The Cyber-Physical Control Room: A Mixed Reality Interface for Mobile Robot Teleoperation and Human-Robot Teaming

Michael E. Walker
University of North Carolina
Chapel Hill, North Carolina, USA
michael.e.walker@unc.edu

Maitrey Gramopadhye
University of North Carolina
Chapel Hill, North Carolina, USA
maitrey@cs.unc.edu

Bryce Ikeda
University of North Carolina
Chapel Hill, North Carolina, USA
bikeda@cs.unc.edu

Jack Burns
University of Colorado Boulder
Boulder, Colorado, USA
jack.burns@colorado.edu

Daniel Szafir
University of North Carolina
Chapel Hill, North Carolina, USA
daniel.szafir@cs.unc.edu

ABSTRACT

In this work, we present the design and evaluation of an immersive Cyber-Physical Control Room interface for remote mobile robots that provides users with both robot-egocentric and robot-exocentric 3D perspectives. We evaluate the Cyber-Physical Control room against a traditional robot interface in a mock disaster response scenario that features a mixed human-robot field team. In our evaluation, we found that the Cyber-Physical Control Room improved robot operator effectiveness by 28% while navigating a complex warehouse environment and performing a visual search. The Cyber-Physical Control Room also enhanced various aspects of human-robot teaming, including social engagement, the ability of a remote robot teleoperator to track their human partner in the field, and opinions of human teammate leadership qualities.

CCS CONCEPTS

• **Human-centered computing** → *Virtual reality*; *Mixed / augmented reality*; *User interface design*; • **Computer systems organization** → **External interfaces for robotics**.

KEYWORDS

Human-robot interaction, robots, field robotics, human-robot teaming, robot teleoperation, immersive displays, virtual reality, mixed reality, interface design, human-computer interaction

ACM Reference Format:
Michael E. Walker, Maitrey Gramopadhye, Bryce Ikeda, Jack Burns, and Daniel Szafir. 2024. 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 (HRI '24), March 11–14, 2024, Boulder, CO, USA*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3610977.3634981

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
HRI '24, March 11–14, 2024, Boulder, CO, USA
© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-0322-5/24/03…$15.00
https://doi.org/10.1145/3610977.3634981

1 INTRODUCTION

As robotic capabilities improve, robots are increasingly supporting human field teams. This is the case for situations in which robots transport specialty equipment or sensors, the environment is hazardous and human presence should be limited (e.g., mobile robots replacing human members of a field team working in radiation zone, search and rescue in disaster environments, etc.), or human presence is costly and highly specialized (e.g., mobile robots teleoperated from orbital platforms to supplement astronaut teams on planetary surfaces).

Although autonomous robotic systems are becoming increasingly robust and able to complete highly complex tasks (e.g., autonomous flying drones, self-driving cars, etc.), situations requiring manual teleoperation or human supervision remain commonplace and will be required for the foreseeable future. These situations include: (1) complex or mission critical tasks that must be executed by expert human operators; (2) robotic systems with insufficient training sets due to inaccessibility or cost; or (3) humans-in-the-loop acting as a fail-safe in case of robotic system failure.

Therefore, it is crucial that interfaces for human operators of remote mobile robots match the improvements seen in modern robot capabilities and sensing. All too often, rich multidimensional data is collected by a field robot, only to be winnowed down into two dimensions on traditional robot interfaces using 2D monitors. This can be seen across industry sectors, even where state-of-the-art robots are deployed in mission critical roles for high-stakes situations such disaster response [9, 12] and space telerobotics for planetary exploration [35]. One way to extend traditional interfaces is by developing mixed reality systems that can fully utilize modern field robot sensor data.

Early efforts to enhance robot teleoperation with virtual reality head-mounted displays (HMDs) saw the creation of 3D immersive interfaces that directly streamed 3D stereo video feeds to the user's eyes. More recently, human-robot interaction researchers have developed two design paradigms that leverage room-scale augmented virtuality environments and decouple robot operators' perspectives from direct-to-eye video streams: Cyber-Physical Interfaces (robot-exocentric perspective via remote environment reconstructions) and Virtual Control Rooms (robot-egocentric perspective via video streams) [31]. To date, interfaces that combine aspects of both design paradigms [3, 39, 58] often targeted stationary robot arms used for manufacturing work and have been limited in terms of the lack

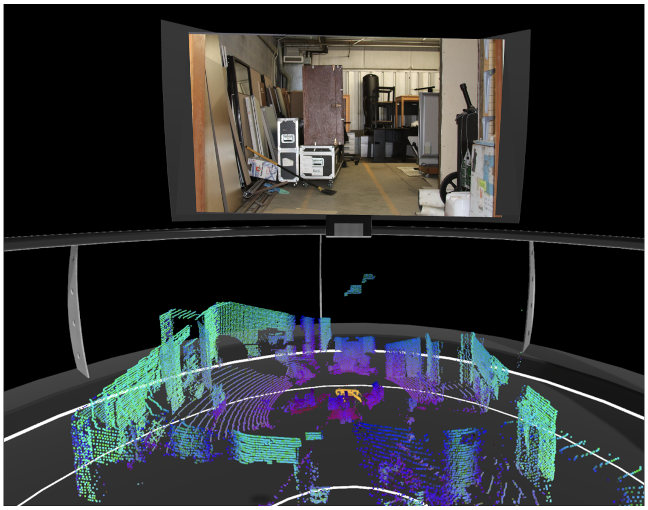

Figure 1: The Cyber-Physical Control Room augmented virtuality HMD interface renders live 3D video streams, a dense 360° 3D point cloud, and a state visualization virtual robot within an immersive virtual environment.

of 3D video streams while being restricted to small-scale and/or offline environmental reconstructions. There has also been early research with this style of interface on humanoid [2, 23] and mobile manipulation robots [30]; however, these interfaces are designed for scenarios without an on-site human presence and make no design considerations for human-robot teaming.

In this work, we present the design and evaluation of an immersive mixed reality interface that is developed for mobile robots that are part of a larger human-robot field team in either known or unknown environments. The interface provides robot operators with live 3D video streams and live large-scale high-resolution dense 3D point clouds for simultaneous viewing of robot-egocentric and robot-exocentric perspectives. We also provide the first deployment and evaluation of a combined perspective mixed reality interface for a mobile robot in a realistic field setting. This evaluation provides insights regarding how such an interface impacts navigation and visual search efficiency as well as various aspects of human-robot teaming between a remote operator and on-site human personnel in a large-scale team-based field experiment.

2 RELATED WORK

Designing interfaces that enable operators to effectively manage and control robots during field deployments has remained an enduring research challenge. Key considerations for such interfaces include how to effectively present users with data about the robot and remote environment (i.e., ensuring sufficient operator situational awareness) and developing appropriate methods for mapping user controls to robot actions (see [53] for a survey). Recently, the modernization of virtual and mixed reality HMDs have enabled promising new methods for enhancing robot operation. For instance, HMDs have enabled streaming stereo video content directly to operators (i.e., showing each eye a separate camera feed), leveraging the human depth cue of stereopsis to enhance immersion and telepresence. Such systems have been deployed on various robotic platforms, including unmanned aerial vehicles [19, 61], unmanned ground vehicles for exploration [41, 51] and manipulation [33], robots for healthcare [18], remotely operated underwater vehicles [7, 49], and humanoid robots [6, 16, 44]. Unfortunately, direct-to-eye video streaming HMD interfaces fully encompass their user's viewpoint and restricts the viewing of non-video sensor data (e.g., environmental maps) and the implementation of additional virtual design elements (VDEs) [57] within the interface (e.g., visualization robots, environment digital twins, etc.). Additionally, robot operators often experience high-levels of nausea due to the abundant proprioceptive system miscues inherent with tethering viewpoints to robots that experience network latencies and do not perfectly match user head and body movements [1, 34].

Researchers in the field of virtual and mixed reality for human-robot interaction (VAM-HRI) have begun to develop alternative designs that attempt to reduce operator nausea and provide additional VDEs absent in earlier direct-to-eye video streaming designs. Two HMD-based mixed reality teleoperation paradigms have positioned themselves at the forefront of this design space and each have shown great promise in enhancing modern robot operator effectiveness: (1) *Virtual Control Rooms*; and (2) *Cyber-Physical Interfaces* [31]. Both design paradigms place users in an *augmented virtuality* environment—real-life imagery rendered within a virtual environment—in which the user's eyes are represented by virtual cameras in the virtual space that move freely with the user's head and body movements. This perspective decoupling helps mitigate nausea caused by communications/hardware delays and/or imperfect mappings between user head motion and robot motion [31].

Virtual Control Rooms place the user within a virtual room that serves as a supervisory command and control center of a remote robot. Within the control room, the user is able to interact with Virtual Control Object VDEs [57] and can view projected 3D stereo video streams that allow users to experience an immersive robot-egocentric perspective with stereoscopic depth. Prior work has explored virtual control rooms for mobile robot teleoperation [27] with Visualization Robot VDEs [57] within the control room to display the remote robot's pose, underwater robot teleoperation [14], and robot manipulation tasks that utilize motion controls and Virtual Control Object VDEs [22, 31].

Cyber-Physical Control Rooms provide a shared mixed reality space (typically with a one-to-one mapping) between: (1) a remote robot and a virtual environment; and (2) a human operator and a virtual environment. Additionally, a 3D reconstruction of the robot's remote environment is rendered within the virtual environment to provide situational context and awareness to the human operator (i.e., a Environment Digital Twin VDE [57]). A virtual robot replica of the remote physical robot is also added to the virtual environment in the same relative location within the virtual environment as in the real remote environment. This virtual robot mimics the remote real robot's state and actions as a Visualization Robot VDE. The user can often interact with the virtual robot to send commands to the remote real robot or visualize the current state or actions being undertaken by the physical robot (changing the Visualization Robot to a Robot Digital Twin VDE [57]). A benefit of this interface paradigm over that of the Virtual Control Room is the robot-exocentric perspective in which users are not restricted to the view from the robot's camera(s) and can observe the environment digital twin and robot digital twin from any viewing angle. This third-person perspective can help mitigate occlusion issues caused by the robot body blocking a camera's field-of-view; however, the sense of immersion of virtually embodying the remote robot is lost. Researchers in this design space have made extensive use of Cyber-Physical interfaces for aiding robotic arm manipulation tasks for manufacturing [8, 43, 54], imitation learning [60], user training [37], manual teleoperation [40, 52, 56], trajectory visualization [45], and remediation of underwater munitions [15]. The prevalence of robot arm-based Cyber-Physical interfaces is primarily due to the stationary nature of the arms and their environments, which reduces the challenge related to recreating the robot's static environment as changes to the scene are small-scale and belong to the robot or manipulated object(s).

Cyber-Physical interfaces for remote mobile robots pose a much greater challenge due to the large-scale environments that should ideally be recreated online at high resolutions in real time. This challenge has limited prior work, which has explored ideas such as low-resolution abstract representations of the mobile robot's environment for urban exploration [4, 25] and planetary exploration [20], offline reconstructions paired with real-time robot and object pose tracking [21, 29], and online SLAM-based RGB 3D model reconstructions, which suffer from numerous large holes in the reconstructed mesh and the inability to capture movement or dynamic changes to the remote robot's environment [50].

Researchers have explored combining robot-egocentric and robot-exocentric perspectives in a single augmented virtuality environment; however, these works do not utilize 3D video streams and are restricted to stationary robotic arms. Small workbench-sized live point clouds have been paired with 2D video feeds either projected on the environment's walls [3] or the robot arms' wrists [58]. Additionally, offline reconstructions of a robot arm's work cell paired with a projected 2D video have also been explored [39]. This work shows promise for robot manipulation tasks, but are not suited for field robots navigating unexplored and dynamic environments.

In our work, we design an immersive Cyber-Physical Control Room interface for remote mobile robots that leverages live HD 3D video streams and 360° room-scale high-resolution 3D reconstructions that can capture, in real time, positions and movements of collocated human teammates anywhere near the remote robot.

3 INTERFACE DESIGN

The design of the augmented virtuality interface began with the creation of a virtual environment that provided enough space for users to comfortably walk around the entirety of a 3D reconstruction of a large indoor room. Additionally, the size of the environment matched the size of the real-life operator environment used in the experiment evaluation. The skybox was shaded black to provide the greatest contrast between the rendered sensor data and the virtual environment background. Stereo video and point cloud data were processed, transmitted, and rendered in real time; therefore, the system did not require prior information about a robot's remote and potentially unknown environment. See Figure 1.

To provide users with a robot-egocentric perspective within the interface, an External Sensor Images and Video VDE [57] was added

Figure 2: Our quadruped robot with an onboard 360° LiDAR and front-facing stereo camera.

to the virtual environment. To make the images 3D to the robot operator, both the left and right video streams, transmitted from a stereo camera mounted on the robot, were rendered within the interface simultaneously. A small horizontal offset was added between the video streams, while the left video stream was masked (i.e., made invisible) to the user's right eye, and the right video stream was masked to the left eye. In this way, each eye only sees one of the offset images, which allowed users to see the video stream as 3D via stereopsis. Additionally, we wanted to ensure robot operators were able to freely view the 3D reconstruction of the environment from any angle without losing sight of the video stream. Therefore, the panels holding the video stream were programmed to slide on the rails that encircled the virtual environment to automatically center in front of the user. See Figures 1 and 3.

An Environment Digital Twin VDE [57] was added to the interface to provide a robot-exocentric perspective within the interface. This digital twin took the form of a dense point cloud, rendered in the center of the virtual environment, that was generated from streamed laser scan data collected by the robot. Users were able to scale the reconstruction up (to allow for more detailed searching) or down (creating a minimap environment summary or top-down bird's eye view) in size. Within the point cloud a Visualization Robot VDE that acted in the role of an External Robot Pose VDE and Robot Location VDE [57]. The Visualization Robot was accurately positioned within the point cloud with an accurate pose that updated in real time to allow users to better understand where the robot is facing and its current operational status. See Figure 3.

The traditional robot interface that served as the baseline in our evaluation was designed to provide the same functionality and visual data as the mixed reality interface. This type of design is representative of commonly used interfaces in modern field robotics [53]. This baseline consisted of two side-by-side windows: (1) the mono video stream; and (2) the dense point cloud. To mimic the ability of viewing the point cloud from any angle (as in the mixed reality interface) a controllable 360° orbit camera was centered on the point cloud and could zoom in and out of the laser scan visualization. Since the Cyber-Physical Control Room design provided users with the ability to control the robot while simultaneously shifting their point-of-view of the point cloud (by moving

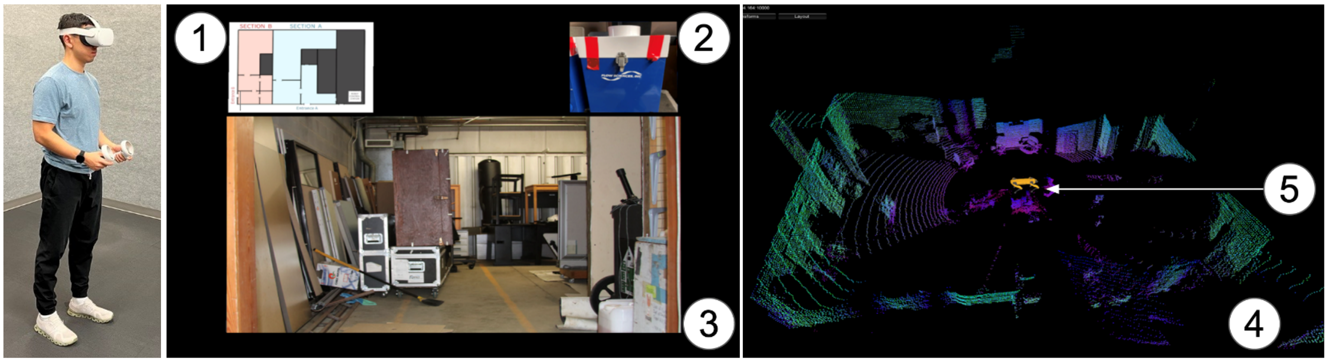

Figure 3: (Left) users wore a virtual reality HMD and two handheld motion controllers. (Middle) [1] warehouse floor plan, [2] image of the target search object, [3] real-time 3D stereo video stream. (Right) [4] real-time 360° dense point cloud, [5] state visualization virtual robot of the Spot embedded within the 3D reconstruction.

their head and body), the baseline interface was designed with similar functionality. Users operated a keyboard with one hand and a mouse with the other hand. The keyboard controlled robot body motion while the mouse controlled the orbit camera in the point cloud window.

4 SYSTEM IMPLEMENTATION

Our fully-implemented Cyber-Physical Control Room interface renders 3D data collected by a Boston Dynamics Spot quadruped robot within an immersive virtual environment. This platform was chosen due to its potential for field robotics, as the robot's four-legged morphology allows it to navigate hazardous terrain inaccessible to wheeled robots. See Figure 2.

The Spot's low FPS grayscale cameras were insufficient for the mixed reality interface; therefore, we mounted a ZED 2i stereo camera [48] to the front of the robot to collect and publish an HD stereo video stream. Additionally, an Ouster OS1 LiDAR [42] was mounted to the top of the robot to collect a 3D point cloud representation of the robot's surroundings. For onboard, online data processing, an NVIDIA Jetson AGX Xavier was also mounted on the Spot. We installed an Ubuntu 20.04 operating system on the Jetson and ran ROS nodes for starting the mounted sensors and establishing communication between the robot and the teleoperation interface. For using the ZED 2i with ROS, we used the ZED ROS wrapper [32]. We used the Ouster SDK [24] for converting the laser scans captured by the LiDAR into a dense 3D point cloud. We used the Ouster ROS driver [38] to access and publish the 3D point cloud as a ROS topic. We employed the Spot ROS Driver [47] to obtain the joint locations of Spot as ROS topics. The Spot ROS Driver also allowed us to control the robot by receiving velocity and pose commands into a ROS subscriber. Finally, the Unity ROS TCP Endpoint [11] allowed the teleoperation interface to communicate with the NVIDIA Jetson AGX Xavier.

We designed the teleoperation interfaces with the Unity engine [55]. Participants using the Cyber-Physical Control Room used the Meta Quest 2 [36] HMD to view and interact with the virtual environment. The robot was remotely controlled in real-time via velocity and pose commands sent from the Quest's handheld controllers (or keyboard if in the Baseline condition).

A WiFi mesh, that connected both interface and robot, was installed at the experiment site comprising of one base router and two mesh nodes, mimicking mesh network setups that may be used in field robotics [13, 17]. The camera streamed compressed HD stereo images from the left and the right cameras at 60 FPS. The 3D point cloud transmitted was calculated from laser scans obtained at a resolution of 1024 points for each height, at a frequency of 10Hz. These settings provided the interfaces with high-resolution data, while maintaining real-time data streaming and rendering.

5 FIELD EXPERIMENT EVALUATION

Based on prior studies of mixed reality interfaces (see §2), we hypothesize that robot operators provided with both a live robot-egocentric perspective and live dynamic room-scale exocentric perspective in a 3D immersive interface will outperform users of traditional robot interfaces that utilizes the same visualizations on a 2D display. We further predict that the immersive interface will not only allow for better simultaneous observation of both robot perspectives, but the immersive nature of the interface will improve aspects of human-robot teaming as well.

To evaluate these hypotheses, we conducted a 2×1 between-subjects experiment to evaluate how our mixed reality Cyber-Physical Control Room design influences operator performance and user experiences during remote robot teleoperation. This design was compared against a traditional robot interface presented on a 2D display, representing the current industry standard for state-of-the-art robotic missions and interfaces [53].

In this experiment, participants were tasked with teleoperating a robot to navigate a cluttered warehouse environment, with the goal of locating objects of interest. During this task, participants operated a remote robot that worked alongside a human teammate in the mock disaster zone, thereby emulating a mixed human-robot field team [28]. With this setup, we evaluated: (1) human-robot field team internal interaction efficiencies; (2) social engagement between robot operator and human field team members; and (3) the robot operator's perceptions of the human field team member's influence and leadership qualities.

We recruited a total of 24 participants (16 males and 8 females, evenly balanced across conditions) from a university campus and surrounding community to take part in our study, approved by the university's IRB. Average participant age was 32.1 (SD = 13.9). On a seven-point scale, participants reported having a moderately high familiarity with virtual reality (M = 5.1, SD = 2.02), possibly reflecting the increasing popularity and pervasiveness of such technologies in today's society.

5.1 Experimental Task

We designed a mock disaster scenario in which a hurricane had damaged a chemical storage facility. To limit human exposure to the hazardous site, a mixed human-robot team was assembled to act as first responders. This team consisted of three members: (1) a human fieldworker, acted out by an experimental confederate, working on-site at the facility; (2) a mobile ground robot working on-site at the facility; and (3) a remote participant that remotely teleoperated the robot.Both team members independently moved through the warehouse to complete their own tasks. The teams' objectives were to: (1) investigate the buildings' safety (structural stability, risk of electrical fires, etc.); (2) ensure there are no chemical leaks; and (3) locate objects of interest within the building for extraction.

The participant was tasked with teleoperating the robot to find objects of interest throughout the building. In a cohesive disaster response narrative, mission command would inform the participant about which objects to locate sequentially. The objects were: (1) a chemical container; (2) electrical switchboard; (3) fire extinguisher; (4) furnace; (5) box of fuses; and (6) computer hard drive. While the robot operator (i.e., experimental participant) searched for objects, the human fieldworker (i.e., experimental confederate) scanned the building's foundations and electrical signals with a handheld device to analyze the current 'on-the-ground' situation, report back to mission command, and receive further instructions based on the information sent. Each time the participant located an object, they had to then search for and locate their fieldworker partner, who was also independently moving throughout the working area, to discuss the next course of action (i.e., what object to find next within the mock disaster narrative). This search and report process repeated six times until all objects were been found which marked the area as secured for the secondary cleanup team.

5.1.1 *Experiment Environment.*

The mock first response scenario took place at an abandoned storage warehouse. The facility was chosen due to its size, complex layout, densely cluttered environment, and dilapidated state that resembled a real-life disaster zone. The experiment was restricted to a 3000 sq ft area on the first floor of the multi-level warehouse. The robot control station was held in a separate area of the warehouse that had it's own outdoor entrance and was isolated from the task working areas by a soundproof metal door. See Figure 3 for an image of one of the rooms within the operational area of the experiment that displays the environment state, size, and clutter.

5.2 Procedure

Participants were randomly assigned to one of the two conditions: Cyber-Physical Control Room or Baseline, with 12 participants in each condition. Participants were prevented from entering or seeing any part of the areas they would be operating the robot in. This was to ensure that all participants completed the task within an unknown environment. Participants were situated in a robot control station isolated from the robot task space by soundproof walls. Additionally, participants were prevented from seeing or meeting their human fieldworker teammate until the training phase was complete and the mock disaster response scenario had begun. This was done to prevent the introduction of confounding priming factors regarding interactions and/or perceptions of the human teammate.

Next, participants watched a training video for their assigned interface to make sure they were properly and uniformity trained. Then, the participants piloted the real robot through a sequential checklist of drills outside of the warehouse. By following these uniform set of procedures, operators of both interfaces were able to demonstrate full proficiency with controlling the robot before the start of the task.

To mark the beginning of the task (and the first of seven *social phases*), the human fieldworker teammate (an experimental confederate) would open the door to the warehouse operational area and introduce themselves as the robot operator's partner for the mission. To provide consistency across both experimental conditions, the fieldworker confederate followed a singular and extensive script to interact with the participants. The same actor was used for all 24 participants and they wore the same orange safety vest in all trials to prevent potential social confounds and keep every experiment trial as uniform as possible.

Upon making introductions, the human fieldworker teammate briefed the participant on the disaster response scenario as well as the roles that command had assigned the two team members. Prior to every *object search phase*, in both interface conditions, an image of the object to be located was uploaded to the teleoperation interface adjacent to the floor plan (see Figure 3). The human fieldworker teammate locked their gaze on the robot's camera during all social phases, regardless of whether or not the robot operator returned the gesture of eye contact.

While the robot operator performed their search, the human fieldworker would follow a predefined path, unknown to participants but consistent across trials. This path was within the experimental working area where the human fieldworker completed their task of structural and electrical scanning. The scanning device was a handheld tablet that secretly contained the robot's emergency stop and control override as a safety backup (neither of which needed to be used during the experiment).

The object search order and hidden object locations were deliberately chosen so that the robot needed to leave the location where their human teammate was working. This enabled the human fieldworker to move (without being seen by the robot's forward-facing camera as it exited the room) to their predefined search-phase positions, which were in blind spots of the robot's camera (e.g., a side room, a side closet, a corner of a room, etc.). This procedure was designed to reveal if human operators were able to simultaneously absorb information from both robot's perspectives (egocentric and exocentric) and notice with the point cloud, while navigating with the video stream, either: (1) when the human teammate moved to their new position in the blind spot while leaving their partner's current room; or (2) the robot operator working in the robot camera's blind spot upon returning to their partner's last known

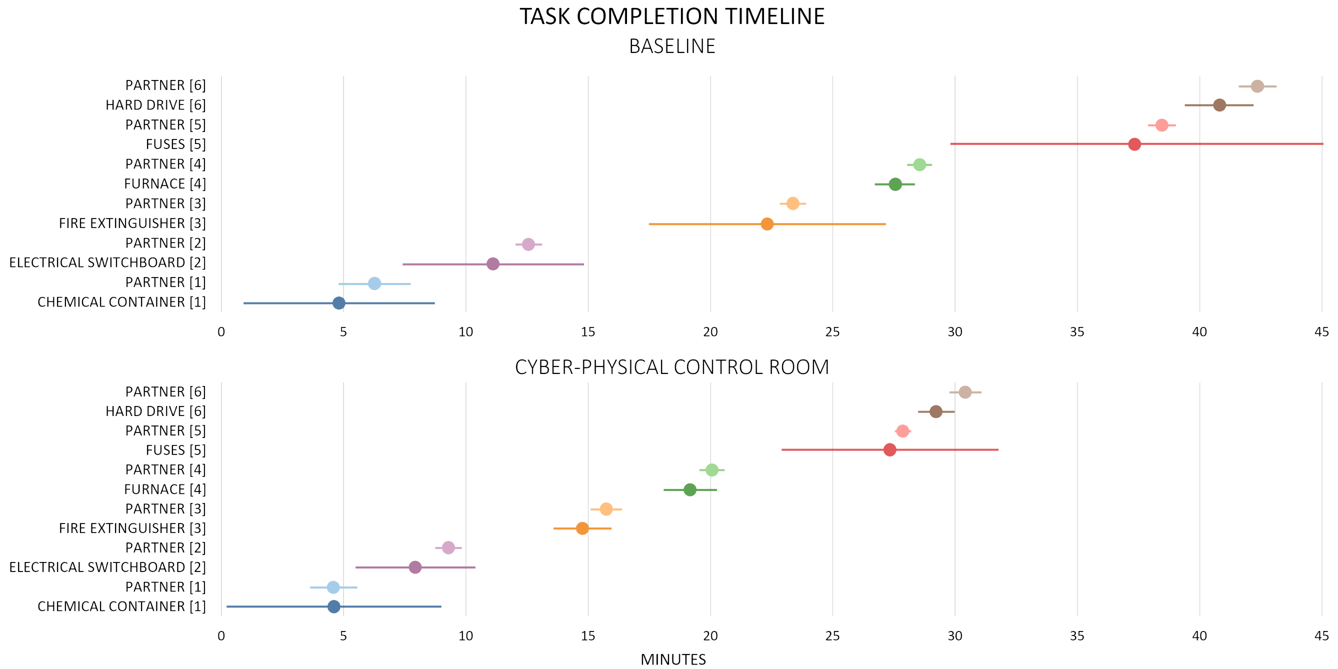

Figure 4: Objective results depicting the time taken to find each object and human teammate for each leg of the experiment (error bars encode standard deviation). Overall task performance was significantly faster in the Cyber-Physical Control Room condition (M = 30.4 minutes, SD = 8.9 minutes) over the Baseline (M = 42.4 minutes, SD = 17.6 minutes), F(1, 22) = 4.47, p = 0.046.

location. The experiment ended when the participant located and reported back to the human fieldworker teammate after finding the final, sixth object. The participant then filled out a survey and participated in a semi-structured interview regarding their experience.

5.3 Measures and Analysis

All objective and subjective measures were analysed using a one-way analysis of variance (ANOVA) with experimental condition (i.e., interface design) as a fixed effect, except for social engagement (measured via gaze frequency as described below), which was analyzed using a Chi-squared test.

5.3.1 *Objective Measures.*

Our primary objective measure was overall *task performance*, measured by how long it took each participant to complete the task. We also examined task sub-components of finding each object and then finding their human teammate partner afterwards.

Beyond standard performance measures, we were also interested in examining the social engagement between robot operator and human fieldworker teammate. Enhancing the social aspects of human-robot teams is a critical objective for interface designs in order to facilitate integration within real-world workplaces. Prior research in HRI has shown that eye contact has a direct impact on social engagement between humans and robots [26] and can impact perceptions of teammates such as trustworthiness and friendliness [59]. Therefore, we measured social engagement by tracking whether or not the robot operator manipulated the robot's body so that the robot camera maintained eye contact with their human partner during the social phases of the task. To make eye contact in this manner requires a deliberate, conscious choice by the operator who, to make eye contact, would need to press and hold the button/joystick for the robot to look up (from its standard knee-level point-of-view) at the human, which eliminates the risk of false positives during results analysis.

5.3.2 *Subjective Measures.*

After using their randomly-assigned interface, participants evaluated perceived interface usability using the System Usability Scale (SUS), an industry standard ten-item attitude survey [5]. SUS scores below 68 are considered below average, scores above 68 are considered above average, and scores above 80.3 are considered in the top 10th percentile. The Interpersonal Dominance Scale (IDS) [10], which consists of 32 statements for which participants provide a seven-point agreement rating, was also administered to measure perceived dominance of the human fieldworker teammate over the robot operator. In addition to the SUS and IDS, we constructed a number of scales from 7-point Likert-style questionnaire items we created to measure participant perceptions of their experience using the teleoperation interface during the task. Scales rated *perceived confidence* (2 items, Cronbach's α = .87) and *perceived task performance* (3 items, Cronbach's α = .65) (scales were constructed according to established methodology [46]). Qualitative feedback was also obtained through open-ended questions in the questionnaire and during a semi-structured interview.

Table 1: Results for each individual object and partner search subtask (in minutes).

	F-VALUE	P-VALUE	CYBER PHYSICAL CONTROL ROOM	BASELINE	PERCENT IMPROVEMENT
[1] OBJECT	F(1, 22) = 0.83	p =.37	M = 3.21, SD = 4.60	M = 4.78, SD = 3.88	32.8%
[2] OBJECT	F(1, 22) = 1.41	p =.24	M = 3.34, SD = 2.51	M = 4.88, SD = 3.72	31.5%
[3] OBJECT	F(1, 22) = 9.19	p < .01**	M = 5.42, SD = 1.16	M = 9.75, SD = 4.81	44.4%
[4] OBJECT	F(1, 22) = 3.91	p = .06	M = 3.43, SD = 1.11	M = 4.20, SD = 0.79	18.4%
[5] OBJECT	F(1, 22) = 0.44	p =.50	M = 7.29, SD = 4.43	M = 8.96, SD = 7.57	18.7%
[6] OBJECT	F(1, 22) = 4.99	p = .03*	M = 1.35, SD = 0.74	M = 2.37, SD = 1.40	43.2%
[1] PARTNER	F(1, 22) = 0.02	p = .89	M = 1.42, SD = 0.95	M = 1.48, SD = 1.49	4.5%
[2] PARTNER	F(1, 22) = 0.21	p = .65	M = 1.35, SD = 0.53	M = 1.45, SD = 0.55	6.9%
[3] PARTNER	F(1, 22) = 0.12	p = .74	M = 1.00, SD = 0.64	M = 1.08, SD = 0.53	7.4%
[4] PARTNER	F(1, 22) = 0.00	p = .99	M = 0.90, SD = 0.51	M = 0.91, SD = 0.52	0.3%
[5] PARTNER	F(1, 22) = 5.64	p = .02*	M = 0.55, SD = 0.30	M = 0.97, SD = 0.54	43.8%
[6] PARTNER	F(1, 22) = 2.11	p = .16	M = 1.16, SD = 0.66	M = 1.58, SD = 0.77	26.8%
TOTAL TIME	**F(1, 22) = 4.47**	**p = .04***	**M = 30.4, SD = 8.90**	**M = 42.4, SD = 17.6**	**28.3%**

6 RESULTS

6.1 Objective Results

We analyzed task performance and social engagement metrics to determine if the Cyber-Physical Control Room design helped participants teleoperate a remote robot more effectively within a mixed human-robot team. We found a significant main effect of interface design on task completion time, F(1, 22) = 4.47, p = 0.046, with the Cyber-Physical Control Room (M = 30.4, SD = 8.9 minutes) improving completion time 28% over the Baseline interface (M = 42.4, SD = 17.6 minutes). We further analyzed each subtask and found significant main effects of interface design on multiple individual object and partner search times (see Table 1).

We also found a significant main effect of interface design on social engagement, $X^2(1,168) = 49.40$, p < .001, with participants that used the Cyber-Physical Control Room making eye-contact substantially more frequently (71 instances) compared to the Baseline interface (26 instances). Examining gaze frequency across social phases of the experiment, we observed that social engagement increased during the task for participants using the Cyber-Physical Control Room, but decreased for participants in the Baseline condition (see Figure 5).

6.2 Subjective Results

We found a significant main effect of interface design on the remote robot operator viewing the human teammate as a decision maker (Interpersonal Dominance Scale item), F(1,22) = 4.82, p = 0.039, with the Cyber-Physical Control Room (M = 5.83, SD = 1.11) enhancing the perception of this leadership quality over the baseline interface (M = 4.5, SD = 1.78). See Figure 5.

A significant main effect was found of interface design on perceived confidence while performing the task, F(1,22) = 5.1, p = 0.034, with the Cyber-Physical Control Room (M = 5.94, SD = 1.16) improving user confidence over the baseline interface (M = 5.00, SD = 0.85). We did not find a significant effect of interface design on user perceived task performance, F(1,22) = 3.33, p = 0.082 between the Cyber-Physical Control Room (M = 5.42, SD = 0.85) and baseline interface (M = 4.64, SD = 1.2).

Interface usability was evaluated with the SUS. We did not find a significant main effect of interface design on usability score, F(1,22) = 0.19, p = 0.67, between the Cyber-Physical Control Room (M = 79.8, SD = 13.6) and the baseline interface (M = 77.3, SD = 14.7). These similar scores provide both interfaces with descriptive adjectives of 'good' and 'acceptable' per SUS grading guidelines.

7 DISCUSSION

The results of our study indicate that the immersive Cyber-Physical Control Room provides significant improvements over the baseline interface in various aspects of object search efficiency, human-robot teaming, and user experience. We did not find support for our usability hypothesis that predicted perceived usability (SUS scores) to be higher for the Cyber-Physical Control Room. However, as both interfaces scored similarly in terms of perceived usability via the SUS (both systems receiving a rating of 'good' and 'acceptable'), meaning participants found both interfaces comparable in terms of ease-of-use, we believe that the results from this study are derived from differences in interface design and not due to usability-related advantages that could potentially make one interface easier to use.

7.1 Navigation and Visual Search Efficiency

When looking at the team-based mock disaster task as a whole, the mixed reality design outperformed the baseline interface in terms of overall efficiency, with operators who used the Cyber-Physical Control Room interface 28.3% more time efficient than participants that utilized the traditional baseline system. Additionally, the search time variance is much smaller in the mixed reality condition, pointing to a higher consistency in operator performance when locating the objects in the complex warehouse environment. See Figure 4.

7.2 Human-Robot Teaming

A goal of this work was to evaluate not only operator performance but explore if interface design could impact aspects of human-robot teaming. An evaluation of this kind in the context of a mobile mixed human-robot team with state-of-the-art immersive mixed reality interfaces, such as our Cyber-Physical Control Room, has not previously been explored. In this work, we found strong evidence suggesting that immersive mixed reality robot interfaces improve not only teleoperation performance but also critical social aspects of human-robot teaming, such as tracking and maintaining a mental model of teammate locations, telepresent social engagement, and robot operator perceptions of human teammates in the field.

7.2.1 *Teammate Position Mental Model.*

In the evaluation, the robot operator's ability to to maintain an accurate mental model of their human teammate's position was tested during the partner search phases of the experiment. The experimental design ensured that users were able to locate their partner if they watched the real-time LiDAR rendering within the interface. The Cyber-Physical Control Room design allowed users to make better use of both the robot egocentric and exocentric perspectives, compared to traditional robotic interfaces that utilize 2D displays. This suggests that immersive mixed reality interfaces

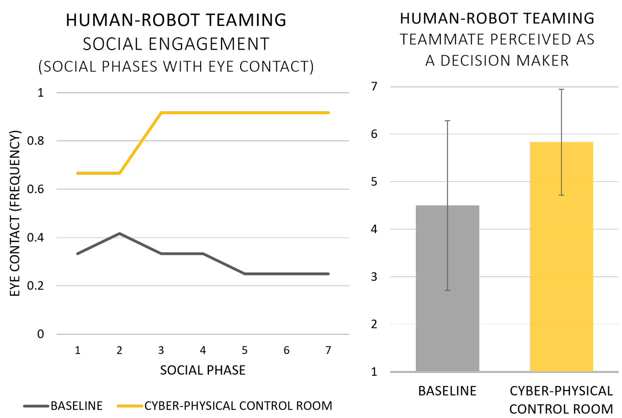

Figure 5: Results depicting (left) eye contact made by robot operators used to measure social engagement during the seven experiment social phases and (right) robot operator opinions of their teammate's decision making capability.

may allow users to more effectively leverage multi-perspective spatiotemporal data during teleoperation tasks See Figure 4.

7.2.2 *Social Engagement.*

To evaluate conversational engagement, both the experiment administrator and the experiment confederate independently recorded eye contact if the participant directed the robot camera toward the human teammate's face during the active social phase. Social phases were marked as containing eye contact if both experimenters' records agreed (100% agreement was found). Interestingly, across all participants during the entire study, we found that this eye contact was "all or nothing" in which the operator either maintained mutual gaze through the robot's camera the entire social phase or not at all (i.e., there was not a single instance of intermittent eye contact). It was found that instances of eye contact were 173% higher in the case of the Cyber-Physical Control Room interface compared to the Baseline. Moreover, we found a divergence in the results during the duration of the experiment. In the Cyber-Physical Control Room condition, eye contact made by the robot operator increased over time, to the point that all but one participant consistently made eye contact with their partner during the social phases of the task from task segment three and on. Conversely, social engagement decreased during the experiment with users assigned to the baseline condition. This finding suggests that remote robot operators' team interactions within mixed human-robot field teams can be enhanced with immersive mixed reality telepresence interfaces, such as our Cyber-Physical Control Room. Robot operators appear to be more engaged with their teammates and experience more natural social interactions that follow human social norms (e.g., eye contact during one-on-one discussion) as if they were in-person with the field team themselves, than if they were to use a traditional robot teleoperation interface. See Figure 5.

7.2.3 *Perception of Human Teammate.*

In addition to enhancing social engagement, the results from the IDS revealed that robot operators of the Cyber-Physical Control Room interface perceived their human fieldworker teammate as a someone people turn to for making important decisions (See Figure 5.). This sentiment reflects a major aspect of leadership, which is a critical component of human teaming and field operations. The effects that interface design and leadership role have on human-robot team interactions are of particular importance in contexts where the human team member's ability to exert authority are critical to the success of the interaction and mission. To our knowledge, this is the first evidence towards the idea that immersive mixed reality interfaces may impact the perceived leadership qualities (in this case, decision-making abilities) of human-robot team members. This finding opens the door to future research that further explores how immersive mixed reality interfaces might enhance and/or manipulate the perceived personality traits within human-robot teams to enhance group cohesion and performance.

8 LIMITATIONS

Although our work has shown promising results in regards to remote mobile robotic teleoperation, it is not without limitations. For instance, our field experiment could not create the degrees of stress that real emergency response missions would likely involve. Future work might objectively capture operator stress via physiological measurements and examine potential links between interface designs, stress, and performance. Additionally, while our users did complete a training procedure prior to their task, robot operators in real-world response scenarios would presumably have substantially more experience and extensive practice. Thus, more work is needed to investigate the effects of learning and expertise. Finally, our work would benefit from further investigations of the potential cognitive burden that egocentric and exocentric coordinate frame switching might place on users.

9 CONCLUSION

In this work, we provide the first deployment and evaluation of a Cyber-Physical Control Room interface for a mobile robot in a realistic field setting. We investigated the effects of interface design on remote robot operator effectiveness while navigating a complex warehouse environment, performing an environmental search, and working with a human fieldworker that was collocated with the robot. The Cyber-Physical Control Room enhanced teleoperation performance and various aspects of teaming, including conversational engagement, teammate positional tracking, and opinions of teammate leadership qualities. These results reveal that an immersive mixed reality interface holds advantages over traditional teleoperation methods for mobile robot teleoperation systems that provide simultaneous robot-egocentric and robot-exocentric perspectives of the remote robot's environment. Additionally, the findings within this work point to a new avenue of research, where mixed reality teleoperation interfaces are used not only as a means to improve task efficiencies, but to strengthen social elements and bonds within human-robot teams.

ACKNOWLEDGMENTS

This work was directly supported by NSF Award #1764092 and the NASA Solar System Exploration Research Virtual Institute (SSERVI) Cooperative Agreement 80ARCC017M0006.

REFERENCES

[1] Wadhah Al-Ashwal, Houshyar Asadi, Shady Mohamed, Shehab Alsanwy, Lars Kooijman, Darius Nahavandi, Ahmad Abu Alqumsan, and Saeid Nahavandi. 2021. Cybersickness measurement and evaluation during flying a helicopter in different weather conditions in virtual reality. In *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2152–2157.

[2] Jordan Allspaw, Gregory LeMasurier, and Holly Yanco. 2021. Implementing Virtual Reality for Teleoperation of a Humanoid Robot. *arXiv preprint arXiv:2104.11826* (2021).

[3] Jordan Allspaw, Jonathan Roche, Nicholas Lemiesz, Michael Yannuzzi, and Holly A Yanco. 2018. Remotely teleoperating a humanoid robot to perform fine motor tasks with virtual reality–. In *Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI)*.

[4] Peter Amstutz and Andrew H Fagg. 2002. Real time visualization of robot state with mobile virtual reality. In *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No. 02CH37292)*, Vol. 1. IEEE, 241–247.

[5] Aaron Bangor, Philip T Kortum, and James T Miller. 2008. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction* 24, 6 (2008), 574–594.

[6] Maxwell Bennett, Tom Williams, Daria Thames, and Matthias Scheutz. 2017. Differences in interaction patterns and perception for teleoperated and autonomous humanoid robots. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 6589–6594.

[7] Josep Bosch, Pere Ridao Rodriguez, Rafael Garcia, and Nuno Gracias. 2016. Towards omnidirectional immersion for ROV teleoperation. In *XXXVII Jornadas de Automática*. Comité Español de Automática, 293–299.

[8] Filippo Brizzi, Lorenzo Peppoloni, Alessandro Graziano, Erika Di Stefano, Carlo Alberto Avizzano, and Emanuele Ruffaldi. 2017. Effects of augmented reality on the performance of teleoperated industrial assembly tasks in a robotic embodiment. *IEEE Transactions on Human-Machine Systems* 48, 2 (2017), 197–206.

[9] Maria Bualat, William Carey, Terrence Fong, Kim Nergaard, Christopher Provencher, Andre Schiele, Philippe Schoonejans, and Ernest Smith. 2014. Preparing for crew-control of surface robots from orbit. In *Space Exploration Conference*.

[10] Judee K Burgoon, Michelle L Johnson, and Pamela T Koch. 1998. The nature and measurement of interpersonal dominance. *Communications Monographs* 65, 4 (1998), 308–335.

[11] Laurie Cheers, Devin M., Amanda, Hamid Younesy, Shuo Diao, Peter Smith, Stephan Hasler, and Tiffany Yau. [n. d.]. ROS TCP Endpoint. https://github.com/Unity-Technologies/ROS-TCP-Endpoint.

[12] Manolis Chiou, Georgios-Theofanis Epsimos, Grigoris Nikolaou, Pantelis Pappas, Giannis Petousakis, Stefan Mühl, and Rustam Stolkin. 2022. Robot-assisted nuclear disaster response: Report and insights from a field exercise. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 4545–4552.

[13] Ludovico Ferranti, Salvatore D'Oro, Leonardo Bonati, Emrecan Demirors, Francesca Cuomo, and Tommaso Melodia. 2019. Hiro-net: Self-organized robotic mesh networking for internet sharing in disaster scenarios. In *2019 IEEE 20th International Symposium on" A World of Wireless, Mobile and Multimedia Networks"(WoWMoM)*. IEEE, 1–9.

[14] Juan C García, Bruno Patrão, Luís Almeida, Javier Pérez, Paulo Menezes, Jorge Dias, and Pedro J Sanz. 2015. A natural interface for remote operation of underwater robots. *IEEE computer graphics and applications* 37, 1 (2015), 34–43.

[15] Zaid Gharaybeh, Howard Chizeck, and Andrew Stewart. 2019. *Telerobotic control in virtual reality*. IEEE.

[16] Liang Gong, Changyang Gong, Zhao Ma, Lujie Zhao, Zhenyu Wang, Xudong Li, Xiaolong Jing, Haozhe Yang, and Chengliang Liu. 2017. Real-time human-in-the-loop remote control for a life-size traffic police robot with multiple augmented reality aided display terminals. In *2017 2nd International Conference on Advanced Robotics and Mechatronics (ICARM)*. IEEE, 420–425.

[17] Abraham Hart, Narek Pezeshkian, and Hoa Nguyen. 2012. Mesh networking optimized for robotic teleoperation. In *Unmanned Systems Technology XIV*, Vol. 8387. SPIE, 523–529.

[18] Satoshi Hashizume, Ippei Suzuki, Kazuki Takazawa, Ryuichiro Sasaki, and Yoichi Ochiai. 2018. Telewheelchair: The remote controllable electric wheelchair system combined human and machine intelligence. In *Proceedings of the 9th Augmented Human International Conference*. 1–9.

[19] Keita Higuchi and Jun Rekimoto. 2013. Flying head: a head motion synchronization mechanism for unmanned aerial vehicle control. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. 2029–2038.

[20] Butler Hine, Phil Hontalas, Terrence Fong, Laurent Piguet, Erik Nygren, and Aaron Kline. 1995. VEVI: A virtual environment teleoperations interface for planetary exploration. *SAE transactions* (1995), 615–628.

[21] Roman Ibrahimov, Evgeny Tsykunov, Vladimir Shirokun, Andrey Somov, and Dzmitry Tsetserukou. 2019. Dronepick: Object picking and delivery teleoperation with the drone controlled by a wearable tactile display. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1–6.

[22] Inmo Jang, Joaquin Carrasco, Andrew Weightman, and Barry Lennox. 2019. Intuitive bare-hand teleoperation of a robotic manipulator using virtual reality and leap motion. In *Towards Autonomous Robotic Systems: 20th Annual Conference, TAROS 2019, London, UK, July 3–5, 2019, Proceedings, Part II 20*. Springer, 283–294.

[23] Steven Jens Jorgensen, Murphy Wonsick, Mark Paterson, Andrew Watson, Ian Chase, and Joshua S Mehling. 2022. *cockpit interface for locomotion and manipulation control of the NASA valkyrie humanoid in virtual reality (VR)*. Technical Report.

[24] kairenw, Ussama Naal, Daniel D'Souza, Pavlo Bashmakov, Michael A. Alcorn, Morten Fyhn Amundsen, Krishtof Korda, Andrey Stepanov, Behrooze Sirang, Federico Proni, Philipp Bandow, and James Snyder. [n. d.]. Ouster SDK. https://github.com/ouster-lidar/ouster_example.

[25] Seungwoon Kim, Yihan Kim, Jeesoo Ha, and Sungho Jo. 2018. Mapping system with virtual reality for mobile robot teleoperation. In *2018 18th International Conference on Control, Automation and Systems (ICCAS)*. IEEE, 1541–1541.

[26] Kyveli Kompatsiari, Francesca Ciardo, Davide De Tommaso, and Agnieszka Wykowska. 2019. Measuring engagement elicited by eye contact in Human-Robot Interaction. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 6979–6985.

[27] Tomáš Kot and Petr Novák. 2014. Utilization of the Oculus Rift HMD in mobile robot teleoperation. *Applied Mechanics and Materials* 555 (2014), 199–208.

[28] Geert-Jan M Kruijff, Ivana Kruijff-Korbayová, Shanker Keshavdas, Benoit Larochelle, Miroslav Janíček, Francis Colas, Ming Liu, Francois Pomerleau, Roland Siegwart, Mark A Neerincx, et al. 2014. Designing, developing, and deploying systems to support human–robot teams in disaster response. *Advanced Robotics* 28, 23 (2014), 1547–1570.

[29] Robert Ladig and Kazuhiro Shimonomura. 2019. High Precision, Intuitive Teleoperation of Multiple Micro Aerial Vehicles Using Virtual Reality. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. IEEE, 2633–2638.

[30] Gregory LeMasurier, Jordan Allspaw, and Holly A Yanco. 2021. Semi-autonomous planning and visualization in virtual reality. *arXiv preprint arXiv:2104.11827* (2021).

[31] Jeffrey I Lipton, Aidan J Fay, and Daniela Rus. 2017. Baxter's homunculus: Virtual reality spaces for teleoperation in manufacturing. *IEEE Robotics and Automation Letters* 3, 1 (2017), 179–186.

[32] Walter Lucetti, Aymeric Dujardin, Raffaello Bonghi, Eric Wieser, Yuki Furuta, Rousseau Vincent, Kevin Jaget, Affonso Guilherme, Faraz Khan, Matt Rousseau, Derek Knowles, Owen Burns, Chris Iverach-Brereton, Błażej Sowa, Tony Baltovski, Jamy Chahal, Moju Zhao, Vishal Prabhu, Vinny Do, Wujun Zhou, and Simone Nardi. [n. d.]. ZED ROS wrapper. https://github.com/stereolabs/zed-ros-wrapper.

[33] Mateusz Maciaś, Adam Dąbrowski, Jan Fraś, Michał Karczewski, Sławomir Puchalski, Sebastian Tabaka, and Piotr Jaroszek. 2020. Measuring performance in robotic teleoperation tasks with virtual reality headgear. In *Automation 2019: Progress in Automation, Robotics and Measurement Techniques*. Springer, 408–417.

[34] Sergo Martirosov, Marek Bureš, and Tomáš Zítka. 2022. Cyber sickness in low-immersive, semi-immersive, and fully immersive virtual reality. *Virtual Reality* 26, 1 (2022), 15–32.

[35] Fumitoshi Matsuno, Noritaka Sato, Kazuyuki Kon, Hiroki Igarashi, Tetsuya Kimura, and Robin Murphy. 2014. Utilization of robot systems in disaster sites of the great eastern japan earthquake. In *Field and Service Robotics: Results of the 8th International Conference*. Springer, 1–17.

[36] Meta. [n. d.]. Meta Quest 2. https://www.meta.com/quest/products/quest-2.

[37] Nadine E Miner and Sharon A Stansfield. 1994. An interactive virtual reality simulation system for robot control and operator training. In *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*. IEEE, 1428–1435.

[38] Ussama Naal, kairenw, Morten Fyhn Amundsen, Jack Geissinger, Krishtof Korda, Andrey Stepanov, Daniel Souza, Federico Proni, and Philipp Bandow. [n. d.]. Ouster ROS driver. https://github.com/ouster-lidar/ouster-ros.

[39] Abdeldjallil Naceri, Dario Mazzanti, Joao Bimbo, Domenico Prattichizzo, Darwin G Caldwell, Leonardo S Mattos, and Nikhil Deshpande. 2019. Towards a virtual reality interface for remote robotic teleoperation. In *2019 19th International Conference on Advanced Robotics (ICAR)*. IEEE, 284–289.

[40] Dejing Ni, Andrew YC Nee, Soh-Khim Ong, Huijun Li, Chengcheng Zhu, and Aiguo Song. 2018. Point cloud augmented virtual reality environment with haptic constraints for teleoperation. *Transactions of the Institute of Measurement and Control* 40, 15 (2018), 4091–4104.

[41] Yeonju Oh, Ramviyas Parasuraman, Tim McGraw, and Byung-Cheol Min. 2018. 360 VR based robot teleoperation interface for virtual tour. In *Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI)*.

[42] Ouster. [n. d.]. Ouster OS1 LiDAR. https://ouster.com/products/scanning-lidar/os1-sensor.

[43] Lorenzo Peppoloni, Filippo Brizzi, Emanuele Ruffaldi, and Carlo Alberto Avizzano. 2015. Augmented reality-aided tele-presence system for robot manipulation in industrial manufacturing. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology*. 237–240.

[44] André Pereira, Elizabeth J Carter, Iolanda Leite, John Mars, and Jill Fain Lehman. 2017. Augmented reality dialog interface for multimodal teleoperation. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 764–771.

[45] Eric Rosen, David Whitney, Elizabeth Phillips, Gary Chien, James Tompkin, George Konidaris, and Stefanie Tellex. 2020. Communicating robot arm motion intent through mixed reality head-mounted displays. In *Robotics Research: The 18th International Symposium ISRR*. Springer, 301–316.

[46] Mariah Schrum, Muyleng Ghuy, Erin Hedlund-Botti, Manisha Natarajan, Michael Johnson, and Matthew Gombolay. 2023. Concerning trends in likert scale usage in human-robot interaction: Towards improving best practices. *ACM Transactions on Human-Robot Interaction* 12, 3 (2023), 1–32.

[47] Michal Staniaszek, Esther Ray, Dave Niewinski, Koki Shinjo, Yoshiki Obinata, Chris Iverach-Brereton, Kei Okada, Michel Heinemann, Naoya Yamaguchi, Tobit Flatscher, Eric Voll, Naoto Tsukamoto, Harel Biggie, Kenji Brameld, Mario Gini, Juan Miguel Jimeno, and Wolf Vollprecht. [n. d.]. Spot ROS Driver. https://github.com/heuristicus/spot_ros.

[48] Stereolabs. [n. d.]. ZED 2i Stereo Camera. https://www.stereolabs.com/zed-2i.

[49] Carol R. Stoker, DR Burch, Butler P Hine, and James Barry. 1995. Antarctic undersea exploration using a robotic submarine with a telepresence user interface. *Ieee Expert* 10, 6 (1995), 14–23.

[50] Patrick Stotko, Stefan Krumpen, Max Schwarz, Christian Lenz, Sven Behnke, Reinhard Klein, and Michael Weinmann. 2019. A VR system for immersive teleoperation and live exploration with a mobile robot. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 3630–3637.

[51] Yunpeng Su, Mohammad Ahmadi, Christoph Bartneck, Frank Steinicke, and Xiaoqi Chen. 2019. Development of an optical tracking based teleoperation system with virtual reality. In *2019 14th IEEE conference on industrial electronics and applications (ICIEA)*. IEEE, 1606–1611.

[52] Da Sun, Andrey Kiselev, Qianfang Liao, Todor Stoyanov, and Amy Loutfi. 2020. A new mixed-reality-based teleoperation system for telepresence and maneuverability enhancement. *IEEE Transactions on Human-Machine Systems* 50, 1 (2020), 55–67.

[53] Daniel Szafir and Danielle Albers Szafir. 2021. Connecting human-robot interaction and data visualization. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 281–292.

[54] Toivo TÄHEMAA, Yevhen BONDARENKO, et al. 2019. Digital twin based synchronised control and simulation of the industrial robotic cell using virtual reality. *Journal of Machine Engineering* 19, 1 (2019), 128–144.

[55] Unity Technologies. [n. d.]. Unity. https://unity.com.

[56] Ievgenii A Tsokalo, David Kuss, Ievgen Kharabet, Frank HP Fitzek, and Martin Reisslein. 2019. Remote robot control with human-in-the-loop over long distances using digital twins. In *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 1–6.

[57] Michael Walker, Thao Phung, Tathagata Chakraborti, Tom Williams, and Daniel Szafir. 2023. Virtual, augmented, and mixed reality for human-robot interaction: A survey and virtual design element taxonomy. *ACM Transactions on Human-Robot Interaction* 12, 4 (2023), 1–39.

[58] David Whitney, Eric Rosen, Elizabeth Phillips, George Konidaris, and Stefanie Tellex. 2019. Comparing robot grasping teleoperation across desktop and virtual reality with ROS reality. In *Robotics Research: The 18th International Symposium ISRR*. Springer, 335–350.

[59] Carrie L Wyland and Joseph P Forgas. 2010. Here's looking at you kid: Mood effects on processing eye gaze as a heuristic cue. *Social cognition* 28, 1 (2010), 133–144.

[60] Tianhao Zhang, Zoe McCarthy, Owen Jow, Dennis Lee, Xi Chen, Ken Goldberg, and Pieter Abbeel. 2018. Deep imitation learning for complex manipulation tasks from virtual reality teleoperation. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 5628–5635.

[61] Jingbo Zhao, Robert S Allison, Margarita Vinnikov, and Sion Jennings. 2018. The effects of visual and control latency on piloting a quadcopter using a head-mounted display. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2972–2979.