IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 3 July 2024, accepted 17 July 2024, date of publication 22 July 2024, date of current version 26 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3432401

RESEARCH ARTICLE

Implementation of a Robotic Arm Control for EOD Applications Using an Immersive Multimodal Interface

WALKER AGUILAR[1], LIZARDO PARI[1], YURI SILVA[2], ERASMO SULLA ESPINOZA[1], LUIS F. CANAZA CCARI[1], RONALD PEÑA[3], AND NICOLÁS O. MEDINA[1]

[1]Electronic Engineering Professional School, Universidad Nacional de San Agustín de Arequipa, Arequipa 04000, Peru
[2]Mechanical Engineering Professional School, Universidad Nacional de San Agustín de Arequipa, Arequipa 04000, Peru
[3]PNP Technology Department, National Police of Peru, Lima 15036, Peru

Corresponding author: Nicolás O. Medina (nmedinac@unsa.edu.pe)

This work was supported by PROCIENCIA through Universidad Nacional de San Agustín de Arequipa under Contract PE501079565-2022.

ABSTRACT The advancement of multimodal interfaces aims to provide an intuitive user interface to improve the performance of various tasks. Teleoperated robotics in Explosive Ordnance Disposal (EOD) tasks require the operator to perform complex maneuvering tasks to control a robotic arm. Problems such as loss of depth perception, degradation of visual perception, system delay, and a high mental workload in the control of the robotic elements make these tasks require extensive training periods, extensive knowledge of the robot operation and demand a great effort from the operator to perform a task efficiently and avoid catastrophic situations in handling explosive packages. To solve this, a multimodal interface is proposed based on the creation of a virtual operating environment, where the operator uses a combination of three interfaces: a visual interface through a Virtual Reality Head Mounted Display (VRHMD), a natural user interface (NUI) and a predictive display based interface. The proposed system is evaluated with the participation of thirteen agents with experience in explosive ordnance disposal tasks; the results obtained are divided into objective and subjective results through time measurements of task completion, success rate, usability through System Usability Scale questionnaire (SUS), mental workload through NASA Task Load Index questionnaire (NASA TLX) in Pick and Place tasks, which constitute the master task type in EOD robotics tasks. The proposed multimodal interface is proven to have a considerably higher efficiency than the conventional keyboard and monitor-based control interface, specifically achieving an improvement in task completion times of 67%. 14%, an improvement in task completion rate of 11.54%, a decrease in overall mental workload of 65.18%, and an improvement in usability of 198.12%.

INDEX TERMS Multimodal interface, immersive environment, EOD tasks, mental workload, teleoperated robotics.

I. INTRODUCTION

Over the last decades, there have been a significant number of cases of explosive attacks with catastrophic results. In response, explosives deactivation units have been implemented in different countries [1], [2], [3], [4], [5], [6]. These specialized agents are responsible for the deactivation of different explosive devices through procedures that require direct handling of the device. However, despite the agents' experience and standardized procedures to avoid risks, fatalities have been recorded in explosive ordnance disposal (EOD) activities [7].

Teleoperated robotics has provided a solution to this problem by developing different robotic systems capable of manipulating objects at distances where the operator is not exposed to catastrophic events [2], [8], [9]. These EOD robots commonly have a mobile chassis and an attached robotic arm that is responsible for interaction with suspicious

The associate editor coordinating the review of this manuscript and approving it for publication was Carmelo Militello.

packages [2], [3], [4], [8], [9], [10]. In [2], the Karo robot is implemented. This robot has a robotic manipulator with seven degrees of freedom, and its control station uses an XBOX controller to send control commands. It has an armed and disarmed mode to prevent unintentional commands from being transmitted to the robot. In [11] an EOD robot is designed with an arm of seven degrees of freedom and independent control of the mobile chassis and each degree of freedom of the robotic arm. In [12], an EOD robot with a five degrees of freedom robotic arm is proposed, and a 3D projection of the workspace is made with the capability to rotate the scene for the operator to have the proper angle of view. All these systems have robotic arms capable of manipulating the object with considerable dexterity. However, they require a complex control system that involves a high mental workload for the operator and long training periods to achieve proper control of each robotic prototype.

On the other hand, in multiple investigations [1], [4], [13], [14], [18], [19], [20], different interfaces of robotic arm control are used. Conventional interfaces for controlling a robotic arm in this type of application use either a keypad or a joystick to independently control each degree of freedom or a particular combination of these. However, this type of control requires operators to have prior experience and a thorough understanding of the operation of the robotic arm. To overcome this problem, many studies have focused on the use of natural user interfaces (NUI). These interfaces focus on making the control of the robotic arm a simple and intuitive task. In [21] the Leap Motion (LM) sensor is used to control a robotic arm in surgical applications. This sensor uses two infrared cameras to constantly detect the user's hand. In [4], this sensor is used in EOD robotic applications, and a Kalman filter is applied to estimate the position of the user's hand more accurately, although physiological noises from the operator's hand are still present. In [1], optimal signal processing is performed to suppress these physiological tremors and estimate the intention of the operator's hand movements. However, there are issues such as the sensitivity of this type of sensor and the mental and physical burden of not having a clear reference for the working space of the LM sensor. In [13], a previous work has been done where the Falcon controller is tested as NUI interface in the control of robotic arms in EOD applications and a comparison is made with the performance of the LM sensor, concluding that the Falcon controller achieves optimal performance in the control of robotic arms in EOD applications.

Another significant disadvantage present in conventional control interfaces that negatively affects operator performance is visual and depth degradation [20], which occurs because these types of interfaces use monocular cameras and computer screens for visual feedback to the operator. An extensive compilation of research studies regarding the effect of interfaces on the performance of specific tasks is presented in [20], highlighting the following important points: Camera Viewpoint, Degraded Perception, Degraded Video Image, and Time Delay. These points are discussed in depth in the Related Work section.

Signal delay is another factor that affects the performance of the operators. In [20] it is found that when the signal delay is around one second, the operator stops trying continuous control and uses a "move and wait" strategy. Because these tasks are as delicate and critical as explosive handling, an adequate interface is required to solve these problems and improve the operators' performance in controlling the EOD robot, the delay perceived by the operator is inherent to all teleoperated systems. However, a possible solution is provided in [15], where a predictive display is used. With this system, the operator can constantly monitor the desired position of the robotic manipulator and its current position. This allows the operator to maintain continuous control of the robotic arm despite system delays.

In [17], a thorough study of multimodal interfaces and the effect of their application in teleoperated robotics is conducted. This study concludes that the application of a virtual reality head mounted display (VRHMD) increases the overall performance by 40% by providing the operator with depth perception through stereoscopic feedback, the application of haptic feedback increases the performance by 10%, and auditory feedback by 5%. In [18], an analysis of different control interfaces of a robotic arm in EOD applications is performed, concluding that a correct choice of the interfaces that compose the multimodal control interface leads to a decrease in the mental workload, which improves the performance of the users in the activity. In [19], the CERNTauro robot, a modular robot with a multimodal control interface, is presented. These features allow the robot to be efficiently teleoperated in hazardous environments by both experienced and novice operators.

In this research we present the implementation and validation of the control of a robotic arm for EOD applications using an immersive and intuitive multimodal interface, this interface is composed of the following components:

- An interface based on VRHMD that allows the operator to have optimal visual feedback.
- NUI control interface based on the Novint Falcon controller, which allows the operator to intuitively and efficiently operate the robotic arm.
- A predictive display that allows the operator to monitor the actual and target position of the robotic arm simultaneously.

To validate this new multimodal interface and evaluate its performance in EOD tasks, the robotic arm of the JVC02 robot, a response EOD robot, is adapted by using magnetic rotary encoders and a communication with RS485 protocol, this implementation is detailed in the methodology section below. In addition, the system is tested by both experienced and novice agents in teleoperated robotics of the Explosives Disposal Unit of Arequipa (UDEX-AQP), in order to assess the performance of the proposed interface in the adaptation times for optimal control of the robotic system, the tests are based on typical tasks in the deactivation of explosives.

This document is divided as follows: In the second section, we provide a review of the major problems that commonly occur during teleoperated robotics tasks. In the third section, we propose our approach to these problems and the solutions provided by other state-of-the-art research. In the fourth section, we state the hypotheses to be tested in the methodology. In the fifth section, we detail the proposed system implementation and testing methodology. In the sixth section, we detail the parameters measured in the experiment. In the seventh section, we detail the results obtained. In the eighth section, we provide an extensive discussion of these results. In the ninth section we show the future work and limitations of this research. Finally, in the tenth section, we present the most relevant conclusions of the study.

II. RELATED WORK

In this section, we explore in detail the problems previously mentioned, considering other papers that also explore these problems in the context of teleoperated robotics.

A. CAMERA VIEWPOINT

In teleoperated EOD robotics, the most typical method of providing visual feedback of the target is by using cameras and an interface that displays live video on a screen, usually located at the control station. An EOD robot with this functionality has been implemented to support explosive ordnance disposal tasks performed by UDEX-AQP agents. The EOD robot JVC02 has three cameras located in the front, turret, and robotic gripper; the video captured by these cameras is transmitted to the control station using a Wi-Fi antenna. In [2], the Karo robot uses a similar interface for visual feedback, as shown in Figure 1. The use of multiple cameras allows the user to have more information about the robot's environment; however, the placement of these cameras in these robots is usually fixed or dependent on the position of another link such as a telescopic camera. This does not allow the operator to change his point of view to more comfortable locations to acquire a better amount of information about the robot's working environment. In [20] it is mentioned that the differences between the point of view of the user and the camera can produce motion sickness to the operator. In general, visual perception from a fixed camera point negatively affects operator performance. Figure 1 shows the conventional visual interface of the JVC02 EOD robot on the left and the Karo robot interface on the right.

B. DEGRADED DEPTH PERCEPTION AND VIDEOIMAGE

Another disadvantage related to the use of cameras and a conventional visual feedback interface is the degradation of image quality. Image quality in teleoperated robotic applications can be affected by two main factors: transmission signal strength and video quality provided by the cameras [22]. In a teleoperated explosive ordnance disposal environment with different obstacles and a considerable

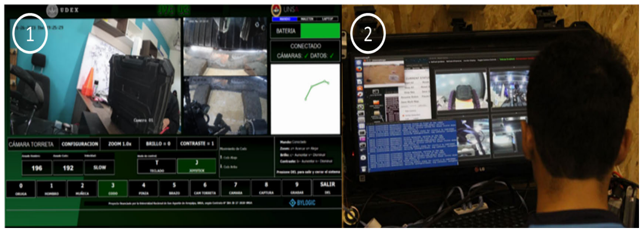


FIGURE 1. Conventional visual interfaces through monocular cameras. (1) JVC02 EOD Robot visual interface. (2) Karo Robot visual interface.

distance between the control station and the robot, the signal strength is compromised. In addition, the quality of the cameras and the video resolution provided by them have a significant effect on the operator's performance when using the robot. In [20] it is determined that a drop in the Frames per second (FPS) below 3 fps makes it impossible to perform a teleoperated robotic operation.

On the other hand, an important problem is the loss or degradation of depth perception when using screens for visual feedback. This is because the information captured by a single camera only transmits 2D information, which does not allow for a disparity effect; therefore, there is no information about the distance from the camera to the object [25]. The lack of depth perception prevents the operator from correctly controlling the robotic arm in the working environment and hinders the correct manipulation of the target [17]. This causes the operator to rely on his experience and skill to correctly operate the robot, using visual cues such as shadows, estimated dimensions, and observed resolution of the objects in the camera.

C. TIME DELAY

The time delay or latency refers to the delay between the instant when the control action occurs and the visible response in a system. Commonly, this value is measured in milliseconds [20]. Studies in virtual environments have shown that people can perceive latencies as low as 10-20 ms [30]. Other studies indicate that when the latency is around 1 s, the operator changes strategy from "operate continuously" to "give a command and wait for a response" [20]. In general, delay has a negative impact on the performance of teleoperated tasks because the operator expects a real time response, where every movement he makes with his environment has an immediate response. However, the delay is inherent to teleoperated systems because of the operation principle itself. Actuators, telecommunications, and electronics add a delay to the system that can not be completely suppressed.

D. ROBOTIC ARM CONTROL

Control of the conventional robotic arm in EOD robots is performed using a keyboard or joystick, which usually controls two different parts: the chassis or mobile platform and the robotic arm or manipulator. Chassis control, although it involves problems such as obstacle avoidance, is usually less complex than teleoperated maneuvering of the robotic arm [4]; however, this control involves a higher level of difficulty because of the multiple joints it presents. The complexity of the control of the robotic arm is aggravated as it has a larger number of degrees of freedom in the robotic arm [23], requiring that the controller not only commands independent movements of each degree of freedom but also simultaneously considers combined movements such as extension and retraction of several links.

III. OUR APPROACH

To overcome the problems detailed in the previous section. A multimodal interface composed of the following elements is introduced:

- A VRHMD-based Interface
- A NUI controller-based interface
- A Predictive display based interface

Figure 2 summarizes these interfaces and their interaction with the user and the EOD robotic arm; in this scheme, the complete system is divided into three parts: The Human-Robot Interaction (HIR) space, the virtual environment, and the EOD environment. Within the HIR environment, the operator uses a NUI controller to send commands to the virtual environment. To visualize these commands, he uses virtual reality glasses (VRHMD). In the virtual environment, the operator can move freely and interact with the virtual robotic arm through the NUI controller, as well as visualize the predictive display based interface. The information from the virtual environment is transmitted to the EOD environment, where the robotic arm finally performs the commands given by the operator.

The operation and implementation of these interfaces to overcome the problems detailed in the Related Work section are explained below:

A. STEREOSCOPIC STIMULATION

Visual perception is crucial in the performance of operators in teleoperated robotic tasks, being the most dominant sense, contributing approximately 70% of human perception [13], [17]. The two main methods used by the operator to obtain visual information in teleoperated operations are monocular monitors and VRHMDs [17]. VRHMDs are capable of providing stereoscopic vision that is lost in conventional interfaces using monocular monitors. Using a stereoscopic vision system, the operator can have a clear depth perception and an enhanced view of the robot's working environment. By improving depth perception, the operator can manipulate objects in obstructed spaces with significantly higher dexterity [24], [25], [28], [29]. Another important point of using VRHMDs is that the virtual environments are represented with a high image quality [25], [26] so that the operator has a clear visual perception of the main components of the working environment. These advantages are used in the proposed system by generating a virtual environment, where there is a representation of the robotic arm, the suspect package to manipulate, and the ground level. This environment is immersively perceived by the operator through the VRHMD, where he can move through the virtual three-dimensional space, have clear depth perception and can see their surroundings with clarity.

B. PREDICTIVE DISPLAY

As mentioned in the Related work section, a critical problem that affects the performance of the operators is the delay present in this type of system [20], [31]. This is produced by different factors such as: Slow speed of the robotic arm motors, delay in telecommunications due to signal latency, high computational load of the software responsible for the display of this video, and processing time of the electronic circuits. Although this delay cannot be completely avoided, it is possible to provide information to the user about the existing delay so that he can make the best decision for the control of the robotic arm. The predictive display system proposed in [15] presents a graphical representation on the screen of the operator station, where two robotic manipulators can be observed, one representing the position desired by the operator and the other representing the real position of the robotic arm. Thus, the operator receives continuous information regarding the command sent to the robotic arm and its response. This system is proposed in our multimodal interface through the representation in the virtual environment of a green ''guide'' robotic arm that the operator controls directly through the NUI interface. This position indicated by the operator is communicated to the arm motors through the RS485 communication protocol. As a second element of the predictive display is implemented a ''follower'' robotic arm is represented in red that reproduces the movements of the real robotic arm through continuous measurement of the encoders placed on the real robotic arm. The operator can observe these two representations constantly, avoiding the need to change the control strategy to ''move and wait''. The operation and implementation of this interface are detailed in the Predictive display interface subsection in the methodology section.

C. NUI CONTROL INTERFACE

As mentioned before, the control of the robotic arm is considered, and multiple studies have aimed to improve its maneuverability using different sensors and controllers. While the conventional control of robotic arms through keyboards or joysticks has proven to be functional [4], [27], multiple investigations have shown that the use of natural user interfaces (NUI) allows a significant increase in operator performance and drastically reduces mental workload [1], [18], [32], [33]. This is mainly because NUI interfaces effectively interpret the operator's intentions; in this regard, multiple robotic arm controller devices have been proposed that follow this approach.

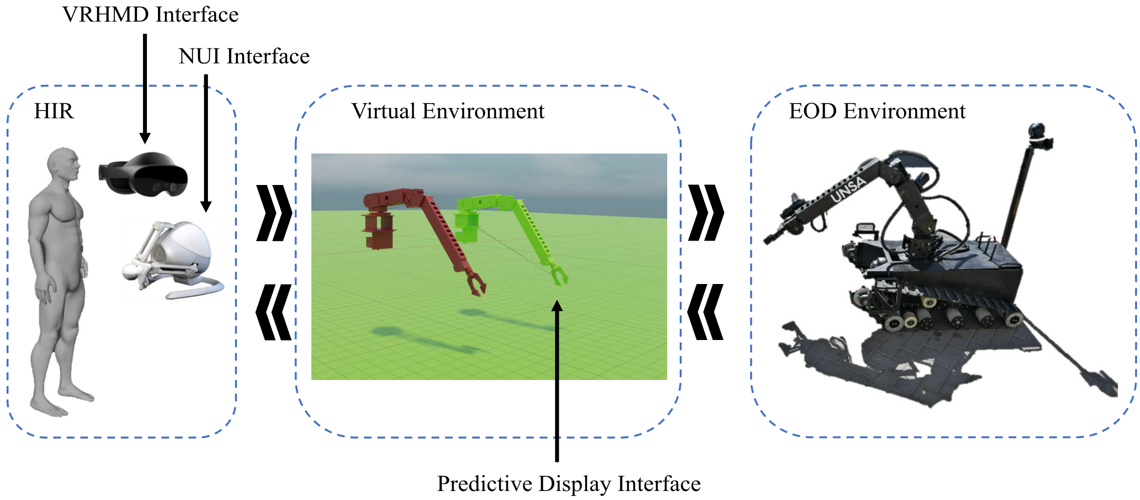


FIGURE 2. Proposed system with the three interfaces that constitute the multimodal interface.

In a previous study, an in-depth analysis of the performance of two NUI interfaces, the Leap Motion sensor and the Novint Falcon controller, was conducted, and it was concluded that the Falcon is optimal for the control of robotic arms in EOD tasks [13]. This control interface has features such as three-dimensional movement, the possibility of exerting force on the operator (Haptic capabilities), and gravitational compensation. Therefore, this interface was selected for controlling the EOD robotic arm. However, as in the [13] case study, the controller is used as a three-dimensional NUI interface without the haptic capability for this research, because making use of this feature requires a larger number of implemented sensors and making major changes to the electromechanical system of the robotic arm in order to measure the forces interacting with it.

IV. HYPOTHESES

To evaluate the performance of the proposed multimodal system, a performance comparison was performed between the proposed multimodal interface and the conventional control interface based on a keyboard and monocular display. The components and operation are detailed in the Existing Architecture subsection of the methodology section. These two systems are evaluated from tests in which a Pick and Place task is performed. The following four hypotheses are proposed on the basis of the measurements taken in the experiments.

- H1: EOD task completion times are significantly reduced using the multimodal interface compared with the conventional interface.
- H2: Less mental workload is experienced performing EOD tasks with the multimodal interface compared with the traditional interface.
- H3: Higher perceived usability is achieved by performing EOD tasks with the multimodal interface compared with the conventional interface.

TABLE 1. Resumen de objetivos y mediciones subjetivas.

Measurement	Type	Metrics
System Usability	Subjective	Questionnaire [Likert Scale]
Cognitive Workload	Subjective	Questionnaire [Likert Scale]
Completion Time	Objective	Seconds [s]
Succeed percentage	Objective	Percentage [%]

To evaluate these parameters in the proposed systems, we use objective and subjective measurements. Table 1 summarizes the measurements considered and the metrics used to evaluate each interface.

V. METHODOLOGY

A. EXISTING ARCHITECTURE

To validate the proposed system in EOD applications, the EOD robot JVC02 was used. This is a response robot composed of a mobile robotic platform with a robotic arm of five degrees of freedom and three cameras located in the front, robotic manipulator, and rear turret. It has a lifting capacity of up to 8 kg with the arm fully extended and more than 20 kg at an operating distance. Its actuators are composed of DC Wiper motors with gearboxes to ensure the necessary torque to lift loads commonly found in EOD operations. It has teleoperated communication through an Access Point and a Wi-fi antenna that allows a stable connection up to 120 m. Figure 3 shows the JVC02 EOD robot being operated during operator training in teleoperated robotics tasks. The movement of this robotic arm is relatively slow due to the high stability required for EOD operations, the speed can be increased but would require heavier gearboxes that would decrease the payload of the EOD robotic arm.

For the control of this robotic system, its control station is used, which is composed of a Getac X500 laptop with 16 GB of RAM, a 15.6” screen with a refresh rate of 60Hz, which is responsible for displaying the three views of the robot’s

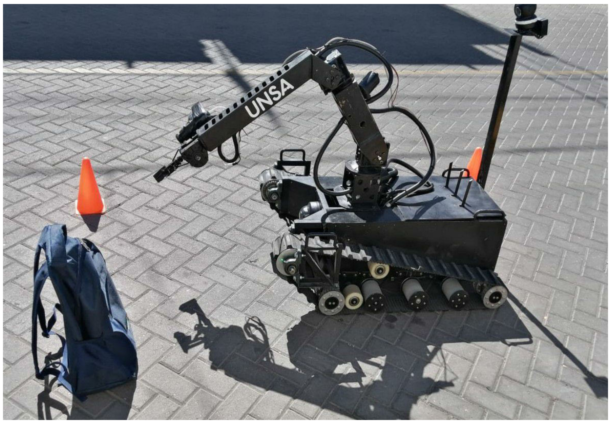


FIGURE 3. EOD Robot JVC02 during training operations.

FIGURE 4. JVC02 robot control station.

cameras; there is also a Wi-Fi antenna that allows radio frequency communication and has a communication range of up to 120 m without line of sight, and 450 m with line of sight. Its internal battery system allows the control station to operate for periods of two hours of uninterrupted work. Figure 4 shows the control station in its respective case.

B. PROPOSED SYSTEM

1) CLOSED LOOP CONTROL IMPLEMENTATION

To implement the proposed multimodal interface, it is necessary to have closed-loop control of each joint of the robotic arm. Feedback for each joint and a control law for the motors must be implemented. The position in three dimensions is given by the Novint Falcon controller, where the rotations of each joint are obtained through inverse kinematics using Blender software. This setpoint is sent by serial communication to an ESP32 DEV-MODULE, which acts as a master device for communication with the rest of the system. The angular position of each joint was measured using magnetic rotational encoders. Likewise, this device computes the PID algorithms for the motors of each joint.

The closed-loop communication and control system for the motors is divided into three sections, as shown in Figure 6. These sections are detailed below:

- **Data processing stage:** This section refers to the ESP32 DEV-MODULE, which is responsible for reading the position data from the three links of the robotic arm. The data array is transmitted with the following syntax: *"Address Angle/n"* where *Address* has a length of two bytes and *Angle* has a length of four bytes, and its values range from 0 to 360. This value represents the angular position of the link in sexagesimal degrees, each manually delimited for the motion range in each joint.
- **Data workflow control stage:** In this section, the master device is a Seeduino based on the ATMega328 controller, which is initially in "transmit mode", that is, the data direction pins are in "high" state. It sends the identifier *$07/n* to all devices and then switches to "receive mode" until it confirms that the received data string is complete and has the expected syntax. The first link corresponding to the elbow has the identifier *$07n* and is initially in "receive mode", that is, with the data direction pins in "low" state; until it receives its identifier, it switches to "transmit mode", reading the AS5600 encoder and transmitting it to the master device. This process is then repeated for the shoulder link with identifier *$08n* and finally for the base link with identifier *$09n*.
- **Control and power stage:** In this stage, the PWM output value is calculated for the PID control implemented for each link, as shown in Figure 5. For the power stage and correct control of the DC motors, BTS7960 H-bridges are used, which can receive PWM signals and can tolerate up to 43 Amperes of peak current.

For hardware implementation, based on the robotic model of the JVC02 robot, additive manufacturing techniques and computer-assisted design software were used to implement three angular position measurement systems for each joint. Each of these systems uses as a central support point the axis of the output of each gearbox of each link and as a second support point the extension of the link that joins the mechanical transmission elements for its movement. In this way, we can use the encoders to read the entire range of angular movement of each link. Each of these measurement systems has the capacity to contain the controller board that is responsible for reading and transmitting data, the AS5600 encoder controller board, and the mechanical transmission mechanisms of the link such as: An 8 mm shaft, the encoder magnet, and its mechanical base. Figure 7 shows the design indicating the elements that have been manufactured by 3D printing.

2) VRHMD INTERFACE

The VRHMD-based interface uses a virtual environment implemented in Blender, a modeling and animation software.

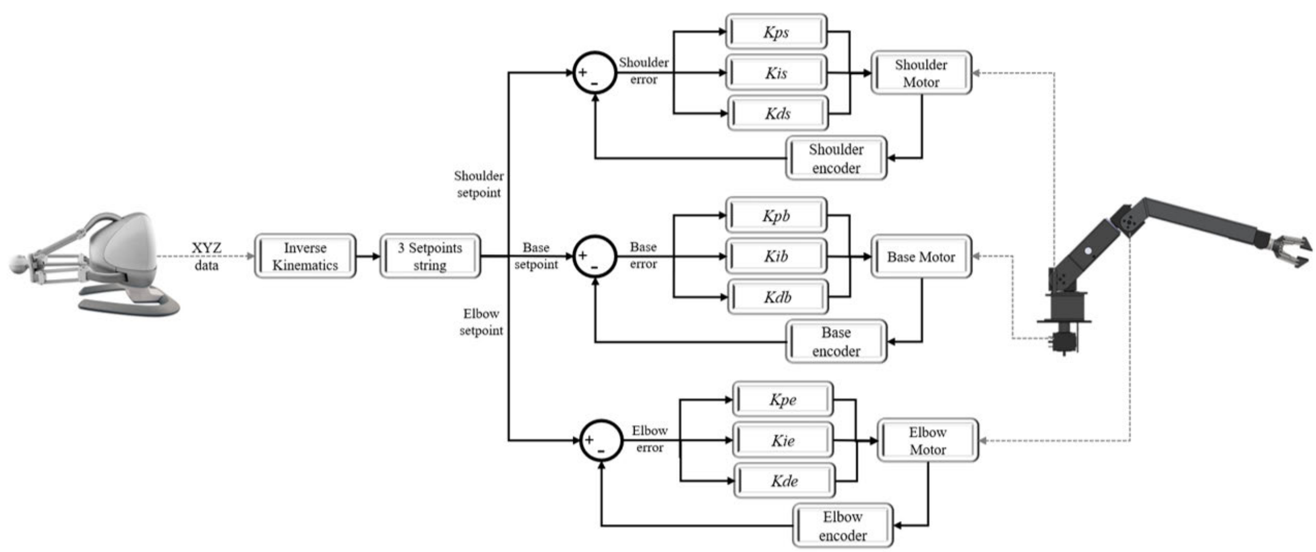


FIGURE 5. Independent PID control system for each joint of the robotic arm using rotary encoders.

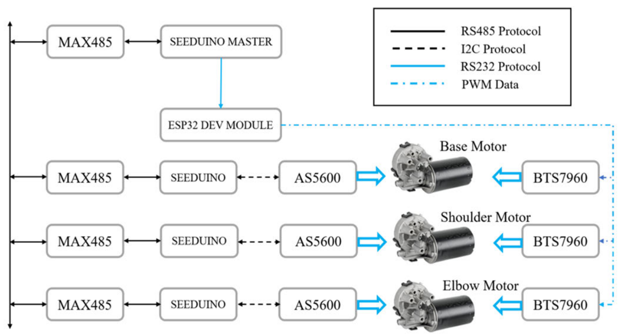


FIGURE 6. Electronic implementation for the closed loop control using rotary encoders.

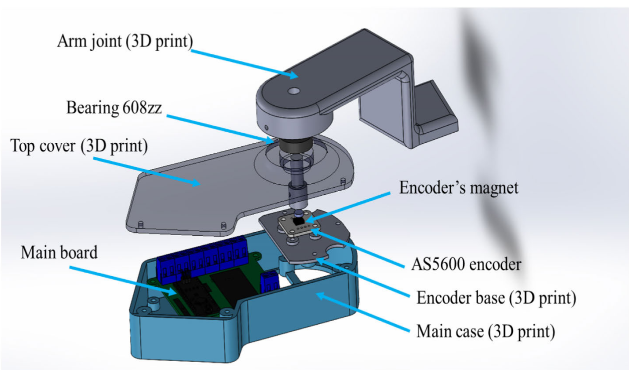


FIGURE 7. 3D design of the encoder adapted to the robotic arm of the EOD JVC02 robot.

The 3D robotic arm design is exported in stereolithography (STL) format and placed in the Blender environment. To make these pieces move coherently, this environment allows the creation of armatures containing bones. In this way, each link of the robotic arm is attached to a bone and can be manipulated directly with the mouse or by an external controller through a Python script. In total, there are four bones conforming to this armature; the first three correspond to the three links of the robotic arm and the last one to the end effector of this robotic arm. For the implementation of this interface, the Meta Quest Pro VRHMD was used with the Air Link connection, which allows the user to move freely without the need for a direct physical connection to the computer. Figure 8 shows the model of the robotic arm on the left, and on the right, the connection with an armature composed of bones. In addition, a reference plane indicating the ground level and a sphere indicating the position of the real-world grenade represented in the virtual environment were included.

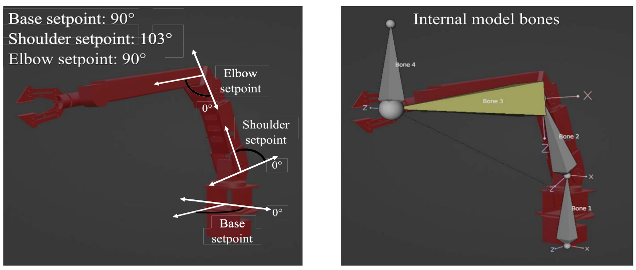


FIGURE 8. 3D design of the robotic arm of the EOD robot JVC02 in Blender.

To provide the operator with an immersive view through the VRHMD, we use the Blender addon called "VR Scene inspection". This addon uses any camera-like object placed in the Blender virtual environment as a viewpoint for the VRHMD. The operator is then able to move through his virtual environment and take different views of the robot with complete clarity due to the virtual representation of the environment in the VRHMD.

3) NUI INTERFACE

For the NUI control interface of the robotic arm, the Novint Falcon controller was used. For its integration with the 3D environment generated in Blender, a Python script and the ForceDimension library were used, which allows serial communication with this controller. The main commands used and their functions are detailed below:

- *forcedimension.drd.open()*: This command allows to establish a reliable communication with the first available device connected and compatible with the Forcedimension library.
- *forcedimension.dhd.setGravityCompensation()*: This command enables gravity compensation, which allows the operator to leave the Falcon effector in any position without it falling due to gravity.
- *forcedimension.dhd.getPosition()*: This command allows to have the Falcon's effector position information in X, Y, and Z dimensions to be continuously obtained, with a default frequency of 10 kHz.

Once the position of the effector of the Falcon controller was obtained, we determined this point in the real three-dimensional space as a Cartesian point in the virtual environment. This point corresponds to the final effector in the representation of the robotic arm, that is, to the last bone of the armor that was implemented in the virtual environment. Since the Falcon controller provides only Cartesian positions, it is necessary to convert this Cartesian position to angles of each link of the EOD JVC02 anthropomorphic robotic arm. To determine the angles of each of the other three bones corresponding to the links of the arm, we used the inverse kinematics provided by Blender from a known position of the end effector. The Novint Falcon workspace is configured within the Python script so that the operator can control the robotic arm without having limitation issues with the EOD robotic arm.To map the workspace of the Novint Falcon onto the robotic arm's control space, the Falcon end effector is brought to its lateral, vertical and horizontal limits, and then these points are programmatically defined as the limits of the controllable workspace on the robotic arm. Figure 9 shows in green the virtual robotic arm that responds to the commands sent by the operator through the Falcon controller and the armature created in Blender.

With this NUI interface, the operator is not required to have extensive knowledge and training with the robotic arm, but instead performs three-dimensional movements of the robot end-effector, and the rest of the robotic arm responds coherently with the kinematics of a three-degree-of-freedom robotic arm. Subsequently, the information from these rotations of this virtual robotic arm "guide" is carried through the system detailed in "Closed loop implementation" with the ESP32 DEV-MODULE to the real robotic arm.

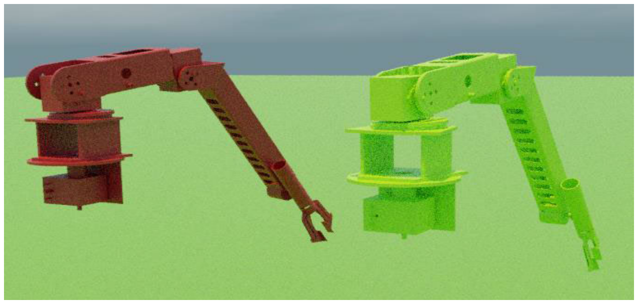


FIGURE 9. Robotic arm "guide" and robotic arm "follower" on virtual interface.

4) PREDICTIVE DISPLAY INTERFACE

The predictive display interface is implemented by creating a copy of the virtual robotic arm "guide" with the same characteristics in red color, called "follower" robotic arm. This "follower" robotic arm receives the rotation information of each link of the real robotic arm and copies these rotations to each of the virtual links.Both models are based on the original 3D design used for constructing the robotic arm, significantly reducing the error between the real and virtual robotic arms. Thus, this robotic arm displays the actual position of the EOD robot as measured by the AS5600 rotary encoders. In this way, the operator visualizes within the virtual environment a "guide" robotic arm that responds instantaneously to the commands generated by the NUI Novint Falcon controller, and a "follower" robotic arm that responds to the movements of the real robotic arm. The operator can then perform continuous maneuvers with the "guide" robotic arm as he receives constant feedback of the actual position of the robotic arm, allowing him to continuously control the robotic arm avoiding the "move and wait" strategy.

C. PARTICIPANTS

Thirteen experienced agents from the Arequipa explosives deactivation unit (UDEX-AQP) participated in the tests, of which ten agents had used the JVC-02 EOD robot at least once and three had never used any EOD robot before.

D. EOD TASK

To evaluate the proposed system in EOD tasks, Pick and Place tasks were performed using the proposed system, and the results were compared with those of the conventional keyboard-based system. For this purpose, a metal prop grenade is placed in the workspace of the robotic arm. Because the objective of this preliminary research is to control the position of the robotic manipulator, only the movements of the three degrees of freedom specified in the Closed Loop Implementation section are considered. To pick up the target using the robotic gripper, neodymium magnets located in the end effector are used. Figure 10 shows the prop grenade and neodymium magnets attached to the end effector.

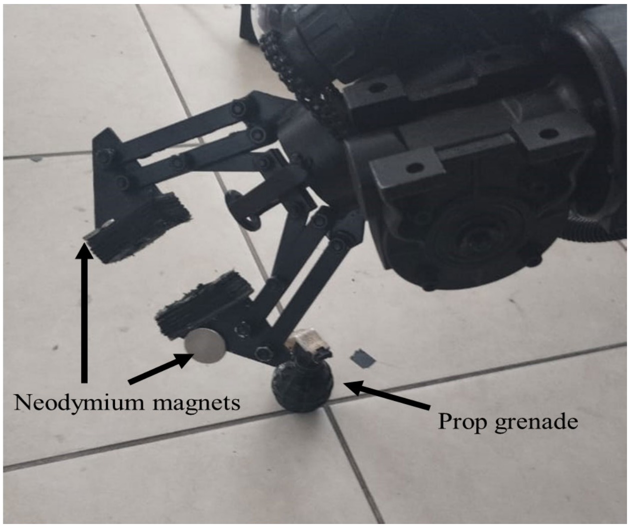


FIGURE 10. Neodymium magnets attached to the end-effector together with the prop grenade.

After a ten minute adaptation time with each of the control interfaces, each participant performed the Pick and Place task with both interfaces twice, the results from both attempts using both interfaces are displayed in Table 2. Each of the Pick and Place tasks starts with the robotic arm in a contracted position with the elbow and shoulder retracted. Once the object has been placed in the work area of the robotic arm, the evaluator instructs the operator to begin the task. The operator should pick up the object using the end effector with attached magnets and place it in the target area. Once the object has touched the target area, the task is considered completed. If the operator is not able to pick up the object or has failed to place the object in the target area, it is considered a failed attempt. The pick and place task is performed using the proposed multimodal system and the conventional keyboard-based system detailed above for comparative evaluation. Figure 11 shows an explosive ordnance disposal unit agent using both interfaces for the proposed Pick and Place task.

VI. EVALUATION

To evaluate the performance of this new multimodal EOD robot control interface, an evaluation of the two systems presented below is proposed:

- **Proposed System**: The proposed multimodal interface system described in the PROPOSED SYSTEM subsection of the METHODOLOGY section.
- **Conventional System**: The conventional system based on a control station that has a keypad and a display for visual feedback. This system is detailed in the EXISTING ARCHITECTURE subsection of the METHODOLOGY section.

To rigorously evaluate the performance of the new interface, subjective and objective measures are proposed as follows:

A. OBJECTIVE MEASUREMENTS

The percentage of successfully completed tasks and task execution time are measured by the evaluators during task execution. To measure this, the first evaluator is in charge of giving instructions to the agents about the Pick and Place task, placing the target, and designating the final placement area, while the second evaluator takes the completion times and records the unsuccessful and successful attempts. Each attempt where the object is successfully picked up and placed in the designated location is considered a successful task; if the object cannot be picked up by the operator, is not placed near the target area, or if the twelve minutes of operation is exceeded, it is considered a failed attempt. Each task was repeated twice with each of the thirteen participants.

B. SUBJECTIVE MEASUREMENTS

To evaluate the performance of the interface in terms of mental workload, the NASA Task Load Index (NASA-TLX) questionnaire was used with six subscales: mental demand, physical demand, performance achieved, effort, and frustration. In addition, the System Usability Questionnaire (SUS) was used with a total of ten questions on a Likert scale from one to five, where one represents "Strongly Agree" and five "Strongly Disagree". Each agent was instructed to complete these questionnaires at the end of the Pick and Place tasks.

VII. RESULTS

A. OBJECTIVE RESULTS

The completion times of the thirteen participants for both attempts with the systems are compiled in Table 2 and visualized in a heat map for better visual representation. In the last column a "No" has been placed when the agent has had previous experience using the EOD robot and a "Yes" when the agent has no previous experience with EOD robotics. The times recorded in this table correspond to the successful attempts made by the operators.

To statistically evaluate these data, we used the Shapiro-Wilks test to verify the normality of each group of data. For the times recorded for the conventional interface, we have a P-value of 5.68×10^{-6} and 3.1×10^{-3} for the case of the times obtained with the proposed interface. In both cases, the P value is much lower than the significance level of 0.05, so the null hypothesis of this test is rejected and it is concluded that these data do not follow a normal distribution. Therefore, the nonparametric Wilcoxon test is used to determine if there are significant differences between these groups of data, considering that each group of data had repeated measures. With the Wilcoxon test applied to each group of repeated measures data, a value $P = 0.347$ is obtained for the conventional interface. Because the null hypothesis cannot be rejected, no significant differences were found between the repeated measures of the conventional interface. For the repeated measures of the multimodal interface, the value $P = 0.008$ is obtained, i.e., the null

FIGURE 11. Interfaces being used in Pick and Place task. (1) Side view of the Proposed Interface used by UDEX Agent. (2) Frontal view of the Proposed Interface used by UDEX Agent. (3) Conventional keyboard interface. (4) EOD Robot executing Pick and Place Task.

TABLE 2. Pick and place task times obtained with both interfaces.

	Conventional Interface		Proposed Interface		Novice
	Time 1 (s)	Time 2 (s)	Time 1 (s)	Time 2 (s)	
Agent 1	178	194	94	59	No
Agent 2	210	144	110	114	No
Agent 3	150	175	73	47	No
Agent 4	68	68	33	27	No
Agent 5	76	125	59	28	No
Agent 6	86	118	24	27	No
Agent 7	96	75	64	62	No
Agent 8	150	121	112	55	No
Agent 9	144	190	63	56	No
Agent 10	75	65	68	51	No
Agent 11	646	226	88	75	Yes
Agent 12	490	196	50	30	Yes
Agent 13	488	77	25	28	Yes

hypothesis is rejected, affirming that significant differences are found in the repeated measures. Finally, the same test is used to check if there is a significant difference between both groups of data, that is, between both interfaces in terms of measured times; obtaining a P-value of $P = 2.98 \times 10^{-8}$. Because this value is much smaller than the significance level of 0.05, it is determined that there are significant differences between the times obtained from both interfaces.

Table 3 shows the failed and successful attempts of the thirteen agents in the Pick and Place tests. Again, as in Table 2, the eleventh, twelfth, and thirteenth agents have no previous experience operating EOD robots. The last two columns represent the percentage of successful tasks performed by each agent using the evaluated interfaces.

To observe the response times of the movements of the base, shoulder, and elbow of the EOD robot, this behavior is plotted in Figures 12, 13, and 14 during one of the adaptation times of the UDEX agents. In the case of the movement of the base joint, it can be observed that there is a time of approximately ten seconds to move ninety sexagesimal degrees on its own axis. The shoulder joint has a time of approximately fifty-two seconds for a movement of ninety sexagesimal degrees on its own axis. Finally, the elbow joint has a time of fifty-six seconds for a movement of ninety degrees sexagesimal on its own axis.

TABLE 3. Rate of task completion of Pick and Place tasks using both interfaces.

Agent	Conventional Interface	Multimodal Interface	Total Attempts	Conventional Interface Success Rate (%)	Multimodal Interface Success Rate (%)
1	2	2	4	100.00	100.00
2	2	2	4	100.00	100.00
3	2	2	4	100.00	100.00
4	2	2	4	100.00	100.00
5	3	2	5	66.67	100.00
6	2	2	4	100.00	100.00
7	2	2	4	100.00	100.00
8	2	2	4	100.00	100.00
9	2	2	4	100.00	100.00
10	2	2	4	100.00	100.00
11	3	2	5	66.67	100.00
12	3	2	5	66.67	100.00
13	4	2	6	50.00	100.00
		Average rate		88.46	100.00

FIGURE 12. EOD robotic arm base response time to different setpoints.

B. SUBJECTIVE RESULTS

The results obtained from the NASA-TLX questionnaire are shown in Figure 15 as a bar chart of the average of the thirteen agents. The questionnaire is given in a Likert-type scale from 1 to 20, where 1 indicates complete disagreement and 20 indicates complete agreement. Mental demand averaged 2.3 points for the proposed interface and 11.3 points for the conventional interface. Physical demand

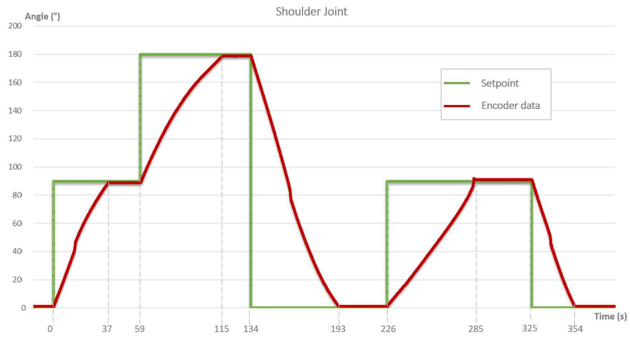


FIGURE 13. Response time of the EOD robotic arm shoulder to different setpoints.

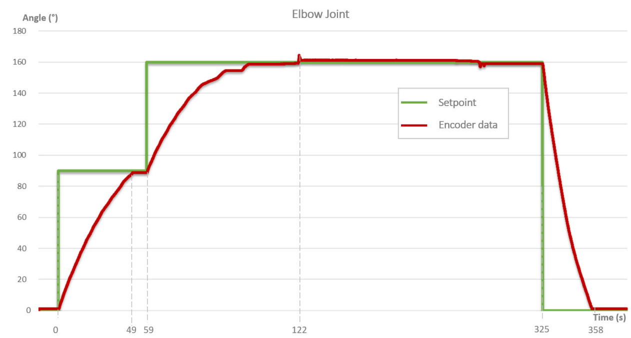


FIGURE 14. Response time of the EOD robotic arm elbow to different setpoints.

averaged 2.3 points for the proposed interface and 10.9 points for the conventional interface. Temporal demand averaged 3.7 points for the proposed interface and 14.4 points for the conventional interface. The performance obtained averaged 14.2 points for the proposed interface and 17.9 points for the conventional interface. The perceived effort averaged 2.8 points for the proposed interface and 15.8 points for the conventional interface. Perceived frustration averaged 1.4 points for the proposed interface and 6.5 points for the conventional interface.

To evaluate these results, the Shapiro-Wilks test was used to verify normality. Where a value $P = 0.433$ is obtained for the results of the NASA-TLX test in the multimodal interface and a value $P = 0.129$ in the conventional interface. None of these values exceed the significance value of 0.05; therefore, the null hypothesis cannot be rejected, indicating that the data follow a normal distribution. Following, it is verified if there is a variance homogeneity through Levene's test, where a value $P = 0.713$ is obtained, because this value is higher than the significance level of 0.05, it is determined that there is a variance homogeneity between these two groups of data. Having determined the normality and homogeneity of variances, we used the T-test to check if there was a significant difference between the two groups of data. We obtained a $P = 2.998\ X10^{-8}$. With this P-value significantly lower at a significance level = 0.05, it is determined that there is a significant difference in the results of the NASA-TLX questionnaire of both interfaces.

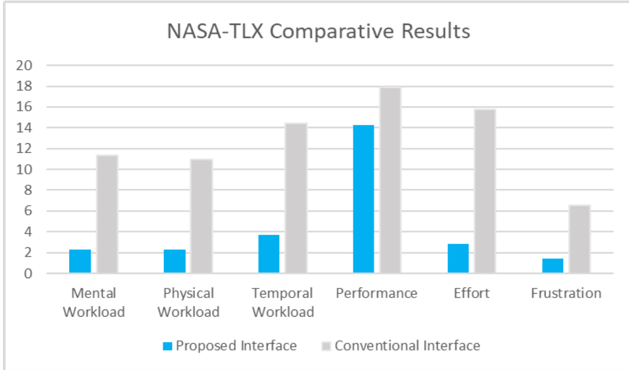


FIGURE 15. Results obtained with the NASA TLX questionnaire in relation to the Pick and Place task.

The results of the SUS questionnaire are shown in Figure 16 as a comparative bar chart. In this case, comparative results are shown for each agent. For the calculation of the SUS questionnaire score, it should be considered that the odd and even questions have inverted scales; that is, the high scores of the odd questions indicate that the system has a higher usability, whereas the high scores of the even questions indicate a lower usability. To determine the total score of the SUS questionnaire, the scores obtained for each question were calculated using the following formula:

$$SUSscore = ((\sum OddQuestions - 5) + (25 - \sum EvenQuestions)) * 2.5 \quad (1)$$

where *SUSscore* is the score of the SUS questionnaire, *Odd Questions* is the score of the odd questions and *Even Questions* is the score of the even questions.

To evaluate the results obtained in the SUS questionnaire, the normality of these groups of data is analyzed through the Shapiro Wilks test, obtaining a value $P = 0.267$ for the results of the multimodal interface and a value $P = 0.076$ for the results of the conventional interface. Because both values are greater than 0.05, the null hypothesis cannot be rejected and both groups are found to have a normal distribution. Then, the homogeneity of variances of both groups is tested through Levene's test, where a value $P = 0.593$ is obtained, since this value exceeds the significance level of 0.05, it is determined that there is homogeneity of variances. Having confirmed the homogeneity of variances and the normality of both groups of data, a T-test is performed to determine if there is a significant difference between these two groups of data. Obtaining a value $P = 1.37 \times 10^{-8}$, as this value is much lower than the significance level it is determined that there is a significant difference between both groups of data.

VIII. DISCUSSION

In this section, we discuss the results presented in the previous section with a focused approach to the proposed interfaces and how they affect operator performance during EOD tasks.

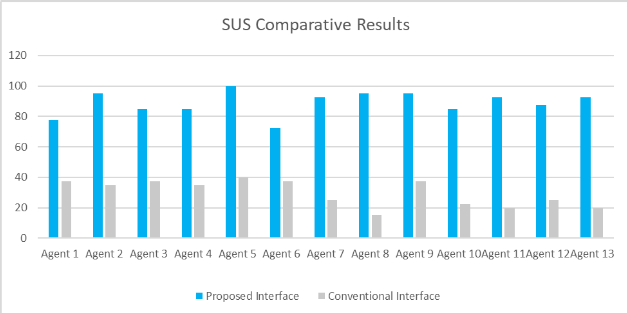


FIGURE 16. Results obtained with the SUS questionnaire in relation to the Pick and Place task.

A. TIME COMPLETION ANALYSIS

One of the main performance indicators in the proposed Pick and Place tasks is the task completion time. These were measured objectively twice for each participant in the two evaluated interfaces, as shown in Table 2, where the times obtained with the conventional interface have an average of 178.1 s, whereas the proposed multimodal interface has an average time of 58.53 s, also it is important to clarify that this calculation does not take into account the failed attempts of the participants. This represents a significant improvement of 67% in task completion times using the proposed multimodal interface. Furthermore, as shown by the statistical tests conducted in Section VIII, there is a significant difference between the times of the two interfaces. In addition, there is a significant difference between the first and second attempt time measurements for the multimodal interface, whereas this significant difference cannot be confirmed for conventional interface attempts. This indicates that the proposed multimodal interface not only greatly improves the completion times of the teleoperated robotics tasks, but there is also a significant improvement between the first and second attempts performed by the UDEX agents. This might indicate that the multimodal interface also allows more accelerated learning of the control system by the operator. However, this requires further analysis regarding the learning curve that this new interface allows.

These significant improvements in the proposed multimodal interface are mostly due to two aspects implemented in this interface: The ease of control of the robotic arm allowed by the NUI interface, which has the same effect as that observed in previous research (1,4,13); and the change of perspective allowed by the VRHMD-based interface. Thus, the operator can correctly identify the position of the target with respect to the robotic arm and quickly bring the robotic arm to the target position and subsequently to the location where it should be placed. This is also clearly evidenced in the comparative results of the mental workload shown in Figure 15, where the perceived effort by the operators is significantly reduced using the multimodal interface compared with the conventional interface. Likewise, frustration is decreased, indicating that the operators can perform the task without having many complications using the multimodal interface. Some comments made by the operators during and after performing the Pick and Place task with both interfaces were: "There is a clear difference in the time it took me to complete the task. It is very easy to operate the arm with the new interface", "I can see perfectly where the target is, that let me finish the task faster".

B. SUCCESS RATE ANALYSIS

Another important indicator of improved performance in teleoperated tasks is the Completion Rate. The Pick and Place tests proposed in this research emulate a common task performed in explosive ordnance disposal situations, where the operator must control the robotic arm to pick up a suspicious package and place it in an area where it will be detonated in a controlled environment or stored if it is safe. This is a critical procedure because it poses a great risk to the people and infrastructure near the EOD operation. The proposed multimodal interface has improved the completion rate of this type of task. This improvement is more evident in users new to teleoperated robotics. As shown in Table 3, novice operators 11, 12, and 13 required up to four attempts to complete the two tasks using the conventional interface, whereas no operator had an unsuccessful attempt with the multimodal interface. This indicates that the multimodal interface provides an easy-to-learn system for novice operators and drastically reduces the risk of a failed operation in a real explosive hazard situation.

This significant improvement in the completion rate is mostly due to the VRHMD-based interface, which allows the operator to have an optimal perspective of the robotic arm and the target, as well as a clear image quality and a stereoscopic view of his surroundings. With this interface, the operator has no problems in guiding the robotic arm to the target and then to the area where it should be placed. Some comments from operators with no previous experience in teleoperated robotics were as follows: "The system is intuitive once you understand how it works", "The control is much easier and I can understand what is going on all the time with the robot".

C. MENTAL WORKLOAD ANALYSIS

The mental workload of a system is a relevant indicator for evaluating the degree of complexity perceived by the system operator. The results obtained using the NASA.TLX questionnaire shown in Figure 15 indicate a significant improvement in the aspects evaluated. Being much more noticeable in mental workload, effort, time demand, and frustration. This indicates that the proposed interface drastically decreases the mental workload of the operator by improving the control of the robot and the feedback received by the operator.

According to operator feedback, the biggest contribution to the decrease in mental workload was the predictive display interface. This interface allowed the operators to continuously control the robotic arm, avoiding the need to wait for visual confirmation as with the conventional interface. As shown in Figures 12 and 13, the time required for the robotic arm to reach the setpoint indicated by the operator is quite long, taking more than a minute for the movement of a link of ninety sexagesimal degrees. This long time is due to the limited speed of the Wiper DC motors used by the JVC02 robot. As Table 2 and Figure 15 show, the delay in which the robotic arm responds causes operators to take time to complete the task and have a high mental workload when using the conventional interface. However, the multimodal interface through the predictive display-based interface has proven to drastically decrease the mental workload and has greatly improved the performance of the operators. Some comments from operators regarding mental workload were as follows: ''Once you understand the movement of the two robotic arms in the virtual environment, it is easy to pick up and place the object'', ''Being aware of where the robotic arm will be placed is a big help''.

D. SYSTEM USABILTY ANALYSIS

The proposed multimodal system has also been shown to have a higher perceived usability by users, as shown in Figure 16, with an average score according to the SUS procedure of 88.84 points, whereas the conventional interface had an average of 29.80 points. According to the questionnaires evaluated, the agents widely prefer to use the proposed multimodal system and consider it to be appropriately integrated. These results obtained in the SUS questionnaire are due to a correctly integrated operation of the multimodal interface, showing that the interfaces that compose it are adequate for this system. However, the results shown in Figure 16 are comparative; therefore, a wider range of interfaces, their combinations, and their effect on operator performance and perceived usability should be considered.

E. MULTIMODAL INTERFACE ANALYSIS

This subsection will discuss relevant details regarding the interfaces that compose the proposed multimodal system and their effect on the performance results obtained.

1) VRHMD INTERFACE ANALYSIS

The VRHMD-based interface positively influences operator performance. The stereoscopic vision provided by this interface, as well as a clear and adaptable virtual environment for the operator, allows users to perform the proposed Pick and Place task with ease. As shown by several studies, improving visual feedback to the operator results in remarkable performance improvement in teleoperated robotics [17], [25], [26], [34].

It is important to mention that after performing the Pick and Place tasks, some operators reported feeling a slight motion sickness when removing the VRHMD. This is common according to other research, which is usually because the pupillary distance is different for each user; therefore, when using the VRHMD, the brain interprets the visual signals to provide a stereoscopic effect to the operators. However, when the glasses are removed, the brain must adapt to the original visual information. This effect can be reduced by considering the pupillary distance of each operator to obtain a customized interface.

2) NUI INTERFACE ANALYSIS

The NUI interface allows operators to have much shorter task completion times than the conventional interface. This was achieved through the Novint Falcon controller and Blender software, in which a virtual robotic arm was recreated and inverse kinematics was implemented so that the operator does not have to deal with making independent movements of each link of the robot. This interface allows the operator to make movements in three dimensions with the controller and the robot to coherently move its joints to reach the indicated position. As seen in previous research [4], [13], [18], [35], this greatly improves the time it takes for the operator to control the robotic arm and drastically reduces the mental workload and frustration of the operator as shown in Figure 15.

It is necessary to mention that the Novint Falcon controller was used only as a system to indicate the target position of the robotic arm. However, this controller also offers the ability to function as a haptic system, i.e., to recreate forces on the operator to obtain more information about what happens with the real robotic arm. This implementation requires a greater number of sensors in the arm, and several studies have indicated an improvement in the performance of tasks in teleoperated robotics.

3) PREDICTIVE DISPLAY INTERFACE ANALYSIS

The predictive display-based interface drastically decreased the mental workload of the operators to complete the Pick and Place tasks, as shown in Figure 15. This is mainly because the operator no longer has to wait for visual feedback of the actual position of the robotic arm but can perform continuous control of the robotic arm [20], [36]. This was achieved by implementing two identical virtual robotic arms within the virtual environment, the ''guide'' robotic arm lets the operator know the position indicated by the operator and to which the actual robotic arm is directed, while the ''follower'' robotic arm lets the operator know the actual position of the EOD robot.

Several comments following the Pick and Place tasks from the operators mention that the system works adequately. However, it is also necessary to have a previous training period in which the operation of these two virtual robotic arms is explained to avoid confusion in their operation.

Finally, according to the statistical analysis carried out in the results section where a significant difference is observed between the times obtained, the percentage of task completion, the mental workload, and the usability of the system between the multimodal interface and the conventional interface, the hypotheses raised in the Hypotheses section are verified as shown in Table 4.

TABLE 4. Validated hypotheses regarding the performance of interfaces in EOD tasks.

Hypotheses	Analysis Method	Support
H1: Less Completion time with proposed interface	Wilcoxon Test Analysis	Full Support
H2: Better Success rate with proposed interface	Analysis of averages	Full Support
H3: Less Mental Workload with proposed interface	T test Analysis	Full Support
H4: Better Usability with proposed interface	T test Analysis	Full Support

IX. LIMITATIONS AND FUTURE WORK

In this study, we presented a multimodal interface focused on improving the performance of teleoperated robotic EOD tasks. However, for this system to be used in a real EOD situation, specific aspects still need to be developed. Some limitations of the proposed system and future work are detailed below:

- The analysis carried out in this research focused on identifying the most suitable interface for the control of an EOD robot. However, the individual performance of each interface that composes the multimodal system has not been evaluated. This analysis will be carried out in detail in a subsequent research.
- The proposed multimodal system uses components that are directly connected to the EOD robot; therefore, further development is required to enable the use of this interface in a real teleoperated EOD situation.
- The NUI interface selected for the multimodal system is the Novint Falcon controller, which has a haptic capability that was not used in this study. The necessary system will be implemented to take advantage of this haptic capability in future research.
- During the tests some operators reported having a slight dizziness of a few seconds when removing the VRHMD, this may be because the individual pupillary distance of each agent was not considered, in future tests should be considered to make this customization of the VRHMD for each user.

These issues will be studied and analyzed in a future research focused on having a multimodal interface applied to a real EOD robotics environment.

X. DESIGN AND RESEARCH COMPLICATIONS

The three interfaces that compose the multimodal system presented in this research have been taken as a reference from multiple investigations that have demonstrated improved performance when applied to teleoperated robotics. The results indicate that this combination of interfaces provides operators with an intuitive and efficient multimodal interface for EOD robotics tasks. Considering that EOD tasks are essentially robotic arm control tasks, this indicates that this interface can also improve user performance in other types of tasks using robots, such as robotic operation applied to part assembly, welding, or rescue robotics.

This new multimodal interface validated through testing by experienced explosive ordnance disposal agents has demonstrated that it is feasible to improve operator performance by implementing interfaces that adapt to human nature such as stereoscopic vision, natural motion and visual feedback. Designers and researchers in this field should focus on interfaces that take into account natural human actions and responses to improve the performance of teleoperated robotics.

XI. CONCLUSION

In this study, a new immersive multimodal interface for controlling an EOD robotic arm was implemented. To validate the proposed system, a comparison of the performance achieved by operators in a Pick and Place EOD task was performed. The tests were evaluated using subjective and objective measures. The objective tests were measured by task completion times and percentage of completion; while the subjective tests were measured by standardized NASA-TLX mental workload and SUS usability questionnaires. This proposed novel multimodal interface has allowed operators to have optimal control of the robotic arm, with an average task completion time of 178.1 s in the conventional interface and 58.53 s in the multimodal interface; an improvement of 67. 14% when using the multimodal interface. Likewise, it is observed that in the percentage of task completion, the multimodal interface had a 100% success rate in the agents, compared to the 86% obtained with the conventional interface, with a significant improvement of 11.54%. In addition, there was a 65.18% decrease in mental workload compared with the conventional interface and a 198.12% increase in perceived usability. However, it should be emphasized that for this positive effect to be evident, a clear understanding of the functioning of the integrated multimodal interface is required, so it is necessary for operators to have a adaptation time to this new proposed interface. This new interface is highly effective in significantly improving the performance of the operators, without the need to make significant changes to the electromechanical system of the robot or to the telecommunications.

ACKNOWLEDGMENT

The authors would like to thank the agents of the explosive deactivation unit of Arequipa UDEX-AQP for their participation in the testing of the interfaces described in this article. They would especially like to thank Emmy Chow for her support with the libraries needed to communicate with the Novint Falcon device using Python.

REFERENCES

[1] N. O. M. Chilo, L. F. C. Ccari, E. Supo, E. S. Espinoza, Y. S. Vidal, and L. Pari, "Optimal signal processing for steady control of a robotic arm suppressing hand tremors for EOD applications," *IEEE Access*, vol. 11, pp. 13163–13178, 2023.

[2] S. Habibian, M. Dadvar, B. Peykari, A. Hosseini, M. H. Salehzadeh, A. H. M. Hosseini, and F. Najafi, "Design and implementation of a maxi-sized mobile robot (Karo) for rescue missions," *ROBOMECH J.*, vol. 8, no. 1, p. 1, Jan. 2021.

[3] J. G. Mamani, P. Pinto, D. V. Goyzueta, E. S. Colquehuanca, E. S. Espinoza, and Y. S. Vidal, "Compilation and analysis of requirements for the design of an explosive ordnance disposal robot prototype applied," in *UDEX-Arequipa*, C. Stephanidis, M. Antona, and S. Ntoa, Eds., London, U.K.: HCI International, 2021.

[4] D. V. Goyzueta, J. G. Mamani, E. S. Espinoza, E. S. Colquehuanca, Y. S. Vidal, and P. Pinto, "Evaluation of a NUI interface for an explosives deactivator robotic arm to improve the user experience," in *Proc. Int. Conf. Hum.-Comput. Interact.*, 2021, pp. 288–293.

[5] F. Luqiao, Y. Xifan, Q. Hengnian, J. Liangzhong, and W. Wei, "A single-hand and binocular visual system for EOD robot," in *Proc. IEEE Int. Conf. Autom. Logistics*, Jinan, China, Aug. 2007, pp. 1930–1935.

[6] T. Horváth, "Emergency cases at countering improvised explosive devices (C-IED), and their potential management," *Land Forces Acad. Rev.*, vol. 24, no. 2, pp. 95–106, Jun. 2019.

[7] M. Bishop, "Analysis of explosive events," in *Blast Injury Science and Engineering*. Cham, Switzerland: Springer, 2022, pp. 115–131.

[8] Y. S. Vidal, C. E. Supo, C. M. Ccallata, G. J. Mamani, M. Betancur, C. B. Pino, P. L. P. Pinto, and E. S. Espinoza, "Analysis and evaluation of a EOD robot prototype," in *Proc. IEEE Int. IoT, Electron. Mechatronics Conf. (IEMTRONICS)*, Jun. 2022, pp. 1–6.

[9] J. Zhao, T. Han, X. Ma, W. Ma, C. Liu, J. Li, and Y. Liu, "Research on kinematics analysis and trajectory planning of novel EOD manipulator," *Appl. Sci.*, vol. 11, no. 20, p. 9438, Oct. 2021.

[10] H. Bećirspahić, H. Basarić, T. Namas, and B. Durakovic, "Minimum viable product: A robot solution to EOD operations," *Defense Secur. Stud.*, vol. 3, pp. 22–31, Jun. 2022.

[11] B. Wei, J. Gao, J. Zhu, and K. Li, "Design of a large explosive ordnance disposal robot," in *Proc. 2nd Int. Conf. Intell. Comput. Technol. Autom.*, 2009, pp. 403–406.

[12] M. Berkemeier, E. Poulson, E. Aston, J. Johnston, and B. Smith, "Development and enhancement of mobile robot arms for EOD applications," in SPIE Proceedings, 2008.

[13] N. O. M. Chilo, L. F. C. Ccari, E. Supo, Y. S. Vidal, E. S. Espinoza, and L. Pari, "A comparison of NUI interfaces as control interfaces for robotic arms in EOD applications," in *Proc. 11th Int. Conf. Control, Mechatronics Autom. (ICCMA)*, Nov. 2023, pp. 221–226.

[14] D. S. Pamungkas and K. Ward, "Tele-operation of a robot arm with electro tactile feedback," in *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatronics*, Jul. 2013, pp. 704–709.

[15] T. Arai, M. Yasuda, and N. Matsuhira, "Development of a teleoperated robot arm system using RSNP: Precise tasks performed using a predictive display," in *Proc. 13th Int. Conf. Ubiquitous Robots Ambient Intell. (URAI)*, Aug. 2016, pp. 128–131.

[16] S. Martin and N. Hillier. *Characterisation of the Novint Falcon Haptic Device for Application as a Robot Manipulator*. Asn.au. Accessed: Apr. 3, 2024. [Online]. Available: https://www.araa.asn.au/acra/acra2009/papers/pap127s1.pdf

[17] E. Triantafyllidis, C. Mcgreavy, J. Gu, and Z. Li, "Study of multimodal interfaces and the improvements on teleoperation," *IEEE Access*, vol. 8, pp. 78213–78227, 2020.

[18] D. V. Goyzueta, J. M. Guevara, A. A. Montoya, E. Sulla, and S. Y. Lester, "Analysis of a user interface based on multimodal interaction to control a robotic arm for EOD applications," *Electronics*, vol. 11, no. 11, p. 1690, May 2022.

[19] G. Lunghi, R. Marin, M. Di Castro, A. Masi, and P. J. Sanz, "Multimodal human-robot interface for accessible remote robotic interventions in hazardous environments," *IEEE Access*, vol. 7, pp. 127290–127319, 2019.

[20] J. Y. C. Chen, E. C. Haas, and M. J. Barnes, "Human performance issues and user interface design for teleoperated robots," *IEEE Trans. Syst. Man Cybern. C, Appl. Rev.*, vol. 37, no. 6, pp. 1231–1245, Nov. 2007.

[21] M. H. Korayem and V. Vahidifar, "Detecting hand's tremor using leap motion controller in guiding surgical robot arms and laparoscopic scissors," *Measurement*, vol. 204, Nov. 2022, Art. no. 112133.

[22] J. B. Van Erp and P. Padmos, "Image parameters for driving with indirect viewing systems," *Ergonomics*, vol. 46, no. 15, pp. 1471–1499, Dec. 2003.

[23] M. J. Massimino and T. B. Sheridan, "Teleoperator performance with varying force and visual feedback," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 36, no. 1, pp. 145–157, Mar. 1994.

[24] S. Livatino, D. C. Guastella, G. Muscato, V. Rinaldi, L. Cantelli, C. D. Melita, A. Caniglia, R. Mazza, and G. Padula, "Intuitive robot teleoperation through multi-sensor informed mixed reality visual aids," *IEEE Access*, vol. 9, pp. 25795–25808, 2021.

[25] Y. Su, X. Chen, T. Zhou, C. Pretty, and G. Chase, "Mixed reality-integrated 3D/2D vision mapping for intuitive teleoperation of mobile manipulator," *Robot. Comput.-Integr. Manuf.*, vol. 77, Oct. 2022, Art. no. 102332.

[26] J. Chen, M. Glover, C. Yang, C. Li, Z. Li, and A. Cangelosi, "Development of an immersive interface for robot teleoperation," in *Towards Autonomous Robotic Systems*. Cham, Switzerland: Springer, 2017, pp. 1–15.

[27] *T7TM Multi-Mission Robotic System*. L3Harris Fast. Forward. Accessed: Apr. 3, 2024. [Online]. Available: https://www.l3harris.com/all-capabilities/t7-multi-mission-robotic-system

[28] J. P. McIntire, P. R. Havig, and E. E. Geiselman, "Stereoscopic 3D displays and human performance: A comprehensive review," *Displays*, vol. 35, no. 1, pp. 18–26, Jan. 2014.

[29] H. Martins and R. Ventura, "Immersive 3-D teleoperation of a search and rescue robot using a head-mounted display," in *Proc. IEEE Conf. Emerg. Technol. Factory Autom.*, Sep. 2009, pp. 1–8.

[30] S. R. Ellis, K. Mania, B. D. Adelstein, and M. I. Hill, "Generalizeability of latency detection in a variety of virtual environments," in *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, 2004, pp. 2632–2636.

[31] J. Corde Lane, C. R. Carignan, B. R. Sullivan, D. L. Akin, T. Hunt, and R. Cohen, "Effects of time delay on telerobotic control of neutral buoyancy vehicles," in *Proc. IEEE Int. Conf. Robot. Autom.*, Sep. 2003, pp. 2874–2879.

[32] B. van Amsterdam, I. Funke, E. Edwards, S. Speidel, J. Collins, A. Sridhar, J. Kelly, M. J. Clarkson, and D. Stoyanov, "Gesture recognition in robotic surgery with multimodal attention," *IEEE Trans. Med. Imag.*, vol. 41, no. 7, pp. 1677–1687, Jul. 2022.

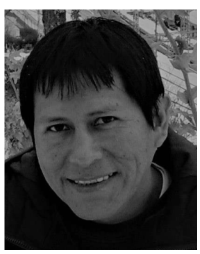
[33] Y. Nam, B. Koo, A. Cichocki, and S. Choi, "GOM-face: GKP, EOG, and EMG-based multimodal interface with application to humanoid robot control," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 2, pp. 453–462, Feb. 2014.

[34] Z. Xuhui, D. Runlin, and L. Yongwei, "VR-based remote control system for rescue detection robot in coal mine," in *Proc. 14th Int. Conf. Ubiquitous Robots Ambient Intell. (URAI)*, Jun. 2017, pp. 863–867.

[35] S. Ikegami, C. Premachandra, B. H. Sudantha, and S. Sumathipala, "A study on mobile robot control by hand gesture detection," in *Proc. 3rd Int. Conf. Inf. Technol. Res. (ICITR)*, Dec. 2018, pp. 1–5.

[36] H. Hu, C. Perez, H.-X. Sun, and M. Jagersand, "Performance of predictive display teleoperation under different delays with different degree of freedoms," in *Proc. Int. Conf. Inf. Syst. Artif. Intell. (ISAI)*, Jun. 2016, pp. 380–384.

WALKER AGUILAR received the B.S. degree in electronic engineering from the National University of San Agustin of Arequipa, Peru, in 2023. He is currently a Researcher in the Project "Design and Development of an EOD Robot with Multimodal Control Interface"-CONCYTEC. His expertise extends as a Hardware Developer for startups in Peru, offering innovative solutions. He designed and developed the multimodal interfaces and hardware used in this article.



LIZARDO PARI received the B.S. and M.S. degrees in electronic engineering from Universidad Nacional de San Agustín de Arequipa, in 1996 and 2005, respectively, and the Ph.D. degree in automation and robotics from the Polytechnic University of Madrid, Spain, in 2007. He is currently a Principal Professor and the Head of the Vision and Robotics Group, Universidad Nacional de San Agustín de Arequipa. His research interests include computer vision, motion tracking, three-dimensional reconstruction, and visual control. He was a member of the IEEE Robotics and Automation Society Peru Section.

YURI SILVA is currently a Mechanical Engineer with experience in aeronautical maintenance and supervision with the Vehicle Technical Inspection Centers, UNSA, and a member with the College of Engineers of Peru (CIP). He is a Co-Investigator with the Multidisciplinary Research Projects of the National Program of Scientific Research and Advanced Studies ProCiencia of CONCYTEC and UNSA-Investiga. He has published scientific articles in the area of robotics. He is a member of the IEEE Robotics and Automation Society.

ERASMO SULLA ESPINOZA received the B.S. and M.S. degrees in electronic engineering from Universidad Nacional de San Agustín de Arequipa, in 1994 and 2013, respectively. He is currently pursuing the Ph.D. degree in mechatronics engineering. He is a Principal Professor of biomedical engineering with Universidad Nacional de San Agustín de Arequipa.

LUIS F. CANAZA CCARI received the B.Sc. degree in electronic engineering from the National University of San Agustin of Arequipa, Peru, in 2022. From 2022 to 2023, he was a Junior Researcher in the Research Project funded by the National Council of Science, Technology and Technological Innovation (CONCYTEC), Peru. His research interests include advanced nonlinear control of unmanned aerial vehicles and robotic manipulators, and formation control of multi-agent systems and cooperative control.

RONALD PEÑA received the Diploma degree in information technology from Universidad Señor de Sipán de Lambayeque, Peru, and the National Police of Peru. He is currently a Systems Engineer with Universidad Señor de Sipán de Lambayeque and a Major with the National Police of Peru, with a focus on applied to fiber optic communications networks, satellite communications, public management and governance of information technology, and telecommunications project management at INICTEL-UNI, has participated in projects, such as the implementation of mobile communications centers of the Peruvian Police.

NICOLÁS O. MEDINA was born in Cusco, Peru, in 1998. He received the Engineering degree in electronic engineering from the National University of San Agustin of Arequipa (UNSA), in 2024. He is currently a Researcher in the Project "Design and Development of an EOD Robot with Multimodal Control Interface"-CONCYTEC. He designed and developed the multimodal interfaces and virtual environment in this article. His research interests include NUI and multimodal interfaces, unmanned vehicles, robotics, and biomedical engineering.