2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
September 27 - October 1, 2021. Prague, Czech Republic

2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) | 978-1-6654-1714-3/21/$31.00 ©2021 IEEE | DOI: 10.1109/IROS51168.2021.9636090

A Mixed Reality Supervision and Telepresence Interface for Outdoor Field Robotics

Michael Walker*, Zhaozhong Chen*, Matthew Whitlock*, David Blair*,
Danielle Albers Szafir*†, Christoffer Heckman*, Daniel Szafir*†
*University of Colorado Boulder
Department of Computer Science
{michael.walker-1, zhaozhong.chen, matthew.whitlock, david.blair-1, christoffer.heckman}@colorado.edu
†University of North Carolina at Chapel Hill
Department of Computer Science
{danielle.szafir, daniel.szafir}@cs.unc.edu

Figure 1: Mixed reality 3D terrain of field teams' operating area seen from a video pass-through HMD. Current position and previous trajectory of human team (blue) and robot (red) are rendered contextually on the table. Above the terrain table, a dense RGB point cloud can be manipulated (position, rotation, scale) with the user's hands.

***Abstract*—Collaborative human-robot field operations rely on timely decision-making and coordination, which can be challenging for heterogeneous teams operating in large-scale deployments. In this work, we present the design of an immersive, mixed reality (MR) interface to support sense-making and situational awareness based on the data collection capabilities of both human and robotic team members. Our solution integrates state-of-the-art methods in environment mapping and MR so that users may gain rapid insights regarding the working environment, the current and previous locations of human and robot team members, and the environment data such team members have collected. We describe the implementation of our system, share lessons learned in collaborating with emergency responders throughout our design process, and offer a vision for the use of immersive displays for human-robot field team deployments in large-scale outdoor environments.**

I. Introduction

Robots hold significant potential for positive societal impact through their ability to support various human field operations involving environmental exploration and data collection. In this paper, we focus on how robots may assist human teams in emergency response operations, such as search-and-rescue and wildland firefighting. Wildfires are increasingly threatening human and wildlife habitats, contributing to global deforestation, and diminishing air quality in and around affected areas [1]. For instance, wildfires have burned approximately 2.8 million acres of land across the United States from January 1 to July 27, 2021[1], an increase of 500,000 acres from 2019.

[1]https://www.iii.org/fact-statistic/facts-statistics-wildfires

To assist in tackling these large-scale and often life-threatening situations, researchers are developing technical solutions to support such activities, where robots play a role. For example, the EU-ICARUS project was a government-funded, multi-nation collaborative effort aimed at decreasing the costs of major crises by equipping emergency responders with a comprehensive and integrated set of unmanned search and rescue tools to increase the situation awareness of human crisis managers such that more rescue work can be done in a shorter amount of time [2]. Emergency response operations often utilize robots to extend the capabilities of a response team by assisting in environmental exploration and data collection (see Murphy [3] for a survey of such work). Integrating data collected by both humans and robots may allow teams to improve operational safety and efficiency by increasing situational awareness and real-time understandings of team locations, paths traveled, data collected, and plans for the remainder of the operation [4].

Our goal is to design and implement solutions that could support emergency responders to have a better cooperation with VR/AR/MR and robot in the field by developing tools for real-time data collection and visualization of mission critical data. Our approach draws on techniques from Virtual, Augmented, and Mixed Reality for Human-Robot Interactions (VAM-HRI), which is an emerging sub-field of Human-Robot Interaction (HRI), the study of interactions between humans and robots. VAM-HRI explores the development of robots that interact with humans in mixed reality, the use of virtual reality for developing interactive robots, the design of new augmented reality interfaces that mediate communication between humans and robots, and best practices for the design of such interactions [5]. Recent VAM-HRI research has shown that mixed reality (MR) head-mounted display (HMD) interfaces may improve human-robot interactions in various significant ways, such as communicating robot motion intent with augmented reality (AR) virtual imagery [6], [7], [8] or enhancing collocated robot teleoperation [9], [10], [11] (see [12] for a survey). VAM-HRI's potential to improve collaboration within mixed human-robot teams is particularly important to large-scale field team operations, where timely decision-making and situational awareness may be supported through immersive MR visualizations [13].

In this research, our major **contributions** are:

1) The design and implementation of an end-to-end system demonstrating how MR interfaces may support emergency response human-robot field operations.
2) Combining vision-based 3D environment mapping with high-level 3D terrain maps to form a novel interface for supervising and operating field robots.
3) Highlighting generalizable design considerations for MR systems targeting emergency response field robotics and sharing our lessons learned through active collaboration with real emergency responders.

II. Related Work

There are many instances in which robots may be located in a remote location relative to their human operators, for instance when fighting wildland forest fires [14], exploring the bottom of the ocean [15], or crawling across the surface of distant planetary bodies [16]. The use of such field robots introduces challenges towards establishing and maintaining human *telepresence*—a user's feeling of being present at a place other than their true location—and *situational awareness*—"the perception of environmental elements and events with respect to time or space, the comprehension of their meaning, and the projection of their future status" [17]—of the remote robot's environment. When mobile robots operate within a remote location, human users may struggle to direct and monitor the robot, understand and analyze incoming data from the robot, and determine appropriate and successful courses of action (for surveys of such issues and potential methods to address them, see Murphy & Tadokoro, 2019 [18] and Szafir & Szafir, 2021 [19]).

Figure 2: Trajectories (lines) and current GPS positions (spheres) of the human field team (blue) and the field robot (red). White cubes represent robot-collected 3D reconstructions that users can select to view and manipulate the corresponding MR dense point clouds above the 3D terrain table.

HRI researchers have explored several methods for combining remote robot sensing and data presentation to local users to enhance user telepresence and situational awareness. For instance, telepresence robots that utilize video streaming technology can provide remote users an acceptable level of telepresence for many applications [20], such as home care assistance [21], education [22], construction [23], and medicine [24]. Although, standard video streaming provides users with some degree of telepresence and situational awareness, advanced sensors such as LiDAR, stereo cameras, or depth cameras may capture larger 3D reconstructions of remote environments (e.g., as dense point clouds, textured mesh reconstructions, etc.) to provide more information and enable users to more fully explore the 3D space. For example, our system enables custom views of the remote scene's 3D reconstruction, independent of the actual camera positions used to capture the event.

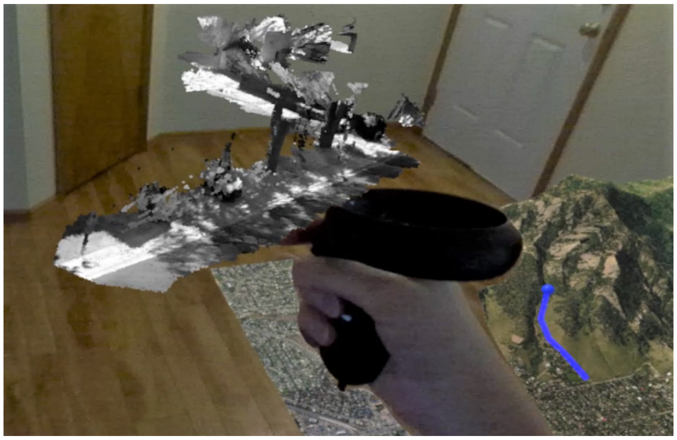

Figure 3: Handheld motion controllers are used to scale, rotate, and translate the 3D reconstructions in the AR interface.

However, mapping and reconstructing 3D environments is still an open area of research in the fields of computer vision and robotics. In our work, we leverage dense SLAM (Simultaneous Localization and Mapping) systems that can estimate a robot's location while reconstructing the surrounding 3D environment. Dense SLAM work conventionally focuses on indoor environments [25], [26], [27], although certain systems, such as Kintinuous [28] and Infinitam [29], can reconstruct large outdoor environments by sacrificing accuracy. However, those large-scale dense mapping systems contain a significant drawback: they are not robust to fast rotation or translation by the robot collecting the data [30]. Most robot state estimation is performed by minimizing photometric errors by comparing pixel intensities between frames and geometric errors by point-to-plane iterative closest point (ICP) optimization [27], [26]. These methods may lose tracking when there is rapid motion. A straightforward solution is to combine a robust pose estimation system with a dense mapping system. Visual-inertial SLAM can provide a robust pose estimation by integrating additional information from an Inertial Measurement Unit (IMU). Several visual-inertial SLAM systems have been developed, including OKVIS [31], VINS-Fusion [32] and ORB-SLAM3 [33]. With the help of visual-inertial SLAM, the dense mapping system can focus on reconstruction rather than pose estimation. With this approach, we combine VINS-Fusion and the dense mapping system Voxblox [34] to increase tracking robustness, which is described in more detail in Section IV.

Researchers have utilized 3D reconstructions for a variety of robot teleoperation interfaces (e.g., [35], [36], [37]), including for emergency response applications such as subterranean tunnel [38] and urban search-and-rescue [39] missions. More recently, HRI researchers have developed methods to stream live 3D reconstructions of robot workstations to HMD interfaces. These cyber-physical interfaces [40] provide improved telepresence for remote assembly and manipulation tasks [41], [42], [43], [44] by allowing operators to explore 3D reconstructions with natural body motions, where users can walk around the robot's remote environment in augmented virtuality environments. In addition to manipulation tasks, HMD interfaces that utilize 3D reconstructions may provide significant improvements in teleoperating robots for navigation tasks [45], [46], [47], [48], [49], [50].

Although the above interfaces may allow human users to better monitor autonomous field robots, they are specifically designed for small-scale settings, whereas in wildland firefighting operations team members may be dispersed across tens of miles. Additionally, very large-scale outdoor areas (e.g., mountain ranges) cannot be completely mapped or reconstructed at high resolutions as seen in MR interfaces that target smaller-scale indoor settings. Some researchers have proposed projector-based MR interfaces for monitoring autonomous outdoor robot teams on 2D AR terrain maps; however, interfaces such as these are not portable or able to display 3D data collected by field robots [51], [52]. In this work, we focus on addressing such challenges through the development of a mixed reality interface that can support large-scale field operations of mixed human-robot teams by utilizing a 3D terrain map visualized in MR.

III. Emergency Response Stakeholders Formative Feedback

To design solutions for emergency response command centers and field teams, we first conducted a semi-structured interview with five members of the *Colorado Center of Excellence for Advanced Technology Aerial Firefighting*, a rural wildland firefighting group based in Rifle, CO. This semi-structured interview method [53] is commonly used in HCI literature to glean information from experts about interface design for domains without significant design guidance. Semi-structured interviews have been used for understanding HRI interface design in emergency response [54], [13], as HRI researchers can gain feedback about interface design without needing to craft a controlled study in a simulated disaster scenario. We sought to understand current practices in wildland firefighting, including examples of specific operations and use of robots. Throughout the interview, the wildland firefighting group explained existing practices and current technologies utilized to combat wildfires. We also discussed how new solutions that integrate mobile devices, ground and aerial robots, and immersive MR might improve local and global operations.

We used these interviews to identify three high-level use cases where VAM-HRI may augment outdoor emergency response operations: *Searching*, *Survey*, and *Response*:

Searching: This use case considers scenarios where persons or objects are missing. For example, responders may search for a person missing in the wilderness, a temporary structure (e.g., hunting stand), or other lost equipment. These scenarios are typically time-sensitive and are often burdened by communication requirements (e.g., routing information back to a remote operations center). Multi-person and robot-aided operations can add complexity as decisions about what to do next depend on data from prior operations. For example, decisions about areas to search next can be supported by knowledge of where team members are and have been. Images and video collected

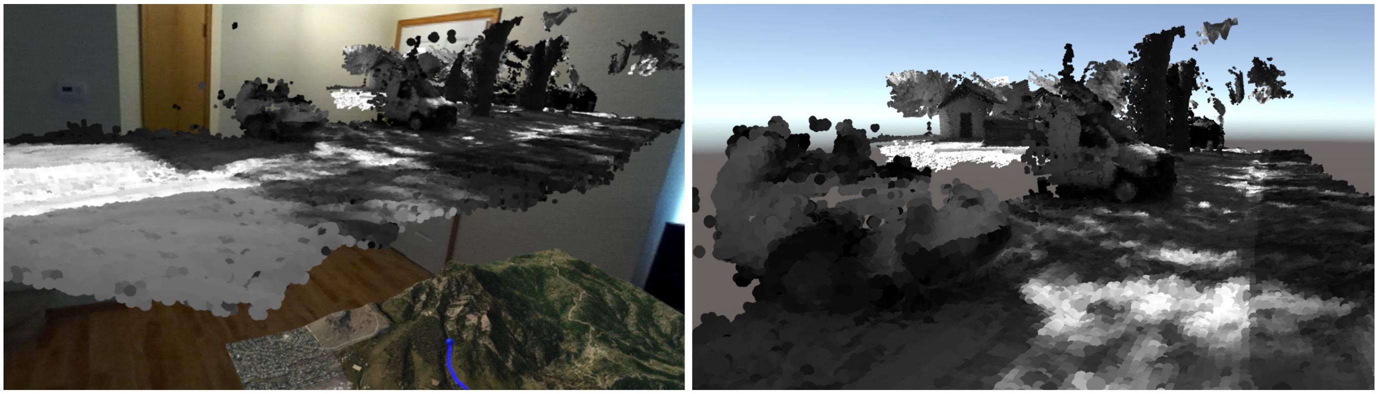

Figure 4: The 3D reconstructions can be viewed in two mixed reality modes: 1.) augmented reality – where the user is able to view the point cloud integrated within their local environment alongside the 3D terrain map; and 2.) augmented virtuality – in which the user can walk in a virtual environment to explore the reconstruction at full-scale as if they were at the robot's remote location themselves.

by robots may provide helpful perspectives of the area and more up-to-date or high-resolution information than static satellite maps. Responders currently rely on a mobile Team Awareness Kit (TAK) application while in the field to help track the locations of other team members. While the location data is helpful, TAK does not necessarily allow team members to also make quick sense of the imagery, and sensor data collected on an operation. By displaying paths traveled and data collected by human and robotic team members, our interface aims to support team members deciding on future areas to search and next steps of an operation.

Surveying: The emergency response group described several activities that involved surveying an unknown or partially known area. For instance, their prior operations had included assessing the burn risk of a particular area and scouting and mapping locations for FEMA (the Federal Emergency Management Agency in the United States). The outcome of such activities was often an annotated map with geotagged images, videos, text, symbols, and data measurements. However, in our discussions, we identified areas for continued improvement by involving more robotic capabilities and a broader array of sensor measures to increase the diversity of information available from the survey of an area. One example presented was the difficulty of managing how when a drone-collected data in real time. In this scenario, it would be helpful to view the collected data and seeing where it has flown on a map view. More generally, the emergency response team mentioned that seeing how the mapped area updates over the course of the operation could help ensure accurate completion of the survey.

Responding: Response operations such as wildland firefighting require real-time assessment and actions out in the field, typically in response to an ongoing emergency. Robots in this scenario may be deployed to perform preliminary fire evaluations, drop fire retardant chemicals, or assess the current and ongoing damage caused by a fire [55]. Response scenarios require a coordinated effort of team members to ensure a timely response and the safety of the team. In this scenario, it is critical that people are aware of challenging terrain (e.g., canyons or mountains) in limited visibility caused by tree cover and/or smoke. With regards to safety of the team, our interviewees discussed a scenario where a team member died when an aerial robot's chemical drop caused trees to fall on the team member. An increased awareness of team members' locations could guide next steps for safe and efficient emergency response.

IV. System Implementation

By synthesizing the formative feedback gathered in §III, we developed a MR interface designed to assist first response field teams in their efforts in *searching*, *surveying*, and *responding*. The interface's design centers around a MR 3D terrain map that provides an overview of positional data for both humans and robot field teams. The MR design of the interface allows first responders to see and be aware of potentially hazardous surroundings (e.g., rocks, cliffs, fires, etc.).

Positional data is visualized on a realistic, scaled 3D terrain map to enhance both field teams' and their associated command center's situational awareness when performing *searching* and *responding* operations. Response missions could be potentially better planned and coordinated when users are contextually provided the current locations of other field teams that are operating in the same geographic region. Also, by identifying locations already travelled, search missions could be made more efficient by preventing redundant searches in areas already explored by other field teams.

Additionally, the interface provides access to 3D environmental reconstructions collected by field robots and visualized to support *searching* and *surveying* tasks. Robotic scouts can provide human team members with 3D reconstructions of remote locations that users can use to enhance telepresence to better assess areas of interest during surveying or searching missions without needing to cross rough, uncharted terrain to be physically present at a location of interest.

A. Mixed Reality HMD Interface

1) Display: We used the Oculus Quest HMD with a Zed Mini pass-through stereo video camera as our MR display.

This configuration allows the user to "see" out of the HMD with accurate stereoscopic depth by projecting the dual live video feeds to the user's left and right eyes, while also allowing for AR imagery to be overlaid on the video feed. Video pass-through is beneficial for a wider range of emergency response operations as it works both indoors and outdoors, whereas see-through HMDs that can effectively operate outdoors are still

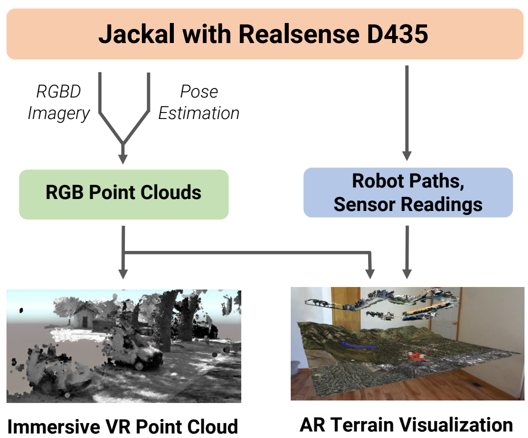

Figure 5: System Overview: The field robot generates point cloud representations of its remote environment that are combined with GPS data visualized within an MR interface.

2) Design: Emergency response field teams operate over various locations and terrains. Therefore, our interface was designed to accommodate any geographical location. Geolocation solutions currently utilized by first response field teams are designed for use on 2D displays. However, 3D interfaces (especially 3D HMD interfaces) may provide better situational awareness and usability than traditional 2D displays for monitoring autonomous robots in large outdoor settings [56]. Our interface provides users with a MR 3D terrain map. The user can use the map to visualize human and robot field team's previous and current locations. GPS coordinates are used to produce an accurate 3D terrain map from publicly available satellite imagery to allow field teams to tailor the 3D terrain map to their current operational area (Fig. 1).

The 3D terrain map tracks remote team members and displays their GPS locations. The HMD application, developed within the Unity engine, receives live GPS coordinates from the field team and robot and translates their latitude, longitude, and altitude values into their corresponding location on the 3D terrain map. GPS data is stored within the Unity application so previous GPS locations can be presented to the user in the form of trajectory lines. The most recent GPS location is displayed as a sphere to indicate the field team human or robot's current location (Fig 2).

While wearing the MR HMD, users can interact with the interface via handheld motion controllers. Users are able to select 3D reconstructions for visualization by tapping on one of the 3D reconstruction markers that are represented as white cubes along the robot's trajectory on the 3D terrain map (Fig. 2). Once selected, the cube's associated 3D reconstruction is rendered above the 3D terrain map in AR (Fig. 1). Since our target users would access the interface from diverse settings (e.g., on a wilderness trail, from a field team command center, etc.), we provide the ability to re-position, rotate, and scale the 3D terrain map and 3D reconstructions to dynamically fit various operating environments. To translate and rotate the point cloud, users grab the visualization with a single hand and "drag" to reorient. To scale the point cloud, users grab the visualization with two hands, bringing their hands together to scale down or moving their hands further apart to scale up. This allows users to dynamically fine-tune the placement of the interface components to custom fit any working space and be seen from any desired viewing angle (Fig 3).

Figure 6: Large-scale 3D reconstruction of the robot's path through a remote outdoor environment. See Figure 7 for a closer view of the reconstruction from the user's perspective.

A disadvantage of the AR visualization is that a visually-cluttered real world environment may make the 3D reconstruction difficult to interpret when overlaid on a complex visual backdrop [57]. Therefore, our interface provides an additional means of viewing the 3D reconstructions. By selecting the point cloud icon again, the video pass-through camera deactivates, and the user enters immersive augmented virtuality to view the 3D reconstruction at full scale overlaid on a simple backdrop for easier comprehension of the 3D data returned by the remote Jackel robot. Users can walk the virtual representation of the environment collected by the robot as if they were at the robot's location, enhancing user telepresence (Fig. 4).

In augmented virtuality, the real world is no longer visible to the user which exposes them to dangers of colliding with, tripping on, or falling of off nearby environmental hazards (i.e., rocks, trees, cliffs, etc.). However, the built-in capabilities of the Oculus Quest HMD allows users to define safe spaces while using the MR application by tracing

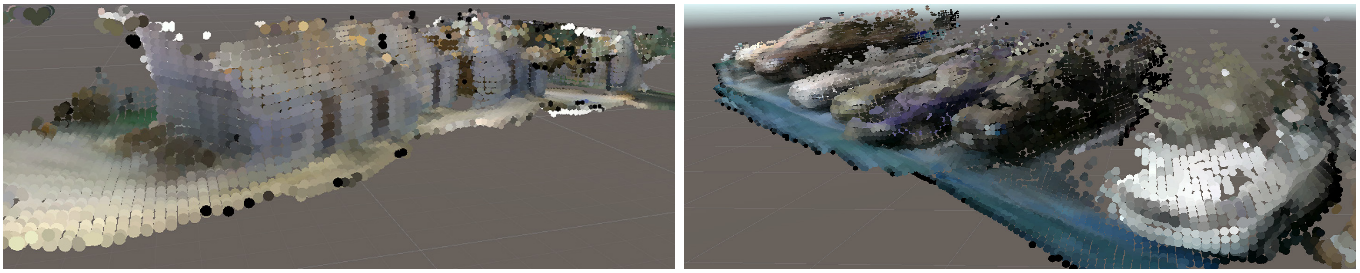

Figure 7: Dense point cloud 3D reconstructions collected by the field robot and their associated details (cars, shadows, houses).

virtual boundaries on the ground with motion controllers. When a user approaches the edge of such a safe space, the pass-through camera automatically activates and brings the user's awareness back to the real world and its associated environmental hazards.

B. Field Robot for Data Collection

We utilized a Jackal unmanned ground vehicle for data collection. This robot was equipped with an on-board computer, GPS, and IMU fully integrated with ROS for outdoor autonomous operations in the field. The robot was additionally outfitted with a Realsense D435 stereo camera for collecting dense point clouds from its surrounding environment.

We built a system for field robot data collection coupled with MR visualizations for telepresence, summarized in Figure 5. We used a Realsense camera D435 with high resolution IMU MicroStrain 3DM-GX5-15 at a 15 Hz frame rate and inertial measurement frequency of 200 Hz. The localization and dense mapping systems run solely on the CPU. For these processes, we used an Intel i7-8850H CPU with 2.6 GHz frequency that can reach ~2Hz mesh update. VINS-Fusion [32] is used for tracking with Voxblox for dense mapping. Voxblox relies on a outside system to provide pose estimation so that it can project the point cloud in the local camera frame to the global frame and merge the local point cloud to the global model [34]. The stereo image and IMU measurement are sent to the VINS-Fusion system for tracking. RGB-D images, together with the estimated pose from the tracking system, are sent to Voxblox for generating RGB point clouds for reconstruction. Voxblox is a voxel-based TSDF update algorithm. The voxel size must be set within a reasonable range according to different environments. Large voxel size yields low resolution while small voxel size decreases the voxel update speed. In pilot testing with our architecture, 30cm voxels offered both reasonable resolution and latency around 300–500 ms. After generated by the Voxblox, the down-sampled 3D reconstruction is sent to the MR headset such that we can render the resulting 3D reconstruction and visualize it in an immersive way. Figures 6 and 7 provide examples of 3D reconstructions for outdoor environments.

C. Robot–HMD Communication

Our simulation framework contains two components: the robot platform (Linux and ROS) and the HMD interface platform (Windows and Unity engine). Our implementation combines the ROS-Sharp package [58] on the Unity side and the ROSBridge WebSocket package [59] on the ROS side. This coordination allows Unity to act as any other ROS node, or collection of nodes, on a ROS system following the standard publisher/subscriber model, with bidirectional data exchanged in a JSON format. Within this framework, ROS issues GPS messages to the Unity node to allow for the rendering of robot and field team locations on the 3D terrain map. Point cloud messages are also sent from ROS to the Unity node to allow for the AR and augmented virtuality renderings of 3D environment data collected by the field robot (Figure 5).

V. Discussion

A. Emergency Responder Summative Feedback

To understand how our system could be applied to large-scale outdoor operations, we conducted a follow-up interview with a member of the emergency response team mentioned in §3. This interview was supported by a video demonstrating the capabilities of our system to help identify specific use cases and applications where this solution might be desirable. We also sought feedback on our visualization design regarding how the MR interface might support the **Searching**, **Surveying**, and **Responding** tasks discussed in §3.

Overall, our stakeholders thought the system could be applied to wildland firefighting, urban search-and-rescue, and land survey use cases, while noting that the system we presented could be adapted to better serve these different applications and team members in different roles. From a hardware perspective, they mentioned preliminary explorations into using AR and VR to support their operations, but noted that constraints such as the weight and power requirements of current headsets make large-scale deployments challenging. They felt that the utility of an MR headset depends on the user's role, explaining that "for some folks it could be a distraction, so we have to be judicious in how much we display and how much data we provide." In particular, people out in the field may find a detailed MR visualization interface less helpful, though "there definitely are use cases where it can help you navigate: when you're off trail or cross-country" but that "subtle augmented reality cues of just the arrow to follow might be more helpful than a full-on virtual map." This is in line with previous work on use of mobile AR for field operations that simple, salient visual cues may be the most helpful during response or field research operations [13]. Command centers, alternatively, have different needs:

"for a command center where you have a god's eye view of everything, that would definitely be more useful", which would be supported by the more detailed visualization of trajectories and point clouds overlaid on the terrain map (Figure 1).

In addition to the role of the team member, our stakeholders indicated the importance usage context when designing MR visualizations. We discussed what types of data and visualization would be helpful to different types of operations and learned that "being able to see the locations of others was the most useful feature" for a wide range of outdoor operations (Figure 2). The usefulness of 3D reconstructions (Figure 7) depended more on the application scenario. For urban scenarios such as "hazmat, collapsed buildings, and law enforcement...[3D reconstructions] would be useful." On the other hand, "for true firefighting what [they] really want is not so much a point cloud or a high resolution reconstruction of the environment, [they] just want to know the forward line of progress of the fire... That really high resolution data would be kind of a distraction." In rural operations, the at-a-glance visualizations for navigation and understanding the locations of other team members would be the only data needed.

B. Future Work

Due to the COVID-19 pandemic, our research team was unable to evaluate our system through user studies. Though our formative and summative qualitative evaluations, we provide guidance for building mixed reality interfaces targeting emergency response. When possible, we plan to run empirical studies to assess our system's usability and performance relative to state-of-the-art systems currently used by field teams.

In extending our current system, we intend to explore rendering AR annotations both on the 3D terrain map and in the real world (i.e., a user could look at either our AR interface or the real world sky to see the trajectory of an aerial robot teammate). These situated annotations would prevent users from having to mentally convert locations from the 3D terrain map to the real world. Also, alternative viewing mediums (e.g., AR tablets and phones) would allow for passive viewing of the interface by team members without HMDs.

While our current system can achieve on-board 3D reconstruction, we plan to explore new ways to increase resolution and efficiency. The new Unity-ROS-HUB with TCP protocol may enable more rapid data transfer through our system. In addition, we currently use Voxblox, which is CPU-based to serve its initial goal for drones; however, a GPU-based system would likely accelerate performance. We also plan to bring other robots with larger payloads into the field to enable the use of LiDAR in conjunction with the visual localization system to enhance robot state estimation.

Finally, as the interface stands, users only receive data from the robot (GPS coordinates and point clouds); however, we envision future iterations of our system allowing for two-way communication between users and robots and commands being sent to the robot from the 3D terrain map (e.g., dropping waypoints, drawing trajectories, assigning tasks, etc.). This visualization-supported workflow could improve communication among teams in the field and remote operations center.

VI. Conclusion

By pairing modern MR technology with state-of-the-art concepts from HRI and robotic perception, we present an interface designed to enhance situational awareness and coordination of mixed human-robot field teams in large-scale outdoor settings. Our system extends prior efforts on improving team situational awareness for indoor environments and research on the design of interfaces for field robotics. Our design, which was informed by discussions with emergency responders regarding their needs both in the field and at operations centers, may provide an extensible basis for future systems that support human-robot teaming in large field settings.

References

[1] J. T. Abatzoglou and A. P. Williams, "Impact of anthropogenic climate change on wildfire across western us forests," *Proceedings of the National Academy of Sciences*, vol. 113, no. 42, pp. 11770–11775, 2016.

[2] G. De Cubber, D. Doroftei, D. Serrano, K. Chintamani, R. Sabino, and S. Ourevitch, "The eu-icarus project: developing assistive robotic tools for search and rescue operations," in *2013 IEEE international symposium on safety, security, and rescue robotics (SSRR)*, pp. 1–4, IEEE, 2013.

[3] R. R. Murphy, S. Tadokoro, and A. Kleiner, "Disaster robotics," in *Springer handbook of robotics*, pp. 1577–1604, 2016.

[4] R. R. Murphy, *Disaster robotics*. MIT press, 2014.

[5] T. Williams, D. Szafir, T. Chakraborti, and H. Ben Amor, "Virtual, augmented, and mixed reality for human-robot interaction," in *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 403–404, 2018.

[6] M. Walker, H. Hedayati, J. Lee, and D. Szafir, "Communicating robot motion intent with augmented reality," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 316–324, 2018.

[7] C. Reardon, K. Lee, J. G. Rogers, and J. Fink, "Augmented reality for human-robot teaming in field environments," in *International Conference on Human-Computer Interaction*, pp. 79–92, Springer, 2019.

[8] C. Reardon, K. Lee, and J. Fink, "Come see this! augmented reality to enable human-robot cooperative search," in *2018 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, pp. 1–7, IEEE, 2018.

[9] H. Hedayati, M. Walker, and D. Szafir, "Improving collocated robot teleoperation with augmented reality," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 78–86, 2018.

[10] M. E. Walker, H. Hedayati, and D. Szafir, "Robot teleoperation with augmented reality virtual surrogates," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 202–210, IEEE, 2019.

[11] J. M. Gregory, C. Reardon, K. Lee, G. White, K. Ng, and C. Sims, "Enabling intuitive human-robot teaming using augmented reality and gesture control," *arXiv preprint arXiv:1909.06415*, 2019.

[12] D. Szafir, "Mediating Human-Robot Interactions with Virtual, Augmented, and Mixed Reality," in *International Conference on Human-Computer Interaction*, pp. 124–149, 2019.

[13] M. Whitlock, K. Wu, and D. A. Szafir, "Designing for mobile and immersive visual analytics in the field," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, pp. 503–513, Jan 2020.

[14] R. N. Haksar and M. Schwager, "Distributed deep reinforcement learning for fighting forest fires with a network of aerial robots," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1067–1074, IEEE, 2018.

[15] T. Hardy and G. Barlow, "Unmanned underwater vehicle (uuv) deployment and retrieval considerations for submarines," in *International Naval Engineering Conference and Exhibition 2008*, 2008.

[16] K. Sanderson, "Mars rover spirit (2003-10): Nasa commits robot explorer to her final resting place," *Nature*, vol. 463, no. 7281, pp. 600–601, 2010.

[17] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human factors*, vol. 37, no. 1, pp. 32–64, 1995.

[18] R. R. Murphy and S. Tadokoro, "User interfaces for human-robot interaction in field robotics," in *Disaster Robotics*, pp. 507–528, 2019.

[19] D. Szafir and D. A. Szafir, "Connecting Human-Robot Interaction and Data Visualization," in *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI'21)*, pp. 281–292, 2021.

[20] A. Kristoffersson, S. Coradeschi, and A. Loutfi, "A review of mobile robotic telepresence," *Advances in Human-Computer Interaction*, vol. 2013, 2013.

[21] F. Michaud, P. Boissy, D. Labonte, H. Corriveau, A. Grant, M. Lauria, R. Cloutier, M.-A. Roux, D. Iannuzzi, and M.-P. Royer, "Telepresence robot for home care assistance.," in *AAAI spring symposium: multidisciplinary collaboration for socially assistive robotics*, pp. 50–55, California, USA, 2007.

[22] O.-H. Kwon, S.-Y. Koo, Y.-G. Kim, and D.-S. Kwon, "Telepresence robot system for english tutoring," in *2010 ieee workshop on advanced robotics and its social impacts*, pp. 152–155, IEEE, 2010.

[23] I. Rae, B. Mutlu, and L. Takayama, "Bodies in motion: mobility, presence, and task awareness in telepresence," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2153–2162, 2014.

[24] P. S. Green, J. W. Hill, J. F. Jensen, and A. Shah, "Telepresence surgery," *IEEE Engineering in Medicine and Biology Magazine*, vol. 14, no. 3, pp. 324–329, 1995.

[25] R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohi, J. Shotton, S. Hodges, and A. Fitzgibbon, "Kinectfusion: Real-time dense surface mapping and tracking," in *2011 10th IEEE International Symposium on Mixed and Augmented Reality*, pp. 127–136, IEEE, 2011.

[26] C. Kerl, J. Sturm, and D. Cremers, "Dense visual slam for rgb-d cameras," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2100–2106, IEEE, 2013.

[27] T. Whelan, S. Leutenegger, R. Salas-Moreno, B. Glocker, and A. Davison, "Elasticfusion: Dense slam without a pose graph," Robotics: Science and Systems, 2015.

[28] T. Whelan, M. Kaess, H. Johannsson, M. Fallon, J. J. Leonard, and J. McDonald, "Real-time large-scale dense rgb-d slam with volumetric fusion," *The International Journal of Robotics Research*, vol. 34, no. 4-5, pp. 598–626, 2015.

[29] V. A. Prisacariu, O. Kähler, S. Golodetz, M. Sapienza, T. Cavallari, P. H. Torr, and D. W. Murray, "Infinitam v3: A framework for large-scale 3d reconstruction with loop closure," *arXiv preprint arXiv:1708.00783*, 2017.

[30] V. Usenko, J. Engel, J. Stückler, and D. Cremers, "Direct visual-inertial odometry with stereo cameras," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1885–1892, IEEE, 2016.

[31] S. Leutenegger, S. Lynen, M. Bosse, R. Siegwart, and P. Furgale, "Keyframe-based visual–inertial odometry using nonlinear optimization," *The International Journal of Robotics Research*, vol. 34, no. 3, pp. 314–334, 2015.

[32] T. Qin, S. Cao, J. Pan, and S. Shen, "A general optimization-based framework for global pose estimation with multiple sensors," *arXiv preprint arXiv:1901.03642*, 2019.

[33] C. Campos, R. Elvira, J. J. G. Rodríguez, J. M. Montiel, and J. D. Tardós, "Orb-slam3: An accurate open-source library for visual, visual-inertial and multi-map slam," *arXiv preprint arXiv:2007.11898*, 2020.

[34] H. Oleynikova, Z. Taylor, M. Fehr, R. Siegwart, and J. Nieto, "Voxblox: Incremental 3d euclidean signed distance fields for on-board mav planning," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017.

[35] P. Beňo, F. Duchoň, M. Tölgyessy, P. Hubinskỳ, and M. Kajan, "3d map reconstruction with sensor kinect: Searching for solution applicable to small mobile robots," in *2014 23rd International Conference on Robotics in Alpe-Adria-Danube Region (RAAD)*, pp. 1–6, IEEE, 2014.

[36] P. Kim, J. Chen, and Y. K. Cho, "Slam-driven robotic mapping and registration of 3d point clouds," *Automation in Construction*, vol. 89, pp. 38–48, 2018.

[37] D. Szafir, B. Mutlu, and T. Fong, "Designing planning and control interfaces to support user collaboration with flying robots," *The International Journal of Robotics Research*, vol. 36, no. 5-7, pp. 514–542, 2017.

[38] N.-B. Jing, X.-M. Ma, and W. Guo, "3d reconstruction of underground tunnel using kinect camera," in *2018 International Symposium on Computer, Consumer and Control (IS3C)*, pp. 278–281, IEEE, 2018.

[39] Z. Zhang, A. Liu, Z. Gao, and C. Liu, "A kinect-based 3d sensing and human action recognition solution for urban search and rescue environments," in *2013 IEEE RO-MAN*, pp. 344–345, IEEE, 2013.

[40] J. I. Lipton, A. J. Fay, and D. Rus, "Baxter's homunculus: Virtual reality spaces for teleoperation in manufacturing," *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 179–186, 2017.

[41] D. Whitney, E. Rosen, E. Phillips, G. Konidaris, and S. Tellex, "Comparing robot grasping teleoperation across desktop and virtual reality with ros reality," in *Robotics Research*, pp. 335–350, Springer, 2020.

[42] D. Whitney, E. Rosen, D. Ullman, E. Phillips, and S. Tellex, "Ros reality: A virtual reality framework using consumer-grade hardware for ros-enabled robots," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1–9, IEEE, 2018.

[43] D. Sun, A. Kiselev, Q. Liao, T. Stoyanov, and A. Loutfi, "A new mixed-reality-based teleoperation system for telepresence and maneuverability enhancement," *IEEE Transactions on Human-Machine Systems*, vol. 50, no. 1, pp. 55–67, 2020.

[44] D. Ni, A. Nee, S. Ong, H. Li, C. Zhu, and A. Song, "Point cloud augmented virtual reality environment with haptic constraints for teleoperation," *Transactions of the Institute of Measurement and Control*, vol. 40, no. 15, pp. 4091–4104, 2018.

[45] J. Allspaw, J. Roche, N. Lemiesz, M. Yannuzzi, and H. A. Yanco, "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)*, 2018.

[46] T. Klamt, M. Schwarz, C. Lenz, L. Baccelliere, D. Buongiorno, T. Cichon, A. DiGuardo, D. Droeschel, M. Gabardi, M. Kamedula, *et al.*, "Remote mobile manipulation with the centauro robot: Full-body telepresence and autonomous operator assistance," *Journal of Field Robotics*, vol. 37, no. 5, pp. 889–919, 2020.

[47] G. Bruder, F. Steinicke, and A. Nüchter, "Poster: Immersive point cloud virtual environments," in *2014 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 161–162, IEEE, 2014.

[48] S. Kim, Y. Kim, J. Ha, and S. Jo, "Mapping system with virtual reality for mobile robot teleoperation," in *2018 18th International Conference on Control, Automation and Systems (ICCAS)*, pp. 1541–1541, IEEE, 2018.

[49] A. Mossel and M. Kroeter, "Streaming and exploration of dynamically changing dense 3d reconstructions in immersive virtual reality," in *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pp. 43–48, IEEE, 2016.

[50] P. Stotko, S. Krumpen, M. Schwarz, C. Lenz, S. Behnke, R. Klein, and M. Weinmann, "A vr system for immersive teleoperation and live exploration with a mobile robot," *arXiv preprint arXiv:1908.02949*, 2019.

[51] S. Omidshafiei, A.-A. Agha-Mohammadi, Y. F. Chen, N. K. Ure, S.-Y. Liu, B. T. Lopez, R. Surati, J. P. How, and J. Vian, "Measurable augmented reality for prototyping cyberphysical systems: A robotics platform to aid the hardware prototyping and performance testing of algorithms," *IEEE Control Systems Magazine*, vol. 36, no. 6, pp. 65–87, 2016.

[52] C. R. Amburn, N. L. Vey, M. W. Boyce, and J. R. Mize, "The augmented reality sandtable (ares)," tech. rep., ARMY RESEARCH LAB ABERDEEN PROVING GROUND MD HUMAN RESEARCH AND ENGINEERING . . . , 2015.

[53] R. Longhurst, "Semi-structured interviews and focus groups," *Key methods in geography*, vol. 3, no. 2, pp. 143–156, 2003.

[54] J. Casper and R. Murphy, "Human-robot interactions during the robot-assisted urban search and rescue response at the world trade center," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 33, no. 3, pp. 367–385, 2003.

[55] E. Beachly, C. Detweiler, S. Elbaum, B. Duncan, C. Hildebrandt, D. Twidwell, and C. Allen, "Fire-aware planning of aerial trajectories and ignitions," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 685–692, IEEE, 2018.

[56] M. Lager and E. A. Topp, "Remote supervision of an autonomous surface vehicle using virtual reality," *IFAC-PapersOnLine*, vol. 52, no. 8, pp. 387–392, 2019.

[57] M. Whitlock, S. Smart, and D. A. Szafir, "Graphical perception for immersive analytics," in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 616–625, IEEE, 2020.

[58] M. Bischoff, "Ros sharp," *Ros Sharp Framework*, 2018.

[59] J. Mace and J. Lee, "Rosbridge suite," *GitHub Repository*, 2013.