Augmented Telepresence for Remote Inspection with Legged Robots

Benjamin Tam†, **Navinda Kottege**†, **Branislav Kusy***
†Robotics and Autonomous Systems Group, CSIRO, Pullenvale, QLD 4069, Australia
*Distributed Sensor Systems Group, CSIRO, Pullenvale, QLD 4069, Australia
{benjamin.tam, navinda.kottege, brano.kusy}@csiro.au

Abstract

In many applications, humans remotely control a robot through teleoperation, to augment its capabilities with the cognitive skills of humans. The arrival of mainstream head mounted displays (HMDs) and 360° cameras have brought about reduction in their size, increase in functionality as well as lowering of cost. This paper presents an augmented telepresence system using a HMD for remote operation of a legged robot via a web application. Low-cost and low-weight hardware were chosen to realise the system on a legged robot platform, but the web technologies used are device-agnostic and work efficiently across multiple platforms. A mock-up remote visual inspection test course was used to experimentally evaluate the HMD system by comparing performance with a fixed camera display approach. A total of 28 individual operators completed the test course with course completion performance recorded along with user feedback. The results show that the HMD increased the operator's perception of the remote environment but there are no significant task performance advantages using a HMD for remote inspection with legged robots.

1 Introduction

As the capabilities of robots improve, the scenarios where they can be deployed expand. Although most research efforts in this domain aim to create fully autonomous robots, there are applications where semi-autonomous systems with a 'human-in-the-loop' are beneficial. Examples for these are search and rescue, confined space inspection and hazardous environment operations. By having a human operator in the loop, the system can utilise the cognitive skills and perception of an experienced operator without putting them at risk [Aracil *et al.*, 2007]. These systems are examples of teleoperation, where an operator remotely controls a robot.

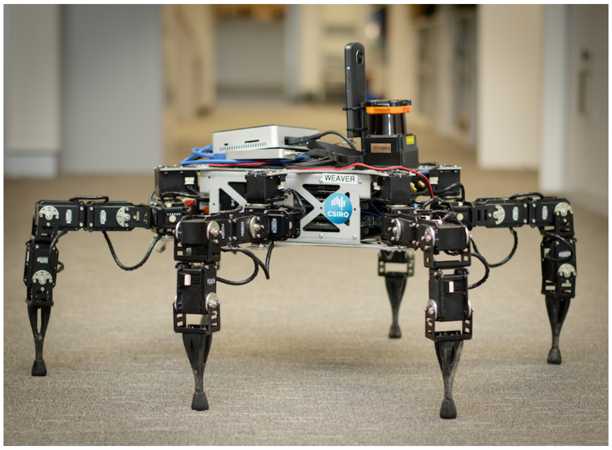

Figure 1: Weaver with a 360° visual system

The operator's performance in completing tasks is impacted by their situational awareness, that is their ability to accurately perceive the remote environment and comprehend the state of the robot. To provide this awareness, an extensive telepresence system that enables a person to feel they are present at a remote location is required. The goal is to allow an operator to critically analyse the situation and manipulate the environment accordingly, just as if they were present in that space.

Legged robots, such as Weaver shown in Figure 1, described in [Bjelonic *et al.*, 2016], [Homberger *et al.*, 2016], [Bjelonic *et al.*, 2017] have the ability to traverse terrain that is challenging for conventional wheeled and tracked robots. They can use their legs to manipulate the environment, step over obstacles and place their feet on small footholds for traversal [Todd, 1985]. This makes them ideal candidates for remote inspection tasks in unstructured terrain or confined spaces; environments where physical interaction is required, complex, and/or requires care to avoid damage. Remote inspection tasks require the operator to navigate the robot across the terrain while avoiding collisions and to visually inspect the mission objective.

Head (or Helmet) mounted displays (HMDs) have benefited from recent consumer products with lower cost and greater functionality. Systems such as the Oculus Rift, HTC Vive and Samsung GearVR allow the user to be immersed in a virtual environment with sensors tracking the user's visual perspective. While the focus of these products is virtual reality (VR), where the user interacts with a computer generated 3D environment, augmented telepresence has also benefited. Augmented telepresence in robotics is superposing additional information as computer-generated graphics in the user's view of an on-board camera. The overlay of information increases the user's perception of the robot's environment, resulting in better situational awareness.

[Fiala, 2005] highlighted the different camera systems used with HMDs; stereoscopic camera pairs on a pan/tilt module or monoscopic 360° cameras. Multiple works have used stereo cameras mounted on a pan/tilt module, controlled according to the HMD's attitude. This provides a stereoscopic view for the user to perceive depth but has a high HMD latency for moving the view. The latency is due to the combination of the delay of sending commands to the pan/tilt motors and for the video to transmit back to the user via the network. Studies on user's perception of latency have shown delays of 15 ms are noticeable [Mania *et al.*, 2004]. In comparison, 360° cameras provide a smooth head tracking response with HMD latency coming from the delay between the HMD tracker and a refresh of the display with the new view point. However, this reduces the effective resolution as the user views only a portion of the captured scene. These cameras provide a monoscopic view, losing depth information.

We propose an augmented telepresence web application capable of running on mobile HMDs to immerse the operator in a remote environment while teleoperating a robot. The system uses a 360° camera over a stereoscopic camera module to eliminate the dependency on low network latency, something not guaranteed on poor networks. This web application is tested using a legged robot platform in a mock-up remote navigation and inspection task.

Section 2 of this paper will discuss previous works related to robot teleoperation. Section 3 and Section 4 will introduce the components of the web application and the hardware used to evaluate the system. The experimental setup and results will be shown in Section 5 and discussed in Section 6. Section 7 will conclude the paper and provide an outline to future work.

2 Related Work

There are various traditional and novel human-robot interfaces (HRI), with [Fong and Thorpe, 2001] describing systems ranging from multiple displays, joysticks and keyboards to gesturing and brainwaves. [Yanco *et al.*, 2004] created guidelines for developing interfaces for HRI based on search and rescue robots. They found that providing: 1) a map for locating the robot, 2) well fused sensor data, 3) minimal control windows, and 4) spatial information about the remote environment, improves the situational awareness of the operator.

Research comparing monoscopic and stereoscopic views with HMDs found that while both had improvements over a screen [Kratz and Ferriera, 2016], the improvement between them varied. A stereoscopic view was beneficial for relative distance tasks [Kratz and Ferriera, 2016], but had no advantage in collision avoidance, navigation or collaborative tasks compared to monoscopic view [Livatino *et al.*, 2009]. This suggests that for tasks performed in remote inspection and navigation, the advantages of a stereoscopic view over monoscopic is negligible.

Previous works with monoscopic camera systems have used 360° panoramic cameras [Swain, 2017], [Fiala, 2005] or spherical cameras to capture the surroundings [Krückel *et al.*, 2015]. To increase the field of view (FoV), multiple wide angle cameras with overlapping views can be stitched together to form a higher resolution image. For situations where bandwidth is limited, sending a partial view and using low bandwidth sensors (other than cameras) to reconstruct the environment virtually [Hosseini and Lienkamp, 2016] as mixed reality (MR) can negate the problem.

Another area of research using HMDs is to use head orientation to control the robot itself and not the camera module. In [Candeloro *et al.*, 2015] the authors have used the HMD to fully control an underwater remotely operated vehicle (ROV) to free up the hands of the operator to control a manipulator. Using the HMD for yaw control was tested in [Martins *et al.*, 2015], with results showing it is confusing and problematic, leading to decreased performance compared with no control.

The development of Robot Web Tools [Toris *et al.*, 2015] allows robots running on Robot Operating System (ROS) [Quigley *et al.*, 2009] to be easily controlled via a web browser. The suite allows images, point clouds and other information to be sent to the operator across compressed JSON streams. The authors mentioned future technologies such as WebRTC can bring greater responsiveness and power to portable web apps.

3 Augmented Telepresence Web Application

Our research goal is to create a cross-compatible system for remote robot telepresence that is low-cost, lightweight and easily deployable on the field. The current state-of-the-art works have all used large and heavy components not suitable for a legged robot where size and weight is limited. Consequently, our two main design

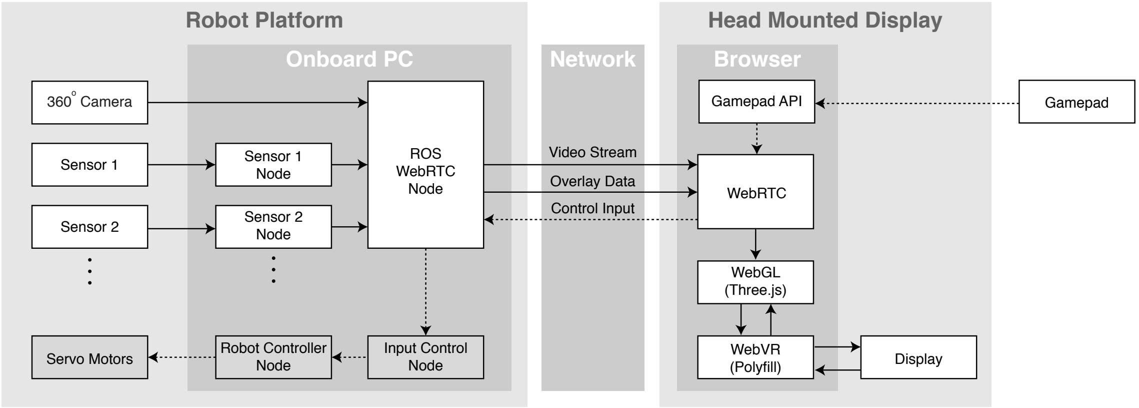

Figure 2: System overview. Various web technologies are used to create the 360° view for a mobile HMD.

decisions are to develop our application using web technologies with ROS integration and being device-agnostic. Although previous work in HMD telepresence have used networking and web apps to control the robot from a base station, none have used a system that is device-agnostic. Web technologies allow for any device with a browser to view and control the robot, reducing the need for custom hardware.

Figure 2 illustrates the different web technologies used in the system. A 360° camera on the robot is connected to the on-board PC. The video stream is then transmitted over a wireless network to the user wearing an HMD. Similarly, we connect several sensors and stream their data in real-time to the HMD using the same network interface. A user remotely controls the robot through a gamepad and the control input is wirelessly streamed back to the robot controller.

The core technology used in the HMD application is WebRTC [Google Inc., 2017a], a framework that allows Real Time Communications in the browser. Using a server to handle signalling (session control), WebRTC enables peer-to-peer (or one-to-many) video and audio streaming. One or more users can access the remote teleoperation video stream by simply opening the appropriate webpage in their browser. We only allow one user to control the robot through the gamepad at any one time to avoid receiving conflicting user inputs.

The system uses 'ros-webrtc' created by Mayfield Robotics [Mayfield Robotics, 2017] to expose Google's implementation of WebRTC to ROS. The client web browser connects to 'ros-webrtc' on the robot to create peer-to-peer video and audio streams. This is a one-way video stream from the robot to the browser, while bidirectional data is exchanged via other data channels. These data channels use rosbridge [Toris *et al.*, 2015] to provide a JSON API to ROS functionality for the web application to interact with the ROS nodes running on the robot.

The video stream received by WebRTC on the browser is passed into a 3D library for mapping and rendering. WebGL allows GPU acceleration for 3D graphics in web browsers, with Three.js [three.js, 2017] being the JavaScript 3D library chosen. A sphere is created with the 360° video UV mapped onto the inside of the sphere. A perspective camera is fixed at the centre of the sphere with orientation being controlled by the user's view.

WebVR API is a draft specification that exposes VR devices to web apps. This experimental technology implemented by web browsers allows web apps to display VR content and interact with the device's head tracking sensors. Due to it's limited browser compatibility at the time of writing, a polyfill [Google Inc., 2017b] was used to make it compatible with current browsers and devices. The polyfill tracks the user's point-of-view via the device's IMU and sends the orientation data to the 3D library to control the camera. A mouse/keyboard interface for non-VR devices such as computers is also provided, allowing the web application to be device-agnostic.

User control information is provided through the Gamepad API, which allows the web application to map any gamepad device input to a standard layout and send the data to the robot. This generic information can subsequently be mapped to the robot's specific controls.

3.1 Augmented Telepresence Overlay

Additional information from sensors can be overlaid onto the view as well. Using a 360° view allows different information to be displayed at various locations, increasing the screen area available. Two overlays developed include a temperature sensor and a Lidar. Information is streamed to the web application through rosbridge data channels. The temperature sensor is located on the robot and we display its data on the robot's body. Lidar information collects data from the environment, which we post-process to detect obstacles. We use blue-purple-red colour coding for regions in the field of view where obstacles are detected, with blue representing an obstacle far away and red representing an imminent collision.

4 Hardware Implementation

The system is implemented onto a hexapod robot platform to test its functionality. The hardware used in testing the effectiveness of the augmented telepresence system for remote inspection tasks is shown in Figure 1 and Figure 4.

4.1 Hexapod Platform

The legged robot platform used is CSIRO's Weaver, a 30 degrees of freedom (DoF) hexapod. Each of its six legs have five joints, allowing efficient and stable locomotion in unstructured environments. Additional capabilities of Weaver are described in [Bjelonic *et al.*, 2016], [Homberger *et al.*, 2016] and [Bjelonic *et al.*, 2017]. The on-board PC is an Intel i5 NUC with 16 GB RAM running ROS in an Ubuntu environment. Weaver's controller allows omni-directional motion (forwards, backwards, strafe left and right), yaw control and body posing. Dimensions of Weaver and mounting location of the payload is given in Figure 5. The single on-board PC runs Weaver's control systems and the WebRTC node.

4.2 360° Visual System

The 360° visual system consists of a Ricoh Theta S [Ricoh Company Ltd., 2017a] mounted above a Hokuyo UTM30LX. The Theta S is selected ahead of other consumer cameras due to its ability to live stream 360° spherical video with minimal delay via USB/HDMI. A HDMI to USB capture device allows a 1080p 30 fps stream from the Theta S (the Theta S is limited to 720p 15 fps streaming via USB) to appear as a webcam to the computer. The Lidar provides forward facing 270° 2D range data for the augmented telepresence overlay.

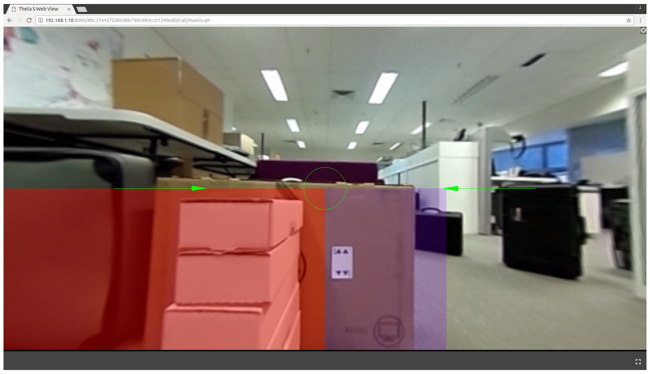

Figure 3: View of the augmented telepresence overlay in providing range data to the user when viewed from a computer display. The green circle and arrows provide the forward direction of the robot when the user looks around. This is fixed to the robot's frame.

Figure 4: Hardware used; Samsung Galaxy S7 mobile phone with Samsung GearVR (left) and Ricoh Theta S 360° camera with a Hokuyo UTM30LX Lidar (right).

Table 1: Gamepad control layout.

Button	Function
Left Joy Up & Down	Forwards & Backwards
Left Joy Left & Right	Strafe Left & Right
Right Joy Left & Right	Rotate Left & Right
B Button	Hide Laser Overlay
D-Pad Up	Reset Camera Position

The live streaming output of the Theta S is a dual-fisheye view, which is sent via WebRTC for the HMD system to stitch together and display. The stitching process occurs within the 3D library and is based on the UV texture provided by Ricoh in [Ricoh Company Ltd., 2017b].

For the Lidar based distance overlays, the full 360° camera view is divided into 12 × 30° longitudinal regions. This results in 6 × 30° regions each on the left and right of the forward direction of the robot. The 270° FoV of the 2D Lidar is aligned on to the 360° camera view such that it covers 135° on either side of the forward direction of the robot. Then the shortest dis-

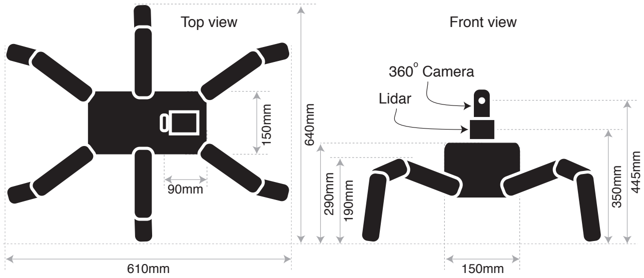

Figure 5: Dimensions of the hexapod robot Weaver along with positioning of the 360° camera and Lidar on the robot platform.

tances in each of the 10 regions (2 × 30° regions do not overlap at the back of the robot) that overlap with the Lidar FoV are mapped to 10 corresponding longitudinal regions in the camera view. Depending on the corresponding shortest distance, each region is overlaid with the following colours: blue (<1.0 m), medium purple (<0.6 m), magenta (<0.5 m) and red (<0.4 m).

4.3 HMD System

To visualise and control the robot, a consumer VR headset and gamepad is used. Smartphone based headsets such as the Samsung GearVR and Google Daydream allow for a portable, low-cost system compared to previous works with bulky systems such as the Oculus Rift. These headsets allow the system to be easily and quickly deployed at a remote site without the need for additional infrastructure. A bluetooth gamepad is connected to the smartphone to send control commands via the web application to the robot. The gamepad layout is summarised in Table 1. The number of controls is limited to suit the HMD as initial testing found users instinctively looked down at the controller if they wanted to press a button (which does not work in HMD mode). Furthermore, body posing was deactivated, which could help orientate the camera. A Samsung GearVR with a S7 Edge smartphone and a Steelseries Stratus XL gamepad is used in the experiments. The system has been successfully tested on Google Daydream with a Pixel smartphone.

The hardware chosen allows a single operator to transport and setup the robot for remote inspection. The robot creates an access point which the operator directly connects their smartphone to for control. This simplifies and reduces deployment time.

5 Experiments and Results

Hardware experiments were conducted to evaluate the performance of the web application.

5.1 Experimental Setup

A test course was created to evaluate the system's performance to reflect a real-life remote inspection task. For comparison, a traditional approach for remote inspection using a forward mounted camera with a wide FoV was simulated with the 360° camera, by displaying a fixed forward view on a 13 inch laptop.

The course, shown in Figure 6, consists of narrow passages (125% of leg span), turns, covered areas, open areas and low obstacles. To simulate inspection points, playing cards were placed at designated positions marked in Figure 6. Most were placed at camera level or higher, with one placed on the underside of an obstacle where the robot can crawl under. The face value of the cards were used to determine accuracy, with each correct identification worth one point. There was a total of 10 cards with participants not knowing the total number of cards beforehand.

5.2 Participants

A total of 28 participants (average age of 24.93, SD = 6.33) split into two groups (13 males and 1 female each), completing the test course with either the display or the HMD. The participants were told beforehand that they would be assessed on completion time, the accuracy of inspection points and the number of collisions; with particular emphasis on not hitting anything for the robot's well-being. Unknown to the participants, only the results from the second test course run was being recorded.

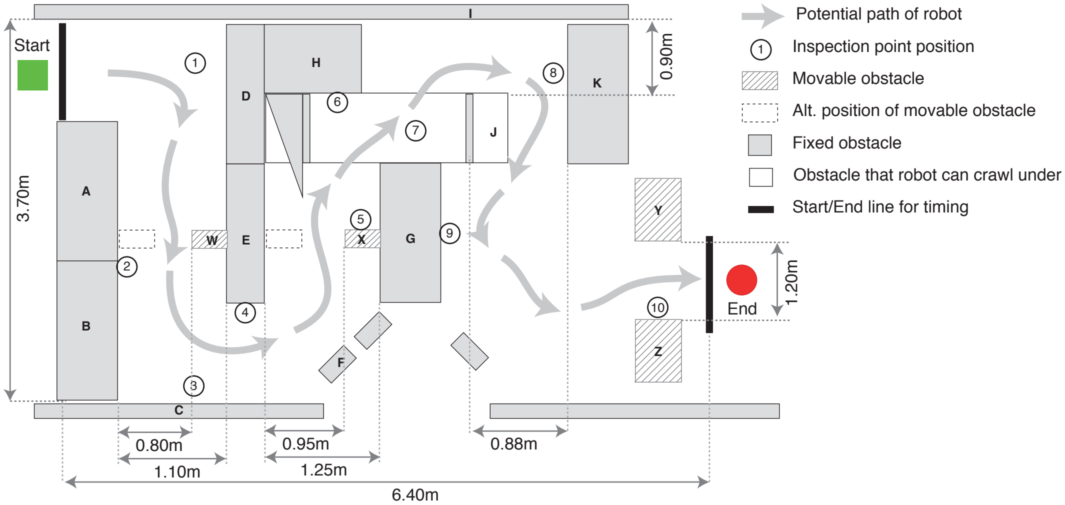

Figure 6: The test course layout with W,X,Y,Z being movable obstacles. W and X were moved to either the left or right side of the passages. The numbered positions indicate inspection point locations, with 7 being on the underside of the obstacle.

Table 2: Technical specifications achieved on Weaver. Brackets indicate minimum and maximum values observed.

Transmitted resolution	Usable resolution	Video frame rate	Avg. latency	Avg. bandwidth usage	Avg. HMD frame rate	Avg. onboard CPU usage
1920×1080	1920×960	30 fps	850 ms (500 ms-1.3 s)	1.2 MB/s (400 kB/s-2.9 MB/s)	53 fps (42-60 fps)	48% (29%-72%)

This removes learning effects and unfamiliarity of controlling the robot. The course was adjusted (boxes W and X shifted) and the playing cards randomly shuffled between the trials to reduce learning effects of the course.

5.3 Measures

To capture the suitability and performance of the system, objective and subjective measures were used. While the objective measures were the main focus in the evaluation, user feedback of the system was collected to guide future improvements.

Objective Measures

Three criteria were chosen to measure the participants' performance using the system: task completion time, number of correct inspection points, and number of collisions. Task completion time was recorded as the time between the front legs passing the start line and crossing the finish line. Inspection points were recorded as the participant verbally said which card they saw. Collisions were manually recorded as instances when any part of the robot touches an obstacle. Additionally, number of topples were recorded for when obstacle W and/or X were pushed over while navigating the narrow passages.

Subjective Measures

Users were asked after the second trial to answer four questions, each on a 1 to 7 scale (1 - strongly disagree, 4 - neither, 7 - strongly agree). The questions were:

A - I found the view helpful in perceiving the robot's environment

B - I found the view comfortable and easy to use

C - I was able to complete the tasks easily

D - I found the laser overlay to be helpful

5.4 Results

The test course was designed to reflect real-life inspection tasks with the combination of measures critical to a successful operation. Participants had repeated trials of the test course to gain familiarity of the system and reduce the novelty of the controls. While learning effects of the test course exist, this is common for both groups. ANOVA was used with a p-value of 0.05 set as the threshold for statistical significance.

System Performance

Technical specifications achieved on the system are provided in Table 2. The bandwidth usage varied depending on the robot's state. While the robot was stationary, bandwidth usage was around 500 kB/s - 800 kB/s; and while it was moving, usage was around 1 MB/s - 2.5 MB/s. WebRTC adjusted the video resolution and bitrate when available bandwidth changed. The latency experienced varied, with WebRTC re-syncing the video stream once latency was high. The value for average latency in Table 2 does not include when the latency times out for a re-sync. The latency is the total delay between a movement command on the gamepad and the video stream showing the robot moving, separate from the HMD latency of moving the view which is limited by the HMD frame rate.

Objective Measures

The authors hypothesise that the HMD technology will provide benefits over traditional display technology that translate to faster completion time, increased inspection

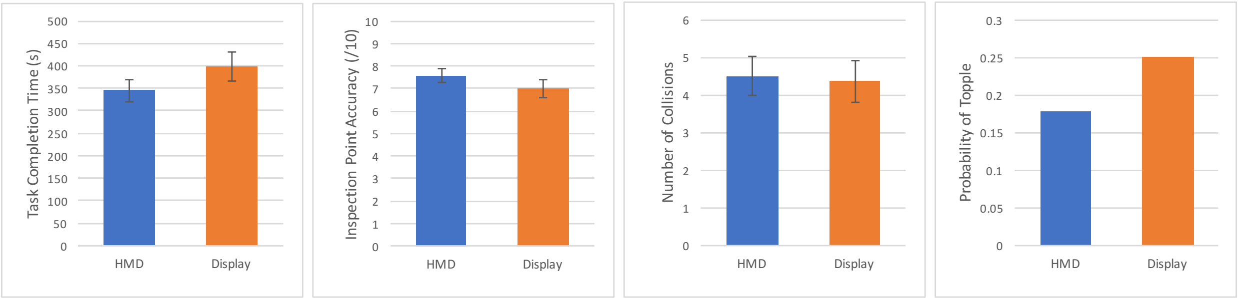

Figure 7: Participant performance completing the test course. The error bars show standard error.

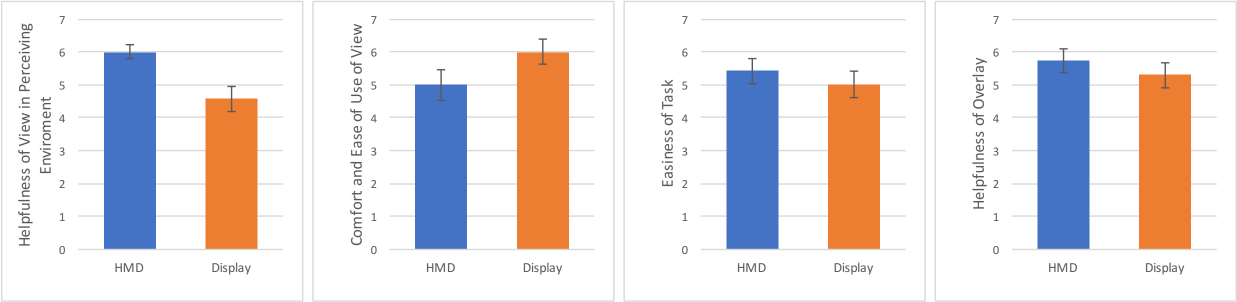

Figure 8: Participant responses to questionnaire on a scale of 1 to 7. The error bars show standard error.

accuracy, and a lower rate of collision. While the results show that HMD provides benefits over traditional display technology, the results are less pronounced than expected (see Figure 7).

There is partial support that task completion time with the HMD is lower than the display. The average time to complete the course in seconds for HMD (M = 345.43, SD = 91.61) and display (M = 399.00, SD = 122.93) supports this hypothesis, but no significant difference was found F(1,26) = 1.71, p = 0.203.

The results for the inspection point accuracy and number of collisions show no significant differences between the two groups. The HMD provides the participants a wider FoV, resulting in a slightly higher inspection point accuracy of M = 7.57, SD = 1.16 compared to display with M = 7.00, SD = 1.41. However, the number of collisions is slightly higher with the HMD, with M = 4.50, SD = 1.91 compared to M = 4.36, SD = 2.10. The results suggests a 360° FoV does not provide an advantage or disadvantage in spotting inspection points, F(1,26) = 1.37, p = 0.253 or avoiding collisions, F(1,26) = 0.04, p = 0.852, respectively.

There is a higher probability of causing a topple using the display (0.25) than that of the HMD (0.18). The HMD offers better situational awareness, resulting in a slight advantage of having less destructive impacts.

Subjective Measures

Subjective measure results are shown in Figure 8. Users were asked at the completion of the second trial of the test course to fill in the questionnaire. Participants found that the HMD helped them perceive the robot's environment better, F(1,26) = 7.83, p = 0.01. Average responses for HMD were more than 1 point higher than the display, M = 6.00, SD = 0.68 and M = 4.57, SD = 1.79 respectively.

While the HMD gave the user greater perception of the robot's environment, this did not translate to a noticeable increase in the easiness of the task. No significant increase was recorded, with F(1,26) = 0.46, p = 0.506. This could be linked to the results of the comfort and ease of use of the view. Participants found the display to be easier to use compared to the HMD, with results of M = 6.00 SD = 1.41 and M = 5.00, SD = 1.30 respectively. This link however is not significant, F(1,26) = 3.79, p = 0.062.

Overall, both groups found the laser overlay (shown in Figure 9 and 10) to be helpful in completing the task. The subjective measure results provides insight into the user's perceived performance of the task.

Figure 9: View rendered by the polyfill for the GearVR showing temperature information.

Figure 10: Laser overlay showing obstacle in close proximity with the middle right leg.

6 Discussion

The experiments show that both viewing methods in our web application, i.e., the forward facing display on a PC screen and the 360° HMD system, allow users to complete a relatively complex remote inspection task successfully. While our hypothesis of HMD technology introducing a significant advantage over a traditional display approach was not confirmed, the HMD technology was able to improve the operators situational awareness of the robot's environment.

The user's perception of the remote environment is increased, though no significant advantages were measured in the tasks. The task completion time, inspection point accuracy and number of collisions all showed no significant improvements of the HMD over the display. The similar task completion times reflect on how the participants are able to control the robot quickly and efficiently. To move the fixed camera of the display, the whole robot is required to be moved. Due to Weaver's ability to rotate within its footprint, the robot can turn around safely without collision. Although the HMD gives a 360° view of the robot, some participants would rotate the robot to see or to walk backwards, without utilising omni-directional walking. These advanced manoeuvres could be performed quickly by an experienced operator. The similar number of collisions suggests that the laser overlay in both approaches and the robot's ability to rotate on the spot reduces the advantages of the 360° view.

The HMD could have caused effects such as motion sickness or fatigue in some participants, with this being reflected in the comfort and ease of use results. As only some participants were affected by this, the decrease in comfort of the HMD was not statistically significant. The results of the easiness of task (question C) summarises the results from both questions A and B. This shows that the perceived easiness of the overall task is influenced by the combination of the situational awareness in the remote environment and the comfort of the viewing interface.

7 Conclusions

This work introduced a web application to teleoperate a legged robot using a mobile HMD. A 360° camera and Lidar is used to capture the remote environment, providing an augmented telepresence view to the user. A test course was built to simulate a remote inspection task. The HMD system was compared with a fixed camera display approach with results suggesting a slight improvement in objective measures. In subjective measures, the HMD system significantly increased the operator's perception of the remote environment, which slightly decreased task difficulty. However, the differences between the two technologies is not significant. Thus, the use of a 360° camera and a HMD gives the user more confidence in understanding the robot's environment but gives no significant advantages in task performance.

To improve the augmented telepresence system and increase autonomy, 2D SLAM from the Lidar will be implemented to provide a map, information on the space/width of passageways the robot can explore and obstacle avoidance. Instructions and navigation directions for the task can also be augmented into the view to aid in mission completion. Further testing of the augmented telepresence system in different scenarios such as stepping over ledges or passageways narrower than the normal leg span, would provide a better understanding on the tasks which the HMD could improve performance.

Acknowledgements

The authors would like to thank Eranda Tennakoon, Andreas Duenser and David Rytz for their support during the project.

References

[Aracil *et al.*, 2007] R. Aracil, M. Buss, S. Cobos, M. Ferre, S. Hirche, M. Kuschel, and A. Peer. The human role in telerobotics. In *Advances in Telerobotics*, pages 11–24, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.

[Bjelonic *et al.*, 2016] M. Bjelonic, N. Kottege, and P. Beckerle. Proprioceptive control of an over-actuated hexapod robot in unstructured terrain. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2042–2049, Oct 2016.

[Bjelonic *et al.*, 2017] M. Bjelonic, T. Homberger, N. Kottege, P. V. K. Borges, P. Beckerle, and M. Chli. Autonomous navigation of hexapod robots using controller adaptation. In *IEEE International Conference on Robotics and Automation (ICRA) (to appear)*, 2017.

[Candeloro *et al.*, 2015] M. Candeloro, E. Valle, M. R. Miyazaki, R. Skjetne, M. Ludvigsen, and A. J. Sørensen. HMD as a new tool for telepresence in underwater operations and closed-loop control of ROVs. In *MTS/IEEE OCEANS*, pages 1–8, 2015.

[Fiala, 2005] M. Fiala. Pano-presence for teleoperation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3798–3802, 2005.

[Fong and Thorpe, 2001] T. Fong and C. Thorpe. Vehicle teleoperation interfaces. *Autonomous Robots*, 11(1):9–18, 2001.

[Google Inc., 2017a] Google Inc. WebRTC, February 2017. webrtc.org.

[Google Inc., 2017b] Google Inc. WebVR Polyfill, February 2017. github.com/googlevr/webvr-polyfill.

[Homberger *et al.*, 2016] T. Homberger, M. Bjelonic, N. Kottege, and P. V. K. Borges. Terrain-dependent motion adaptation for hexapod robots. In *International Symposium on Experimental Robotics (ISER)*, 2016.

[Hosseini and Lienkamp, 2016] A. Hosseini and M. Lienkamp. Enhancing telepresence during the teleoperation of road vehicles using hmd-based mixed reality. In *IEEE Intelligent Vehicles Symposium (IV)*, pages 1366–1373, 2016.

[Kratz and Ferriera, 2016] S. Kratz and F. Rabelo Ferriera. Immersed remotely: Evaluating the use of head mounted devices for remote collaboration in robotic telepresence. In *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 638–645, 2016.

[Krückel *et al.*, 2015] K. Krückel, F. Nolden, A. Ferrein, and I. Scholl. Intuitive visual teleoperation for UGVs using free-look augmented reality displays. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 4412–4417, 2015.

[Livatino *et al.*, 2009] S. Livatino, G. Muscato, and F. Privitera. Stereo viewing and virtual reality technologies in mobile robot teleguide. *IEEE Transactions on Robotics*, 25(6):1343–1355, Dec 2009.

[Mania *et al.*, 2004] K. Mania, B. D. Adelstein, S. R. Ellis, and M. I. Hill. Perceptual sensitivity to head tracking latency in virtual environments with varying degrees of scene complexity. In *Symposium on Applied Perception in Graphics and Visualization (APGV)*, pages 39–47, New York, NY, USA, 2004. ACM.

[Martins *et al.*, 2015] H. Martins, I. Oakley, and R. Ventura. Design and evaluation of a head-mounted display for immersive 3D teleoperation of field robots. *Robotica*, 33(10):2166–2185, 12 2015.

[Mayfield Robotics, 2017] Mayfield Robotics. ros-webrtc, February 2017. github.com/mayfieldrobotics/ros-webrtc.

[Quigley *et al.*, 2009] M. Quigley, K. Conley, B. P. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng. ROS: an open-source robot operating system. In *ICRA Workshop on Open Source Software*, 2009.

[Ricoh Company Ltd., 2017a] Ricoh Company Ltd. Ricoh Theta S, February 2017. theta360.com/en/about/theta/s.html.

[Ricoh Company Ltd., 2017b] Ricoh Company Ltd. Ricoh Video Streaming Samples, February 2017. github.com/ricohapi/video-streaming-sample-app.

[Swain, 2017] R. Swain. Remote Excavation: Using WebRTC and Real-Time Video with an Eye on 5G - Ericsson Research Blog, February 2017. www.ericsson.com/research-blog/5g/remote-excavation-using-webrtc-real-time-video-eye-5g/.

[three.js, 2017] three.js. Three.js, February 2017. threejs.org.

[Todd, 1985] D. J. Todd. A brief history of walking machines. In *Walking Machines: An Introduction to Legged Robots*, pages 11–33, Boston, MA, 1985. Springer US.

[Toris *et al.*, 2015] R. Toris, J. Kammerl, D. V. Lu, J. Lee, O. C. Jenkins, S. Osentoski, M. Wills, and S. Chernova. Robot web tools: Efficient messaging for cloud robotics. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4530–4537, 2015.

[Yanco *et al.*, 2004] H. A. Yanco, J. L. Drury, and J. Scholtz. Beyond usability evaluation: Analysis of human-robot interaction at a major robotics competition. *Human–Computer Interaction*, 19(1-2):117–149, 2004.