Analysis of the Intuitive Teleoperated System of the TxRob Multimodal Robot

Jeyson Carpio A.[*,1], Samuel Luque C.[*,2],
Juan Chambi C.[*,3], Jesús Talavera S.[*,4]
Universidad Nacional de San Agustín
de Arequipa, Perú[*]

***Abstract*—Natural disasters such as earthquakes, avalanches, landslides, among others, leave in their path people who may be trapped in the rubble, which are hardly found by rescue agents, so a reliable system in the operation of an exploration and rescue robot is essential. This paper aims to evaluate the systems proposed for the operation of the TxRob exploration robot. The teleoperated control systems that were developed for the manipulation of the robot are: a multimodal system feedback with information through different sensors, and a GUI control system using joystick buttons. These systems were analyzed using subjective metrics such as NASA-TLX, Scale Utility System (SUS) and Microsoft Reaction Cards, which provide interesting data when evaluating the performance of an interface, as well as the workload, user satisfaction and usability; these aspects are used to conclude which system is the most intuitive when performing rescue operations in case of a disaster, among others. 15 operators were evaluated to validate this system; the age range of the operators was between 20 and 43 years old and 20% of them had previously used VR headsets. Priority is given to the most immersive, easy to use and the most efficient system to perform the task of handling the robot.**

Keywords*—Multimodal interface; immersive teleoperation; exploration robot; gyroscope and subjective measurements

I. INTRODUCTION

In Peru there is a constant interaction between the Nazca and South American plates that generate natural disasters such as earthquakes, which put the population at risk of possible trapping due to landslides. According to the “Instituto Geofísico del Perú” (IGP), a total of 811 earthquakes were reported in 2020, 834 earthquakes in 2021 and approximately 581 earthquakes so far in 2022. The average of these earthquakes is of magnitude greater or equal to 4.5 ML (Richter scale), and also that the average number of earthquakes that occur per month is greater or equal to 58 earthquakes [1]; on average there are at least 2 earthquakes per day which is a constant risk to the Peruvian population. In the event that an earthquake has a greater magnitude, it would result in trapped people, uninhabitable houses or houses on the verge of collapse, at any time generating greater difficulty for the various rescue groups such as firefighters or members of the police who are also exposed to these dangers. Because of these problems, several robots are currently used for exploration and rescue, which reduce the exposure of the rescue agents, but do not reduce the mental load of the operators when making decisions [2].

For a successful operation there must be trust between the operator and the robotic partner, the cobots are designed with the purpose that the robot can support the various tasks of the operator, for them the human-robot interaction (HRI) is paramount. HRI covers the various fields such as design, understanding and evaluation of robotic systems, which involve humans and robots interacting through communication [3], [4]. The trust with a cobot can decrease drastically if it provides constant wrong information in risky situations, therefore a good robotic system that increases the trust with cobots is necessary [5].

For a robotic system in rescue operations to be classified as good, it should not miss proprioceptive sensors, nor should it miss good image processing, otherwise it would decrease the usability of search and rescue (SAR) robots and the performance of the operators by increasing the work performed using minimal information from the sensors. SAR robots should also have modular sensors and modular payloads increasing the usability of the robots and robot feedback can create more levels of competence [6]. The need for robotic interventions in hazardous environments is high, due to the presence of dust, fire, pressurized water or radioactivity; for that reason, robotic platforms must be reliable and user interfaces appropriate due to the complexity of the environment, for them the use of different inputs such as visual user interfaces help a better understanding of the environment for the operator [7], [8].

This paper presents an analysis of the user interface developed for the operation of the TxRob robot using subjective metrics such as NASA-TLX, SUS and Microsoft Reaction Cards [9]. It is believed that a developed bidirectional multimodal feedback system will help to decrease the operator’s workload, resulting in less stress and better teleoperation. The TxRob robot was presented by our team in a previous article [10]; this robot has the advantages of being low cost, its compact size allows it to enter into confined spaces for the search of possible people trapped in the rubble, it also has a sensor feedback system, in addition to a graphical interface that generates a greater immersion in his teleoperation.

The distribution of information in this document is divided as follows: In Section II the works related to this research are presented, the methodology is developed in Section III, the description of the proposed multimodal interface and the interfaces used for comparison are described in Section IV, the description of the experimentation used in this paper is presented in Section V, the results and discussions obtained are presented in Section VI, and the conclusions of this research are found in Section VII.

II. Related Work

According to [11], the combination of different tools such as stereo vision, haptic feedback and auditory feedback increases the manipulation performance of a robotic system. For this, the multiple information that we can provide to the operator would generate a more immersive system for the teleoperation of a robot. For that reason, the use of multimodal interfaces generates a higher performance in robotic teleoperation. In [12] they propose a method to organize the presentation of information and a set of visual assistance to facilitate the visual communication of data in teleoperation control panels, seeking to make them understandable and not generating a greater workload to the operator to understand the different information of the system where their interface combined immersive visualization, three-dimensional mixed reality and visual assistance.

With a gyroscope it is possible to know, maintain or change the orientation in space of an object; a gyroscope provides angular velocity readings usually in °/s on the x, y and z axes to determine the rotational speed of the object on which it is located; in addition, the use of the gyroscope in immersive virtual reality systems can be seen in [13] where making use of the gyroscope of a mobile device the manipulation of objects in a virtual environment is achieved, in [14] where to obtain continuous visual effects a gyroscope was incorporated, being very helpful for the detection of the direction of vision of a person in a virtual environment and in [15] where the gyroscope is used to estimate the posture of an operator to integrate the virtual environment with the real environment.

In [16], they use a haptic interface for robotic control in which immersive interaction with the person being rescued significantly improves task performance compared to other control systems; and visual feedback further increases user performance metrics. In [17], they evaluate different multimodal interfaces to find the system with the most intuitive and least operator workload; using visual interfaces, GUI interfaces and NUI interfaces. It is shown that a multimodal interface helps considerably in an intuitive teleoperation improving robotic efficiency. In [18] they present a new haptic mediator interface for teleoperative mobile robotic platforms that have a variety of manipulators and functions; employing identical bilateral master-slave teleoperation of robotic manipulators is achieved by representing them in virtual reality and allowing the operator to interact with them using a multipoint haptic device. As a result, the operator can perform a wide range of control functions and achieve functionality similar to that of conventional teleoperation schemes using a single haptic interface.

In [19], they measure the effects of working with autonomous robots on perceived workload and work performance, measuring both objective and subjective tests such as NASA-TLX, with the results of the study showing the viability of applying fully autonomous mobile robots to improve the productivity of the human-robot team while prioritizing physical safety and reasonable increases in mental workload.

In [20], they present an HMI which they developed to be able to control virtual forearm prostheses for extended periods of time. By calculating the physical performance for the given tasks, they develop an algorithm that adapts to the mental states of a user, thus improving its usability. According to the NASA-TLX evaluation, a better physical performance was improved.

III. Methodology

This document explains the two systems that will be analyzed for the manipulation of the TxRob robot. The control system for the movement of the robot is done by the joystick control which has functions that will be detailed later. For the vision system we present two interfaces:

- Traditional Interface: Uses a display screen that shows in real time the graphical interface developed for the TxRob.
- VR interface: Uses virtual reality (VR) glasses for visualization of the developed graphical interface and a gyroscope system for feedback, see Fig. 1.

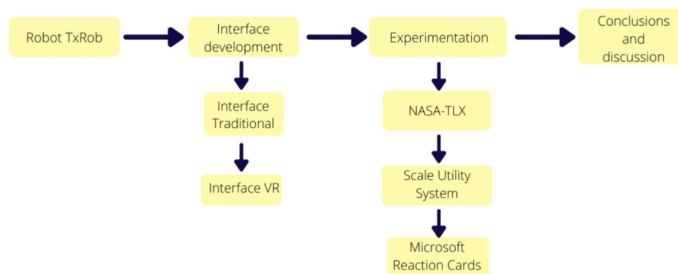

Fig. 1. Methodology Performed.

In addition, the evaluation procedure to be applied to the operators in the teleoperation of a simulated rescue is explained. The objective is to find an immersive and intuitive multimodal interface for the operator, understanding all the information obtained by the proximity and gas sensors, the images from the cameras and the microphone of the TxRob robot. To determine which of the proposed interfaces is better for the user, both interfaces are tested to 15 operators, which are evaluated by subjective metric tests that help to obtain the value of the workload generated by each interface in the immersive teleoperation; the usability as well as the utility of the system and the satisfaction of the operator when manipulating the proposed systems.

IV. Proposed Multimodal Interfaces

Fig. 2, shows the distribution of the two developed interfaces. The traditional interface is the HMI interface that performs the motion control of the TxRob robot by means of a joystick control and the vision is given by a graphical interface displayed on a screen. On the other hand, the VR interface performs the same motion control as the previous interface but the vision is provided by VR glasses and a feedback system generated by a gyroscope and a two-axis turret for the camera.

A. *TxRobt Control Interface*

The motion control of the TxRob was previously developed by means of a joystick control as shown in Fig. 3(a), in this work we have optimized the motion control by increasing the

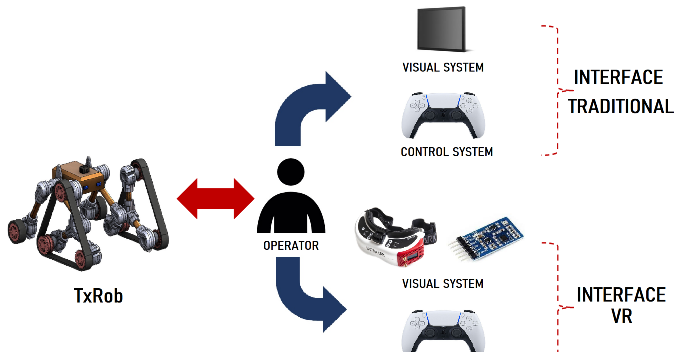

Fig. 2. Visuals and Control Interfaces.

control of a turret where the cameras are located. The turret was implemented in order to obtain the video in real time and to be able to manipulate it in two axes to cover the largest possible space of vision for the operator in the occlusion points without the need to move the TxRob. In this way, images of the upper part of the TxRob robot were obtained.

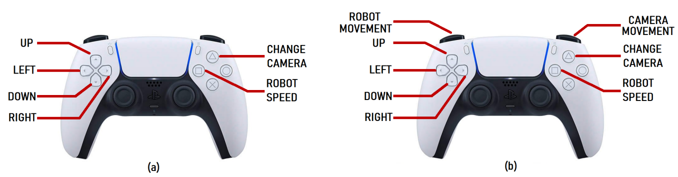

Fig. 3. Button Assignments. (a) First Configuration. (b) Last Configuration.

In the section (a), the change of cameras is controlled by pressing the triangle button; to choose the speed of movement of the robot the square button must be pressed and then by means of the two direction buttons (Up and Down) we regulate the speed of the TxRob. In section (b), to choose the robot movement control, press the L1 key (Robot Movement) and control by means of the four direction buttons (Up, Down, Left and Right); to choose the turret movement control, press the R1 key (Camera Movement) and control by means of the four direction buttons. The rest of the functions of the first configuration were maintained.

B. Traditional Teleoperation Multimodal Interface

In most teleoperated robot vision systems, the interfaces presented consist of cameras and a screen where the video can be observed in real time. In the previous work [10], a control system was realized by means of a screen, which shows the developed interface. This interface shows the images captured from the cameras in addition to the values sampled by the various sensors incorporated.

In Fig. 4, the previously presented interface is improved, in this new interface the collected images are shown on the left side together with the developed motion detection system. The upper right side shows the real-time measured values of the gas sensors (CO_2, O_2, NO_X) and the lower part shows an anti-shock system, which indicates the proximity of different objects around the TxRob.

Fig. 4. Improved Teleoperation Interface.

C. Proposed Multimodal Teleoperation Interface

This developed system works by means of specular imitation, that is to say that all the movement made by the operator in the X and Y axis detected by the gyroscope of the VR lenses, will be replicated in the camera turret. This turret consists of two servomotors, each servomotor manages to control the movement in each respective axis, see Fig. 5.

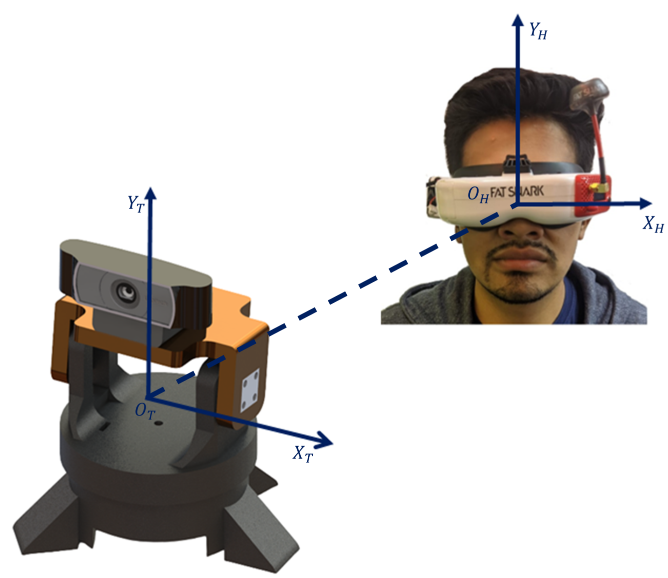

Fig. 5. Coordinates System.

When using VR lenses, the visual interface developed can become overloaded for the operator due to the FPV resolution that this system uses and at the same time the viewing space is reduced, our priority is to have an immersive and intuitive multimodal interface for the user. According to the literature reviewed, the overwhelming amount of data in teleoperation tends to confuse the user, decreasing the effectiveness of the task. For them, a more intuitive interface was developed for this system as shown in Fig. 6.

The interface proposed for the VR glasses is a first-person-view (FPV) interface, this format makes the proposed interface

Fig. 6. Display Shown to the User.

more immersive and the immersion is complete; to avoid saturating the operator's view we prioritized the real-time image of the cameras and in the corners we displayed the data from the different sensors of the robot that will help the teleoperation.

V. Experimentation

In order to confirm that the VR system with gyroscope feedback can achieve a high performance, it was evaluated by taking measurements in two stages: the first stage consists of the gyroscope attached to the VR device, the measurements recorded in the three axes for 20 seconds; the second stage consists of the measurement of the gyroscope attached to the camera, the movement of the motors are regulated with encoders after receiving the signal from the first gyroscope. In order to compare these two proposed multimodal systems, a debris entrapment rescue environment was simulated. A mirror environment was created for each multimodal interface. The proposed circuit is in the shape of an "S" for the traditional interface and for the proposed interface it is in the shape of an "inverse S" so that the experience is equal in both. This teleoperation test of the TxRob was performed on 15 operators, who ranged in age from 20 to 43 years old and 20% indicated that they had experience in handling VR systems. Each operator must manipulate the two multimodal systems to find the person trapped in the rubble, for this purpose they were given 5 minutes of tests so that they could familiarize themselves with the interfaces to be evaluated. At the time of the real tests, the time it takes to perform the search with each interface is measured, the time increases due to the presence of debris in the test scenario. At the end of the test, each operator performs the three subjective metric tests: NASA-TLX, SUS and Microsoft Reaction Cards; these data are used to obtain the values and scores of both multimodal systems.

VI. Results and Discussion

A. Gyroscope Measurements

Fig. 7 shows each of the gyroscope axes with the VR lenses on the operator; where the operator moved his head for 20 seconds, making a total of 500 samples taken for each axis of the gyroscope. These measurements correspond to the angular velocity of motion recorded during the experimental phase. Fig. 8 shows each of the gyro axes with the turret camera on the TxRob .Where the motors simultaneously triggered the movement of the camera according to the data recorded by the gyroscope attached to the VR device.

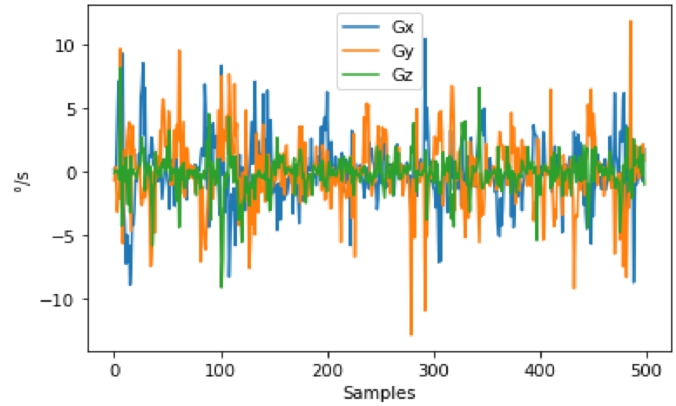

Fig. 7. 3-Axis Plot for the First Gyroscope.

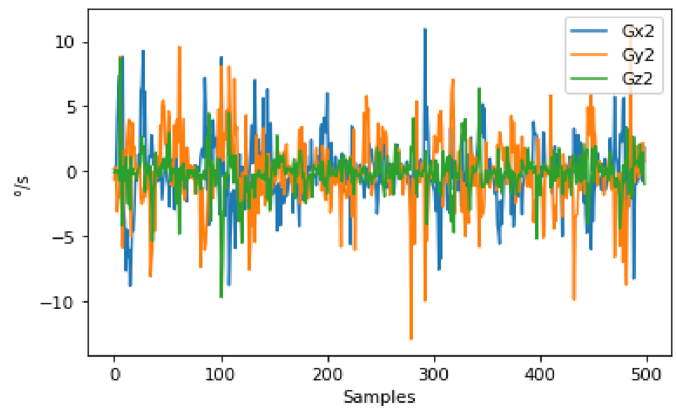

Fig. 8. 3-Axis Plot for the Second Gyroscope.

A comparison was also made between the measurements of both gyroscope divided on each axis defined as Gx, Gy and Gz. Data obtained from both the X-axis lens and turret camera gyroscopes are shown in Fig. 9, while Y-axis data are shown in Fig. 10 and Z-axis data are shown in Fig. 11. On average, a 4.13% variation is found due to the configuration of the motors driven by the encoders.

B. Subjective Measurements

The following values were obtained from Fig. 12, where an improvement in time per task performed with the new proposed multimodal interface is noticeable. We consider this improvement to be relevant because in human rescue operations, time is the most important element. As long as an operation is performed in the shortest time, the possibility of saving more people increases considerably.

The overall workload per operator evaluated was reduced when using the multimodal VR interface. The traditional interface showed an average of 60.66 of overall workload appreciated in performing the experiment as opposed to the VR interface which showed 50.56 of overall workload presented. It

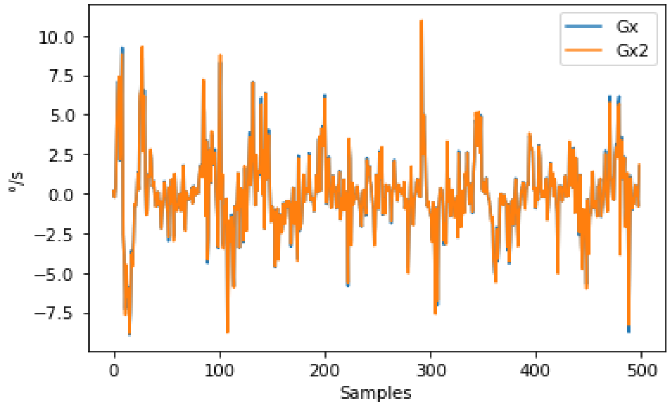

Fig. 9. Comparative Graph between Gx of Both Gyroscopes.

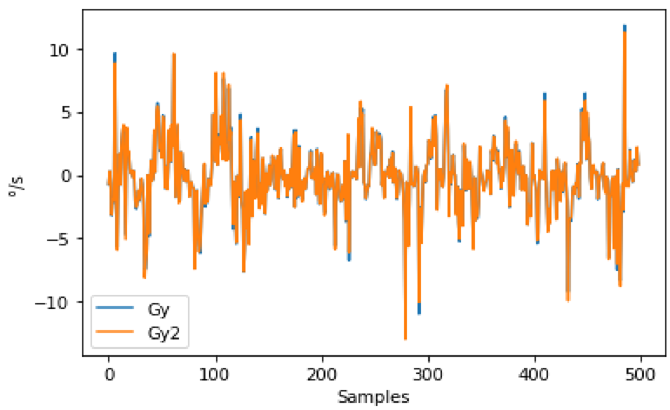

Fig. 10. Comparative Graph between Gy of Both Gyroscopes.

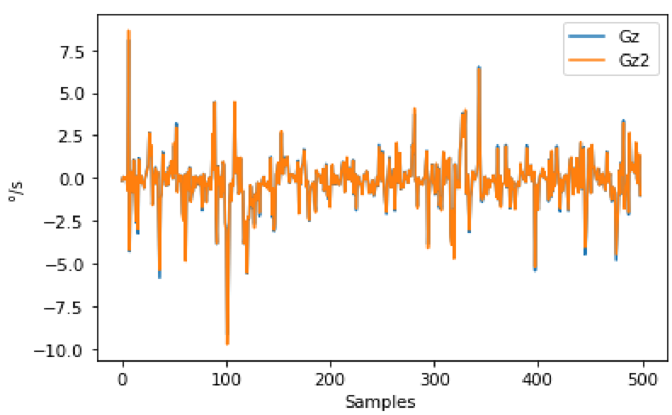

Fig. 11. Comparative Graph between Gz of Both Gyroscopes.

was also noted that the overall workload was higher for people with no experience in manipulating teleoperated systems and older people. This reduced workload helps make teleoperation more intuitive and easier for operators to manipulate. In Fig. 13 the blue bar is the first interface evaluated and the orange bar is the VR interface.

In Fig. 14 we present the dimensions evaluated by the NASA-TLX subjective method such as mental demand, physical demand, temporal demand, effort, performance and level of frustration that may occur in the task. We have taken the average of the 15 operators to evaluate each dimension independently. According to the results obtained, the physical demands are lower because both systems are teleoperated and do not require physical labor. The mental demand is higher because the systems require most of the senses of people to manipulate both interfaces. When comparing the results of both interfaces, a reduction is observed when using the VR interface, in some dimensions this difference is more noticeable, such as the effort and frustration presented by the operators.

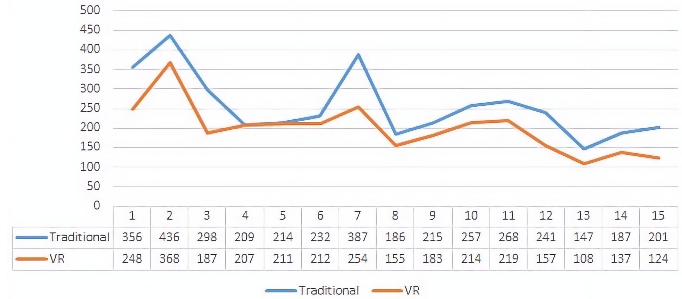

Fig. 12. Time Taken in the Corresponding Tests.

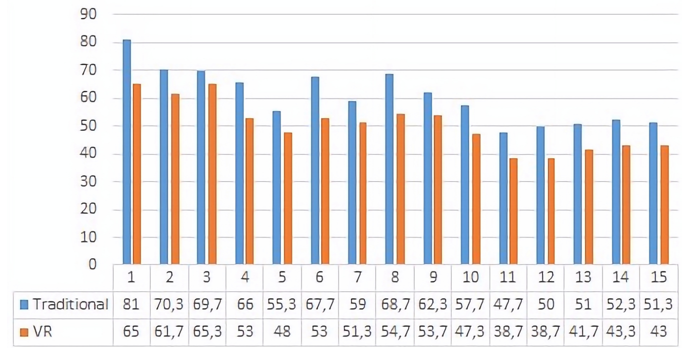

Fig. 13. General Workload - NASA TLX.

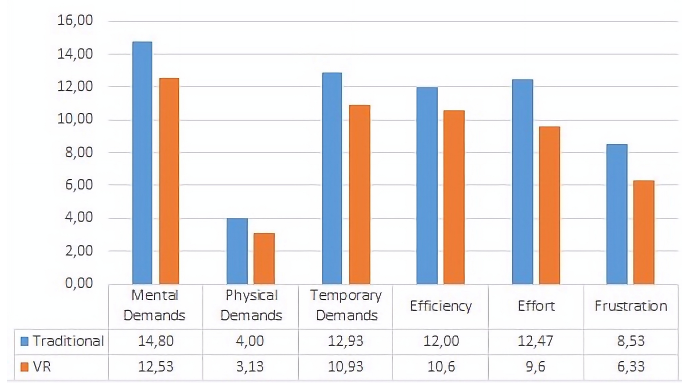

Fig. 14. Average Dimensions - NASA TLX.

The operators were also evaluated with the SUS metric, which will help us to see the usability of the systems, for the first interface an average of 85.33 was obtained, which indicates that the interface has a wide usability in the tasks of the experiments carried out. In addition, the second interface obtained an average of 95.66, a much more favorable result; in Fig. 15 you can see the scores obtained by each operator, it can be seen that most of them are above 80 points in the traditional interface and above 90 points in the VR interface.

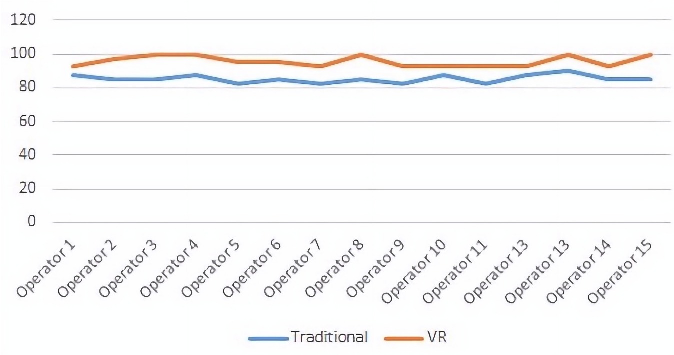

Fig. 15. SUS Results.

Microsoft's product reaction cards are, in essence, a list of 118 words that are used to describe a product or a completed good. The list includes positive words such as 'Useful', 'Comfortable' and 'Innovative', along with negative words, such as 'Stressful', 'Complex' and 'Dull'. The test consists of asking the robot operators to choose words from the list that they would use to describe the robot and, for each one, asking them why they chose that particular word. In these tests each of the 15 operators were limited to choosing 5 words that they felt best qualified the traditional system with joysticks to control the robot (light blue color), while on the other hand, in orange color are the results of the proposed system, see Fig. 16; only the cards with scores were considered, the rest of the words were ignored.

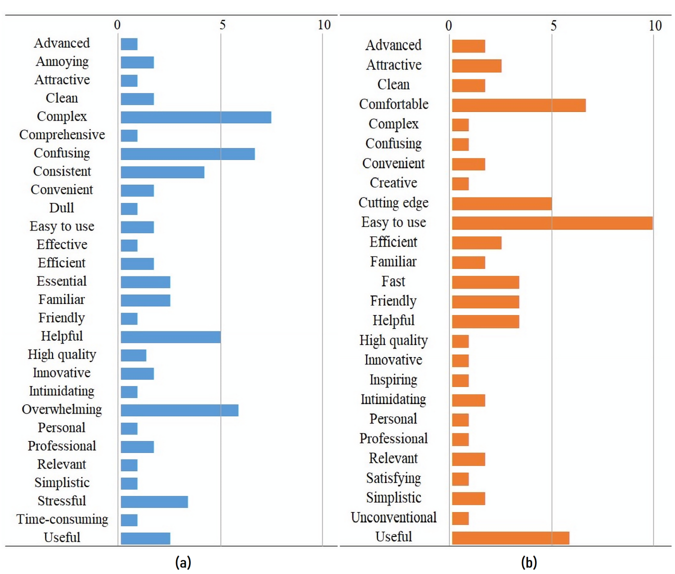

Fig. 16. Results of Microsoft Reaction Cards: (a) Traditional System. (b) VR System.

It can be seen in Fig. 16 that the words that were most chosen by the users were: "Complex", "Confusing", "Overwhelming", "Helpful" and "Consistent". Although some of the words are positive, it is clear that most of the operators feel that the traditional system with joysticks is not as comfortable to use, unlike the proposed system which in its absence most of the words selected were positive, of which the most chosen were: "Easy to use", "Comfortable", "Useful", "Cutting edge", "Fast", "Friendly" and "Helpful". This information shows that the proposed system is the one preferred by the operators who carried out the TxRob robot handling tests.

VII. Conclusion

In this paper an evaluation of the multimodal interface developed for the operation of the TxRob robot was performed. The teleoperated control systems that were developed for the manipulation of the robot are: multimodal system fed back information through the different sensors and a GUI control system using joystick buttons. These interfaces were analyzed using subjective metrics such as NASA-TLX, Scale Utility System (SUS) and Microsoft Reaction Cards, which provided us with data such as workload, experiences such as user satisfaction and usability of each system in order to designate the most intuitive when performing rescue operations in case of a disaster. The results of the NASA-TLX tests showed that the multimodal system with gyroscope feedback for camera control is the best because it greatly reduces stress generation when manipulating robots. In the SUS tests, the proposed system obtained an average score of 95 points, demonstrating that it is a better interface than the joystick control system, which only obtained an average score of 85 points. Finally, the Microsoft Reaction Cards tests showed that the majority of operators chose the proposed interface as more comfortable and easier to use, unlike the majority of words chosen to qualify the traditional joystick system.

Acknowledgment

The authors would like to thank Universidad Nacional de San Agustin de Arequipa.

References

[1] Geophysical Institute of Peru, 2022, Reported Earthquakes.

[2] R. R. Murphy, "Human-robot interaction in rescue robotics," in IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 34, no. 2, pp. 138-153, May 2004, doi: 10.1109/TSMCC.2004.826267.

[3] Garrett M. Clayton. Mechatronics for humanitarian explosive ordnance disposal in cambodia. ASME Mechanical Engineering, 140:S4–S10, 2018. http://dx.doi.org/10.14569/IJACSA.2022.01307104.

[4] Deborah R. Billings, Kristin E. Schaefer, Jessie Y.C. Chen, and Peter A. Hancock. 2012. Human-robot interaction: developing trust in robots. In Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction (HRI '12). Association for Computing Machinery, New York, NY, USA, 109–110. https://doi.org/10.1145/2157689.2157709.

[5] P. Robinette, A. M. Howard and A. R. Wagner, "Effect of Robot Performance on Human–Robot Trust in Time-Critical Situations," in IEEE Transactions on Human-Machine Systems, vol. 47, no. 4, pp. 425-436, Aug. 2017, doi: 10.1109/THMS.2017.2648849.

[6] J. Casper and R. R. Murphy, "Human-robot interactions during the robot-assisted urban search and rescue response at the World Trade Center," in IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 33, no. 3, pp. 367-385, June 2003, doi: 10.1109/TSMCB.2003.811794.

[7] Angulo, A.M.; Pinto, L.P.; Espinoza, E.S.; Vidal, Y.S.; Colquehuanca, E.S. Assisted Operation of a Robotic Arm Based on Stereo Vision for Positioning Near an Explosive Device. Robotics 2022, 11, 100. https://doi.org/10.3390/robotics11050100.

[8] J. Carff, M. Johnson, EM El-Sheikh y JE Pratt, "Human-robot team navigation in visually complex environments", Proc. Internacional IEEE/RSJ. Conf. Intel. Sistema de robots, pp. 3043-3050, octubre de 2009.

[9] Brooke, John. (1995). SUS: A quick and dirty usability scale. Usability Eval. Ind. 189.

[10] Rafael Verano M, Jose Caceres S, Abel Arenas H, Andres Montoya A, Joseph Guevara M, Jarelh Galdos B and Jesus Talavera S, "Desarrollo de un robot explorador teleoperado de bajo costo (TXRob)" International Journal of Advanced Computer Science and Applications (IJACSA), 13(7), 2022. http://dx.doi.org/10.14569/IJACSA.2022.01307104.

[11] E. Triantafyllidis, C. Mcgreavy, J. Gu and Z. Li, "Study of Multimodal Interfaces and the Improvements on Teleoperation," in IEEE Access, vol. 8, pp. 78213-78227, 2020, doi: 10.1109/ACCESS.2020.2990080.

[12] S. Livatino et al., "Intuitive Robot Teleoperation Through Multi-Sensor Informed Mixed Reality Visual Aids," in IEEE Access, vol. 9, pp. 25795-25808, 2021, doi: 10.1109/ACCESS.2021.3057808.

[13] Hudak, M., Korecko, S., Sobota, B., Balluch, T., & Grib, J. (2018). Walking Pad and Gyroscope-Based Object Manipulation for Virtual Reality CAVE. 2018 IEEE 18th International Symposium on Computational Intelligence and Informatics (CINTI). doi:10.1109/cinti.2018.8928223.

[14] Shanmin, C., Tao, N., & Ke, W. (2011). Motion Control of Virtual Human Based on Optical Motion Capture in Immersive Virtual Maintenance System. 2011 International Conference on Virtual Reality and Visualization. doi:10.1109/icvrv.2011.24.

[15] Fang, D., Xu, H., Yang, X., & Bian, M. (2019). An Augmented Reality-Based Method for Remote Collaborative Real-Time Assistance: from a System Perspective. Mobile Networks and Applications. doi:10.1007/s11036-019-01244-4.

[16] M. Aggravi, A. A. S. Elsherif, P. R. Giordano and C. Pacchierotti, "Haptic-Enabled Decentralized Control of a Heterogeneous Human-Robot Team for Search and Rescue in Partially-Known Environments," in IEEE Robotics and Automation Letters, vol. 6, no. 3, pp. 4843-4850, July 2021, doi: 10.1109/LRA.2021.3067859.

[17] Goyzueta, Denilson V., Joseph Guevara M., Andrés Montoya A., Erasmo Sulla E., Yuri Lester S., Pari L., and Elvis Supo C. 2022. "Analysis of a User Interface Based on Multimodal Interaction to Control a Robotic Arm for EOD Applications" Electronics 11, no. 11: 1690. https://doi.org/10.3390/electronics11111690

[18] Q. -Z. Ang, B. Horan and S. Nahavandi, "Multipoint Haptic Mediator Interface for Robotic Teleoperation," in IEEE Systems Journal, vol. 9, no. 1, pp. 86-97, March 2015, doi: 10.1109/JSYST.2013.2283955.

[19] Y. Chen, C. Yang, Y. Gu and B. Hu, "Influence of Mobile Robots on Human Safety Perception and System Productivity in Wholesale and Retail Trade Environments: A Pilot Study," in IEEE Transactions on Human-Machine Systems, vol. 52, no. 4, pp. 624-635, Aug. 2022, doi: 10.1109/THMS.2021.3134553.

[20] I. M. Rezazadeh, M. Firoozabadi, H. Hu and S. M. R. H. Golpayegani, "Co-Adaptive and Affective Human-Machine Interface for Improving Training Performances of Virtual Myoelectric Forearm Prosthesis," in IEEE Transactions on Affective Computing, vol. 3, no. 3, pp. 285-297, July-September 2012, doi: 10.1109/T-AFFC.2012.3.