Msc. Thesis
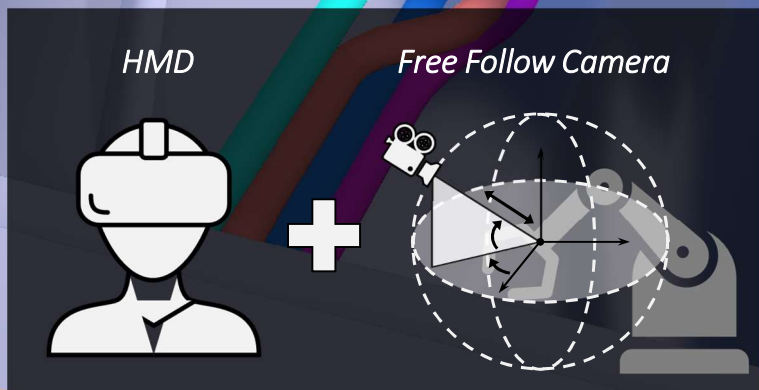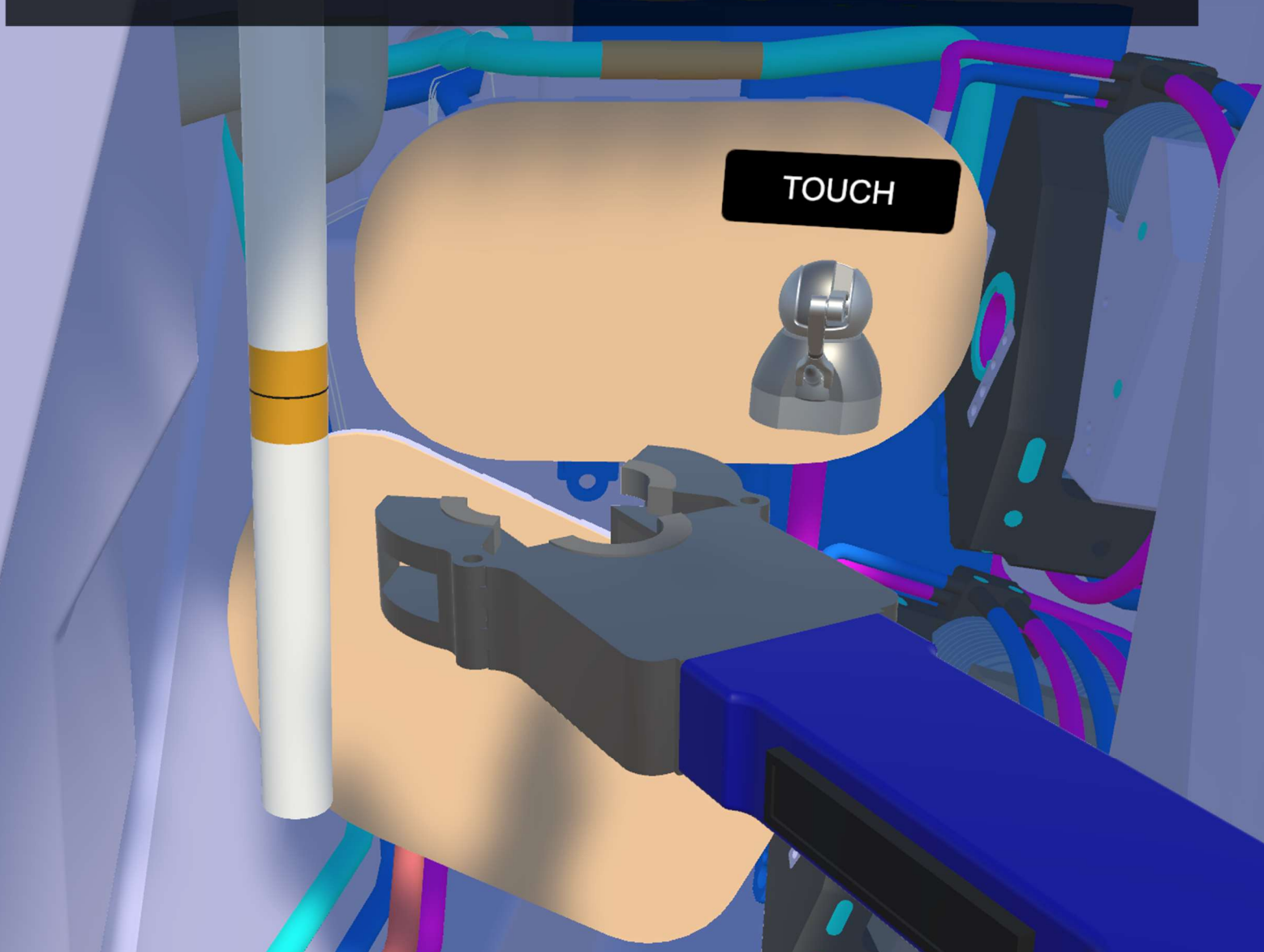
Orbital HMD: An Intuitive Visual Interface for Robot Teleoperation

With Robotic Welding as Use Case

Sjoerd Kuitert

TOUCH

HMD

Free Follow Camera

TUDelft

HiT
Heemskerk Innovative Technology

Orbital HMD: An Intuitive Visual Interface for Robot Teleoperation

Sjoerd Kuitert

A thesis submitted for the degree of Master of Science

Faculty of Mechanical Engineering, Delft University of Technology, Delft
In collaboration with Heemskerk Innovative Technology B.V., Delft

Student ID:	4092449	
Master Track	BioMechanical Design	
Thesis committee:	Prof. dr. ir. D.A. Abbink,	TU Delft, committee chair and supervisor
	Dr. ir. L. Peternel,	TU Delft, supervisor
	Ir. J. Hofland	Heemskerk Innovative Technology B.V., supervisor
	-	TU Delft, external member

TU Delft
March, 2022

Orbital HMD: An Intuitive Visual Interface for Effective Robot Teleoperation

Sjoerd Kuitert

Supervised by: Jelle Hofland, Cock J. M. Heemskerk, David A. Abbink, Luka Peternel

Abstract—Robotic teleoperation has gained popularity in various applications, including the nuclear industry, where the experience and intelligence of a human operator is necessary for making complex decisions that are beyond the autonomy of robots. Human-robot interfaces that help strengthen an operator's situational awareness without inducing excessive cognitive load are crucial to the success of teleoperation. This paper presents an intuitive visual interface that allows operators to simultaneously control a 6DOF camera platform and a robotic manipulator whilst experiencing the remote environment through a virtual reality head-mounted-display (VR-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. Whilst orbital head-mounted display excelled in usefulness, the standard HMD interface allowed operators to perform robotic welding tasks significantly faster. Still participants reported they preferred OHMD due to the increased assistance and freedom in viewpoint selection.

Index Terms—Teleoperation, Virtual Reality, Head-Mounted Display, Camera Control, Visual Feedback.

I. INTRODUCTION

ROBOTIC teleoperation systems enable operators to execute dexterous tasks in hazardous remote environments such as nuclear facilities. In these environments, a human-in-the-loop system is required to perform unpredictable tasks such as on-site maintenance [1]. Telerobotic operations rely predominantly on visual feedback to provide operators with information about the remote process [2] [3]. As a result, visual support is critical to the success of teleoperation [4].

Conventional systems display several camera views on multiple monitors to enable the operator to fully monitor a remote process. However, research has shown that this method induces high cognitive load at the cost of teleoperation performance [5]. The operator needs to divide his attention across multiple displays, piece together the visual information, and perform mental transformations due to misalignment between the camera and robot reference frames [6] [7]. Furthermore, in real-world scenario's, cameras have to be placed in sub-optimal positions due to lack of access and obstructions that decrease the field of view. In fact, the robot belonging to the industrial process (hereinafter referred to as *task robot*) frequently occludes the cameras' line-of-sight [8].

A commonly used alternative is the so-called eye-in-hand (EIH) camera. This is a camera attached to the end-effector of the robot belonging to the process (hereinafter task robot). While such a setup allows more freedom in the selection of viewpoints, the viewpoint adaptions are coupled with the task-robot actions. When the task robot moves to perform an action, the viewpoint moves along with it. As a result, the operator needs to perform visual inspections and manipulation sequentially. Furthermore, during grasping or manipulation tasks the EIH-view is often occluded by an object that is being manipulator/grabbed. Additionally, such a camera does not provide sufficient information about the surrounding workspace [9].

In the past 10 years people started to adopt a secondary robot (hereinafter monitoring robot) to monitor the task robot, the process, and the remote workspace. Like many robotic systems, so too can the control of the monitoring robot take any shape between a completely manual approach and a fully autonomous one. Since controlling the monitoring robot is not a primary task its control interface should intuitive so that it does not overload the operator.

Wilde et al. [10] proposed a manually controlled robotic arm equipped with a camera to monitor a spacecraft docking task. Operators indicated that they preferred the system over a tradition eye-in-hand camera and experienced better spatial awareness. They also claimed that a more intuitive control interface for the monitoring arm and assistance in selecting viewpoints would better support them during the execution of the main docking task. In [11] the authors proposed an autonomous monitoring robot arm 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 depending on the state of the task robot [12]. An important contribution was made by Rakita et al. [13]. They also used a task manipulator and a monitor manipulator.

Contrary to the previously mentioned works, Rahnamaei et al. [14] proposed a shared control approach. They reasoned that because the optimal camera orientation is task dependent

S. Kuitert is with the Department of Biomechanical Engineering, Faculty 3mE; Delft University of Technology, 2628 CD Delft, The Netherlands, and with Heemskerk Innovative Technology B.V., 2629 HD Delft, The Netherlands (e-mail: sjoerdkuitert@hotmail.com)

J. Hofland and C. J. M. Heemskerk are with Heemskerk Innovative Technology B.V., 2629 HD Delft, The Netherlands

D. A. Abbink and L. Peternel are with the Department of Cognitive Robotics, Faculty 3mE; Delft University of Technology, 2628 CD Delft, The Netherlands

[15]–[17], the operator should control the camera orientation while an autonomous component moves the camera to prevent obstructed views. Furthermore, they used head-tracking to send velocity commands that rotate the camera. While the head-tracking input allowed users to operate the camera hands-free, the rotation of their head with respect to the display (i.e. view rotation) introduced additional cognitive load [5].

In [18] authors argued that operators should dictate task specific viewpoint adaptions by giving spare input to nudge a viewpoint algorithm towards specific solutions. Subsequently, the monitoring robot would move its camera towards that solution. Furthermore, the authors added a mode that allowed operators to take more control of the robot by manually selecting viewpoints to enable visual exploration of the remote site. In a user study, their new system proved to be better than their previous fully automatic system [13]. Participants reported they found the new system easier to use despite having to control more viewpoint parameters.

In the work of Sita [19], four challenges in the field of monitoring robots are outlined:

- If available, select and optimal viewpoint for the user based on the state of the task;
- Increase perception and understanding of the task without overloading the operator;
- Find a way to efficiently map the user control input when the camera viewpoint changes;
- Design a human-machine interface that allows for natural and efficient control of both camera viewpoint and the robot.

In this study, we will focus on the fourth challenge.

A. Research Objectives

To tackle this challenge, we propose a visual interface called Orbital Head-Mounted Display (OMHD) based on a third person free follow camera (i.e. a camera with a spherical constraint), like in [20] [21], and combine it with a head-mounted display (HMD) to control a camera viewpoint using head rotations which have been shown to be an effective method of controlling a cameras gaze [22]. The OHMD interface will be evaluated on task performance, perceived workload and interface acceptance in a one-factor within-subjects experiment design for a remote welding task of a cooling pipe in a simulated environment. We compare the OHMD interface to a conventional array of multiple fixed cameras (Array) and a standard head-mounted display interface.

B. Hypothesis

It is hypothesized that the Orbital HMD and the standard HMD interfaces improve task performance, reduce perceived workload, and increase self-reported interface acceptance during remote manipulation tasks compared to an array of multiple fixed cameras (H1). Furthermore, it is hypothesized that the Orbital HMD interface will improve self-reported interface usefulness compared to a standard HMD implementation, without sacrificing task performance and increasing the perceived workload (H2).
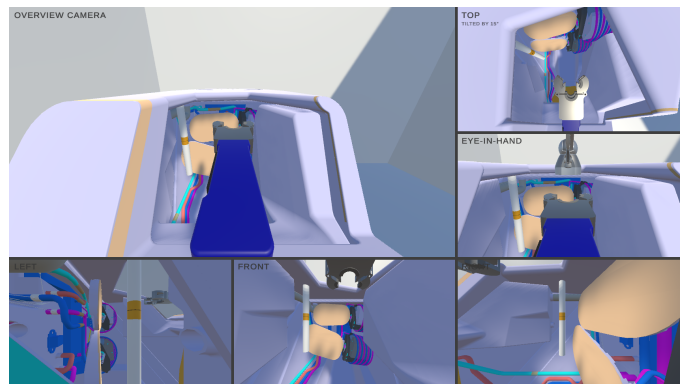

Fig. 1. A screen capture of a typical operator view when using the Array visual feedback method.

II. VISUAL INTERFACES

This section outlines the three visual interfaces that are compared in this study through simulation are outlined.

A. Array of Multiple Fixed Cameras

For remote teleoperation the standard approach to visual feedback is the use of multiple cameras placed in fixed positions in the remote environment. The camera feeds are then displayed to the operator across multiple monitors at the control station. In this study, the camera array comprises of 6 views displayed to the operator on a single monitor. The screen layout comprises of a 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. Lastly it includes the feed of an eye-in-hand camera.

Note 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. In virtual environments, however, cameras can be placed anywhere. Nevertheless, to increase the field of view of each camera and at the same time avoid occlusion culling (i.e. not rendering objects close to the camera), some of the cameras required a small translation and/or rotation with respect to their ideal configurations.

B. Standard Head-Mounted Display

In a typical Virtual Reality (VR) format, a user wearing a head-mounted display (HMD) views a stereoscopic image of a 3D environment through small embedded display optics located in front of each eye. During use, motion tracking sensors pick up the user's head movements and adjust the view accordingly. Thus, users can look around in a (virtual) 3D space as if they are actually there, experiencing changing viewpoints and perspectives that are related to their own head movements just like in real-life.
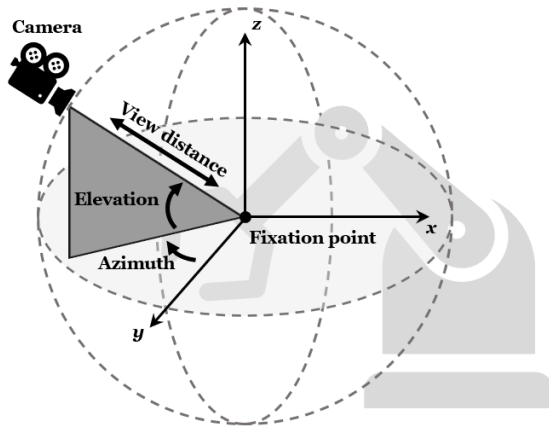
Fig. 2. A diagram illustrating the coordinate system used by the Orbital Head-Mounted Display method.

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. Like the name suggests, such a camera automatically positions itself relative to the player-character to provide a third-person view. As a result, the camera chases the player-character whenever it moves. Players can pan the camera around their character and change the view distance (i.e. zoom) using camera controls that are normally mapped onto a gamepad, keyboard or mouse.

Instead of using a conventional interface, OHMD uses inputs from motion tracking sensors of a HMD to control a free follow camera. Specifically, it utilizes the user's head rotation to set 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. As a result, the fixation point is always centered in the view. In this study the fixation point is located at the task robot's end-effector. The zoom functionality can be mapped onto any input device that has a single-axis (composite) binding (i.e. a pair of buttons or a joystick). If available, unused input bindings on the robot control interface can be used for this.

III. METHODOLOGY

To test our hypotheses, the effect of the visual interface on task performance, perceived workload and interface acceptance was evaluated in a one-factor within-subjects experiment design. During the experiment, participants use each visual interface described in Section II. to perform teleoperated welding tasks in a counterbalanced order. Thus there are three experiment conditions: Array, HMD and OHMD. In this section, the experiment setup, design, and procedure are summarized.

A. Participants

Twelve participants (10 male, 2 female) between 26 and 62 years old (M = 31.2, SD = 9.9) were recruited via e-mail. Participants included TU Delft students, family members, friends, colleagues and interns. 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. Each had healthy hand control and normal/corrected vision. The setup and experiments were approved by the Human Research Ethics Committee (HREC) of the Delft University of Technology.

B. Experimental Setup

The experimental setup used for each visual interface of this study is shown in Fig.3. For the Array interface (Fig. 3a), a standard LCD monitor was used to display visual feedback. The monitor was a 24.1 inch Dell IPS display, with a 1920x1200 resolution and 60Hz refresh rate placed 60-80cm from the participant. A VR-HMD was used for the HMD and OHMD interfaces (Fig. 3b,c). The VR-HMD was an Oculus Quest 2 128GB with a 3664x1920 resolution fast-switch LCD display (1832x1920 per eye), 72Hz refresh rate, 89° FOV, and adjustable interpupillary distance (58, 63 and 68mm). An NVIDIA GTX 1060 3GB was used to ensure an average frame rate of around 100.

A 3D Systems Touch haptic device was used to send movement commands to the robot end-effector. The device has 6DoF position sensing (stylus gimbal), 3DOF force feedback (translation only) up to 3.3N, and a workspace of 160x120x70mm (WxHxD). To control the camera during the VR conditions, the inside-out head tracking of the Oculus Quest 2 HMD was used. The joystick on the Oculus Motion Controller was used to control the view distance (i.e. zoom) during the OHMD condition.

C. Use Case

The use case that was selected for this study is a robotic welding task of a cooling pipe within the Upper Port Launcher (UPL) of the ITER nuclear fusion reactor. Due to the high risk, complex, and hazardous environment this is a typical scenario in which teleoperation is used. The use case helps to motivate the task requirements, and subsequently the metrics used to evaluate the performance of each visual feedback method, such as collision avoidance, workload, and usability. Furthermore, the UPL represents a real-world environment in which acquiring effective views is challenging due to the limited space and the presence of (visual) obstacles.

D. Software and Simulated Environment

In this study participants conducted teleoperation tasks in a virtual environment built in the Unity3D engine (v2021.1.1.19f1). In this section, the core features of the simulation are described.

1) Environment: In the experiment, the participants performed remote manipulation tasks in a virtual environment. As such, a Unity scene that replicates a hot cell (i.e. radiation containment chamber) that contains a 3D model of the UPL was created. All physics interactions between objects are computed by the Unity Physics engine using a Temporal Gauss Seidel solver.
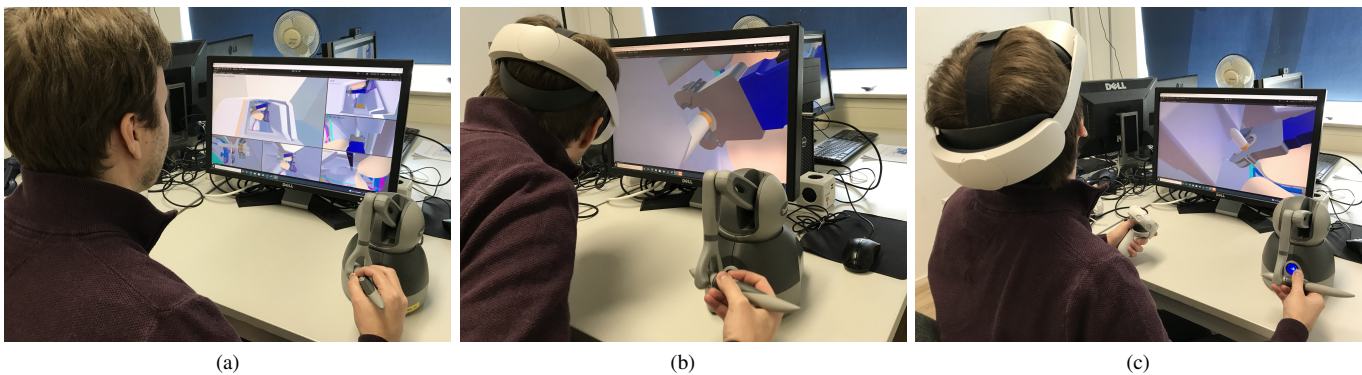
(a)   (b)   (c)

Fig. 3. Pictures of the experimental setup. In each image the participant uses the Touch haptic device in combination with one of the visual interfaces to perform the robotic welding task. (a) Array; the participant looks at multiple camera feeds displayed on a single monitor. (b) HMD; the participant wears a VRHMD to perceive the virtual environment and interact with the camera. (c) OHMD; In addition to the HMD setup, the participant holds a motion controller to zoom in and out with the joystick.

2) Welding Tool: To weld the cooling pipes inside the UPL a modified version of the ORBIWELD 38S TIG orbital welding tool is used. A CAD model with a custom three-part clamping mechanism and convex collision meshes was made using SOLIDWORKS and Fusion360 and added to the simulation environment. The operator commands the movement of the tool as if its handle is connected to the end-effector of a robotic manipulator.

3) Haptic Control Implementation: The Touch haptic device is implemented in Unity3D using a custom plugin that connects to the device's drivers through the OpenHaptics software package. Participants need to press and hold the front button on the Touch's stylus to couple the stylus and the tool. Once coupled, information of the stylus' pose is used to control the movement of the tool with respect to the world coordinate frame (i.e. relative to a front view of the UPL). When the button is released, the stylus and tool are decoupled, allowing users to reset the Touch's workspace (i.e. re-index or clutch) and take a comfortable hand/wrist posture. Note, however, that the translational movement input is scaled such that the welding task can be completed without re-indexing.

4) Virtual 3D Systems Touch: When wearing a VR headset, the operator cannot see the physical Touch device. Because of this it is very difficult to recognize the Touch's pose, especially while re-indexing. To aid the participant, a virtual model of the Touch appears above the tool while the operator is re-indexing (Fig. 4). The model is aligned with the world coordinate frame and mimics the configuration of the physical device. Even though the physical Touch is on the desk beside the operator, it falls outside his/her field of view when he/she is looking at a monitor. Therefore, the virtual Touch is also displayed in the EIH camera view.

E. Task Description

Each trial required the participants to move the welding tool from a set starting point towards a cut in a cooling pipe indicated by an orange marker (Fig. 5a). Once close-by, they had to fit the profile of the tool's head around the pipe-cut. When the alignment was within the margin of error (5mm for x,y,z, and 2° for pan and tilt) a visual cue (Fig. 5) appeared to notify participants they had completed the task.

Fig. 4. Image of the virtual Touch. The virtual Touch appears above the welding tool when the human-operator is re-indexing (i.e. clutching) to show the configuration of the physical Touch haptic device.
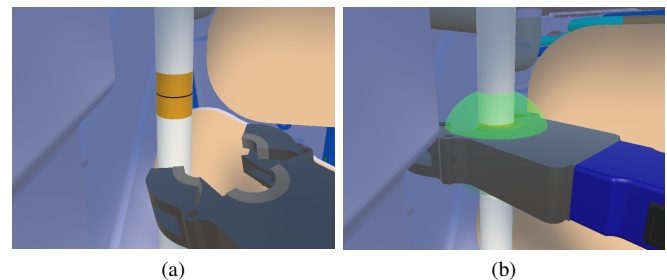

(a)   (b)

Fig. 5. (a) Visual indication of the pipe cut location. (b) Visual cue to indicate the alignment is within task success margins.

Cooling pipe placements varied between 3 locations (Fig. 6) to introduce different levels of complexity, and to change the visual information provided by the camera views [23]. The risks of damaging the components of the UPL and/or the robot were explained to the participants. They were then told to perform the task as fast as possible, but to avoid collisions with any components other than the cooling pipe, and to minimize the force exerted by the welding tool onto the pipe.

F. Data Acquisition and Metrics

Objective data, including simulation data and device read-outs, was recorded within the Unity simulation at a rate of 50Hz. Self-reported measures were collected through ques-
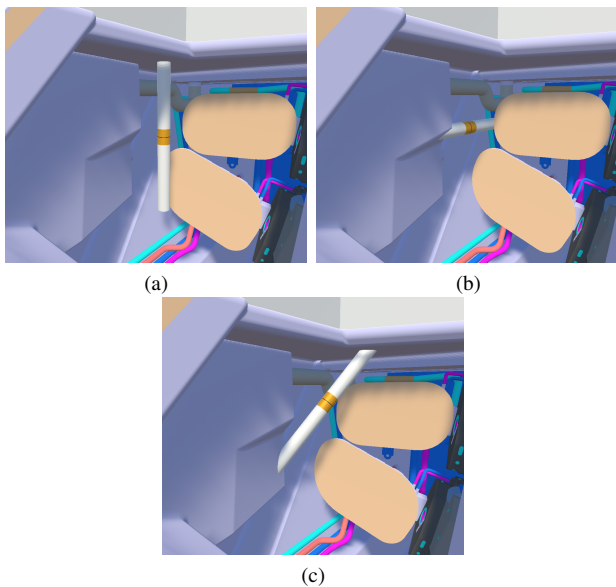
Fig. 6. Figure showing the different cooling pipe locations. (a) Pipe A. (b) Pipe B. (c) Pipe C.

tionnaires. The following metrics were used to assess the three visual interfaces:

1) Interface acceptance:

- *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 [24] 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. The usefulness score derives from 5 components: average of useful/useless, good/bad, effective/superfluous, assisting/worthless, and raising alertness/sleep-inducing. The satisfying score derives from 4 components: pleasant/unpleasant, nice/annoying, likeable/irritating, and desirable/undesirable. Usefulness and satisfying scores are calculated by averaging the respective components.

2) Perceived workload:

- *Task load rating*. The NASA Task Load Index (TLX) questionnaire [25] was used to capture the subjectively 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 sub-scales: mental demands, physical demands, temporal demand, own performance, effort and frustration.

3) Performance:

- *Time to completion (ttc)*. The amount of time it takes a participant to complete the remote welding task. Effective views enable operators to perform teleoperation tasks faster. However, participants may require additional time to select a view, control the camera or process the visual information.
- *Number of collisions*. The amount 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.

G. Experimental Procedure

A flowchart of the experimental procedure is shown in Fig. 7. Prior to commencing the experiment, participants were briefed on the goal of the study, gave written consent and were handed a pre-experiment questionnaire with questions regarding personal details and prior experience. Then they were asked to read the experiment instructions and to examine the Acceptance Scale and NASA-TLX questionnaires to get familiar with the scales. Subsequently, the researcher checked for understanding by asking questions and reciting key points of the instructions.

Next, participants were allowed 5 minutes to get acquainted with the simulation environment, the force feedback and the haptic device used to control the end-effector. Note that during this session they were not exposed to any of the experimental conditions. Instead, they were using a single view on a monitor and practiced on a cooling pipe that was positioned in a unique location.

Afterwards, participants performed the following steps for each condition. First, they were given 3 minutes to acclimate to the visual interface and camera controls. Then, for the three pipe locations, participants performed a practice trial followed by three recorded trials. Subsequently, they were asked to complete the Acceptance Scale and NASA-TLX questionnaires before moving on to the next condition. 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. Once they were finished with all the recorded trials for all conditions, they were asked to fill out a post-experiment questionnaire with questions regarding their experience of the study and their preferred visual interface.

H. Data Analysis

To analyze our results repeated measures analysis of variance (RM-ANOVA) was used. 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 subject. 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. Non-parametric data and parametric data that violates ANOVA assumptions, like normality, was transformed using the Aligned Rank Transform (ART) [26] [27] to allow the use of a RM-ANOVA.
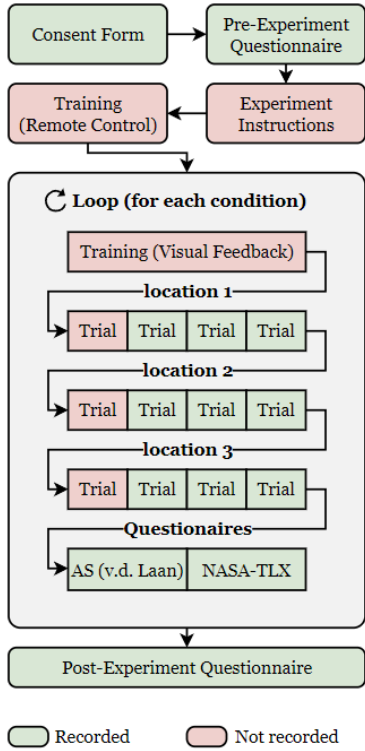
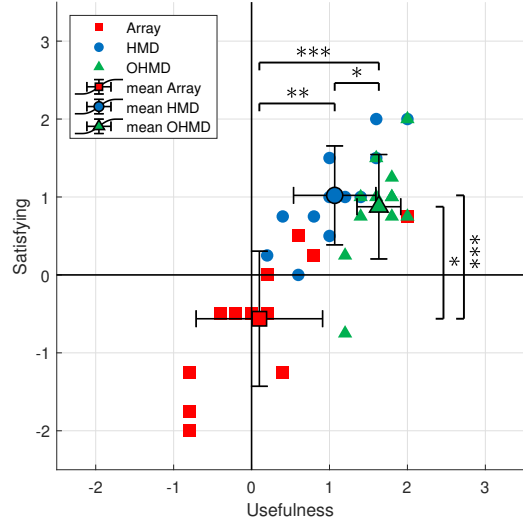Fig. 7. Diagram showing the experiment procedure.

Fig. 8. Overall Van der Laan's acceptance scores [24] 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.

IV. RESULTS

One-way RM-ANOVA's were conducted on metric data to asses for significant differences between the interfaces. A summary of the results is given by Table I. Observed differences were considered statistically significant at p-values of 0.05 or less. Hereinafter, 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.

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 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 self-reported perceived workload (Fig. D.1) 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 objective time to completion and number of collision metrics violated the normality assumption according to the Shapiro-Wilk test. Thus the ART transformation was used to transform the data so that the standard RM-ANOVA procedure could be applied to test for significant differences. The results for the objective performance metrics are presented in Fig. 10, Table I and Table II.

For time to completion, a significant effect was found, $F(2, 22) = 48.77, p < 0.001$. Interestingly, participants were

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

Metric		Visual interface			RM-ANOVA, (F2,22)	Post-hoc analysis		
		Array	HMD	OHMD	Effect: interface	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 [28]. This data was transformed using an *Aligned Rank Transform* (ART) for multifactor contrast tests [26] [27] so that the ANOVA and post-hoc procedures could be applied.
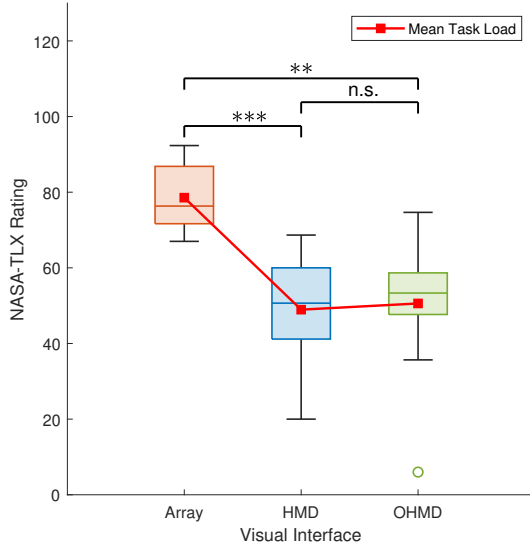

Fig. 9. Weighted NASA-TLX ratings for each visual interface. NASA-TLX values range between 0 and 100, with higher values indicating higher task load. The marks '∗ ∗ ∗', '∗∗', and '∗' denote a significance of $p \leq 0.001$, $p \leq 0.01$, and $p \leq 0.05$, respectively, and 'n.s.' denotes no significance.

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 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 a 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

This study proposed a novel visual interface that uses the motion tracking and stereoscopic display capabilities of an HMD to control the view provided by a free-follow camera platform. The interface was designed to provide operators the ability to acquire more effective views during teleoperation without sacrificing task performance and causing cognitive overload.

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 load. 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 avarage.

The results outlined in Section IV will be discussed in this section. Answers of post-experiment questionnaires, participant comments during the experiment, and observations made by the researcher were used to motivate the discussion of these results. Furthermore, a small supplementary analysis on learning effects is included to discuss participants claims.

A. Effect on Interface Acceptance

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
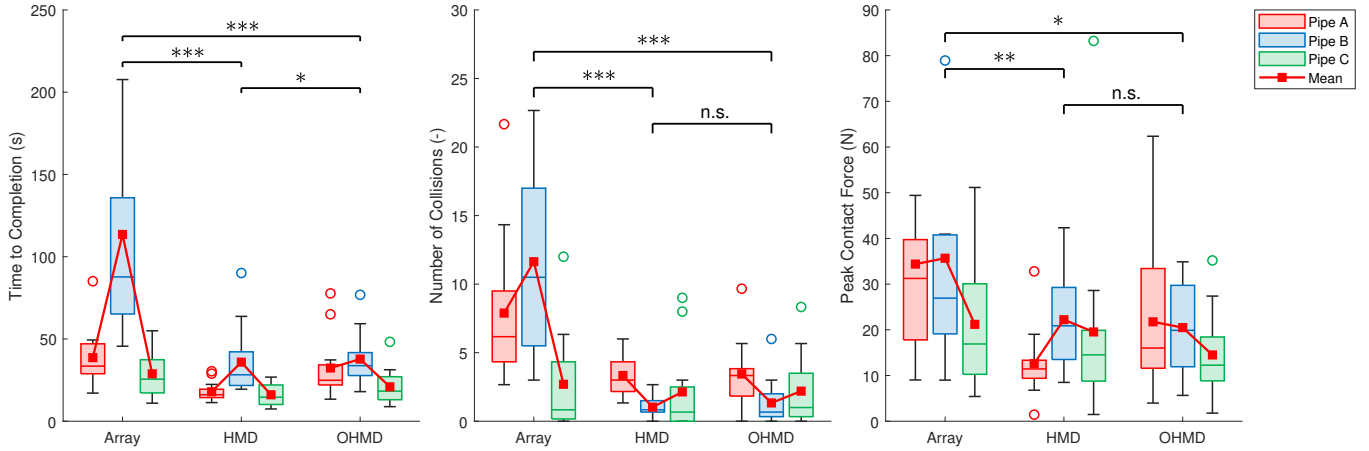
Fig. 10. Results for the objective task performance metrics. The significance lines above the box-plots indicate differences in mean between the visual interfaces. The marks '∗∗∗', '∗∗', and '∗' denote a significance of $p \leq 0.001$, $p \leq 0.01$, and $p \leq 0.05$, respectively, and 'n.s.' denotes no significance.

TABLE II
MEAN VALUES AND STANDARD DEVIATIONS OF ALL OBJECTIVE PERFORMANCE METRICS ACROSS ALL INTERFACE AND PIPE LOCATION COMBINATIONS.

Factors		Completion Time (s)		Number of Collisions (-)		Peak Contact Force (N)	
Interface	Pipe	Mean	Std. D.	Mean	Std. D.	Mean	Std. D.
Array	A	38.65	±17.64	7.89	±5.43	34.38	±25.51
Array	B	113.54	±77.49	11.64	±7.10	35.67	±25.67
Array	C	28.82	±15.37	2.69	±3.66	21.20	±15.27
HMD	A	17.92	±6.12	3.33	±1.54	12.63	±7.62
HMD	B	35.90	±21.43	1.03	±0.78	22.22	±10.52
HMD	C	16.07	±6.49	2.14	±3.13	19.54	±21.39
OHMD	A	32.34	±19.30	3.44	±2.45	21.75	±16.63
OHMD	B	37.78	±16.35	1.33	±1.76	20.49	±10.44
OHMD	C	20.93	±11.15	2.19	±2.62	14.50	±9.24

interface. However, they also reported it was less pleasant to use and not as likable. This aligns with participant responses in the post-experiment questionnaire. The acceptance of the HMD interface aligned with results from other studies [2].

Two participants mentioned they suffered from minor nausea after using the OHMD interface. However, they did not feel nauseous while using the standard HMD interface. It is worth mentioning that both of these participants reported to have little prior experience with virtual reality (approximately 1 hour). The visually induced motion sickness (VIMS) is likely caused by the illusory of self motion (i.e. vection). When the participant rotates his/her head, the camera rotates ánd translates. The perceived camera movement does not match with a previously stored experience (i.e. sensory pattern). Hence there is a sensory conflict which causes VIMS [29], [30].

There are differences between participants that may cause certain individuals to be more susceptible to motion sickness. [29], [30]. This could also explain why the participants did not suffer from nausea when using the standard HMD interface, as its active camera control method is close to the human vision (i.e. a known pattern). Therefore, as the participants get more accustomed to OHMD, the sensory conflict may reduce and therefore reduce the change of motion sickness occurring [31] [32].

Another source of discomfort that was reported by participants is camera shake while using the OHMD method.

Further analysis has found three possible explanations. Firstly, participants seem to make small unconscious head movements. When the participant is zoomed out, these small head movements cause relatively large camera translations. Furthermore, participants don't hold the stylus of the haptic device perfectly still. Thus, the welding tool linked with the stylus is also shaking unless the clutch is engaged. Once more these small movements are exaggerated, this time by the input scaling. Finally, the motion sensors in the HMD have some noise. In future, these effects can be mitigated by applying a filter to the input (e.g. a dead-zone of frequency filter) or by introducing a moderate amount of damping.

B. Perceived Workload

According to the NASA-TLX questionnaire workload components, the Array interface caused a lot of frustration, induced high mental workloads and furthermore the trials took a lot of effort to complete. This aligns with findings in literature [5]. While the mean weighted task load rating for the OHMD and HMD interfaces was similar, the individual components of the NASA-TLX questionnaire show some intriguing results. Notably, participants rated that the OHMD interface required more physical demand and effort than the standard HMD display implementation. During the experiments, very few participants where actively moving while using the HMD interface. The stereoscopic view of the HMD already provided

enough contextual and depth information. This result aligns with findings in [9] who found that for most task a passive (i.e. fixed) stereoscopic view is sufficient, and that only more complex tasks require dynamic views. This still leaves the question why OHMD induced more physical demand. Firstly, because the OHMD viewpoint followed the welding tool, the operator was forced to readjust the camera position in order to avoid occlusions. Because of this, the researcher observed that the operators were more involved in the visual task (i.e., they were more actively looking for better views. This also explains the higher 'Raising Alertness' score in the Acceptance Scale results as a dynamic view provides more depths cues such as motion parallax and optical flow [33]–[35].

To keep robot control consistent between the camera systems, the robot's control frame was aligned to the world frame of the simulation environment. However, for OHMD's free follow camera and the Array's eye-in-hand camera it may be preferable to continuously transform the robot's Cartesian motion coordinates such that they align with the axes of the camera view. In-fact, this was suggested by four participants because they had difficulty understanding the link between the robot control input and the welding to movement, particularly its rotations. Some participants even stopped re-indexing, valuing the understanding of the control link over a comfortable hand posture.

C. Task Performance

Several participants experienced difficulty in locating the contact points during collisions due to the lack of torque force feedback and auditory cues. This was especially troublesome when participants attempted to adjust the position of the welding tool after it had been fitted around the pipe. In such an instance, the rotational movement of the tool around its longitudinal axis (i.e. roll) is constrained. When the operator attempts to roll the tool, a torque is generated as the tool's jaws start to grip the pipe. The tool hasn't visually rotated at this point, so it was hard for the operators to understand what was going on. Of course this negatively affected the task completion time.

In the post-experiment questionnaire, participants reported that they found the trials very difficult to complete when using the Array system and that they considered pipe location B (Fig. 6) to be most difficult. The results indeed show that the simulated welding task for pipe location B took the most time to complete regardless of which interface was used. However, when looking at the number of collisions, pipe B seemed to be easier than pipe A and pipe C when the OHMD or HMD was used. Because the pipe 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 very difficult to understand the link between the robot control input and the eye-in-hand camera view. Just like described in literature, participants were focused all their attention the the immersive egocentric view provided by the eye-in-hand camera even though the view was ineffective [36], [37].

Interestingly, when the number of collisions are plotted against time, there appears to be a positive, yet insignificant, correlation. This means that participants that took more time to complete the task did not cause less collisions. Some participants even mentioned they were still getting used to the OHMD system during the trials. According to them, more extensive training would increase their task performance.

D. Limitations and Future Work

This study had several limitations. Firstly, the experiment took participants 2 hours to complete on average. Because of this, the number of trials recorded for each combination of system and pipe location had to be limited to 3 trials. This made it impossible to identify a significant learning effect in the objective data that supports the claim of participants that they needed more training. A study with more participants and a between-participant design would allow each participant to complete more trials and also remove cross-over effects.

Additionally, the OHMD interface has only been evaluated for a single task robot control interface; the 3D systems Touch. While such a device is common for robot teleoperation, there might be other interfaces that have better synergy with OHMD.

Furthermore, during the trials the robotic manipulator that holds the tool was not rendered due to the lack of a robust inverse kinematics controller with collision- and joint limit avoidance. In real world scenario's the robotic manipulator would frequently block camera views. This also meant that one of the key advantages of the OHMD design, its freedom in selecting viewpoint, was less essential for successful task completion. In theory interface designers could actively choose not to render the robotic manipulator in 3D virtual copies or reconstructions of the remote environment. However, monitoring the robot's behavior is important to ensure safe task execution.

Also, the cameras in the simulation were not modeled as physical objects and were not constrained in their movement. This meant that in practice participants were able to move the camera to positions that a monitoring robot would not be able to reach. Adding camera movements constraints would be a nice step towards applying the system on a real-world monitoring robot.

The parameters of the human-machine interface, such as the OHMD camera movement sensitivity, task robot control sensitivity, and task robot control frame of reference, were set according to results of a small pilot study. Different parameter settings could affect the results. A more extensive study that investigates the effect of these parameters on task performance may improve the system. Based on the analyses of the results and the questionnaires, the following features will potentially improve the system:

- Allow operators to set the camera movement sensitivity to their preference.
- Allow operators to change the task robot control frame of reference from the world coordinate system to the OHMD camera's local coordinate system.
- Introduce a dead-zone or moderate damping to reduce camera shake.

Acknowledgments

The author greatly acknowleges the help and support from Prof.dr.ir. D.A. Abbink, dr.ir. L. Peternel, ir. J. Hofland and dr.ir. C.J.M. Heemskerk. This work was done in collaboration with Heemskerk Innovative Technology B.V. (HIT). The company HIT also provided the hardware required for the experiment.

VI. Conclusion

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

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

The results from this study indicate that the OHMD interface has potential benefits compared to a standard HMD. Participants preferred the OHMD interface due to its increased flexibility in viewpoint selection despite reporting they needed more time to take full advantage of the its capabilities. Hence, for tasks that aren't time critical, OHMD is a valid alternative to a standard HMD. While interface designers should be hesitant to use OHMD in its current state of development over a standard HMD setup, the interface does have some inherent attributes that might benefit teleoperation in other applications. For example, the follow-constraint makes it suitable for teleoperating mobile robots or manipulators with a large workspace. To conclude, this study takes an important step towards integrating natural and intuitive control inputs like head movement in human-machine interfaces with the goal of improving situation awareness and remote operator performance.

References

[1] A. Loving, P. Allan, N. Sykes, S. Collins, and P. Murcutt, "Development and application of high volume remote handling systems in support of JET and ITER," *Fusion Engineering and Design*, vol. 87, no. 5-6, pp. 880–884, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.fusengdes.2012.02.039

[2] E. Triantafyllidis, C. McGreavy, J. Gu, and Z. Li, "Multimodal Interfaces for Effective Teleoperation," *arXiv*, mar 2020. [Online]. Available: http://arxiv.org/abs/2003.14392

[3] J. L. Burke, M. S. Prewett, A. A. Gray, L. Yang, F. R. Stilson, M. D. Coovert, L. R. Elliot, and E. Redden, "Comparing the effects of visual-auditory and visual-tactile feedback on user performance: A meta-analysis," *ICMI'06: 8th International Conference on Multimodal Interfaces, Conference Proceeding*, no. January, pp. 108–117, 2006.

[4] 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, aug 2014. [Online]. Available: http://www.mdpi.com/2218-6581/3/3/310

[5] B. P. DeJong, J. E. Colgate, and M. A. Peshkin, "Mental Transformations in Human-Robot Interaction," Tech. Rep., 2011.

[6] 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.

[7] M. S. Prewett, R. C. Johnson, K. N. Saboe, L. R. Elliott, and M. D. Coovert, "Managing workload in human–robot interaction: A review of empirical studies," *Computers in Human Behavior*, vol. 26, no. 5, pp. 840–856, sep 2010. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0747563210000506

[8] D. Clark, "The Effect of Camera Placement and Display Configuration on a Remote Manipulation Task," 2006. [Online]. Available: http://commons.erau.edu/db-theses/33/

[9] S. YU, J. LEE, 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. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1738573315301170

[10] 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. [Online]. Available: http://ieeexplore.ieee.org/document/6705697/

[11] 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.

[12] 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.

[13] D. Rakita, B. Mutlu, and M. Gleicher, "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. [Online]. Available: https://dl.acm.org/doi/10.1145/3171221.3171279

[14] 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. [Online]. Available: http://link.springer.com/10.1007/s10846-014-0028-7

[15] G. T. McKee and P. S. Schenker, "Visual acts for remote viewing during teleoperation," *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 1, pp. 53–58, 1995.

[16] X. Xiao, J. Dufek, and R. R. Murphy, "Tethered Aerial Visual Assistance," *arXiv*, jan 2020. [Online]. Available: http://arxiv.org/abs/2001.06347

[17] A. Saran, B. Lakic, S. Majumdar, J. Hess, and S. Niekum, "Viewpoint selection for visual failure detection," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, vol. 2017-Septe. IEEE, sep 2017, pp. 5437–5444. [Online]. Available: http://ieeexplore.ieee.org/document/8206439/

[18] D. Rakita, B. Mutlu, and M. Gleicher, "Remote Telemanipulation with Adapting Viewpoints in Visually Complex Environments," in *Robotics: Science and Systems XV*. Robotics: Science and Systems Foundation, jun 2019. [Online]. Available: https://github.com/uwgraphics/relaxedhttp://www.roboticsproceedings.org/rss15/p68.pdf

[19] E. Sita, "Monitoring Companion for Industrial Robotic Processes," Ph.D. dissertation, University of the West of England, 2021. [Online]. Available: https://uwe-repository.worktribe.com/output/6004547

[20] 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. [Online]. Available: https://ieeexplore.ieee.org/document/8968218/

[21] E. Sita, T. Thomessen, A. G. Pipe, F. Dailami, and M. Studley, "Robot Companion for Industrial Process Monitoring Based on Virtual

Fixtures," in *IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society*, no. l.   IEEE, oct 2018, pp. 6051–6056. [Online]. Available: https://ieeexplore.ieee.org/document/8591105/

[22] D. Zhu, T. Gedeon, and K. Taylor, ""Moving to the centre": A gaze-driven remote camera control for teleoperation," *Interacting with Computers*, vol. 23, no. 1, pp. 85–95, jan 2011. [Online]. Available: https://academic.oup.com/iwc/article-lookup/doi/10.1016/j.intcom.2010.10.003

[23] 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. [Online]. Available: http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1009107

[24] 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. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0968090X96000253

[25] 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. [Online]. Available: http://journals.sagepub.com/doi/10.1177/154193120605000909

[26] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins, "The aligned rank transform for nonparametric factorial analyses using only anova procedures," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.   New York, NY, USA: ACM, may 2011, pp. 143–146. [Online]. Available: https://dl.acm.org/doi/pdf/10.1145/1978942.1978963?casa_token=nkCTUlppBoYAAAAA:eoddwwvko9Upk7uy3xwb_6JSMHEYEqcpWxE6vsePd0nQ7xquhJmlBFr_8AVeYyJAIdlryCbFzEVL4Qhttps://dl.acm.org/doi/10.1145/1978942.1978963

[27] 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.

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

[29] B. Keshavarz, B. E. Riecke, L. J. Hettinger, and J. L. Campos, "Vection and visually induced motion sickness: how are they related?" *Frontiers in Psychology*, vol. 6, no. APR, pp. 1–11, apr 2015. [Online]. Available: http://journal.frontiersin.org/article/10.3389/fpsyg.2015.00472/abstract

[30] J. T. Reason, "Motion Sickness Adaptation: A Neural Mismatch Model," *Journal of the Royal Society of Medicine*, vol. 71, no. 11, pp. 819–829, nov 1978. [Online]. Available: http://journals.sagepub.com/doi/10.1177/014107687807101109

[31] Y. Mizukoshi, R. Sato, T. Eto, M. Kamezaki, A. Matsuzaka, L. Yang, A. Namiki, A. Imai, T. Matsuzawa, K. Hashimoto, A. Takanishi, and H. Iwata, "A Low Cognitive Load and Reduced Motion Sickness Inducing Zoom Method Based on Typical Gaze Movement for Master-Slave Teleoperation Systems with HMD," in *2020 IEEE/SICE International Symposium on System Integration (SII)*.   IEEE, jan 2020, pp. 28–33. [Online]. Available: https://ieeexplore.ieee.org/document/9026260/

[32] Y. Chen, X. Wang, and H. Xu, "Human factors/ergonomics evaluation for virtual reality headsets: a review," *CCF Transactions on Pervasive Computing and Interaction*, vol. 3, no. 2, pp. 99–111, 2021. [Online]. Available: https://doi.org/10.1007/s42486-021-00062-6

[33] G. Smets and K. Overbeeke, "Trade-off between resolution and interactivity in spatial task performance," *IEEE Computer Graphics and Applications*, vol. 15, no. 5, pp. 46–51, 1995. [Online]. Available: http://ieeexplore.ieee.org/document/403827/

[34] 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. [Online]. Available: https://direct.mit.edu/pvar/article/18/4/304-320/18754

[35] G. P. Bingham and C. C. Pagano, "The necessity of a perception–action 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. [Online]. Available: http://doi.apa.org/getdoi.cfm?doi=10.1037/0096-1523.24.1.145

[36] 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, oct 2001. [Online]. Available: http://journals.sagepub.com/doi/10.1177/154193120104500415

[37] C. D. Wickens, "Multiple resources and performance prediction," *Theoretical Issues in Ergonomics Science*, vol. 3, no. 2, pp. 159–177, jan 2002. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/14639220210123806

APPENDIX

APPENDIX A
UNITY PROJECT

The Unity3D scene used in the simulation is meant to replicate a hot cell (i.e. a radiation containment chamber). The hot cell contains a model of the front part of the Upper Port Launcher (Fig. A.1a). The environment is kept relatively plain in order to optimize for frame-rates. During the trials, the operator is only focused on the inside part of the port plug. Here, more details are added. There are several game-objects that act as obstacles during the experiment. In Fig.A.1b, a cooling pipe with an orange marker that indicates the location of the pipe-cut is located behind (into the plane) the welding tool. During the experiments, the participants are allowed to collider with the cooling pipe, as it is the target of the manipulation. All collisions between the welding tool and other object in the environment are track and contribute to the final number of collisions metric for the active trial. For more details about the task environment, please check the participant instruction (Appendix C.2).

A. Unity scene

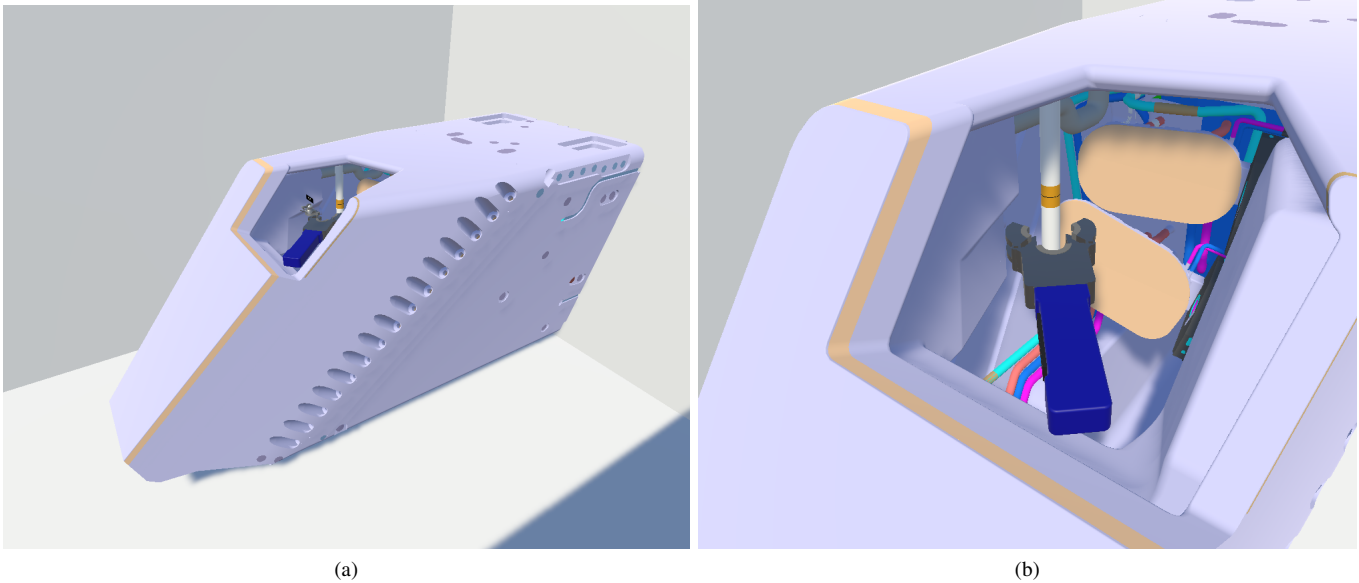(a)   (b)

Fig. A.1. (a) An overview show of the simulated task environment (b) An outside camera looking into the port plug.

B. Custom implementation of the OpenHaptic API in Unity

Because this study involves a remote manipulation task, accurate physics are needed to render good haptic feedback. In Unity3D 2021, several additional physics options and scripts are available that are more suitable for simulating contact transitions. To make use of these physics features but still be able to use a haptic device for robot control, either a lot of existing code needed to be altered, or a new plugin for the haptic device had to be made. Since the available plugin from OpenHaptics contains scripts with a lot of dependencies, it is hard to change or add a feature without breaking another. Therefore, all but the *.dll* file that allows Unity to call functions from the OpenHaptics API were removed. Then, three scripts were written each with their own functionality. These are outlined below.

1) Haptic Settings: This scripts manages the connection between Unity and the 3D Systems Touch through the Openhaptics software (Fig. A.2). The script first creates a connection with the device and subsequently reads the configuration. This includes button states, position of the stylus, orientation of the stylus and several transformations. These parameters are made available to other scripts. Including the object that hold the connection.
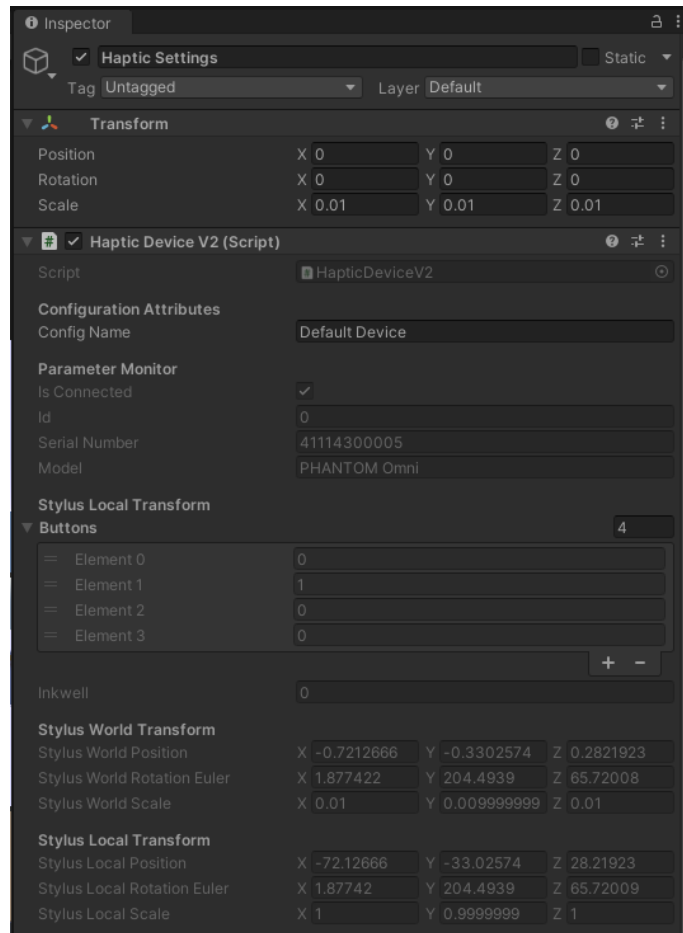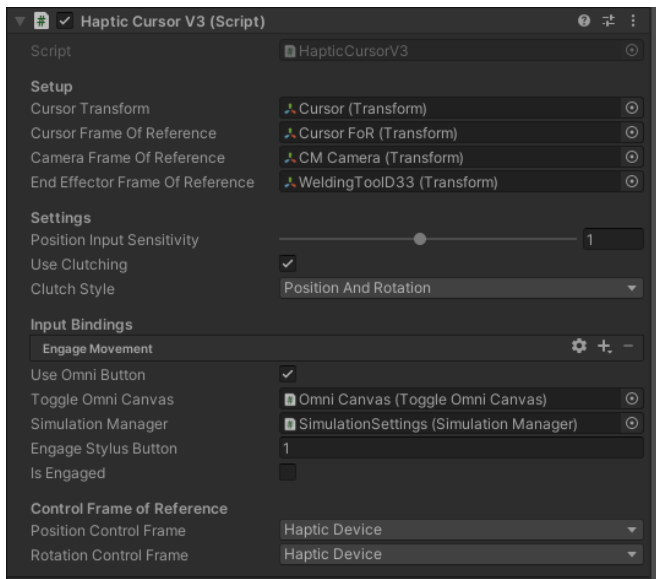

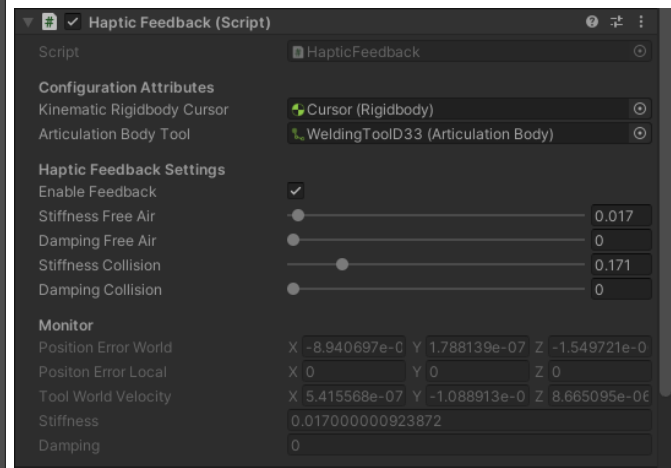Fig. A.2. Unity editor view of the Haptic Settings script.

2) Haptic Cursor: Next a script was written to use the omni readouts to control a rigidbody object (i.e. virtual cursor). This script includes several attributes that can be selected in the editor (Fig. A.3. These include an input sentivity, control reference frame, and a clutch button that enables and disables the input.

C. Haptic Feedback

Finally, a script was written to control a tool and send haptic force to the device. An added feature compered to the standard OpenHaptics plugin is that the virtual spring stiffness of the spring that connects the tool with the cursor can be set to different values depending on whether the tool is colliding or not. So when the tool is in free air, the spring can be set very stiff to get a very snappy response to the input, but then when the tool collides the spring can be set to a lower stiffness such that the tool becomes compliant. Force of the virtual spring are directly sent to the haptic device to provide haptic feedback to the operator.

Fig. A.3. (a) Unity editor view of the Haptic Cursor script. (b) Unity editor view of the Haptic Feedback script.

D. Virtual Touch haptic device

During preliminary testing of the Orbital-HMD concept it soon became apparent that commanding the movement of a robotic manipulator with a 3D systems Touch haptic device was difficult while wearing a head-mounted display. To solve this issue, a virtual replica of the Touch was developed. First, CAD drawings of an older generation Phantom Omni were downloaded[1]. Then, the model was edited in SOLIDWORKS. Some of the major changes involved changing the coordinate frames of each part to joint origins to make the model compatible with Unity3D's articulation bodies. Additionally, some buttons were added to the stylus. Next, the final model was imported to Unity3D and put together using articulation bodies. Figure A.4a shows the model of the virtual Touch device.

Next, custom scripts where written to add the following functionalities (Example in Fig. A.4b):

- A tracker script that updates the virtual omni's stylus to match the state of the real-world Touch.
- A button script that added button functions.
- A script to set the reference frame. For example, the world coordinate frame of the camera's reference frame.
- A UI script that enabled the Virtual Omni to spawn as a 3D 'hologram' in VR at a specified position.
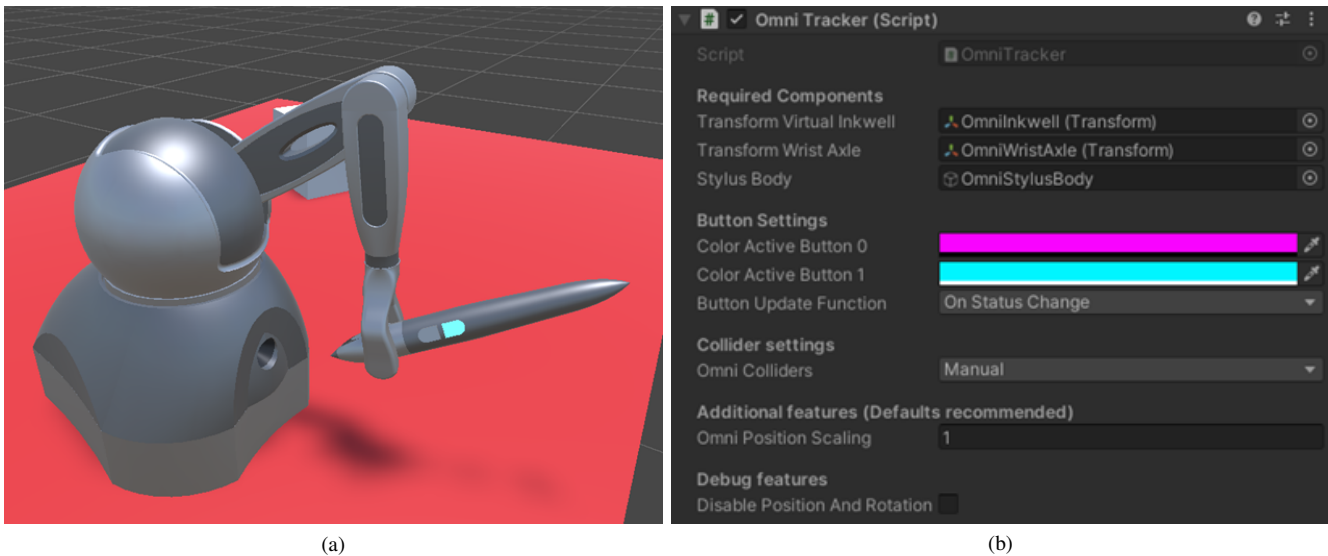- A script that hides/reveals the Omni depending on whether the operator is clutching (i.e. re-indexing)

(a)   (b)

Fig. A.4. (a) Image of the virtual Touch haptic device. (b) Image of the script in the Unity editor.

---

[1]https://grabcad.com/library/omni-phantom-1

E. ORBIWELD enclosed welding tool

For the remote robotised welding task at ITER a tool similar to the ORBIWELD 38S[2] is used. This is a TIG pipe welding tool that is specifically designed for applications where space is very restricted. The tool is supposed to be clamped onto the pipe that needs welding. By doing so, an enclosed chamber is created. Gas is pumped into this chamber as the arc welding process starts and the electrode starts to turn (i.e. orbit). Unfortunately, the design of the clamps of the original product made it unsuitable. Therefore, a CAD model of the tool with a custom clamp system was created in SOLIDWORKS (Fig. A.5.a). To ensure the welding tool would interact with the Unity3D physics, separate convex colliders needed to be assigned to the tool's assemble as shown in Fig. A.5.b. The tool was then exported as `.FBX` and imported into the Unity3D project. Using Unity3D's articulation bodies the various components of the tool were put together. Additionally, a script was written to be able to open and close the clamps so the tool could latch itself onto a pipe.
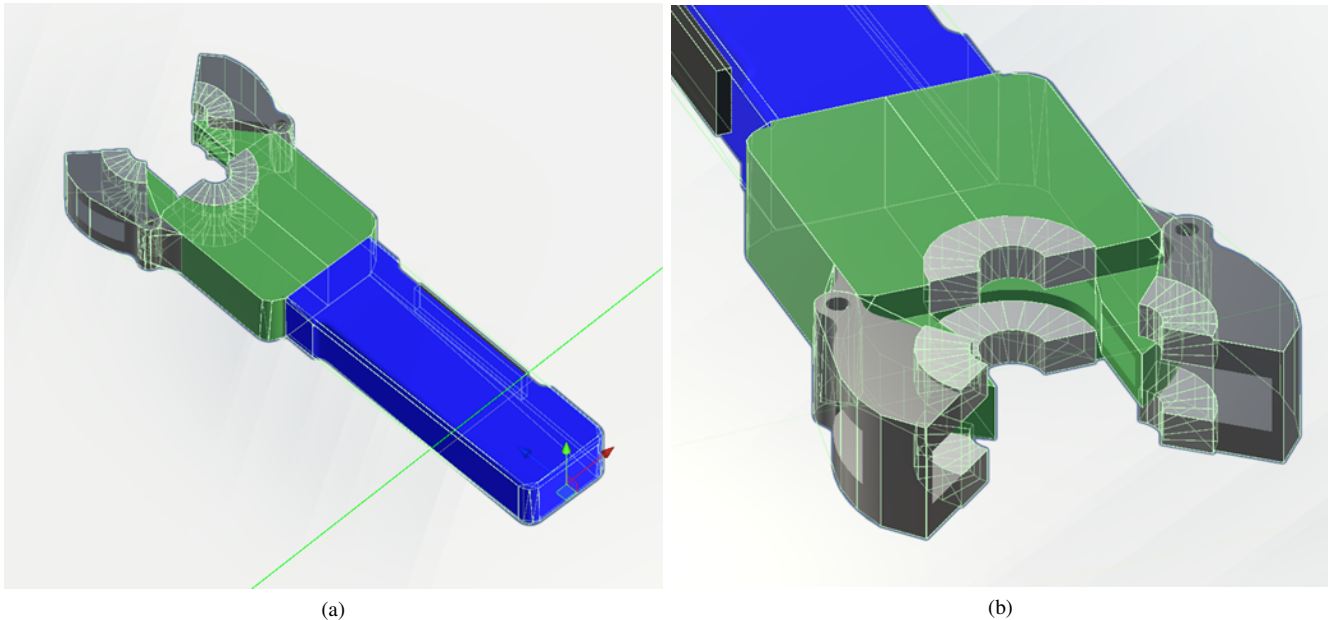

(a)

(b)

Fig. A.5. (a) Image of the custom ORBIWELD welding tool CAD model. (b) Image showing the convex collision meshes of the custom ORBIWELD CAD model.

F. Data Logger

To record objective data in unity, a custom data log plugin was written using the CsvHelper .NET library [3]. A manager script was written start logging, stop logging, and saving data-logs to files. The implementation is very scale-able and allows users to create their own data-logs by creating a new class object containing the parameters than need to be logged. Additionally, it requires an update function that lets the plugin know which game-objects in the project contain the required data..

[2]https://www.orbitalum.com/en/products/product-details/orbiweld_ow_38s.html

[3]https://joshclose.github.io/CsvHelper/

APPENDIX B
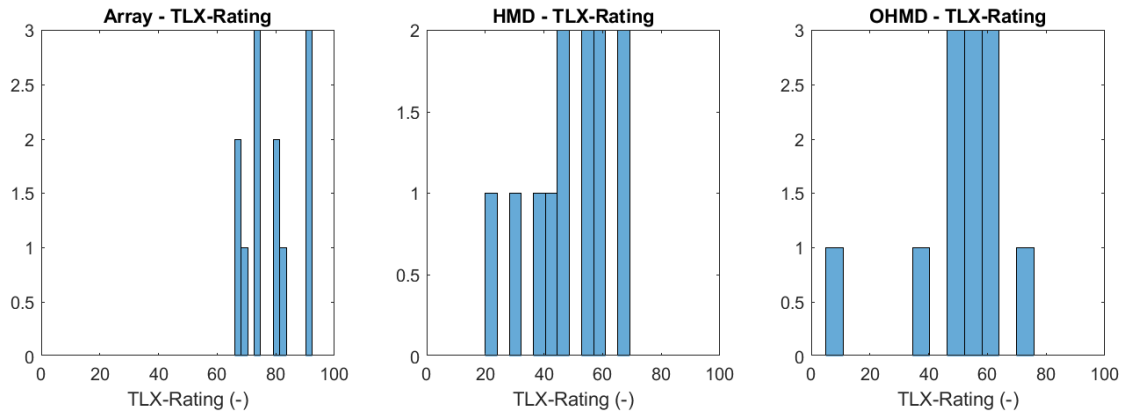STATISTICAL ANALYSES

A. Histograms

Fig. B.1. Histograms of the weighted NASA-TLX rating across the three visual interfaces.
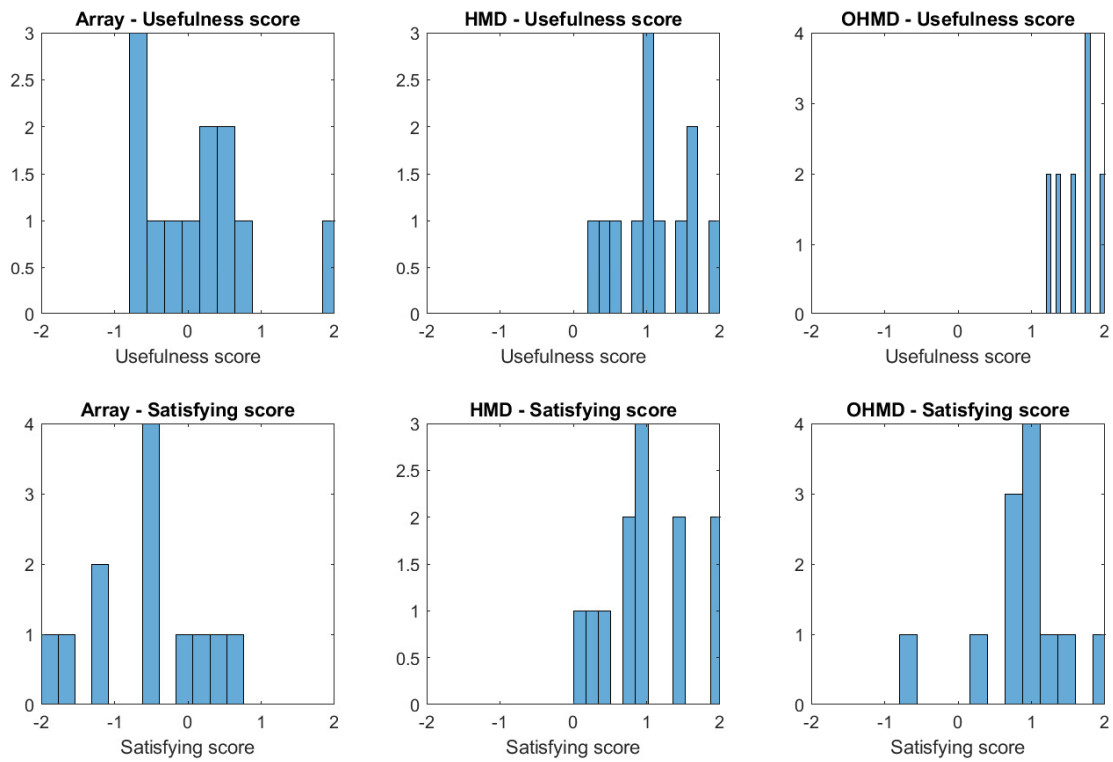

Fig. B.2. Histograms of the usefulness score and satisfying score across the three visual interfaces.
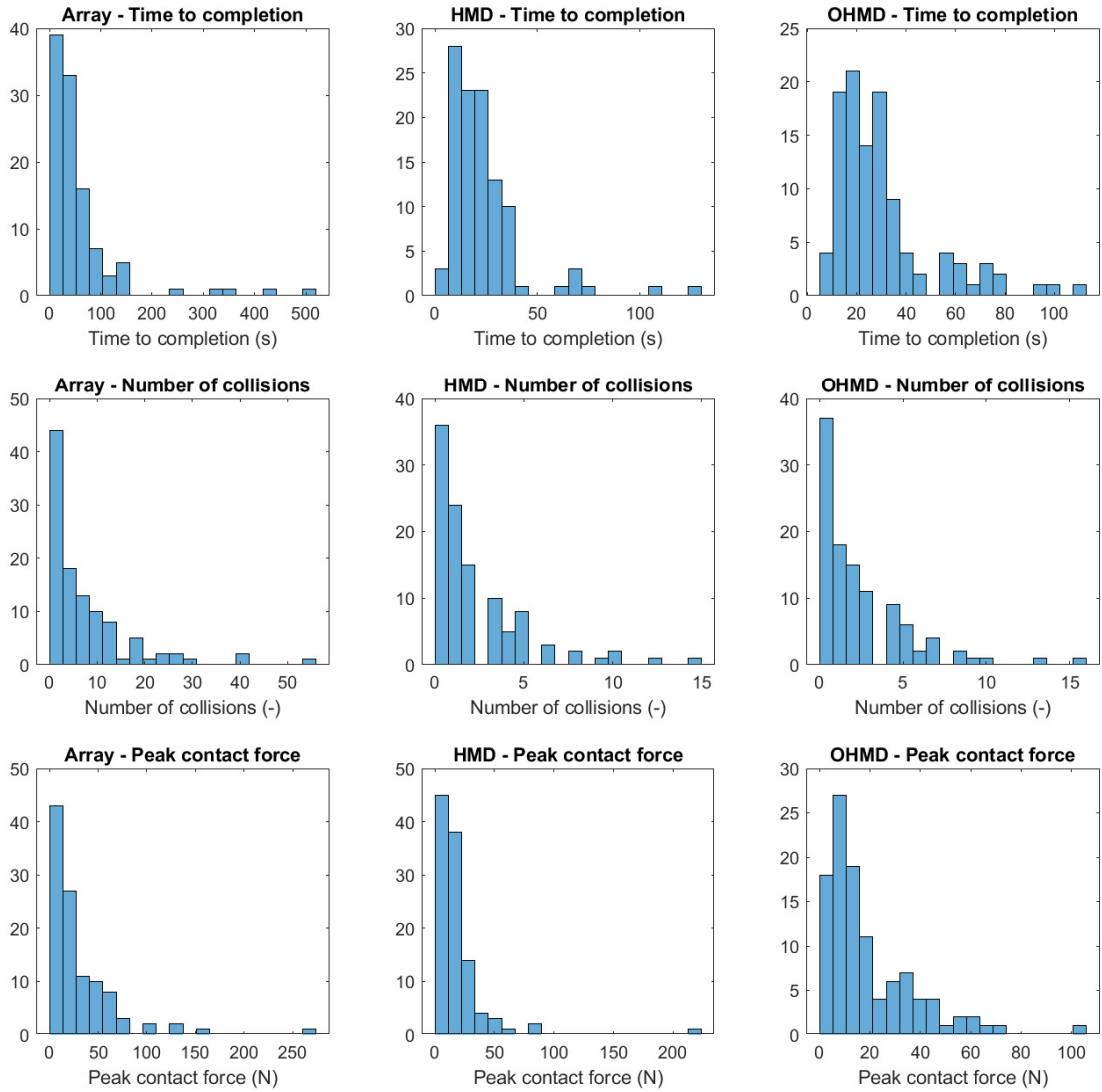
Fig. B.3. Histograms of the objective metrics across the three visual interfaces.
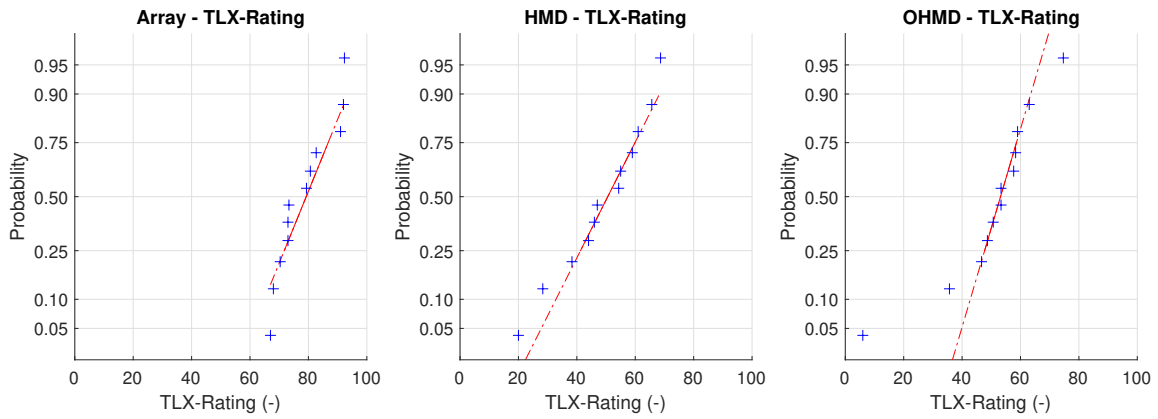
B. Normal probability plots

Fig. B.4. Normal probability plots of the weighted NASA-TLX rating across the three visual interfaces.
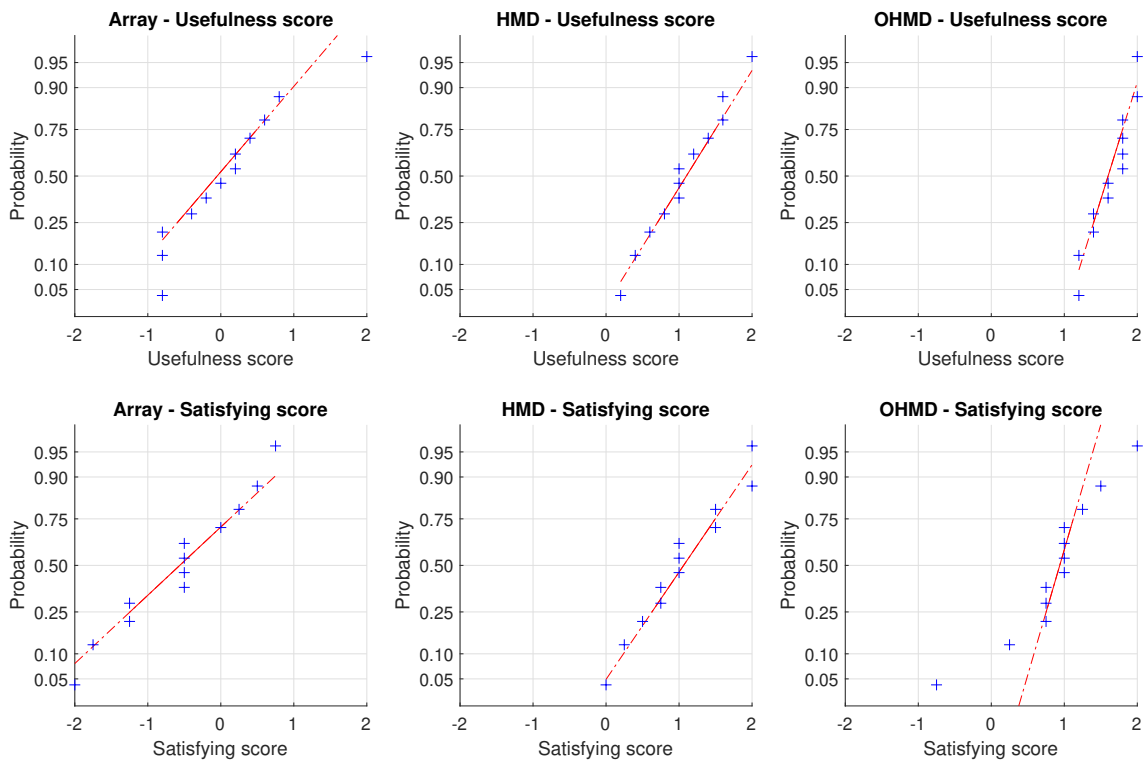

Fig. B.5. Normal probability plots of the usefulness score and satisfying score across the three visual interfaces.
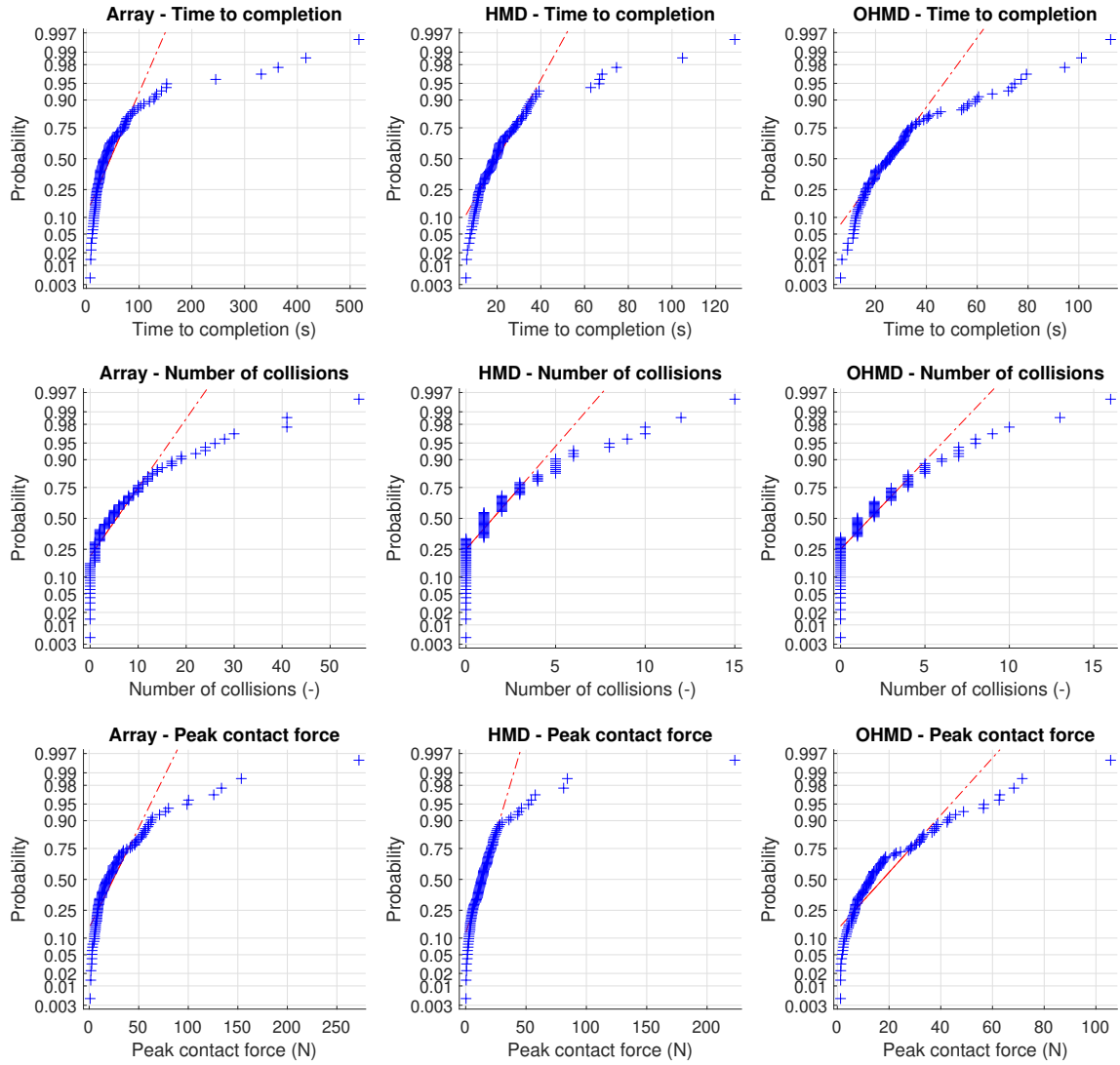
Fig. B.6. Normal probability plots of the objective metrics across the three visual interfaces.

C. Shapiro-Wilk Test for normality

The Shapiro-Wilk test [28] was used to test whether the collected data fit a normal distribution. As expected from the histograms for the objective metrics (Fig. B.3), the results of the test shown in Table B.3 reveal that the measured data for *time to completion*, *number of collisions* and *peak contact force* is heavily skewed and violates the normality assumption of ANOVA. Thus, this data was transformed using an *Aligned Rank Transform* (ART) for multifactor contrast tests that the ANOVA and post-hoc procedures could be applied citeWobbrock2011 [27].

TABLE B.1
SHAPIRO-WILK TEST FOR NORMALITY RESULTS FOR EACH METRIC.

Metric	Interface	Obervations	Skewness	Kurtosis	W	Z	p
TLX-rating	Array	12	0.390	1.738	0.894	1.117	0.132
	HMD	12	-0.558	2.428	0.955	-0.570	0.716
	OHMD	12	-1.456	5.255	0.848	1.819	0.034
Usefulness score	Array	12	0.892	3.521	0.908	0.846	0.199
	HMD	12	0.056	2.230	0.979	-2.052	0.980
	OHMD	12	-0.305	1.848	0.906	0.888	0.187
Satisfying score	Array	12	-0.163	2.017	0.951	-0.402	0.656
	HMD	12	0.147	2.118	0.952	-0.412	0.660
	OHMD	12	-0.888	4.382	0.894	1.109	0.134
Time to completion	Array	108	3.601	17.500	0.565	8.123	0.000
	HMD	108	3.082	15.065	0.685	7.404	6.62E-14
	OHMD	108	1.792	6.286	0.814	6.229	2.34E-10
Number of collisions	Array	108	2.425	10.351	0.734	7.028	1.05E-12
	HMD	108	2.042	7.870	0.756	6.830	4.25E-12
	OHMD	108	1.979	8.091	0.778	6.619	1.80E-11
Peak contact force	Array	108	3.464	19.895	0.661	7.563	1.98E-14
	HMD	108	5.644	44.514	0.513	8.372	0.00E+00
	OHMD	108	1.894	7.484	0.806	6.319	1.31E-10

D. Additional subjective result figures
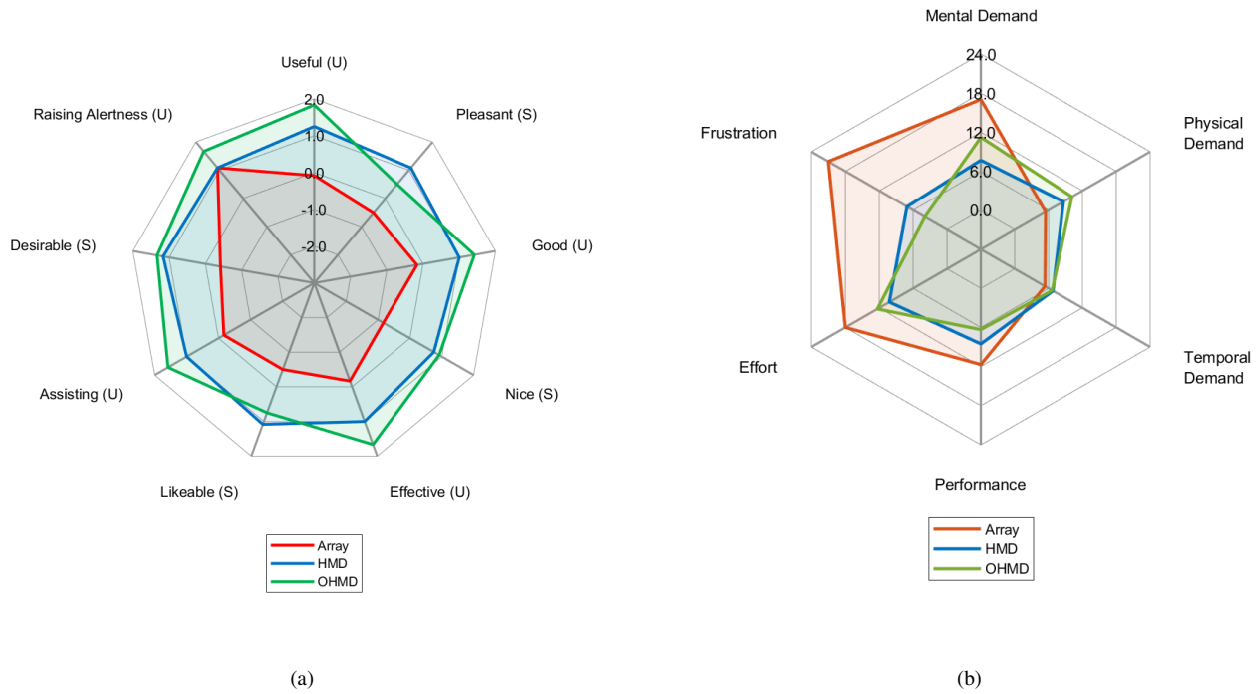

(a)

(b)

Fig. B.7. Spider-charts of showing scores for individual questionnaire components. (a) Scores for individual Acceptance Scale Likert items across the three visual interfaces [24]. A larger surface indicates higher interface acceptance. (U) marks components thats contribute to the usefulness scale and (S) marks components that contribute to the statisfying scale. (b) Weighted task load ratings of individual NASA-TLX components across the three visual interfaces [25]. A larger surface indicates a higher perceived workload.
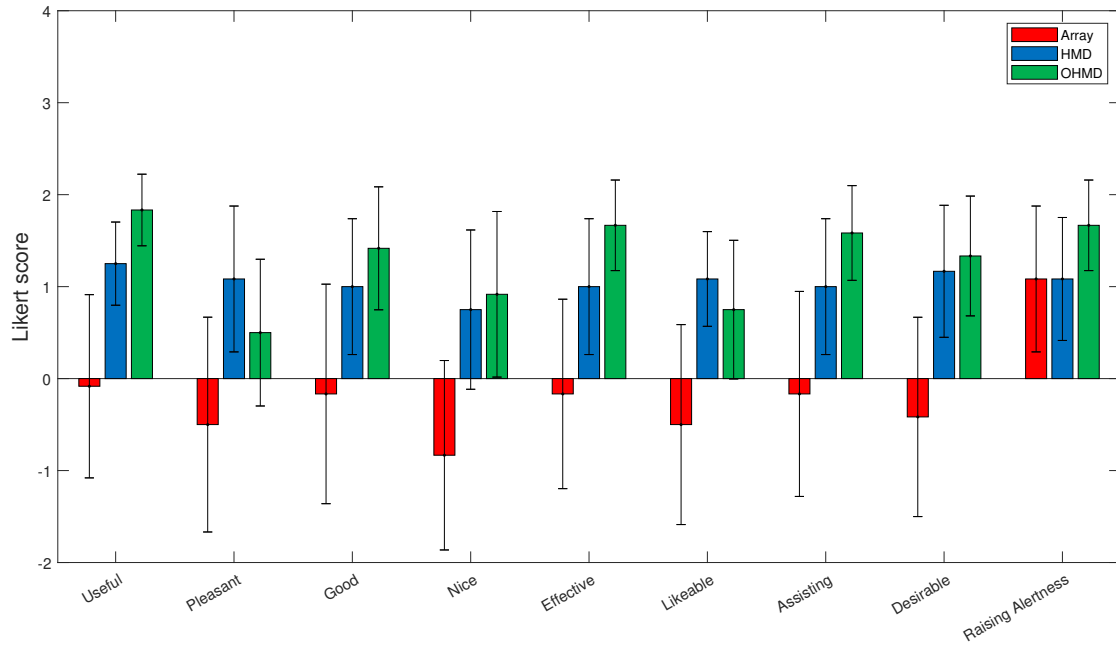
24

Fig. B.8. A bar-chart of the individual Acceptance Scale Likert items across the three visual interfaces [24]. The errorbars indicate the standard deviation.
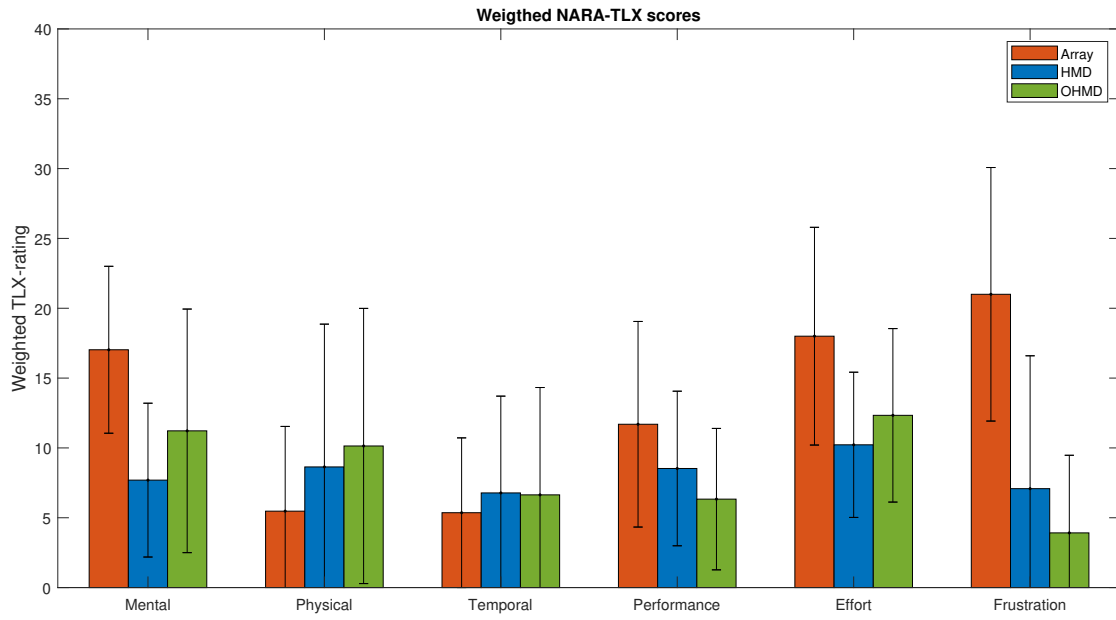
Fig. B.9. A bar-chart of the NASA-TLX components across the three visual interfaces [25]. The errorbars indicate the standard deviation.

A. Experiment Protocol

Experimental Protocol

This document provides an overview of all the steps that are necessary to perform the experiment. Detailed information can be found in the **Study Information, Informed Consent** and the **Participant Instruction** forms.

1. Contacting participants

6 to 10 participants will be contacted via e-mail and asked if they would like to participate in the experiment. Potential participants must have previous experience with haptic input devices. These include co-workers and TU-Delft students.

They will receive the following documents:
1. **Study Information –** Explaining the purpose of the study, risks/ benefits, safety precautions/procedures, withdrawal procedures and data management.
2. **Informed Consent Form –** Which can be signed when willing to participate. Either by replying to the mail or signing it upon receival of the participants.
3. **Experiment Instructions –** This document contains a detailed instruction of the experiment and what is expected from the participants.

2. Receiving participants

Once the participant agrees to participate, they are welcomed in the office of Heemskerk Innovative Technology provided he/she does not show symptoms of corona or other illness. The participant will be reminded on the corona related measures (e.g. wash hands) for which the website of the RIVM website is regularly checked. These measures will also be explained in the **Study Information** document. After a brief spoken introduction of the study, the participant will receive an informed consent form to sign and additional written information about the study. These documents can be read aloud when necessary. Furthermore, the safety and withdrawal procedures are repeated before the experiment commences.

3. Experiment instructions

When the informed consent is signed, the participant will be seated, and the experiment specifics are explained. This is done via the **Experiment Instructions** document which explains the layout of the experiment, how to use the devices and how to behave during the experiment. Furthermore, a video instruction explaining how to operate the difference input devices is shown to the participant.

4. The Experiment

Before starting, information about the participant is entered on the computer (age, gender, experience, etc.) in a form format. Next, the participants are allowed 5 minutes to get acquainted with the Haptic Input Device, the Head-Mounted Display and Motion Controller used in the experiment. Subsequently, 3 experiments will be performed, all preceding a practice run. In between the three conditions, a small break is scheduled. After the three conditions are concluded, participants will fill in a questionnaire containing questions about the acceptance and useability of the three systems.

5. Participants leaving

When the experiment is finished, the participants are thanked for participating and reminded to wash their hands again before leaving. Next, all surfaces and devices are cleaned and made ready for the next participant to enter. Enough time between the participants is scheduled such that proper cleaning is possible.

Final remarks

Participants are allowed to ask questions during the whole experimental protocol. Furthermore, at any time, participants can withdraw from participation of the experiment without having to give reason. As mentioned, this is also explained in the **Study Information** document.

B. Participant Instructions

EXPERIMENT INSTRUCTIONS

Today, you are the remote operator of an industrial robot in a nuclear facility. You are responsible for the maintenance of the Upper Launchers that stabilize plasma by focusing high frequency beams onto it.
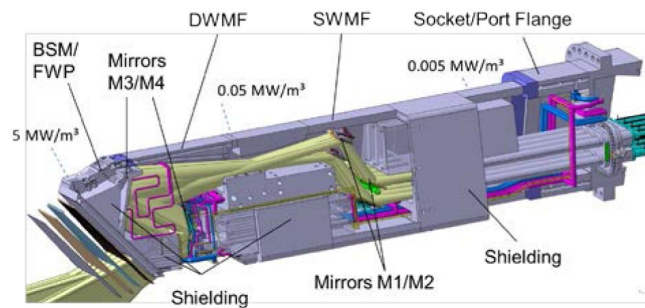
To access components within the Upper launcher, a front shielding needs to be removed. This requires unbolting the shield module and cutting several cooling pipes. When the maintenance of the inner components is done, the shield module is bolted back onto the launcher structure. To finish the job, the cooling pipes need to be rewelded. This last part will be your task today.

Like any teleoperation task, this task involves two major activities: **remote control** and **remote perception**. While a lot of research has been done on remote control, less has focused on the remote perception aspect. Yet, visual perception is critical for the success of teleoperation. Today, you will compare the usability and performance of 3 camera control systems during a remote welding task.

Experiment

The experiment will approximately take **2 hours**.

Experiment set-up

- You will be presented with the setup where you will receive information on how to handle the equipment and interpret the simulation.
- Next, you will get **three minutes** to familiarize yourself with the movement commands that control the position and rotation of the welding tool. You are allowed to ask questions.

Experiment overview

- For each camera system, you start with a **three-minute training session** to get accustomed to the camera system (and its controls). During the session you are allowed to ask questions.
- Subsequently, with each of the three camera systems you will be asked to perform the remote welding task for **3 pipe locations**.
  - For each pipe location, you will start by performing a **practice trial** to familiarize yourself with this specific scenario.
    - *The practice trial will not be evaluated. However, it is important that you try to complete the trial to the best of your ability.*
  - Afterwards you will perform **three trials** for each pipe location.
  - After each pipe location (every 4 trials) you will have a short break.
- After completing all the trials for a camera system, you will be asked to fill out two questionnaires (NASA-TLX and Acceptance Scale) in which you are asked to evaluate your experience.

Trial overview

- In each trial the objective is the same: You will align the position and orientation of the clamp fitting of the welding tool with the marked pipe cut.
- You are expected to perform this task in a time efficient manner. So, try to perform this task as fast as possible while avoiding unwanted collisions and minimizing the force exerted by the welding tool onto the pipe.
- The trial will end under one of these conditions:
  - Proper alignment between the pipe and the clamp fitting
  - You decide to withdraw from the trial/experiment.
  - An error in the simulation
  - The welding tool gets stuck.

The environment

Of course, you won't be operating in a real nuclear facility. Instead, you will perform the remote welding task in a virtual environment created in Unity. This allows some simplification to the scene. For example, during the experiment the robotic arm holding the tool is invisible.
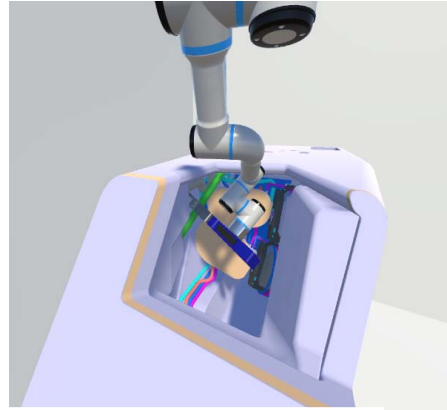
During the experiments only the welding tool is visible

Image of the inside of the Upper Port Launcher.

Remote control

A haptic device (3D Systems Touch) is used to command the movement of the welding tool. By moving the *stylus* of this device, the user can send movement commands to the **welding tool**.

The **control frame of reference** is the **world frame** (i.e., relative to the front view of the plug).

By default, movement commands are disabled. The user can enable movement commands by pressing and holding the *dark grey button* on the stylus.

Stylus

- *Light grey button* to start the trial
- *Dark grey button* to enable movement commands

- **If enabled**; the welding tool will respond to movements of the stylus.
- **If disabled**; the stylus can be moved freely without affecting the welding tool. Furthermore, a *virtual copy of the Touch* will appear in the view so the operator can check the pose of the Touch.

Note that the reach of the stylus is limited. Therefore, you likely won't be able to complete the task in one smooth motion of the stylus before it reaches the end of its workspace. To reset the workspace of the touch:

1. Disable movement commands
2. Move the stylus to a more comfortable position.
3. Re-enable movement commands

TIP: Use the virtual Touch to align the longitudinal axis of the stylus with the longitudinal axis of the welding tool before re-enabling the movement commands. This way, the movement of the welding tool as a response to your input is most transparent (i.e., it feels like you are holding the tool in your hand).
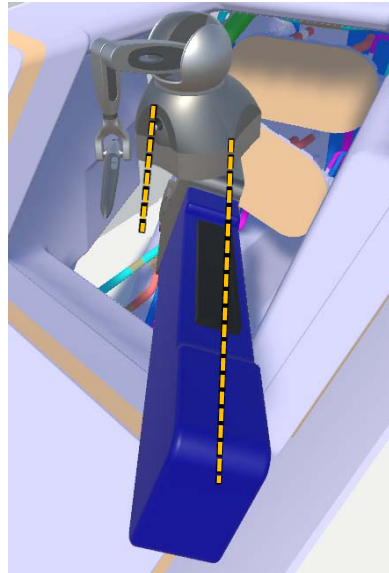
Remote perception

You will use three different camera systems while performing the remote welding task.

Array of camera views (Array)

Six views of the virtual environment will be displayed on a monitor.

- Five of these views are captured by cameras in fixed positions, providing *left-, front-, right-, top-,* and *overview* shots of the scene.
- One '*eye-in-hand*' view is captured by a camera that is attached to the welding tool.
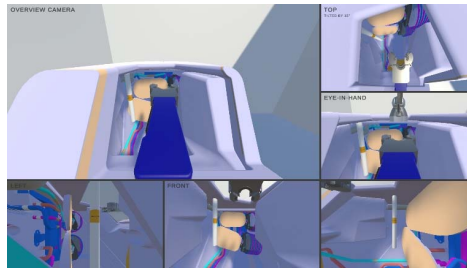
Standard head-mounted display (HMD)

You will perceive the virtual environment through an immersive head-mounted display (HMD).

- You can change your viewpoint position and viewing direction by making head movements just like in real life.

sHMD is renamed to OHMD

Spherical head-mounted display (sHMD)

You will perceive the remote environment through an immersive head-mounted display (HMD).

- Your head rotation controls the viewing direction relative the welding tool, which will be kept at the center of your view.
  (i.e., rotating your head to the right will provide left side view of the welding tool)
- The viewpoint position will be adjusted automatically to keep the welding tool centered in the view
- You can control the distance between the camera and the welding tool using the joystick on the motion controller.

Task instructions

Once the virtual environment has loaded, do the following to start the trial:

1. Sit straight up in your chair, facing forwards (towards the monitor).
2. Press the **light grey stylus button** to start the trial.

In each trial the objective is the same: You will align the position and orientation of the **clamping insert** of the welding tool with the marked **pipe cut**.

The **pipe cut** that needs to be welded.

The **clamping insert.**
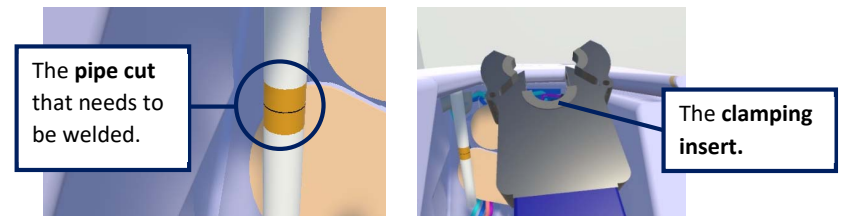
The task can be completed when proper alignment is achieved. This is when:

- The pipe cut (marked by a black ring) is within a **5mm margin of error** from the centre of the clamping insert.
- The angular alignment (roll and tilt) between the pipe and the tool is within a **2 degree margin of error**
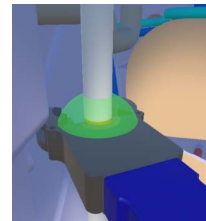
To indicate proper alignment has been achieved, a **green sphere** will appear around the pipe cut. The trial can then be ended by releasing the **dark grey stylus button.**

Safety

To avoid damage to the cooling pipes and the welding tool the force exerted by the robot onto the pipe must be kept to a minimum. In the simulation, this can lead to instability. Therefore, it is important to use the visual feedback (views) provided by the camera systems to get a good alignment between the pipe and clamping insert of the welding tool.

Furthermore, all components of the Upper Launcher are surface treated (i.e., copper coated) to withstand the immense thermal loads. Any collision with these components is unwanted as it may lead to damage to the coating and/or the welding tool.

Evaluation Criteria

You are expected to perform this task in a time efficient manner. So, try to:

- perform the task as fast as possible
- avoid unwanted collisions
- minimize the force exerted by the welding tool onto the pipe

C. Consent Form

Consent Form for participating in an HMD Orbital Camera System experiment

Please tick the appropriate boxes   **Yes   No**

Taking part in the study

I have read and understood the study information dated 25/11/2021, or it has been read to   ○   ○
me. I therefore understand that the research in done in collaboration with Heemskerk
innovative Technologies. I have been able to ask questions about the study as well as the
company's involvement, and my questions have been answered to my satisfaction.

I consent voluntarily to be a participant in this study and understand that I can refuse to   ○   ○
answer questions and I can withdraw from the study at any time, without having to give a
reason.

I understand that taking part in the study involves gathering and storing of non-direct personal   ○   ○
identifiable data as well as this consent form containing my name and signature.

Objective data metrics will be recorded by the computer simulation while the participant is
performing the described experiment tasks. Subjective data is obtained through a
questionnaire form completed by the participant.

Risks associated with participating in the study

I understand that taking part in the study involves the following risks: ○   ○

- Physical discomfort while looking at a computer screen
- Physical discomfort such as nausea while using a Head Mounted Display
- Physical discomfort while using a haptic input device (3D systems Touch)

Use of the information in the study

I understand that information I provide will be used for analysis in this research and that the   ○   ○
corresponding results (papers, theses, reports, and articles) are published.

I understand that personal information collected about me that can identify me, such as [e.g.   ○   ○
my name or where I live], will not be shared beyond the study team. Data gathered during the
experiment will be linked to this consent form using a unique participant number such that the
results are pseudonymous. The link between participant number and participant name
(consent form) will be recorded in a key document to which only the research team has
access. This key will be destroyed as soon as the research is concluded, and the corresponding
results (papers, theses, reports) are published. Data will we stored according to the TU Delft
Research Data Framework Policy. During the active phase of research, the project leader from
TU Delft will oversee the access rights to data (and other outputs), as well as any requests for
access from external parties.

Future use and reuse of the information by others

I give permission for the non-direct personal identifiable data that I provide to be archived in   ○   ○
the Project Storage system of the TU-Delft, the 4TU.ResearchData archive so it can be used for
future research and learning. This includes anonymized data recorded by the simulation in
.CSV format and questionnaire forms completed by the participant in .PDF format.

Signatures

_____   _____   _____
Name of the participant [printed]   Signature   Date

I have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

_____   _____   _____

Researcher name [printed]   Signature   Date

Study contact details for further information:
Name: Sjoerd Kuitert
Phone number: +31(0)6 25 42 88 48
Email address: s.kuitert@student.tudelft.nl

QUESTIONAIRES

A. Pre-Experiment Questionnaire

PRE-EXPERIMENT QUESTIONNAIRE

What is your participant number?

What is your age?

What is you gender?

☐ Male   ☐ Female   ☐ Other

Are you left- or right-handed?

☐ Left   ☐ Right

What is your field of study/work?

How experienced are you with videogames? Please mark on of the options.

☐ None   ☐ 1 hour   ☐ 10 hours   ☐ 1 day   ☐ 1 week   ☐ 10 weeks   ☐ More

How experienced are you with virtual reality?

☐ None   ☐ 1 hour   ☐ 10 hours   ☐ 1 day   ☐ 1 week   ☐ 10 weeks   ☐ More

How experienced are you with (virtual) remote-control/teleoperation?

☐ None   ☐ 1 hour   ☐ 10 hours   ☐ 1 day   ☐ 1 week   ☐ 10 weeks   ☐ More

---

B. Post-Experiment Questionnaire

POST-EXPERIMENT QUESTIONNAIRE

Participant number:

Please answer these final questions:

1. What, if anything, did you think was difficult about this experiment? Please provide examples when you can.

2. Which condition did you like the best?

3. Why did you like this condition the best? Please provide examples when you can.

4. If you have additional remarks, please write them here…

C. Van der Laan Acceptance Scale

A simple scale with nine Likert items that asseses system acceptance on two dimensions, a *usefulness* scale and a *satisfying* scale [24].

ACCEPTANCE SCALE

Participant number: Camera system:

My judgements of the <u>current camera system</u> (as a whole) are …

(Please tick a box on every line)

1.	Useful	□	□	□	□	□	Useless
2.	Pleasant	□	□	□	□	□	Unpleasant
3.	Bad	□	□	□	□	□	Good
4.	Nice	□	□	□	□	□	Annoying
5.	Effective	□	□	□	□	□	Superfluous
6.	Irritating	□	□	□	□	□	Likeable
7.	Assisting	□	□	□	□	□	Worthless
8.	Undesirable	□	□	□	□	□	Desirable
9.	Raising Alertness	□	□	□	□	□	Sleep-inducing

D. NASA Task Load Index

To administer and record the NASA Task Load Index (NASA-TLX) [25] its official iOS app[4] was used. This made it easy to set up, capture, and analyze NASA TLX results. Participants used an Ipad Pro 2020 to fill in the form. Data was saved to the internal storage as `.csv` files.
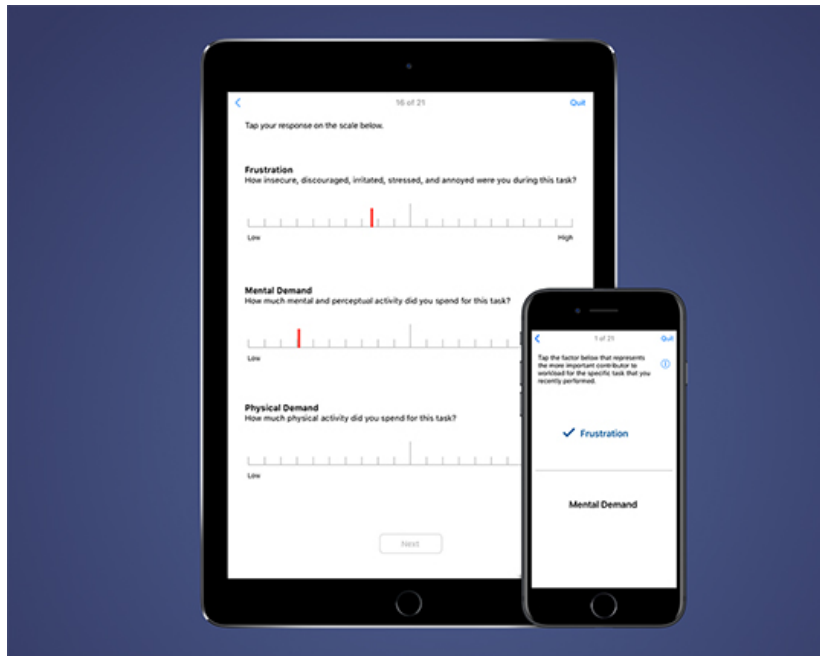

Fig. D.1. An image of the the NASA-TLX app on iPad and iPhone[5]