Orbital Head-Mounted Display: A Novel Interface for Viewpoint Control during Robot Teleoperation in Cluttered Environments

Kuitert, S.; Hofland, Jelle; Heemskerk, Cock J.M.; Abbink, D.A.; Peternel, L.

DOI
10.1109/IROS55552.2023.10341733

Publication date
2023

Document Version
Final published version

Published in
Proceedings of the 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)

2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 1-5, 2023. Detroit, USA

Orbital Head-Mounted Display: A Novel Interface for Viewpoint Control during Robot Teleoperation in Cluttered Environments

Sjoerd Kuitert[1,2], Jelle Hofland[2], Cock J. M. Heemskerk[2], David A. Abbink[1], and Luka Peternel[1]

2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) | 978-1-6654-9190-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/IROS55552.2023.10341733

Abstract— **Robotic teleoperation is used in various applications, including the nuclear industry, where the experience and intelligence of a human operator are necessary for making complex decisions that are beyond the autonomy of robots. Human-robot interfaces that help strengthen an operators situational awareness without inducing excessive cognitive load are crucial to the success of teleoperation. This paper presents a novel visual interface that allows operators to simultaneously control a 6-DoF camera platform and a robotic manipulator whilst experiencing the remote environment through a virtual reality head-mounted display (HMD). The proposed system, Orbital Head-Mounted Display (OHMD), utilizes head rotation tracking to command camera movement in azimuth and elevation directions around a fixation point located at a robot's end-effector. A human factor study was conducted to compare the interface acceptance, perceived workload, and task performance of OHMD with a conventional interface utilizing multiple fixed cameras (Array) and a standard head-mounted display implementation (HMD). Results show that both the OHMD and HMD interfaces significantly improve task performance, reduce perceived workload and increase interface acceptance compared to the Array interface. Participants reported they preferred OHMD due to the increased assistance and freedom in viewpoint selection. Whilst OHMD excelled in usefulness, the standard HMD interface allowed operators to perform robotic welding tasks significantly faster.**

I. Introduction

Teleoperation systems enable human operators to execute dexterous tasks through robots in hazardous remote environments such as nuclear facilities. In these environments, the experience and intelligence of a human operator is necessary for making decisions that are beyond the autonomy of robots. In teleoperation, operators rely predominantly on visual feedback to gain real-time information about the remote process [1]. As a result, visual feedback interfaces are critical to the success of teleoperation [2].

Conventional feedback systems display several camera views on multiple monitors, which induces high cognitive load at the cost of remote task performance [3]. The underlying reason is that operators need to divide their attention over multiple displays, piece together the visual information, and perform mental transformations due to a misalignment between the camera and robot reference frames [4]. Furthermore, fixed placement of cameras is limited or not possible in cluttered environments and can be occluded [5].

[1]Cognitive Robotics, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: sjoerdkuitert@hotmail.com)
[2]Heemskerk Innovative Technology B.V., 2629 HD Delft, The Netherlands

A common alternative is an eye-in-hand (EIH) camera, which is a camera attached to the end-effector of the teleoperated robot that performs the task (hereinafter referred to as *task robot*). While this approach enables more adaptation of the viewpoint, the camera pose is coupled with the task execution. Therefore, the operator needs to perform visual inspections and manipulation sequentially. Additionally, the EIH-view is often occluded by an object that is being manipulated. Furthermore, information about the surrounding workspace is typically very limited [6].

Some of these drawbacks can be avoided by the introduction of a secondary robot (hereinafter *monitoring robot*) to track the task robot, the process, and the remote workspace. In two extreme cases, the monitoring robot can be controlled autonomously or manually by the operator. The study in [7] proposed an autonomous monitoring robot to monitor a remote-controlled robotic manipulator during offshore operations. The operator was only able to select the desired view distance. Later a learning system was added that remembers previously set view distances [8]. However, this still limits the operator's choice, thus giving the operator direct control might be preferred. The method in [9] proposed a manually controlled monitoring robot for a space docking task, where the results indicated that operators preferred the system over a traditional eye-in-hand camera and experienced better spatial awareness. Nevertheless, the interface should not overload the operator and affect the primary task.

A compromise between autonomous and manual solutions can be achieved through a shared control approach. The method in [10] used a shared control approach where the operator controlled the camera orientation while an autonomous component moved the camera to prevent obstructed views. The operator controlled the monitoring camera via a head-tracking system. While the head-tracking input allows hands-free camera operation, the head rotation with respect to the display (i.e. view rotation) can induce additional cognitive load [3].

In [11] authors argued that operators should dictate task-specific viewpoint adaptions by giving spare input to nudge a viewpoint algorithm towards specific solutions, while a monitoring robot should move its camera towards that solution. Furthermore, the authors added a mode that allowed operators to take manual control of the monitoring robot to enable visual exploration of the remote site. A user study indicated that their new system performed better than the autonomous system [12], as participants found the new system easier to use despite having to control more viewpoint parameters. Nevertheless, the design of a human-machine

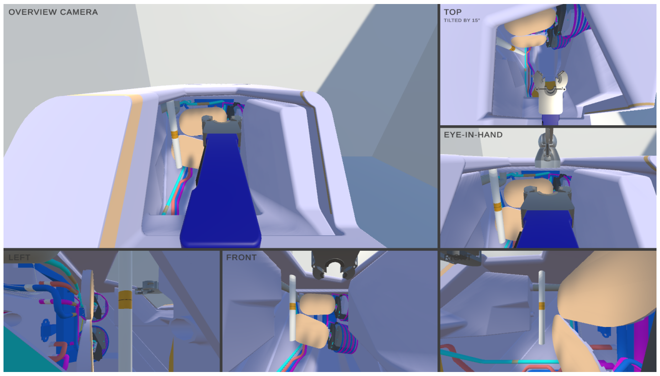


Fig. 1. A typical operator view when using an array of cameras.

interface that allows for natural and efficient control of both the camera viewpoint and the task robot is still an open research challenge [13].

To tackle this challenge, we propose a visual interface called Orbital Head-Mounted Display (OHMD). This method combines a third-person free-follow camera and a head-mounted display (HMD) to control a camera viewpoint using head rotations. While a third-person free-follow camera has been used before for monitoring a remote process [13], [14], it has not been interfaced with an HMD and head-tracking for viewpoint control. To gain new insights regarding the usability of the developed OHMD interface, we compare it to a conventional array of multiple fixed cameras (Array) and a standard HMD interface. The camera interfaces are evaluated based on task performance, perceived workload and interface acceptance in a one-factor within-subjects experiment for a remote welding task of a cooling pipe in a simulated environment.

We hypothesize that the proposed OHMD and the standard HMD interfaces improve task performance, reduce perceived workload, and increase self-reported interface acceptance compared to an array of multiple fixed cameras (H1). Furthermore, we hypothesize that the proposed OHMD interface will improve self-reported interface usefulness compared to a standard HMD implementation, without sacrificing task performance and increasing the perceived workload (H2).

II. Methods

This section outlines the three visual interfaces that are compared in this study through simulation, with two being baselines (Array and HMD) for the proposed method.

A. Array of Multiple Fixed Cameras

The standard approach for providing visual feedback from a remote environment in teleoperation is the use of multiple cameras placed in fixed positions. Arrays of the camera feeds are presented to the operator on LCD displays. In this study, the array comprises 6 views, which are displayed to the operator on a single monitor. The layout includes one large section scaled to a resolution of 1280x720 surrounded by 5 smaller sections set to a lower resolution of 640x360 along its bottom- and right edge as shown in Fig. 1. The large section shows the feed of a line-of-sight overview camera. The additional views include feeds from orthogonally positioned cameras along the bottom edge (left-, front-, and right-side views) and top-right corner (top-down view) of the screen. It should be noted that perfectly orthogonal cameras are better than what can be achieved in most real-world scenarios where lack of access and obstructions often yield sub-optimal camera positions. The final view includes the feed of an eye-in-hand camera.

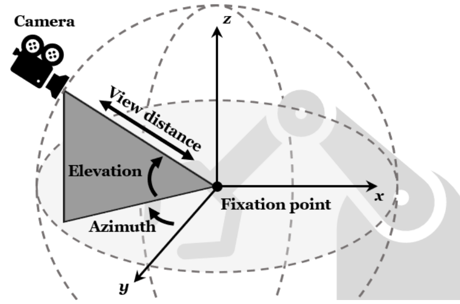


Fig. 2. An illustration of the Orbital Head-Mounted Display interface.

B. Standard Head-Mounted Display

In a standard VR setup, an operator wears an HMD and views a stereoscopic image of a 3D remote environment through small embedded display optics located in front of each eye. The operator can then adjust the view by head rotations measured by motion tracking sensors, thus they experience virtual viewpoint adaptations similar to real-life.

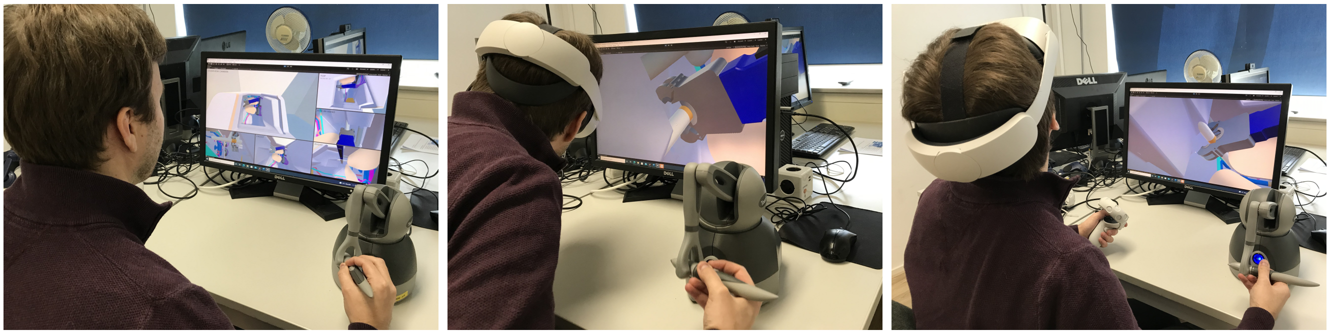
C. Orbital Head-Mounted Display

The concept of the OHMD method is based on the so-called *third-person free follow camera* used in video games. This camera automatically positions itself relative to the character to provide a third-person view and the camera follows the character whenever it moves. Players can pan the camera around their character and change the view distance (i.e. zoom) using camera controls.

Instead of using a conventional interface, the proposed OHMD uses motion tracking sensors of an HMD to enable the operator to control a free-follow camera with their head movements. Specifically, the operator's head rotation controls the camera's azimuth and elevation angle with respect to a fixation point as if the camera is constrained to the surface of a sphere (Fig. 2). Additionally, a look-at constraint ensures the camera is always pointed towards the fixation point, which is located at the task robot's end-effector in this study. The view distance was bound to a hand-held joystick.

III. Experiments

To test our hypotheses, we evaluated the effect of the visual interface on task performance, perceived workload and interface acceptance in a one-factor within-subjects experiment. During the experiment, each participant used each visual interface described in Section II to perform a teleoperated welding task in a cluttered space, thus resulting in three experiment conditions: Array, HMD and OHMD. The conditions were presented in a counterbalanced order. The setup and experiments were approved by the Human



(a) (b) (c)

Fig. 3. Experimental setup for a teleoperated welding task. Images show a desktop haptic device in combination with one of the visual interfaces: (a) Array, (b) standard HMD, and (c) the proposed OHMD, where the operator additionally holds a joystick to control the view distance.

Research Ethics Committee (HREC) of the Delft University of Technology. Twelve participants (10 male, 2 female) between 26 and 62 years old (M = 31.2, SD = 9.9) participated in the study. Prior to the experiment, participants gave their informed consent and completed questions on a 6-point Likert scale about their gaming, VR and remote control experience.

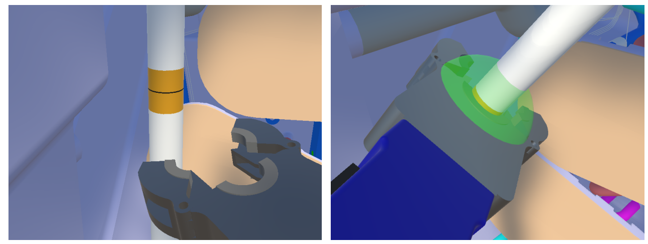
A. Experimental Setup

The experimental setup is shown in Fig.3. For the Array interface (Fig. 3a), we used a standard 24.1 inch LCD monitor with 1920x1200 resolution to display visual feedback and was placed 60-80cm from the participant. An Oculus Quest 2 VR-HMD was used for the HMD and OHMD interfaces (Fig. 3b,c). During the OHMD condition, the joystick on the Oculus Motion Controller was used to control the view distance. We use a 3D Systems Touch haptic device for the operator to send 6-DoF movement commands to the task robot end-effector and receive 3-DoF force feedback.

B. Use Case and Task Description

We selected a robotic welding task of a cooling pipe within the Electron Cyclotron Heating Upper Port Launcher (ECH UPL) of the ITER nuclear fusion reactor [15] as a use case task. This use case has several elements of complexity that make it interesting for the experiment, such as a high risk of damaging equipment, working in a cluttered and confined space, and operating in a hazardous remote environment. Such use cases typically have limited sensory information from the remote environment and thus operators rely on or benefit from a digital twin of the environment. Therefore, we developed and studied the use case in a simulated environment.

In each trial, the participants had to move the welding tool from a set starting point towards a cut in a cooling pipe indicated by an orange marker (Fig. 4a). After reaching the pipe, they had to fit the profile of the tool's head around the cut. A visual cue (Fig. 4) appeared to notify participants they had completed the task when the alignment was within 5 mm for x, y, z, and 2° for tilt and roll. To introduce different levels of complexity, and to change the visual information provided by each viewpoint [16], we varied the cooling pipe placements between 3 locations (Fig. 5, 4a, and 4b). The participants were instructed to perform the task as fast as possible. Also, they were told not to damage the equipment by avoiding collisions with any components other than the cooling pipe and to minimize the force exerted by the welding tool onto the pipe.



(a) (b)

Fig. 4. (a) Visual indication of the pipe cut location. (b) Visual cue to indicate the alignment is within task success margins.

C. Software and Simulated Environment

The use case and remote environment were simulated in a virtual environment built in Unity3D. For this purpose, we created a 3D model loosely based on the UPL. All physics interactions between objects are computed by the Unity Physics engine using a Temporal Gauss Seidel solver. To weld the cooling pipes inside the UPL we used a modified version of the ORBIWELD 38S TIG orbital welding tool. The operator commanded the movement of the tool as if its handle is connected to the end-effector of a robotic manipulator.

To couple the movement of the stylus and the tool, participants had to press and hold the front button on the haptic device's stylus. When the button was released, the stylus and tool were decoupled, which allowed users to re-index the haptic device's workspace and take a comfortable hand/wrist posture. The translational movement input was scaled such that the welding task could be completed without re-indexing.

When wearing a VR headset (HMD and OHMD conditions), the operator cannot see the physical haptic device, which can make it difficult to recognize the haptic device pose, especially while re-indexing. To aid the participant with this visual feedback, we added a virtual model of the haptic device above the tool while the operator is re-indexing (Fig. 5). The model was aligned with the world coordinate frame

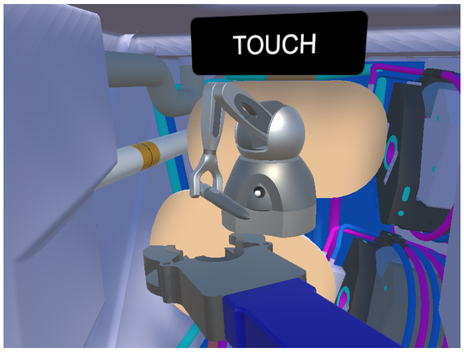


Fig. 5. The virtual model of the haptic device (3D Systems Touch) that appears above the welding tool when the human-operator is re-indexing.

and mimicked the configuration of the physical device. While during the Array condition the physical haptic device can be seen on the desk, it can fall outside the operator's field of view while looking at the monitor. Therefore, the virtual Touch haptic device was also displayed in the EIH-view.

D. Data Acquisition and Metrics

We collected subjective data through questionnaires. Additionally, objective data of the haptic device, HMD and Unity simulation was recorded. The three visual interfaces were then assessed using the following metrics:

- *Preferred visual interface*. In the post-experiment questionnaire, participants were asked which visual interface they preferred and to explain their choice.
- *Usefulness and satisfying scores*. The Van der Laan Acceptance Scale [17] was used to capture the usefulness and satisfying score of the visual interfaces. Participants rated nine components on a five-point Likert scale from -2 to 2 for each visual interface.
- *Task load rating*. The NASA Task Load Index (TLX) questionnaire [18] was used to capture the perceived workload. The NASA-TLX is a multi-dimensional rating procedure that provides an overall workload score based on a weighted average of ratings on six subscales: mental demands, physical demands, temporal demand, own performance, effort and frustration.
- *Time to completion (ttc)*. The amount of time it takes a participant to complete the remote welding task. While effective views may enable operators to perform the task faster, they may spend additional time to select a view, control the camera or process the visual information.
- *Number of collisions*. The number of undesired collisions with the environment. This involves all collisions between the tool (end-effector) and UPL components excluding the cooling pipe.
- *Peak contact force (pcf)*. The peak contact force denotes the highest measured contact force between the tool and the cooling pipe. This gives an indication of how smooth the profile of the tool slips around the pipe. It is expected that the participant is able to fit the tool more easily when views that show the profile's alignment with respect to the cooling pipe are available.

E. Experimental Procedure

Prior to the experiment, participants were briefed on the goal of the study and were given experiment instructions. Participants were then allowed 5 minutes to get acquainted with the setup and simulation, without exposing them to any of the three conditions. They practised on a cooling pipe that was positioned in a unique location and feedback was given by a single view on a monitor.

The main experiment included the following steps for each condition. First, they were given 3 minutes to get used to the visual interface and camera controls. Then, for the three pipe locations, participants performed a practice trial followed by three recorded trials. After each condition, they were asked to complete the subjective questionnaires. The order in which participants were exposed to the visual interfaces and pipe locations was randomized according to the Balanced Latin Square method to reduce carry-over effects. After they completed all conditions, they were asked to fill in a post-experiment questionnaire with questions regarding their experience of the study and their preferred visual interface.

F. Data Analysis

We used repeated measures analysis of variance (RM-ANOVA) to analyze the results. Additionally, post-hoc analyses with a Bonferroni correction were performed for pairwise comparison of the three different visual interface conditions. The performance metrics were averaged over the three measured trials for each combination of visual interface and pipe location per participant. A moving-average filter was used on collision data to remove sudden peaks caused by inaccuracies in the physics simulation. The Shapiro-Wilk test was used to test whether the collected data fit a normal distribution. To enable the use of an RM-ANOVA, we transformed non-parametric data and parametric data that violates ANOVA assumptions (e.g., normality) using the Aligned Rank Transform (ART) [19].

IV. Results

We conducted one-way RM-ANOVA on metric data to check for significant differences between the interfaces. A summary of the results is given in Table I. Observed differences were considered statistically significant at p-values of 0.05 or less.

A. Interface Acceptance

A one-way RM-ANOVA revealed a highly significant difference in mean usefulness across the visual interfaces, $F(2, 22) = 24.29, p < 0.001$. Post-hoc analysis showed that OHMD improves the mean interface usefulness score by 1.53 compared to Array ($p = 0.001$) and 0.56 compared to HMD ($p = 0.036$) conditions. While the standard HMD interface scored lower than OHMD, it did improve by 0.97 compared to the Array ($p < 0.001$).

The results show that the Array interface is the least satisfying to use. Compared to Array, the HMD and OHMD conditions significantly improve the mean satisfying score by 1.58 ($p \leq 0.001$) and 1.44 ($p = 0.010$) respectively. There

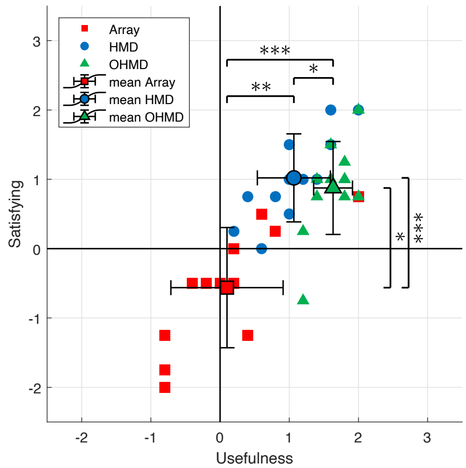


Fig. 6. Overall Van der Laan's acceptance scores [17] on two dimensions; a usefulness scale (horizontal) and a satisfying scale (vertical). The error-bars indicate the standard deviations from the mean. The marks '$***$', '$**$', and '$*$' denote a significance of $p \leq 0.001$, $p \leq 0.01$, and $p \leq 0.05$, respectively, and 'n.s.' denotes no significance.

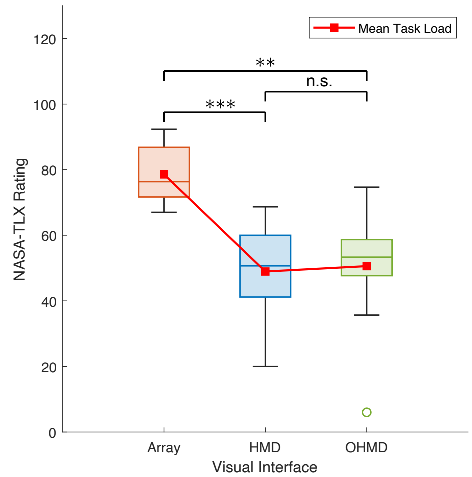


Fig. 7. Weighted NASA-TLX ratings for each visual interface. NASA-TLX values range between 0 and 100, with higher values indicating higher task load. The marks '$***$', '$**$', and '$*$' denote a significance of $p \leq 0.001$, $p \leq 0.01$, and $p \leq 0.05$, respectively, and 'n.s.' denotes no significance.

is a small but insignificant decrease in interface satisfaction when comparing OHMD to HMD.

Also for the mean satisfying score significant differences were found, $F(2, 22) = 13.44, p < 0.001$. A Bonferroni multi-comparison test revealed significant differences in mean usefulness between Array and HMD ($p < 0.001$), and between Array and OHMD ($p \leq 0.001$). No significant difference was found between HMD and OHMD.

According to the post-experiment questionnaire responses, 8 out of 12 participants preferred the OHMD visual interface because of the increased flexibility in selecting viewpoints (4 mentions), its zoom feature (2 mentions), intuitiveness and assistance (2 mentions). The other 4 participants preferred the regular HMD implementation because it felt the most natural and simple to understand (4 mentions) and it gave a sense of embodiment (2 mentions). Four participants mentioned the remote welding task was most difficult to perform when using the Array interface which caused frustration and provoked a trial-and-error approach.

B. Perceived workload

Statistical analysis results for the self-reported perceived workload (Fig. 7) revealed a significant difference in mean TLX-ratings, $F(2, 22) = 15.11, p < 0.001$. On average the perceived workload decreased when a VR-HMD was used instead of the Array interface. This is indicated by decreases in TLX-rating by 38% ($p \leq 0.001$) and 36% ($p = 0.002$) for HMD and OHMD respectively. No significant difference in perceived workload was found between the two VR interfaces.

C. Task performance

The results for the objective performance metrics are presented in Fig. 8 and Table I. For time to completion, a significant effect was found, $F(2, 22) = 48.77, p < 0.001$. Interestingly, participants were able to perform the simulated welding task much quicker with the HMD interface, requiring only 23.29 seconds on average compared to 60.34 seconds with the Array interface ($p < 0.001$) and 30.35 seconds with the OHMD interface ($p = 0.010$).

The visual interface also affected the number of collisions made, with approximately 2 collisions per trial for the VR-HMD interfaces and 7 collisions per trial for the Array interface. Additionally, out of 108 trials per interface, only 18 trials (17%) were completed without any undesired collisions when the Array interface was used. With a VR-HMD the amount of collisions-free trials increased to 36 trials (33%) for the HMD and 37 trials (34%) for the OHMD interfaces. Participants applied lower peak force onto the cooling pipes when using the HMD and OHMD interfaces instead of a camera array.

V. Discussion

In summary, the findings are that both the OHMD and HMD systems improve the self-reported system acceptance compared to the standard array. None of the participants preferred the Array interface over either OHMD or HMD. Additionally, the OHMD and HMD interfaces reduce the perceived workload by 36% and 38% respectively and significantly outperform the Array system on all recorded objective metrics. This confirms the first hypothesis. The second hypothesis is only partially confirmed. OHMD indeed scored a higher usefulness rating compared to HMD. Furthermore, OHMD and HMD induced similar task loads. However, there was a significant difference in one of the objective performance metrics. While using the OHMD interface the tasks took 30% longer to complete on average. We discuss these results below.

According to the results both HMD and OHMD have an increased interface acceptance compared to the conventional Array setup. Individual components of the Acceptance Scale questionnaire indicate that participants found the OHMD interface more useful, effective and assisting than the HMD interface. However, they also reported it was less pleasant to use and not as likeable. This aligns with participant responses

TABLE I

RM-ANOVA RESULTS, MEAN VALUES (M) AND STANDARD DEVIATIONS (SD) FOR EACH METRIC.

Metric		**Visual interface** *Array*	*HMD*	*OHMD*	**RM-ANOVA, (F2,22)** *Effect: interface*	**Post-hoc analysis** *Array-HMD*	*Array-OHMD*	*HMD-OHMD*
NASA-TLX rating	M	78.56	48.94	50.58	$p < 0.001$	$p < 0.001$	$p = 0.002$	$p = 1$
	SD	9.29	14.75	16.95	$F = 15.11$			
Usefulness score	M	0.10	1.07	1.63	$p < 0.001$	$p = 0.001$	$p < 0.001$	$p = 0.036$
	SD	0.81	0.53	0.28	$F = 24.29$			
Satisfaction score	M	-0.56	1.02	0.88	$p < 0.001$	$p < 0.001$	$p = 0.010$	$p = 1$
	SD	0.87	0.63	0.67	$F = 13.43$			
Preferred interface	Count	0	4	8	–	–	–	–
Time to completion*	M	60.34	23.29	30.35	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p = 0.010$
	SD	79.72	18.56	20.49	$F = 48.77$			
Number of collisions*	M	7.41	2.17	2.32	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p = 1$
	SD	9.53	2.79	2.88	$F = 31.11$			
Peak contact force*	M	30.42	18.13	18.91	$p < 0.001$	$p = 0.001$	$p = 0.034$	$p = 1$
	SD	36.94	24.87	18.18	$F = 8.03$			

* Observations violate the normality assumption according to the Shapiro-Wilk test [20]. This data was transformed using an *Aligned Rank Transform* (ART) for multifactor contrast tests [19] so that the ANOVA and post-hoc procedures could be applied.

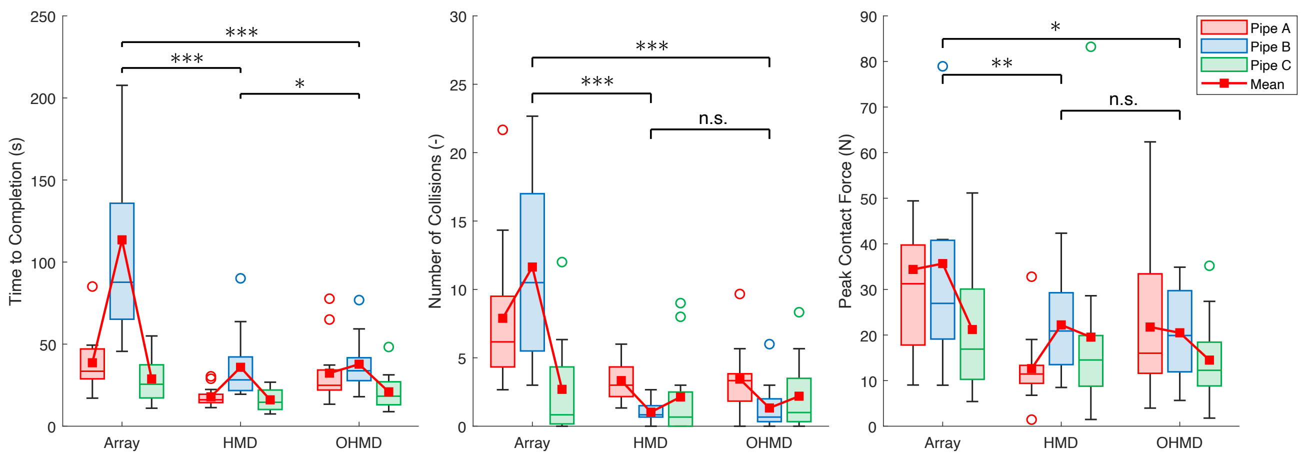


Fig. 8. Results for the objective task performance metrics. The significance lines above the box-plots indicate differences in mean between the visual interfaces. The marks '$***$', '$**$', and '$*$' denote a significance of $p \leq 0.001$, $p \leq 0.01$, and $p \leq 0.05$, respectively, and 'n.s.' denotes no significance.

in the post-experiment questionnaire. The acceptance of the HMD interface aligned with results from other studies [1].

A few participants reported a slight camera shake while using the OHMD method. Further analysis indicated three possible explanations. First, participants might have made small unconscious head movements (also observed in [21]). At a large view distance, small head movements can cause relatively large camera translations. Second, most participants did not hold the stylus of the haptic device perfectly still. Thus, the welding tool linked with the stylus can also shake. Since the used haptic device was small and position command scaling was used to better exploit the remote robot workspace, tremor on the operator side can amplify on the remote side. Finally, the motion sensors in the HMD have some noise. In future, this can be mitigated by filtering the input (e.g., dead zone or frequency filter) or by adding a moderate amount of damping.

According to the NASA-TLX questionnaire, the Array interface caused a lot of frustration, induced high mental workloads and the trials took a lot of effort to complete. This aligns with findings in literature [3]. On the other hand, the overall task load rating for the OHMD and HMD interfaces was much lower than for Array. Participants rated the physical demand and effort higher for the OHMD interface compared to the standard HMD implementation. During the experiments, very few participants were actively moving while using the regular HMD interface. The stereoscopic view of the HMD already provided enough contextual and depth information. This result aligns with findings in [6] which found that for most tasks a passive (i.e. fixed) stereoscopic view is sufficient and that only more complex tasks require dynamic views.

A potential explanation why OHMD induced more physical demand compared to HMD is that the OHMD viewpoint followed the welding tool and the operator was forced to re-adjust the camera position in order to avoid occlusions. Because of this, we observed that the operators were more involved in the visual task (i.e., they were more actively searching for better views). This is perhaps positively reflected in the higher 'Raising Alertness' score for OHMD compared to HMD, as a dynamic view provides more depth cues such as motion parallax and optical flow [22], [23].

In the post-experiment questionnaire, participants reported that the trials were very difficult to complete when using the

Array interface and that they considered pipe B (Fig. 5) to be the most difficult. The task performance results indeed show that the task for pipe B took the most time to complete regardless of which interface was used. Because pipe B was positioned horizontally in the UPL, participants had to roll the welding tool by 90 degrees around its longitudinal axis. As a result, the view from the eye-in-hand camera was also rotated while the views of the HMD and OHMD interface were unaffected by this rotation. This made it more difficult to understand the link between the robot control input and the eye-in-hand camera view. However, based on the number of collisions, pipe B seemed to be easier than pipe A and C when the OHMD or HMD was used. In line with the literature [24], participants paid considerable attention to the immersive egocentric view provided by the eye-in-hand camera even though the view was ineffective.

Based on the analyses of the results, OHMD could be improved by allowing operators to set the camera movement sensitivity to their preference. Additionally, allowing operators to change the task robot's control frame of reference from the world coordinate system to the OHMD camera's local coordinate system could be another improvement. Finally, in real-world scenarios, other factors such as signal noise and delays could affect all three systems differently, so future studies could examine these effects.

VI. Conclusion

This study proposed a novel visual interface called Orbital Head-Mounted Display that uses the motion tracking of a stereoscopic HMD to control the view provided by a free-follow camera platform. The developed interface has been found to provide benefits during teleoperated manipulation tasks. Its effects were evaluated in a one-way human factors experiment for a remote welding task of a cooling pipe in a simulated environment. From the results, we can conclude:

- The OHMD and standard HMD interfaces improve task performance, have an increased interface acceptance, and lower perceived workload compared to a conventional array of cameras.
- The OHMD interface is perceived more useful than the standard HMD interface.
- The OHMD and HMD interfaces induce a similar perceived workload.
- Task completion takes approximately 30% longer when using the OHMD interface instead of a standard HMD implementation.

The follow-constraint attribute of the OHMD interface also makes it suitable for teleoperating mobile robots or manipulators with a large workspace.

References

[1] E. Triantafyllidis, C. McGreavy, J. Gu, and Z. Li, "Multimodal Interfaces for Effective Teleoperation," *arXiv*, mar 2020.

[2] A. Pandya, L. Reisner, B. King, N. Lucas, A. Composto, M. Klein, and R. Ellis, "A Review of Camera Viewpoint Automation in Robotic and Laparoscopic Surgery," *Robotics*, vol. 3, no. 3, pp. 310–329, 2014.

[3] B. DeJong, J. Colgate, and M. Peshkin, "Mental Transformations in Human-Robot Interaction," *Mixed Reality and Human-Robot Interaction*, pp. 35–51, 2011.

[4] C. D. Wickens, J. W. Keller, and R. L. Small, "Left. no, right! Development of the frame of reference transformation tool (FORT)," *Proceedings of the Human Factors and Ergonomics Society*, vol. 2, pp. 1022–1026, 2010.

[5] D. Clark, *Effect of camera placement and display configuration on a remote manipulation task*. Embry-Riddle Aeronautical University, 2006.

[6] S. Yu, J. Le, B. PARK, K. Kim, and I. Cho, "Ergonomic Analysis of a Telemanipulation Technique for a Pyroprocess Demonstration Facility," *Nuclear Engineering and Technology*, vol. 46, no. 4, pp. 489–500, aug 2014.

[7] M. Bjerkeng, K. Y. Pettersen, and E. Kyrkjebø, "Stereographic projection for industrial manipulator tasks: Theory and experiments," *IEEE International Conference on Intelligent Robots and Systems*, pp. 4676–4683, 2011.

[8] S. A. Fjerdingen, M. Bjerkeng, A. A. Transeth, E. Kyrkjebø, and A. Røyrøy, "A learning camera platform for remote operations with industrial manipulators," *IFAC Proceedings Volumes (IFAC-PapersOnline)*, vol. 1, no. PART 1, pp. 39–46, 2012.

[9] M. Wilde, Z. K. Chua, and A. Fleischner, "Effects of Multivantage Point Systems on the Teleoperation of Spacecraft Docking," *IEEE Transactions on Human-Machine Systems*, vol. 44, no. 2, pp. 200–210, apr 2014.

[10] S. Rahnamaei and S. Sirouspour, "Automatic Viewpoint Planning in Teleoperation of a Mobile Robot," *Journal of Intelligent & Robotic Systems*, vol. 76, no. 3-4, pp. 443–460, dec 2014.

[11] D. Rakita, B. Mutlu, and M. Gleicher, "Remote Telemanipulation with Adapting Viewpoints in Visually Complex Environments," in *Robotics: Science and Systems XV*, jun 2019.

[12] ——, "An Autonomous Dynamic Camera Method for Effective Remote Teleoperation," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. New York, NY, USA: ACM, feb 2018, pp. 325–333.

[13] E. Sita, "Monitoring Companion for Industrial Robotic Processes," Ph.D. dissertation, University of the West of England, 2021.

[14] M. Talha and R. Stolkin, "Preliminary Evaluation of an Orbital Camera for Teleoperation of Remote Manipulators," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, nov 2019, pp. 2754–2761.

[15] D. Strauss *et al.*, "Nearing final design of the iter ec h&cd upper launcher," *Fusion Engineering and Design*, vol. 146, pp. 23–26, 2019.

[16] B. G. Brooks and G. T. McKee, "Toward the implementation of visual acts for automated viewing during teleoperation," in *Sensor Fusion and Networked Robotics VIII*, P. S. Schenker and G. T. McKee, Eds., vol. 2589, no. September 1995, sep 1995, pp. 230–241.

[17] J. D. Van Der Laan, A. Heino, and D. De Waard, "A simple procedure for the assessment of acceptance of advanced transport telematics," *Transportation Research Part C: Emerging Technologies*, vol. 5, no. 1, pp. 1–10, feb 1997.

[18] S. G. Hart, "Nasa-Task Load Index (NASA-TLX); 20 Years Later," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 50, no. 9, pp. 904–908, oct 2006.

[19] L. A. Elkin, M. Kay, J. J. Higgins, and J. O. Wobbrock, "An Aligned Rank Transform Procedure for Multifactor Contrast Tests," *UIST 2021 - Proceedings of the 34th Annual ACM Symposium on User Interface Software and Technology*, pp. 754–768, 2021.

[20] S. Shapiro and M. Wilk, "An analysis of variance test for normality (Complete samples)," *Biometrika*, vol. 52, no. 3-4, pp. 591–611, 1965.

[21] D. Zhu, T. Gedeon, and K. Taylor, "Head or gaze? Controlling Remote Camera for Hands- Busy Tasks in Teleoperation: A Comparison," in *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction*. New York, New York, USA: ACM Press, 2010, p. 300.

[22] J. A. Gomer, C. H. Dash, K. S. Moore, and C. C. Pagano, "Using Radial Outflow to Provide Depth Information During Teleoperation," *Presence: Teleoperators and Virtual Environments*, vol. 18, no. 4, pp. 304–320, aug 2009.

[23] G. P. Bingham and C. C. Pagano, "The necessity of a perceptionaction approach to definite distance perception: Monocular distance perception to guide reaching." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 24, no. 1, pp. 145–168, 1998.

[24] L. C. Thomas and C. D. Wickens, "Visual Displays and Cognitive Tunneling: Frames of Reference Effects on Spatial Judgments and Change Detection," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 45, no. 4, pp. 336–340, 2001.