A Cost-Effective Immersive Telexistence Platform for Generic Telemanipulation Tasks

Reem Al-Remaihi[1], Aisha Al-Raeesi[1], Reem Al-Kubaisi[1],
Mohammed Al-Sada[1,2(✉)], Tatsuo Nakajima[2], and Osama Halabi[1]

[1] Department of Computer Science and Engineering,
Qatar University, Doha, Qatar
alsada@dcl.cs.waseda.ac.jp, ohalabi@qu.edu.qa
[2] Department of Computer Science and Communications Engineering,
Waseda University, Tokyo, Japan
tatsuo@dcl.cs.waseda.ac.jp

Abstract. Robots are expected to become an essential part in tackling the global corona virus pandemic. However, acquiring robotic technology in economically-challenged nations is very difficult for a variety of factors. This work aims at developing a low-cost Telexistence systems that can be used for remote manipulation. We propose a system comprising components on the local site, which comprise a commonly available VR headset, a tracking system and a 3D-printed haptic exoskeleton. The remote site comprises a small robot arm. We discuss the implementation specifics of our approach. We evaluated our system by comparing usability and performance in two tasks while using the VR headset and a monitor. Users were able to complete the two tasks successfully with our system. Moreover, results show superiority of the VR headset, how-ever, the screen also show promising results, and thereby offer a cheaper deployment option.

Keywords: Telexistence · Teleoperation · VR · Haptics · HRI

1 Introduction and Motivation

Robots are expected to become powerful instruments in tackling global issues such as the challenges caused by the global Corona virus epidemic that severely overwhelmed the world and transformed our lives. The resulting lockdowns have tremendous impact on our work and daily lives. A critical challenge is the difficulty of people's presence in work environments due to emergency constraints, which has impact tasks and activities that require physical interaction like operation of industrial machinery, taking care of patients at hospitals or hands-on training contexts.

Previous works in telexistence and telepresence technologies [1–5] have long been investigated as approaches that utilize robots to accomplish physical tasks with high sense of presence. However, the complexity and cost of such systems have constrained their deployment within various industries, especially within developing countries that face challenges to access, develop and deploy such technology.

This work attempt to bridge the gap into developing a cost-effective teleoperation robotic system that is able to accomplish a variety of physical manipulation capabilities, while being easy to use and providing high sense of presence and agency in the remote locations. Accordingly, we explain the specifics of our implementation and follow with an evaluation of the system usability and the effect of the level of immersion, using an HMD and a desktop screen, on task accuracy in two physical manipulation tasks. The overall results are encouraging to pursue deeper evaluations. Lastly, we provide our conclusion and future research direction.

Our contribution is summarized as follows: 1) we design and implement a telepresence robotic system based on off-the-shelf components to ensure ease of accessibility and deployment in developing countries. 2) Evaluate the effect different levels of immersion on the accuracy achieving different tasks.

2   Related Works

Telexistence and telepresence has long been investigated in previous literatures. Telexistence is a term that refers to a group of technologies and approaches that focus on high sense of presence within remote environments [1, 2, 3, 4, 5]. Various previous works proposed robotic platforms that enabled locomotion and multimodal interactions within remote environments. Generally, these works focused on enabling high sense of presence, therefore, engaging multiple senses such as vision, olfactory and haptic to deliver high sense of agency and presence with the relayed remote environment. Haptics is indispensable for such systems as it increase the sense of presence and accuracy of doing tasks [6, 7]. It also enhances the memory retention that lead to better performance [8]. A telepresence presents a group of technologies to enable a person to feel as if they are in other locations [9, 10]. Research has thoroughly investigated a variety of methods for telepresence, with varied levels of immersion, interaction modalities and sense of presence with the remote environments. Telepresence robotics focuses on enabling users to remotely access and interact with a remote environment [10].

Despite the robustness of previous efforts, we believe that most existing robotic telepresence systems and telexistence are inaccessible to economically-challenged nations, whether in terms of cost, attainability of equipment or availability within such nations. In comparison to previous works, our approach focuses on bridging the gap in deployability and cost for telexistence systems. Therefore, we use off-the-shelf components that are commonly available, without much reliance on industrial or custom-made components. Second, we use cost-effective components to reduce the cost as much as possible. These two aspects ensures that our implementation can be duplicated based on easily available components, or similar alternative components. Likewise, our software infrastructure uses commonly available technologies, such as Unity3D [11], SteamVR [12], and network connectivity based on websockets, which are all commonly available and attainable for free (with varied licensing for commercial use).

3 System Design and Implementation

The system is divided into two sub systems, local site (controller site) and remote site (robot site), as shown in Fig. 1. Next sections will explain every sub system:

3.1 Local Environment

The local site is designed to enable high sense of presence, similar to telexistence systems. The local site comprises a 3D-printed haptic exoskeleton that can both sense user's finger locations and deliver haptic feedback (based on [13]), as shown in Fig. 2. The exoskeleton is controlled through a Pololu Mini Maestro controller [14]. The user can see the remote site by wearing virtual reality (VR) head mounted display (HMD) with trackers placed on the user's hands to track their hand movements (HTC Vive system).
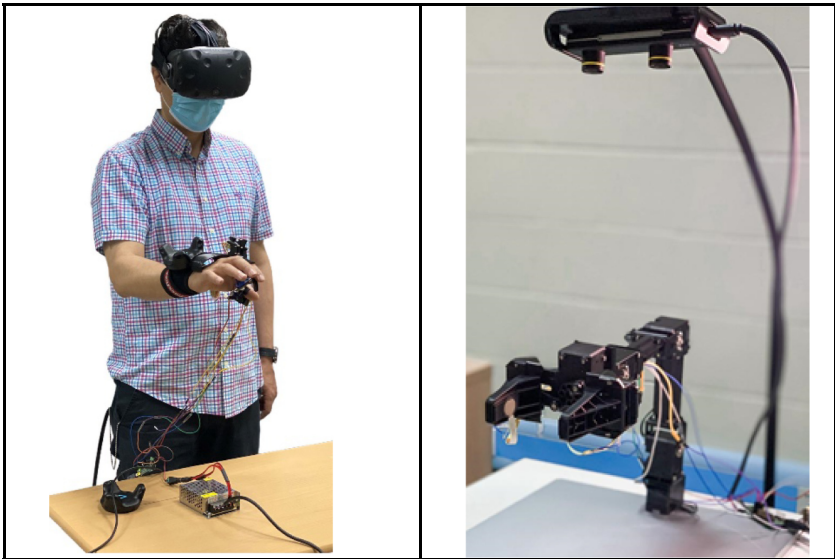
Fig. 1. Left) The exoskeleton haptic glove with attached tracker and HMD in Local environment. Right) The stereoscopic camera and the robot.

| Before holding an object	After holding an object
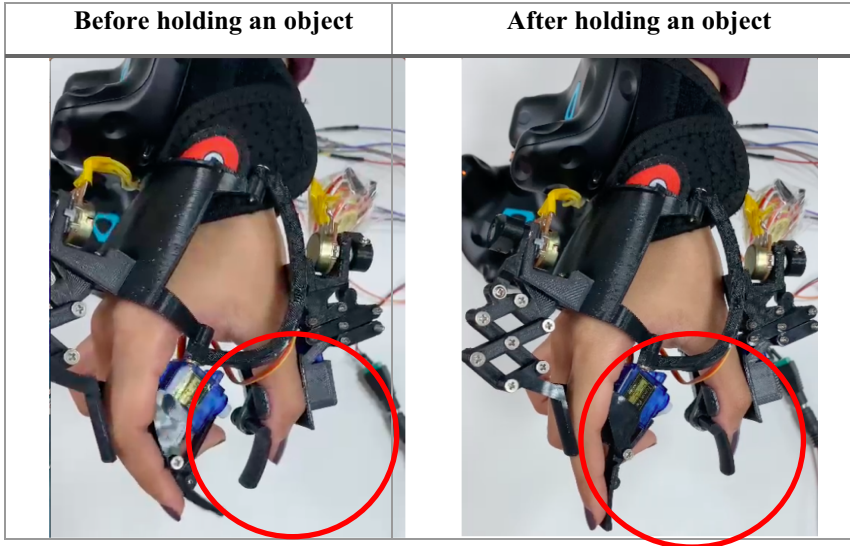
Fig. 2. This image shows the haptic exoskeleton system with the tracker, which is used to manipulate the robot. To the left, the user is attempting to hold an object by closing his fingers, when an object is detected, the robot provides haptic feedback by pressing against their fingers.

Overall, the local environment is created using Unity3D, which integrates the HMD, tracking system provides robot connectivity to the remote environment. In order to control the robot, we first created a 3D model of the robot arm and imported it to our unity3d project. Next, we utilized an inverse kinematic (IK) solver [15] setting the user's arm location as an objective for the robot's model to reach. Finally, we extracted the calculated joint angles produced by the IK solver and sent them to the remote environment. The remote environment finally receives the angles and executes them on the robot similar to previous works [16]. Figure 3 shows snapshots from the local and the remote environment for the robot.

3.2 The Remote Environment

The remote site comprises a stereoscopic camera that is positioned above the robot arm to provide live streaming to the local site, and a robot arm based on Robotis Manipulator X with a gripper end-effector equipped with force-sensitive sensors (FSR) to detect touch forces, see Fig. 4. The client-server architecture was used to enable controlling the robot in remote locations. Overall, the control system of the remote environment is deployed on a PC, while the FSR's data is captured through a Pololu Mini Maestro controller, which is also connected to the PC to forward the data to the local environment. Lastly, we used a ZED Mini Stereoscopic Camera to transmit the stereoscopic feed through WebRTC [17].
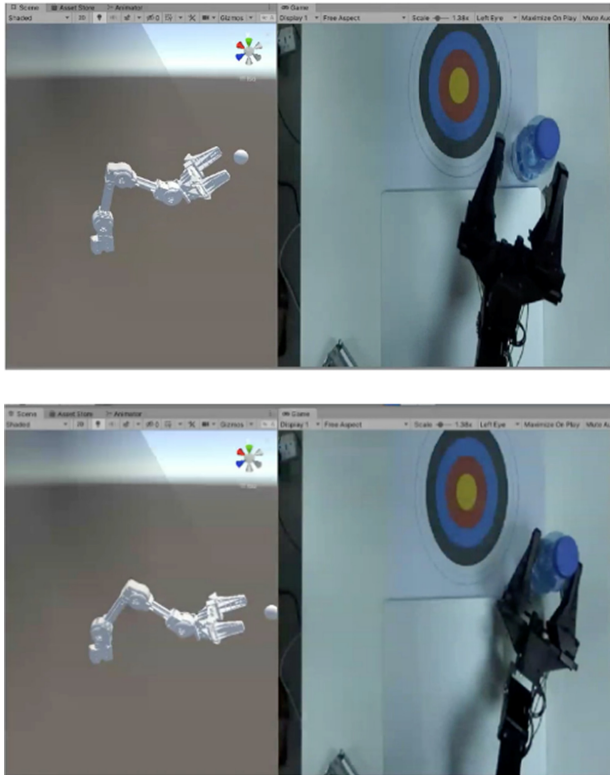
Fig. 3. Screenshots from the local and the remote environment. The left side of each picture shows the Unity3D scene with the imported robot model and IK solver. The right side of the image shows the camera feed from the remote environment with the robot attempting to follow the user's movement in the local environment.
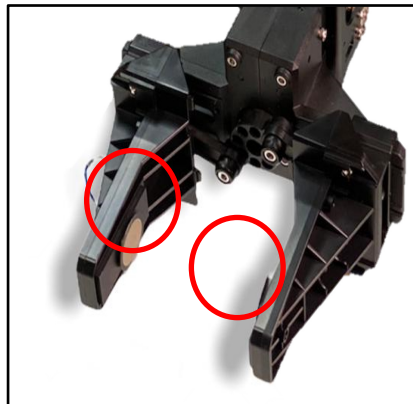
Fig. 4. An FSR is attached on each side of the end-effector to provide the ability to sense the applied pressure on the manipulated objects. The captured pressure data is sent to the local environment where they are conveyed as haptic feedback to the user.
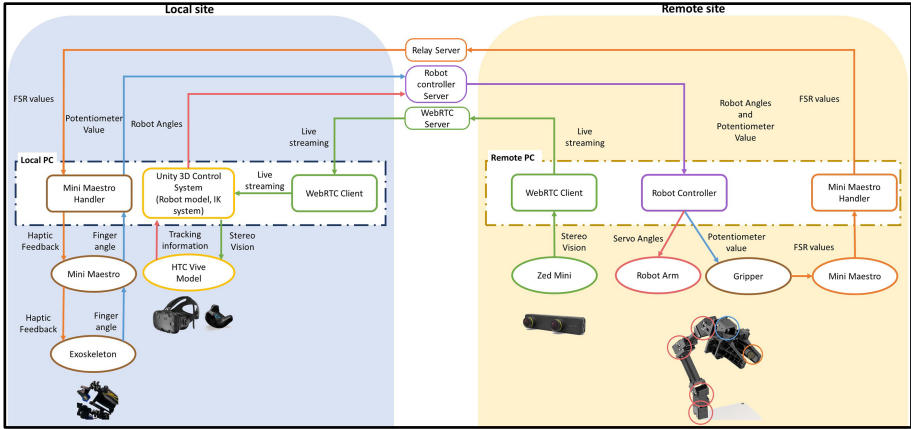
Fig. 5. The overall architecture of the system. (Color figure online)

The overall architecture of the system is explained in Fig. 5. There are four main modules that comprise data flow within the system (color coded in the diagram for clarity):

Haptic Feedback: The Mini Maestro in the remote PC will read the values from the FSR sensors that are attached on the gripper and send it to the Relay Server using websockets. Since the PC at the local environment is connected to the relay server, it would receive the FSR data, and map such readings into servomotor angles that control the feedback magnitude on the haptic exoskeleton.

Robotic Gripper Control: The Mini Maestro connected to the PC at the local environment will read the values from the potentiometer sensors that are attached on the exoskeleton and send them to the Robot Controller Server using websockets. Accordingly, the remote site's system will map the potentiometer readings into one angle that controls the gripper motor actions (open and close) using the robotic control system used in our previous work [16].

Robot Arm Control: The 3D model Unity3d is implemented and the IK solver to find out the robot's angles for each posture is used. Also, the tracker's location on the user's hand is used to determine the target location the robot's model should move to. Accordingly, the calculated robot angles are these angles are transferred through websockets to the remote site's system. The remote site's system will read the angles and control the robot using the received data.

Stereoscopic Video Streaming: The ZED mini camera will provide the system a live stereoscopic streaming, integrated within unity using webRTC server. The stereoscopic image is processed and shown to the user in the local environment.

The system is generally comprised of both off-the shelf components or 3D printed ones. The local environment utilizes consumer-level and 3D-printed components that can easily be obtained and are cost-effective. The chosen robot arm is also capable of lifting up to 500 g, with an effective workspace of 450 mm, and its end-effector exchangeable to meet other application domains (e.g., 5-finger hand). Therefore, we believe our design and implementation is easy to follow and customizable to match different applications. Moreover, with the falling costs of HMDs and actuators in the market, we believe a system based on the proposed architecture would be very cost-effective yet efficient for a variety of tasks, serving both industrial, hobbyist or research applications.

4   Evaluation

4.1   Objectives and Design

The main objective is to investigate the overall usability of the system as well as explore the effect of using different immersive displays on the task accuracy. In the experimental setup, HMD used as very high immersive display and a desktop display as a lower immersive level as can be seen in Fig. 6. Two types of manipulation tasks were considered in the study: 1) Lift and place an object (a plastic bottle of water) from one location and placing it to different locations on the table (T1), which is a common task within object manipulation contexts [16], 2) holding an instrument and pointing it towards a specific target (T2), which resembles taking a Corona virus swab. Both tasks are shown in Fig. 7 and Fig. 8.
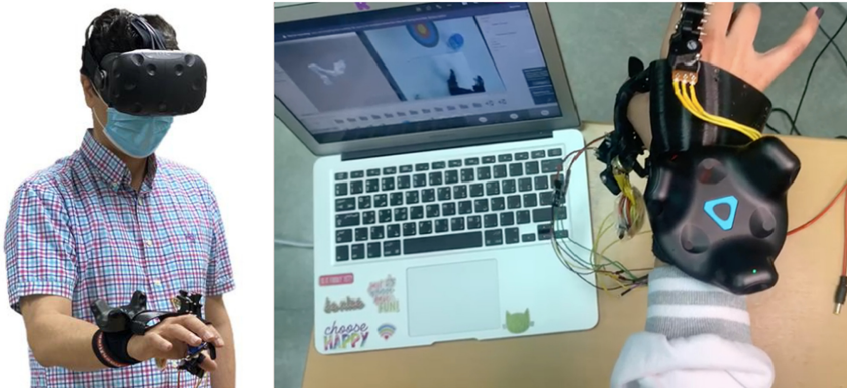

Fig. 6. The conditions of the experiment. We compared the user's accuracy and usability of our system on the HMD (Left) and a monitor (Right).

The accuracy was measured in terms of percentage of ± 5 mm range in both tasks. As shown in Fig 7, if the user places the bottle in the middle of the target (yellow circle), they get 100, and with each measured 5 mm from the center, the user loses 5 points.

We used the same calculation method for both T1 and T2. Moreover, we set some more rules to calculate the percentages, for example, if the bottle/poking device falls from the robot during operation, the accuracy will be counted as 0%. Also, users were not allowed to modify the position of the bottle or poke the target after their initial touch or poke. In addition to measuring the task accuracies, we also measured the time to complete each trial.

In addition to demographic data, we created a questionnaire to measure user's impressions regarding the two tasks under the HMD and monitor conditions. We also evaluated our system's usability using the SUS questionnaire [18, 19].
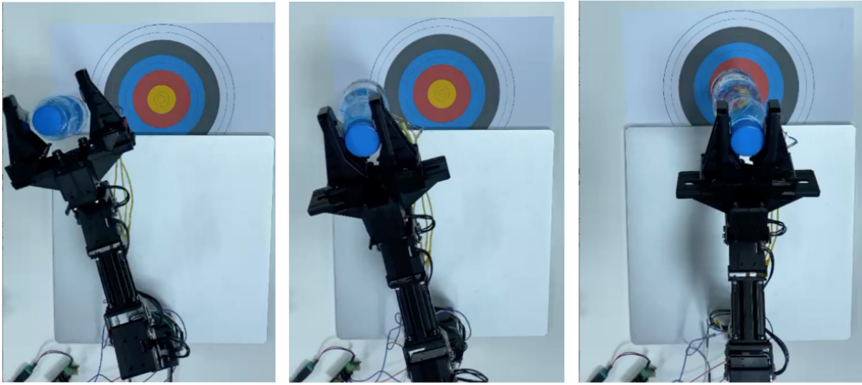
Fig. 7. Task 1 (T1): users had to pick up a bottle and place it at the middle of the target. This task was executed based on three conditions depending on the starting location of the bottle, which are in-front of the target location, to its left, or to its right. (Color figure online)
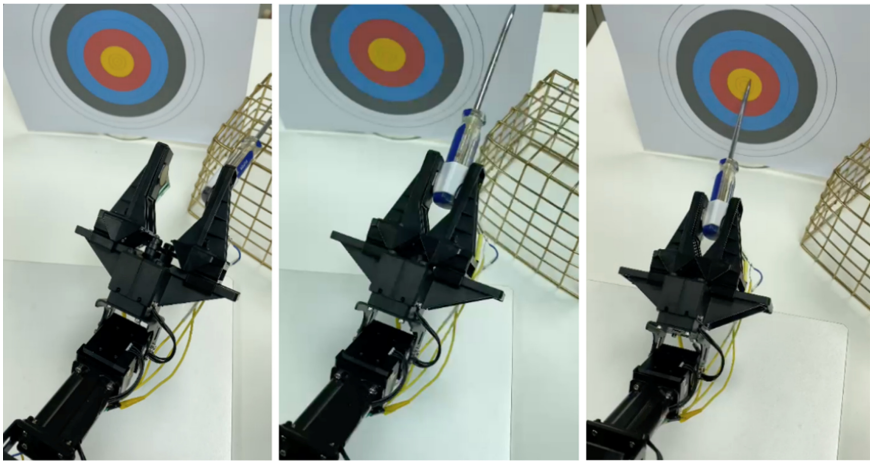
Fig. 8. Task 2 (T2) resembles a swab test. Participants had to pick up the screw-driver and poke the target in the middle. T2 was also executed in two basic conditions that alternated the locations of the screw-driver, placing it to the left or right of the target mark.

4.2   Participants and Procedure

Participants: We recruited 10 participants (all females, age m = 23.1, std = 2.96), who were students from various disciplines. Six participants indicated that they were familiar with VR, while the rest were not.

Flow: After a brief familiarization session, each participants took a demographic questionnaire, followed by the user study conditions. Participants did task 1, then task 2, using the HMD and followed by the monitor conditions, where each trial was repeated 3 times similar to previous works [20]. Overall, each participants undergone 10 trials (6 for T1, 4 trials for T2), where the experiment took approximately 60 min per participant. Upon finishing the trials, participants took a questionnaire that gauged their overall impression of the system and tasks, in addition to the SUS questionnaire.

4.3   Results and Analysis

Task Accuracy: The accuracy of manipulating the two tasks was higher when using HMD display than standard monitor as shown in Fig. 9. This demonstrates the superiority of immersive stereoscopic view in telexistence systems.

As shown in Table 1, the overall accuracy and time needed to accomplish both T1 and T2 was better in the HMD than the monitor. In T1, the average accuracy was slightly higher in the HMD than the monitor. However, in T2, the accuracy was remarkably higher in the HMD. We believe such results are in-line with previous telexistence results that showed the superiority of stereoscopic vision during physical manipulation tasks as users are able to perceive the depth of objects and thereby control the robot arms efficiently.
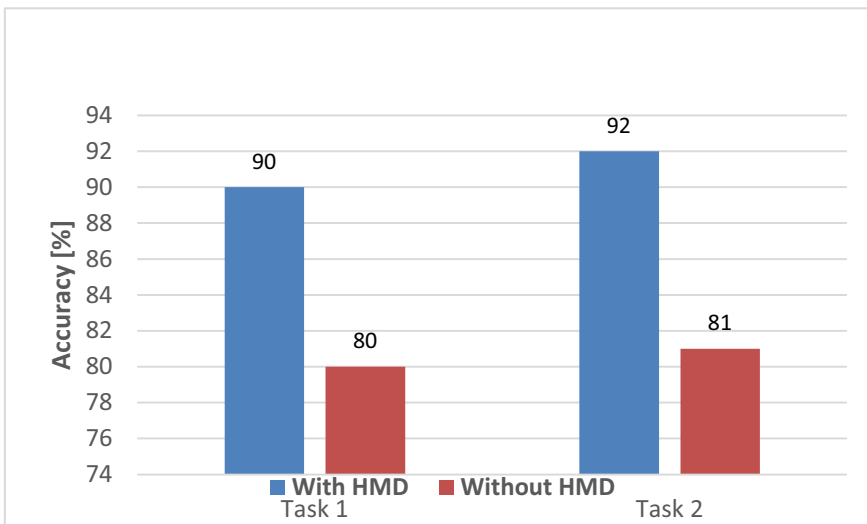

Fig. 9. The accuracy of achieving tasks for different levels of immersion.

Table 1. Average time on task and accuracies for each condition, with standard deviation values between brackets.

Tasks	Average time on task seconds (SD)	Average accuracy % (SD)
T1 (Monitor)	161 (44.8)	87.66 (8.89)
T1 (HMD)	145 (28.4)	88.83 (7.62)
Task2 (Monitor)	236 (18.2)	80.5 (8.87)
Task2 (HMD)	202 (35.3)	91 (8.52)

User Impressions: The participants were asked whether they preferred the HMD or the screen to accomplish the tasks. Seven participants preferred the HMD, mentioning that the sense of depth, better dimensions and realism of the remote objects seen through HMD and stereoscopic feed made them accomplish the tasks better. Participants who preferred the monitor mentioned aspects of dizziness and clarity of vision since these participants could not use their glasses with the HMD. When tasked to rate the difficulty of both tasks (5 means very difficult), participants rated T1 and T2 in terms of difficulty with 1.20 (std = 0.42) and 2.80 (std = 0.91), respectively. These results indicate that the participants generally though that the tasks were generally easy to accomplish.

SUS Questionnaire Results: Participants liked the system and found it very useful and usable. High scores in the questions related likeness, well integrated, easy to use, and confidence to use the system with scores over 80%. However, the lowest scores were in the questions related to complexity, inconsistency with scores less than 40% (Fig. 10). Nevertheless, final SUS score was 73.5, which is considered good as the average score for SUS to be considered usable is 68.
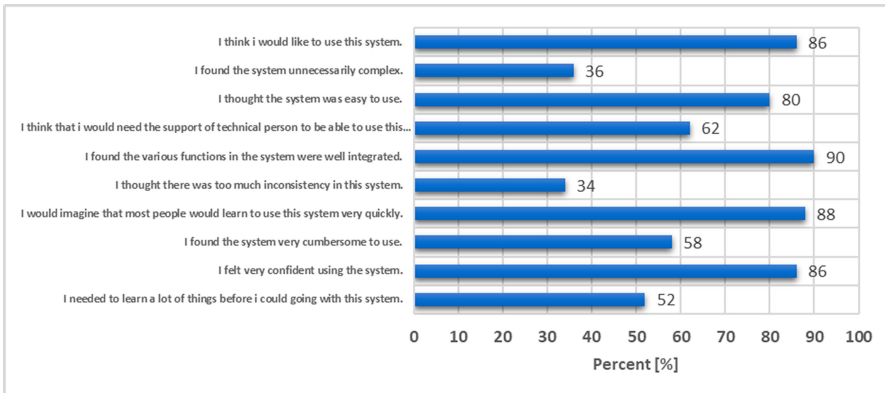
Fig. 10. The SUS questionnaire results.

Overall, we believe the results are encouraging to pursue further work. Users were able to accomplish two fundamental physical manipulation tasks remotely, therefore, we believe the control system was generally successful. Although the HMD proved to be superior to the monitor, we believe that using the monitor is a viable option and users had generally good results. Therefore, we believe that the results show the flexibility of deploying similar systems without the use of HMDs, which can contribute to reducing the overall cost.

5   Conclusion

This paper presents the design, implementation, and evaluation of a low-cost telexistence robotic system. We believe the presented architecture is cost-effective yet highly capable in terms of task accuracy and usability. The evaluation results also revealed the superiority of the HMD and stereoscopic vision for the evaluated physical manipulation tasks. However, replacing the HMD with a screen is also a viable option, and it produced acceptable results. Therefore, we believe that both deployment options can accomplish a variety of tasks, and future work should further evaluate the advantages of using a monitor for telemanipulation tasks.

A critical direction to research is to explore potential applications of telepresence robotics within daily usage contexts [16]. Moreover, as the context of use in economically challenged nations could be different than developed nations, we believe that focus groups and brainstorming workshops should be held within target countries. These workshops would enable us to capture design requirements or application domains that we have not initially thought about [21]. Accordingly, such data can be used as bases to advance our architecture to fit specific deployment domains.

We hope that this work will inspire further research that contributes to advancing of cost-effective technologies and exploring potential deployment domains that addresses the daily life constraints imposed by the corona pandemic.

References

1. Tachi, S.: Telexistence: enabling humans to be virtually ubiquitous. IEEE Comput. Graph. Appl. **36**, 8–14 (2016). https://doi.org/10.1109/MCG.2016.6
2. Tachi, S.: Telexistence. J. Robot. Soc. Japan (2015). https://doi.org/10.7210/jrsj.33.215
3. Alhalabi, M.O.O., Horiguchi, S.: Haptic cooperative virtual workspace: architecture and evaluation. Virtual Reality **5**, 160–168 (2000). https://doi.org/10.1007/BF01409421
4. Inoue, Y., Kato, F., Tachi, S.: Master-slave robot hand control method based on congruence of vectors for telexistence hand manipulation. In: 2019 22nd 2019 IEEE International Symposium on Measurement and Control in Robotics: Robotics for the Benefit of Humanity (ISMCR), ISMCR 2019, pp. 1–4 (2019). https://doi.org/10.1109/ISMCR47492.2019.8955688

5. Fernando, C.L., Furukawa, M., Minamizawa, K., Tachi, S.: Experiencing ones own hand in telexistence manipulation with a 15 DOF anthropomorphic robot hand and a flexible master glove. In: 2013 23rd International Conference on Artificial Reality and Telexistence, pp. 20–27 (2014). https://doi.org/10.1109/icat.2013.6728901

6. Halabi, O., Halwani, Y.: Design and implementation of haptic virtual fixtures for preoperative surgical planning. Displays **54**, 9–19 (2018). https://doi.org/10.1016/j.displa.2018.07.004

7. Halabi, O., Bahameish, M.A., Al-Naimi, L.T., Al-Kaabi, A.K.: Response times for auditory and vibrotactile directional cues in different immersive displays. Int. J. Hum.-Comput. Interact. **35**, 1578–1585 (2019). https://doi.org/10.1080/10447318.2018.1555743

8. Halabi, O., Al-Mesaifri, F., Al-Ansari, M., Al-Shaabi, R., Miyata, K.: Incorporating olfactory into a multi-modal surgical simulation. IEICE Trans. Inf. Syst. E97.D, 2048–2052 (2014). https://doi.org/10.1587/transinf.E97.D.2048

9. Björnfot, P., Kaptelinin, V.: Probing the design space of a telepresence robot gesture arm with low fidelity prototypes. In: ACM/IEEE International Conference on Human-Robot Interaction, Part F1271, pp. 352–360 (2017). https://doi.org/10.1145/2909824.3020223

10. Yamen Saraiji, M., Sugimoto, S., Lasantha Fernando, C., Minamizawa, K., Tachi, S.: Layered telepresence: simultaneous multi presence experience using eye gaze based perceptual awareness blending. In: Proceedings of ACM SIGGRAPH 2016 Emerging Technologies (2016)

11. Unity3D: Unity Game Engine. https://unity.com

12. Steam: Stream VR. https://store.steampowered.com/app/250820/SteamVR/

13. youbionic: youbionic. https://www.youbionic.com/handexohand

14. Pololu Corporation: Pololu. https://www.pololu.com/product/1354

15. Starke, S., Hendrich, N., Zhang, J.: Memetic evolution for generic full-body inverse kinematics in robotics and animation. IEEE Trans. Evol. Comput. 23 (2019). https://doi.org/10.1109/TEVC.2018.2867601

16. Al-Sada, M., Höglund, T., Khamis, M., Urbani, J., Nakajima, T.: Orochi: investigating requirements and expectations for multipurpose daily used supernumerary robotic limbs. In: Proceedings of the 10th Augmented Human International Conference 2019 on - AH2019, pp. 1–9. ACM Press, New York (2019)

17. StereoLabs: ZED mini. https://www.stereolabs.com/zed-mini/

18. Brooke, J.: SUS: a "quick and dirty" usability scale. In: Usability Evaluation in Industry, pp. 207–212. CRC Press (1996)

19. Brooke, J.: SUS : a retrospective. J. Usability Stud. **8**, 29–40 (2020)

20. Al-Sada, M., Jiang, K., Ranade, S., Kalkattawi, M., Nakajima, T.: HapticSnakes: multi-haptic feedback wearable robots for immersive virtual reality. Virtual Reality **24**(2), 191–209 (2019). https://doi.org/10.1007/s10055-019-00404-x

21. Al Sada, M., Khamis, M., Kato, A., Sugano, S., Nakajima, T., Alt, F.: Challenges and opportunities of supernumerary robotic limbs. In: CHI 2017 Workshop on Amplification and Augmentation of Human Perception (Amplify 2017), Denver, CO, USA (2017)