TeLeMan: Teleoperation for Legged Robot Loco-Manipulation using Wearable IMU-based Motion Capture

Chengxu Zhou[1], Christopher Peers[1], Yuhui Wan[1], Robert Richardson[1], and Dimitrios Kanoulas[2]

Abstract— **Human life is invaluable. When dangerous or life-threatening tasks need to be completed, robotic platforms could be ideal in replacing human operators. Such a task that we focus on in this work is the Explosive Ordnance Disposal. Robot telepresence has the potential to provide safety solutions, given that mobile robots have shown robust capabilities when operating in several environments. However, autonomy may be challenging and risky at this stage, compared to human operation. Teleoperation could be a compromise between full robot autonomy and human presence. In this paper, we present a relatively cheap solution for telepresence and robot teleoperation, to assist with Explosive Ordnance Disposal, using a legged manipulator (i.e., a legged quadruped robot, embedded with a manipulator and RGB-D sensing). We propose a novel system integration for the non-trivial problem of quadruped manipulator whole-body control. Our system is based on a wearable IMU-based motion capture system that is used for teleoperation and a VR headset for visual telepresence. We experimentally validate our method in real-world, for loco-manipulation tasks that require whole-body robot control and visual telepresence.**

I. INTRODUCTION

In the area of defence and security, there are often tasks that may be too risky for a human to attempt to complete. For instance, in the UK, there are more than $2,500$ yearly operations for Explosive Ordnance Disposal (EOD). The human expertise involved in completing such tasks is at an impressive level. Therefore, due to human safety and well-being purposes, a robotic alternative is sought by defence departments around the world.

Robotic platforms, that are able to navigate in challenging environments and manipulate objects, would be ideal for operating in the aforementioned demanding tasks. Legged quadrupedal robots have become cheaper, more robust in unstructured environments (e.g., rough terrains and stairs), with high computational capabilities (e.g., GPUs), and a wide range of on-board sensors (e.g., IMUs, cameras, force/torque sensors). Very recently, researchers started embedding robotic arms on the top of quadruped robots to

Fig. 1. Teleoperative legged manipulator system: a teleoperator with VR headset and wearable IMU-based motion capture, controlling locomotion and manipulation.

enhance them with manipulation capabilities. For instance, such robots include Spot (by Boston Dynamics), ANYmal (by ANYbotics), HyQ and CENTAURO (by IIT). Such a platform–a legged quadruped manipulator–is ideal in completing loco-manipulation tasks.

While a big part of the robotics community has focused on autonomous scene understanding and action performance, achieving human-level skill quality is not trivial. Robust high-level decision-making for several complex tasks, is still under heavy research. Thus, controlling the high-level actions of robots via teleoperation and telexistence is a good compromise on the path towards full autonomy. Even though robot teleoperation and telexistence is not a new topic of study [1], it has attracted the interest of the research community again, given that only recently, the computational units, sensors, and wireless connection capabilities have been developed to such a level that might allow for real-time robust performance, up to an extent.

In this paper, we firstly introduce hardware adjustments to assemble a low-cost, but yet competent, quadrupedal manipulator (Fig. 1). It includes a quadruped robot (Unitree Laikago), a modified 5 Degrees-of-Freedom (DoF) robotic arm (ViperX 300), and an RGB-D camera (Intel Realsense). Secondly, we introduce a unified system for whole-body loco-manipulation control, using a wearable IMU-based motion capture system for teleoperation (Perception Neuron Motion Capture), via 5 GHz Wi-Fi 6, for real-time communication. Based on this system, a teleoperator can control simultaneously the robot locomotion and manipulation in an intuitive manner, while receiving visual feedback from the depth camera, via a Virtual Reality (VR) headset (HTC Vive Pro). The goal is to allow easy real-time robot teleoperation

and telexistence for risky defence and security tasks, such as the EOD. Finally, real-world experiments for EOD tasks were used to evaluate the full system.

Overall, the main contributions of this paper are as follows:

1) introducing a novel hardware development for an inexpensive quadrupedal manipulator, by also reconfiguring the robotic arm into a lightweight one, saving payload effort for the robot,

2) introducing a new unified teleoperation system for loco-manipulation, based on a wearable IMU-based motion capture system; a system that has not been used in the research quadrupedal legged robotics community before, and

3) demonstrating via real-world experiments the superiority of the telexistence and teleoperation capability to complete Explosive Ordnance Disposal (EOD) tasks, compared to traditional gamepad methods.

The remainder of the paper is structured as follows. In Section II, we present the related work in robot teleoperation and telexistence, followed by the hardware system description in Section III. In Section IV, we present the whole-body teleoperation system, while in Section V we present the experimental validation of the system on the real robot. Finally, we discuss conclusions and future work in Section VI.

II. RELATED WORK

Joystick control was the standard for legged/wheeled robot teleoperation [2], [3] until recently. Even though joysticks are ideal when a teleoperator needs to control the position or velocity of a single part of a robot, e.g., its body on flat ground, it becomes challenging and non-trivial when: 1) more complex movements are required, e.g., the 6DoF motion of a manipulator's end-effector (which can be resolved with motion capture hand controllers [4]), or 2) multiple parts of a robot need to be controlled simultaneously, e.g., the position/velocity of its body and a different position/velocity of its end-effector. We have extensively noticed these disadvantages on legged manipulators. For instance, two examples include our work on the IIT-Centauro animaloid robot [5] and the work on ETH-ANYmal legged manipulation robot [6]. For this reason, the research community has focused on different devices and sensors for teleoperating complex legged manipulators.

A common alternative method to teleoperate complex multi-DoF robots is to create a wearable exoskeleton twin for the teleoperator. Even though this is an expensive alternative, it has been only successful with manipulation tasks [7]. While usually navigation is controlled firstly with a joystick, disengaging the two tasks makes the system too complex and challenging to be used. On the other side, locomotion has been handled in the past by using treadmills [8], which is a rather expensive solution to the problem.

Several researchers are using motion capture systems to control robots, mainly again for manipulation tasks only. For instance, in [9], motion capture for the position of the arm, along with a hand/fingers haptic device, was used to teleoperate a quadrupedal manipulator. Although, again locomotion was not demonstrated via these devices, thus the robot navigation was controlled before the manipulation task. Moreover, motion capture systems have the disadvantage that they need to be installed in a room or space, while ideally a human teleoperator should be able to move anywhere to operate the robot.

Following a new trend in robotics, originating from the animation industry, our work is focused on using an IMU-based wearable suit to teleoperate the robot. Previous works show that IMU-embedded wearable motion capture devices can be used reliably enough to generate high and low-level control for robots with a high degree of accuracy, furthering its potential in complex manipulation tasks. One example is presented in [10], where the high-level walking motion is used to control a bipedal robot, with the teleoperator using foot positioning to determine the robot actions. Another example, presented in [11], demonstrates the use of a gesture-based control scheme, via an IMU embedded device on the hand and arm, to control a wheeled robot. Likewise, wearable IMU-based motion capture systems are also possible to reliably control 7DoF robotic arms [12]. Wearable IMU-based motion capture devices have also been demonstrated to be accurate and robust in domestic scenarios, e.g., for a dementia-care robot [13]. In [14], the TALOS bipedal robot's whole-body is controlled as a fixed-base system (i.e., without locomotion), using an Xsens 3D motion tracking suit. In the system realisation, the hands and head tracking is used with an absolute Cartesian approach to achieve robotic teleoperation. The teleoperator's motion capture suit data is completely overlaid on the humanoid robot. An additional human model is used as a method to format the suit data to fit the robot. Our work, is the first we are aware of, that uses an IMU-suit to control simultaneously the navigation and manipulation of a quadrupedal legged manipulator.

Regarding telexistence, the majority of the current methods include VR headsets [15], and haptic devices [16], to give enough visual, audio, or haptic feedback to the teleoperator, in order to control the robot efficiently and accurately. We use similar vision-based VR-based methods in this work to provide the teleoperator with telexistence capabilities, leaving haptic feedback for future work.

III. HARDWARE SYSTEM DESCRIPTION

In this section, we provide a description of all hardware that was used for robot teleoperation and telexistence, as well as its modifications. In Fig. 2, an overall structure of the system is visualised, including: 1) an integrated robotic platform, consisting of a quadruped robot, a robotic arm with a gripper, and an RGB-D sensor, and 2) a human teleoperation system, consisting of a VR headset and an IMU-based body motion capture system.

The robot hardware can be separated into two parts. The first part is the robot base, which consists of the Unitree Laikago quadruped robot, with 5 kg of payload. The robot base is powered by an on-board Intel NUC i5 computing
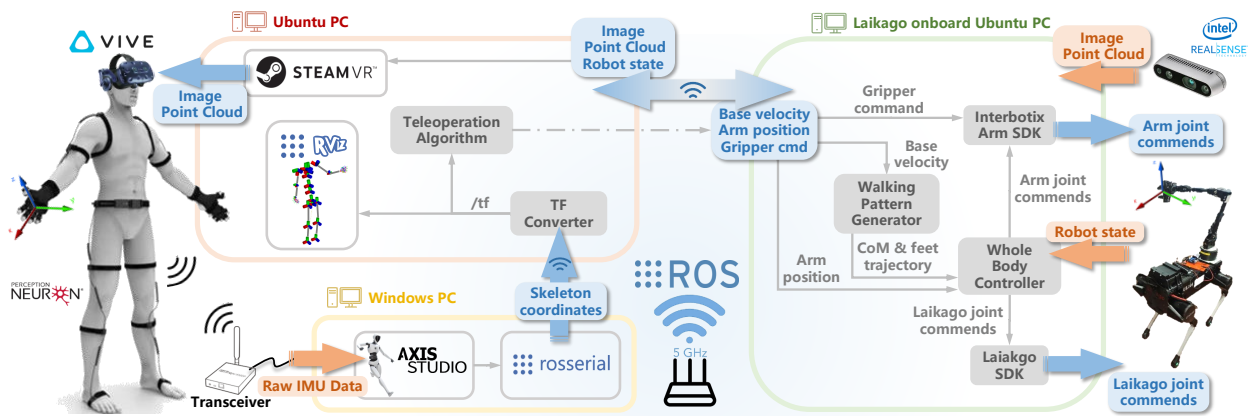
Fig. 2. System overview: the teleoperator (left) controlling the robot base and arm (right), via wearable IMU-based motion capture, while getting feedback from the robot's RGB-D camera via the VR headset.

system. On top of the quadruped robot base is the second part, which is a ViperX 300 Robot Arm and a visual Intel RealSense RGB-D camera. The robotic arm was redesigned in order to be mounted on top of the quadruped robot. Also, the RGB-D camera is able to rotate with the arm around the yaw-axis. A full simulated model of the integrated robot parts was generated within the Robotic Operating System (ROS) and Gazebo, to test all its capabilities before experimenting on the real robot.

For teleoperation, Noitom's Perception Neuron Studio, a wearable motion capture system based on inertia sensing, was used for precisely estimation of human body postures. The system consists of 16 IMU sensors, which are attached to each link of a human body, and a pair of Studio Motion Capture Gloves for the finger estimation. This system provides accurate estimation of the human skeleton's positions and orientations in world coordinates, which include 19 body segments and 40 hand segments, at 100 Hz with a minimum resolution of 0.02 degrees. Calibration is needed for every new user which only takes about 30 s.

For the VR system, a HTC VIVE Pro headset was used, which visualises the robot model, the human teleoperator as a transformation tree, as well as the point clouds and RGB images from the robot's Intel RealSense camera. Last, to be able to operate in a real-time sensitive environment, a 5 GHz Wi-Fi router was used for the communication between all devices.

A. Hardware Redesign

One of the biggest constraints of using a quadruped robot base platform is its payload limit. A heavy payload not only increases the workload on the motors, but also increases the height of the center of mass. Therefore, it has a considerable impact on the robot's stability and maneuverability. To minimise this impact, various methods were utilised to reduce extra loads placed on the robot. The heaviest object in the design was the 5 DoF ViperX 300 robotic arm. In order to reduce the weight of the arm, while maintaining its capability, several parts have been redesigned. This involves, firstly, replacing the aluminum box-section links of the arm with carbon fibre rods. Secondly, redesigning the structure of

the shoulder joints to maintain the robotic arm's capability with the new carbon fibre rods. Given the aforementioned changes, we managed to reduce the overall weight of the arm from 4.1 kg to 2.5 kg, a 39% weight reduction. With the newly designed arm, the Laikago quadruped robot base can maintain strong stability and maneuverability, while still having an allowance for extra manipulation workload on the arm. This robot arm was mounted onto the Laikago's rear parallel carbon fiber rods, using a 3D-printed easily interchangeable slide-in mounting system. A diagram of the redesigned robot arm and the 3D-printed Laikago mounting system is shown fully rendered in Fig. 3.

An RGB-D Intel RealSense camera was used as visual feedback. This allows the operator to achieve telexistence via the RGB image and coloured depth-based point cloud representation of the world. In particular, the RGB images provided are essential for both locomotion and manipulation teleoperation tasks, however, the addition of the point clouds enables manipulation tasks that require a more accurate measure of distance. Point clouds allow the operator to visually confirm and pinpoint the three-dimensional locations of specific details in the environment and could provide great assistance while dealing with complex tasks that require a heightened sense of target and surrounding environments, such as opening boxes or cutting wires. Various mounting positions for the RGB-D camera were tested during the experimental validation to locate the most suitable position and angle for teleoperation. We mounted the RGB-D camera onto a part of the robot arm that allows the camera to maintain its perspective, regardless of the robot arm's yaw angle. We tested three mounting positions: 1) on the base of the robot arm, in the center pitched slightly upwards, 2) on the wrist joint of the robot arm behind the end-effector, and 3) to the side of the arm at a set height rotated towards the arm slightly in the yaw axis. A visual representation of these locations along with their respective 3D-printed mounting parts is shown in Fig. 3.

During real-world experiments, we noticed that, when a manipulation task was aimed (e.g., when attempting to open a box cover or pull out a wire), it was difficult for the operator to perceive whether the end-effector was properly positioned, due to the arm itself was obstructing the vision

(a) Exploded view   (b) Arm assembly   (c) Camera views from three locations
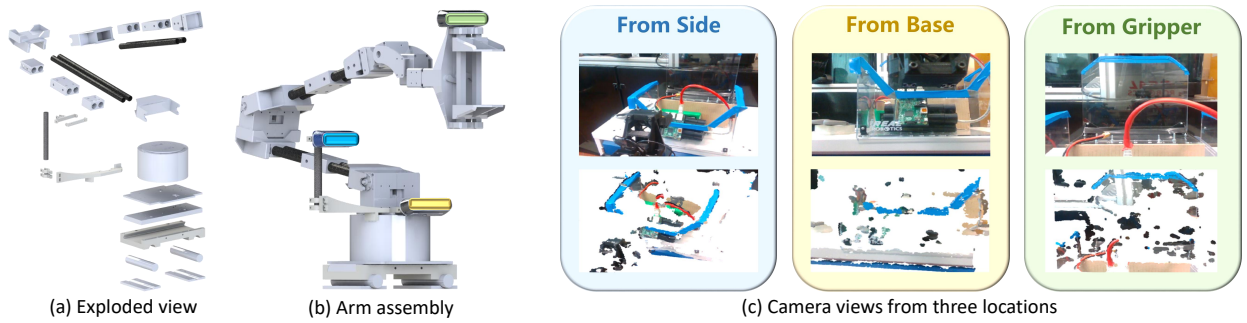
Fig. 3. (a) Exploded view of all redesigned components in the robotic arm, (b) full assembly of the robotic arm with three possible camera mounting positions coloured and (c) the corresponding RGB and depth camera image from each mounting location on the robotic arm.

of the top/bottom end-effector gripper respectively, making it extremely difficult to determine the vertical location of end-effector. This issue was however solved with the side camera positioning. This camera location gave a "third-person perspective" to the operator. In this way, the operator can easily determine both the horizontal and vertical location of the end-effector. This camera position also allowed for all key objects that were manipulated to be in the frame along with the end-effector, without any obscuring problems. The quality of the point cloud however differs in each position. It was observed that, while viewing from the base and gripper locations, the quality of the point cloud began to decrease due to the camera being too close to the object, which causes some points to not be readable.

IV. SOFTWARE SYSTEM DESCRIPTION

Having introduced the hardware design that was developed for teleoperation and telexistence above, this section describes the software architecture of the system. The system can be split into three parts: robot motion generation, IMU-sensing for teleoperation with strategies, and the generation of the VR environment that enables telexistence.

A. Gait Generation and Whole-Body Control

Based on the virtual leg concept [17], by considering a pair of legs that move simultaneously as a single one, several quadrupedal robotic gaits, e.g., pacing, trotting, and bounding, can be realised by adopting the gait pattern generators developed for bipedal robots. Therefore, we have extended our previous walking algorithm [18] to generate the quadrupedal trotting gait by considering the diagonal legs as two virtual legs. As shown in Fig. 2, the generated trajectories are sent to a whole-body controller for coordinated loco-manipulation of the developed legged manipulator.

The whole-body controller that considers the full dynamics of the developed legged manipulator is formulated as a quadratic programming problem:

$$\min_{\boldsymbol{\mathcal{X}}} \quad \sum_{i=1}^{n} \omega_i \left\| \boldsymbol{A}_i \boldsymbol{\mathcal{X}} - \boldsymbol{b}_i \right\|^2, \qquad (1)$$

where the sum of n task's costs is minimised to obtain the optimal value of the target variable, $\boldsymbol{\mathcal{X}} = [\ddot{\boldsymbol{q}}, \boldsymbol{\lambda}]^T$, which consists of the generalised acceleration $\ddot{\boldsymbol{q}}$ and contact wrenches $\boldsymbol{\lambda}$. The ith task is defined by an objective matrix

and vector, $\boldsymbol{A}_i$ and $\boldsymbol{b}_i$, and a weight ω_i that determines the soft priorities between tasks. Please refer to [19] for detailed formulations of individual tasks.

During the optimisation process, the following constraints are considered: i) floating base dynamics $\boldsymbol{M}_f \ddot{\boldsymbol{q}} + \boldsymbol{h}_f = \boldsymbol{J}_f^T \boldsymbol{\lambda}$, ii) joint torque limits $\boldsymbol{M}_a \ddot{\boldsymbol{q}} + \boldsymbol{h}_a - \boldsymbol{J}_a^T \boldsymbol{\lambda} = \boldsymbol{\tau} \in [\boldsymbol{\tau}_{\min}, \boldsymbol{\tau}_{\max}]$, iii) non-slip contact constraints $\boldsymbol{J} \ddot{\boldsymbol{q}} + \dot{\boldsymbol{J}} \dot{\boldsymbol{q}} = \boldsymbol{0}$, and iv) contact force constraints to keep each contact force within a linearised friction cone. $\boldsymbol{M}$ is the inertia matrix, $\boldsymbol{h}$ is the sum of Coriolis, centrifugal and gravitational terms, $\boldsymbol{J}$ is the contact Jacobian matrix corresponding to the contact wrenches $\boldsymbol{\lambda}$ at all contact points, the subscript f represents the top six rows of a matrix for the floating base and a corresponds to the actuated DoFs, $\boldsymbol{q} = [\boldsymbol{q}_f, \boldsymbol{q}_a]^T$ are the generalised coordinates which consists of the 6 DoF floating-base coordinates $\boldsymbol{q}_f$, and the actuated joints $\boldsymbol{q}_a$, $\boldsymbol{\tau}$ is a vector of joint torques.

The optimised result $\boldsymbol{\mathcal{X}}$ is then used to calculate the joint torque, joint position, and joint velocity references for lower-level control of the developed legged manipulator [19].

B. Teleoperation Strategies

The human body motion is captured by a wearable motion capture system (Noitom Perception Neuron), as appears in Fig. 2-left. This provides stable and accurate human body segments pose estimations. The accompanied software (Noitom Axis Studio) can read full skeletal data, including fine finger movements. This data is then broadcasted to ROS through the *rosserial* protocol. Since the human tele-operator and the robot are kinematically dissimilar, directly connecting them at a joint level is impractical. Thus, we have developed four strategies to intuitively teleoperate the legged manipulator in locomotion and manipulator tasks. As shown in Fig. 4, the teleoperator uses hand closure as a trigger to send commands to the robot:

- *Gripper trigger* is activated when the left hand closes above the waist. The gripper on the manipulator will remain closed until this trigger is released.
- *Walking trigger* is activated when the left hand closes below the waist. The teleoperator can then send base velocity references to the walking pattern generator by taking a step forward/backward, left/right, counter-clockwise/clockwise, or with a combination of directions. The step size or the rotation angle defines the
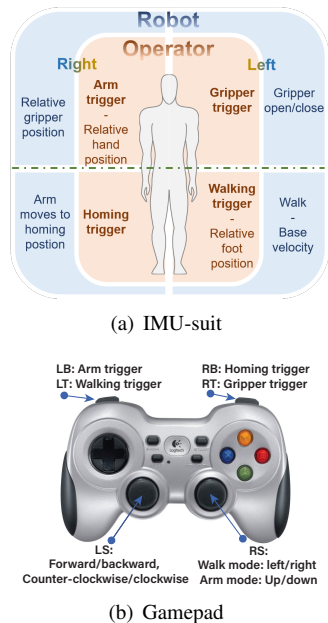
(a) IMU-suit

(b) Gamepad

Fig. 4. Details of mapping from interfaces to trigger and argument strategies: (a) IMU-suit, (b) gamepad.

magnitude of velocity that the robot will move at. The robot stops when this trigger is released.

- *Arm trigger* is activated when the right hand closes above the waist. Then, the teleoperator's right arm movement will generate relative Cartesian arm position references to the whole body controller. The arm stops moving when this trigger is released.
- *Homing trigger* is activated when the right hand closes below the waist. The robot joints will move to a predefined configuration.

The left and right-hand closure triggers can be combined, for instance, using left hand to enable *walking trigger* and right hand for *arm trigger* at the same time, thus performing simultaneous locomotion and manipulation control of the legged manipulator.

Relative scaled pose relationship is used for connecting the teleoperator and robot movements, where at time t, the *arm* and *walking triggers* are modelled as:

$$\begin{cases} \boldsymbol{a}_{rd}^t = \boldsymbol{a}_{rd}^0 + \boldsymbol{\mu}(\boldsymbol{a}_e^t - \boldsymbol{a}_e^0) & \textit{arm trigger} \\ \dot{\boldsymbol{w}}_{rd}^t = \dot{\boldsymbol{w}}_{rd}^0 + \boldsymbol{\mu}(\boldsymbol{w}_e^t - \boldsymbol{w}_e^0) & \textit{walking trigger} \end{cases} \quad (2)$$

where $\boldsymbol{a} = [x, z, \theta_{\text{roll}}, \theta_{\text{pitch}}, \theta_{\text{yaw}}]$, $\boldsymbol{w} = [x, y, \theta_{\text{yaw}}]$ define the displacements in sagittal (x), lateral (y) and vertical (z) directions and rotation about the front-to-back axis (θ_{roll}), the side-to-side axis (θ_{pitch}) and the vertical axis (θ_{yaw}) for arm and base movement, respectively, subscripts e refers to the teleoperator, r to the robot, and d to a desired value, superscript 0 refers to the initial timing when the trigger is activated, $\boldsymbol{\mu}$ is used to scale the motions between the teleoperator and the robot. The VR glove's readings are used to detect the hand's closures for activating the triggers to teleoperate the robot.

This trigger-based relative motion strategy eliminates the long-term sensor drifting issue, especially for the yaw angle, which normally imposes negative impact on the controllers that rely on absolute yaw sensing.

C. VR-based Integration for Telexistence and Teleoperation

In many real-world applications, a direct line of sight between the robot and the teleoperator may not always be possible. Some missions even require the teleoperator to keep a safe distance on-scene, for instance, in confined space, HAZMAT, and EOD tasks. Therefore, a remote way to monitor the robot movement and surrounding environment is necessary for these situations. Furthermore, since the wearable motion capture system uses human body motion as a control input signal from the teleoperator, it will not be ergonomic to look at a fixed display while the teleoperator's body turns around. VR also introduces a full telexistence experience, which provides more information than traditional display and allows robust and efficient robot control. In the experiment, the teleoperator cannot perceive any aspect of the experiment except for the information displayed within the VR headset to maintain the teleoperative aspect of the experiment.

To generate the VR environment, the view of the robot's and teleoperator's state along with visual RGB-D feedback is generated and integrated into the teleoperator's view. First, the robot's RGB-D camera provides the teleoperator with a coloured 3D point-cloud and a 2D RGB image embedded above it (see Fig. 3-(c)). The RGB image generates a straightforward view of the surrounding environment, while the point cloud provided a more detailed structure. Notice that, the coloured images were compressed before communication to achieve higher transmission performance. The point clouds were cropped into the surrounding 2 m distance only, to give a clearer and fast view of the setting. Various methods of filtering the point cloud were tested, such as downsampling (e.g., voxelisation) or statistical outlier removal. Although, it is hard to balance precision with minimal representation, objects such as wires, still occupy very few points in the cloud, and thus they cannot be expressed if such filtering takes place.

Secondly, the robot's model and the teleoperator's IMU-based skeleton view were visualised in the VR headset. In this way, the teleoperators were able to see their own body structure, the robot, and the visual environment around it for a full telexistence experience. The data between the robot and the VR headset were transmitted via 5 GHz Wi-Fi to ensure bandwidth. Several experiments, e.g., VR-based manipulation, were also performed on a mobile manipulator at the early stage of development (UCL MPPL [20]).

V. EXPERIMENTAL RESULTS

A gamepad is another widely used control method for robots, and many quadruped manufacturers consider them as the default controller for commercial customers. In this research, we compare the performance difference between the wearable motion capture system and the gamepad. As a control, the teleoperation strategies are the same between the two controllers. They both have velocity control for
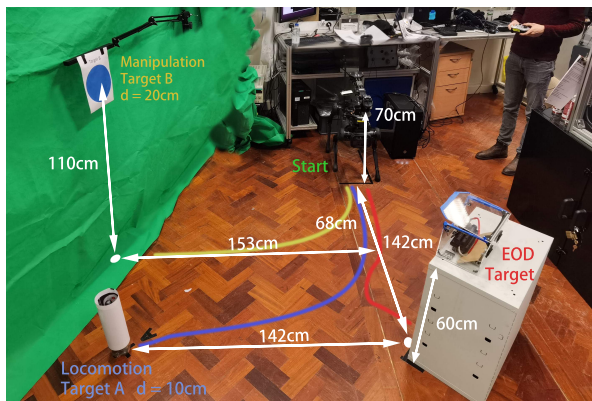
Fig. 5. Detailed map of the experiment area with the locations of each target.

locomotion and position control for manipulation. In detail, the trigger and bumper buttons on the gamepad are mapped to triggers in the teleoperation strategy and the joysticks are mapped to relative position and velocity. During the experiment, the user teleoperates a legged manipulator robot to complete four tasks with both types of controllers. The time took to finish each task is measured as their performance.

A. Experimental Design

The experiment has been divided into four tasks, as shown in Fig.5. Users firstly perform basic manoeuvre of robot teleoperation to have a standard view with both controllers, then they try to accomplish a locomotion task, a manipulation task, a combined loco-manipulation task and an EOD task. The order of the first two tasks is randomised, half of the users perform the locomotion task first, and half perform the manipulation task first. After that, they perform the EOD task again in the same conditions, without direct sight through VR.

1) The locomotion task requires the robot to walk from starting point to target A only.
2) The manipulation task starts with the robot standing below target B, and it requires moving the robot arm from the home pose to reach target B hanging in the air.
3) The locomotion and manipulation task requires the robot to start from the starting point and walk towards target B, then use its arm to reach target B in the air.
4) The last task is the EOD task, a simulated real-world application to disassemble a bomb inside a box. In detail, the robot needs to walk from the starting point to the front of the bomb box. Then, the robot opens the box and pulls out a red wire from the bomb, using its manipulator.

Failure conditions include the robot's movement deviating outside the experiment area, the robotic arm moving or breaking the box, and also any sort of hardware failure such as the control system, robot or robotic arm becoming non-functional.

B. User Profiles

In total, we have six volunteer users to perform all the above-mentioned four tasks. These users cover different ages,

TABLE I
AVERAGE TIME TO COMPLETE EACH TASK AND ITS STANDARD DEVIATION IN SECONDS

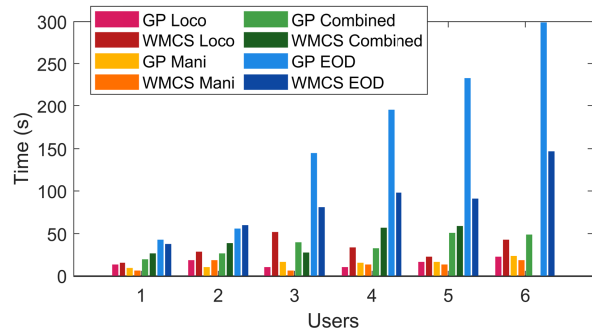|	Gamepad		IMU		IMU+VR
	Avg.	Stdev.	Avg.	Stdev.	Avg.
Locomotion	14.8	4.8	32.5	13.6	
Manipulation	15	5.4	12.3	5.4	
Loco+Mani	18.2+18.3	12.5	34.2+7.8	15.4	
EOD	161.8	100.6	85.8	37.1	118

Fig. 6. Movement time each user takes to complete tasks with gamepad (GP) and IMU-based wearable motion capture system (WMCS).

gender, and experience. Half of these users had experience with gamepads. Furthermore, two of them even have robot operating experience. On the other hand, three users have no engineering or gamepad experience. After basic instruction on controllers, they were directly assigned to tasks without additional practice or training.

C. Gamepad vs IMU Suit Performance Comparison

The time taken for all six users to complete each experiment stage is shown in Fig. 6. User 6 did not complete the combined test due to tactical difficulty, and all other users completed all planned tasks. Table I shows the average time of the six users using different controllers to complete each task. From the table, we see that the gamepad has higher performance in the locomotion task and that the wearable motion capture system has the advantage in complex tasks. Also, the user's performance with the wearable motion capture system improves as they practice throughout the experiment.

In the locomotion task, the wearable motion capture system took almost twice as long as the time of gamepad, to finish the task. While the time is similar for both controllers in the manipulation task, with only less than 20% difference. At this point, the gamepad proved to have the advantage of performing better the locomotion tasks.

For the combined locomotion and manipulation task, we see a similar total time between the gamepad and the wearable motion capture system. However, we can break it into two separate parts: the locomotion part and the manipulation part. In this way, we can see that the gamepad has an equal distribution between two parts (locomotion of 18.2 s and manipulation of 18.3 s), and the motion capture system experiences difficulty in the locomotion part, but has a significantly better performance in the manipulation part (locomotion of 34.2 s and the manipulation of 7.8 s). One reason for such a result is the difference in the Degrees
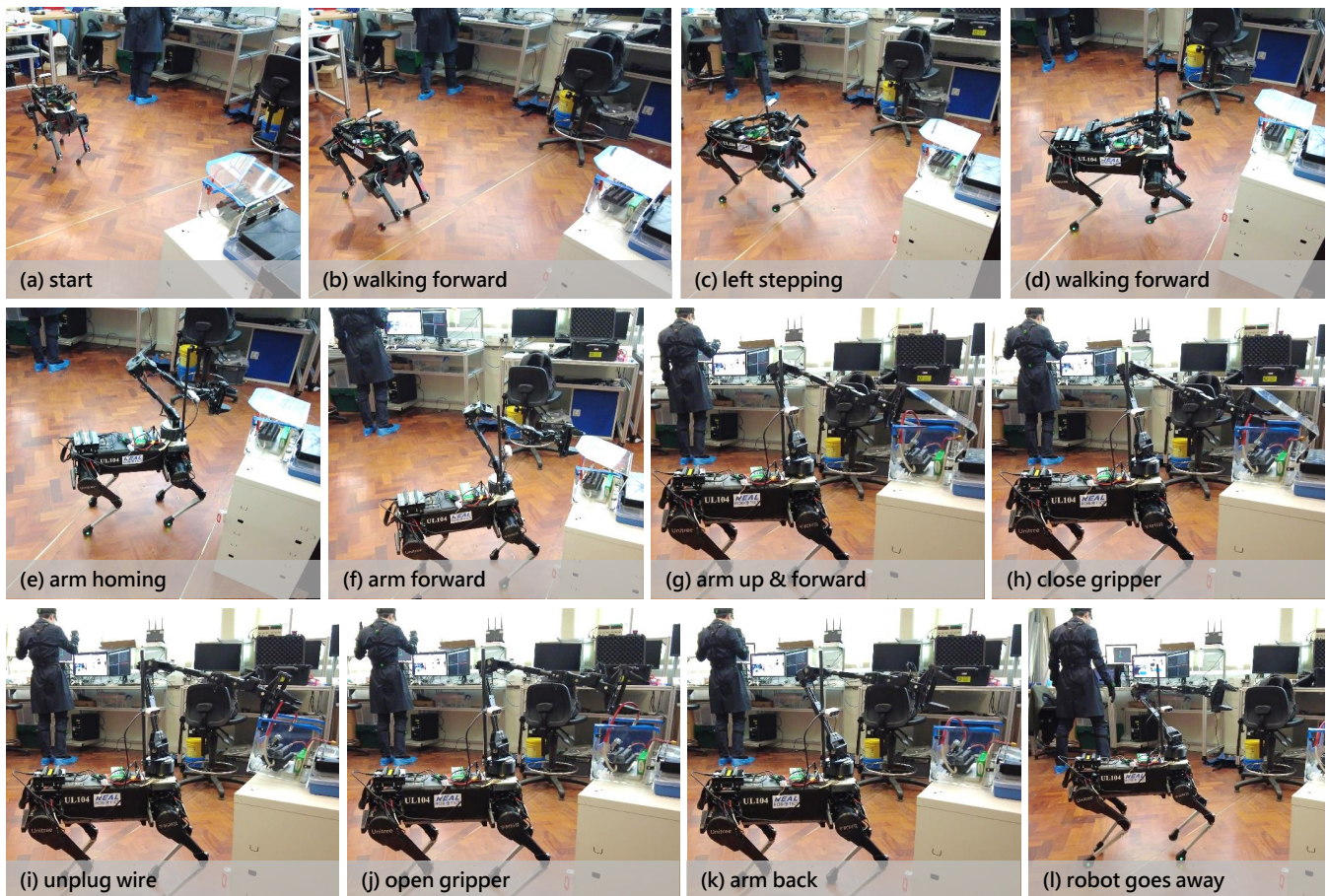
Fig. 7.   Chronological snapshots of the teleoperator's command with the corresponding teleoperated system's action throughout the live experiment.

of Freedom, where locomotion is a 2D movement, and manipulation is a 3D movement, which causes manipulation to be more complex than the locomotion movement in this setup. The traditional joysticks on a gamepad usually have 2DoF control, which makes them suitable for 2D movements. In contrast, the motion capture system captures human arm movement in 3D space, which can easily map to a manipulator's 3D movement. It is also more natural and user-friendly for human teleoperators to use their arm to control the robotic arm in teleoperation. This hardware characteristic makes the wearable motion capture system more suitable for real-world applications.

One of the example missions is the EOD task, where the robot needs to perform both locomotion and manipulation, however, the manipulation part in this task is more challenging. From Table I, we can see that the gamepad took almost twice the amount of time to complete the EOD mission in comparison to the wearable motion capture system on average. Some missions similar to this example can be time-sensitive, and timely response can prevent further property damage, personal injury, even fatality. With the advantage of high efficiency, the proposed system that uses wearable motion capture system can be helpful in many real-world applications for teleoperating legged manipulators.

As mentioned earlier, the experiment performs in the order of simpler tasks to more challenging tasks. With this order, users are getting more and more familiar with the system as they practice through the experiment. By examining Table I for each controller's performance in completing the tasks, we can see that the wearable motion capture system has the tendency to improve performance. In particular, the manipulation part in the combined task has an improvement of over 30% when compared to the earlier manipulation task. However, for the gamepad, there is no evidence of improvement with practice. Furthermore, the performance difference between the two controllers also grows with more practice in favour of the motion capture system. Using the wearable motion capture system, we have one user with only less than two hours further training, who was eventually able to finish the EOD task within 45 seconds, which almost reduces the time by half. Snapshots of one training trial are shown in Fig. 7. This further proves the practice trend for the motion capture system.

The manipulation tasks we aim at solving with this system requires precision at the level of centimeters. We have successfully experimented with EOD tasks that roughly needed a 5 cm movement accuracy, which demonstrated the effectiveness of the proposed IMU-based teleoperation.

D. VR Usability

During the experiment, all users could complete the EOD task with the wearable motion capture system and VR-based integration. From Table I, they finished the task with an average of 118 s, which is slower than direct sight control

with the wearable motion capture system but still faster than direct sight control with the gamepad. As the camera is located on the side of the arm to have a closer view of the manipulating object, both users think the camera angle is too narrow, especially during locomotion. However, none of the users had any discomfort during the tests. Most of the users believe the VR is a good choice in a human inaccessible environment.

VI. Conclusions

Quadrupedal legged manipulation is an exciting new area of research and development in robotics. In combination with teleoperation, it makes the ideal case of allowing telexistence to replace humans in dangerous or hard tasks. In this paper, we developed such a system and tested it for a typical EOD task. Telexistence is an area of current development in real-world robotics, but also a very challenging one. There are several challenges in providing a reliable teleoperation/VR system on a real robot.

Some of these challenges were regarding hardware and software integration. The main hardware issue was that legged robots of this scale have a low payload. In our case, the maximum payload of the quadruped robot is only 5 kg. Thus, only lightweight robot/sensory parts could be installed on its body. We resolved this problem by redesigning the manipulator that was embedded on the robot's back. The main challenge with the software integration was that all submodules installed on three computers (Fig. 2) had to be synchronised. Some devices were supported in Ubuntu (quadruped, manipulator, RGB-D sensor), while others only in Windows (VR and wearable IMU-based motion capture), and thus they must be bridged synchronously. Our system structure and communication over 5 GHz Wi-Fi handles this. The wearable IMU-base motion capture system requires calibration before being used, which takes only a few seconds for the teleoperator. Potential drifts due to slight miss-calibration were not preventing the teleoperator from driving the robot accurately, through the use of intuitive relative movement adjustments which eliminate the long term sensor drifting effects.

Throughout the experiment, the operators have only visual feedback provided either by direct sight or VR. Therefore, safe physical interactions between the robot and the environment heavily rely on the low-level joint torque controller. We are looking at introducing haptic feedback for the operator in future work for a better HRI experience.

There are several directions for advancing the proposed system. Firstly, we are looking into developing code that allows all the devices to work in the same operating system. In this way, the amount of computers that are required will be minimised, and synchronisation issues might decrease. Secondly, we are looking into semi-autonomous teleoperation, where one-to-one mapping between the teleoperator and the robot is avoided. Instead, the robot can use autonomy for some cognitive tasks, while the teleoperator configures the sequences. Last, we aim at using the system for more real-world tasks, such as inspection and monitoring in industrial settings.

In the experiment, we had users with gamepad experience and users without gamepad experience. Despite this difference, they are from the same user group: college students and researchers. This unicity contributes to the limitation of the result, as another group of users with different backgrounds may generate a different result.

References

[1] M. T. Mason, "Creation Myths: The Beginnings of Robotics Research," *IEEE RAM*, vol. 19, no. 2, pp. 72–77, 2012.
[2] J. Chestnutt, P. Michel, K. Nishiwaki, J. Kuffner, and S. Kagami, "An Intelligent Joystick for Biped Control," in *IEEE International Conference on Robotics and Automation*, 2006, pp. 860–865.
[3] T. Klamt *et al.*, "Supervised Autonomous Locomotion and Manipulation for Disaster Response with a Centaur-Like Robot," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2018, pp. 1–8.
[4] T. Zhang, Z. McCarthy, O. Jow, D. Lee, X. Chen, K. Goldberg, and P. Abbeel, "Deep Imitation Learning for Complex Manipulation Tasks from Virtual Reality Teleoperation," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2018, pp. 5628–5635.
[5] E.-J. Rolley-Parnell, D. Kanoulas *et al.*, "Bi-manual articulated robot teleoperation using an external RGB-D range sensor," in *ICARCV*, 2018, pp. 298–304.
[6] C. D. Bellicoso, K. Krämer, M. Stäuble, D. Sako, F. Jenelten, M. Bjelonic, and M. Hutter, "ALMA - articulated locomotion and manipulation for a torque-controllable robot," in *IEEE International Conference on Robotics and Automation*, 2019, pp. 8477–8483.
[7] T. Klamt *et al.*, "Remote Mobile Manipulation with the Centauro Robot: Full-body Telepresence and Autonomous Operator Assistance," *ArXiv*, vol. abs/1908.01617, 2020.
[8] M. Elobaid, Y. Hu, J. Babic, and D. Pucci, "Telexistence and Teleoperation for Walking Humanoid Robots," *CoRR*, vol. abs/1809.01578, 2018. [Online]. Available: http://arxiv.org/abs/1809.01578
[9] N. Deshpande, J. Ortiz *et al.*, "Next-generation collaborative robotic systems for industrial safety and health," *WIT Transactions on The Built Environment*, vol. 174, pp. 187–200, 2018.
[10] S.-K. Kim, S. Hong, and D. Kim, "A Walking Motion Imitation Framework of a Humanoid Robot by Human Walking Recognition from IMU Motion Data," in *IEEE-RAS IJHR*, 2009, pp. 343–348.
[11] S.-O. Shin, D. Kim, and Y.-H. Seo, "Controlling Mobile Robot Using IMU and EMG Sensor-Based Gesture Recognition," in *International Conference on BWCCA*, 2014, pp. 554–557.
[12] C. Yang, J. Chen, and F. Chen, "Neural learning enhanced teleoperation control of baxter robot using IMU based motion capture," in *ICAC*, 2016, pp. 389–394.
[13] H. Lv, G. Yang, H. Zhou, X. Huang, H. Yang, and Z. Pang, "Teleoperation of collaborative robot for remote dementia care in home environments," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 8, pp. 1–10, 2020.
[14] E. Dalin, I. Bergonzani, T. Anne, S. Ivaldi, and J.-B. Mouret, "Whole-body teleoperation of the Talos humanoid robot: Preliminary results," in *ICRA 2021 - 5th Workshop on Teleoperation of Dynamic Legged Robots in Real Scenarios*, 2021.
[15] M. Elobaid, Y. Hu, G. Romualdi, S. Dafarra, J. Babic, and D. Pucci, "Telexistence and teleoperation for walking humanoid robots," in *Intelligent Systems and Applications*, vol. 1038, 2019, pp. 8477–8483.
[16] S. Tachi, K. Mlnamlzawa, M. Furukawa, and C. L. Fernando, "Haptic Media Construction and Utilization of Human-Harmonized "Tangible" Information Environment," in *International Conference on Artificial Reality and Telexistence*, 2013, pp. 145–150.
[17] I. E. Sutherland and M. K. Ullner, "Footprints in the Asphalt," *International Journal of Robotics Research*, vol. 3, no. 2, pp. 29–36, 1984.
[18] C. Zhou, X. Wang, Z. Li, and N. Tsagarakis, "Overview of gait synthesis for the humanoid COMAN," *Journal of Bionic Engineering*, vol. 14, no. 1, pp. 15–25, 2017.
[19] Y. You, S. Xin, C. Zhou, and N. Tsagarakis, "Straight Leg Walking Strategy for Torque-controlled Humanoid Robots," in *IEEE International Conference on Robotics and Biomimetics*, 2016, pp. 2014–2019.
[20] J. Liu, P. Balatti *et al.*, "Garbage collection and sorting with a mobile manipulator using deep learning and whole-body control," in *IEEE-RAS International Conference on Humanoid Robots*, 2020.