Improving Immersive Telepresence Locomotion by Using a Virtual Environment as an Interface to a Physical Environment (VEIPE)

Eetu Laukka[1], Evan G. Center[1], Basak Sakcak[1], Steven M. LaValle[1], Timo Ojala[1], Matti Pouke[1]

***Abstract*— Immersive mobile robotic telepresence enables humans to feel present in a remote environment. These systems often use 360-degree panoramic cameras to stream video over a network to a head-mounted display (HMD) where the video feed is rendered to the user. This enables the user to freely look around in a remote environment. A drawback of using highly immersive technologies instead of a more traditional computer screen is that users often experience virtual reality (VR) sickness. Therefore, sometimes the users are only able to use these systems for brief durations. Moreover, the increase in bandwidth requirements of panoramic cameras and the time necessary to process the 360-degree panoramic view contributes to an often unacceptable amount of latency between the user's actions and the observed reaction of the mobile robot, which can be referred to as perception-actuation loop. We present a novel method to mitigate these problems in immersive mobile robotic telepresence systems. We call this method *virtual environment as an interface to a physical environment (VEIPE)*. In VEIPE, a digital twin of the remote environment is used to interface with the telepresence robot in the real remote environment. We present a study comparing teleportation through VEIPE as a locomotion method against a more traditional joystick-based continuous locomotion method for controlling a telepresence robot. Our results indicate that VEIPE induces less VR sickness compared to the joystick condition as measured by the simulator sickness questionnaire (SSQ) and users perform about 31 percent better in a simple navigation task. Furthermore, the users subjectively prefer teleportation through VEIPE over the joystick. We also present exploratory data about cognitive load measured with the NASA task-load-index (NASA-TLX) questionnaire, presence measured with the Slater-Usoh-Steed (SUS) questionnaire, and accumulated yaw in the navigation tasks.**

I. INTRODUCTION

Telepresence [1] enables humans to feel present in a remote environment. Mobile robotic telepresence (MRP) enables the user to embody and actuate a mobile robot, allowing them to move and act within the remote environment [2]. The remote environment is captured by sensors and relayed back to the user. A panoramic camera can be used to capture a 360-degree point of view from the remote environment. This can be achieved by using two or more opposite-facing wide-angle cameras. The video from the multiple cameras is then fused into a spherical panoramic video by stitching the frames together [3].

The panoramic video can then be rendered in a head-mounted display (HMD) to allow the user to look around freely practically without delay due to there being no observable latency in rotating the viewpoint. Another option is to use a pan-tilt camera that follows or predicts the user's head movements [4]. This type of immersive approach to telepresence can be called immersive MRP. However, some tasks, like navigation, can cause more cyber- or simulator sickness in more immersive settings, compared to more traditional, less immersive monoscopic screens [5]. Moreover, in virtual reality (VR), the different locomotion techniques have been identified as a primary factor in causing this sickness [6]. This is believed to be due to sensory mismatches, such as a mismatch between visual and vestibular sensory systems [7], [8]. Experiencing motion sickness can be common for some people, for example, while reading a book in a car. The symptoms are very similar to this feeling, but since there is no physical motion and rather just visually induced sickness, this can be referred to as visually induced motion sickness (VIMS) [7], [8]. There are multiple theories on what causes this nausea, and it can be a combination of multiple components [9]. In the context of VEs experienced via HMDs, the specific phenomenon has been referred to as VR sickness [10]; therefore, this is the term used in this paper.

Whereas telepresence robots can navigate autonomously using motion planning methods, their motion can also be manually controlled by a user through some user interface that maps the user input to a control command. In the case of autonomous navigation, the user can specify a target position or a sequence of waypoints as an input. Furthermore, the user can manually control the robot by mapping analog joystick commands to the robot's control commands.

Since telepresence relies on video streaming through a network to provide the remote environment visually to the user, it inherently possesses some delay between the user's intended actuation of the robot and the corresponding sensing of that actuation. This latency, which can reach up to seconds [11] in the perception-actuation loop, is naturally one of the issues to be dealt with in telepresence applications, and the problem is amplified in immersive telepresence systems since they inherently possess much more delay caused by larger bandwidth requirements and added processing.

We present a novel locomotion method for immersive telepresence to address the main issues of immersive telepresence locomotion: VR sickness and latency. We call it *Virtual environment as an interface to a physical environment* (VEIPE). VEIPE is the 3D reconstructed VE, or a digital twin, of the physical remote environment. We refer to the 3D reconstructed environment of the physical remote room

This work was supported by Infotech Oulu ;the European Research Council (project ILLUSIVE 101020977); and the Tauno Tönning Foundation

[1]Author is with Faculty of Information Technology and Electrical Engineering, University of Oulu, P.O. Box 4500, Oulu, FI-90014, Finland `firstname.lastname@oulu.fi`

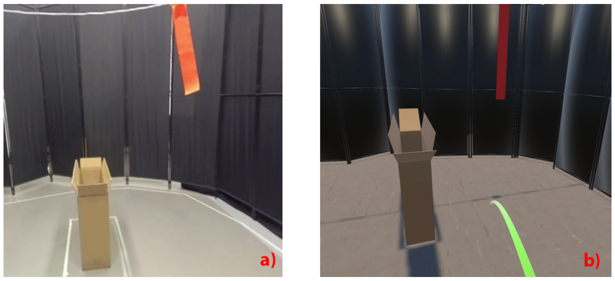

Fig. 1. a) User's view of the real remote environment, b) User teleporting through VEIPE in the digital twin of the remote environment (teleportation indicator in green).

as just *VE*, and the physical remote room as the *remote environment*. VEIPE can be created from a high fidelity VE, seen in Fig. 1b modeled from the remote environment seen in Fig. 1a, or it can also be a simple approximation, such as a point cloud generated using depth enabled cameras and/or artificial intelligence (AI). The remote environment and VEIPE must be mapped in a way that the Cartesian coordinates of the VE can be mapped to the corresponding position in the remote environment. This enables us to transfer the immersive telepresence user to the virtual representation of the physical environment, where the user can teleport around, similar to many commercial VR applications.

Teleportation through VEIPE is depicted in Fig. 1b, where the green indicator is the pointer the user uses to select a target position. While the user is navigating in the VE, the MRP system navigates to match the user's desired location. Once the MRP system has reached the user's target position, the system can transfer the user to the remote room by again showing the camera's live view in the remote environment. Since the navigation is done in the VE without any delay, the panoramic camera video streaming delay cannot affect the controlling of the robot. Assuming the robot can perform robust localization, mapping, and autonomous navigation–key requirements for VEIPE–it will enable the use of all the same locomotion methods proposed in conventional VR application research, such as teleportation and natural walking, which are known to induce less VR sickness than simple joystick translation [12], [13], [14], or even redirected walking [13], enriching the possibilities of current immersive telepresence applications.

II. Related Work

Immersive telepresence has some advantages over less immersive telepresence applications, particularly in achieving greater perceptual abilities such as depth perception [15], improved task performance [16] and situational awareness [15]. Additionally, media with higher immersion seems to consistently produce a heightened sense of presence [17]. However, in certain applications, it can lead to an increase in cognitive load or even cognitive overload [18], [19]. In some cases, a monoscopic screen may be sufficient and even preferable. However, specific tasks such as object height estimation can be significantly impaired on monoscopic screens [20], although there have been promising recent advances in creating rich 2D interfaces to help solve problems in less immersive versions of teleoperation or telepresence [21].

One of the major obstacles preventing mass adoption of immersive technologies is VR sickness, which also applies to telepresence or teleoperation applications [22], [19]. Some approaches have been developed to cure, detect, or prevent VR sickness. Locomotion is a big factor in inducing VR sickness in VR applications and telepresence [23]. Matching the user's physical motion in VR seems to be the most effective way of preventing sickness, but it inherently carries the major downside of requiring a physical space as large as the VE. Moreover, it should also have a similar shape. Therefore, often some manipulation of the user's viewpoint in the VE is required to allow them to traverse different environments. From those methods, instant viewpoint jumps, or teleportation, tend to cause the least amount of VR sickness, at a level comparable to natural walking, since there is no continuous motion that would induce a visual sensory mismatch [12], [13]. However, it should be noted that a significant drawback of teleportation is that it can disorient the user [24], [25].

The delay between the actuation and perceived action of the robot has been shown to significantly affect completion times of tasks like manipulation or navigation in teleoperation [26]. Moreover, it has been reported that latency has a negative effect on physical demand in a teleoperated search task [27]. Humans can perceive delays as small as 10-20 ms [28]. In the context of teleoperated robots, the system latency should not exceed 250 ms in driving-like tasks [29]. It has been observed that humans stop trying to compensate for this delay at around 1 second and rather start incorporating slower tactics like "move and wait," which can have a crippling effect on navigation performance [28], [30]. Suggestions of predictive displays to address this problem have existed for a long time [31]. With a predictive display, the operator is shown an estimate that something will happen or is happening to the robot based on the control actions and computed predictions.

Some advances have been made in creating a 3D reconstruction of a remote environment to display to teleoperators for improved performance [11], [32], [33], [34]. These have been used as an additional interface, often in a cockpit-style setting to show an exocentric view of the teleoperated robot. However, there are few existing approaches that rely on transforming an immersive telepresence user's environment into a representation of the physical remote environment to interface with the remote robot. In a paper by Kato, a similar VE interface was used to operate a real telepresence robot [35]. Their interface consisted of an exocentric view of the approximated position of the telepresence robot in the VE interface. Our method extends this idea that a VE can be used as a general interface to the physical remote environment, especially in the immersive telepresence setting.

In telepresence, teleportation has not been used as a locomotion method mainly because there is no way to actually teleport the physical robot. This challenge is addressed by our novel approach, teleportation through VEIPE, which provides a solution to this limitation.

III. HYPOTHESES

The goal of the study was to evaluate whether teleportation through VEIPE is a viable candidate as a locomotion method for immersive MRP. Traditionally, telepresence robots are controlled with a directional control to rotate and translate the robot [36]. This control method was chosen as the baseline for comparison to our method. VEIPE was designed to overcome sickness and to mitigate negative effects of latency. Thus, the following directional predictions and accompanying analysis methods were preregistered in the Open Science Framework (OSF) at https://osf.io/qg2sc. Our primary hypothesis (H1) was that VR sickness would be significantly lower when using teleportation through VEIPE as a non-continuous locomotion method compared to the baseline condition's joystick-based continuous locomotion method, where the user sees the physical environment only, as measured by the total weighted simulator sickness questionnaire (SSQ) score [37]. Additionally, we hypothesized (H2) that task efficiency would be increased when using teleportation through VEIPE as a locomotion method compared to the baseline condition's joystick-based continuous locomotion method, as measured by the number of reached navigation goals in the given time. Finally, regarding preference, we hypothesized (H3) that the users would prefer teleportation through VEIPE over joystick as measured by the forced choice question, "Which method would you choose for this task if you had to pick one?"

The baseline condition, where users navigated the remote environment using the joystick-based continuous locomotion method, will be referred to as the *joystick* condition. The experimental condition, where users used teleportation through the VEIPE as a non-continuous locomotion method, will be referred to as the *VEIPE* condition.

IV. EXPERIMENT

A. Participants

Based on the typically large effects reported in previous literature comparing continuous locomotion and teleportation as locomotion, the conservative effect size chosen for an SSQ total score difference between these two locomotion methods was Cohen's $dz = 0.5$. *A priori* power analysis indicated a need for 31 participants to detect effects of this magnitude or larger with 80% power for our selected SSQ analysis method. In order to fill the counterbalancing groups, one more participant had to be collected; therefore, data was collected from 32 participants. The participants were recruited from the wider university community via the University of Oulu Sona system, and they were rewarded with compensation worth around 17 euros, consisting of a coffee gift card and University of Oulu merchandise.

Ultimately, data from 38 participants were collected since six participants had to be excluded, following our preregistered exclusion criteria. Three participants did not show up to the second session of the study. In two participants' sessions, technical problems with the robot prevented a full run of the study. One participant was unable to follow the task instructions and had to be excluded. Out of the 32 valid participants, 20 were male, 11 were female, and 1 preferred not to say. The participants' ages ranged from 18 to 67, with most participants being aged 18 to 34. The mean age of the sample was 28.97 ($sd = 11.35$). All of the participants had normal or corrected-to-normal vision.

Most of the participants had some prior experience with VR systems. The participants were asked: "Have you used any VR systems previously, and how many times?" There was only one participant who had no experience beforehand, 12 participants who had used VR systems once or twice, 9 who use VR systems once or twice a year, 6 who use VR systems once or twice a month, 2 who use VR systems once or twice a week, and 2 who use VR systems several times a week. There were no participants who reported using VR systems daily. The participants were also asked about their video game background by asking: "How often do you play or used to play video games?" Two participants answered never, 5 answered once or twice, 6 answered once or twice a year, 5 answered once or twice a month, 2 answered once or twice a week, 8 answered several times a week, and 4 participants reported playing video games daily.

B. Procedure

Protocols were approved by the University of Oulu Ethical Review Board. The study was conducted using a within-subjects design. The two main conditions, joystick and VEIPE, were counterbalanced by condition and navigation goal path orientation (clockwise or counterclockwise), resulting in four counterbalancing groups. Each participant was randomly assigned to one of these groups.

In the joystick condition, participants constantly saw the 360-degree panorama view of the remote environment live-streamed to the user's HMD. The participant had to reach different navigation goals by controlling the telepresence robot using a game controller with two thumbsticks while dealing with the robot's input latency and the camera's live stream feedback latency.

In the VEIPE condition, participants used teleportation through the VEIPE as the locomotion method. In this condition, the participant could control the view to be switched to the VEIPE counterpart of the remote environment where there is no latency between the user movement and the perceived movement. The participant would navigate the telepresence robot to the navigation goals by selecting a target position through VEIPE. This teleportation action would instantly transfer the participant's viewpoint to the selected position, after which the telepresence robot would navigate autonomously to this position. When the robot reached the users' desired position in the VEIPE condition, the panoramic camera live stream was shown again to the user automatically.

The main task for the participants was to reach navigation goals in either clockwise or counter-clockwise order. The ribbons of different colors (as seen in Fig. 2b) visualized the

Fig. 2. Lab environment visualized. (a) User using teleportation through VEIPE. b) Robot navigating the remote environment. c) Close-up of the telepresence robot.

navigation goals. The path orientation was counterbalanced in order to mitigate learning effects.

After a participant had reached a navigation goal, which was a 20 cm radius around the ribbon, the participant would hear a sound indicating that the navigation goal had been reached and could then move on to the next navigation goal. In the VEIPE condition, after reaching a navigation goal, the participant would have to wait until they were transferred to the remote environment from the VE in order to hear the "navigation goal reached" -sound. In both conditions, the participants had a 1-minute practice session to learn to control the telepresence robot, and following this training session, they had 4 minutes to navigate to as many navigation goals as they could.

Before each session, a consent form was collected from the participants, informing them what kind of data would be collected and how it would be processed. Between the conditions, participants completed questionnaires, always beginning with the SSQ questionnaire, followed by the NASA-TLX questionnaire [38], and finally a revised version of the SUS presence questionnaire [39], [40], (similar to the one used by Pan et al. [41]). After the second session, the users would additionally answer two forced-choice questions: "Remembering the two times you did the tasks, which one of them felt easier to complete?" with the options "first session" and "second session" and "If you would need to control a telepresence robot, which method would you choose: Teleporting or continuous control?" Finally, open-ended questions on how they would improve each method were collected for exploratory analysis. The conditions took place on two separate days, within 1 to 7 days between each session. This was designed to mitigate any carryover effects of VR sickness [42].

C. Setup

We implemented a fully functional immersive MRP system by attaching an Insta 360 Pro 2, a 360-degree panoramic camera, on top of Husarion ROSbot XL, an autonomous mobile robot platform. The telepresence robot is seen in Fig. 2c. The Husarion ROSbot XL uses the Robot Operating System (ROS) [43]. The robot is able to map the surrounding environment through Light Detection and Ranging (LIDAR). A 2D map of the experiment room was obtained using the LIDAR on board the robot. The mapping was done offline to avoid possible errors that sometimes occurred during the live mapping. This ensured more stable data collection. This 2D map is essential for the robot's motion planning. The Husarion ROSbot XL also comes with ready-made software packages for autonomous navigation that rely on ROS 2 Nav2 software package [44]. The setup was configured to use A* to compute a path to the goal and a Regulated Pure Pursuit Controller to track the computed path. This enabled the telepresence robot to navigate to the user-selected position through VEIPE. The speed of the robot was set to 0.75 m/s for both conditions. This was a choice made by the authors in pilot testing, which represents an attempt to balance the speed of the robot such that it was fast enough in the VEIPE condition not to annoy users waiting for the robot arrive at the teleport destination, but slow enough not to be overly error prone to overshooting turns in the joystick condition.

The 3D reconstructed VE, or the digital twin of the remote room, was modeled beforehand into the VEIPE application using Unity 2021.3.19f1 Game Engine. This model was carefully adjusted to match the size of the physical remote room. The robot's starting location was used as the frame of reference for the VE. This means the origin of the Cartesian coordinates in the VE was the starting position of the robot, yielding matching frames of reference between environments.

The panoramic video was streamed wirelessly over the Insta 360 Farsight wireless control system, which is a transmitter and receiver pair that communicates over 5.18 GHz Wi-Fi. This video stream was then rendered through the VEIPE application. The video stream is a separate pipeline from the ROS and is responsible for most of the system latency. The measured latency of the actuation-perception loop was approximately 1.6 s. Even though the use of a stereoscopic view to enable better depth perception and more than 1920x1080 resolution was possible with the camera on hand, we opted to forgo these options because we wanted to keep the latency minimal.

The users could view the remote room through Valve Index VR headset. Two different controllers were used: Valve Index Controllers for the VEIPE condition, and a Logitech Wireless

Gamepad F710 controller for the joystick condition with directional controls. A normal video game controller was chosen for the joystick condition to maximize familiarity with this control method. It had two separate thumbsticks for the thumbs to control the rotation and translation of the robot individually. Holding one controller with two hands was thought to be a more natural method for the joystick condition than holding two Valve Index Controllers separately. Furthermore, these two locomotion methods are completely in separate domains, so the choice to use different input methods helped to emphasize the difference between the locomotion methods and prevent potential carryover effects.

V. Results

A. VR sickness (H1)

We used the SSQ questionnaire as our measure of VR sickness in both conditions. A Wilcoxon-signed-rank test (one-sided, alpha = .05) was used to compare the SSQ total scores between the two conditions. To avoid carryover effects, the conditions took place in two different sessions within one to seven days apart. VEIPE condition ($Mdn = 0, sd = 15.32$) was found to induce significantly less VR sickness compared to the joystick condition ($Mdn = 20.57, sd = 36.45$), $Z = 4.02, p < 0.001, r = 0.71$. The SSQ results are plotted in Fig. 3, demonstrating the large advantage of using teleportation with VEIPE for an immersive telepresence system over the joystick-based continuous locomotion method.

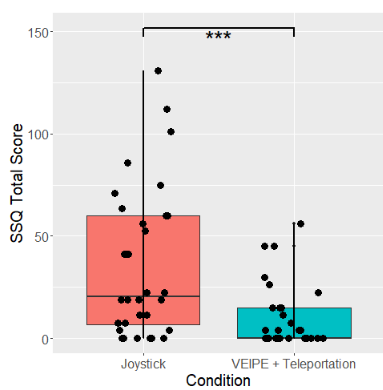

Fig. 3. SSQ Total Score between conditions

B. Performance (H2)

The participants had four minutes to reach as many navigation goals as they could in both conditions. These scores were compared with a Wilcoxon-signed-rank (one-sided, alpha = .05) test for matched pairs. Navigation goals reached in the VEIPE condition ($Mdn = 10, sd = 2.49$) indicated a significant improvement of around 31 percent over the joystick condition ($Mdn = 7, sd = 2.53$), $Z = 3.82, p < 0.001, r = 0.68$. Fig. 4 depicts the improvement in the navigation goals reached.

C. Preference (H3)

The participants were asked: (Q1) "If you would need to control a telepresence robot, which method would you choose: Teleporting or continuous control?" with the response options: "Continuous control/Joystick" or "Teleporting" and (Q2) "Remembering the two times you did the tasks, which one of them felt easier to complete?" the answer options being: "The first time" or "The second time." Our results show that VEIPE was preferred in Q1, with 25 out of 32 choosing the answer: "Teleportation." This was confirmed with an exact binomial test (one-sided, alpha = .05), $p < 0.001$. Q2 indicated an even stronger effect of people finding the VEIPE method easier to use, with 29 out of 32 answering in favor of the VEIPE condition. This was confirmed with an exact binomial test (one-sided, alpha = .05), $p < 0.001$.

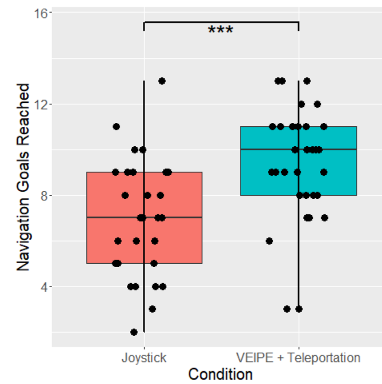

Fig. 4. Comparison of navigation goals reached in both conditions

D. Exploratory Results

On top of the primary predictions, exploratory data was gathered. We measured cognitive load using the NASA-TLX questionnaire. The six components of the NASA-TLX were not weighted, so we present here a variant of the NASA-TLX questionnaire called Raw-TLX (RTLX) [45]. This has become popular since it is easier to implement. However, it should be noted that the Raw-TLX has been shown to perform at times worse [46] and at times better [47] than the original NASA-TLX [38].

We compared each component individually between the conditions with a Wilcoxon-signed-rank (two-sided, alpha = .05) test for matched pairs. From these components, Effort ($Z = 4.53, p < 0.001, r = 0.80$), Frustration Level ($Z = 3.74, p < 0.001, r = 0.66$), and Mental Demand ($Z = 3.53, p < 0.001, r = 0.62$) were significantly improved in the VEIPE condition in comparison to the joystick condition. However, the remaining three components, Performance ($Z = 0.68, p = 0.25, r = 0.12$), Physical Demand ($Z = 1.42, p = 0.08, r = 0.25$), and Temporal Demand ($Z = 0.15, p = 0.44, r = 0.03$) were not significantly different. The results from the NASA-TLX questionnaire are illustrated in Fig. 5.

The users' feeling of presence in the real remote environment was measured using the SUS presence questionnaire. Each user's SUS score is calculated by counting the number of 6 or 7 answers in all 6 presence-related questions. This results in each user's SUS score ranging from 0 to 6. The

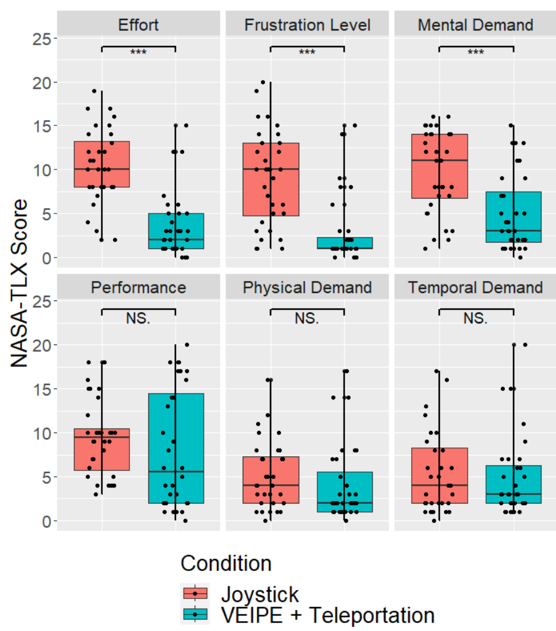

Fig. 5. Comparison of the NASA-TLX individual components, significant results at the top and non-significant results on the bottom

scores were compared using the Wilcoxon-signed-rank (two-sided, alpha = .05) test for matched pairs. The SUS scores of the VEIPE condition ($Mdn = 1, sd = 1.28$) did not differ from the SUS scores of the joystick condition ($Mdn = 1, sd = 1.48$), $Z = 0.89, p = 0.19, r = 0.16$. This indicates no evidence for the VEIPE to affect the user's feeling of presence.

An inductive thematic analysis of the qualitative responses revealed that the most common reason for preferring teleportation over joystick locomotion was that participants found it easier to use with 53 percent of answers including this theme. Delay was seen as the second biggest reason to opt out of the joystick condition with 28 percent of answers including this theme. Finally, lower VR sickness was mentioned in 21 percent of the answers as the reason to choose teleportation. Of participants that preferred the joystick control method, one mentioned that it was annoying to wait for the robot to come to your location in the VEIPE condition and another thought that the switching of the different views between the real remote environment and the VEIPE counterpart was confusing.

Finally, we measured the accumulated yaw of the telepresence robot throughout each condition. In order to make this value meaningful, the accumulated yaw was averaged by the navigation goals reached by the user. This calculation yielded the average yaw per navigation goal reached for each condition. These results were then analyzed using the Wilcoxon-signed-rank (two-sided, alpha = .05) test for matched pairs. The accumulated yaw per goal for the VEIPE condition ($Mdn = 162.31, sd = 91.98$) was found to be significantly lower than the corresponding value for the joystick condition ($Mdn = 285.47, sd = 125.36$), $Z = 3.64, p < 0.001, r = 0.64$.

VI. Discussion

The results show that all our main hypotheses H1, H2 and H3 were supported, demonstrating that in certain tasks using a VE to interface with the physical environment can provide benefits over using a more traditional immersive telepresence setup. Our results suggest that teleportation using VEIPE could be extremely advantageous for immersive telepresence applications where VR sickness is considered a high priority. This result supports the finding that non-continuous locomotion methods are superior to continuous locomotion methods in regards to VR sickness [14]. Furthermore, having users switch between a real panoramic video and a corresponding virtual copy of the world does not seem to cause VR sickness. This, combined with no evidence of degraded presence, suggests that VEIPE as a locomotion method can be a useful way to navigate remote environments with an immersive telepresence robot.

Beyond achieving very low VR sickness levels, we show that VEIPE can improve navigation performance in immersive telepresence applications by around 31 percent, when the actuation-perception loop latency reaches 1.6 s, but we predict that this effect is amplified in higher latency applications. Currently, streaming 360-degree video over the Internet can introduce delays of up to 3 to 4 seconds [48], [49]. Precise control of the robot would become very difficult under those conditions, when using a manual joystick control. A slightly lower improvement in performance was reported in a study of a search and rescue robot, where they compared controlling the robot with and without latency of 0.5 s [27]. This suggests that joystick control could emerge as the more efficient and preferred method, when the delay approaches the tolerable thresholds for interactive applications and teleoperations. Moreover, the search and rescue study also reports a significant increase in physical demand. However, our study did not find the same significant increase in physical demand. This difference might be due to our experiment's task being much shorter in duration, indicating that the same overshooting errors observed did not have sufficient time to compound to exert a great amount of physical demand in the first place. We believe that latency is the biggest component for this performance drop. However, VR sickness can significantly hinder performance as well, as indicated by previous research [50], [51]. More research should be conducted with lower latency immersive telepresence systems as they become available. However, the latency experienced with current commercially available 360-degree cameras will be significant, especially over long distances.

The main improvement suggested by the participants in their open-ended feedback was that they wanted the delay to be significantly lower. This implies a strong need for research to mitigate latency in both immersive and non-immersive telepresence applications. The latency seemed to greatly impact the accuracy of controlling the telepresence robot. The accumulated yaw per reached navigation goal was significantly higher in the joystick-based control. This difference arises from the participants constantly overshooting their control commands to the robot. This finding supports the idea that a predictive display of the robot's movements is needed to help address this problem [31]. Some participants were observed to rely on the "move and wait" technique because of the 1.6 s latency, which replicates previous findings for this behaviour [30].

The participants preferred teleportation through VEIPE as a locomotion method. They reported that it was easier to use and that it had no delay. Moreover, the participants did not prefer the joystick control method because it was more nauseating.

To conclude, when the latency exceeds acceptable thresholds, other predictive displays or some other form of support should be accessible to the operator [30]. VEIPE can be an ideal option for some immersive telepresence applications until the delay can be reduced significantly.

A. Limitations and future research

Our results are limited by a high amount of latency for real-time applications. The latency of 1.6 seconds is nowhere close to usable thresholds for these kinds of real-time navigation controls, but as mentioned before, the current immersive telepresence application delay is normally higher than in our laboratory setting [48]. If latency was significantly lower, the users could have likely performed better against the VEIPE condition. However, conventional immersive telepresence locomotion would still induce vection through optical flow which in turn would most likely cause higher VR sickness anyway [10], [8]. This implies that performance could still be negatively affected, considering the amount of previous literature on VR sickness and performance.

Furthermore, the robot introduced some vibrations to the camera due to not being fully stable in the joystick condition. This might have had an effect on the SSQ score as a compounding factor. In reality, mobile telepresence systems will always have vibrations and instability, so this issue must be considered. The video stream should be stabilized to a great extent, but this might add to the perception-actuation delay. Investigating this trade-off could be an interesting research direction.

The navigation task used in the experiment was relatively simple, as users were only required to navigate towards straightforward and visible goals. Additionally, the navigation goals were not very far apart. Further research should be done with a more complex navigation task and environment. Moreover, if continuous observation of the remote environment is required, a pre-made VE is not feasible since it hides the remote environment view during the robot's navigation. An interesting research topic would be to investigate whether observation in real-time 3D reconstructed VE could match or even surpass the effectiveness of observing through a 360-degree camera.

Finally, the robot had a fixed height for the camera lenses. This height was rather short, at around 110 cm. This viewpoint could affect the perception of the users. However, given the simplicity of the environment, the participants did not seem to pay attention to this. If the task includes height or distance estimation, however, this is a critical factor that should be addressed. Future systems should impose an adjustable height for the camera, keeping in mind that stability should not be compromised.

We recommend future research to further test different locomotion methods through the use of VEIPE and to conduct system usability and user experience research to further develop the system. We aim to develop our system to work in general environments and not just in a lab environment to further validate the robustness of our method. In order to be able to navigate any remote environment, the telepresence robot must be able to reconstruct the remote environment into a 3D virtual environment in real-time, for example by using simultaneous localization and mapping (SLAM) [52]. This would require using depth cameras or AI assisted algorithms, to recreate the remote environment into the VE. Newly entered environments will have inherently incomplete chunks in the VE recreation since the camera's view will be obscured by objects and the system must be able to update the VE while the MBR is moving through the remote environment. This allows the observation of changes in the remote environment through VEIPE. Moreover, it should be able to recognize humans and other moving entities in the remote environment.

ACKNOWLEDGMENT

We would like to thank our colleagues Juho Kalliokoski and Alexis Chambers, and all our participants for supporting this research.

REFERENCES

[1] M. Minsky, "Telepresence," 1980.
[2] A. Kristoffersson, S. Coradeschi, and A. Loutfi, "A review of mobile robotic telepresence," *Advances in Human-Computer Interaction*, vol. 2013, p. 3, 2013.
[3] T. Jokela, J. Ojala, and K. Väänänen, "How people use 360-degree cameras," in *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia*, pp. 1–10, 2019.
[4] T. Aykut, C. Zou, J. Xu, D. Van Opdenbosch, and E. Steinbach, "A delay compensation approach for pan-tilt-unit-based stereoscopic 360 degree telepresence systems using head motion prediction," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3323–3330, 2018.
[5] S. Martirosov, M. Bureš, and T. Zítka, "Cyber sickness in low-immersive, semi-immersive, and fully immersive virtual reality," *Virtual Reality*, vol. 26, no. 1, pp. 15–32, 2022.
[6] A. Kemeny, J.-R. Chardonnet, and F. Colombet, *Getting Rid of Cybersickness: In Virtual Reality, Augmented Reality, and Simulators*. 5 2020.
[7] J. T. Reason and J. J. Brand, *Motion sickness.* Academic press, 1975.
[8] L. Rebenitsch and C. Owen, "Review on cybersickness in applications and visual displays," *Virtual Reality*, vol. 20, no. 2, pp. 101–125, 2016.
[9] K. Stanney, B. D. Lawson, B. Rokers, M. Dennison, C. Fidopiastis, T. Stoffregen, S. Weech, and J. M. Fulvio, "Identifying causes of and solutions for cybersickness in immersive technology: reformulation of a research and development agenda," *International Journal of Human–Computer Interaction*, vol. 36, no. 19, pp. 1783–1803, 2020.
[10] S. M. LaValle, *Virtual reality*. Cambridge university press, 2023.
[11] D. B. Kaber, J. M. Riley, R. Zhou, and J. Draper, "Effects of visual interface design, and control mode and latency on performance, telepresence and workload in a teleoperation task," in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 44, pp. 503–506, 2000.

[12] F. Buttussi and L. Chittaro, "Locomotion in Place in Virtual Reality: A Comparative Evaluation of Joystick, Teleport, and Leaning," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 1, pp. 125–136, 2021.

[13] H. Cherni, N. Métayer, and N. Souliman, "Literature review of locomotion techniques in virtual reality.," *International Journal of Virtual Reality*, 2020.

[14] S. Ang and J. Quarles, "Reduction of cybersickness in head mounted displays use: A systematic review and taxonomy of current strategies," *Frontiers in Virtual reality*, vol. 4, p. 1027552, 2023.

[15] H. Martins and R. Ventura, "Immersive 3-d teleoperation of a search and rescue robot using a head-mounted display," in *2009 IEEE conference on emerging technologies & factory automation*, pp. 1–8, 2009.

[16] J. C. García, B. Patrão, L. Almeida, J. Pérez, P. Menezes, J. Dias, and P. J. Sanz, "A natural interface for remote operation of underwater robots," *IEEE computer graphics and applications*, vol. 37, no. 1, pp. 34–43, 2015.

[17] J. J. Cummings and J. N. Bailenson, "How immersive is enough? A meta-analysis of the effect of immersive technology on user presence," *Media psychology*, vol. 19, no. 2, pp. 272–309, 2016.

[18] G. Makransky and G. B. Petersen, "The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality," *Educational Psychology Review*, vol. 33, no. 3, pp. 937–958, 2021.

[19] P. Breves and J.-P. Stein, "Cognitive load in immersive media settings: the role of spatial presence and cybersickness," *Virtual Reality*, vol. 27, no. 2, pp. 1077–1089, 2023.

[20] A. K. Kanduri, G. Thomas, N. Cabrol, E. Grin, and R. C. Anderson, "The (in) accuracy of novice rover operators' perception of obstacle height from monoscopic images," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 35, no. 4, pp. 505–512, 2005.

[21] E. Akita, G. Zaidner, and M. Pryor, "Improved Situational Awareness and Performance with Dynamic Task-Based Overlays for Teleoperation," in *Proceedings of the 2024 International Symposium on Technological Advances in Human-Robot Interaction*, pp. 65–73, 2024.

[22] T. Kot and P. Novák, "Application of virtual reality in teleoperation of the military mobile robotic system TAROS," *International journal of advanced robotic systems*, vol. 15, no. 1, p. 1729881417751545, 2018.

[23] J. Mayor, L. Raya, and A. Sanchez, "A comparative study of virtual reality methods of interaction and locomotion based on presence, cybersickness, and usability," *IEEE Transactions on Emerging Topics in Computing*, vol. 9, no. 3, pp. 1542–1553, 2019.

[24] D. A. Bowman, D. Koller, and L. F. Hodges, "Travel in immersive virtual environments: an evaluation of viewpoint motion control techniques," in *Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality*, pp. 45–52, 1997.

[25] N. Bakker, P. Passenier, and P. Werkhoven, "Effects of Head-Slaved Navigation and the Use of Teleports on Spatial Orientation in Virtual Environments," *Human factors*, vol. 45, pp. 160–169, 4 2003.

[26] J. C. Lane, C. R. Carignan, B. R. Sullivan, D. L. Akin, T. Hunt, and R. Cohen, "Effects of time delay on telerobotic control of neutral buoyancy vehicles," in *Proceedings 2002 IEEE international conference on robotics and automation (Cat. No. 02CH37292)*, vol. 3, pp. 2874–2879, 2002.

[27] A. Khasawneh, H. Rogers, J. Bertrand, K. C. Madathil, and A. Gramopadhye, "Human adaptation to latency in teleoperated multi-robot human-agent search and rescue teams," *Automation in Construction*, vol. 99, pp. 265–277, 2019.

[28] S. R. Ellis, K. Mania, B. D. Adelstein, and M. I. Hill, "Generalizeability of latency detection in a variety of virtual environments," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 48, pp. 2632–2636, 2004.

[29] S. Neumeier, E. A. Walelgne, V. Bajpai, J. Ott, and C. Facchi, "Measuring the feasibility of teleoperated driving in mobile networks," in *2019 Network Traffic Measurement and Analysis Conference (TMA)*, pp. 113–120, 2019.

[30] J. Y. C. Chen, E. C. Haas, and M. J. Barnes, "Human performance issues and user interface design for teleoperated robots," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1231–1245, 2007.

[31] T. B. Sheridan, "Space teleoperation through time delay: Review and prognosis," *IEEE Transactions on robotics and Automation*, vol. 9, no. 5, pp. 592–606, 1993.

[32] M. E. Walker, M. Gramopadhye, B. Ikeda, J. Burns, and D. Szafir, "The Cyber-Physical Control Room: A Mixed Reality Interface for Mobile Robot Teleoperation and Human-Robot Teaming," in *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 762–771, 2024.

[33] D. Huber, H. Herman, A. Kelly, P. Rander, and J. Ziglar, "Real-time photo-realistic visualization of 3D environments for enhanced teleoperation of vehicles," in *2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops*, pp. 1518–1525, 2009.

[34] M. D. Moniruzzaman, A. Rassau, D. Chai, and S. M. S. Islam, "Teleoperation methods and enhancement techniques for mobile robots: A comprehensive survey," *Robotics and Autonomous Systems*, vol. 150, p. 103973, 2022.

[35] Y. Kato, "A remote navigation system for a simple tele-presence robot with virtual reality," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4524–4529, 2015.

[36] P. Björnfot, "Evaluating Input Devices for Robotic Telepresence," in *Proceedings of the 32nd European Conference on Cognitive Ergonomics*, pp. 1–8, 2021.

[37] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness," *The international journal of aviation psychology*, vol. 3, no. 3, pp. 203–220, 1993.

[38] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," in *Human Mental Workload* (P. A. Hancock and N. Meshkati, eds.), vol. 52 of *Advances in Psychology*, pp. 139–183, North-Holland, 1988.

[39] M. Usoh, E. Catena, S. Arman, and M. Slater, "Using presence questionnaires in reality," *Presence*, vol. 9, no. 5, pp. 497–503, 2000.

[40] M. Slater, M. Usoh, and A. Steed, "Depth of presence in virtual environments," *Presence: Teleoperators & Virtual Environments*, vol. 3, no. 2, pp. 130–144, 1994.

[41] X. Pan, M. Slater, A. Beacco, X. Navarro, A. I. Bellido Rivas, D. Swapp, J. Hale, P. A. G. Forbes, C. Denvir, A. F. de C. Hamilton, and others, "The responses of medical general practitioners to unreasonable patient demand for antibiotics-a study of medical ethics using immersive virtual reality," *PloS one*, vol. 11, no. 2, p. e0146837, 2016.

[42] N. Dużmańska, P. Strojny, and A. Strojny, "Can simulator sickness be avoided? A review on temporal aspects of simulator sickness," *Frontiers in psychology*, vol. 9, p. 410742, 2018.

[43] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, A. Y. Ng, and others, "ROS: an open-source Robot Operating System," in *ICRA workshop on open source software*, vol. 3, p. 5, 2009.

[44] S. Macenski, F. Martín, R. White, and J. G. Clavero, "The marathon 2: A navigation system," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2718–2725, 2020.

[45] S. G. Hart, "NASA-task load index (NASA-TLX); 20 years later," in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50, pp. 904–908, 2006.

[46] Y. Liu and C. D. Wickens, "Mental workload and cognitive task automaticity: an evaluation of subjective and time estimation metrics," *Ergonomics*, vol. 37, no. 11, pp. 1843–1854, 1994.

[47] K. C. Hendy, K. M. Hamilton, and L. N. Landry, "Measuring subjective workload: when is one scale better than many?," *Human Factors*, vol. 35, no. 4, pp. 579–601, 1993.

[48] Z. Yan and J. Yi, "Dissecting Latency in 360 Video Camera Sensing Systems," *Sensors*, vol. 22, no. 16, p. 6001, 2022.

[49] H. Liu, X. Zhang, X. Zhou, T. Zhang, Y. Li, Z. Li, and G. Xie, "When 360-degree video meets 5G: a measurement study of delay in live streaming," in *Proceedings of the 17th ACM Workshop on Mobility in the Evolving Internet Architecture*, pp. 7–12, 2022.

[50] J. M. Mittelstaedt, J. Wacker, and D. Stelling, "VR aftereffect and the relation of cybersickness and cognitive performance," *Virtual Reality*, vol. 23, pp. 143–154, 2019.

[51] K. J. Mimnaugh, E. G. Center, M. Suomalainen, I. Becerra, E. Lozano, R. Murrieta-Cid, T. Ojala, S. M. LaValle, and K. D. Federmeier, "Virtual reality sickness reduces attention during immersive experiences," *IEEE Transactions on Visualization and Computer Graphics*, 2023.

[52] R. C. Smith and P. Cheeseman, "On the representation and estimation of spatial uncertainty," *The international journal of Robotics Research*, vol. 5, no. 4, pp. 56–68, 1986.