electronics

Article

Virtual Reality Interface Evaluation for Earthwork Teleoperation

Léa Saunier [1,*,†], Nicolas Hoffmann [1,2,†], Marius Preda [1] and Catalin Fetita [1]

[1] SAMOVAR, Telecom Sud-Paris, Institut Polytechnique de Paris, 91000 Evry, France; nicolas.hoffmann@telecom-sudparis.eu (N.H.); marius.preda@telecom-sudparis.eu (M.P.); catalin.fetita@telecom-sudparis.eu (C.F.)
[2] HERACLES Robotics, 78440 Gargenville, France
* Correspondence: saunier.lea@telecom-sudparis.eu
† These authors contributed equally to this work.

Abstract: Automation and robotics are destined to play a critical role in the Industry 4.0 revolution, as illustrated by the emergence of autonomous machinery in earthwork operations. Despite rapid progress, autonomous agents will always require human supervision to instruct their mission and to guaranty safety when unexpected problems arise. Traditional human supervision requires an operator to physically enter each machine at risk and manually take control. This approach is time-consuming and requires highly qualified personnel capable of operating various machines. This process can be hastened and simplified by means of teleoperated supervision, which itself requires the appropriate interface. In this paper we evaluate a virtual reality (VR)-based interface using hybrid interactions and an immersive digital-twin compared to a real-life control. We compare these interfaces through control tasks performed by expert and non-expert operators, analyzing time and precision, as well as user feedback. The preliminary results show that the VR interface brings equivalent and satisfactory performances for experts and improves the efficiency of apprentices. Therefore, not only does everyone performs well in the virtual environment, but also the training time can be shortened significantly as non-experts can perform similarly under the same conditions.

Keywords: virtual reality; teleoperation; digital twin; industry 4.0; human–machine interactions; robotics; excavation

check for updates

Citation: Saunier, L.; Hoffmann, N.; Preda, M.; Fetita, C. Virtual Reality Interface Evaluation for Earthwork Teleoperation. *Electronics* **2023**, *12*, 4151. https://doi.org/10.3390/electronics12194151

Academic Editors: Grzegorz Zwoliński and Dorota Kamińska

Received: 15 September 2023
Revised: 25 September 2023
Accepted: 3 October 2023
Published: 6 October 2023

1. Introduction

Automation and teleoperation are rapidly changing the face of earthwork and construction industries. While automation aims for the complete removal of the human-in-the-loop operator, teleoperation keeps the operator but shifts the control type to a remote one, removing the need for a human on-site operator. Advances in both of those technologies allowed for greater efficiency, accuracy, and safety in the operation of heavy machinery. The use of autonomous and semi-autonomous earthwork machines has seen a significant increase in recent years due to the benefits they offer, including reduced labor costs, increased productivity, and improved safety [1]. However, these machines are not immune to unforeseen events and technical failures, which could lead to potential safety risks, delays, and additional costs. To address these concerns, teleoperation systems have been developed to allow operators to remotely control the machines in real-time from a safe and comfortable location [2]. This approach has gained significant attention in recent years and is expected to become more prevalent in the near future. Automation and teleoperation are already well developed in industry, especially for factory work [3], but they are limited by laws in the earthwork field. Furthermore, virtual reality (VR) technologies offer new, more extensive interfaces in the world of teleoperation by providing operators with a sense of presence and spatial awareness [4]. VR, through specific devices like headsets, provides new immersive 3D environments with new ways of visualising and interacting with virtual objects. With our project, we aim to improve on those features and use the full extent of the VR environment to enhance teleoperators' experience and facilitate precise machine control.

Our project features a VR interface for the remote operation of these machines. The primary objective is to provide a fail-safe solution in situations where unforeseen events occur during automated operations. Our focus is on creating a universal and user-friendly interface for remote earthwork that minimizes the need for extensive training and complex skill requirements.VR training for teleoperation has already proven its efficiency [5], as well as for direct control [6]. By enabling operators to assume the direct control of the machines from a remote location, we eliminate the need for constant on-site presence. In this paper, we present our primary investigation on comparing the performance of operators using our VR interface and direct control of the real machine, with operators having varying levels of experience.

Other studies also experimented with various virtual environments for teleoperation for industrial purposes, earthwork being no exception [7]. Diverse interfaces and technologies were tested with positive results, like augmented-reality [8], 2D interfaces [9], digital-twins [10], and advanced environments reconstructions [11]. In our case, we chose to use a VR interface to make use of its immersiveness and space awareness, as well as a digital-twin for accurate representation and control in a given situation. Another motivation is that VR enables shifts in point of view, which is massive in excavating work, as operators have a rather limited field of view when inside their cabin. This teleoperation study acts as a bridge between direct control and full automation and is being conducted in parallel with automation work, which is related to a previous paper [12].

This primary study shows encouraging results regarding the use of a VR interface for teleoperation and as a training tool. Not only do experts and non-experts perform similarly in the VR interface, but training in the VR interface also provides better results when working on the real machine afterwards. The study also provides insight on how to optimise the current VR interface in terms of design and technical developments. Despite the advice of experts on both technologies regarding the ergonomics of the application, we also demonstrate that a hybrid form of control—with a real-machine logic but a VR/gaming button mapping—is not optimal. Finally, we found that the time and performance achieved in VR are lesser compared to the real machine due to the current system limitations—in terms of speed and lag mostly—but that VR contributes to safer operations.

2. Materials and Methods

2.1. Hardware and Software

For the real interface, the real machine (RM) used in this experiment is a fully functional Cat323 excavator—see Figure A1, Appendix A for schematics—, rigged with a control-by-wire system connected to a SEMIL-1700 and made accessible remotely with a router. It uses ROS2 [13] for supervising the machine's information and sending it remote commands. For the participants, however, this machine appears and is operated as any other excavator of the same model, without any added interface element. Both the machine and its software are provided by HERACLES Robotics, autonomous earthwork https://www.heracles-robotics.com/ (accessed on 1 September 2023)), an earthwork automation company working in partnership on this project.

For the VR interface, we use a HTC Vive Pro Eye VR head-mounted device (HMD) from the HTC Vive company (Xindian District, New Taipei City, Taïwan). It provides 6 degrees of freedom (DOFs) of control, as well as eye-tracking (ET) data for future optimisation. The VR application is developed with Unity 2020.3.1f1, and the connection between the two interfaces is achieved using the Unity Robotics Hub plugin [14]. This allows us to have a real-time digital twin of the machine with its current state and get control over it directly inside the VR HMD (see Figure 1). The controllers mapping between the virtual and real machine interfaces is described in Figure A2, Appendix B. However, the current machine is not equipped with terrain sensors, and therefore only the machine state is rendered in the 3D application. We use predictive deformable terrain to bypass this issue, using the Open Construction Simulator package [15]. An overview of this feature can be observed in Figure A3, Appendix C. Recording of the experiments data was achieved using ROSbags

(ROS information on the state of the machine over time) and video recording—with the knowledge and approval of the participants. All the data are anonymised.

(**a**) (**b**)

Figure 1. User views in the different interfaces. (**a**) VR View; (**b**) RM view.

2.2. Experimental Protocol

Our study seeks to evaluate the following properties of the VR interface:

- **Operational performance:** measured by the task's completion speed and the overall dexterity of the control. The task completion time is critical in an industrial context and must, therefore, be kept short by any tool destined to be used in such a context. Also, since novice operators have a tendency to shock and damage the machine, we propose to evaluate and monitor user dexterity by measuring the overall acceleration of the articulations of the arm;
- **Ergonomics and ease of use**: obtained through the analysis of compound motions and user feedback in the NASA-TLX surveys [16]. The NASA Task-Load-Index test lets the user rate their perceived workload on the different tasks and give invaluable first-hand insight to the evaluators. It still gives subjective inputs, however, and must be coupled to other types of data for more extensive results. Compound motions are a symbol of expertise in excavator control, and they are necessary to achieve the highest performances. They are heavily influenced by the mappings between the real and virtual environment, and thus play a big role in control design. On the other hand, the NASA-TLX form allows the user to provide direct feedback;
- **Progression and ease of learning**: measured by the completion time evolution as a candidate repeats an operation. We not only observed that performances in VR improve fast, but also that VR training translates into a performance increase on the real machine.

The expertise of the participants plays a crucial role in our analysis and is measured through a numerical assessment of their familiarity with excavator manipulation and VR tools. We evaluated our VR interface by measuring and monitoring 10 participants while they perform a set of exercises with the Real Machine (RM) and with the VR interface (see Figure 2). All participants performed all the exercises both in RM and VR, with the order of the interfaces randomised to avoid learning bias.

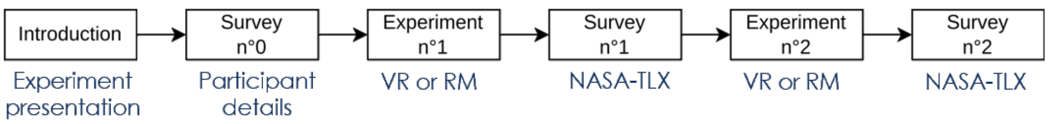


Figure 2. Experimental Protocol.

The protocol—displayed in Figure 2—starts with a short introduction to the interfaces, safety concerns, and data recording setup—video and machine motion. The participant then fulfills a first survey assessing their familiarity with excavators and their manipulation.

Once completed, the controls of a first interface are explained to the participant—RM or VR, randomly chosen—, they manipulate the machine on their own for 60 s and then complete the exercises with a first interface. Before moving onto the other interface, the participant submits their feedback in the form of the NASA-TLX survey, which has already proven fruitful for analogous studies. They also fill a user survey about themselves to get the participants profiles, and an open feedback survey for more insight. Once filled, the participants terminate the experiment by completing the same exercises with the other interface and submitting their feedback with the same forms.

2.3. Computing Speed

While a participant p executes an exercise, we measure the position, velocity, and acceleration of each articulation of the excavator, such as $Position(p, Boom)$, $Velocity(p, Boom)$, and $Acceleration(p, Boom)$. We also measure its overall completion time $CompletionTime(p)$ as well as the pause time and downtime (time used for explanations).

The speed of a participant p is measured by their overall time to complete all the exercises, excluding pauses and downtime (1).

$$OverallTime(p) = \sum_{Exercise=1}^{3} CompletionTimeExercise(p) \tag{1}$$

2.4. Computing Clumsiness and Dexterity

Novice operators tend to be clumsy, slow or brutal with the machine, damaging it quickly. This concern is omnipotent in the conception of the teleoperation controls and we aim to measure the consequences of this by comparing the participants' dexterity with and without the VR interface, independent of their skill level.

We measure the clumsiness of a participant p with an articulation a, noted $Clum(p, a)$, by monitoring the acceleration of the articulation. Because we want to measure rare and extreme events, such as brutally stopping an articulation when hitting the ground or a mechanical stop, we define the clumsiness as the standard deviation of the acceleration of an articulation over all exercises, noted $StdAcceleration(p, a)$, for example, with the boom (2):

$$ClumRaw(p, a) = StdAcceleration(p, a) \tag{2}$$

In order to aggregate the clumsiness of a participant for each articulation, we normalise their amplitudes, with respect to the average value over all participants (3).

$$Clum(p, a) = \frac{ClumRaw(p, a)}{Average_p(ClumRaw(p, a))} \tag{3}$$

We observed that the normalised clumsiness values of a given participant for each articulation are strongly correlated, telling us that we can measure a meaningful feature of the participant by averaging his clumsiness for all articulations (4):

$$Clum(p) = (Clum(p, Boom) + Clum(p, Stick) + Clum(p, Bucket))/3 \tag{4}$$

All the values are re-scaled in $[0, 1]$ and we define the dexterity of a participant, $Dexterity(p)$, as the opposite of his clumsiness, such that a higher value is desirable (5) and (6):

$$Clum_{Rescaled}(p) = \frac{Clum(p) - min_p(Clum(p))}{max_p(Clum(p)) - min_p(Clum(p))} \tag{5}$$

$$Dexterity(p) = 1 - Clum_{Rescaled}(p) \tag{6}$$

2.5. Exercises

We based our evaluation on 3 exercises of increasing difficulty: a tutorial exercise composed of primary movements, a target-reaching exercise, and a digging exercise. They are performed in the same conditions in VR and RM. Also, while doing the exercises in VR, the users control a digital-twin of the real machine, which makes the real machine move in real-time too. The primary objective of this study is to assess participants' performance in real-life exercises, reflecting actions that are commonly performed in everyday tasks, as well as primary movements.

2.5.1. Exercise 1—Tutorial

The first exercise focuses on elementary movements, the rotation of single degree of freedom (DOF) articulations. These exercises serve a dual purpose: as introductory tutorials to familiarize users with the controls and interface, and as analytical tools to gather data on the ease or complexity of different primitive actions with the different interfaces. By progressively mastering these basic movements, participants will be better prepared for more complex tasks.

The user must move one given articulation at a time, in a cycle of full opening: 3 s break—full closing—3 s break—default (middle) position. Each cycle is repeated 3 times for each articulation, in the order: boom, stick, bucket, turret (see Figure 3).

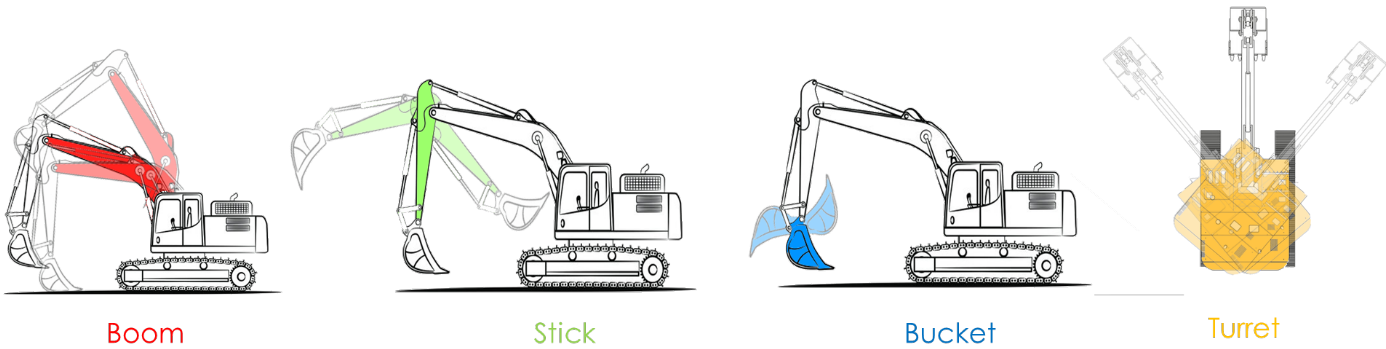


Figure 3. Exercise 1—Tutorial.

2.5.2. Exercise 2—Target Practice

The second exercise is closer to a real-life scenario. The user has to reach 6 targets placed on the ground with the tip of the bucket. The user is free to choose the path to complete the exercise, which articulations are used, how many are moved at the same time, and at which speed. The idea is to let the participant find their own way of controlling the machine in total freedom.

The user has to reach each target in a given order, with increasing level of difficulty. In between each target, the user must come back to a default neutral position, as represented in Figure 4. Validation of the target reaching is carried out in RM by the examiners, and in VR with a system of collision detection attached to the target object. Targets and default position are the same for the rest of the experiment.

2.5.3. Exercise 3—Dig and Fill

The third exercise is based on the most common real-life scenario of an excavator operator—digging and filling holes. The participant has to perform digging cycles: first they dig in front of them—at target 1—, then unload their bucket on the side—at target 2—, then load the previously unloaded earth at the same spot—target 2—and unload it where it was first taken—target 1. Then, the user comes back to the default position (see Figure 5). This cycle is repeated 3 times.

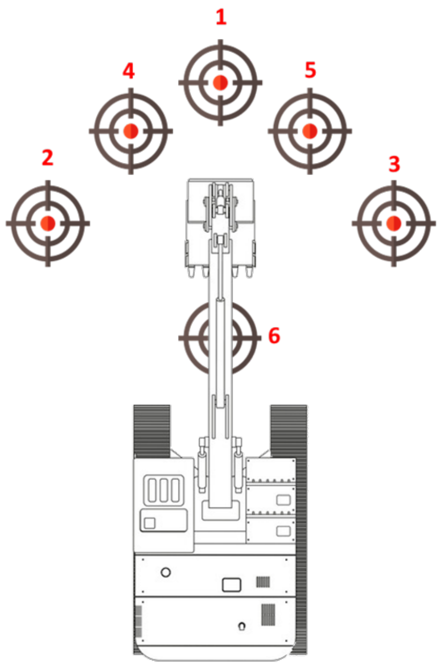


Figure 4. Exercise 2—Target Practice. Excavator is shown in default neutral position and targets in order 1 to 6.

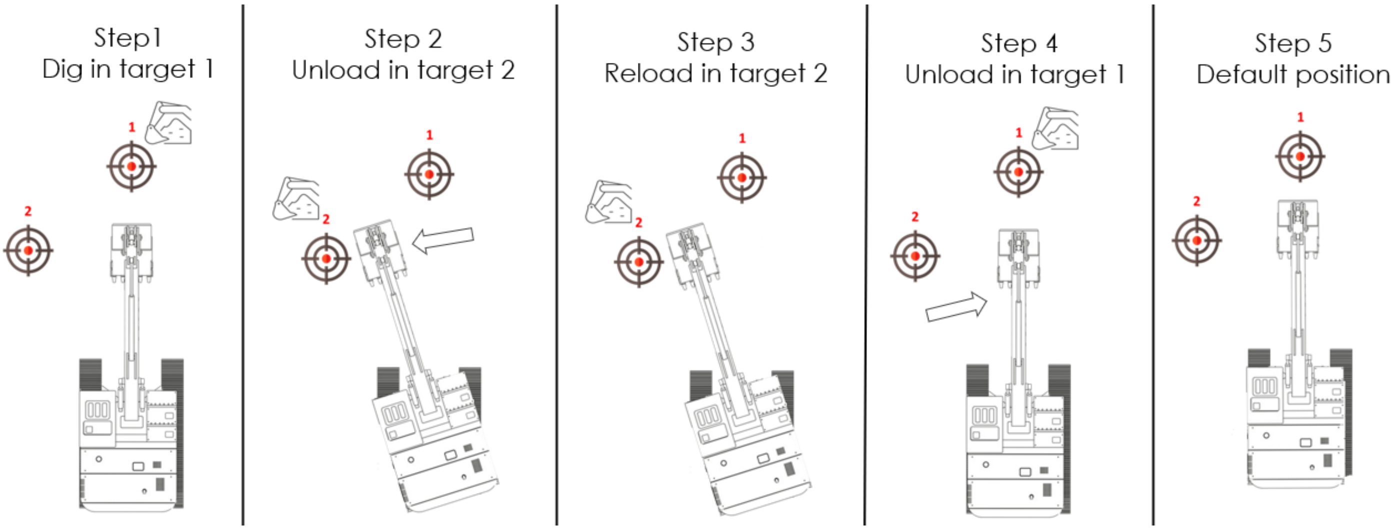


Figure 5. Exercise 3—Dig and Fill.

2.6. Participants

A total of 10 participants were included in this study, of various ages (between 22yo and 45yo with an average of 29.2yo), backgrounds (including students, engineers, and vehicle operators) and degrees of experience with both the real and the virtual technologies. Each participant was given an experience level score for each technology depending on their practical knowledge of the technologies. The scale varies between 0 and 1, with 0 being "no practical knowledge", 0.5 being "practical experience in the global technology (virtual environments/vehicles) but not on the specific machine (VR HMD/excavator)"—at least 100 h of global experience—and 1 being "extended practical knowledge of the specified technology (VR HMD/excavator)"—at least 100 h on the technology itself—with values ranging in between depending on their hours of practice.

The idea behind this score is distinguishing expert-operators (today's excavator operators) from non-operators (tomorrow's excavator teleoperators) and evaluate how differently they perform with both interfaces. Our measures and results illustrate that the VR interface offers the same quality of control to both groups, despite their disparity in experience with earthwork machinery.

3. Results

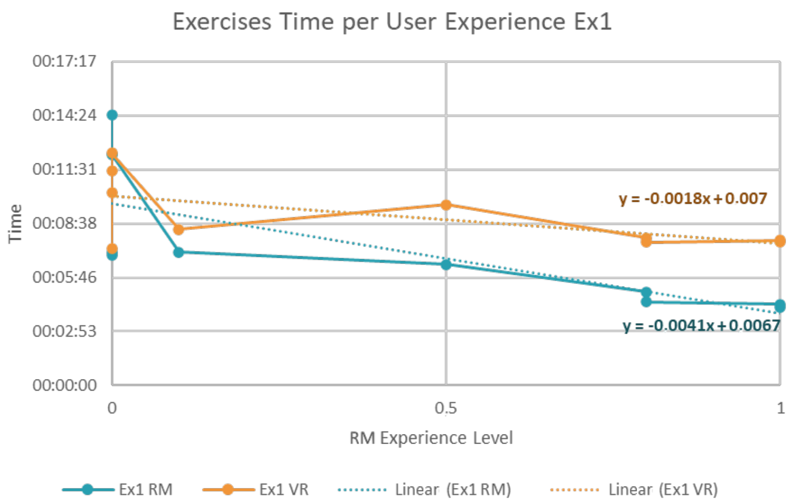
Our evaluation is based on several metrics, quantified data, and user feedback:

- **Time**: completion time of the exercises, as well as each cycle separately;
- **Progression**: progression rate of the user obtained by comparing the time between each cycle within exercises;
- **Errors**: numbers of command errors committed by the users;
- **Multi-control**: percentage of time spent per exercise using at least two commands at once;
- **User feedback**: carried out mostly through the NASA-TLX survey, completed using an open survey where the users were able to give varied personal feedback.

3.1. Time

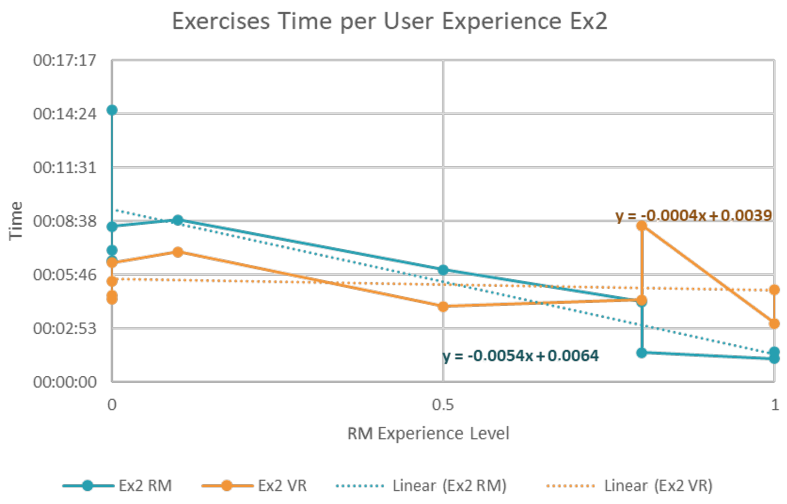
To measure the participants' performances in both interfaces, we timed the exercises—and made the participants aware of that fact. For each exercise, we compared the completion time of all participants, depending on their experience level with a real excavator. For each interface, we also draw the linear trend to evaluate the differences between the experienced and non-experienced participants (see Figure 6).

These results show that the time between interfaces is comparable. RM is usually faster than VR, which is partly due to the system limitations. It is noteworthy to mention that the VR-RM bridge is dependent on the Unity-ROS connection, which can provoke some lag in the control-feedback loop. Also, for human, environment, and machine safety, as well as servo-control issues, the maximum speed attainable by VR control has been reduced in our experiments to around half the full speed of the machine.

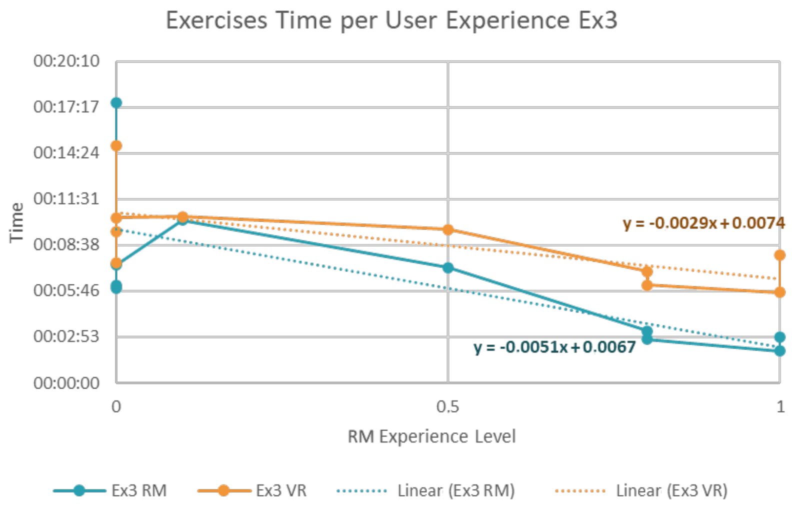


(**a**)

Figure 6. *Cont.*



(**b**)



(**c**)

Figure 6. Exercises completion time on RM/VR interfaces versus RM experience level. (**a**) Exercise 1; (**b**) Exercise 2; (**c**) Exercise 3.

The difference in linear trend is also meaningful. While it is obvious that the more experienced a participant is at operating an excavator, the better their performance on the real machine is, the results showed that the VR interface tends to flatten the differences between users regardless of their prior operating experience. The trend lines' coefficients never reach 0—which would be our final objective—but are always lower in VR than in RM, showing how VR is less dependent on experience, a supportive result for the development of this interface.

3.2. Dexterity

The best excavator operators are not only the fastest—doing a lot of work quickly—but also the smoothest—preserving the machine in the long run. The idea of smoothness is deeply embedded in the design and implementation of the remote control system, and we show in this section that our interface allows even the most novice of operators to operate the excavator in complete safety with regards to the machine.

We measured the normalised smoothness and speed (as defined above in Sections 2.3 and 2.4) of all the participants and visualised their performance in Figure 7. The higher the dexterity and speed the better; it means not only that the users are fast in their movements—better performance—but that the way they perform them is smoother—not damaging the machine, keeping it in better condition in the long run.

We computed the smoothness and speed for both interfaces, as seen in Figure 7.

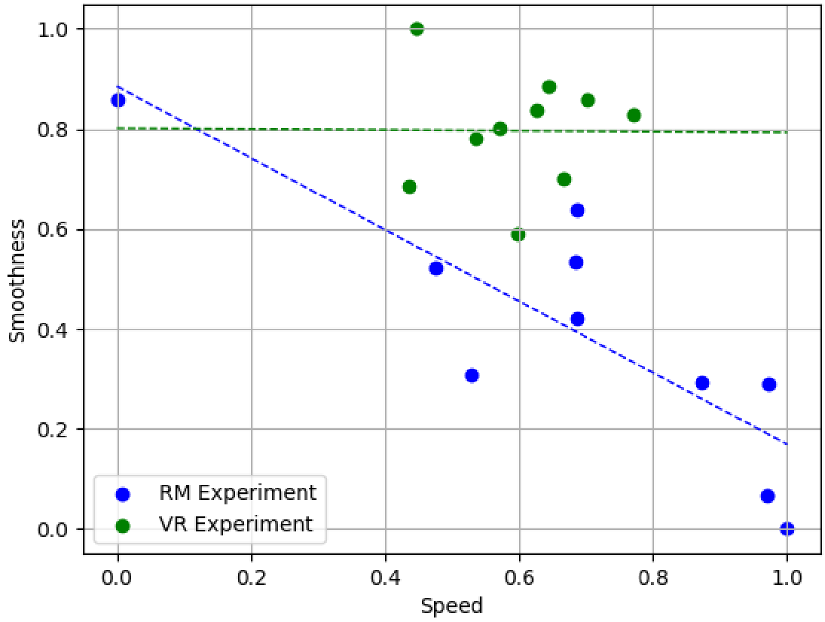


Figure 7. Smoothness and speed comparison in both interfaces.

The ideal control behaviour, which optimises both speed and smoothness, is situated at the upper-right corner of the graph. As depicted in the graph, users operating the RM at higher speeds tend to exhibit lower smoothness in their movements. While this allows for excellent performance in terms of completing tasks quickly, it poses a risk of damaging the machine over time. In contrast, the VR interface, on average, operates at a slightly slower pace but maintains a consistently high level of smoothness. Tasks performed in VR yield longer-lasting benefits for the machine's condition, as well as enhanced precision, albeit with a minor trade-off in execution speed compared to RM.

3.3. Progression

We evaluated the progression rate of each participant throughout the different cycles of each exercise. We compared analogous cycles: for exercise 1—each articulation (boom, stick, bucket, turret); for exercise 2—matching targets (targets 2 and 3 are Lateral 1, targets 4 and 5 are Lateral 2, and targets 1 and 6 are Forward targets, see Figure 4); and for exercise 3—the full dig cycles.

The obtained data are sorted for both interfaces in three categories: global rates for all participants, interface-only rates (the participants that started with the interface without training on the other first), and dual-interface rates (participants that started with the other interface before carrying out the experience with the specified one, see Table 1).

Table 1. Progression rates.

Interface	Interface Order	Boom	Stick	Bucket	Turret	Lateral 1	Lateral 2	Forward	Dig
RM	Global	19.2%	22.1%	13.0%	12.5%	0.6%	−8.0%	−2.1%	12.2%
	RM-only	11.5%	21.9%	12.8%	13.9%	−2.3%	−19.3%	1.8%	12.5%
	VR → RM	26.9%	22.3%	13.2%	11.2%	3.5%	3.3%	−5.9%	11.8%
VR	Global	21.8%	3.2%	21.7%	10.6%	−61.6%	−3.2%	−41.4%	2.2%
	VR-only	18.9%	−18.7%	23.4%	2.8%	−59.1%	4.0%	−24.1%	4.6%
	RM → VR	24.6%	25.1%	19.9%	18.8%	−64.1%	−10.3%	−58.7%	−0.2%

- **Global Progress**: the global progression rates are higher in RM compared to VR. This discrepancy arises because most individuals with no prior experience in excavator operation, often labelled as non-operators, are complete novices when it comes to operating excavators. However, many of them have some familiarity with computer simulations, enabling them to adapt relatively well to the VR interface;
- **RM Progress**: training with VR before using RM does bring improvement in the progress rate—on average 4.2% compared to starting directly on the RM—and this is quite consistent across the different tasks;
- **VR Progress**: conversely, training in RM not only fails to enhance progression in VR but sometimes even hinders it, resulting in an average regression rate of −0.7% when individuals first train in RM before using VR. Furthermore, this effect varies significantly depending on the specific task at hand.

3.4. Trials and Errors

Starting with the second exercise, participants are free to complete the task according to their preferences, in terms of speed, composition of movements, and trajectory followed. Such control freedom facilitates compound movements, from which we extract two measures: command errors and command combinations.

A "command error" is defined as any instance where an articulation movement deviates from its intended execution, for example, turning the turret left when the intended movement was to the right. It is important to note that this evaluation focuses exclusively on articulation control errors and not on the correctness of actions such as digging movements.

Command combinations are obtained by dividing the time spent using two or more articulations at the same time by the total completion time of the exercise. This metric is only recorded on exercise 2 and 3, as exercise 1 states explicitly to move only 1 articulation at a time—and the participants complied with the instructions correctly.

3.4.1. Errors

Errors are generally more prevalent in VR compared to RM. Nevertheless, when we examine error counts based on the order of interface usage, as shown in Table 2, we observe the positive impact of VR training. Conversely, training in RM for subsequent use in VR does not demonstrate any notable improvement in error rates.

Table 2. Errors comparison over interfaces order.

Interface	Interfaces Order	Ex1 Errors	Ex2 Errors	Ex3 Errors
RM	Global	0.4	4.5	3.2
	RM-only	0.8	5	4.8
	VR → RM	0	4	1.6
VR	Global	2.2	4.5	3.8
	VR-only	1.4	4.8	4.4
	RM → VR	3	4.2	4

When we examine the number of errors with respect to user profiles, as illustrated in Table 3, distinguishing between RM-operators (those with high RM experience, labeled RM-OP) and VR-operators (those with high VR experience, labeled VR-OP), we observe that the error count is more strongly correlated with experience rather than the order in which interfaces were used. Another noteworthy observation is that, while one might expect RM-operators to excel in RM compared to VR-operators, it is striking that RM-operators achieve nearly equivalent error rates in VR as well.

Table 3. Errors comparison over users profile. RM-OP—RM expert operators. VR-OP—VR expert operators.

Interface	Operators Profiles	Ex1 Errors	Ex2 Errors	Ex3 Errors
RM	Global	0.4	4.5	3.2
	RM-OP	0.2	1	1.6
	VR-OP	0.6	8	4.8
VR	Global	2.2	4.5	3.8
	RM-OP	2.2	4.6	3.2
	VR-OP	2.2	4.4	4.4

3.4.2. Compound Movements

The first noticeable observation is the significant reduction of compound movements rate in VR (see Table 4). Users are more hesitant to use multiple articulations at a time in VR, even when they start with the RM. However, starting with the VR gives them more ease and boldness when it comes to controlling the RM, underlying once again the positive training impact of the VR interface.

Table 4. Compound movements rate over interfaces order.

Interface	Interfaces Order	Ex2 Multi-Control	Ex3 Multi-Control
RM	Global	20.0%	25.3%
	RM-only	14.2%	14.6%
	VR $\rightarrow$ RM	25.8%	36.0%
VR	Global	7.0%	18.9%
	VR-only	12.2%	18.2%
	RM $\rightarrow$ VR	3.6%	19.8%

Similarly, akin to the error rates, the frequency of compound movements is more strongly associated with user experience rather than the order of interface usage. RM-operators notably outperform VR-operators, demonstrating their proficiency in machine control, even in a virtual context (see Table 5).

Table 5. Compound movements rate over users profiles. RM-OP denotes RM expert operators. VR-OP denotes VR expert operators.

Interface	Operators Profiles	Ex2 Multi-Control	Ex3 Multi-Control
RM	Global	20.0%	25.3%
	RM-OP	30.8%	42.9%
	VR-OP	9.2%	7.6%
VR	Global	7.0%	18.9%
	RM-OP	12.2%	29.3%
	VR-OP	1.7%	8.6%

However, it is important to note that VR exhibits a higher overall error count and a lower percentage of multi-control, indicating a requirement for ergonomic enhancements. For a comprehensive summary table, please refer to Appendix D Table A1.

3.5. User Feedback

User feedback was collected trough two rounds of NASA surveys, one after each interface experiment. The participants had the possibility to change their prior answers after the second experiment. The average results obtained are displayed in Table 6.

Table 6. NASA-TLX average results on a [1,20] scale.

Interface	Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration
RM	7.1	6.4	8.2	15.4	8.2	5.9
VR	11.7	6.7	5.3	13.5	9.5	5.1

While there are a few disparities, such as in temporal demand and performance, the most notable contrast between the two interfaces is on the mental demand. On average, users perceive VR as imposing a higher cognitive burden than RM. However, when assessing diverse user experiences, a discernible pattern emerges where RM-operators tend to experience the highest cognitive load, with the exception of those who already possess prior VR knowledge, as depicted in Figure 8. Participants were aware of being timed, which might have introduced a performance bias. However, according to their feedback, in both interfaces the temporal demand is rather low, suggesting the resulting low effect on the experiment.

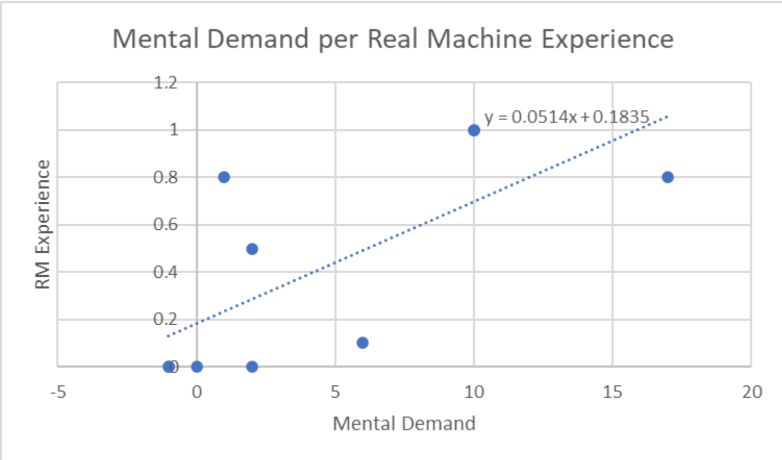


Figure 8. Mental demand per user experience.

4. Discussion

This initial investigation yields encouraging outcomes for VR interfaces in earthwork teleoperation. It demonstrates that both current operators (RM operators) and future teleoperators (non-experts) perform comparably using the interface. Moreover, it serves as an effective training tool for RM control. Nonetheless, it is important to acknowledge that the current VR interface does not meet the performance standards of the real machine, particularly in terms of speed.

Teleoperation played a hefty part in the earthwork field in recent years, more often than not coupled with the other rising star of Industry 4.0, robotic automation [17]. In terms of interface, however, to our knowledge, no other VR interface is used for direct earthwork teleoperation, making it difficult to compare our results with other studies. It is used for visualisation and/or planning [18], but not for direct control. The 2D methods for teleoperation show great results, as evoked earlier in reference [9], especially in terms of ease-of-use control, but only when the operators are on-site with visual feedback on the machine, which defies our purpose of removing the human on-site operator. VR interface for teleoperation is nonetheless a subject already investigated by others [19] with promising prospects, a trend to which our work aims to contribute.

4.1. Limitations

To establish a viable teleoperation system, the interface must match the real one in terms of both speed and safety, providing identical information about the machine and its environment. However, the current VR interface faces certain technical limitations that hinder its effectiveness:

- **Unity-ROS bridge**: to enable VR interface usage, a bridge between VR (Unity) and RM (ROS) software is essential. While the bridge is currently operational, it heavily relies on network connections. To ensure the swiftest data flow, the loop involving *receiving commands in VR, transferring them to ROS, translating them for the machine, moving the machine, and sending back the updated machine state to the VR twin* could benefit from optimisation in terms of speed. Presently, there is a slight delay between the user issuing a command and the virtual machine's response, resulting in movement inertia that can pose safety concerns. If left as is, this issue could impact the generalisation of our results to practical real-world scenarios. To mitigate this, the maximum user speed is restricted, which, in turn, slows down the entire process but increases safety. Additionally, lag can trigger motion sickness in some cases, significantly degrading the user experience;
- **Terrain state**: the current interface offers a real-time digital twin of the machine and control over it. However, for full usability, it also requires a real-time (or near-real-time) digital twin of the surrounding terrain so that the teleoperator can comprehend all the relevant factors of the current situation. Currently, the system employs a deformable terrain for prediction, and some work has been done to obtain terrain data through LiDAR, albeit in a one-time map update only. Further development should include real-time terrain updates, necessitating more real-time data transmission and thus demanding additional technical optimisation.

Outside of our method, other limitations exist. Network speed will always be an obstacle for any visualisation and/or control tool, as the better the network, the more information sent, the better the quality of the virtual environment, and the less the lag in the system. All in all, the better the network, the better the tool itself. Another point worth mentioning is legislation. As of today, governments are very reticent to allow autonomous machines on dig sites, or even remote-controlled ones, as they are not yet optimised enough to act on their promises of safety. Research in the field in thus limited to private dig sites, with research laboratories having to collaborate with private companies to gain access to real machines and dig sites for their experiments.

Potential solutions may emerge with the advent of 5G networks, which allow for enhanced data transfer capabilities. With such technology, considerations like 360° video streaming could be explored to enhance VR visualisation. As for laws, only time and technological improvements will tell.

4.2. Optimisation

Apart from technical enhancements, our system calls for design refinements based on both our findings and user feedback. Several key areas for improvement can be identified:

- **Virtual Annotations (VAs)**: our VR application already incorporates a variety of virtual annotations (VAs), as illustrated in Figure 1a. These VAs take the form of displayed mappings, a reticle projecting the bucket's position onto the ground, and task instructions. While these VAs were generally deemed useful by most participants as compensatory aids for the VR HMD's limited field of view (FOV), there is room for enhancing their positioning and utility. Our next significant project phase should center on the design and optimisation of existing and new VAs, primarily in a visual form, but also exploring auditory [20] and haptic [21] cues. VAs play a critical role in enhancing the user experience, particularly by improving control precision and environmental awareness. They should function as assistive tools, providing comprehensive task-related information to enhance safety and efficiency.
  To select the most pertinent VAs, we intend to leverage the integrated eye-tracking (ET) capabilities of our VR HMD for visual annotations [22], while user feedback will guide the development of other forms of VAs. Research on VAs has demonstrated their effectiveness, yet it also raises an important challenge: how to strike a balance

between improving the user experience and managing the additional cognitive load they may introduce. Addressing this conundrum is essential, and we plan to evaluate VAs' impact on cognitive load using electroencephalogram (EEG) analysis [23];

- **Cognitive Load**: the issue of cognitive load, particularly noticeable among experienced operators, must be addressed in the current VR interface. Our NASA-TLX surveys have highlighted this concern. While there will inherently be increased cognitive demand when individuals adopt a new and complex technology like VR, our design can be optimised to mitigate this effect. Also, the increase of information provided by the VR interface further inflates the mental load of the user. Therefore, EEG analysis is essential for evaluating which design elements and tasks are more mentally demanding [24], finding the root causes of the increased mental charge and thus enabling us to make targeted improvements;
- **Controllers and Mapping**: participants made more errors in VR than in RM, despite having the controls mapping displayed at all times. Our VR controls mapping was designed using a hybrid approach, aligning with the real machine's logic (e.g., pulling down on both touchpads to bring the bucket closer, akin to the RM's joysticks) but incorporating gaming-style buttons, as illustrated in Figure A2, Appendix B. We rejected natural movements like grabbing virtual commands due to their taxing, imprecise, and unsafe nature in our context. Furthermore, similar VR simulations have received negative feedback regarding natural movements [25]. Nearly all users expressed dissatisfaction with the control schemes: operators found them "not close enough to reality" (while adhering to the correct logic), whereas non-operators found them "illogical", particularly the need to pull down to raise the boom (despite using buttons), even among those who began with the RM interface featuring the same control logic. It is evident that a reevaluation of the controls mapping is imperative, as the hybrid approach failed to satisfy both experts and novices. Going a step further, exploring various controllers, beyond adjusting mappings, such as gamepads, keyboards, joysticks, and more, should be considered.

5. Conclusions

This article presents an initial study investigating the utilisation of a VR interface for earthwork teleoperation. The assessment encompassed both an actual excavator and its virtual counterpart. Ten participants, each with varying degrees of experience in machine operation, were tasked with completing three exercises of increasing complexity, progressively approaching real-world scenarios. The findings, derived from data analysis and user feedback, indicate that VR delivers comparable performance for both seasoned operators and novices. Furthermore, VR training enhances skill development on the real machine for both groups.

These promising outcomes emphasise the potential of VR interfaces in the teleoperation of intricate machinery, especially in the field of earthwork which has, as of yet, been scarcely touched by those kind of interfaces, provided that refinements are made in the design of virtual annotations, the reduction of cognitive load, and the selection of appropriate control schemes tailored to specific tasks.

Author Contributions: Conceptualisation, L.S., N.H., M.P. and C.F.; methodology, L.S. and N.H.; software, L.S. and N.H.; validation, L.S. and N.H.; formal analysis and N.H. and L.S.; investigation, L.S.; resources, M.P.; data curation, M.P.; writing—original draft preparation, L.S.; writing—review and editing, N.H., M.P. and C.F.; visualisation, L.S.; supervision, M.P. and C.F.; project administration, M.P. and C.F.; funding acquisition, M.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to ongoing research.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

VR	Virtual Reality
RM	Real Machine
HMD	Head-Mounted Device
DOF	Degree of Freedom
VA	Virtual Annotation
FOV	Field of View
ET	Eye-tracking
EEG	Electroencephalogram

Appendix A. Excavator Articulation Overview

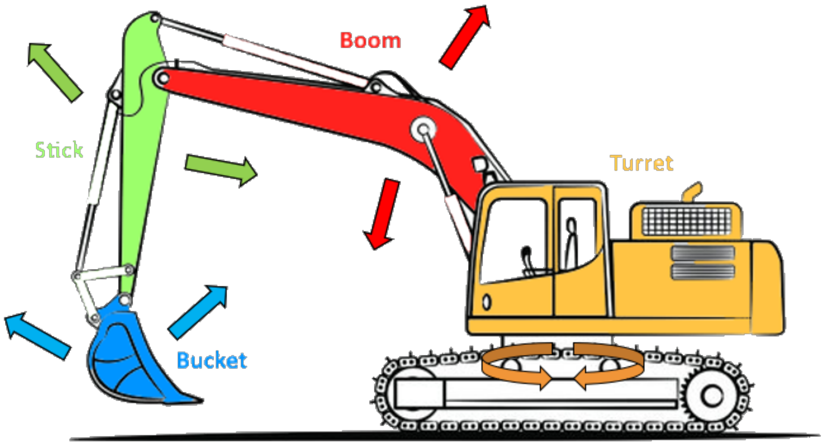


Figure A1. Excavator articulations.

Appendix B. VR Interface Mapping

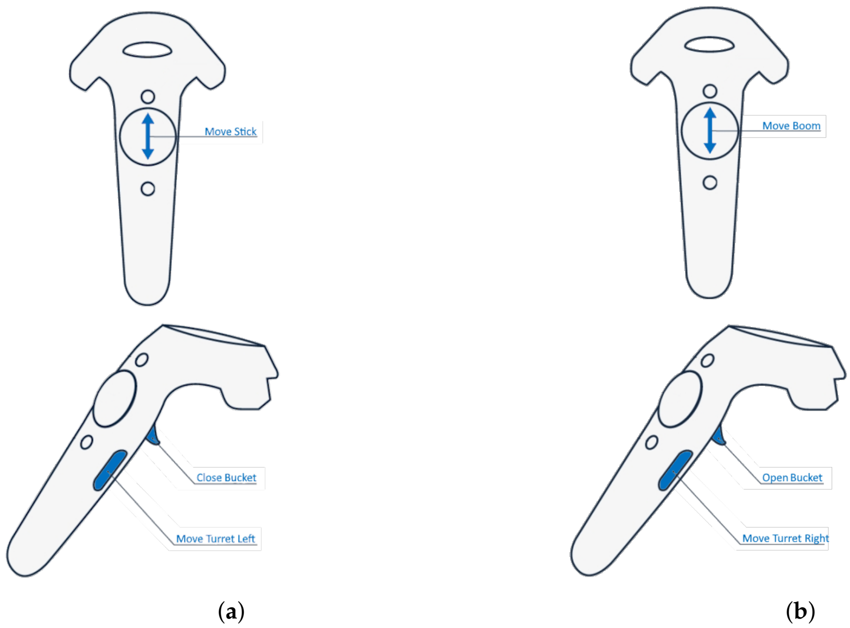


Figure A2. VR controllers mapping. (**a**) Left-hand mapping; (**b**) Right-hand mapping.

Appendix C. Simulated Terrain Overview

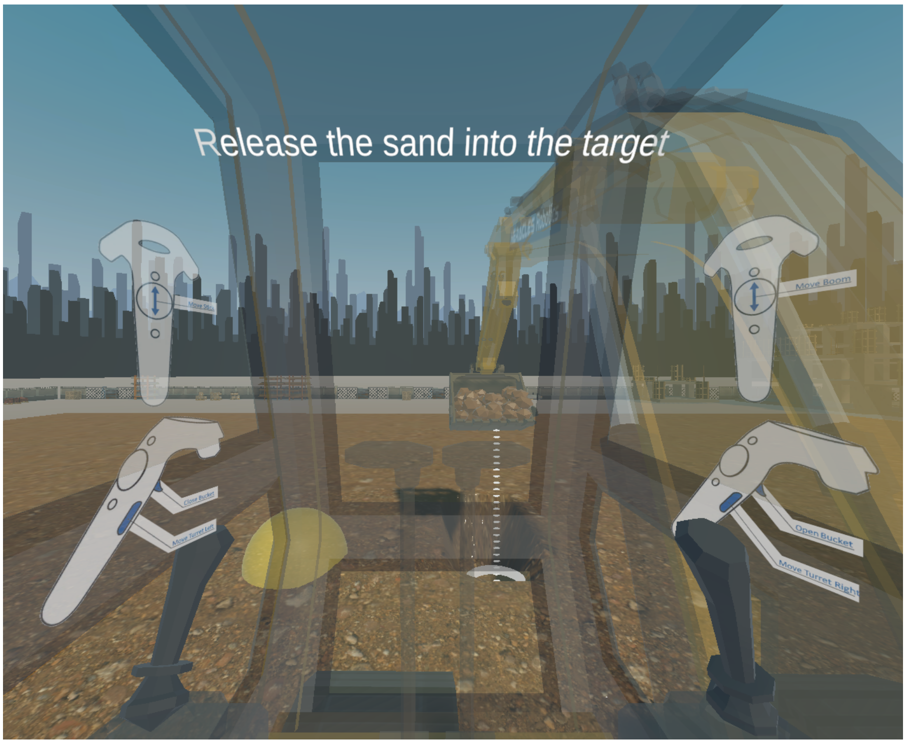


Figure A3. Deformable terrain using open construction simulator.

Appendix D. Summarised Errors and Dexterity Results

Table A1. Errors and multi-control rates.

Interface	Interface Order	Ex1 Errors	Ex2 Errors	Ex3 Errors	Ex2 Multi-Control	Ex3 Multi-Control
RM	Global	0.4	4.5	3.2	20.0%	25.3%
	RM-only	0.8	5	4.8	14.2%	14.6%
	VR → RM	0	4	1.6	25.8%	36.0%
	RM-OP	0.2	1	1.6	30.8%	42.9%
	VR-OP	0.6	8	4.8	9.2%	7.6%
VR	Global	2.2	4.5	3.8	7.0%	18.9%
	VR-only	1.4	4.8	4.4	12.2%	18.2%
	RM → VR	3	4.2	4	3.6%	19.8%
	RM-OP	2.2	4.6	3.2	12.2%	29.3%
	VR-OP	2.2	4.4	4.4	1.7%	8.6%

References

1. Sulaiman, H.; Saadun, M.N.A.; Yusof, A.A. Modern manned, unmanned and teleoperated excavator system. *J. Mech. Eng. Technol.* **2015**, *7*, 1.
2. Lee, J.S.; Ham, Y.; Park, H.; Kim, J. Challenges, tasks, and opportunities in teleoperation of excavator toward human-in-the-loop construction automation. *Autom. Constr.* **2022**, *135*, 104119. [CrossRef]
3. González, C.; Solanes, J.E.; Munoz, A.; Gracia, L.; Girbés-Juan, V.; Tornero, J. Advanced teleoperation and control system for industrial robots based on augmented virtuality and haptic feedback. *J. Manuf. Syst.* **2021**, *59*, 283–298. [CrossRef]
4. Eswaran, M.; Bahubalendruni, M.R. Challenges and opportunities on AR/VR technologies for manufacturing systems in the context of industry 4.0: A state of the art review. *J. Manuf. Syst.* **2022**, *65*, 260–278.
5. Adami, P.; Rodrigues, P.B.; Woods, P.J.; Becerik-Gerber, B.; Soibelman, L.; Copur-Gencturk, Y.; Lucas, G. Effectiveness of VR-based training on improving construction workers' knowledge, skills, and safety behavior in robotic teleoperation. *Adv. Eng. Inform.* **2021**, *50*, 101431. [CrossRef]

6. Stotko, P.; Krumpen, S.; Schwarz, M.; Lenz, C.; Behnke, S.; Klein, R.; Weinmann, M. A VR system for immersive teleoperation and live exploration with a mobile robot. In Proceedings of the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Macau, China, 3–8 November 2019.
7. Lee, J.S.; Ham, Y. Exploring Human-Machine Interfaces for Teleoperation of Excavator. *Constr. Res. Congr.* **2022**, *2022*, 757–765.
8. Hedayati, H.; Walker, M.; Szafir, D. Improving collocated robot teleoperation with augmented reality. In Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, Chicago, IL, USA, 5–8 March 2018.
9. Okishiba, S.; Fukui, R.; Takagi, M.; Azumi, H.; Warisawa, S.I.; Togashi, R.; Kitaoka, H.; Ooi, T. Tablet interface for direct vision teleoperation of an excavator for urban construction work. *Autom. Constr.* **2019**, *102*, 17–26. [CrossRef]
10. You, K.; Zhou, C.; Ding, L.; Chen, W.; Zhang, R.; Xu, J.; Wu, Z.; Huang, C. Earthwork digital twin for teleoperation of an automated bulldozer in edge dumping. *J. Field Robot.* **2023**. [CrossRef]
11. Zhou, T.; Zhu, Q.; Du, J. Intuitive robot teleoperation for civil engineering operations with virtual reality and deep learning scene reconstruction. *Adv. Eng. Inform.* **2020**, *46*, 101170. [CrossRef]
12. Hoffmann, N.; Saunier, L.; Prouten, S.; Serroukh, O.; Le Floch, J.C.; Preda, M.; Fetita, C.; Zaharia, T. Industrial Use-Case: Digital Twin for Autonomous Earthwork in Virtual-Reality. In Proceedings of the 27th International Conference on 3D Web Technology, Evry, France, 2–4 November 2022; pp. 1–4.
13. ROS—Robotic Operating System. Available online: https://docs.ros.org/en/foxy/index.html (accessed on 1 September 2023).
14. Unity Robotics Hub. Available online: https://github.com/Unity-Technologies/Unity-Robotics-Hub (accessed on 1 September 2023).
15. Open Construction Simulator. Available online: https://github.com/Field-Robotics-Japan/OpenConstructionSimulator (accessed on 1 September 2023).
16. Akyeampong, J.; Udoka, S.; Caruso, G.; Bordegoni, M. Evaluation of hydraulic excavator Human–Machine Interface concepts using NASA TLX. *Int. J. Ind. Ergon.* **2014**, *44*, 374–382. [CrossRef]
17. Ha, Q.P.; Yen, L.; Balaguer, C. Robotic autonomous systems for earthmoving in military applications. *Autom. Constr.* **2019**, *107*, 102934. [CrossRef]
18. Xu, S.; Fu, D.; Xie, Y.; Hou, L.; Bu, S. Integrating BIM and VR for highway construction site layout planning. In Proceedings of the CICTP 2020, Xi'an, China, 14–16 August 2020; pp. 1068–1079.
19. Salimov, M. Design and Manufacturing of a Virtual Reality-Assisted Teleoperated Excavator Platform. Master's Thesis, University of Nevada, Reno, NV, USA, 2020.
20. Ding, R.; Cheng, M.; Han, Z.; Wang, F.; Xu, B. Human-machine interface for a master-slave hydraulic manipulator with vision enhancement and auditory feedback. *Autom. Constr.* **2022**, *136*, 104145. [CrossRef]
21. Zhu, Q.; Du, J.; Shi, Y.; Wei, P. Neurobehavioral assessment of force feedback simulation in industrial robotic teleoperation. *Autom. Constr.* **2021**, *126*, 103674. [CrossRef]
22. McNamara, A.; Boyd, K.; George, J.; Jones, W.; Oh, S.; Suther, A. Information placement in virtual reality. In Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, 23–27 March 2019.
23. Hong, Z.; Zhang, Q.; Su, X.; Zhang, H. Effect of virtual annotation on performance of construction equipment teleoperation under adverse visual conditions. *Autom. Constr.* **2020**, *118*, 103296. [CrossRef]
24. Chen, J.; Taylor, J.E.; Comu, S. Assessing task mental workload in construction projects: A novel electroencephalography approach. *J. Constr. Eng. Manag.* **2017**, *143*, 04017053. [CrossRef]
25. Excavator Simulator VR—Steam Page Reviews. Available online: https://store.steampowered.com/app/1526380/Excavator_Simulator_VR/#app_reviews_hash (accessed on 1 September 2023).