2024 I-RIM Conference
October 25-27, Rome, Italy
ISBN: 9788894580556
10.5281/zenodo.14730987

Teleoperation of Boston Dynamics Spot Robot using Meta Quest 2 Virtual Reality Headset

Ali Yousefi[1], Zoe Betta[1], Giovanni Mottola[1], Carmine Tommaso Recchiuto[1], and Antonio Sgorbissa[1]

***Abstract*—This extended abstract investigates an immersive teleoperation system where a human operator controls a Spot robot. The robot, equipped with a stereo camera, provides the operator with visual feedback via virtual reality (VR) eyewear, allowing them to perceive the environment from the robot's perspective. The operator's head movements control the robot's pitch and yaw angles, while joysticks are used for locomotion. The study seeks to determine whether this immersive control system enhances the operator's ability to identify important environmental features and includes a user study to test the effectiveness of the approach.**

***Index Terms*—Teleoperation, Quadruped Robots, Virtual Reality**

I. INTRODUCTION

Teleoperation of mobile robots with immersive control has been explored in various studies, showing that stereo cameras with depth information paired with a Head-Mounted Display (HMD) enhance situational awareness and immersion compared to 2D cameras and monitors [1]. This approach has proven effective in search and rescue scenarios [2], with studies often focusing on controlling a pan and tilt camera mounted on the robot [1]. For quadruped robots, immersive control can be more intuitively and efficiently implemented directly on the robot. Recent research on quadrupeds with Extended Reality (XR) systems has primarily focused on manipulation tasks [3], rather than inspection. Halder et al. [4] have explored inspection tasks using a Boston Dynamics Spot robot with a 360° camera. Their method lacked depth information, suggesting that stereo cameras with depth sensors could be beneficial. Our work introduces a novel control system for quadruped robots using a stereo camera with depth sensors to provide 3D feedback and head-tracking, compared against the standard controller of the robot.

II. METHOD

In the proposed method, the head pose tracking task is implemented using a closed-loop velocity control system shown in Figure 1, focusing on angular velocities along the robot's y and z axes denoted as $\omega_y^{ROB}(t)$ and $\omega_z^{ROB}(t)$. The system tracks reference angular velocities from an HMD denoted as $\omega_x^{HMD}(t)$ and $\omega_y^{HMD}(t)$, though the HMD and robot frames have different axis definitions as it is shown in Figure 2. The control input signals, $C_y(t)$ and $C_z(t)$, derived from Equation 1, are used to control the robot with gains K_y and K_z, and are constrained within safe boundaries.

[1] RICE lab, University of Genoa, via Opera Pia 13, Genova, Italy.
Contact author: ali.yousefi@edu.unige.it

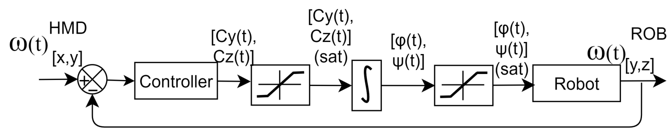
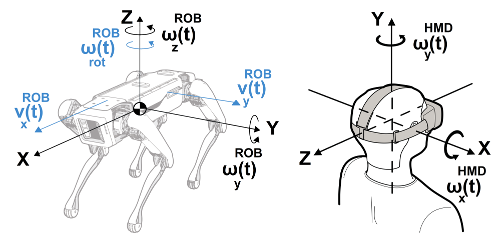


Fig. 1: Block diagram of the control system.



(a) Robot base frame (b) HMD frame

Fig. 2: Frame definitions.

$$
\begin{aligned}
C_y(t) &= K_y(-\omega_x^{HMD}(t) - \omega_y^{ROB}(t)),\\
C_z(t) &= K_z(\omega_y^{HMD}(t) - \omega_z^{ROB}(t)).
\end{aligned} \tag{1}
$$

Since the robot command service only accepts angular values for pitch and yaw, these are computed via discrete integration of the control signals. The pitch ($\phi(t)$) and yaw ($\psi(t)$) are also limited by the robot's joint constraints.

For locomotion, thumbstick inputs control the robot's linear velocities along the x and y axes denoted as $v_x^{ROB}(t)$ and $v_y^{ROB}(t)$, and rotational velocity along the z axis denoted as $\omega_{rot}^{ROB}(t)$. Once the thumbstick passes a dead-zone, the robot receives a constant velocity input for movement. The system allows the operator to control the robot's movement and orientation simultaneously using the HMD inputs, enhancing the operator's sense of presence in the environment, a feature not possible with the standard tablet-based Spot controller due to its limited control options and mode-switching delays.

The system architecture of this study, shown in Figure 3 consists of two main components: the User Side and the Robot Side. On the Robot Side, a Spot robot is equipped with a Stereolabs ZED 2 stereo camera, an NVIDIA Jetson Nano, and a Raspberry Pi 4. The ZED 2 captures stereo images at 720p60 with H.265 compression, and the Jetson Nano processes these images for real-time transmission. Since the Jetson Nano lacks wireless capabilities, the Raspberry Pi creates a hotspot for network communication. On the User Side, an Oculus Quest 2 is used for immersive control, requiring an additional Windows-based PC (OMEN HP) to render the stereo images

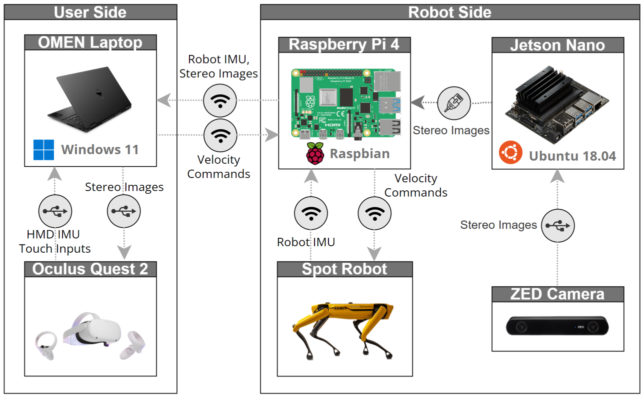


Fig. 3: System architecture.

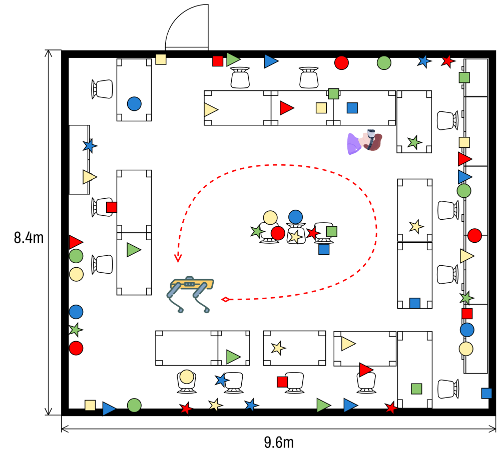


Fig. 4: Each configuration is represented with a different color (1: yellow, 2: green, 3: blue, 4: red).

and handle control inputs like yaw, pitch, and locomotion commands.

A user study was conducted to evaluate a proposed system architecture, involving 20 participants with ages ranging from 24 to 38 years ($M = 28.00, SD = 3.93$). The study took place in an 8.4 × 9.6 m² environment with obstacles and 16 hidden shapes (circles, squares, triangles, stars) placed at different heights. Participants controlled the robot to identify these shapes using two methods: the proposed immersive control system (method A) and the standard tablet controller (method B). They had to navigate a path and verbally identify shapes. Each participant completed two trials for each control method, with different shape configurations as it is shown in Figure 4. The participants were divided into four groups, with each group following a different sequence of control methods and configurations to avoid bias. Performance was measured by the number of shapes found per minute and the time taken to complete the path.

Additionally, participants completed a Presence Questionnaire (PQ) to assess environmental awareness and a Simulator Sickness Questionnaire (SSQ) to evaluate discomfort before and after using the HMD.

III. Results

The data analysis involved pairwise comparisons. Since the sample size was under 30, normality was tested using the Shapiro-Wilk method. Normally distributed data were analyzed using the paired-samples t-test, while non-normal data were analyzed using the Wilcoxon signed-rank test, with a significance level of $\alpha = 0.05$. Key findings include the HMD outperforming the tablet in terms of the number of shapes found per minute in both trials, and improved performance for both control methods in the second trial. The time on task was shorter for the HMD compared to the tablet, and both control methods had shorter time on task in the second trial. Considering PQ results, the HMD had higher involvement, sensory fidelity, and immersion scores, but no difference was found in interface quality. The SSQ results showed that oculomotor disorientation, nausea, and disorientation symptoms were significantly higher after using the VR system.

IV. Discussion

The presented approach shows that immersive control of a quadruped robot using advanced virtual reality eyewear with 3D visual feedback is both feasible and advantageous over the standard Spot controller. This method improves operational efficiency by allowing better adjustment of the robot's attitude during locomotion, reducing operation time and enhancing maneuverability. User studies revealed that participants performed tasks more efficiently and in less time with immersive control, showing better outcomes after an initial learning period. However, participants with no VR experience and those using the system for extended periods experienced increased distress. Limitations include a small participant sample, the disparity in field of view between the HMD and tablet, and constraints related to wireless bandwidth that should be considered in future works.

Acknowledgment

This work was carried out within the framework of the project "RAISE - Robotics and AI for Socio-economic Empowerment" Spoke 4 and has been supported by European Union – NextGenerationEU.

References

[1] B. Illing, M. Westhoven, B. Gaspers, N. Smets, B. Bruggemann, and T. Mathew, "Evaluation of immersive teleoperation systems using standardized tasks and measurements," *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, Aug 2020, doi:10.1109/ro-man47096.2020.9223497.

[2] H. Martins, I. Oakley, and R. Ventura, "Design and evaluation of a head-mounted display for immersive 3d teleoperation of field robots," *Robotica*, vol. 33, no. 10, p. 2166–2185, May 2014, doi:10.1017/s026357471400126x.

[3] J. Humphreys, *et al.*, "Teleoperating a legged manipulator through whole-body control," *Towards Autonomous Robotic Systems*, p. 63–77, 2022, doi:10.1007/978-3-031-15908-4_6.

[4] S. Halder and K. Afsari, "Real-time construction inspection in an immersive environment with an inspector assistant robot," *EPiC Series in Built Environment*, 2022, doi:10.29007/ck81.