness results after experiments with the HMD setting.

	Flat Terrain		Sloping Terrain	
	Mean	SD	Mean	SD
General Discomfort	0.53	0.8	0.97	1.09
Stomach Awareness	0.03	0.18	0.53	0.8
Difficulty Concentrating	0.06	0.25	0.69	0.82
Dizziness with Eyes Open	0.16	0.45	0.81	0.93
Fullness of Head	0	0	0.88	1.01
Difficulty Focusing	0.03	0.18	0.66	0.87
Blurred Vision	0.28	0.52	0.41	0.8
Fatigue	0.34	0.34	0.69	0.86
Eye Strain	0.13	0.67	0.72	0.85

0: None, 1: Slightly, 2: Moderate, 3: Severe

the teleoperator's manipulation, while in the sloping terrain, the rotation of the surrounding environment doesn't move parallel to the manipulation direction, rather has a difference of an angle referring to the slope of the terrain. From the teleoperator's viewpoint, as they execute a 'swing right' manipulation (as Fig. 26 (b)), the excavator's body parts rotate in parallel, while the surrounding environment rotates in alignment with the angle of terrain slope (Fig. 26 (c)). This disparity between inclination of surrounding environments with excavation body parts results in a divergence between the teleoperator's visual perspective and the actual spatial arrangement of objects or elements on the terrain. These results could give us significant implications for the design of visual interfaces tailored for teleoperators in construction scenarios, considering crucial human factors. First, teleoperators' simulation sickness could be highly related to work environment-related factors. This could highlight the necessity of site-specific interface design to mitigate simulation sickness. Interfaces designed to accommodate the variations in a sloping terrain might encounter misalignment between the global axis and local axis on the teleoperator, demanding adaptable viewpoint adjustments to mitigate the misalignment given from different slope of terrains. Addressing this misalignment requires interface designs that bridge the gap between the teleoperator's viewpoint and the true spatial arrangement of obstacles, possibly through enhanced depth perception cues or adaptive viewpoints. Second, the interface should aim to balance scanning of work environments with the guidance of their fixations, by providing cues that could aid in targeted

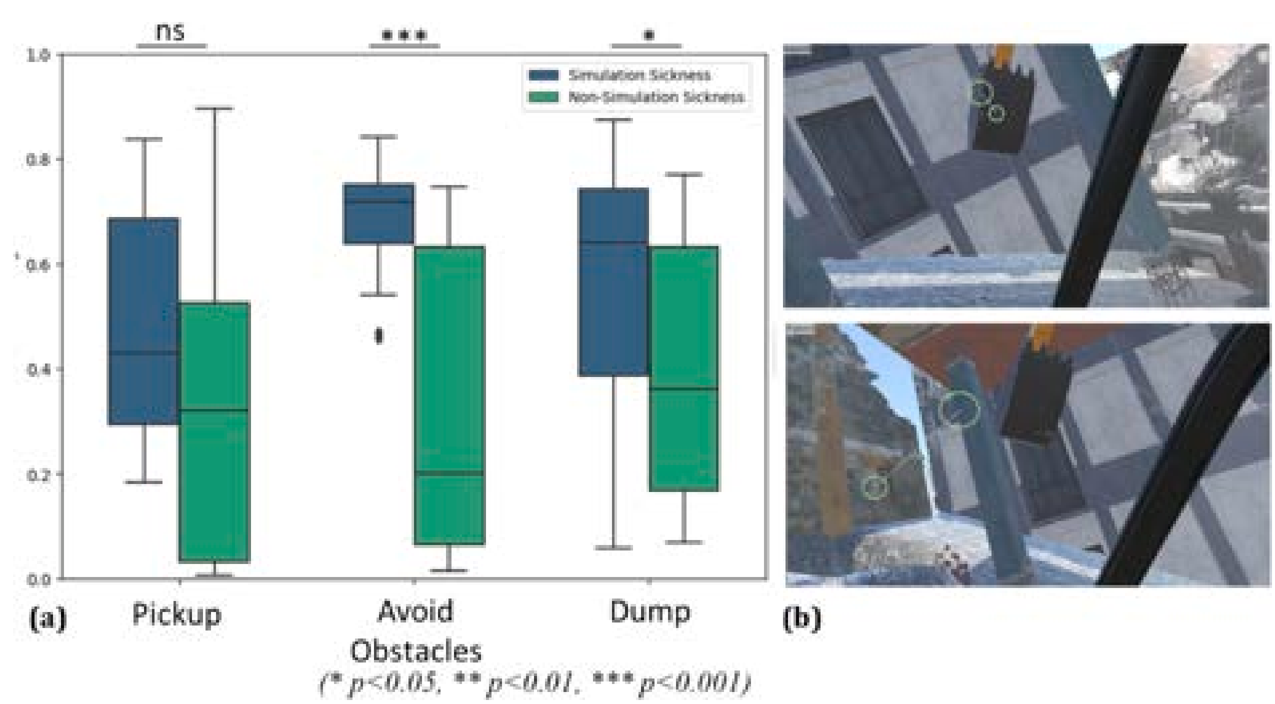

Fig. 25. Eye gaze pattern during the collision avoidance task in challenging environments: (a) environment scanning ratio for the participants experiencing simulation sickness and those who didn't (b) an example of eye gaze pattern.

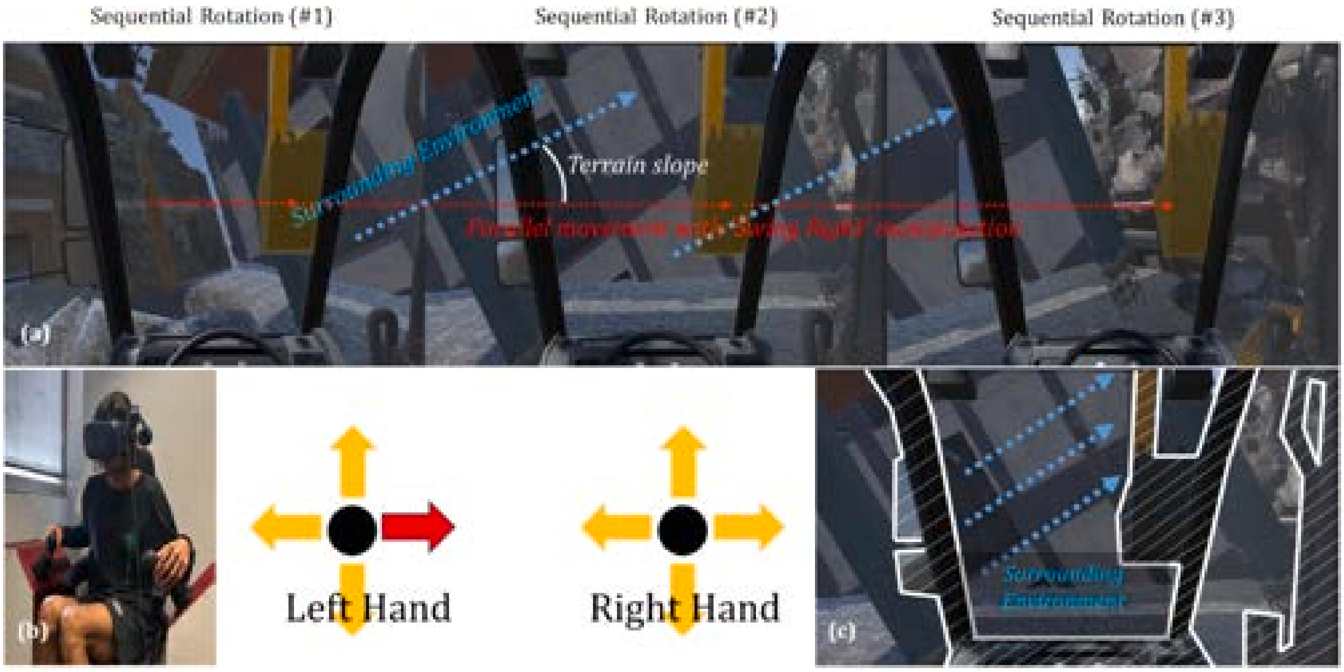

Fig. 26. Misalignment between teleoperator's perspective and terrain slope, (a) Misalignment due to terrain slope during heavy machinery rotation by teleoperators, (b) Teleoperator's rotation manipulation, (c) Area surrounding the excavator from cabin view (teleoperator's perspective).

visual exploration without inducing discomfort.

5. Conclusions

This paper analyzes the divergent impact of visual interface in teleoperation, a crucial aspect for teleoperators, within challenging work environments. Examining human factors, we analyze how each visual interface affects teleoperators and work performance. To address spatial awareness challenges in teleoperation, especially in demanding environments like sloping demolition sites, interface design should integrate depth-related information, such as additional viewpoints or immersive 3D displays, to aid the operation of heavy machinery. This paper focuses on the divergent impact of this provision during teleoperation especially in challenging jobsites, when additional challenges arise in work environments (e.g., sloping terrain and limited visibility). We investigate distractions and diverse user preferences in multiple viewpoint settings across work environments. While fixed multiple viewpoints can offer benefits, as teleoperators may become accustomed to using each viewpoint by developing their own strategies, they also pose challenges, such as data overload problems. Additionally, our findings reveal that immersive displays with HMD can induce simulation sickness, especially during excessive environment scanning in challenging workplace settings. Furthermore, this paper underscores the distinctive patterns and dependencies of visual interfaces in challenging work environments. These findings provide valuable insights into the development of intuitive and informative interfaces tailored to teleoperator needs. The findings of this paper suggest when designing the interface, it is critical to consider how to convey information intuitively to teleoperators and rigorously evaluate its impact on their performance. Still, how to provide the spatial data effectively to teleoperators needs to be further examined, which is currently being explored as part of our ongoing research.

This paper exclusively focuses on visual modality. However, its impact on teleoperators may vary in teleoperations with multi-modality conditions. Incorporating additional sensory inputs should be considered in future research to understand how both the visual interface and operational interface as a whole affect teleoperators. Meanwhile, this paper was conducted in challenging jobsite environments characterized by sloping terrain and closely located obstacles, necessitating delicate operations under limited visibility conditions. However, real-world construction sites may present additional components that compromise the teleoperator's spatial awareness, which also needs to be considered in future research to generalize the findings to real-world scenarios. Additionally, this paper was conducted with undergraduate and graduate students. To generalize the findings to a wider demographic range, further research is needed. Finally, exploring the impact of multiple viewpoints in HMD can be analyzed in further study.

CRediT authorship contribution statement

Yeon Chae: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Samraat Gupta:** Visualization, Methodology, Investigation, Data curation. **Youngjib Ham:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

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

[1] Y. Hiramatsu, T. Aono, M. Nishio, Disaster restoration work for the eruption of Mt Usuzan using an unmanned construction system, Adv. Robot. 16 (2002) 505–508, https://doi.org/10.1163/156855302320535836.

[2] J.S. Lee, Y. Ham, 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, in: Proceedings of the 1st Future of Construction Workshop at the International Conference on Robotics and Automation (ICRA 2022), International Association for Automation and Robotics in Construction (IAARC), 2022, https://doi.org/10.22260/ICRA2022/0017.

[3] J. Rasmussen, Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models, IEEE Trans. Syst. Man Cybern. SMC-13 (1983) 257–266, https://doi.org/10.1109/TSMC.1983.6313160.

[4] T. Zhou, P. Xia, Y. Ye, J. Du, Embodied robot teleoperation based on high-Fidelity visual-haptic simulator: pipe-fitting example, J. Constr. Eng. Manag. 149 (2023) 04023129, https://doi.org/10.1061/JCEMD4.COENG-13916.

[5] K. Shigematsu, T. Tsubouchi, S. Sarata, Tip-over prevention system based on motion prediction for teleoperated excavator, Adv. Robot. 35 (2021) 1438–1449, https://doi.org/10.1080/01691864.2021.2004223.

[6] G. Lunghi, R.M. Prades, M.D. Castro, An Advanced, Adaptive and Multimodal Graphical User Interface for Human-robot Teleoperation in Radioactive Scenarios, in: Proceedings of the 13th International Conference on Informatics in Control, Automation and Robotics, SCITEPRESS - Science and and Technology Publications, Lisbon, Portugal, 2016, pp. 224–231, https://doi.org/10.5220/0005971402240231.

[7] C.W. Nielsen, M.A. Goodrich, R.W. Ricks, Ecological interfaces for improving Mobile robot teleoperation, IEEE Trans. Robot. 23 (2007) 927–941, https://doi.org/10.1109/TRO.2007.907479.

[8] M. Kamezaki, J. Yang, H. Iwata, S. Sugano, Visibility enhancement using autonomous multicamera controls with situational role assignment for Teleoperated work machines, J. Field Robot. 33 (2016) 802–824, https://doi.org/10.1002/rob.21580.

[9] D. Liu, J. Kim, Y. Ham, Multi-user immersive environment for excavator teleoperation in construction, Autom. Constr. 156 (2023) 105143, https://doi.org/10.1016/j.autcon.2023.105143.

[10] A. Naceri, D. Mazzanti, J. Bimbo, D. Prattichizzo, D.G. Caldwell, L.S. Mattos, N. Deshpande, Towards a Virtual Reality Interface for Remote Robotic Teleoperation, in: 2019 19th International Conference on Advanced Robotics (ICAR), 2019, pp. 284–289, https://doi.org/10.1109/ICAR46387.2019.8981649.

[11] X. Wang, P.S. Dunston, Mixed Reality — Enhanced Operator Interface for Teleoperation Systems in Unstructured Environment, 2012, pp. 1–8, https://doi.org/10.1061/40830(188)93.

[12] H. Hedayati, M. Walker, D. Szafir, Improving Collocated Robot Teleoperation with Augmented Reality, in: Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, Association for Computing Machinery, New York, NY, USA, 2018, pp. 78–86, https://doi.org/10.1145/3171221.3171251.

[13] M. Kamezaki, J. Yang, R. Sato, H. Iwata, S. Sugano, A situational understanding enhancer based on augmented visual prompts for teleoperation using a multi-monitor system, Autom. Constr. 131 (2021) 103893, https://doi.org/10.1016/j.autcon.2021.103893.

[14] H. Song, T. Kim, J. Kim, D. Ahn, Y. Kang, Effectiveness of VR crane training with head-mounted display: double mediation of presence and perceived usefulness, Autom. Constr. 122 (2021) 103506, https://doi.org/10.1016/j.autcon.2020.103506.

[15] P. Carayon, A.S. Hundt, B. Karsh, A.P. Gurses, C.J. Alvarado, M. Smith, P. F. Brennan, Work system design for patient safety: the SEIPS model, Qual. Saf. Health Care 15 (2006) i50–i58, https://doi.org/10.1136/qshc.2005.015842.

[16] S. Lee Pazuchanics, The effects of camera perspective and field of view on performance in Teleoperated navigation, Proc. Human Fact. Ergon. Soc. Ann. Meet. 50 (2006) 1528–1532, https://doi.org/10.1177/154193120605001603.

[17] F. Ferland, F. Pomerleau, C.T. Le Dinh, F. Michaud, Egocentric and exocentric teleoperation interface using real-time, 3D video projection, in: Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction, Association for Computing Machinery, New York, NY, USA, 2009, pp. 37–44, https://doi.org/10.1145/1514095.1514105.

[18] S. Okishiba, R. Fukui, M. Takagi, H. Azumi, S. Warisawa, R. Togashi, H. Kitaoka, T. Ooi, Tablet interface for direct vision teleoperation of an excavator for urban construction work, Autom. Constr. 102 (2019) 17–26, https://doi.org/10.1016/j.autcon.2019.02.003.

[19] R. Sato, M. Kamezaki, M. Yamada, T. Hashimoto, S. Sugano, H. Iwata, Experimental investigation of optimum and allowable range of side views for Teleoperated digging and release works by using actual construction machinery, in: 2019 IEEE/SICE International Symposium on System Integration (SII), 2019, pp. 788–793, https://doi.org/10.1109/SII.2019.8700330.

[20] H. Son, H. Seong, H. Choi, C. Kim, Real-time vision-based warning system for prevention of collisions between workers and heavy equipment, J. Comput. Civ. Eng. 33 (2019) 04019029, https://doi.org/10.1061/(ASCE)CP.1943-5487.0000845.

[21] S. Livatino, D.C. Guastella, G. Muscato, V. Rinaldi, L. Cantelli, C.D. Melita, A. Caniglia, R. Mazza, G. Padula, Intuitive robot teleoperation through multi-sensor informed mixed reality visual aids, IEEE Access 9 (2021) 25795–25808, https://doi.org/10.1109/ACCESS.2021.3057808.

[22] H.A. Yanco, J.L. Drury, J. Scholtz, Beyond usability evaluation: analysis of human-robot interaction at a major robotics competition, Human–Comp. Interact. 19 (2004) 117–149, https://doi.org/10.1080/07370024.2004.9667342.

[23] S. Opiyo, J. Zhou, E. Mwangi, W. Kai, I. Sunusi, A review on teleoperation of Mobile ground robots: architecture and situation awareness, Int. J. Control. Autom. Syst. 19 (2021) 1384–1407, https://doi.org/10.1007/s12555-019-0999-z.

[24] M. Moniruzzaman, A. Rassau, D. Chai, S.M.S. Islam, Teleoperation methods and enhancement techniques for mobile robots: a comprehensive survey, Robot. Auton. Syst. 150 (2022) 103973, https://doi.org/10.1016/j.robot.2021.103973.

[25] H.J. Ferguson, I. Apperly, J.E. Cane, Eye tracking reveals the cost of switching between self and other perspectives in a visual perspective-taking task, Q. J. Exp. Psychol. 70 (2017) 1646–1660, https://doi.org/10.1080/17470218.2016.1199716.

[26] J. Jarmasz, C.M. Herdman, K.R. Johannsdottir, Object-based attention and cognitive tunneling, J. Exp. Psychol. Appl. 11 (2005) 3–12, https://doi.org/10.1037/1076-898X.11.1.3.

[27] R. Sato, M. Kamezaki, S. Niuchi, S. Sugano, H. Iwata, Cognitive untunneling multi-view system for teleoperators of heavy machines based on visual momentum and saliency, Autom. Constr. 110 (2020) 103047, https://doi.org/10.1016/j.autcon.2019.103047.

[28] B. Han, F. Leite, Measuring the impact of immersive virtual reality on construction design review applications: head-mounted display versus desktop monitor, J. Constr. Eng. Manag. 147 (2021) 04021042, https://doi.org/10.1061/(ASCE)CO.1943-7862.0002056.

[29] A. Hosseini, M. Lienkamp, Enhancing telepresence during the teleoperation of road vehicles using HMD-based mixed reality, in: 2016 IEEE Intelligent Vehicles Symposium (IV), 2016, pp. 1366–1373, https://doi.org/10.1109/IVS.2016.7535568.

[30] H. Martins, I. Oakley, R. Ventura, Design and evaluation of a head-mounted display for immersive 3D teleoperation of field robots, Robotica 33 (2015) 2166–2185, https://doi.org/10.1017/S026357471400126X.

[31] Y. Luo, J. Wang, R. Shi, H.-N. Liang, S. Luo, In-device feedback in immersive head-mounted displays for distance perception during teleoperation of unmanned ground vehicles, IEEE Transact. Haptics 15 (2022) 79–84, https://doi.org/10.1109/TOH.2021.3138590.

[32] P. Xu, G. Yan, H. Fu, T. Igarashi, C.-L. Tai, H. Huang, Global beautification of 2D and 3D layouts with interactive ambiguity resolution, IEEE Trans. Vis. Comput. Graph. 27 (2021) 2355–2368, https://doi.org/10.1109/TVCG.2019.2954321.

[33] M. Candeloro, E. Valle, M.R. Miyazaki, R. Skjetne, M. Ludvigsen, A.J. Sørensen, HMD as a new tool for telepresence in underwater operations and closed-loop control of ROVs, in: OCEANS 2015 - MTS/IEEE Washington, 2015, pp. 1–8, https://doi.org/10.23919/OCEANS.2015.7404466.

[34] P. Adami, P.B. Rodrigues, P.J. Woods, B. Becerik-Gerber, L. Soibelman, Y. Copur-Gencturk, G. Lucas, Effectiveness of VR-based training on improving construction workers' knowledge, skills, and safety behavior in robotic teleoperation, Adv. Eng. Inform. 50 (2021) 101431, https://doi.org/10.1016/j.aei.2021.101431.

[35] F. Nakamura, M. Murakami, K. Suzuki, M. Fukuoka, K. Masai, M. Sugimoto, Analyzing the effect of diverse gaze and head direction on facial expression recognition with photo-reflective sensors embedded in a head-mounted display, IEEE Trans. Vis. Comput. Graph. 29 (2023) 4124–4139, https://doi.org/10.1109/TVCG.2022.3179766.

[36] Y. Ohmura, S. Yano, J. Katsuhira, M. Migita, A. Yozu, T. Kondo, Inclination of standing posture due to the presentation of tilted view through an immersive head-mounted display, J. Phys. Ther. Sci. 29 (2017) 228–231, https://doi.org/10.1589/jpts.29.228.

[37] K.P. Kisi, O. Snosi, V. Sharma, Creating a Construction Safety VR Game with Decision-Making Scenarios for an Interactive Learning Experience, 2024, pp. 452–460, https://doi.org/10.1061/9780784485248.054.

[38] A.L. Gardony, R.W. Lindeman, T.T. Brunyé, Eye-tracking for human-centered mixed reality: promises and challenges, in: Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR), SPIE, 2020, pp. 230–247, https://doi.org/10.1117/12.2542699.

[39] Q. Zhu, J. Du, Y. Shi, P. Wei, Neurobehavioral assessment of force feedback simulation in industrial robotic teleoperation, Autom. Constr. 126 (2021) 103674, https://doi.org/10.1016/j.autcon.2021.103674.

[40] Y. Liu, M. Habibnezhad, H. Jebelli, Brain-computer interface for hands-free teleoperation of construction robots, Autom. Constr. 123 (2021) 103523, https://doi.org/10.1016/j.autcon.2020.103523.

[41] X. Zhao, X. Chen, Y. He, H. Cao, T. Chen, Varying speed rate controller for human–robot teleoperation based on muscle electrical signals, IEEE Access 7 (2019) 143563–143572, https://doi.org/10.1109/ACCESS.2019.2944794.

[42] D. Kent, C. Saldanha, S. Chernova, Leveraging depth data in remote robot teleoperation interfaces for general object manipulation, Intern. J. Robot. Res. 39 (2020) 39–53, https://doi.org/10.1177/0278364919888565.

[43] K.D. Katyal, C.Y. Brown, S.A. Hechtman, M.P. Para, T.G. McGee, K.C. Wolfe, R. J. Murphy, M.D.M. Kutzer, E.W. Tunstel, M.P. McLoughlin, M.S. Johannes, Approaches to robotic teleoperation in a disaster scenario: from supervised autonomy to direct control, in: 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2014, pp. 1874–1881, https://doi.org/10.1109/IROS.2014.6942809.

[44] 1926 Subpart P App B - Sloping and Benching Occupational Safety and Health Administration, (n.d.). https://www.osha.gov/laws-regs/regulations/standardnumber/1926/1926SubpartPAppB (accessed December 26, 2023).

[45] Y. Chae, S. Gupta, Y. Ham, Design and Evaluation of Human-Centered Visualization Interfaces in Construction Teleoperation, 2024, pp. 109–118, https://doi.org/10.1061/9780784485262.012.

[46] X. Wei, Z. Ran, X. Lu, DCPB: Deformable Convolution Based on the Poincare Ball for Top-View Fisheye Cameras, 2023, pp. 13308–13317. https://openaccess.thecvf.com/content/ICCV2023/html/Wei_DCPB_Deformable_Convolution_Based_on_the_Poincare_Ball_for_Top-view_ICCV_2023_paper.html (accessed April 28, 2024).
[47] J. Wang, H.-N. Liang, D. Monteiro, W. Xu, J. Xiao, Real-time prediction of simulator sickness in virtual reality games, IEEE Trans. Games 15 (2023) 252–261, https://doi.org/10.1109/TG.2022.3178539.
[48] R.M. Taylor, Situational Awareness Rating Technique (Sart): The Development of a Tool for Aircrew Systems Design, Situational Awareness, Routledge, 2011.
[49] R.S. Kennedy, N.E. Lane, K.S. Berbaum, M.G. Lilienthal, Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness, Int. J. Aviat. Psychol. 3 (1993) 203–220, https://doi.org/10.1207/s15327108ijap0303_3.
[50] E.B. Nash, G.W. Edwards, J.A. Thompson, W. Barfield, A review of presence and performance in virtual environments, Intern. J. Human–Comp. Interact. 12 (2000) 1–41, https://doi.org/10.1207/S15327590IJHC1201_1.
[51] E.L. Schuurink, A. Toet, Effects of third person perspective on affective appraisal and engagement: findings from SECOND LIFE, Simul. Gaming 41 (2010) 724–742, https://doi.org/10.1177/1046878110365515.
[52] W. Barfield, T.A. Furness, Virtual Environments and Advanced Interface Design, Oxford University Press, USA, 1995 (ISBN 978-0-3879-7772-0).).
[53] R.S. Kennedy, N.E. Lane, K.S. Berbaum, M.G. Lilienthal, Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness, Int. J. Aviat. Psychol. 3 (1993) 203–220, https://doi.org/10.1207/s15327108ijap0303_3.