Robot Teleoperation through Virtual Reality Interfaces: Comparing Effects of Fixed and Moving Cameras

Mehdi Sobhani
mehdi.sobhani@brl.ac.uk
Bristol Robotics Laboratory
University of the West of England
Bristol, UK

Manuel Giuliani
manuel.giuliani@brl.ac.uk
Bristol Robotics Laboratory
University of the West of England
Bristol, UK

Alex Smith
alex.smith@brl.ac.uk
Bristol Robotics Laboratory
University of the West of England
Bristol, UK

Anthony G. Pipe
tony.pipe@brl.ac.uk
Bristol Robotics Laboratory
University of the West of England
Bristol, UK

ABSTRACT

Robot teleoperation is a traditional form of human-robot interaction. One of the challenges in this type of interaction is for the robot operator to have appropriate situational awareness of the robot's surroundings. Streaming videos from cameras to screen displays has been the main approach so far for understanding the remote environment and controlling the robot. In this work, a virtual reality interface for controlling a remote robot is analysed through a user-study with 40 participants. The task of the participants in the study is to teleoperate a robot to pick and place five barrels on predefined target positions. The aim of the study is to analyse and compare the effectiveness of three different camera setups, namely: one-fixed camera, two-fixed cameras with the ability for the user to switch between the two, and a moving camera. The setups are compared subjectively using NASA Task Load Index and System Usability Scale. As an objective measure, the participants' performance is also measured based on their precision and task completion time. The study results suggest that using a virtual reality interface with a moving stereo camera can significantly improve teleoperation precision and completion time. Using a moving camera also significantly decreased the workload of the participants. The results also show that the participants were able to use the interface with minimum training, suggesting this interface is intuitive for operators.

CCS CONCEPTS

• **Computer systems organization → Embedded systems**; *Redundancy*; Robotics; • **Networks** → Network reliability.

KEYWORDS

Teleoperation, Interface, Virtual Reality, Human-Robot Interaction

ACM Reference Format:
Mehdi Sobhani, Manuel Giuliani, Alex Smith, and Anthony G. Pipe. 2020. Robot Teleoperation through Virtual Reality Interfaces: Comparing Effects of Fixed and Moving Cameras. In *Proceedings of VAM-HRI 2020*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Remote robot teleoperation is one of the first forms of human-robot interaction. Traditionally, an operator in the control room uses a joystick to control the robot and uses streaming videos to screen displays from cameras, either mounted on the robot or at a different point in the remote environment to gain situational awareness. Such a setup has been used in many sensitive applications such as search and rescue missions and remote inspection tasks. Any mistake of the operator can cause serious damage to the robot or destruction in the environment. Hence, it is very important for the operator to have the highest possible situational awareness.

In our research, the main application is decommissioning of old nuclear facilities. According to the World Nuclear Association, there are over 115 commercial reactors, 48 experimental reactors and 250 research reactors that are no longer functional and need to be dismantled and decommissioned [19]. In the UK alone, "the 2019 forecast is that future clean-up across the UK will cost around £124 billion spread across the next 120 years or so" [17]. Using robots for this application will increase efficiency as well as the safety of human operators [14]. However, the structure of a typical nuclear facility imposes the constraint of having minimal visibility of the robot's environment, as there are thick walls between the operators and the robot. Therefore, gaining appropriate situation awareness is a challenge for the operators. Using a multiple camera system and switching between cameras to monitor the environment from different viewpoints could be a way forward. Such systems have been successfully used in other applications; however, changing viewpoints will increase cognitive load and slows the process [15].

As an alternative to multiple fixed camera systems, researchers have investigated novel systems with moving cameras for teleoperation. As such, a camera is mounted on a separate robotic arm and this arm movement, either autonomously controlled [15, 20] or manually controlled by a human operator [4], provides the best

---

VAM-HRI 2020, March 23, 2020, Cambridge, UK

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

viewpoint on a screen display. Such novel interfaces have significantly improved teleoperation performance [15].

Augmented Reality (AR) has also been used for robot control. So far, most interfaces using AR were mostly screen-based or tablet-based. For instance, "TouchMe" was developed to remotely control a mobile robot equipped with a manipulator, through a touchscreen [9]. This interface was tested in an experimental setup for a remote pick-and-place task in which participants, controlling the robot, had not seen the setup of the remote room in which the robot was operating. As a result, among 12 participants all but one could complete the task without training. In another research work, an AR interface was developed to control and program an industrial ABB IRB140 robot arm [3]. The interface shows the robot view on a computer screen and when the operator gives a command using keyboard and mouse, simulated actions of the robot for completing the task is displayed on the screen so the operator can observe the trajectory, motion and the result of the action. The operator can then either confirm or cancel the command. This interface was tested for stacking and sorting of small block tasks in which inexperienced operators completed the tasks with no significant time difference from professional operators. This system could be safely deployed in our application; however, a clear downside of such an interface could be increased time of the operation, as the operator first needs to evaluate the simulated action before performing the actual one. In addition, while these studies reported the number of people who successfully completed the task, the performances of the participants in terms of their precision in manipulating objects is not reported. Most recently, Lin et al. [12] developed an AR interface to remotely control a Kinova JACO Assistive Robotic Arm using hand motion and gestures. Their teleoperation interface is designed so that the operator can handle objects virtually and the gesture, created as a result of virtual object manipulation, is used as a command for manipulating the real object or executing a specific behaviour. This provides a more intuitive grasping approach compared to using a joystick to control a robot arm. However, with a very small workspace this setup is not suitable for large scale industrial applications like nuclear decommissioning.

Virtual Reality (VR) has also been investigated as an immersive medium to tele-control robots. While many AR cases are screen-based, VR interfaces are implemented using Head-Mounted Displays (HMDs). For example, researchers at MIT have developed Baxter's Homunculus, an interface for remote controlling a Baxter dual-arm robot over a long distance [13]. The interface uses a commercial off-the-shelf game engine, an Oculus Rift VR headset and touch controllers. The operator does not work directly with the robot in the virtual environment, as the environment is designed to be only like a virtual control room inside the robot similar to a cockpit of a plane. Baxter's Homunculus was reported to outperform automated and in-hand object localisation systems in grasping, pick-and-place and assembly tasks. Another similar research work was done at the Brown University Humans to Robot Laboratory using a Baxter dual-arm robot and HTC Vive HMD [18]. In this work, the operator was allowed to switch between two viewpoints: either the robot view or a third-person view, which is similar to our two camera experiment condition. This system was tested and compared to conventional interfaces like a joystick, or keyboard and mouse for a cup stacking task. The result of this comparison was a 16% increase in accuracy with a 62% decrease in the time to complete the task, compared to the next best system using conventional interfaces. Closely related to our approach, these works are good examples of VR application for teleoperation, provided that the interface is used with a more precise robotic arm.

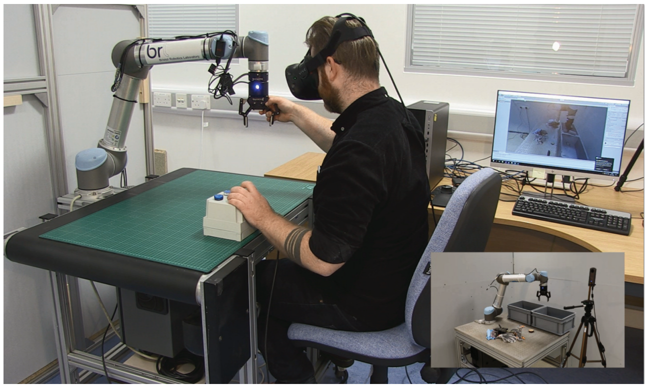

Figure 1: Teleoperation using 360°camera video stream on HMD.

There are also many more examples of VR application in robot control like Toyota's third-generation T-HR3 humanoid, which is also being controlled through VR using a HTC Vive and an exoskeleton interface, to provide a full-body intuitive control so that the robot can mirror the operator's movements [2]. In the field of medical robotics, aiming to provide surgeons with a third hand controlled by their foot during laparoscopic surgeries, researchers at EPFL introduced a third hand along with two others (representing the surgeon's "own" hands) in a series of VR experiments to analyse if people can coordinate these three virtual hands to complete a task [1]. Further examples are: VR interfaces being investigated for teleoperation and robot compliance control using a haptic feedback device [10], wheeled mobile robot for remote real-time assembly [11], and for inspection and maintenance processes in the aviation industry [7].

In our own previous work [16], we used a Ricoh Theta V 360° camera placed in front of a robot arm to stream video to a VR headset, in a mirrored robot teleoperation setup (Figure 1). This setup was designed considering our project requirement for nuclear decommissioning using robotic arms to sort nuclear waste which is mostly stored in barrels. However, considering the monocular nature of the Ricoh 360 camera, we found that the operator struggled to safely perform a waste sorting task due to the lack of depth perception. Considering the sensitivity of the task, a monocular 360 camera was considered insufficient. One of the advantages of using VR headsets is the possibility of having a stereo camera streaming video directly on a HMD, which increases the immersion of the operator with the remote robot. Having stereo vision can improve the operator's depth perception.

In this paper we compare the efficiency of three stereo camera setups through a user study. In the firs setup, the user has only one stereo camera at a fixed position. In the second setup, the user has two-fixed stereo cameras with the possibility to change between them. In the third setup, the user has a moving camera, which corresponds to having the camera mounted on another robot arm that is synchronised with the VR HMD, so that the operator's movements are translated to camera movements. In all three setups, the operator has a third person view with respect to the robot manipulating objects and there is no camera mounted on the teleoperated robot arm. This covers all current screen-based methods apart from first-person view, where the camera is mounted on the robot. In the following section, we introduce the details of the user study carried out to compare the effectiveness of these camera placements for a teleoperation task.

2 METHOD

To investigate different interfaces for a teleoperation scenario, a virtual environment was developed to test the three camera setups: one fixed camera, two-fixed cameras with switching, and a moving camera synchronised to the operator's head movements. This section describes the experimental setup, experiment hypotheses, user tasks, measurements, and participant demographics.

2.1 Experiment Setup

The experiment virtual environment was designed using the game engine *Unity 3D* by integrating open source CAD models of a 6DoF robotic arm and of industrial complexes. The Unity basic physics properties, such as gravity and collision, were used to make the environment more realistic. The manipulator end-effector position was controlled with a joystick-like 3Dconnexion Spacemouse, with robot joint values calculated using inverse kinematics. Two keys of the Spacemouse were used for grasping (closing the gripper) and releasing (opening the gripper) objects. Five cylinders of different colours were used as objects to be moved in the pick-and-place task. The environment was designed with similar features for all the conditions, apart from the cameras. For the single and double fixed camera, HMD movement did not result in camera movements as it normally does in VR environments. Participants could switch between the front and side (having a camera 90°rotated and placed at the left of the scene 3.2m from the robot arm base and 1.2m above the floor) cameras at any time in the two-fixed camera condition, using an allocated key on the 3Dconnexion mouse. The front camera position was the same for the single and two-fixed camera conditions, and for the starting point of the moving camera condition located 1.5m to the left and 0.8m behind the robot arm base, and 1m above the floor. The position of the second camera in the two-camera condition was chosen based on participant behaviour from a preliminary experiment with our 360°camera setup (Figure 1) in which people tended to bend to the left when trying to get a side-view of the scene. The experiment setup and virtual environment are depicted in Figures 2 and 3.

2.2 Hypotheses

The main aim of our study was to investigate the effect of the different camera setups on task efficiency, workload and system usability. Based on this, we formulated the following two hypotheses:

H1 The moving camera interface will outperform the setup with one fixed camera.

H2 Since teleoperation systems traditionally use setups with more than one screen, we expect that the performance of the two-fixed camera setup is similar to the setup with the moving camera.

Figure 2: Experimental setup.

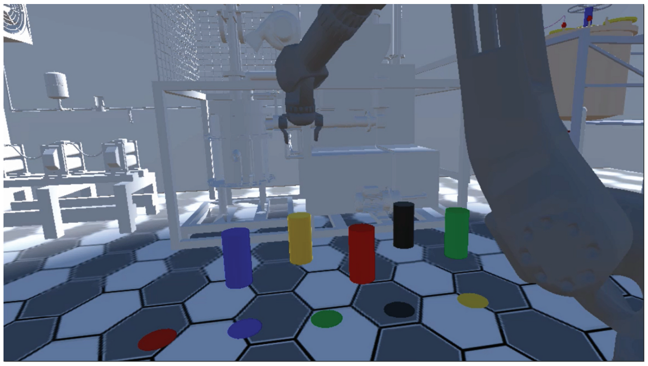

Figure 3: Virtual environment view from the front fixed camera.

2.3 Task, Instructions and Procedure

Participants were asked to wear an HTC Vive VR Headset and to use a 3Dconnexion mouse to move the robot arm and pick five coloured cylinders and place them on five coloured circles. Participants were instructed to move each cylinder to the circle of the same colour and place it as accurately as possible to cover the circle. This task was designed in a way to be similar to a nuclear waste sorting procedure in which a high degree of accuracy is expected. Participants were told their performance was measured based on time and how precise they place each cylinder on the target circle. Before starting the main experiment, participants had a chance to test the interface and to learn how to use the 3D mouse and get used to the VR environment with the two-fixed camera condition. After that, each person completed all three experiment conditions (each condition once) in a randomised order to counterbalance any learning effects. At the end of each condition, participants were asked to fill out two questionnaires (as explained in Section 2.4). After completing all three experiment conditions, participants were debriefed and interviewed with the following three questions:

- Can you rank the interfaces, please?

- Which interface do you recommend to be used in the real industrial application?

If the answer to the second question was the moving camera interface the next question was:

- Considering that implementing a moving camera adds significantly to cost and complexity of the system, as it needs a separate robotic arm to move the camera, which interface do you recommend to be used in the real industrial application?

If they chose any other interface when answering the second question they were only asked why they chose that interface.

2.4 Objective and Subjective Measures

In terms of objective measures the precision and task completion time were considered. The cross-sectional area of the cylinders was the same as the area of the target circles. Participants scored precision points, between 0 and 100, proportional to the ratio of intersection area of these two circles with respect to the target circle area, meaning full coverage of the target circle by cylinder resulted in 100 points. The overall score was calculated by averaging the score for all cylinders.

In terms of subjective measures, two questionnaires, namely the System Usability Scale (SUS) [6] and NASA Task Load Index (NASA TLX) [8] were used. The SUS is a questionnaire that measures the usability of systems, including hardware or software. It consists of 10 item that are rated on a 5 point Likert scale ranging from 1 (“strongly disagree”) to 5 (“strongly agree”). NASA TLX is an assessment tool used to measure workload in different tasks and systems. It includes a workload rating scale of 6 categories: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. Participants also ranked contribution of these categories in the next phase, providing an overall weighted score for the workload.

2.5 Participants

40 participants took part in the experiment 30 of which were male. Participants were staff and student members of the university and had an average age of 31.78 ($STD = 10.17$) ranging between 22 and 63 years old. Participants were mainly right-handed, apart from 4 reported left-handed and 2 ambidextrous participants. All participants reported normal or corrected-to-normal eyesight (17 wearing glasses). Before taking part, participants were asked to rank their pre-existing experience in 5 categories on a scale ranging from 0 (for no experience) to 100 (for highly experienced). The experience categories were VR ($Ave = 33.50$, $STD = 28.72$), 3D gaming ($Ave = 28.50$, $STD = 33.38$), 3D CAD Design ($Ave = 46.13$, $STD = 34.30$), 3Dconnexion Spacemouse ($Ave = 14.00$, $STD = 22.42$) and Teleoperation ($Ave = 19.88$, $STD = 26.66$). Informed consent was obtained from all individual participants included in the study. The study was reviewed and approved by the Ethics Committee of the University of the West of England, Bristol.

3 RESULTS

A one-way analysis of variance (ANOVA) and further Tukey test were performed on the data to compare the mean precision scores of participants and task completion time for each interface. For the precision score, visualised in Figure 4, tests revealed a significant difference between interfaces ($F = 159.55, p < 3.75e-34$) with the precision score for the moving stereo camera interface being significantly higher than the one-fixed camera and two-fixed stereo camera interfaces, and the two fixed camera having a significantly higher precision score than the fixed camera interface. In terms of mean completion time, visualised in Figure 5, an ANOVA showed a significant difference between interfaces ($F = 8.51, p < 0.0004$) with the moving camera interface significantly faster than the one fixed camera and two fixed camera interfaces, and the two fixed camera setup being significantly slower than both other interfaces.

For the subjective measures, a one-way ANOVA was carried out for the results of the NASA TLX and SUS questionnaires. For the NASA TLX, visualised in Figure 6, a difference in the workload was revealed ($F = 53.27, p < 3.56e-17$) with the moving camera interface having significantly less workload than both other interfaces. The analysis of the SUS scores, visualised in Figure 7, shows that

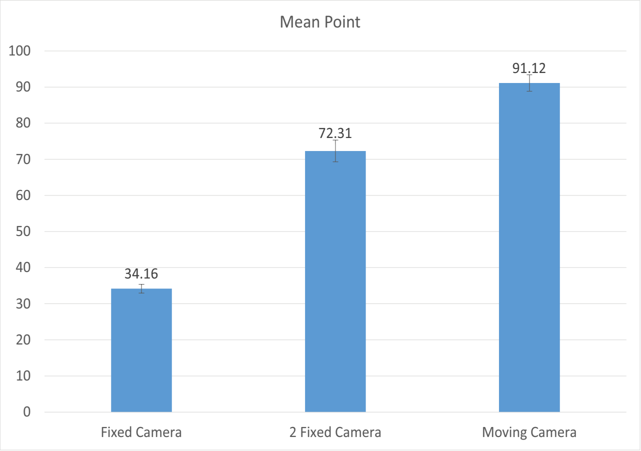

Figure 4: Mean precision scores of three camera conditions. Error bars are ±1SEM.

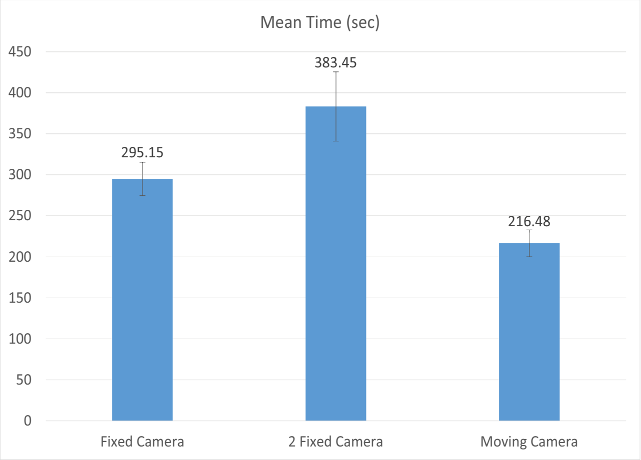

Figure 5: Mean task completion time of three camera conditions in seconds. Error bars are ±1SEM.

Table 1: Mean values, standard deviation and standard error of mean for all objective and subjective measures.

Measures	Fixed Camera			2-fixed Camera			Moving Camera		
	Mean	STD	SEM	Mean	STD	SEM	Mean	STD	SEM
Precision Score	34.16	19.13	1.2	72.31	14.5	3.02	91.12	7.57	2.3
Completion Time (sec)	295.15	127.86	20.22	383.45	267.27	42.26	216.48	103.28	16.33
NASA-TLX Workload	70.49	16.03	2.53	54.71	20.00	3.16	29.4	17.63	2.79
SUS Score	46.88	13.65	2.16	63.94	20.13	3.18	88.81	11.06	1.75

the usability of the moving camera setup was rated significantly higher compared to both other interfaces ($F = 74.72, p < 1.23e{-}21$). The SUS scores can also be interpreted semantically based on the mean score [5] and, therefore, the moving camera interface is considered as "Excellent", the two-fixed camera interface as "OK" and the one-fixed camera interface as "Poor".

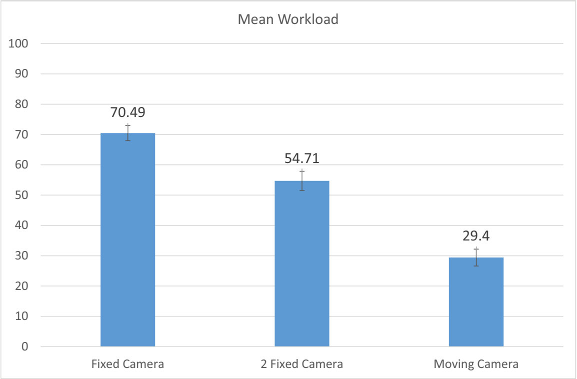

Figure 6: Mean workload scores of three camera conditions. Error bars are ±1SEM.

The mean, standard deviation and standard error of mean (SEM) of the all objective and subjective measures are also presented in Table 1.

Participants answered all three interview questions at the end of the experiment, with 37 out of 40 choosing the moving camera interface, and 3 chose the two-fixed camera setup as the best interface. Those three selecting the two fixed camera as the best setup chose the moving camera as the second best, along with another 31 participants who chose the the two fixed camera setup as the second best. Six others chose the single fixed camera interface as the second best. Consequently, 34 participants selected the single fixed camera and 6 others chose the two-fixed camera interface as the worst interface.

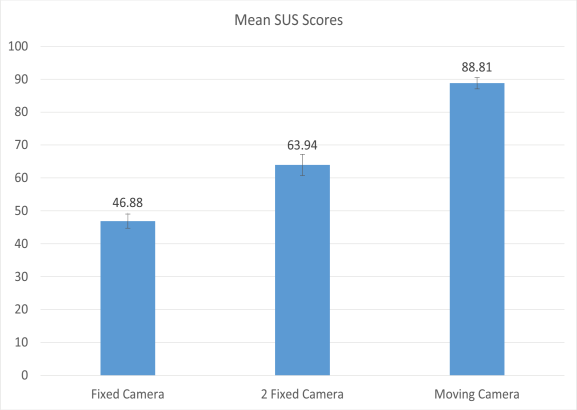

Figure 7: Mean system usability scores of three camera conditions. Error bars are ±1SEM.

Answering to the second question, 35 out 40 participants recommended the moving camera interface for the real industrial application, while 5 others chose the two fixed camera setup, giving their main reasons as "no need for physical movement" and "easier implementation". Finally, 35 participants who recommended the moving camera interface answered to the third question with 22 people still choosing the moving camera and 13 others decided to change to the two fixed camera as a "cheaper" option.

4 DISCUSSION

The experiment results prove our first hypothesis and showed the moving camera interface outperforms both other two interfaces. Even for a simple pick-and-place task in this experiment, the results strongly suggest that having a moving camera will significantly improve teleoperation performance. Not only did participants have better depth perception (hence higher precision) but also they could complete the task faster. It is noteworthy that faster task completion can be interpreted to cost reduction, and higher precision means safer teleoperation when it comes to safety-critical applications. In addition, the moving camera interface scored the lowest mean workload of 29.40 ($STD = 17.63$) and highest SUS mean score of 88.81 ($STD = 11.06$). The moving camera interface seems to be the most intuitive one, as participants could easily perform the task in a faster and more precise way without any previous training. Interestingly, further analysis showed no correlation between participants performance and past experiences in the five areas (see Section 2.5) in which they had assessed their experience level before the experiment. The two fixed camera interface was second in terms of precision score; however, it was significantly slower than both the fixed camera and moving camera interface. Our second hypothesis, therefore, is being rejected as the two fixed camera interface was not close to the moving camera interface in any of the objective and subjective measures.

While 37 out of 40 chose the moving camera interface as the best, when answering the second post experiment questions, 5 people recommended the two fixed camera interface for an industrial application: three giving "no need of physical movement" as their reason and two mentioned "easier implementation" of the two-fixed camera setup as their reason. Considering their age ($Ave = 41.80$, $STD = 11.32$), it can potentially mean that older operators with experience of traditional screen-based systems may find VR interfaces physically demanding, as they need to move around to gain situation awareness rather than just pressing a button to change a camera view.

In terms of complexity and cost, it is important to note that even if a moving camera system requires more expensive equipment, such as an extra robotic arm, a deeper analysis is required to assess the net cost of all the systems. For instance, the moving camera interface has reduced the task completion time (meaning reduced man hours cost), while the two fixed camera interface requires longer time, hence adding to the cost. Also, as many participants mentioned, the two-fixed camera interface needs more training. This is despite the fact that all the participants had a training session with the two fixed camera interface, and yet they needed more training. Extra training can also increase the net cost of the interface as well. Responding to the last post-experiment question, even though 18 participants recommended the two-fixed camera interface as a cheaper option, it might not necessarily be truly cheaper as it was significantly the slowest and required more training.

In terms of system limitation, apart from the equipment cost, the moving camera interface can potentially add complexity. It is important to assess the complexity of this system as having two robotic arms could increase probability of a collision. Hence, an evaluation and validation of the system is required to assure safety of the interface in the nuclear environment. One solution is to have an array of stereo cameras, instead of the second arm, and develop an algorithm to switch at appropriate times between camera views to provide a virtual moving camera. Nonetheless, even such a virtual moving camera system will still be complex and computationally extensive.

One may argue that testing the interfaces in the VR environment does not produce a realistic result, as communication delays are likely in real systems. However, our previous experiment using a ZED stereo camera to stream to a VR environment with a NVIDIA GeForce GTX1060 GPU provided a good example of near real-time video stream, therefore, justifying the use of a VR environment.

5 CONCLUSION

In this paper, a user study for a virtual reality interface for teleoperation is presented. The virtual environment is used to simulate real applications in which the operator has no or little direct visual information of the remote workspace. It was expected that having a stereo camera streaming video to the HMD will improve the operator's depth perception compared to a traditional screen display. Therefore, different camera conditions were compared with the operator wearing a HMD and using a 3D mouse (similar to a traditional joystick) for teleoperation control. Participants performance was measured based on their precision and task completion time, along with subjective measures from the questionnaires. Considering both objective and subjective measures, the study results show that the moving camera interface outperforms the other tested interfaces. This is, therefore, depending on the application and its required level of safety, precision and time that one can choose the right interface for a specific task. In terms of our application of nuclear decommissioning, a moving camera interface has a clear advantage as safety is of high importance. So, in the next steps, this interface will be implemented using two robotic manipulators and will be thoroughly examined and evaluated in terms of real net cost and any safety requirement to comply with nuclear industry standards.

ACKNOWLEDGMENTS

This work was supported by UK Engineering and Physical Sciences Research Council (EPSRC No. EP/R02572X/1) for the National Centre for Nuclear Robotics (NCNR).

REFERENCES

[1] Elahe Abdi, Etienne Burdet, Mohamed Bouri, and Hannes Bleuler. 2015. Control of a supernumerary robotic hand by foot: An experimental study in virtual reality. *PloS one* 10, 7 (2015), e0134501.

[2] Evan Ackerman. 2017. Toyota Gets Back Into Humanoid Robots With New T-HR3. *IEEE Spectrum* (Nov 2017).

[3] B Akan, B Çürüklü, et al. 2010. Augmented reality meets industry: Interactive robot programming. In *Proceedings of SIGRAD 2010: Content aggregation and visualization; November 25–26; 2010; Västerås; Sweden.* Linköping University Electronic Press, 55–58.

[4] Andreas Angerer, Andreas Bareth, Alwin Hoffmann, Andreas Schierl, Michael Vistein, and Wolfgang Reif. 2012. Two-arm Robot Teleoperation using a Multi-touch Tangible User Interface.. In *ICINCO (2).* 327–332.

[5] A. Bangor, P. Kortum, and J. Miller. 2009. Determining what Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of Usability Studies* 4, 3 (2009), 114–123.

[6] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.

[7] Henrik Eschen, Tobias Kötter, Rebecca Rodeck, Martin Harnisch, and Thorsten Schüppstuhl. 2018. Augmented and virtual reality for inspection and maintenance processes in the aviation industry. *Procedia manufacturing* 19 (2018), 156–163.

[8] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology.* Vol. 52. Elsevier, 139–183.

[9] Sunao Hashimoto, Akihiko Ishida, Masahiko Inami, and Takeo Igarashi. 2011. Touchme: An augmented reality based remote robot manipulation. In *The 21st International Conference on Artificial Reality and Telexistence, Proceedings of ICAT2011*, Vol. 2. Citeseer.

[10] Cheng-Peng Kuan and Kuu-Young Young. 2001. VR-based teleoperation for robot compliance control. *Journal of intelligent and robotic systems* 30, 4 (2001), 377–398.

[11] Janus S Liang, Kuo-Ming Chao, and Paul Ivey. 2013. VR-based wheeled mobile robot in application of remote real-time assembly. *The International Journal of Advanced Manufacturing Technology* 64, 9-12 (2013), 1765–1779.

[12] Yuan Lin, Shuang Song, and Max Q-H Meng. 2016. The implementation of augmented reality in a robotic teleoperation system. In *2016 IEEE International Conference on Real-time Computing and Robotics (RCAR).* IEEE, 134–139.

[13] Jeffrey I Lipton, Aidan J Fay, and Daniela Rus. 2017. Baxter's homunculus: Virtual reality spaces for teleoperation in manufacturing. *IEEE Robotics and Automation Letters* 3, 1 (2017), 179–186.

[14] Naresh Marturi, Alireza Rastegarpanah, Chie Takahashi, Maxime Adjigble, Rustam Stolkin, Sebastian Zurek, Marek Kopicki, Mohammed Talha, Jeffrey A Kuo, and Yasemin Bekiroglu. 2016. Towards advanced robotic manipulation for nuclear decommissioning: A pilot study on tele-operation and autonomy. In *2016 International Conference on Robotics and Automation for Humanitarian Applications (RAHA).* IEEE, 1–8.

[15] Daniel Rakita, Bilge Mutlu, and Michael Gleicher. 2018. An autonomous dynamic camera method for effective remote teleoperation. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction.* ACM, 325–333.

[16] Alex Smith, Manuel Giuliani, Mehdi Sobhani, Aravinda Srinivasan, Marc Hanheide, Gerhard Neumann, and Tony Pipe. 2019. Enhancing Teleoperation for Nuclear Decommissioning. In *IROS 2019 Workshop on Legacy Disruptors in Applied Telerobotics: Improving the Machine, the Interface, and the Human.* (in press).

[17] UK Government Report. 2019. Corporate report - Nuclear Provision: the cost of cleaning up Britain's historic nuclear sites. (2019). https://www.gov.uk/government/publications/nuclear-provision-explaining-the-cost-of-cleaning-up-britains-nuclear-legacy/nuclear-provision-explaining-the-cost-of-cleaning-up-britains-nuclear-legacy Last accessed 13 September 2019. Available at: https://www.gov.uk/government/publications/nuclear-provision-explaining-the-cost-of-cleaning-up-britains-nuclear-legacy/nuclear-provision-explaining-the-cost-of-cleaning-up-britains-nuclear-legacy.

[18] David Whitney, Eric Rosen, Elizabeth Phillips, George Konidaris, and Stefanie Tellex. 2017. Comparing robot grasping teleoperation across desktop and virtual reality with ros reality. In *Proceedings of the International Symposium on Robotics Research.*

[19] World Nuclear Association. November 2018. Decommissioning Nuclear Facilities. (November 2018). http://www.world-nuclear.org/information-library/nuclear-fuel-cycle/nuclear-wastes/decommissioning-nuclear-facilities.aspx Last accessed 03 April 2019. Available at:http://www.world-nuclear.org/information-library/nuclear-fuel-cycle/nuclear-wastes/decommissioning-nuclear-facilities.aspx.

[20] Keith Yerex, Dana Cobzas, and Martin Jagersand. 2003. Predictive display models for tele-manipulation from uncalibrated camera-capture of scene geometry and appearance. In *2003 IEEE International Conference on Robotics and Automation (Cat. No. 03CH37422)*, Vol. 2. IEEE, 2812–2817.