HALO: a Rock Scaling Mobile Manipulator with Interactive Virtual Reality Live Digital Twin

Dinh Tung Le, Sheila Sutjipto, Dac Dang Khoa Nguyen, Thanh Long Vu,
Pham Dinh Thong Vu, Nuwan Munasinghe, Gavin Paul *

Abstract

The High Access Localised Operations (HALO) system is a mobile manipulator that amalgamates advancements in digital twins, and virtual reality (VR) to enable safer rock scaling operations in the mining industry. Currently, the essential geotechnical activity of rock scaling is performed by certified workers suspended on the side of the rock wall, who perform the physically demanding task of removing loose rock debris. The HALO system enables a remote operator immersed in a VR environment that visualises the digital twin of the robot and its sensor data in real-time to interact with the robot intuitively. This eliminates the need for people to be exposed to the hazards associated with performing manual rock scaling, while enabling them to apply their existing expertise when operating the robot in VR. In this work, we present a summary of the HALO hardware and its interaction architecture, encompassing a framework for real-time remote scene reconstruction and natural interaction. Findings from preliminary site trials are also be presented to provide preliminary evaluation of the system.

1 INTRODUCTION

Rock scaling involves the removal of unstable or loose rock from cliffs or slopes that have deteriorated over time due to natural factors such as rainfall or human-induced vibrations. This operation is essential in mining to mitigate rockfall risk and ensure the safety of workers and equipment in open-cut mines. A range of slope stability technologies exist for predicting and detecting rockfalls, such as radar monitoring systems [1], 3D laser scanning [2], and 3D seismic surveying [3], which analyse changes in rock mass movement and provide early warning of instability. However, these systems are ineffective in detecting more minor rock failures that still pose a fatal threat to workers. Therefore, rock scaling is still the most effective and in-demand method to manage rockfalls.

Figure 1: An operator performing rock scaling remotely in VR with HALO.

Researchers continue to explore the use of semi-autonomous and teleoperated robots to perform dull, dirty, and dangerous tasks, such as screw-fixing installation in construction [4], steel bridge inspection [5], abrasive blasting for metal preparation [6] and bridge maintenance [7]. However, abseiling on rock walls using ropes is a unique and uncommon locomotion modality. Most abseiling and rock scaling robots are lightweight and equipped only for inspection, lacking a manipulator with enough force to remove rocks from a wall. An example is NASA's *LEMUR* [8], which uses microspine grippers to assist with its climb on rocky surfaces but would only work in reduced-gravity environments like Mars.

In recent years, remote-controlled equipment, such as bulldozers, excavators, and mobile drill rigs, have become popular in the mining industry to improve safety in high-risk tasks. One example is the *Jayben S60 Scaler* [9], a prototype 18-ton remote scaling machine based

---

[*] Authors are with the UTS Robotics Institute, University of Technology Sydney, Australia `robotics.uts.edu.au`

on a *Caterpillar 308 E2CR* excavator. The machine, equipped with an 2D cameras for visualisation of the work area, is lowered down the rock face using two mounted hi-torque winches. During real-world testing on a mine, researchers found that it had limited manoeuvrability on flat benches due to its added mass, and its hydraulic slew failed since typical excavators are not designed to work at a 90-degree vertical angle. Another major challenge that retrofitted equipment like the *Jayben S60* face is the usability and effectiveness of the data visualisation and control interface. Most remote-controlled heavy machinery relies on an enclosed remote control station with computer monitors to view camera feeds and control panels with joysticks, a keyboard, and a mouse. This setup hinders industry's shift from manual to automated or assisted task execution with remote robots, demanding considerable training while removing an operator's crucial spatial perspective for decision-making.

One promising approach to address these issues is utilising *Extended Reality* (XR) technologies as the interaction medium. Operators in XR can leverage intuitive actions and expressions, spatial reasoning, and controllability with exceptional accuracy and speed facilitated by the tracked headset and controllers. Conversely, the pose of a robot model and external sensor 3D data, such as point clouds, are best presented in native 3D. *Augmented Reality* (AR) has been shown to outperform 2D interfaces in local Human-Robot Interaction (HRI) scenarios, enhancing both task performance and user experience in collision detection [10], learning by demonstration [11], trajectory modification [12], and high-level goal setting [13] tasks. Operators can co-localise with their remote robotic partners by means of a digital twin in *Virtual Reality* (VR), which displays the robot and environment that updates in real-time. VR has been employed for remote HRI in *egocentric* mode for dexterous manipulation tasks with humanoid robots [14] and *exocentric* mode for setting waypoints of mobile robots [15], drones [16], and mobile manipulators [17]. Although the field is constantly evolving, there are still ongoing challenges to address. One major challenge is constructing the dynamic environment digital twin from the robot's external data, which the operator relies on to make decisions based on spatial relativity between themselves, the robot, and the remote environment. To provide an immersive VR experience, the dynamic environment digital twin must be visualised in near real-time, meaningfully presented, and potentially accompanied by assistive virtual annotations.

This paper combines advancements in digital twins and VR to describe the HALO robot system, which enables safer rock scaling operations in the mining industry as shown in Fig. 1. The breakdown of this paper is as follows: Section 2 presents an overview of the system. Section 3 describes the interactive architecture of the system, which enables the operator to be virtually *in situ* and intuitively interact with the system. Section 4 analyses the system's capabilities through laboratory and field testing results. Section 5 identifies the limitations and numerous on-site challenges, which lead to a discussion of the critical outcomes and the ongoing research and development. Finally, Section 6 presents the conclusions and future work.

2 SYSTEM OVERVIEW

The HALO system shown in Fig. 2 is a prototype custom-built mobile manipulator designed to achieve the modes of locomotion and manipulation required for rock scaling operations. The HALO platform, which weighs approximately 117 kg with a 1.5 x 1 m footprint, is made of waterjet-cut steel and aluminium plates, and off-the-shelf aluminium extrusions framing. The system is equipped with RGB-D cameras, motor encoders, and internal and external sensing to accommodate non-light-of-sight remote control from a safe VR station via its live digital twin.

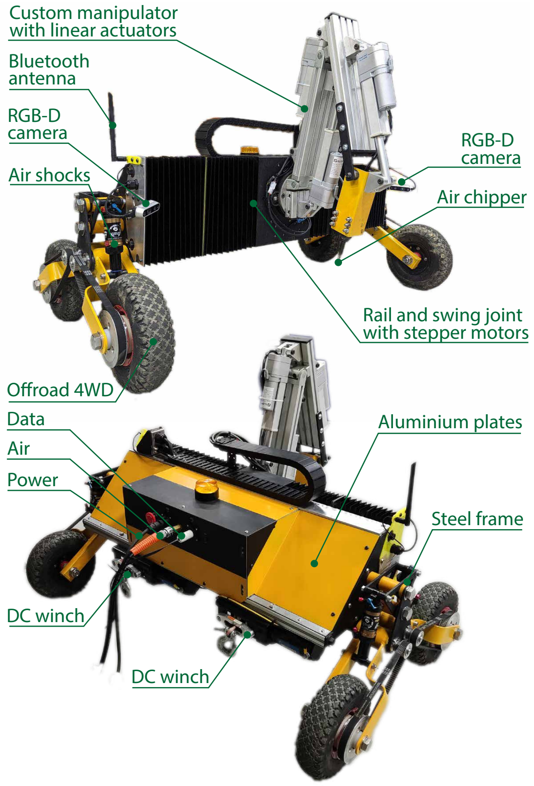


Figure 2: Locomotion, manipulation, sensing and remote communication features of HALO - designed for all-terrain traversal, rock wall abseiling and scaling

2.1 Mobile Platform with All-Terrain Abseiling Locomotion

The design decisions of the mobile platform were aimed at maximising performance and reliability for the use case of rock scaling while minimising complexity and potential points of failure. The platform is a four-wheel differential drive system with each pair of wheels on one side of the robot being connected to the same drivetrain driven by a 180 W worm-geared DC motor, that can produce reliable high torque up to 20 N.m at 150 RPM with minimal back drive. The use of wheeled systems on independent angle-adjustable swing arms was chosen over treads to better facilitate the transition from flat 0° to inclined or vertical 90° surfaces. The platform's passive suspension system with air shocks was found to be sufficient to maintain stability while driving and scaling.

The platform is also equipped with two high-powered winches with synthetic ropes, rated for 4500 lbs line pull with a 153:1 gear reduction ratio, to be attached to anchors at the top and scales down a rock wall by releasing ropes up to 10 metres. As the robot descends a cliff, rope tension can create moments that lift the front wheels or tip the platform forward, depending on its current centre of mass. An early solution explored was adjustable winch positions relative to the robot's body using a hinge joint with linear actuators. However, this resulted in unwanted body twisting. In the final design, the moments are balanced by driving the contacting wheels forward or backward, creating a counteracting moment, thus keeping all wheels on the wall.

2.2 Excavator-inspired Manipulator with Air Chipper End-Effector

The development of a custom robotic manipulator was necessary to meet the unique and demanding requirements of the rock scaling application. Using an off-the-shelf option, such as a serial collaborative or industrial robotic manipulator, would not have provided the necessary power-to-weight ratio for the scale of the platform. The manipulator needed to be powerful enough to lift the robot's weight and leverage it to remove heavy rocks. Inspired by the arm of mini excavators, the custom manipulator featured DC linear actuators with worm drives instead of hydraulics, while retaining the link length ratios and actuator placements. The linear actuators are capable of moving loads up to 1000 N, with a maximum speed of 1.5 cm/s, and hold their position under static loads up to 450 kgf when unpowered. To extend the manipulator's horizontal range and approach angle, a linear rail and swing joint were added, both of which are driven by stepper motors, coupled to a ball screw and reduction gearbox, respectively. To aid with breaking down rocks into smaller chunks before removal, a 100-psi air chipper was incorporated, controlled with an internal solenoid valve. The manipulator has 5 Degrees-of-Freedom (5-DOF) in total and has a reach of 1 metre in an approximately 1 m x 1 m workable area. All motors are equipped with encoders to allow for velocity and position control with the embedded controller.

2.3 Battery Add-on and Direct Line-of-sight Control

The need for direct line-of-sight control with easy start-up capabilities became evident during development. This feature significantly reduces deployment time, especially for positioning the robot on flat benches before scaling. A battery add-on was developed to address this need, providing a convenient untethered power source for the robot. The battery add-on has a capacity of 100 Ah and can be easily attached and removed from the robot, enabling operators to switch between tethered and untethered modes. The robot has a Bluetooth antenna with a 100 m range to facilitate direct wireless control. Operators can operate the robot wirelessly using a DualShock 4 gamepad, with a mapping for intuitive acceleration and steering controls.

2.4 Sensing for Remote Non-line-of-sight Control

The robot is equipped with sensors to provide real-time feedback to the user via its digital twin, enabling remote non-line-of-sight control. An extrinsically calibrated camera system - two Intel RealSense D435 RGB-D cameras with different viewpoints - generates visual information about the external environment in form of coloured point clouds. The cameras were mounted such that they achieved a high visual coverage and limited the data loss from occlusions caused by the manipulator during motion. The internal encoders of the motors in the 5-DOF manipulator and IMU enable the embedded controller to calculate the robot's joint positions and orientation, respectively. The robot's non-visual internal status is also made available with internal humidity, temperature, and power sensors.

2.5 Computing, Networking and VR Station Setup

The onboard computer is a NUC 11 with 16GB RAM and an Intel i5 processor. The DC motor drivers, which are equipped with embedded controllers, communicate with the mobile platform via USB. For the arm, Teensy MCU and stepper motor drivers are used, along with DC motor drivers. ROS is used as the middleware to manage the robot's control nodes, with *MoveIt!* and ros_control used for arm planning and control respectively. In the tethered mode, a wired LAN connection is used to send data to and from the VR control station. The VR control station consists of a laptop with a GTX 1080 GPU, and a Meta Quest 2 headset and controllers.

3 INTERACTION ARCHITECTURE

The interaction architecture aims to enable the user to be virtually *in situ* with the robot and its environment on the rock wall. This medium allows users of the robot, specifically robot access technicians, to utilise their prior knowledge and experience when operating the robot to perform rock scaling. The objective was to create a natural and intuitive interaction interface, facilitating a seamless transition from manual rock scaling to robot operation.

The interaction architecture consists of two primary components: *RealityStream* – a framework to recreate the remote robot and its environment in VR in real-time; and an intuitive control interface with natural interactions and affordance feedback for intent visualisation.

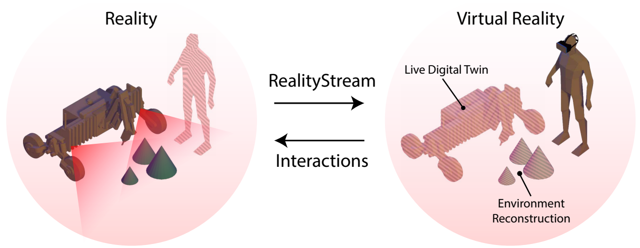


Figure 3: The Interaction Architecture - the user is virtually *in situ* with the remote robot and its environment

3.1 *RealityStream*: Towards Real-time Remote Scene Reconstruction with Point Cloud Streaming to VR for Robotics Applications

Similar to other remote field robots, HALO is equipped with color and depth cameras that can create a color point cloud of their environment. However, point clouds are not efficient for data transmission, and reconstructing and rendering them are computationally intensive. This could lead to a slower refresh rate, which is not suitable for VR. To solve these challenges, *RealityStream* was developed as a framework to produce a dynamic digital twin of a remote robot and its environment in VR. The framework was designed to be compatibile with ROS, and optimised to meet the high performance and fidelity requirements of VR.

Live Digital Twin of the Robot

The robot's digital twin is created in the Unified Robotics Description Format (URDF), with kinematic, dynamic, and visual models extracted from its CAD. The URDF was then imported into Unity using the URDFImporter from Unity Robotics Hub[1]. In Unity, the digital twin is used for visualisation and user interaction. Hence, its simulated physics properties like gravity and collisions are disabled, and its joint positions are updated by subscribing to the joint state messages from the real robot. The digital twin's orientation, most importantly the inclination on the rock surface, is also updated from the robots IMU data. Two RGB-D cameras on the robot are extrinsically calibrated, and their two 3D point clouds with overlapping views are reconstructed and rendered in VR in real-time, as detailed below.

GPU-Accelerated Point Cloud Reconstruction and Rendering

Each RGB-D camera equipped on HALO generates at 30 Hz an 8-bit RGB colour image and 16-bit gray-scale depth image, which is aligned to the colour image. The 16-bit depth image is first split into two 8-bit grey-scale images, containing the upper and lower bits. The three images are then compressed with JPEG, with the colour images at 80% quality and the depth images at 100% quality to retain as much detail as possible.

In Unity, after JPEG decompression, each coloured point in the point cloud can be constructed from data of three pixels from the same coordinate in the three images. As each group of three pixels can be processed independently of all other groups, a GPU's parallelisation capability can be utilised to extensively optimise this process. A *Compute Shader*, written in HLSL, was designed to process a block of 8×8 threads, or 64 threads in each thread group, with each thread processing a three-pixel group in a chunk of the colour and depth images simultaneously. This parallelism ensures that for images of size $N \times M$, the *Compute Shader* would attempt to run in $\frac{N \times M}{64}$ thread groups, making the process vastly faster compared to serial computation. In each thread, the following computations are executed.

Firstly, the original 16-bit depth value Z_i is restored from the lower and upper 8-bit depth pixels D_L and D_U respectively, and subsequently converted to metres:

$$Z_i = \left(D_L \times 256 + D_U \times 256^2\right) \times 0.001 \tag{1}$$

Then, the position of the 3D point P_i in the camera coordinate system is calculated using a pinhole camera model:

$$(X_i, Y_i) = \frac{(v_i, u_i) - (c_x^d, c_y^d)}{(f_x^d, f_y^d)} \times Z_i \tag{2}$$

where (c_x^d, c_y^d) is the depth image's principal point, $[f_x^d, f_y^d]$ are the depth camera's focal lengths, (u_i, v_i) is the coordinate of the corresponding depth pixel. By representing the point P_i in the camera frame by its homogeneous coordinates $\mathbf{p}_{\text{cam}} = [X_i, Y_i, Z_i, 1]$, it can be transformed to the world frame using the extrinsically

---

[1] `https://github.com/Unity-Technologies/Unity-Robotics-Hub`

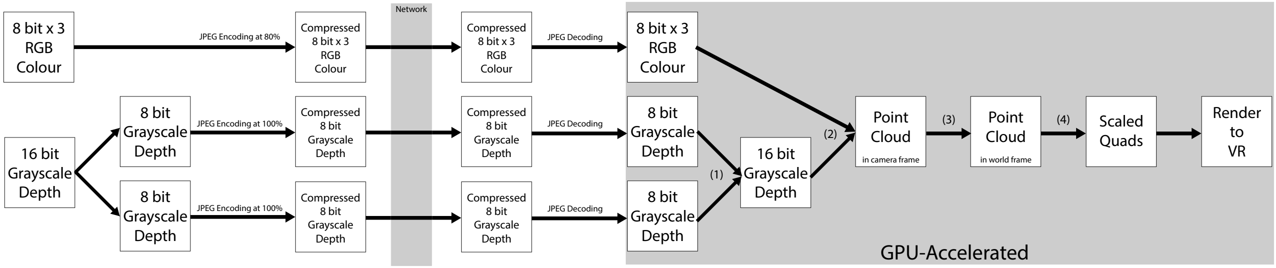


Figure 4: RealityStream: GPU-Accelerated point cloud streaming to VR.

calibrated depth camera's transformation matrix $\mathbf{T}_{\text{origin}}$ with

$$\mathbf{p}_{\text{world}} = \mathbf{T}_{\text{origin}} \cdot \mathbf{p}_{\text{cam}} \tag{3}$$

The colour information of each 3D point is then extracted from the same coordinate (u_i, v_i) of the colour pixel in the RGB image.

The computed world positions of all points, and their corresponding colours, are finally passed to a *Geometry Shader* via an internal GPU buffer. For each point $\mathbf{p}_{\text{world}}(x, y, z)$, four vertices are created to form the four corners of a square (quad):

$$\begin{aligned} p_1 &= (x - \frac{s \times d}{2}, y - \frac{s \times d}{2}, z) \\ p_2 &= (x + \frac{s \times d}{2}, y - \frac{s \times d}{2}, z) \\ p_3 &= (x - \frac{s \times d}{2}, y + \frac{s \times d}{2}, z) \\ p_4 &= (x + \frac{s \times d}{2}, y + \frac{s \times d}{2}, z) \end{aligned} \tag{4}$$

where d is the distance of the point to depth camera's origin, and s is a scale factor. The quad consists of two triangles ((p_1, p_2, p_3) and (p_2, p_3, p_4)) with normals perpendicular to the screen plane. The use of the distance d is to render points that are further away with larger quads to fill in gaps, and points that are closer with smaller quads to preserve details.

Preliminary Evaluation

Two user studies were previously conducted using variations of the point cloud reconstruction pipeline to investigate the effects of sensor modalities and various visualisation settings on user performance in VR.

The first user study investigates the effect of variances in real-time VR-based sensor data visualisation on user performance in an assembly task [18]. In this study, the point cloud reconstruction steps were similar, however it was executed serially on CPU. Therefore, the fastest rate that a complete point cloud could be visualised from an RGB-D camera was 5 Hz. The results indicated that the colour image stream was the most preferred modality over fully-coloured and mono-coloured variations of the non-optimised point cloud rendering.

The second user study investigated the effect of annotations on participant performance when executing a remote robot teleoperation task in VR [19]. In this study, the HALO robot was used in a controlled indoors setting, with two overlaying point clouds generated using the aforementioned optimised GPU-accelerated method. The study showed that the non-annotated coloured point clouds alone were sufficient for completing the manipulation task. The results also indicated that simultaneously visualising two modalities, point cloud and 2D image streams, is better at assisting the user.

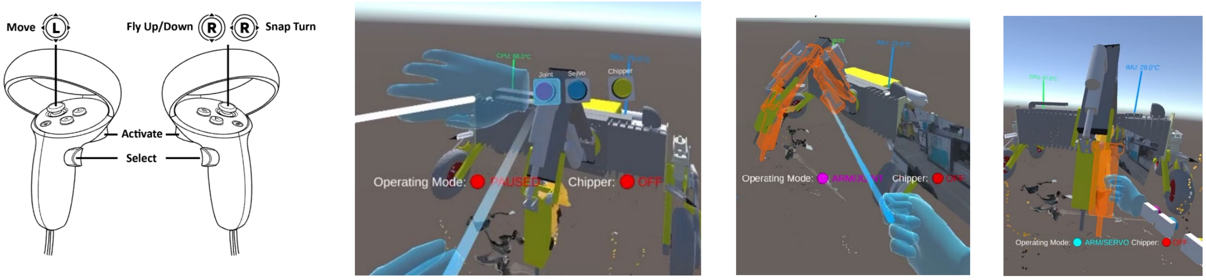
3.2 VR Control Interface with Natural Interactions and Affordance Feedback for Intention Visualisation

Users immersed in VR can exist virtually *in situ* with a 1:1 scale live digital twin of the remote robot and its environment with two overlapping point clouds, enabled by *RealityStream.* Users view the robot digital twin from an exocentric perspective, and they can control their position and orientation with natural physical movements, or with virtual movements by "flying" using the joysticks on the controllers.

Interactable virtual objects, that are contextual and exist in the virtual world-space, are placed in the scene and act as the main points of interaction. These include the links of the robot itself for individual joint jogging; the end-effector for seroving; and push-buttons attached to the user's wrists for mode switching or triggering an emergency stop.

Users can use the VR controllers (displayed as virtual "hands") to "*select*" and "*activate*" interactable objects that are in the scene, with the mapping shown in Fig. 5. The interface also includes "laser beams" that extend out of the virtual hands so that users can interact with objects at a distance.

The interactable objects, shown in Fig. 5, have the following feedback behaviours to signify their affordances: When a user hovers either their virtual hands or the laser beam on an interactable object, it will be high-



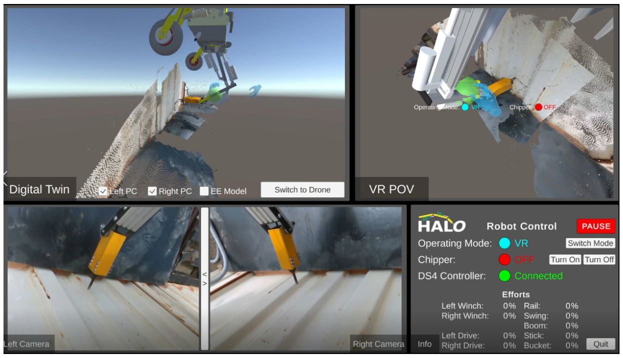
(a) (b) (c) (d)

Figure 5: (a) VR controllers' *Select*, *Activate* and movement controls mapping. Interactable objects with their affordance feedback: (b) virtual push-buttons for mode switching; (c) serial links of the robot in joint control mode; and (d) the chipper end-effector in servoing mode

lighted blue, and when the user *selects* the object, it will be highlighted orange. When interacting with the robot, the *select* state acts as a prior visualisation of the user's intention. This allows the user to verify their intention before executing the action.

The *Select-View-Activate* sequence is common in the interaction design of other VR applications. Placement of the *Select* button allows the user's natural grabbing motion to be registered, and placement of the *Activate* button acts as the trigger to confirm the intended action.

3.3 2D Control Interface with Gamepad



Operating Mode: VR
Chipper: OFF
DS4 Controller: Connected

Efforts
Left Winch: 0%
Right Winch: 0%
Left Drive: 0%
Right Drive: 0%
Rail: 0%
Swing: 0%
Boom: 0%
Stick: 0%
Bucket: 0%

Figure 6: 2D Interface of the HALO system

In addition to the VR control interface, a 2D interface, shown in Fig. 6, was developed. The live digital twin view provides operators with a quick overview of the robot's current state and its surroundings. During operation, if the rendered point clouds provide insufficient information, operators can rely on the 2D live camera feeds. To control the robot, a gamepad is used as a simple input device. In driving and scaling modes, analog joysticks are mapped to steering; and in arm control modes, they mimic the controls of an excavator arm. This interface and control scheme is useful for quick actions, or those that do not require 3D manipulation with the arm, such as driving on the bench from the deployed location to the edge of the rock wall.

4 FIELD TESTING

Field testing was conducted throughout the duration of the project to ensure that each system component functioned as intended. In particular, the aspects of the system that were evaluated included the manipulator's ability to remove rocks, maneuverability of the base on varying surfaces, the VR interaction intuitiveness, and the system as a whole. Testing was performed numerous times in the university's car park off the back of a truck and at a real-world open-cut sandstone quarry for outdoor, rough-terrain testing, shown in Fig. 7. The final full system test was conducted at a mine site, the system's intended environment.

4.1 Testing Arm Maneuverability and Rock Chipping

Early tests focused on the dexterity of the manipulator, as the air-chipper used for the end-effector works most effectively when strategically placed on the rock surface. During these tests an operator moved the end-effector around a object ensuring that it was able to be reached from all its faces.

The capability of air chipper was tested on objects of varying hardness and sizes. Using a constant pressure of 100 psi, the air chipper mounted on the arm could be maneuvered to break objects constructed of terracotta, concrete and sandstone. Additionally, it was observed that when the manipulator is in contact with the ground, it is capable of lifting its own weight. By pushing the end-effector into the ground, the robot pivots about the two rear wheels, lifting the front two off the ground.

These tests were conducted using gamepad teleoperation within line-of-sight. It was noticed that users found controlling individual joints of the robot with this method to place the chipper and specific position and

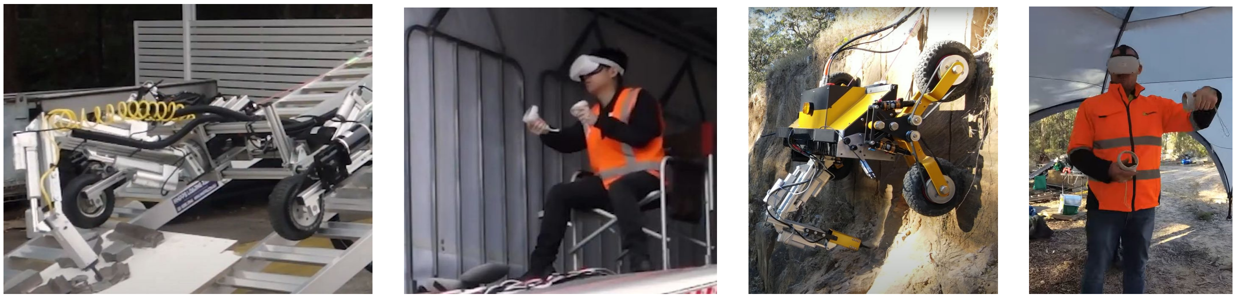

(a) (b) (c) (d)

Figure 7: Field testing: a) In the car park by deploying HALO off a truck, (b) Operator is seated in the back of the truck, (c) Real-world trial at a sandstone quarry, (d) Operator is free to walk around within the command tent.

orientations challenging. The VR control interface was later developed for a more intuitive method of interacting with the robot.

4.2 Testing Different Rock Scaling Locomotion Modes

In parallel to testing the capability of the manipulator, the platform's locomotion was verified to ensure its traversal capability on different rough surfaces and the ability to support the manipulator during rock scaling. Consequently, the tests involved traversing over flat ground, sloped ramps of varying degrees then scaling vertical walls. For each of these inclines the robot was evaluated on differing surfaces, including concrete, asphalt, gravel and sandstone.

The dynamics between the platform's two locomotion systems, the driving wheels and the two winches, were investigated. On an inclined surface, the winches are responsible for securing the robot, with the driving wheels being the main drivers for motion. In scenarios with limited traction, contact or stability, such as when the robot scales a near vertical wall, the winches are the main drivers. Therefore its reliability were carefully examined to ensure that the robot never becomes a falling-object hazard. Tests of the winching system involved attaching the winches to a crane that lifted the robot into the air. The winches would raise and lower the robot vertically ensuring that it could lift the robots unsupported weight. In addition to maneuvering the platform, these tests included having an operator simultaneously manipulate the arm to conduct rock-scaling and control the mobile platform to ensure stability.

Although the performed as expected indoors, extensive outdoors tests caused the motor drivers for the wheels and winches to overheat. One strategy explored was to employ an active cooling solution. The existing air line to the mobile manipulator was utilised. Additionally to powering the air chipper, the compressed air flowed through a vortex tube producing cool air for the enclosed space. However, this caused condensation to build up in the enclosure presenting differing challenges. It was evaluated that the alternative challenges presented outweighed the benefits and a passive cooling solution would be favourable for the current iteration of the system.

4.3 Field Testing at a Sandstone Quarry

Two field tests were conducted at a sandstone quarry such that the environment would simulate the intended deployment setting. For this type of deployment, the robot and its accessories are required to be self-sufficient. As a result, these field tests included running the mobile manipulator, air compressor and computing equipment off a generator.

During this testing an experienced rope access technician and a member of the project team operated the mobile robot. For two different runs, the users drove the robot from a starting location on flat ground to the top of the face of the sandstone wall. From there, the winches were manually secured to a post, and the robot was subsequently lowered down a near vertical sandstone wall by a user via line of sight. Once an ideal location was determined, the operators began conducting the rock-scaling process in VR from a safe location away from the robot and sandstone wall.

The technician provided positive feedback regarding the VR interface, particularly with the ability to visualise the sensor data as point clouds and as two 2D images. For fast arm movements, there was a preference to operate the robot with the gamepad as it provided more immediate control and required less setup time than the VR control station. Furthermore, the field trials provided valuable insight when considering the acquisition and visualisation of sensor data. It was found that point clouds from RealSense cameras do not yield optimal results when used outdoors due to interference from the sun's IR noise. Additionally, when scanning a rock wall, there may not be enough distinctive features in the point cloud to distinguish between different rock objects. Consequently, allowing users to switch between data visualisations is essential. A combination of 3D point clouds and 2D image streams proved to be integral for obtaining the desired results. Given this evaluation, it is recognised that feedback from the intended end-user is vital for system improvement.

4.4 Field Testing at a Mine Site

A final field test was conducted at a mine site to collect data on the robot's performance in a real-world setting operated by mining employees across various operational divisions. The trial spanned two days; on the first day HALO was unloaded, drove on the bench to the desired location, anchored to the pit wall, and lowered over the crest line to perform scaling operations. The system remained anchored overnight, and on the second day, additional scaling operations were performed.

The total operational time was 6 hours and 6 minutes, where the longest continuous operational duration was 1 hour and 45 minutes. During this time, all the robot's functionalities were tested: driving on the open-pit mine's rocky surfaces, winching down the pit wall, arm manipulation, and operating the air chipper to perform rock scaling. Initially, the robot was operated visually when it was within line-of-sight. However, once it went over the edge, control was shifted to either the VR interface or the 2D interface, with the operator relying entirely on sensor data feedback.

Analysis of the Select-View-Activate interaction design

During the two day site trial, 5 users engaged with the VR interface to control the robot, collectively spending 42 minutes and 30 seconds in VR (average of 8 minutes and 30 seconds per user). For all users, this was their first time using the interface and interacting with the robot. Among the participants, 3 out of 5 users indicated that this was their first time using a VR headset, while the remaining 2 out of 5 users reported limited prior experience with VR. The following data of each user shows the number of times *Select* and *Activate* were pressed, and the number of times they were pressed in a specific sequence. The *Select-View-Activate* sequence was specified to be the "correct" way to control the robot, where after selection, the user can visualise the future motion of the robot before executing it. Based upon this information and the data shown in Table 4.4, the interaction style of each user can be analysed. For a sequence of actions to be recognised in this table, a minimum of 1.0 second must be recorded between the select and activate presses. After the initial instructions were given, users operated the robot without further guidance, this resulted in a varying number of attempts for each user.

User	1	2	3	4	5
Select	15	120	21	82	28
Activate	16	121	18	84	33
Activate-then-Select	1	12	0	42	2
Select-then-Activate	10	103	18	38	27
Select-View-Activate	2	55	16	2	26
Select-View-Activate (%)	20	53.4	88.9	5.3	96.3

Table 1: Data for each user showing the number of times *Select* and *Activate* were pressed, and the number of times they were pressed in a specific sequence

- User 1 opted for more immediate action execution, ignoring the intention visualisation
- User 2 opted for a mixed of immediate action execution and intention visualisation before action execution
- User 3 and User 5 mostly used the intention visualisation before action execution
- User 4 used a more immediate action execution, however was successful less than 50% of the time, as the *Activate* state is only valid if it follows the *Select* state. With almost an equal number of "Activate-then-Select" and "Select-then-Activate" sequences, it is posited that User 4 pressed the buttons randomly until it worked.

5 DISCUSSION

The field testing results verify a scaled prototype of the HALO system, the first of its kind, capable of sensing and remotely controlling a robot performing rock scaling operations via a digital twin in VR. By assisting workers to do the difficult, hazardous and laborious rock scaling task, the system is able to reduce the risk of injury, and reduce workers' exposure to heights, crush hazards, fine dust particles, and dangerous tools. Thus, there is a significant health, safety and economic impact. The system can be supervised and interacted with remotely by an operator who is kept outside of the danger zone, at a safe distance away from the rock scaling operations.

Overall, the productivity of the robotic rock scaling operations was found to be lower than the continuous operation of one worker. The main physical limitation of the current robot design is the fixed size of the chipping tool mounted on the low degree of freedom single manipulator. The platform is suitable for the real-world quarry and mine environments where it was tested. However, it needs to be ruggedised before it can experience prolonged exposure to high temperatures in direct sunlight and high-pressure water cleaning to remove dust and debris after use.

The system faces several sensing challenges which must be investigated. Due to the size and complexity of the environment, complete maps are difficult to attain either *a priori* or through robot exploration and mapping. Therefore, a more holistic visualisation is predicted to help an operator make more informed decisions. Additionally, techniques could be used to improve the fidelity of the map of the environment based on processes that use inference and prior knowledge.

Another concern is the feasibility of scaling up this proof-of-concept remote mobile manipulator with a live VR digital twin to practical equipment. For example, whether joints should be driven hydraulically or pneumatically remains a question for future consideration. Furthermore, the interaction architecture can remain the same or be iterated from the current version, as the virtual environment and VR interactions can be easily changed. Furthermore, it is clear that the HALO platform is a proof of concept for remotely controlled field robots with live VR digital twins. This design approach can be expanded and used in broader applications within the mining and other high-risk industries.

Another research question that arises is the level of automation that a VR-based digital twin platform should provide to aid the human operator. A highly automated VR-based digital twin platform can potentially reduce the cognitive load subjected on to the human operator, enabling them to focus on higher-level decision-making tasks. However, it is possible that such a platform can impact on the level of situational awareness of the operators, if they rely heavily on the ability of the platform.

From Section 4.4, the data highlights challenges with the current interaction design, which point towards improvements that can be made. Only two out of five users utilised the visualisation before execution. To help elucidate this design, the interaction states can be better signified with additional feedback, such as audio or animations. In the case where a user attempted random button inputs, a potential solution is to replace virtual hands with virtual controllers to show button mappings with virtual pop-ups, sacrificing a small level of immersiveness for clarity. Additionally, an introductory VR tutorial is expected to aid the development of a baseline skill level for new operators.

In terms of RealityStream, further development is required to mitigate its current limitations. One of the significant issues is JPEG decompression. This choke point causes the frame rate to drop significantly with 720p point clouds, which is undesirable since lower resolution point clouds may not capture all the necessary information required for the operator to control the HALO platform effectively.

A concern associated with the integration of robots into society is their potential impact on the workforce as their presence displaces human workers. HALO has been designed to be a platform that eliminates the risk of injury that the operator is subjected to when performing rock-scaling, rather than an autonomous system replacing the rope access technician. In its current iteration the platform leverages the knowledge and control of the operator to manoeuvre the robot base and end-effector. However, during field tests, operators demonstrated a positive sentiment for the use of a robot as a tool, rather than a robot taking away their job.

6 CONCLUSIONS

We present a first-of-its-kind rock scaling mobile manipulator that can be remote controlled by means of intuitive interaction with its live digital twin in VR. The HALO robotic system, with its high-level autonomous remote teleoperation, can be used for rock scaling a range of mining environments. The immersive and intuitive nature of VR facilitates human interaction, for visualising and examining the proposed system's sensor data, as well as providing an interface for control and decisions. The project has manifested in a manufactured proof of concept prototype, with control schemes, and on-site proof-of-concept testing in real-world outdoor environments. Future work will require improving HALO in three key ways: increasing the machine's physical robustness; increasing operator awareness by providing a more holistic vision of the surrounding environment; and enabling high level forms of intuitive interaction to future the assist the operator to remotely perform rock scaling.

ACKNOWLEDGMENT

We would like to thank Raegan Rumbold and Clint Ricetti for their support. This research is supported by UTS:RI, The Commonwealth of Australia's Department of Industry, Science, Energy and Resources (Innovative Manufacturing CRC Ltd) and Ausdrill (Perenti). D.T. Le, S. Sutjipto, T.L. Vu and D.D.K. Nguyen are supported by Australian Government Research Training Program Scholarships.

The authors would like to acknowledge the support received through the ARC Industrial Transformation Training Centre (ITTC) for Collaborative Robotics in Advanced Manufacturing under grant IC200100001.

References

[1] G. Macqueen, E. Salas, and B. Hutchison, "Application of radar monitoring at Savage River Mine, Tasmania," *Slope 2013: 2013 International Symposium on Slope Stability in Open Pit Mining and Civil Engineering, 2013 25-27 September, Brisbane*, pp. 1011–1020, 9 2013.

[2] R. Fanti, G. Gigli, L. Lombardi, D. Tapete, and P. Canuti, "Terrestrial laser scanning for rockfall stability analysis in the cultural heritage site of Pitigliano (Italy)," *Landslides*, vol. 10, no. 4, pp. 409–420, 8 2013.

[3] T. Dean, M. Grant, and M. Pavlova, "The efficient acquisition of high-resolution 3D seismic surveys for shallow open-cut mining," *First Break*, vol. 39, no. 8, pp. 43–49, 8 2021.

[4] D. D. Khoa Le, G. Hu, D. Liu, R. Khonasty, L. Zhao, S. Huang, P. Shrestha, and R. Belperio, "The QUENDA-BOT: Autonomous Robot for Screw-Fixing Installation in Timber Building Construction," *IEEE International Conference on Automation Science and Engineering*, vol. 2023-August, 2023.

[5] P. Ward, P. Manamperi, P. Brooks, P. Mann, W. Kaluarachchi, L. Matkovic, G. Paul, C.-H. Yang, P. Quin, D. Pagano, D. Liu, K. Waldron, and G. Dissanayake, "Climbing Robot for Steel Bridge Inspection: Design Challenges," in *Proceedings for the Austroads Publications Online*. ARRB Group, 10 2014, pp. 1–13.

[6] M. G. Carmichael, S. Aldini, R. Khonasty, A. Tran, C. Reeks, D. Liu, K. J. Waldron, and G. Dissanayake, "The ANBOT: An Intelligent Robotic Co-worker for Industrial Abrasive Blasting," in *IEEE International Conference on Intelligent Robots and Systems*, 2019, pp. 8026–8033.

[7] G. Paul, S. Webb, D. Liu, and G. Dissanayake, "A robotic system for steel bridge maintenance: Field testing," in *Proceedings of the 2010 Australasian Conference on Robotics and Automation, ACRA 2010*, 2010.

[8] A. Parness, N. Abcouwer, C. Fuller, N. Wiltsie, J. Nash, and B. Kennedy, "LEMUR 3: A limbed climbing robot for extreme terrain mobility in space," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 5467–5473, 7 2017.

[9] M. Anderson and C. Johnson, "Scaling the heights: the background and development of a novel remote highwall scaling machine for use at the Savage River Mine," *2020 International Symposium on Slope Stability in Open Pit Mining and Civil Engineering*, pp. 1203–1212, 5 2020.

[10] E. Rosen, D. Whitney, E. Phillips, G. Chien, J. Tompkin, G. Konidaris, and S. Tellex, "Communicating Robot Arm Motion Intent Through Mixed Reality Head-Mounted Displays," *Springer Proceedings in Advanced Robotics*, vol. 10, pp. 301–316, 2020.

[11] C. Wang and A. Belardinelli, "Investigating explainable human-robot interaction with augmented reality," in *5th International Workshop on Virtual, Augmented, and Mixed-Reality for Human-Robot Interactions (VAM-HRI 2022)*, 2022.

[12] C. P. Quintero, S. Li, M. K. Pan, W. P. Chan, H. Machiel Van der Loos, and E. Croft, "Robot Programming Through Augmented Trajectories in Augmented Reality," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 10 2018, pp. 1838–1844.

[13] J. D. Hernandez, S. Sobti, A. Sciola, M. Moll, and L. E. Kavraki, "Increasing robot autonomy via motion planning and an augmented reality interface," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1017–1023, 2020.

[14] 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)*. IEEE, 10 2018, pp. 1–9.

[15] K. Li, R. Bacher, W. Leemans, and F. Steinicke, "Towards Robust Exocentric Mobile Robot Tele-Operation in Mixed Reality," in *5th International Workshop on Virtual, Augmented, and Mixed-Reality for Human-Robot Interactions (VAM-HRI 2022)*, 7 2022.

[16] N. T. T. Le, H. Y. Zhu, and H. T. Chen, "Remote visual line-of-sight: A Remote Platform for the Visualisation and Control of an Indoor Drone using Virtual Reality," *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*, 12 2021.

[17] D. T. Le, S. Sutjipto, Y. Lai, and G. Paul, "Intuitive Virtual Reality based Control of a Real-world Mobile Manipulator," in *16th IEEE International Conference on Control, Automation, Robotics and Vision, ICARCV 2020*. Institute of Electrical and Electronics Engineers Inc., 12 2020, pp. 767–772.

[18] T. L. Vu, D. T. Le, D. D. K. Nguyen, S. Sutjipto, and G. Paul, "Investigating the effect of sensor data visualization variances in virtual reality," *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*, 12 2021.

[19] T. L. Vu, D. D. K. Nguyen, S. Sutjipto, D. T. Le, and G. Paul, "Investigation of Annotation-assisted User Performance in Virtual Reality-based Remote Robot Control," in *2022 Australasian Conference on Robotics and Automation (ACRA 2022)*, 2022.