



Damien Jean Maxime Mazeas

Key Principles for Assessing and Implementing Remote Inspection
with Telexistence Capability

Cranfield Manufacturing and Materials

PhD

Academic Year: 2022 - 2023

Supervisor: Professor John Ahmet Erkoyuncu (Cranfield University)

Associate Supervisor: Professor Frédéric Noël (Grenoble University)

December 2023



Cranfield Manufacturing and Materials

PhD

Academic Year 2022 - 2023

Damien Jean Maxime Mazeas

Key Principles for Assessing and Implementing Remote Inspection
with Telexistence Capability

Supervisor: Professor John Ahmet Erkoyuncu (Cranfield University)
Associate Supervisor: Professor Frédéric Noël (Grenoble University)

December 2023

© Cranfield University 2023. All rights reserved. No part of this
publication may be reproduced without the written permission of the
copyright owner.

Academic integrity declaration

I declare that:

- the thesis submitted has been written by me alone.
- the thesis submitted has not been previously submitted to this university or any other.
- that all content, including primary and/or secondary data, is true to the best of my knowledge.
- that all quotations and references have been duly acknowledged according to the requirements of academic research.

I understand that to knowingly submit work in violation of the above statement will be considered by examiners as academic misconduct.

Abstract

This thesis investigates assessing and implementing telexistence capabilities for enhancing remote inspection and control in hazardous environments. Telexistence, allowing virtual presence via robotic systems or avatars, is crucial for performing high-risk tasks in fields like marine engineering.

The research integrates virtual reality, industrial robotics, and sensors to improve human abilities in maintenance, repair, and overhaul tasks, aiming for safer and more efficient methods in hazardous settings. A key aspect is the design and evaluation of telexistence interfaces for remote visual inspection, focusing on interfaces that enhance operators' sense of presence and spatial awareness, crucial for decision-making in risky conditions.

A proposed system architecture is validated with a real-world use case and used as a basis for two user studies. The first study compares the effectiveness of a 2D and a VR interface in remote visual inspection in machinery spaces, leading to recommendations for enhancing telexistence in maintenance scenarios. The second study examined the impact of interface design on operators' telexistence capabilities at a cognitive workload level, using modalities like 2D feeds, 3D point clouds, augmented virtuality, and mixed modalities.

A central theme is using VR to bridge the gap between human operators and robots. VR's immersive nature improves operators' remote-control abilities, with a digital representation in a VR control room mirroring the robot's actions in real-time, enhancing remote-control effectiveness and providing a more immersive telexistence experience.

Key findings demonstrate the advantages of immersive VR interfaces over traditional 2D interfaces in remote visual inspection tasks. VR interfaces offer a heightened sense of presence, reduced cognitive workload, and improved usability, leading to more efficient and accurate task execution. Additionally, VR users experienced fewer errors and completed tasks faster. Different visualisation modalities also showed varied impacts on cognitive load, with 3D modality imposing the most substantial demands.

Keywords:

Telexistence, Virtual Reality, Remote Inspection, Remote Control, Human Factors, Human-Computer Interaction

Acknowledgements

First, I would like to thank the Defence Science and Technology Laboratory (Dstl) and the French Ministry of Defence's Defence Innovation Agency (AID) for operating and funding this UK-France joint PhD programme. I want to thank Ian Greening, a marine engineer at Dstl, for his questioning and feedback regarding my work throughout the 36-month duration of this programme.

I thank the PhD project supervisors, Professor John Ahmet Erkoyuncu and Professor Frédéric Noël, for making this research project possible, for their supervision, and for allowing me to work on it.

I also want to express my appreciation for the opportunity to spend six months at the Laboratoire G-SCOP in Grenoble, France, with Professor Frédéric Noël and Camilo Medina. This experience has enriched my research and provided me with knowledge that has significantly contributed to the success of my experiments.

Lastly, I would like to thank the students and staff at Grenoble and Cranfield for their time and willingness to participate in my experiments.



Table of Contents

Academic integrity declaration.....	i
Abstract	ii
Acknowledgements	iv
List of Figures.....	viii
List of Tables	xi
List of Equations.....	xii
List of Abbreviations	xiii
1 Introduction.....	14
1.1 Context of the PhD thesis	14
1.2 Problem statement.....	15
1.3 Motivation behind telexistence for maintenance context.....	18
1.4 Aim and objectives.....	22
1.5 Research questions	23
1.6 Research contributions and publications	26
1.7 Thesis structure	29
2 Literature review.....	30
2.1 Virtual reality and telexistence	30
2.1.1 Introduction	30
2.1.2 A brief history of virtual reality	31
2.1.3 Telexistence capability	37
2.1.4 Key terminologies.....	39
2.1.5 Presence and immersion.....	42
2.1.6 Factor contributing to the presence	44
2.1.7 Cognitive workload and presence	45
2.2 Telexistence for remote inspection	46
2.2.1 Immersive interface for teleoperation	48
2.2.2 Visualisation techniques.....	52
2.2.3 Control techniques	54
2.2.4 AI in monitoring Human decisions.....	57
2.3 Literature review conclusions.....	58
3 Research methodology	60
3.1 Capturing context and industrial requirements.....	60
3.2 Comparative analysis of communication protocol.....	64
3.3 Use case of remote visual inspection.....	68
3.4 Experiment 1 method.....	69
3.4.1 Experiment setup	70
3.4.2 Participants	70
3.4.3 Data collection.....	72
3.4.4 Procedure.....	73
3.4.5 Data analysis.....	76

3.5 Experiment 2 method	77
3.5.1 Experimental setup.....	77
3.5.2 Participants	78
3.5.3 Data collection.....	79
3.5.4 Procedure.....	82
3.5.5 Data analysis.....	84
3.6 Ethical considerations	85
4 Enabling remote operations: A comprehensive framework for telexistence capability	86
4.1 Design Principles for Telexistence Capability Framework	86
4.2 System design	87
4.3 System architecture based on DDS	88
4.4 Immersive interface design	93
4.5 Use case application: Fanuc industrial robot	98
4.5.1 Control of the physical robot.....	98
4.5.2 Digital representation of the robot	99
4.5.3 Coupling the physical and digital twins.....	99
4.6 Conceptual design refinement of the telexistence-based framework....	104
4.7 Latency and time-delay discussion	105
4.8 Limitations.....	109
5 Experiment 1: Evaluating remote visual inspection interfaces.....	111
5.1 Introduction	111
5.2 Tasks and interfaces design	112
5.2.1 Gamified task creation.....	112
5.2.2 Real-time feedback and kinematic control of collaborative robots..	115
5.2.3 Interfaces designs	118
5.3 Hypotheses.....	125
5.4 Results.....	126
5.4.1 Presence	126
5.4.2 Cognitive workload.....	130
5.4.3 Usability.....	136
5.4.4 Performance.....	138
5.4.5 Participants feedback during the experiment	142
5.5 Summary	144
6 Experiment 2: Impact of data visualisation modalities on cognitive workload in VR-based industrial robot teleoperation	148
6.1 Introduction	148
6.2 Experiment and modalities setups	149
6.2.2 Modalities setups	151
6.3 Hypotheses.....	153
6.4 Results.....	153
6.4.1 Cognitive workload.....	155

6.4.2 Performance.....	158
6.5 Summary	163
7 Discussion.....	167
7.1 Principles for assessing and implementing remote inspection with telexistence capability	167
7.2 Key actors of the telexistence industry.....	169
7.3 Enhancing safety and efficiency through AI and machine vision	171
7.4 Future steps of telexistence in maintenance.....	172
8 Conclusions.....	175
8.1 Addressing the aim	175
8.2 Addressing the objectives	176
8.3 Answering the research questions.....	178
8.3.1 Research question 1	179
8.3.2 Research question 2	180
8.3.3 Research question 3	181
References	183
Appendices.....	204
Appendix A Ethical approval letter	204
Appendix B Expert 1 questionnaire responses	205
Appendix C Expert 2 questionnaire responses	209
Appendix D Inverse Kinematics for UR16e Robot	213
Appendix E Fanuc remote control framework.....	215
Appendix F Analysis of the Fanuc trajectory sample	217

List of Figures

Figure 1-1 Context and associated main concepts of the PhD	15
Figure 1-2 Machinery spaces of the XV Patrick Blackett (www.navylookout.com)	19
Figure 1-3 Current and future CBM of vessels	21
Figure 2-1 VR categories: non-immersive VR (1), semi-immersive VR (2), and fully immersive VR (3).....	30
Figure 2-2 NASA VIVED and DataGlove.....	32
Figure 2-3 CAVE systems concept.....	33
Figure 2-4 Comparison between 3Dof and 6Dof	35
Figure 2-5 Telepresence Robots	37
Figure 2-6 XR-based robot teleoperation interfaces; (1) stereoscopic telepresence (Ballantyne and Moll, 2003), (2-3) AR (Quintero et al., 2018; Walker, Hedayati and Szafir, 2019), (3) VR (Omarali et al., 2020)	51
Figure 2-7 Egocentric view (left), exocentric view (right)	52
Figure 2-8 Direct control in robotics.....	54
Figure 2-9 Shared control in robotics (Wang et al., 2022)	55
Figure 2-10 Supervised control in robotics (Wonsick and Padir, 2020)	56
Figure 3-1 Research methodology approach to the dissertation	60
Figure 3-2 Participants during the experiment with the VR interface (left), and the 2D interface (right).....	70
Figure 3-3 Participants and experiment 1.....	74
Figure 3-4 Participant during the experiment 2.....	77
Figure 3-5 Pupils diameters measurement from the camera of the Varjo XR-380	
Figure 3-6 Scoring framework for participants performance.....	81
Figure 3-7 JSON configuration files in charge of the experiment flow	83
Figure 4-1 Elements to consider as part of a telexistence system.....	87
Figure 4-2 DDS Infrastructure	89
Figure 4-3 Architecture and components overview of the proposed framework	90
Figure 4-4 Proposed user interface	91

Figure 4-5 Interaction methods in the 3D environment for moving around (1), interact with the UI (2), and interact with the robot (3)	94
Figure 4-6 Components of the interface: UI of the framework, 3D buttons (1), 2D video feed viewer (2), point cloud (3), and robot pose values (4)	95
Figure 4-7 Pipeline for transmitting the point-cloud data	97
Figure 4-8 Fanuc M-20iA (left) and R-30iB mate cabinet (right)	98
Figure 4-9 Hierarchy robot elements setup in the software Unity 3D	99
Figure 4-10 Digital twin for remote inspection task.....	100
Figure 4-11 Coordinate systems of Unity 3D (left) and Fanuc (right)	101
Figure 4-12 High-level system architecture of the refined framework and data flow overview	104
Figure 4-13 Robot trajectory recording: 133 data samples.....	107
Figure 5-1 Description of the two designed scenarios for each Interface	113
Figure 5-2 Flowchart of the data for real-time feedback in the proposed process	116
Figure 5-3 Forward and Inverse kinematics representation	117
Figure 5-4 Top-left panel	119
Figure 5-5 Top-right panel	119
Figure 5-6 Bottom-right panel.....	120
Figure 5-7 Bottom-left panel	120
Figure 5-8 Screenshot of the VR Interface	122
Figure 5-9 Presence profile of the 2D and VR interface regardless of the order effect.....	127
Figure 5-10 Comparison of IPQ factors across 2D and VR modalities based on testing order.....	128
Figure 5-11 Comparison of cognitive workload between 2D and VR interfaces	131
Figure 5-12 Impact of testing order on cognitive workload assessment in 2D and VR	133
Figure 5-13 Comparative usability scores: 2D vs. VR interfaces.....	137
Figure 5-14 Boxplots of Completion Times by Order and Interface	139
Figure 6-1 Dual Check Safety (DCS) areas (in green)	150
Figure 6-2 Actions taken for the safe operation of the robot.....	151

Figure 6-3 Modalities setup and positions of areas to inspect (t1, t2, t3).....	152
Figure 6-4 Screen capture of the modalities during the experiment	155
Figure 6-5 NASA-TLX workload assessment across modalities.....	156
Figure 6-6 Change in pupil dilation from baseline across modalities	157
Figure 6-7 Box plot of task completion times by modality.....	158
Figure 6-8 Mean completion time of modalities by order	159
Figure 6-9 Examples of incorrect and correct photos from participants.....	161
Figure 6-11 Comparison of mean scores across modalities by presentation order	162
Figure 6-10 Distribution of scores (from the scoring framework introduced in section 6.2.4.2) across visualisation modalities	162
Figure 7-1 Telexistence's model-T robot (June 2021)	169
Figure 7-2 Extend Robotics low latency VR teleoperation demonstration between the UK and Turkey (October 2023).....	170
Figure 7-3 Sarcos Guardian XT demo (May 2022).....	171
Figure 7-4 Example of expert-worker remote collaboration with telexistence capability.....	173

List of Tables

Table 2-1 Suggested factors contributing to the sense of presence (Witmer and Singer, 1998)	44
Table 2-2 Overview of immersive teleoperation interfaces	48
Table 3-1 Comparison of some of the most popular communication protocols (created by the author)	64
Table 3-2 Flow of the experiment (blue are conducted two times)	76
Table 3-3 Flow of the experiment (blue steps are repeated four times)	84
Table 4-1 Data structures description of samples transmitted.....	92
Table 4-2 Iterations done for optimising the point-cloud processing speed.....	96
Table 4-3 FANUC M-20iA specification and motion range	98
Table 5-1 Example of participant photos taken during VR and 2D Interface tasks	114
Table 5-2 D-H parameters for the UR16e robot (Universal Robots, 2022)	117
Table 5-3 Descriptive statistics of IPQ questionnaire results.....	126
Table 5-4 Paired T-test Results Comparing Workload Perceptions Based on Testing Order for 2D and VR Interfaces	134
Table 5-5 Descriptive Statistics of Errors Based on Interface and Testing Order	140

List of Equations

(3-1)..... 80

(4-1)..... 102

(4-2)..... 102

(4-3)..... 102

(4-4)..... 103

(4-5)..... 103

(4-6)..... 103

List of Abbreviations

AID	Defence Innovation Agency
AR	Augmented Reality
CAVE	Cave Automatic Virtual Environment
CBM	Condition-Based Maintenance
DDS	Data Distribution Service
D-H	Denavit-Hartenberg
DoF	Degrees of Freedom
Dstl	Defence Science and Technology Laboratory
HMD	Head-Mounted Display
HMI	Human-Machine Interface
HRI	Human-Robot Interface
IPQ	iGroup Presence Questionnaire
MR	Mixed Reality
MRO	Maintenance, Repair and Overhaul
QoS	Quality of Service
ROS	Robot Operating System
RTDE	Real-Time Data Exchange
RTT	Round-Trip Time
SUS	System Usability Scale
TCP	Tool Center Point
USV	Unmanned Surface Vehicles
UUV	Unmanned Underwater Vehicles
UVS	Unmanned Vehicle Systems
VR	Virtual Reality
XR	Extended Reality

1 Introduction

1.1 Context of the PhD thesis

This PhD thesis was part of the UK-France joint PhD programme, which aims to develop research in key areas of mutual interest to France and the UK. It was jointly managed by the Defence Science and Technology Laboratory (Dstl) and the French MOD's Defence Innovation Agency. The G-SCOP laboratory of Grenoble University partnered to conduct this PhD over 36 months, with 30 months spent at Cranfield University and 6 at Grenoble University.

Dstl had demonstrated interest in telexistence concepts for military personnel to undertake dangerous tasks in hazardous environments without physically being present (www.gov.uk, 2022). Telexistence is a capability that allows individuals to be present in a remote location, controlling robotic systems or avatars as if they were there themselves. This PhD focused on developing a telexistence-based approach to augment human capabilities and remotely complete Maintenance, Repair, and Overhaul (MRO) activities that risk causing physical injury and harm to the maintainer.

This PhD thesis explores the intersection of Virtual Reality (VR), industrial robotics, and sensors, primarily focusing on their applications in telexistence, real-time interaction, and remote control, respectively. These three domains merge to address the future of augmented remote inspection work, allowing enhanced human capability in MRO activities and hazardous environments, as shown in Figure 1-1. By leveraging advances in these fields, we can develop a safer and more efficient approach to these tasks, pushing the boundaries of human capabilities while minimising risk.

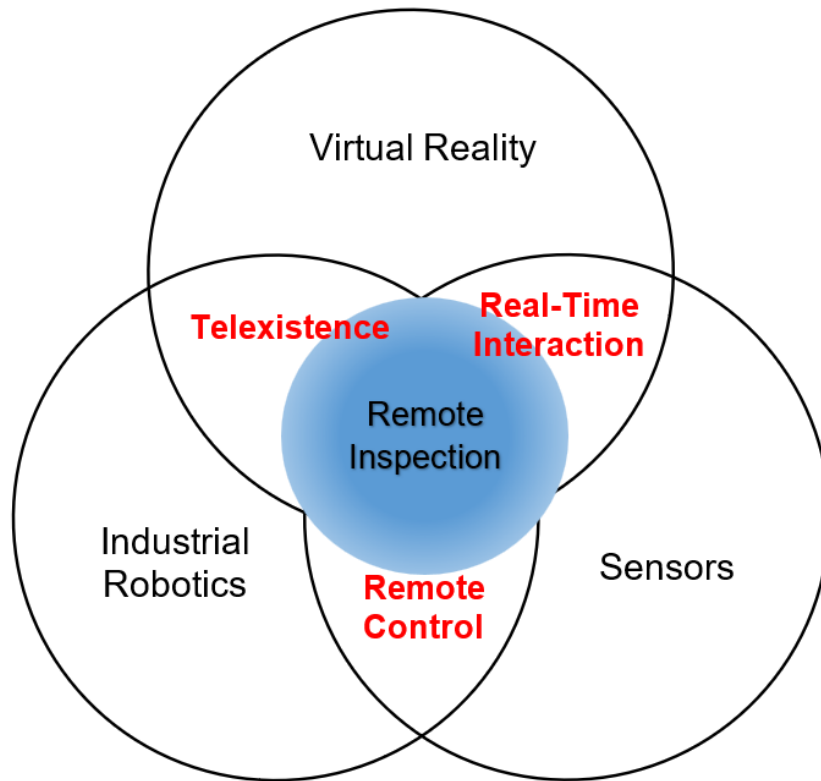


Figure 1-1 Context and associated main concepts of the PhD

1.2 Problem statement

Recent advancements in robotics, AI, and immersive technologies have made realising effective telexistence systems possible and amplified their necessity in the increasingly digital and interconnected society. Telexistence systems, leveraging robotic avatars, are ready to play a crucial role in the future, especially highlighted by the global COVID-19 pandemic. This crisis has fundamentally altered lives, severely restricting physical interactions and compelling rapid adaptation to new norms. The pandemic-induced shift towards remote work, telemedicine, and distance learning has accelerated the adoption of telepresence technologies. These technologies have transformed various sectors and redefined human interaction, underscoring the growing need for telexistence systems to enable efficient, safe operations in various fields. Moreover, as we continue to navigate through and beyond the pandemic, the need for robust Human-Robot Interfaces (HRIs) capable of facilitating seamless distant interaction and meeting industrial safety requirements in hazardous

environments becomes more critical. Therefore, This PhD project focuses on tackling these challenges, aiming to advance the development of innovative telexistence systems essential for evolving societal needs.

Industries where human operators interact with heavy machinery and dangerous materials are particularly affected (Marvel, Falco and Marstio, 2015). In the space industry, teleoperated robots are crucial in performing MRO in outer space's hostile and inaccessible environment. Current teleoperation methods, however, suffer from signal delays and communication disruptions, presenting risks for mission-critical operations. In the healthcare sector, telepresence robots are used for remote patient care, enabling healthcare providers to interact with patients from a distance. Yet, these robots often lack the tactile feedback for more nuanced interactions, such as physical examinations or surgical procedures. In disaster response scenarios, operators employ remotely operated robots to enter disaster-stricken areas unsafe for human responders. However, these robots often need help navigating complex terrains and managing unfamiliar objects, limiting their effectiveness in these critical operations. In nuclear facilities, teleoperated robots are essential for handling radioactive materials and conducting maintenance tasks in highly radioactive environments. These robots, however, often have difficulty performing delicate tasks due to the limitations of current teleoperation interfaces. Even with the central role of robotic automation in modern smart factories, these systems often fail to manage unexpected scenarios, leading to productivity losses and financial setbacks (Benotsmane, Kovács and Dudás, 2019).

It underscores the urgency of developing secure and efficient human-robot collaboration systems for future production and maintenance. If unresolved, these issues risk compromising human safety, reducing efficiency, and inflating costs for organisations reliant on remote maintenance in hazardous environments.

A specific challenge lies in designing and evaluating telexistence interfaces for remote visual inspection tasks requiring precision and environmental awareness (Nostadt et al., 2020; Wonsick and Padir, 2020).

Focusing on inspection tasks in the context of teleexistence and remote maintenance makes sense for the following reasons:

- **Providing human operators with a sense of presence in the remote environment.** This is important for ensuring human operators can make accurate decisions and act appropriately in hazardous environments (Hagiwara, 2018; Villani et al., 2018). Providing a sufficient level of spatial and environmental condition (e.g., temperature, pressure, acoustic signature) awareness is critical in teleexistence and robotic avatars for a comprehensive understanding of the remote physical context.
- **Guaranteeing safe working conditions for the human operator in hazardous environments.** This is important for preventing accidents and injuries (Villani et al., 2018). Inspection tasks frequently occur in dangerous environments, such as maritime, defence, and nuclear industries or during disaster response. Teleexistence can improve operator safety by enabling remote task execution, which reduces the risk of human exposure to these hazardous environments.
- **Improving the performance and user experience of HRIs for remote maintenance tasks.** Teleexistence can enable improvements in the efficiency of conducting inspection-related tasks. The digital twin in the VR control room provides real-time data, enabling swift detection and addressing of issues. This can lead to time, cost savings, and increase the overall productivity of maintenance (Senft et al., 2021). A digital twin is a virtual representation of a physical product instance, and it replicates its operation and defects. It is a real-time, dynamic, digital counterpart of a physical object or process. In teleexistence, a digital twin could represent a remotely operated robot (Grieves and Vickers, 2017; Lo, Chen and Zhong, 2021; El Saddik, 2018).
- **Real-time Interaction and Feedback.** Teleexistence systems provide real-time interaction with the environment and immediate feedback, essential for effective inspection tasks. Robotic systems could relay visual, auditory, or even haptic feedback to the operator, improving the quality of inspections compared to traditional remote monitoring or teleoperation.

The concept of telepresence addresses these challenges, enabling human operators to perform complex tasks such as inspection, diagnostics, and repair with robotic assistance while maintaining a safe distance from hazardous conditions. Telepresence, a concept introduced by Susumu Tachi in 1980 (Tachi, 2015a), goes beyond telepresence robots (Tsui et al., 2011a). It refers to systems that allow a person to have a real-time sensation of being in a place other than where they are and to interact with the remote environment.

However, despite the growing interest in telepresence, a research gap exists in understanding how to design and evaluate telepresence interfaces (Wonsick and Padir, 2020), especially in MRO activities. This research addresses the primary challenge of developing, implementing, and assessing HRIs, given the telepresence capability in remote maintenance tasks.

1.3 Motivation behind telepresence for maintenance context

The advent of humanoid robots operating within human environments represents a promising step towards the next generation of autonomous systems. While existing robotic systems, equipped with rigid joints and complex pre-programmed instructions, perform well in simple, repetitive tasks, they often must improve when faced with activities outside their set parameters. Despite the integration of artificial intelligence, these machines need help executing complex, intellectually demanding activities, often requiring human intervention and oversight.

This underscores the need for and potential benefits of surrogate robots. Designers may create some surrogate robots to be humanoid, while others may not follow this design. Surrogate robots, a specific subclass of humanoid robots, are designed to serve as human stand-ins, enabling users to perform tasks remotely. This is achieved by mirroring human actions and transmitting sensory data to the operator. This unique trait makes surrogate robots an attractive solution for scenarios that require human-like performance yet pose risks for direct human involvement.

However, while the concept of surrogate robots is compelling, the challenge lies in achieving a degree of autonomy and adaptability that matches human

capabilities. Telexistence may pave the way for more effective collaboration between human operators and remote robot platforms, offering safer and more efficient alternatives to tasks in hazardous environments.

The motivation for this research lies within these challenges. By improving telexistence capabilities and understanding the optimal design and evaluation of interfaces for remote tasks, this work strives to bring us closer to a future where the current robotic platform, equipped with advanced telexistence capability, becomes familiar and used.

Later in the thesis, the potential benefits of telexistence capabilities for marine engineers are discussed. Their responsibilities range from performing routine maintenance tasks—such as inspections, cleaning, and lubrication—to engaging in more complex problem-solving activities, such as diagnosing and troubleshooting engine issues through advanced condition monitoring techniques. These tasks are conducted in spaces that contain the ship's primary propulsion machinery, fuel units, generators, boilers, pumps, and other auxiliary equipment, as illustrated in Figure 1-2. The space is typically located on the ship's lower decks and designed to be heat and sound insulated. The machinery space is a complex environment with high noise levels, high temperatures, and potentially dangerous machinery. Regular maintenance and inspections are required to ensure the ship's systems operate efficiently and safely. Traditional



Figure 1-2 Machinery spaces of the XV Patrick Blackett
(www.navylookout.com)

vessels are manned ships, meaning they have a crew onboard to operate and maintain the ship's systems. The crew is responsible for navigating the ship, operating the engines, performing maintenance and repairs, and handling emergencies. The crew also manages cargo operations and ensures the safety and well-being of all persons on board.

In contrast, unmanned or autonomous ships are designed to operate without a crew. They use advanced technologies like artificial intelligence, machine learning, and remote sensing to navigate, operate machinery, and perform maintenance tasks. Unmanned vessels can be remotely controlled from a land-based control centre or operate autonomously using pre-programmed instructions.

The primary distinction between traditional vessels and unmanned vessels lies in crew presence. While unmanned vessels can operate without an onboard crew, reducing the risks linked to human error on the vessel and potentially enhancing efficiency, they introduce different challenges. Notably, there are risks associated with errors from humans who remotely operate or monitor these vessels. These remote operators, despite not being physically present on the vessel, can still make critical decisions and actions that impact the vessel's operation. Therefore, understanding and mitigating these remote human error risks is crucial in unmanned vessel operations. However, they require advanced technologies and robust communication systems to operate effectively. On the other hand, traditional vessels rely on the skills and experience of the crew. In this context of remote maintenance with telexistence capability, the machinery space of an unmanned vessel would be equipped with sensors, cameras, and robotic systems that can be controlled remotely. For unmanned vessels, telexistence Condition-Based Maintenance (CBM) could revolutionise maintenance operations (illustrated in Figure 1-3). By enabling engineers to virtually inhabit and interact with the machinery spaces of vessels from a remote location, telexistence technologies can overcome the challenges posed by the absence of onboard personnel. This approach enhances the safety and efficiency of maintenance operations and enables immediate response to machinery issues,

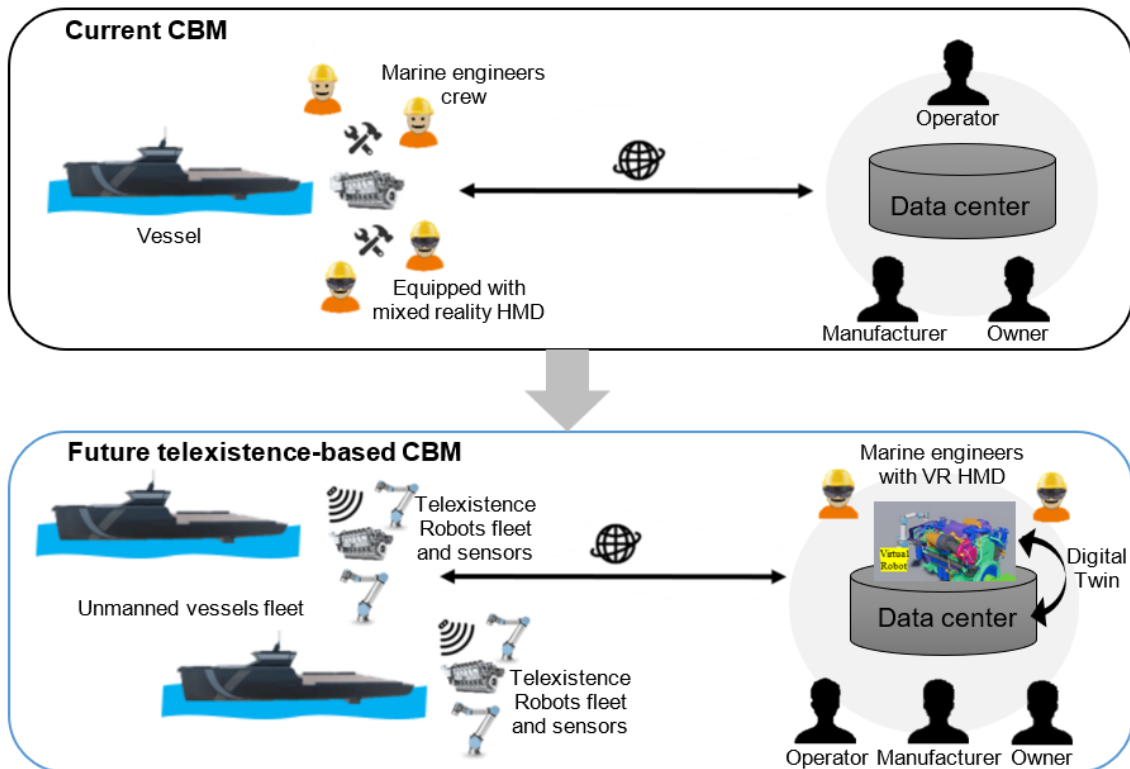


Figure 1-3 Current and future CBM of vessels

thereby minimising downtime, and potential damage. Furthermore, the real-time visual and sensory feedback provided by telexistence technologies can enhance the accuracy and effectiveness of condition monitoring, leading to more precise and timely maintenance interventions.

The maintenance of a vessel's machinery space usually involves a series of activities conducted by marine engineers (Kandemir and Celik, 2021). Remote maintenance can improve the efficiency and safety of maintainers in charge of unmanned ships. It can also reduce the time and cost associated with maintenance, as the maintainer can perform multiple tasks simultaneously and does not need to travel to the ship. The defence sector is interested in deploying Unmanned Vehicle Systems (UVS) to limit human resource use in hostile operational environments (Finn and Scheduling, 2010), but UVS must be maintained. The remote maintenance solution strives to reduce maintenance duration and make it more effective.

1.4 Aim and objectives

This thesis investigates and assesses the fundamental principles for conducting remote inspections using telexistence capability. It addresses the need for efficient and safe remote maintenance in the maritime and defence industries by applying telexistence technology. Specifically, it targets the field of remote visual inspection tasks requiring precision and environmental awareness in hazardous environments, such as the ones that operate unmanned vessels.

The aim of this research is to **develop a repeatable process for designing and deploying telexistence capability that enhances the sense of presence in remote maintenance contexts**. This, in turn, is intended to enable effective collaboration and interaction between human operators and robotic systems within virtual environments. It addresses the earlier problem of needing more efficient and safer HRIs for remote maintenance tasks in hazardous environments. The hypothesis proposes that providing telexistence to the remote maintainer will improve the performance and productivity of a teleoperated robot system (Maciaś et al., 2020; Naceri et al., 2021; Shin et al., 2021).

The first objective through the literature review is to analyse the limitations and challenges of traditional teleoperation systems in remote inspections, with a focus on the labour-intensive nature of conventional inspection techniques and the inherent lack of comprehensive sensory feedback in current teleoperation approaches. This review will identify the critical factors that impact operators' effectiveness, safety, and cognitive workload in remote inspection scenarios, particularly in marine vessel maintenance. By synthesising findings from previous research, the objective is to highlight the necessity for enhanced teleoperation strategies incorporating advanced immersive technologies like VR, AR, and MR to improve situational awareness, reduce cognitive burdens, and ensure a safer and more intuitive remote operation environment. The goal is to collate design criteria and best practices that address the identified limitations, paving the way for developing teleoperation interfaces and control systems that offer comprehensive environmental feedback and support efficient decision-making processes in remote inspections.

The second objective is to create a framework that allows for the consistent and scalable design and implementation of telepresence capability. This process aims to enhance the sense of presence for human operators in remote maintenance contexts, enabling effective collaboration and interaction between human operators and robotic systems within virtual environments. The goal is to address the need for more efficient and safer human-robot interfaces for remote maintenance tasks in hazardous environments. It involves assessing whether a telepresence-based framework can provide a sense of presence in a virtual environment and allow operators to manage an industrial robotic arm's digital representation seamlessly. The focus is evaluating the natural interaction between the operator and the robotic arm, emphasising the importance of communication protocols and system architecture to support intuitive and uninterrupted control.

The third objective is to compare different interface modalities, particularly immersive and non-immersive interfaces, in their effectiveness for remote visual inspection tasks. This objective includes applying a human-centred methodology to evaluate the interfaces, considering factors of presence, workload, usability, and performance metrics. The goal is to determine which interface type best supports telepresence capability and enhances the user experience in remote maintenance settings.

The fourth objective is to study how various data visualisation approaches (such as 2D feeds, 3D point clouds, augmented reality, and mixed modalities) impact the cognitive workload of the maintainer. This involves assessing the operator's telepresence capability associated with these choices, how they affect task performance and the sense of presence. The study aims to identify visualisation modalities that optimise cognitive workload without compromising task efficiency and safety in remote maintenance tasks.

1.5 Research questions

The main research question of this PhD research is the following:

How can telexistence capability improve remote inspections' efficiency, user experience, and safety in maintenance settings?

Three research questions were formulated from this overarching question: RQ1, RQ2, and RQ3.

RQ1: Can a scalable telexistence-based framework provide seamless interaction with an industrial robotic arm?

RQ1 is addressed in Chapter 4, focusing on evaluating the natural interaction between the operator and the industrial robotic arm through a telexistence-based framework. The central premise assesses if this framework fosters a genuine sense of presence in a virtual environment and permits operators to manage its digital representation seamlessly. The investigation begins by exploring the enablement of seamless interaction with industrial robotic arms. The term "seamless" in the context of integrating telexistence with industrial robotics refers to creating an interaction that is smooth, uninterrupted, and without any perceivable discontinuities. The telexistence system should integrate with industrial robotics naturally and intuitively. Operators should be able to control and interact with the robotic systems as if they were in the same physical space, without needing to overcome technical hurdles or adapt to complex interfaces.

It goes beyond sending commands to the system, as it establishes seamless communication between humans and machines. Multiple factors influence this relationship. This paper analyses VR's capability to support real-time visualisation and interaction with industrial robots designated for maintenance tasks. A notable outcome of this analysis emphasises VR's potential to enhance the user's sense of presence and immersion.

Despite VR's advanced features, the foundation of an effective system rests on the communication protocols in place. Various protocols were analysed to identify the one that best facilitates coherent communication between the operator, sensors, and the robot. This investigation is imperative, given that communication quality can profoundly impact the intuitiveness of interactions.

Finally, a performance analysis was conducted on the proposed system architecture and industrial robot use case.

RQ2: How does the interface choice impact telepresence capability during the remote visual inspection?

This question informed the second main objective of Chapter 5, which is to provide a human-centred evaluation of interfaces for telepresence in remote visual inspection tasks. Two interface modalities, VR and 2D (immersive and non-immersive VR), were investigated using a design methodology that involved real-time feedback, hybrid simulation settings, and gamification. A baseline non-immersive VR interface (2D), alongside the immersive one (VR), was proposed for a human-centred evaluation and comparative approach. This approach addressed the challenges in remote maintenance scenarios within unmanned vessels. A methodology that combines gamification, real-time feedback from a physical industrial robot, and a hybrid simulated setting was introduced. This combination offered an immersive and realistic environment for evaluating interface designs and enabled the collection of valuable user feedback. Factors considered included user experience elements such as presence, workload, and usability and performance metrics such as task completion time and collision rates. A systematic comparison of VR and 2D interfaces provided insights into their strengths and weaknesses for remote visual inspection tasks. This comparison was expected to inform the design of more effective telepresence systems for remote visual inspection. The validation user study was conducted on remote visual inspection in machinery spaces, demonstrating the practical applicability and effectiveness of the proposed methodology. Based on the experimental insights, recommendations for enhancing the implementation of telepresence capabilities in remote maintenance scenarios were presented.

RQ3: How does data visualisation impact the telepresence capability of the maintainer at a cognitive workload level?

The third research question investigated the impact of various interface design choices in the real-world remote operation context. It assessed the telepresence capability of the operator associated with these choices. In Chapter 6, a user

study was conducted employing four modalities, including 2D feeds, 3D points clouds, augmented virtuality (OpenCV marker in the real environment to retrieve position data information of elements), and mixed modalities to present information to the maintainer and evaluate these design choices and their effects on cognitive workload, impacting presence (Harris, Wilson and Vine, 2020). Task performance and the cognitive workload, an essential factor in developing efficient telepresence capability that affects the sense of the operator's presence, were studied (Harris, Wilson, and Vine, 2020). The first modality served as a baseline, where users relied solely on a live 2D camera feed to perform the inspection task. In the second modality, users were given a real-time 3D point cloud visualisation of the remote environment. The third modality was a hybrid visualisation combining the live 2D camera feed and the 3D point cloud. Finally, the fourth modality used augmented virtuality. The experiment included novice participants in robotic programming, and to control potential confounding factors such as learning effects or fatigue, four different orders of modalities were used. The data visualisation with the associated real environment was set up with safety measures for participants and equipment. Participants underwent a trial period offline on a simulator to learn how to manage the robot in VR. Task performance was measured regarding inspection accuracy, operation time, and cognitive load via average pupil dilation and a cognitive load questionnaire. Finally, the results from the study were discussed, considering faster task completion times, lower error rates, and reduced mental load indicators of a better telepresence capability.

1.6 Research contributions and publications

The research contributes to the field of telepresence and remote maintenance by developing a repeatable and scalable process for designing and deploying telepresence capabilities that enhance presence in remote maintenance tasks. Through rigorous research and experimentation, it provides comprehensive, evidence-based recommendations for implementing immersive technologies in controlling industrial robots for remote inspection. Addressing three sub-research questions, the study validates a scalable telepresence framework for seamless industrial robotic interaction, compares non-immersive and immersive interface

impacts on telexistence capability, and assesses how data visualisation affects cognitive workload during remote maintenance. Key findings reveal the superiority of VR interfaces in enhancing user engagement and task efficiency, the importance of bi-directional communication in telexistence systems, and the nuanced impact of data visualisation modalities on cognitive workload without compromising task performance. These insights offer significant implications for designing and implementing future telexistence systems, promoting enhanced operational efficiency, safety, and user experience in industrial maintenance tasks.

In summary the contributions of this thesis are the following:

- Developed a novel teleoperation framework for industrial robotic arms with telexistence capabilities.
- Investigated the use case of remote visual inspection with telexistence capabilities.
- Investigated the added value of a VR-based teleoperation interface for remote visual inspection over the traditional 3D interface.
- Investigated the effects of data visualisation modalities on the operator's cognitive workload with pupillometry during VR-based teleoperation and remote visual inspection tasks.

The research presented in this thesis resulted in the following papers:

- **Mazeas, D.**, Erkoyuncu, J.A., Noël, F. (2023). A Telexistence Interface for Remote Control of a Physical Industrial Robot via Data Distribution Service. In: Noël, F., Nyffenegger, F., Rivest, L., Bouras, A. (eds) Product Lifecycle Management. PLM in Transition Times: The Place of Humans and Transformative Technologies. PLM 2022. IFIP Advances in Information and Communication Technology, vol 667. Springer, Cham. https://doi.org/10.1007/978-3-031-25182-5_38
- **D. Mazeas**, J. A. Erkoyuncu and F. Noël, "Telexistence-Based Remote Maintenance for Marine Engineers" 2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), Shanghai, China, 2023, pp. 407-413, doi: <https://doi.org/10.1109/VRW58643.2023.00089>
- S. C. Medina Galvis, **D. Mazeas**, F. Noël and J. A. Erkoyuncu, "Designing Immersive Tools for Expert and Worker Remote Collaboration" 2024 CIRP Design Conference, UK, Cranfield [accepted]
- **D. Mazeas**, I. Greening, F. Noël, and J. A. Erkoyuncu, "Enhancing remote inspection in ship machinery spaces with telexistence capability" 2024 International Naval Engineering Conference, Liverpool, United Kingdom [accepted]

1.7 Thesis structure

The thesis has the following structure:

- **Chapter 1** presents the PhD's context, problem statement and motivation. This chapter also presents the aim and objectives of the central research questions of this PhD.
- **Chapter 2** presents the current knowledge in VR and telexistence, key terminologies, the design of telexistence systems, its application in remote maintenance, and existing research gaps.
- **Chapter 3** presents the research methodology for addressing the objectives and research questions, focusing on enabling telexistence capability, evaluating remote visual inspection interfaces, and assessing the impact of data visualisation modalities.
- **Chapter 4** presents a comprehensive framework for telexistence capability and discusses developing and implementing a novel framework for telexistence capability, with a case study demonstrating its application in a marine engineering context.
- **Chapter 5** evaluates immersive and non-immersive interfaces for remote visual inspection, including the system refinement process, data collection and analysis, results, and discussion.
- **Chapter 6** presents the investigation of the influence of data visualisation on telexistence systems, outlining the procedure, system refinement, data collection and analysis, results, and discussion.
- **Chapter 7** provides an overview of the principles for assessing and implementing remote inspection with telexistence capability. It also presents key actors in the telexistence industry and reflects on future steps for maintenance powered by telexistence capability.
- **Chapter 8** provides the conclusions of the PhD, addressing the objectives and answering the research questions.

2 Literature review

2.1 Virtual reality and telexistence

2.1.1 Introduction

VR is a unique digital frontier, offering a comprehensive, three-dimensional, computer-generated environment (Jason Jerald, 2015). Unlike conventional user interfaces that limit users to passive viewing, VR encourages active participation, inviting users to enter and interact within a dynamic 3D environment. This environment can replicate the real world or present an entirely different, artificial reality (Petroski, 2018; Slater, 2003). By mimicking sensory inputs – such as visual, auditory, and occasionally tactile sensations – VR transcends traditional computing, taking the user into an immersive realm where the only constraints are computational power and content availability (Gobbetti and Scateni, 1998).

VR can be broadly classified into three primary categories: non-immersive, semi-immersive, and fully immersive simulations.

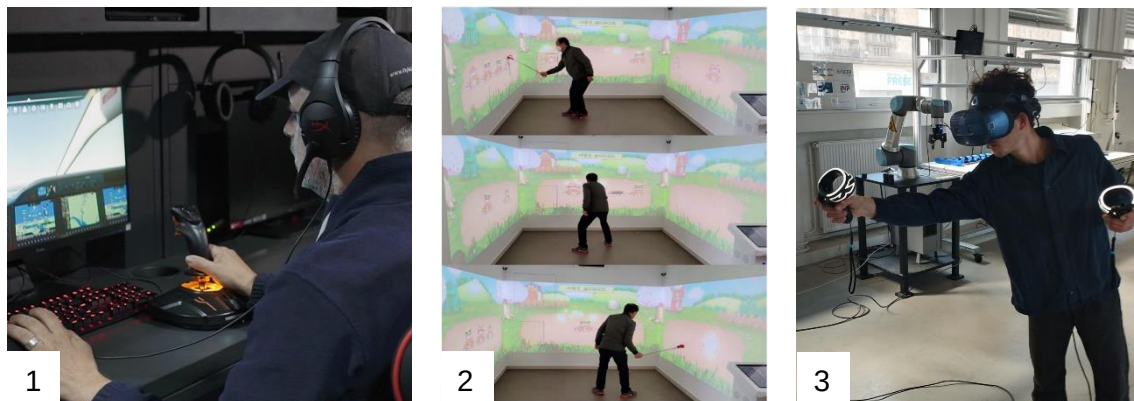


Figure 2-1 VR categories: non-immersive VR (1), semi-immersive VR (2), and fully immersive VR (3)

Non-immersive VR, often underestimated as a category of VR, is already a part of our day-to-day experiences. It creates a computerised environment while allowing users to remain aware and in control of their physical surroundings. Conventional devices like computers or gaming consoles, displays, and input peripherals like keyboards, mice, and controllers typically facilitate the interaction

with non-immersive VR (Wu, Houben and De Bra, 2000). A typical video game is an apt illustration of a non-immersive VR experience.

Semi-immersive VR, on the other hand, offers users a partially virtual environment. It provides an impression of being in an alternate reality when the user focuses on the digital content but still retains a connection with their physical environment (Alqahtani, Daghestani, and Ibrahim, 2017). Detailed 3D graphics or vertical reality depth enhances the authenticity of semi-immersive VR. This category of VR is frequently employed in education and training (Lorusso et al., 2020). Combining high-resolution displays, powerful computers, projectors, and complex simulators enables users to engage with the system through hand movements. It occasionally requires them to wear glasses or data gloves.

Lastly, fully immersive VR provides the most authentic simulation experience, encompassing visual and auditory sensory inputs. Users require specialised equipment, such as VR glasses or an HMD, to engage with this immersive VR. The display is typically divided between the user's eyes, generating a stereoscopic 3D vision, and combined with input tracking to create a convincingly immersive experience (Mihelj, Novak and Beguš, 2014).

VR's immersive capabilities, which encapsulate users in a digitally simulated world, synergise seamlessly with the concept of telepresence (Tachi, 2015b). This technology enables users to have a real-time sensation of being in a place other than where they are and to interact with the remote environment. This convergence broadens the scope of user experience and potentially transforms numerous sectors, especially in remote work, by transcending spatial limitations. In this thesis, the primary emphasis will be on fully immersive VR as part of the telepresence capability, exploring its various aspects, potential applications, and challenges in implementation. Unless stated otherwise, the term "VR" in the following chapters will refer to fully immersive VR.

2.1.2 A brief history of virtual reality

In the early 1960s, Morton Heilig developed the Sensorama, a multimedia artist who envisioned the "Cinema of the Future." The Sensorama was a mechanical

device like an arcade-style video game cabinet. It was designed as a "theatre cabinet" that stimulated multiple senses to fully immerse the individual in the experience (Heilig, 1962). Later, in 1968, Ivan Sutherland proposed a concept that set today's VR's foundation. Sutherland envisioned an "Ultimate Display" that could simulate reality to the point of being indistinguishable from it. It would involve a virtual world that could *"look real, sound real, feel real, and respond realistically to the viewer's actions."* His concept was about visual representation and sensory experiences, such as force feedback and sound (Sutherland, 1968).

The period between 1975 and 1990 witnessed the development of several ground-breaking projects for VR (Gobbetti and Scateni, 1998). The creation of VIDEOPLACE by Myron Krueger in 1975 can be seen as a precursor to modern VR systems. This technology projected the silhouettes of users on a large screen, enabling interaction with each other through image processing techniques that computed their positions in a 2D space. This was followed by the Visually Coupled Airborne Systems Simulator (VCASS) development by Thomas Furness at the US Air Force's Armstrong Medical Research Laboratories in 1982. VCASS was a revolutionary flight simulator that enhanced a fighter pilot's view with graphical overlays displaying targeting or optimal flight path data. The technology used an HMD, a significant advancement in VR technology at the time. Two years



Figure 2-2 NASA VIVED and DataGlove

later, in 1984, the NASA Ames Research Center introduced the Virtual Visual Environment Display (VIVED). This was a stereoscopic monochrome HMD built using readily available technology, further progressing the use of VR in diverse settings. The mid-eighties saw the emergence of VPL Research, a company that manufactured the DataGlove (1985) and the Eyephone HMD (1988). VPL Research democratised the term "virtual reality" in everyday language. These were among the first VR devices commercially available, marking a milestone in the commercialisation and accessibility of VR technologies. The decade's end brought about the Binocular Omni-Orientation Monitor (BOOM) commercialisation by Fake Space Labs in 1989. BOOM, a small box containing two CRT monitors viewable through eyeholes, allowed users to navigate the virtual world while a mechanical arm tracked the position and orientation of the box. The UNC Walkthrough project was undertaken in the late 1980s at the University of North Carolina. This project centred around the development of an architectural walkthrough application. It led to the creation of several VR devices, such as HMDs, optical trackers, and the Pixel-Plane graphics engine, to enhance the system's quality.

Carolina Cruz-Neira demonstrated the CAVE (Cave Automatic Virtual Environment) system in 1992 (Figure 2-3). The CAVE provided a much more realistic VR experience, notably a holodeck allowing individuals to see their bodies in the room. The system included rear projection screens, a downward

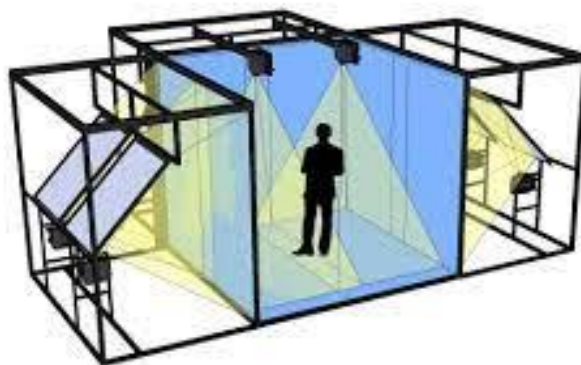


Figure 2-3 CAVE systems concept

projection system, and an upward projection system to create the illusion of reality in the physical space. A person standing in the CAVE and wearing 3D glasses could see floating objects and move around the room (Cruz-Neira et al., 1992).

The primary objective of CAVE was to solve a significant problem: early versions of HMDs were bulky. They posed practical restrictions, mainly when applied in research. The pioneering iteration of this technology used electromagnetic sensors to monitor movements and, later, infrared technology. The motion capture software utilised data from sensors embedded within the 3D glasses and the room to track movements, consistently adapting, and adjusting the video accordingly. The CAVE's projection system ensured accurate images were perceived in each eye by continuously synchronising the glasses. Various speakers also equipped the room with 3D audio. The result was an enveloping virtual space, facilitating the user's interaction and control over 3D objects, more akin to the portable and mobile systems we see today. The CAVE concept achieved rapid success.

The new generation of VR devices was introduced in 2013 with the Oculus Rift DK1, followed by the Oculus Rift DK2 in 2014. Oculus was an American start-up founded by Palmer Luckey, who, with John Carmack, started to develop the Oculus Rift in 2012. It was acquired by Meta (Facebook) in 2014 for \$2 billion. In 2016, HTC, a Taiwanese mobile phone manufacturer, partnered with the US game publisher Valve to manufacture a commercial VR solution: the HTC VIVE. This system employs an "outside-in" approach, with several external units, known as "lighthouses," emitting beams of laser and LED lights. These beams are received by sensors located on the headset, enabling the calculation of their position and orientation within the room with 6DoF (Figure 2-4).

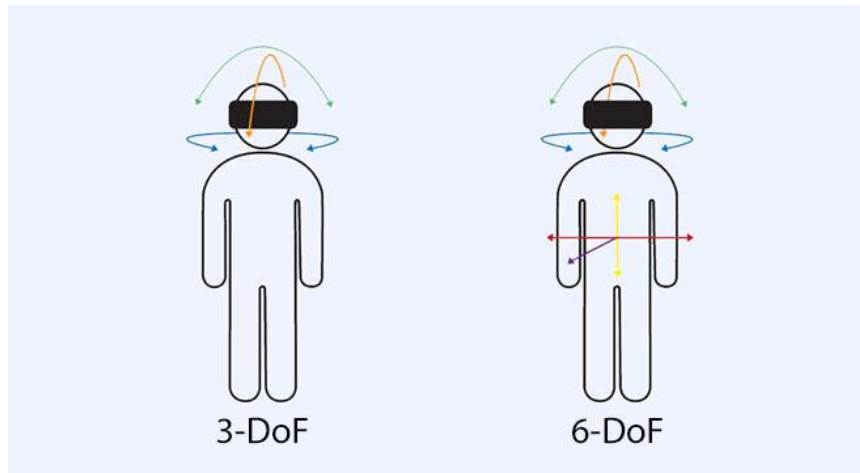


Figure 2-4 Comparison between 3Dof and 6Dof

The HTC VIVE has been sold at prices under one thousand dollars (compared to the 100,000 dollars cost for a CAVE), making VR more accessible. It is estimated that between 2016 and 2023, approximately 26 million VR HMDs were sold, according to Statista (Statista, 2023). Microsoft introduced the HoloLens in 2016, marking a significant advancement in mixed reality (MR) technology. Unlike VR headsets, the HoloLens combines elements of the physical and digital worlds, allowing users to interact with holographic images overlaid on their real-world environment. This was a step towards seamlessly blending virtual and real-world elements, providing a unique MR experience. The Apple Vision Pro MR device developed by Apple Inc. was released on February 2, 2024. The Vision Pro is heralded as a "spatial computer," integrating digital media with the physical world through advanced features such as 6DoF inside-out tracking, photogrammetry, hand gesture recognition, eye tracking, and voice input, all powered by Apple's M2 and R1 chips. With its high-resolution dual micro-OLED displays, providing approximately 3660x3200 pixels per eye, and a unique "EyeSight" system for external interaction.

X-Reality (XR), today's term for encapsulating all immersive and interactive 3D technologies, has sparked substantial interest in research and industry, primarily due to the promising avenues they open for human interfacing with machines. From constrained to simulated training experiences, XR technologies function as communication, control, and creation conduits. They encompass a spectrum of

real and virtual environments and human-machine interactions engendered through advancements in computer technology and wearable devices (Dwivedi et al., 2022; Rauschnabel et al., 2022). More recently, XR technologies have served as gateways to three-dimensional (3D) society-like constructs or "Metaverses", integrative spaces blending real and virtual environments (Dwivedi et al., 2022).

The concept of VR, however, is often a subject of debate owing to the multitude of definitions found in the literature, thereby leading to potential ambiguities. Some individuals equate playing a video game on a 2D screen or viewing 360° videos on an HMD to experiencing VR. However, VR fundamentally provides an avenue for real-time interaction within a virtual environment or 3D space. This realm may or may not simulate the scale of the physical world but offers a virtual platform for humans to project, act, and interact freely. Furthermore, incorporating features like stereoscopic vision, head/hands/body tracking, and sound accentuates the immersive experience. Mihelj et al. describe VR as an interactive computer simulation that amplifies the user's sense of presence, providing a sense of immersion in the virtual environment (Mihelj, Novak and Beguš, 2014). Sherman et al. outlined four critical components of VR: the virtual environment, virtual presence, sensory feedback, and interactivity (Sherman and Craig, 2003). VR represents a computer-engineered simulation of an alternative world or reality, crafting an immersive and interactive experience.

Although VR has witnessed a surge in popularity in recent years, primarily due to commercially available and affordable VR headsets, the concept has deep roots, with its evolution tracing back several decades. Computer graphics have become integral to various facets of our lives, especially architecture, engineering, and interior design (Foley, 1995). The rapid advancement of microprocessor technology has rendered computers more robust and faster, concurrently reducing costs. This democratisation of technology has empowered average users to plunge into computer graphics, often beginning with computer games and gradually expanding to explore unique dimensions and experiences beyond

real-life confines. This fascination for interaction with the virtual world, beyond simple passive observation, has led to the advent of VR.

2.1.3 Telexistence capability

Real-time interaction, 3D spaces, and self-projection are fundamental elements for a system pretending to provide telexistence capability. Telexistence is a concept first introduced in 1980 by Professor Tachi, which aims for a "highly realistic sensation of existence in remote places without actual travel" (Tachi, 2015b). VR and telexistence are the same concepts; nevertheless, VR is usually reserved for 3D immersive simulation in a virtual environment, and telexistence is for the real world. Telexistence seeks to provide an operator with a sense of presence in the real world, where a robotic system operates within a simulated virtual environment or a combination of the two.

Telexistence in the real world relies on the use of a robotic system of any kind (e.g., 3-axis arm, 6-axis arm, ground robot, aerial robot, under-water robot, humanoid) to facilitate the operator's presence in the actual remote environment and enable him to perform a task most efficiently. Additionally, the real environment and the robot must be enhanced using sensors to provide the operator with the best real-time sensation of being present.

Telexistence can be wrongly associated with telepresence robots, as illustrated in Figure 2-5, which refers to telecommunication via video-sounds-based technology and mobile platforms for conferencing (Koh, 2010), social activities



Figure 2-5 Telepresence Robots

with friends and family or under exceptional circumstances such as during the COVID-19 pandemic (Isabet et al., 2021). There is not a widespread consensus on the definitions of telepresence; Marvin Minsky is often cited with the first definition of telepresence as “presence in a remote environment through technology” (Minsky, 1980). Telepresence robots are also used by people with physical or cognitive disabilities (Shishehgar, Kerr and Blake, 2019). The disadvantage of telepresence robots is that they do not give users the sensation of self-existence in a remote place.

Telepresence robotics has been used for many years to help remote workers by providing a virtual presence in any location at any time. Telepresence's main application is teleconferencing, which is used in educational situations where a robot can replace the teacher or a student. However, telepresence systems provide a limited sensation of being present in the virtual environment or real life through a virtual environment because of the lack of awareness in the remote workplace. Telepresence robots usually only provide visual and sound feedback with the ability to move around in some instances (Tsui et al., 2011b).

Telexistence, however, has a broader scope of usage, including the virtual environment, real environment, and remote robotic systems. In the book *Telexistence* (Tachi, 2015c), Professor Tachi describes two types of telexistence:

- Standard telexistence technology seeks to replicate, in real-time or with minimal delay, the presence of an operator by utilising a robot equipped with human-like visual and sensory functions.
- Augmented telexistence aims to provide enhanced functions to the operator when the tasks or activities are different and must overcome human capabilities. Furthermore, this encompasses scenarios where the distance or network limitations prevent real-time communication between the operator and the robot. Augmented telexistence can be applied under specific use cases in MRO, such as using more advanced sensors.

Despite the similarity between telepresence and telexistence, telepresence has been the most used term by researchers in robotics, who have established it as the sensation of being "transported" to a distant physical location through

technology. In the early 1990s, this idea was extended to VR; before being transported to a remote physical environment, the operator is immersed in a virtual environment with the sensation of being in the place generated and represented by VR devices (Held and Durlach, 1992). The term "presence," used as an abbreviation for telepresence, was quickly adopted by the scientific community and evolved from its initial meaning of "transportation" to a distant existing place to the idea of "being there" in a virtual environment real or not (Steuer, 1992). Slater explores the psychological impact of immersive VR on body representation, demonstrating that embodying different virtual bodies can alter perception, attitudes, and behaviour. It introduces the concepts of place illusion and plausibility illusion as mechanisms by which VR enhances the sense of presence, leading to realistic responses within virtual environments (Slater and Sanchez-Vives, 2014).

Over time, presence has become a multidisciplinary concept, encompassing not only researchers in VR and robotics but also those in media and communication (Biocca, 2006), industrial engineering (Draper, Kaber and Usher, 1998), or cognitive sciences (Sanchez-Vives and Slater, 2005). The increasing multidisciplinary nature of the field has resulted in the development of fresh interpretations of the concept tailored to the specific domains of researchers and their goals.

2.1.4 Key terminologies

Key terminologies of telepresence include:

- **Virtual Environment:** Interactive computer-generated environment created by a VR system (Stuart, 1996). This term is essential because it defines the foundational setting in which telepresence operates. Understanding the characteristics and limitations of virtual environments is crucial for exploring how telepresence functions within these spaces.
- **Immersion:** A distinct concept from presence, immersion refers to the objective level of sensory fidelity a VR system provides (Slater, 2003). The degree of immersion a virtual environment provides directly impacts the

effectiveness and experience of telepresence. It is an essential metric for assessing the quality of a virtual experience.

- **Presence:** A subjective experience of being in an environment or place, even when physically situated in another location (Witmer and Singer, 1998). Differentiating presence from immersion is critical. Presence is about the subjective experience and feeling of being in a different place, a core aspect of telepresence, highlighting technology's psychological and perceptual aspects.
- **Interaction:** Refers to the natural interaction between the user and the virtual scene. Feedback gives users the same sensation as in the real world (LI et al., 2019). Focusing on interaction emphasises the dynamics between the user and the virtual environment. Understanding how users can manipulate and navigate within a telepresence system is essential.
- **Operator:** User that uses their intelligence and experience about the operating situation by working at a distance. Defining the operator's role provides insight into the human aspect of telepresence systems, focusing on the user's experience and responsibilities.
- **Remote control systems:** The system interfaces with the operator, receives messages from the sensors at the remote location, and controls the robot to perform the required operation based on the immersive virtual environment in which the operator evolves. This term is crucial for understanding the technical framework that enables telepresence. It encompasses the hardware and software facilitating remote operation, a cornerstone of telepresence technology.
- **Sensors:** Devices in charge of collecting several types of data remotely or not. Sensors are fundamental components in telepresence systems, as they gather crucial data that inform and drive the virtual experience. Understanding their role is critical to comprehending how telepresence systems interact with the real world.
- **Robot:** A system that receives commands from the operator in different manners through the remote-control system to mimic the operator's action. Robots are often the physical embodiment of telepresence systems in the

real world. They are the agents through which operators interact with distant environments, making this a key term.

- **Human-robot interaction:** Human-robot and robot-human communication can be audio, visual, or any other data, including haptics or tactile (Yanco and Drury, 2004). This term is vital as it explores the interface between humans and robots, which is central to telexistence. It encompasses the methods and challenges of communication and control in these systems.
- **Human-robot collaboration:** "Working jointly with others or together, especially in an intellectual endeavour" can define the collaboration between the robot and the human (Nass, Steuer and Siminoff, 1994). It is essential to differentiate between interaction and collaboration. Collaboration implies a more complex, two-way relationship, increasingly relevant in advanced telexistence systems.
- **Head-Mounted Display:** A head-mounted display device worn on the head or as part of a helmet with a small display optic in front of one or both eyes (Sutherland, 1968). As a primary interface for many telexistence systems, understanding head-mounted displays is crucial for discussing how users visually engage with virtual environments.
- **Haptics:** The concept of haptics encompasses both actual and simulated tactile engagements involving robots, humans, and environments that may be real, remote, or virtual. This includes a variety of interactions across different scenarios and combinations, highlighting the diverse nature of touch-based communication and manipulation in technologically advanced settings. Haptics can be categorised into three primary types: tactile, force feedback, and kinesthetic. Tactile haptics involve surface sensations such as texture, temperature, or vibration, providing detailed information about the contact surface. Force feedback haptics deliver resistance or pressure against the user's actions, enabling them to feel virtual objects' weight, inertia, or resistance. Kinesthetic haptics focuses on the position, movement, and force exerted by the user's muscles and

joints, offering a sense of body movement and spatial orientation in the virtual environment (Hayward et al., 2004).

- **Augmented reality:** While AR alone does not constitute teleexistence, it significantly enhances the capabilities and experiences possible within teleexistence systems. The combination of AR with robotics, VR, and sensory feedback technologies is critical to developing advanced teleexistence systems that provide rich, immersive, and interactive experiences of being present in a remote environment. AR can offer additional information and interaction capabilities (Inoue et al., 2019), more intuitive control and feedback mechanisms for manipulating avatars (Pakanen et al., 2022), and facilitating collaborative and educational experiences across distances (Oyama et al., 2021).

2.1.5 Presence and immersion

Since 1992, when Sheridan and Furness decided to use it as the title for their journal dedicated to exploring virtual reality systems and teleoperations, the term "presence" has attracted interest within the field of VR research.

Researchers in various fields have suggested a wide range of presence dimensions. For instance, Reiner and Hecht have defined "object presence" as the feeling of being able to touch and manipulate virtual objects in haptic virtual environments (Reiner and Hecht, 2009). Other dimensions include "environmental presence" (Heeter, 1992), "connected presence" (Schroeder, 2006), and "temporal presence" (Power, 2016). However, teleexistence focuses on the three dimensions of "self-presence", "social presence", and "spatial presence" due to their relevance in the domains of robotics and VR (Schultze, 2010).

Presence questionnaires are designed to quantify the participant's level of presence (Schwind et al., 2019). This can be a complex construct to measure, as it is inherently subjective. They allow researchers to collect data in a standardised form, enabling comparison across different studies or virtual environments. The degree of presence can significantly impact the user's experience, affecting factors such as engagement, performance, and emotional response. The results

from presence questionnaires can inform designers and engineers about the effectiveness of virtual environment design elements, leading to technological improvements. Different questionnaires may focus on various aspects of presence, such as physical presence (the sense of being physically situated in a virtual space), social presence (the sense of social interaction with virtual or real agents), and co-presence (the sense of being in the company of others). By analysing questionnaire responses, researchers can understand how different variables, such as immersion and interactivity, contribute to the sensation of presence.

Apart from the subjective assessment of presence offered by the questionnaires. Presence can also be measured objectively. Previous studies have used physiological measurements like skin conductance (Caldas, Aviles, and Rodriguez-Guerrero, 2020), heart rate (Kiryu and So, 2007) and electroencephalography (Grassini et al., 2021).

Immersion is sometimes defined as “a quality of the system’s technology, an objective measure of the extent to which the system presents a vivid virtual environment while shutting out physical reality” (Diemer et al., 2015; Slater and Wilbur, 1997). Immersion is a mental state where an individual feels encompassed by a part of and actively engaged with an environment that delivers an unbroken flow of stimuli and experiences (Witmer and Singer, 1998).

2.1.6 Factor contributing to the presence

A VR system is deemed immersive, including the user's body within virtual settings. This immersion can foster a sense of presence, where the user feels like they are in the virtual environment. Table 2-1 Categorises factors contributing to a sense of presence as suggested by Witmer and Singer (Witmer and Singer, 1998).

Table 2-1 Suggested factors contributing to the sense of presence (Witmer and Singer, 1998)

Control Factors	Sensory Factors	Distraction Factors	Realism Factors
Degree of control	Sensory modality	Isolation	Scene realism
Immediacy of control	Environmental richness	Selective attention	Information consistent with the objective world
Anticipation of events	Multimodal presentation	Interface awareness	Meaningfulness of experience
Mode of control	Consistency of multimodal information		Anxiety/disorientation
Physical environment modifiability	Degree of movement perception		

The research work done by Witmer and Singer highlights various factors influencing the sense of presence in virtual environments. Empirical studies indicate that update rates, the field of view limitations, and meaningful arrangements in virtual settings can significantly impact the user's sense of presence. Theoretically, presence is influenced by a blend of Control, Sensory, Distraction, and Realism Factors. Control over the environment, immediate feedback, anticipation, and natural interaction modes enhance presence. Sensory factors like modality richness, multimodal coherence, and active search capabilities play a crucial role. Distractions reduce presence, while factors like scene realism, consistent information, and meaningful experiences enhance it. These insights underscore the complexity of creating immersive virtual environments and the necessity of multifaceted approaches to enhance presence (Witmer and Singer, 1998).

Moreover, Gary Fontaine defines presence as a state of consciousness with distinct characteristics. This state of consciousness includes experiences of "realness, vividness, and feeling very much alive," attention to the immediate situation, "a perception of thinking and acting in new and innovative ways," and "a broad awareness of everything around" (Fontaine, 1992). Fontaine claims that the user focus enables the individual to be broadly aware of the entire task environment and is also necessary for a high level of presence.

2.1.7 Cognitive workload and presence

Cognitive workload refers to a user's mental effort to perform a task or interact with a system. The intricacy of tasks within a virtual setting can impact how information is processed, overall performance, and user focus (Ma and Kaber, 2006). High cognitive workload can lead to user fatigue, reduced performance, and decreased overall user satisfaction. This indicates that virtual environments can influence users' perceptions (Milleville-Pennel and Charron, 2015).

The relationship between cognitive workload and presence is complex and bidirectional. On the one hand, a well-designed immersive environment that promotes a high sense of presence can reduce cognitive workload by providing intuitive interactions and naturalistic experiences. Users who feel more present in the virtual environment can often perform tasks more naturally and with less cognitive effort. On the other hand, an excessive cognitive workload can detract from the sense of presence. If a user is overwhelmed by the complexity of the tasks or the interface, their attention may be diverted from the immersive experience, thus reducing their sense of presence. This can happen when the virtual environment is cluttered, interactions could be more intuitive, or the system has a steep learning curve.

Therefore, managing cognitive workload is essential for ensuring users can interact with the system efficiently and comfortably. Numerous systems for recording pupil size are being used and depend on the application. The size of the pupils ranges between 2-8 mm. Pupillometry is a widely used response system that reflects diverse cognitive states ranging from arousal, interest, and effort (Eckstein et al., 2017). Apart from

pupillometry, other aspects of eye movement, such as blink rate, saccades, and fixations, can also provide information about cognitive workload. Electroencephalography (EEG) measures the electrical activity of the brain. Changes in brainwave patterns can indicate variations in cognitive workload. EEG is beneficial for detecting real-time changes in mental effort (Tremmel et al., 2019). Functional Magnetic Resonance Imaging (fMRI) can track brain activity by measuring changes in blood flow. Research often uses it to identify which brain areas are engaged during specific tasks (Beck et al., 2007). Near-Infrared Spectroscopy (NIRS) is a non-invasive method that measures brain blood flow and oxygenation, providing insights into cognitive activity and workload (Qu, Zhang, and Bu, 2023). Heart Rate Variability (HRV) measures changes in heart rate variability that can reflect stress and cognitive effort. HRV is especially useful when emotional or physical stress affects cognitive workload (Malińska et al., 2015). Galvanic Skin Response (GSR), also known as skin conductance, GSR measures the electrical conductance of the skin, which changes with sweat gland activity, an indicator of physiological arousal related to cognitive load (van 't Wout et al., 2017). Questionnaires and scales like the NASA Task Load Index (Harris, Wilson, and Vine, 2020; N. H. P. R. Group et al., 1987) allow users to self-assess their perceived cognitive workload. While subjective, these measures can provide valuable insights into user experiences.

Designers of immersive systems need to strike a balance between cognitive workload and presence to optimise cognitive workload. This involves creating realistic and engaging environments that foster a sense of presence while being user-friendly and not overly demanding regarding cognitive load. Simplifying interactions, providing clear feedback, and designing intuitive navigation are some strategies that can help achieve this balance.

2.2 Telexistence for remote inspection

Continuous monitoring and inspection of remote assets provide operational efficiency benefits, reducing downtime and costs associated with maintenance (Ramezani et al., 2020). Using robots designed to access and navigate facilities, combined with a diverse array of sensors, allows for comprehensive and precise

monitoring and inspection of remote assets. Today's envisioned systems are highly autonomous and conduct inspections with minimal intervention by human operators. The robots aim to achieve navigation into the remote place, accomplish crucial tasks, and sensing.

Significant strides have indeed been made in autonomous systems, but achieving high levels of autonomy remains challenging, particularly in complex or unstructured environments. Besides autonomous operations, teleoperation is another critical aspect of remote asset monitoring and inspection. This approach is particularly relevant when human judgment is essential, or the environment is too complex for full autonomy.

Teleoperation could be defined as manipulating a system to accomplish at a distance. The distance refers to the physical distance between the operator and the robotic system, whether close or far. A surgeon can use a micromanipulator system (Avgousti et al., 2016), or a nuclear engineer can use robotic arms through a protective shield at a short distance (Piolain et al., 2010). Alternatively, scientists can operate a robot on another planet at a great distance (Fong et al., 2013), implying no direct or visual contact with the teleoperated robot and communication delays. Teleoperations are composed of two main components, the human operator (master) and the remote robot (slave) in a remote location from the operator and connected through a communication channel.

(Jaselskis et al., 2015) Introduces a telepresence approach to monitor construction projects. The technique uses bonded streaming technology, enabling real-time audio and video feeds from portable cameras located anywhere on the construction site. This allows off-site team members to guide the camera operator, navigate the site remotely, and perform their responsibilities virtually or "tele-present". The combination of telepresence and telerobotics can provide a powerful solution for industries requiring regular inspections or maintenance work in hazardous or hard-to-reach areas. This solution is becoming possible thanks to VR technologies. VR is about engaging content that can be performed in real-time 3D; it can be either immersive through HMD or not through a conventional monitor, as discussed earlier. Immersive VR has evolved

rapidly since 2016, and the kick-off of commercial VR HMD devices is designed for broader consumer adoption, primarily on 360-degree video and gaming. VR brings a wide range of application areas due to its flexibility in various research and industrial use cases. VR in robotics enhances the user experience by adding immersion and presence (Mütterlein, 2018) that meet the principles of the telepresence concept. Most previous research on teleoperation has investigated VR as a helpful tool for providing operators with visual feedback on the robot and the environment via HMD or immersive projection technologies. VR headsets allow total visual immersion in 3D and simulate training situations (Bric et al., 2016).

However, few publications focus on VR as an interface for operating a remote robot with a sense of presence in that remote inspection context. Immersive teleoperation, providing telepresence capability, still presents many challenges for researchers and industrial players, resulting in a lack of implementation in companies. Effective human-robot interactions and seamless integration of data into immersive interfaces are required.

2.2.1 Immersive interface for teleoperation

XR-based robot teleoperation can be categorised on the level of user interface immersion (Figure 2-6). Table 2-2 presents an overview.

Table 2-2 Overview of immersive teleoperation interfaces

Interface type	Constraints	Use Cases
Stereoscopy	- Limited field of view, bound to stereo camera's perspective. - Risk of motion sickness from discrepancies in operator's head movements and camera agility. - Dependency on camera quality and positioning for depth accuracy. - Potential latency in image transmission affecting real-time operation.	- Procedures involving minimally invasive robotic surgery. - Virtual tours. - Road condition monitoring.
AR	- Proximity is required to overlay virtual elements in the real world.	- Close-range human-robot interaction scenarios.

	- Supervision-centric due to AR hardware limitations. - Limited field of interaction, constrained by the AR device's range. - Potential occlusion of important real-world cues by virtual overlays. - Dependence on environmental conditions for adequate overlay visibility.	- Maintenance and repair guidance in industrial settings. - Enhanced user experience in entertainment and gaming.
VR	- Inaccuracies in remote environment reconstruction using RGBD sensing. - Ambiguities in the impact of VR interfaces on precise robot control. - Developmental stage of VR reconstruction manipulation techniques. - Potential latency in image transmission affecting real-time operation. - Possibility of disorientation or VR-induced symptoms in users. - High processing power is needed to render complex environments realistically.	- Remote robot control is available in inaccessible or hazardous areas. - Training simulations for high-risk operations.

Firstly, there are stereoscopic telepresence systems with master-slave devices (Ballantyne and Moll, 2003). The operator uses a master device to control a remote slave robot. The stereoscopic system provides depth perception by presenting slightly different images to each user's eyes, creating a 3D visualisation of the remote environment. Stereoscopic technology has been instrumental in advancing road safety through model-driven 3D lane detection systems and collaborative approaches for maintaining safe driving distances (Benmansour et al., 2008; Olaverri-Monreal et al., 2019).

Secondly, AR-based teleoperation (Quintero et al., 2018; Walker, Hedayati and Szafir, 2019). In AR-based teleoperation, the operator views the real environment, usually through a head-mounted display or a handheld device, onto which computer-generated images of the remote robot's operations are superimposed. This approach allows for seamless real and virtual information integration, enhancing the operator's situational awareness and providing

intuitive visual cues for remote control tasks. In AR-based telerobotics configurations, it is typically necessary for the operator to have direct visual access to both the robot and its surrounding environment using their sight. Most AR-based systems utilise the Microsoft HoloLens headset, renowned for tracking the user's head and hands, recognising gestures and speech, and performing Simultaneous Localisation And Mapping (SLAM) of the immediate surroundings. These functionalities are particularly advantageous for supervised teleoperation. Operators can employ gestures and speech commands to set various markers in relation to the environmental geometry for the robot's motion planning. This setup allows them to preview the motion plan and then execute it. By superimposing computer-generated images of a robot's operations onto the real environment viewed through AR devices, operators gain a heightened level of situational awareness and intuitive control over remote tasks (Nuzzi et al., 2020). This enables direct visual and operational synchronisation with the remote robot and its surroundings and leverages gesture and speech recognition for precise command execution.

Thirdly, VR-based teleoperation (Omarali et al., 2020; Whitney et al., 2018a). Using a VR headset, operators can experience a simulated representation of the remote site, allowing for intuitive control and navigation as if they were physically present. This fully immersive experience is particularly beneficial in complex or dangerous environments where physical presence is impossible or unsafe. VR teleoperation systems can also incorporate haptic feedback, giving the user a sense of touch, which is crucial for tasks requiring fine motor skills and tactile sensation (Mihelj and Podobnik, 2012). VR-based teleoperation allows the display of multiple operator and robot points of view in a single virtual environment in addition to enhanced data visualisation modalities such as point cloud or 3D reconstruction, for instance. Their adaptive interfaces and customising controls distinguish VR teleoperation systems and display specific user needs and tasks, enhancing usability and task efficiency. They also excel in real-time data processing and visualisation, transforming complex sensor inputs into comprehensible 3D models and overlaying vital information such as risk assessments within the virtual environment. The scalability and customisation of

these systems allow for a wide range of applications, from individual robotic control to large-scale automated operations. Moreover, VR environments facilitate collaborative teleoperation, enabling multiple operators to work in unison within a shared virtual space, regardless of their physical locations.

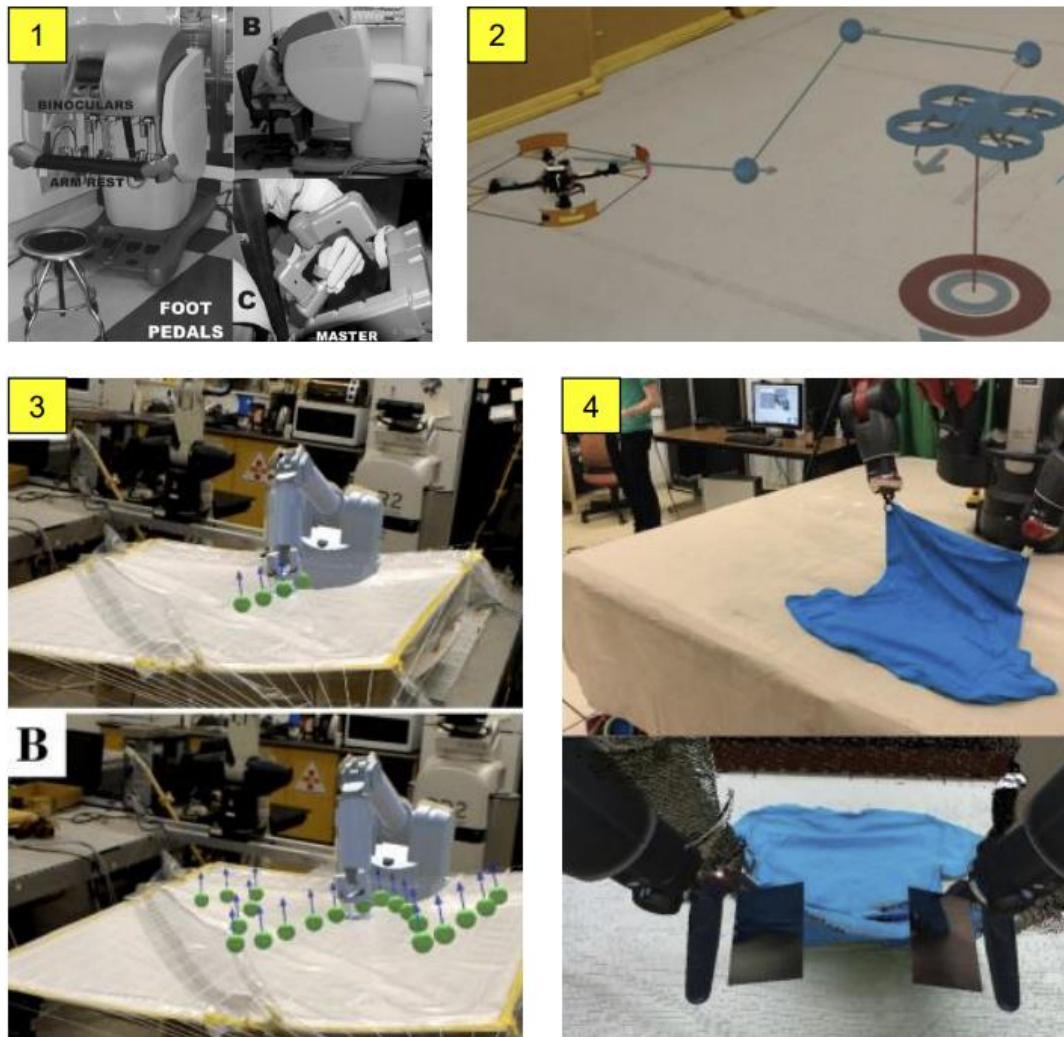


Figure 2-6 XR-based robot teleoperation interfaces; (1) stereoscopic telepresence (Ballantyne and Moll, 2003), **(2-3) AR** (Quintero et al., 2018; Walker, Hedayati and Szafir, 2019), **(3) VR** (Omarali et al., 2020)

2.2.2 Visualisation techniques

Visualisation is critical for the operator's situation awareness and ability to orientate the robotic system within the remote environment. Conventional visualisation methods in telerobotics involve placing one or more cameras or end-effectors in the robot environment to provide the operator with a video stream of the remote location. It causes the limitation of the FOV and increases the cognitive workload of the operator. However, the operator's situational and spatial awareness is limited. Therefore, new methods must be implemented or combined with regular video feed feedback to exploit the full potential of immersive VR. This requires using sensor information in the robot's environment and providing the operator with relevant data seamlessly integrated into the immersive 3D scene. Choosing a fast and intuitive user interface is essential to communicate all the necessary live information for performing the work without overloading the user's cognitive abilities. The user interface is the central element of the telepresence robot. The two major types of 3D interfaces for displaying information are egocentric and exocentric views Figure 2-7.



Figure 2-7 Egocentric view (left), exocentric view (right)

The egocentric view, often used with humanoid robots, offers a first-person perspective of the environment to mimic the same fields of view. However, this method has occlusion problems caused by the environment or the robot. It is also uncomfortable to use if the robot is in motion, which can lead to cybersickness for the operator. The exocentric view provides a broader view of the remote working environment, including the digital model of the robot.

Livatino et al. proposed a mixed reality visual aids solution to provide the operator of a remote ground robot. Their solution offers the possibility to switch between egocentric and exocentric views. In an egocentric view environment, information

such as suggested travel trajectory and guide direction arrow is superposed on the robot camera stream. In the exocentric view, thanks to the reconstructed 3D environment where the robot operates, the user can get additional data (altitude, slope, distance, and ability cost) by using a virtual pointer targeting area on the 3D virtual environment (Livatino et al., 2021), (Saakes et al., 2013).

In teleexistence systems, understanding and navigating remote environments, which can be unstructured or structured, is crucial. Unstructured remote environments are unpredictable and need more prior knowledge, as exemplified by DARPA's robotics challenge simulating disaster clean-up scenarios. These environments are typically visualised in VR as point clouds (Gray et al., 2017). Structured remote environments, in contrast, are at least partially known beforehand and are common in goal-specific applications or simulations. They are usually represented in VR using 3D meshes for clarity and efficiency in communication and computation (Egashira et al., 2021).

Point clouds, though effective in capturing unstructured environments, demand substantial bandwidth and computational power. Techniques like replacing some elements with predefined 3D meshes can significantly reduce bandwidth usage and computational load (Whitney et al., 2018b; Zhou, Zhu, and Du, 2020). For instance, Kohn et al. demonstrated a reduction in communication needs from 760MB/s to 4MB/s by removing point clouds of certain elements (Kohn et al., 2018). Downsampling techniques, like voxelising and using Octo Maps, also aid in reducing the processing demands for large point clouds.

Capturing these environments often involves RGBD cameras, which combine RGB video with depth information. Despite their simplicity, they need help with noise on metallic surfaces and limited textural information (Rusu and Cousins, 2011). Multi-camera setups can provide more comprehensive reconstructions but increase bandwidth requirements and necessitate robust calibration methods (Wei, Huang, and Li, 2021); another technique was employed for a near 360-degree reconstruction (Okura et al., 2014).

Regarding material property detection in remote environments, which can be helpful for remote inspection contexts, machine learning, particularly

convolutional neural networks, has been pivotal in classifying and segmenting materials from images (Sean et al., 2015). In challenging environments with limited lighting or visibility, tactile exploration methods combining different sensing techniques have proven useful (de Boissieu et al., 2009).

Navigating VR reconstructions in teleoperation is another critical aspect. Various methods exist, but each has its limitations, like motion sickness with joystick control (Buhler, Misztal and Schild, 2018), disorientation with transportation methods (Hoeg et al., 2017), and spatial requirements for physical walking (Park and Jang, 2019). Hence, choosing suitable navigation and exploration methods is essential for effective teleoperation.

2.2.3 Control techniques

2.2.3.1 Direct control in robotics

Direct control (Figure 2-8) in robotics is characterised by the operator's continuous, hands-on specification of a robot's actions and movements. This approach is akin to guiding the robot's end-effector or adjusting its joint angles manually. Direct control can be further categorised into unilateral and bilateral control.

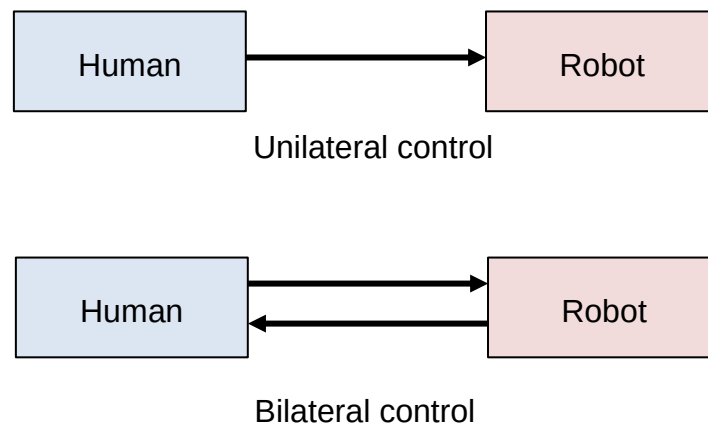


Figure 2-8 Direct control in robotics

Unilateral Control: In unilateral control, operators guide a robot's movements without experiencing any reciprocal effects. This form of control is particularly advantageous in settings where real-time responsiveness is crucial, but force feedback is not necessary or feasible. Human motion-capture-based control

significantly advances this category, especially in VR-based teleoperation (Jang et al., 2019). This method involves translating human movements directly to the robot. Technologies like inertial sensors and visual tracking systems are pivotal in achieving this, making control more intuitive and aligned with natural human movements than traditional devices like joysticks (Kofman et al., 2005). This intuitiveness is especially beneficial in complex tasks like surgical procedures or precision manufacturing.

Bilateral Control: Bilateral control is more intricate, particularly in VR environments, due to the operator's limited physical freedom. It typically employs model-mediated techniques, generating force feedback based on virtual models rather than direct physical interaction (Al-Sada et al., 2020; Mihelj and Podobnik, 2012). This type of control is instrumental in tasks that require a high degree of precision and tactile feedback, such as teleoperated bomb disposal or underwater exploration. The challenge lies in accurately simulating real-world physics and ensuring timely feedback to the operator.

2.2.3.2 Shared control

Shared control (Figure 2-9) synergistically combines user input with the robot's onboard autonomous systems. This approach is often utilised for assistance in navigation, like path planning and obstacle avoidance. While shared control can enhance the accuracy and efficiency of operations, it may also increase the

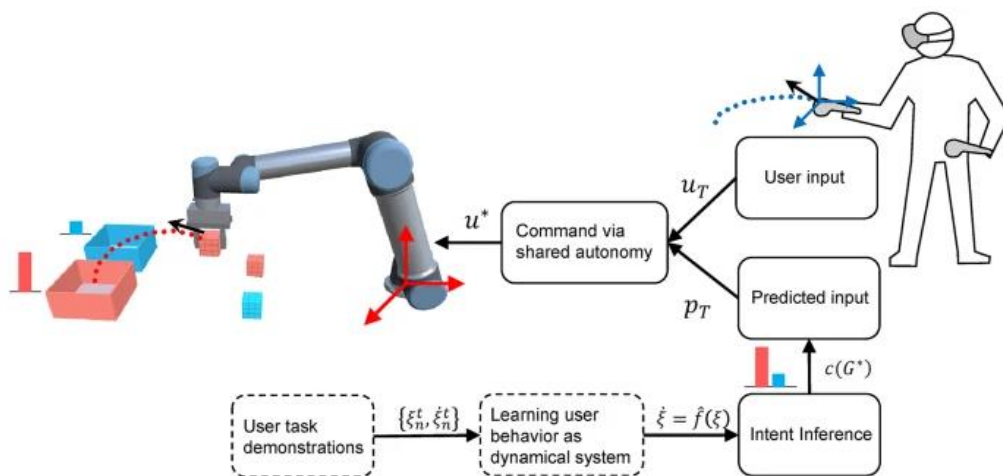


Figure 2-9 Shared control in robotics (Wang et al., 2022)

cognitive load on the operator (Wang, Reed, and Fey, 2018). The key is to find a balance where the robot's autonomy supplements human decision-making without overwhelming it (Nikolaidis et al., 2017; Xi et al., 2019). This control type is increasingly relevant in scenarios like autonomous driving (Johns et al., 2016) and complex industrial automation (Ramacciotti et al., 2016).

2.2.3.3 Supervised control

Supervised control (Figure 2-10) involves robots executing tasks primarily under human oversight. Operators set objectives and oversee the execution, while the robots use their sensing and control systems to perform the tasks (Gaschler et al., 2014; Leeper et al., 2014; Makhal, Thomas and Gracia, 2017). This control type demands high trust in the robot's capabilities and a clear understanding of its limitations. It is prevalent in AR-based teleoperation and effective in VR settings (Gaschler et al., 2014). Supervised control is handy in scenarios where complete autonomy is not feasible or desirable, such as remote exploration or hazardous environment interventions.

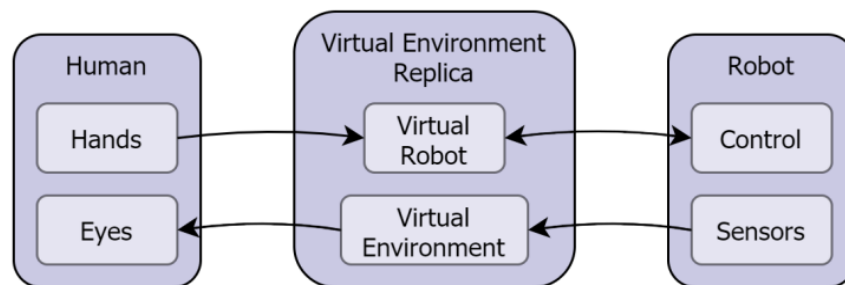


Figure 2-10 Supervised control in robotics (Wonsick and Padir, 2020)

The use of VR in robot teleoperation is rapidly evolving. There is a growing trend towards leveraging supervised control paradigms and direct control to train autonomous machine-learning algorithms. This dual approach harnesses the precision of human control and the efficiency of automated systems. Another solution is to use Supervisory Control and Data Acquisition (SCADA) systems in telepresence frameworks, specifically for remote inspection, which offers potential enhancements in operational control, data acquisition, and real-time feedback mechanisms. SCADA is a control system architecture comprising computers, networked data communications, and graphical user interfaces for high-level

process supervisory management. Through SCADA, operators could remotely monitor environmental conditions, machinery status, and system performances in real time, enabling immediate responses to detected issues or anomalies. This capability is particularly beneficial for telepresence applications that require important levels of precision and reliability in data handling and operational control over distant or inaccessible locations.

2.2.4 AI in monitoring Human decisions

Integrating AI into organisational decision-making signifies a pivotal shift towards a harmonious coexistence between human intelligence and technological advancement. Central to this paradigm is the notion that AI does not supplant human decision-making but enhances it by marrying AI's computational and analytical prowess with the nuanced, intuitive judgment of human expertise. This concept finds resonance in the domain of autonomous ships and their remote maintenance, where the collaborative interplay between AI and human capabilities is beneficial and essential. For remote maintenance, AI's ability to process and analyse extensive datasets enables predicting potential failures, optimising maintenance schedules, and diagnosing issues that traditional methods may not easily uncover. Such capabilities are invaluable in maintaining autonomous vessels' operational integrity and efficiency. Yet, it is crucial to recognise that the role of human technicians remains paramount. Their interpretative skills are indispensable, particularly when evaluating AI-generated recommendations against the vast backdrop of a ship's operational history, current environmental conditions, and the intricate knowledge of machinery that comes with firsthand experience. Thus, the employment of AI in monitoring human decisions in remote maintenance should be seen as an opportunity to augment human skill sets rather than a strategy for their obsolescence (Jarrahi, 2018).

As opposed to full automation, AI augmentation offers a pathway to building Trustworthy AI by enhancing human capabilities without replacing human involvement, particularly in high-risk areas like robot teleoperation. This approach leverages the strengths of both humans and AI to improve safety, efficiency, and

decision-making quality, maintaining transparency, accountability, and robustness essential for fostering trust in AI systems. A developed teleoperation testbed illustrates AI augmentation's application, using a multifaceted sensor system to inform and refine robot operations through AI algorithms. By keeping humans in the loop and integrating AI as a supportive mechanism, the testbed aligns with Trustworthy AI principles, enhancing operational precision and mitigating risks in remote control environments. The project underscores the importance of human-centred design in AI development, aiming to balance technological capabilities with human intuition and judgment (Marino et al., 2020).

2.3 Literature review conclusions

Traditional inspection techniques are labour-intensive and necessitate in-person on-site (Pour Rahimian et al., 2020). For example, conventional marine vessel maintenance inspection techniques are labour-intensive and require in-person on-site. Routine inspections are essential to maintain marine vessels' safe and efficient operation. They inspect various vessel components, including their structural integrity, mechanical and electrical systems, navigational equipment, and safety systems (Serebrenniy, Lapin and Mokaeva, 2019). Teleoperation strategies could address these limitations. The need for more awareness of traditional teleoperation systems is an issue that impacts their effectiveness and safety. They often provide limited sensory feedback to the operator. This can include a lack of haptic feedback, limited visual feedback due to low resolution (resulting in lousy orientation), or narrow field-of-view cameras. As a result, operators may not completely understand the environment in which the remote robot is operating (Roldán et al., 2017). Operators may have trouble understanding the spatial orientation of the robot, especially in complex or unfamiliar environments. This can be exacerbated by camera angles that do not match the operator's natural viewpoint. Traditional systems might not provide sufficient information about the environmental conditions around the robot, such as temperature, humidity, or unseen obstacles. This lack of context can hinder decision-making and operational efficiency. Operating a robot remotely can be

mentally taxing, leading to fatigue. The need to constantly interpret limited or poor-quality information can increase cognitive load, potentially leading to mistakes. Without a comprehensive awareness of the robot's environment and capabilities, there is an increased risk of accidents, which can be harmful in scenarios like surgery, hazardous material handling, or exploration in dangerous environments. While there were studies assessing the human performance of VR-based teleoperation systems, they were often based on software evaluations with no consideration for critical human factors, such as the cognitive demands imposed on the user by the interface design.

3 Research methodology

Figure 3-1 summarises the research methodology proposed to address the objectives and answer these three research questions.

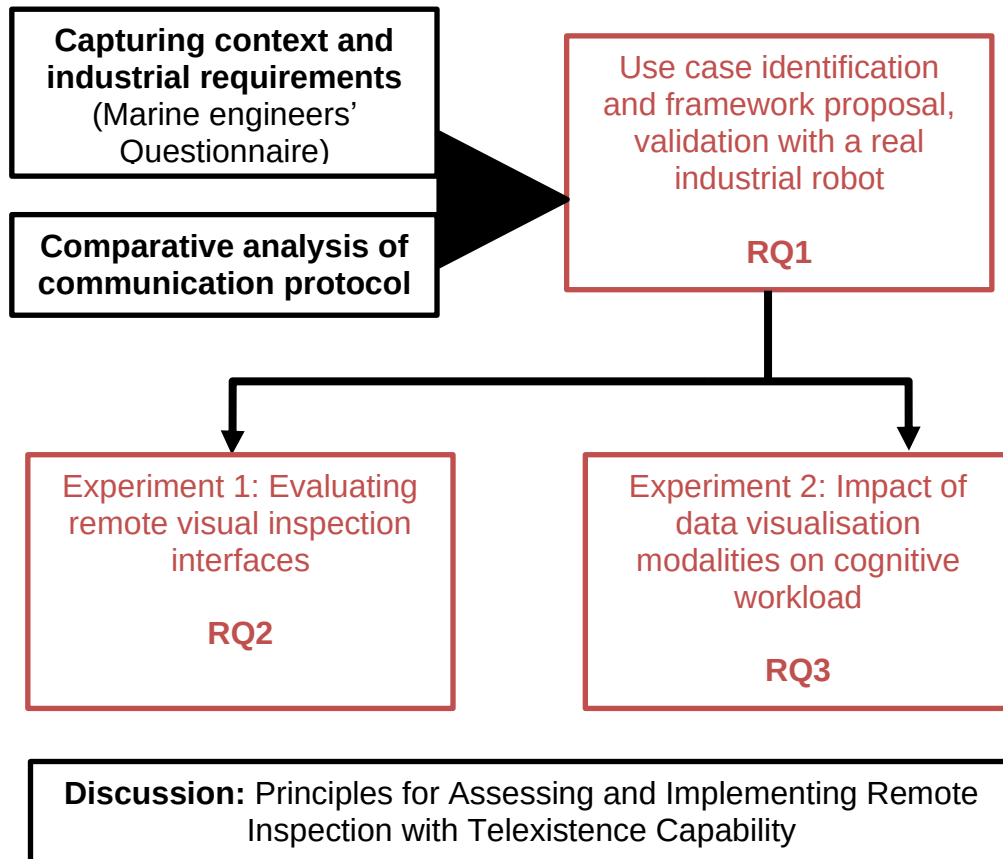


Figure 3-1 Research methodology approach to the dissertation

3.1 Capturing context and industrial requirements

A questionnaire was conducted with two marine engineering experts as part of a comprehensive study focusing on the upkeep of machinery spaces in marine vessels. The aim was to gain insights into their daily responsibilities, challenges, and prospects in marine engineering and maintenance. This questionnaire also sought their views on the design of user interfaces for maintenance on marine vessels to use the collected information to inform the development of more effective and user-friendly telexistence systems for these applications.

The questionnaire consisted of the following questions focused on the job and activities of marine engineers and aimed to reflect on future working interfaces:

- Number of years of experience in the field of marine engineering/maintenance?
- Past/Current responsibilities?
- What does a typical day on the job look like for you as a marine engineer in charge of maintenance?
- What everyday maintenance tasks do you perform in the machine room?
- How do you diagnose and troubleshoot problems with the engines and other machinery in the engine marine room?
- What tools and equipment do you use in your work? How do you ensure that these tools are properly maintained and calibrated?
- What safety precautions do you take when working in the engine marine room?
- What training and experience are required for a marine engineer to be proficient in maintaining the engine marine room?
- What challenges do you face in your job, and how do you overcome them?
- Are there any new technologies or practices you see adopting to maintain marine engines and machinery in the future?
- What are the most important factors to consider when designing a user interface for maintenance on a marine vessel?
- How do maintenance needs and requirements differ across different types of marine vessels?
- How do you envision the user interface being used in practice? What kind of information and functionality should it provide?
- Are there any existing user interfaces for maintenance on marine vessels that you think are exceptionally well-designed? What elements of these interfaces could be incorporated into the design of a new user interface?

The first respondent, a Marine Engineer at Dstl, part of the United Kingdom Ministry of Defence, has 31 years of experience in the field. His responsibilities include working on ship and small craft concept designs and technology studies

into related equipment. He also studies the implementation of modern and advanced maintenance and repair techniques and technologies for Royal Navy surface ships, along with implementation in other environments (responses are available in Appendix B). The second responder is also a marine engineer with 37 years of experience in the field. His professional background is rooted in the Royal Navy, where he held various positions, including Deputy Marine Engineer Officer of a Type 42 Destroyer, Propulsion Section Head and Senior Watchkeeper of aircraft carriers, and Marine Engineer Officer (Chief Engineer) of a Type 22 Frigate and Type 45 Destroyer (responses are available in Appendix C). Both respondents have experience in marine engineering, indicating a depth of knowledge and expertise that is difficult to find. Their roles and responsibilities cover a wide range of experiences within marine engineering, from operational duties on diverse types of ships to conceptual design and technology studies. This breadth of experience offers rich, qualitative insights for the study's focus on telexistence systems. A limitation of this approach is the question of how representative these two opinions are for the entire sector. The marine engineering is diverse, and practices can vary significantly across different organisations, vessel types, and geographical locations. The insights gained, while valuable, may only partially capture this diversity.

The following summarises the key points that informed our framework development and maintenance scenario creation:

- **Day-to-day job:** Marine engineers' daily responsibilities involve reporting overdue maintenance items to the marine engineering department and coordinating with the command for machinery shutdown. Using logged machinery performance data and observations from watchkeepers, they detect imminent failures or systems that require further investigation.
- **Common maintenance tasks:** Common maintenance tasks in the machinery space include daily inspections, data logging if not recorded automatically, cleaning, tightening connections, replacing gaskets, and routine oiling and greasing.

- **Diagnose onboard the vessel:** Troubleshooting onboard is performed using watchkeeping senses such as sound, sight, smell, and touch. Engineers also use condition monitoring techniques like vibration and oil analysis to diagnose problems.
- **Safety:** Safety in the machinery space is maintained by tracking personnel with identification discs, implementing the buddy system, which is never alone in the operating area, and wearing personal protective equipment, including fire-resistant clothing, ear protection, and steel-toed safety boots. Hazards such as hot surfaces and slippery decks should also be considered before conducting the work. Extra precautions if the ship is manoeuvring or in a high sea state.
- **Training:** Marine engineers receive training through shore-based instruction and onboard mentorship. Task books record the completion of necessary activities.
- **Main challenges:** Challenges in marine maintenance include understanding the design intent, system-based fault finding, and determining the material state.
- **Maintenance differences across different vessels:** Maintenance practices vary between naval and non-naval vessels, with the priority being the safety of personnel on civilian vessels and mission continuation on naval vessels.
- **Design of the maintenance user interface:** First, the presentation of the priority of alarms. Current systems present alarms and warnings as the system interrogates them. This results in a mass of alarms you cannot see, and the alarm order does not reflect how the fault occurred. Secondly, to give the user the correct information, when needed, in a format they can use. Users might take a while to gain confidence in the system, so it must be as helpful and correct as possible before being rolled out to sea. Also, as the amount of data produced by naval vessel systems goes up dramatically, the data must be processed and presented to the user without drowning them.

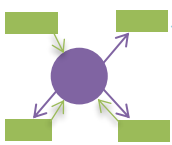
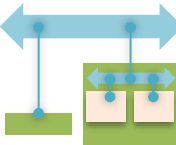
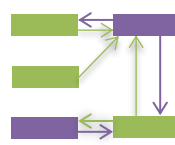
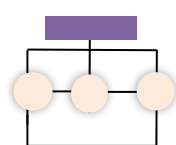
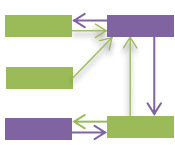
- **Envision future maintenance practices:** (1) Provide technical information at the point of work. (2) Condition monitoring that interventions to prevent failure as opposed to telling you why something has failed. (3) Subsystem digital twin at a sufficient level to allow both warnings of imminent failure or issues and conduct 'what if' games to see the effect on system performance and failure of proposed operations, operating regimes, and profiles.
- **Envision of the future user interface:** (1) Provide automation to remove the need to do mundane tasks; there should be no manual recording of running parameters at the machine, for example. (2) Overlay running parameters on a view of a machine; enhance with analysis, e.g., where is this pump on its pressure/flow curve and what is its performance trend. (3) Present known defects and allow new ones to be added, e.g., leaks. (4) Material state information – more for maintainer/supervisor, e.g., latest condition monitoring results, when last run for standby machinery. (5) The provision of operational guidance to a platform while underway, in peacetime and wartime conditions, via awareness of expected versus actual Hull, Mechanical and Electrical (HM&E) system performance, likelihood of what system failures with different operating profiles.

3.2 Comparative analysis of communication protocol

Considering the development of a telexistence framework for remote inspection, a meticulous examination of various communication protocols was conducted, as presented in Table 3-1. The main criteria for selection included the capacity for real-time applications, scalability, the ability to discover data automatically, and the persistence of data. In addition, the protocol's suitability for a client-server model was also considered, as this architecture is highly compatible with a telexistence setup, which demands centralised control with remote execution.

Table 3-1 Comparison of some of the most popular communication protocols (created by the author)

<i>Parameters</i>	<i>MQTT</i>	<i>DDS</i>	<i>OPC UA</i>	<i>ROS 1</i>	<i>CoAP</i>
-------------------	-------------	------------	---------------	--------------	-------------

	(Message Queuing Telemetry Transport)	(Data Distribution Service)	(OPC Unified Architecture)	(Robot Operating System)	(Constrained Application Protocol)
Target application types	Device to server, data Centre, back office, IT cloud, IoT	Within and between devices, dedicated systems, real-time applications, scalable systems	PLCs, industrial network settings, data acquisition, IoT	Robotics applications	Low-power devices, wireless and mobile networks are the preferred choice
Architecture Clients Servers Databases Nodes					
Model	Publish-Subscribe	Publish-Subscribe	Publish-Subscribe	Publish-Subscribe	Publish-Subscribe
Transport protocol (base specification)	TCP	TCP, UDP	TCP, UDP	TCP, UDP	UDP
Quality of Service (QoS) availability	Best Effort/ Guaranteed Message Arrival	Parameters on durability/lifespan/destination order	Best Effort	No	Yes, confirmable/non-confirmable
Security, Encryption	Not by default, TLS/SSL options	Security Plug-ins	Security policies	Not by default, TLS/SSL options	Not by default, can be combined with DTLS (Datagram Transport Layer Security)
Data discovery	Manual	Automatic	No	Automatic	Manual and registration
Data persistence	No	Yes	No	No	No

Standardisation status	Open standard maintained by the Organization for the Advancement of Structured Information Standards (OASIS) consortium	Open standard maintained by the Object Management Group (OMG)	Open standard supported by the OPC foundation	Open-source software framework maintained by the Open Robotics organisation	Open standard maintained by the Internet Engineering Task Force (IETF)
-------------------------------	---	---	---	---	--

Table 3-1 explains the architecture parameter with clients, servers, databus, and nodes. Here is a brief explanation of each common knowledge of network protocols:

- **Client:** In a client-server model, a client refers to a hardware component or software application that uses a service provided by a server. Although not a fixed rule, it is typical that the server is located on a separate computer system. In such scenarios, the client taps into the server's service via a network. This terminology applies to software programs and hardware devices within the framework of a client-server architecture.
- **Server:** A server is a computer hardware component or software application that delivers services to other software applications or devices, termed "clients." This is encapsulated within the client-server architectural model. Servers are multifunctional, with capabilities ranging from data or resource distribution among several clients to conducting computations on behalf of a client. Interestingly, one server can cater to numerous clients; inversely, a single client can employ services from multiple servers.
- **Databus:** A databus represents a system integrated within a computer or device, comprising a connector or a collection of wires, which acts as a conduit for data transmission. The evolution of personal computers and other hardware components has led to the emergence of several types of data buses. A databus serves as a communication

mechanism, enabling data transfer among components within a computer or between multiple computers.

- **Node:** In networking, a node is any device that can send, receive, or forward information. This includes computers, servers, smartphones, and other devices. In a distributed system like a blockchain, a node might represent a point where a copy of the distributed ledger is stored.

DDS excels in real-time data handling, a critical attribute for telexistence applications where immediate response and seamless data exchange are paramount. This capability surpasses MQTT and CoAP, which are more oriented towards IoT applications with less stringent time constraints. Additionally, DDS's scalability is particularly advantageous for the evolving nature of telexistence systems (Xiang et al., 2004). It adeptly expands networks of devices and inspection tasks, a feature that could be more efficiently managed by OPC UA or ROS 1.

A pivotal aspect of DDS is its automatic data discovery, which simplifies the integration and expansion of telexistence systems, a complex process when using protocols like MQTT or OPC UA (Bienstock, Mentzer and Bird, 1997). Furthermore, the inherent data persistence in DDS is invaluable for maintaining continuity in remote inspections, especially under network fluctuations—a challenge that protocols like ROS 1 and CoAP are less equipped to manage.

While DDS's security capabilities require additional configurations, like ROS 1 and CoAP, its compatibility with robust security plug-ins makes it a strong contender in environments where data security is paramount. Moreover, the support for both TCP and UDP transport protocols in DDS offers unmatched flexibility, accommodating the diverse needs of telexistence scenarios, whether prioritising reliability or speed (Kang, Kapitanova and Son, 2012).

The standardised nature of DDS, maintained by the Object Management Group (OMG), ensures reliability and consistent community support, contrasting with the more open-ended frameworks of protocols like ROS 1. This standardisation, coupled with DDS's proven history in similar high-performance applications, affirms its suitability for the intricate demands of telexistence frameworks,

particularly in ensuring efficient, reliable, and scalable remote inspection operations.

3.3 Use case of remote visual inspection

The insights derived from the marine engineering experts' questionnaire responses provide a compelling rationale for the focus of this PhD research on remote visual inspection. The experts' daily tasks underscore the critical need for real-time information in effective maintenance. This need aligns seamlessly with the capabilities of remote visual inspection, which provides real-time visual data, thereby enhancing the ability of engineers to monitor and maintain machinery.

Condition monitoring emerged as a theme in the experts' responses. They highlighted its importance in maintaining machinery health. Remote visual inspection serves as condition monitoring, allowing operators to visually assess the state of machinery and identify any signs of wear or damage. Thus, the technology is well-positioned to address the needs expressed by the experts.

Safety measures were another critical theme in the experts' responses. Remote visual inspection can enhance safety by enabling operators to inspect machinery from a safe distance, reducing the risk of accidents. This safety enhancement aligns with the experts' emphasis on the importance of safety measures in their work.

The experts also provided valuable insights into the design of user interfaces for maintenance on marine vessels. These insights can inform the design of user interfaces for remote visual inspection systems, ensuring that they present visual data to operators clearly and understandably.

Finally, the experts expressed a forward-looking view of the field, seeing the potential for adopting new technologies and practices. Remote visual inspection fits within this vision, suggesting that it could play a vital role in the future of marine engineering and maintenance. This alignment with future trends further justifies the choice of remote visual inspection as a use case for this PhD research. The preliminary evaluation criteria will be investigated, including natural interaction, interface choice and its impact, and data visualisation's impact.

The chapter 4 introduces and explores the proposed framework aimed at enhancing telepresence capabilities through the design and implementation of virtual representations. These virtual entities are equipped with real-time information states and interaction capabilities, enabling maintenance personnel to engage in telepresence effectively. The framework's design process underwent several iterations and included a use case with a physical industrial robot to highlight its practical application in unmanned vessel maintenance. A significant part of the framework involves sharing sensor data via a Data Distribution Service (DDS) and leveraging an immersive VR interface to control and monitor an industrial robot. The Unity 3D engine is utilised for real-time data visualisation through a digital twin of the robot. This approach was validated using a FANUC M20-IA industrial robot and Azure Kinect RGBD cameras, showcasing reliable performance in remote operation scenarios. Furthermore, the chapter discusses the application of telepresence technology in maritime settings, particularly for Unmanned Vehicle Systems (UVS) maintenance, and delves into a performance analysis focusing on managing time delay and latency in data transmission.

3.4 Experiment 1 method

The experiment one addressed in Chapter 5 of the research focuses on the comparative analysis of maintenance interface modalities, specifically between non-immersive (2D) and immersive (VR) interfaces, in the context of unmanned surface vessels (USVs).

The research methodology involves creating gamified tasks to enhance engagement and motivation among participants, who must perform inspection tasks on various engine parts through both VR and 2D interfaces. The study aims to evaluate the effectiveness of these interfaces in remote visual inspection scenarios, measuring performance metrics and human factors such as presence, workload, and usability. The experimental design includes real-time feedback mechanisms and kinematic control of collaborative robots to simulate a realistic maintenance scenario within a hybrid simulated setting.

Participants in the study undertake a mission-like task involving the inspection of engine parts, with the process designed to emulate real-world maintenance

operations. This includes navigating a simulated environment, identifying specific parts for inspection, and documenting observed conditions. The task is designed to compare the effectiveness and user experience between VR and 2D interface modalities, with an element of gamification to engage participants actively.

3.4.1 Experiment setup

The experiments were conducted in March 2023 at Grenoble INP. In the experimental setup shown in Figure 3-2, a Universal Robot UR16e (number 1) collaborative robot was utilised, an industrial robot arm featuring six degrees of freedom, a payload capacity of 16 kg, and a reach of 900 mm. For both the 2D and VR interfaces, an MSI GT63 Titan 15.6-inch laptop (number 4) was employed as the primary computing device. This laptop has an Intel Core i7-8750H processor, an NVIDIA GeForce GTX 1080 graphics card, and 16GB DDR4 RAM. The HTC VIVE Cosmos headset (number 2) was used for the VR interface within a free-tracked area of approximately 4m x 4m. The headset featured a combined resolution of 2880 x 1700 pixels, a 90Hz refresh rate, and a 110-degree field of view. The VIVE Cosmos has two 6DoF (degree of freedom) controllers (number 3).



Figure 3-2 Participants during the experiment with the VR interface (left), and the 2D interface (right)

3.4.2 Participants

The G*Power software (Faul et al., 2007), a statistical power analysis tool, was used to determine the necessary sample size of participants for the experiment.

G*Power was configured using ANOVA, precisely the "Repeated measures, within-between interaction" model, to facilitate the comparison of the two interfaces. The input parameters were set with an effect size ($f = 0.4$), deemed acceptable in anticipation of a difference between the 2D and VR groups and conditions in the study; alpha error probability (0.05); Power $1-\beta$ error probability (0.95); the number of groups set to 2 (Order 1: 2D first, Order 2: VR first); the number of measurements set to 2, since there were two conditions within each group (2D and VR); correlation among repeated measures (0.5); and nonsphericity correction ϵ (1). The result gave us a sample size of 24 or 12 participants per group. The choice of these parameters and the resulting sample size of twenty-four participants were based on standard practices in statistical analysis to ensure the experiment was adequately powered to detect the expected differences between the 2D and VR interfaces while controlling for error probabilities. Similar experimental user studies for assessing a VR-based framework with a 2D interface were conducted with twenty-four or fewer participants (de Boer, de Winter and Eisma, 2023; Stotko et al., 2019).

Twenty-four people were recruited. All participants were volunteers from Grenoble INP. They were not maintenance experts but were familiar with VR and robotics at varying usage frequencies. The group included 14 males and 10 females, students (both master's and PhD level), and staff engineers, with a mean age of **24.08 $\pm$ 2.19**.

Participants' familiarity with VR ranged from novices who had never used a VR system to seasoned users who had experienced VR in leisure (e.g., gaming) and professional contexts (e.g., simulations or design work). About **10%** of participants claimed to use VR regularly, **50%** had tried it a few times, and **40%** had no prior VR experience.

Given the institution's technical focus, several participants (around **60%**) were exposed to robotics through coursework, research, or personal interest. However, only about **15%** had hands-on experience operating or programming robots. The remaining participants knew robotics as a concept but had limited direct interaction with them.

As the study employed gamified tasks, understanding participants' backgrounds in gaming was essential. Approximately **40%** identified as regular gamers, playing video games at least once a week. Another **35%** played occasionally, while the remaining **25%** rarely or never played video games.

Given their affiliation with Grenoble University's industrial engineering department, all participants had a baseline level of technical proficiency. They were comfortable using computers, software applications, and digital tools.

3.4.3 Data collection

In the proposed experiment, the independent variables were the HMI technology used, either the laptop and mouse or the VR HMD with controllers. The objective dependent variables were performance metrics, including the time to complete the tasks in seconds and the collision rate. They reflect the easiness of use and the perception of local immersion. The subjective, dependent variables were collected using three human factors: presence, workload, and usability.

The iGroup Presence Questionnaire (IPQ) is a commonly used tool in research to measure a user's sense of presence when interacting with a virtual environment (iGroup, 2006) based on the model of the presence of (Schubert, Friedmann and Regenbrecht, 2001). The IPQ consists of fourteen items categorised into four subscales: spatial presence, involvement, experienced realism, and general presence. Each item in the questionnaire is rated on a 7-point Likert scale, ranging from “-3” (entirely disagree) to “+3” (fully agree). The sub-elements evaluating the virtual environment are comprised of spatial presence (sense of being physically present), involvement (measuring the attention), experienced realism (measuring the subjective experience of realism), and general presence (the subjective sensation of “being there”). The responses are then summed to produce a total score for each subscale and can also be combined to produce an overall measure of presence. Schaik et al. first presented its use for mixed reality environments (Schaik et al., 2004).

The NASA-TLX (N. H. P. R. Group et al., 1987) is a multi-dimensional approach to assess the perceived workload in an individual engaged in a task. NASA's

Human Performance Research Group developed it and has been widely used across various fields and disciplines. The index offers a subjective yet robust method for evaluating the workload experienced by users in human-machine interaction scenarios. The NASA-TLX comprises six dimensions, each reflecting a different aspect of the perceived workload: mental demand, physical demand, temporal demand, performance, effort, and frustration level. In application, the NASA-TLX uses a two-step process. Individuals rate their experience on each of the six dimensions on a scale from 0 (very low) to 100 (very high), except for the performance dimensions from 0 (perfect) to 100 (failure).

The System Usability Scale (SUS) is a simple tool used to measure the usability of a system, service, or product (Brooke, 1996). It was developed by John Brooke in 1986 to evaluate various interface technologies, including hardware, software, mobile devices, websites, and applications. The SUS consists of a 10-item questionnaire, where each item is a statement to which the user rates their level of agreement on a 5-point Likert scale, with one being “strongly disagree” and 5 being “strongly agree.” Although it is not a detailed usability assessment tool like others, the SUS is globally recognised and widely used because of its simplicity and ability to provide reliable results with just a small sample of users. It allows researchers and designers to quickly assess the usability of a design, compare various systems or design versions, and monitor changes in usability over time.

Finally, participants' remarks and suggestions about the overall experiment and interfaces were collected through a final open-ended question.

3.4.4 Procedure

The experiment was structured to ensure participants' understanding, comfort, and optimal engagement (Figure 3-3).

- **Informed consent:** Upon arrival, the priority was to ensure participants understood the nature of the study and their rights. They were handed the consent form detailing the experiment's objectives, methods, and potential risks. This form adhered to the Cranfield University Research Ethics

System guidelines (8.3.3Appendix A). Participants were encouraged to read it thoroughly and ask any questions before signing.

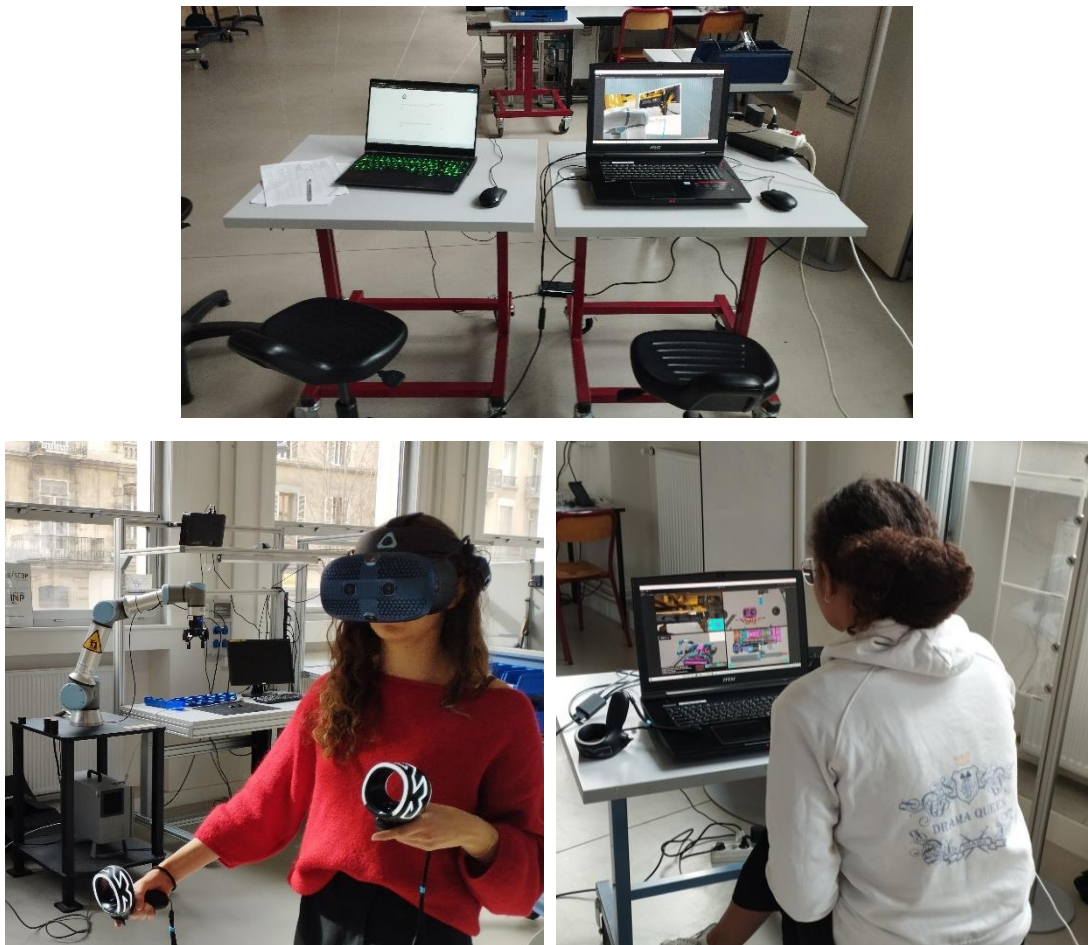


Figure 3-3 Participants and experiment 1

- **Introduction and background collection:** Once consent was secured, participants were given a brief overview of the study's purpose and the technologies with which they would interact. This context-setting helped ensure that participants approached the tasks with a clear understanding. Subsequently, they filled out an online form capturing demographic information and gauging their prior experience with virtual reality and robotics.
- **Interface familiarisation:** Recognising that participants had varied levels of exposure to VR and robotics, a practice session was introduced. This allowed them to navigate, explore, and get a feel for the VR and 2D interfaces. During this phase, which lasted around 3 minutes for each

interface, participants could seek clarifications or ask questions, ensuring they felt comfortable before the tasks commenced.

- **Conducting the tasks:** Participants were then directed to perform specific tasks using the interfaces. These tasks, aligned with the study's objectives, were designed to simulate real-world scenarios and challenges. Each task session lasted approximately 6 minutes per interface.
- **Post-task questionnaires:** Participants were given a series of questionnaires after completing tasks on each interface. These questionnaires aimed to capture their sense of presence, perceived workload, usability of the interface, and other specific feedback. Each set of questionnaires required about 7 minutes to complete.
- **Debriefing:** Participants were debriefed after completing tasks on both interfaces and filling out all questionnaires. This session, which lasted around three minutes, allowed participants to share their preferences between the two interfaces, provide additional feedback, and ask questions about the study. Insights from this session proved invaluable for understanding user preferences and improving future iterations of the interfaces.

The sequence in which participants interacted with the interfaces was varied to control for any potential order effects. Half of the participants (12 individuals) began with the VR interface, while the other half started with the 2D interface. This counterbalancing ensured that any systematic biases or learning effects were minimised. Table 3-2 provides a step-by-step breakdown of the experiment's flow. Notably, the total expected duration for each participant was approximately **39 minutes**, making it a short yet insightful engagement as participants were not paid.

Table 3-2 Flow of the experiment (blue are conducted two times)

 Video Recording	Activities	Duration
	Welcome & informed consent signature	2 min
	Pre-test questionnaire	2 min
	Modality Practice	3 min (x2)
	Tasks conducted on the modality	6 min (x2)
	Presence questionnaire	7 min (x2)
	Workload questionnaire	
	Usability questionnaire	
	Our questionnaire	
	Debriefing preference questions and comments	3 min
The total planned duration of the experiment 1: 39 min		

3.4.5 Data analysis

The data analysis gathered during the experiment was meticulously structured to comprehensively understand participants' experiences with the VR and 2D interfaces. The approach encompassed both quantitative and qualitative dimensions. For the quantitative data, descriptive statistics were derived for each questionnaire: IPQ, NASA-TLX, and SUS. This provided insights into the overall scores and individual items, emphasising the central tendencies and dispersions. Similarly, performance metrics, such as the time taken to complete tasks and the collision rate, were analysed through their mean and standard deviation. Paired-sample t-tests were deployed to identify the nuanced differences in user experiences across the two interfaces. This enabled us to compare the scores of presence, workload, usability, and performance metrics. Further depth was added with repeated measures ANOVA, which illuminated potential interactions between the type of interface (VR vs. 2D) and the order in which participants encountered them.

3.5 Experiment 2 method

Experiment two, addressed in Chapter 6, advances the study of immersive VR interfaces for remote visual inspection by exploring the effectiveness of various data visualisation modalities to represent the spatial information of a remote operating space. Focusing on four distinct visualisation methods—2D live camera feed, 3D live point cloud, a combination of both, and augmented virtuality with OpenCV markers—the research evaluates how these modalities influence cognitive workload and task performance in a VR environment. Through a user study employing pupillometry to measure cognitive workload, the chapter aims to assess the cognitive demands associated with each modality and the impact of the order in which they are evaluated on task effectiveness and cognitive strain.

3.5.1 Experimental setup

The experiments were conducted in July 2023 at the Centre for Digital Engineering and Manufacturing of Cranfield University. In the experimental setup shown in Figure 3-4, a Fanuc M-20iA was used (5) with its control unit (1) and control interface (3). It is an industrial robot arm featuring six degrees of freedom, a payload capacity of 20 kg, and a reach of about 1800 mm. The robot operated within fences (9) with two Microsoft Azure Kinects (6) for video and point-cloud streaming. The desktop computer used (4) has an AMD Ryzen PRO 3975WX 32-

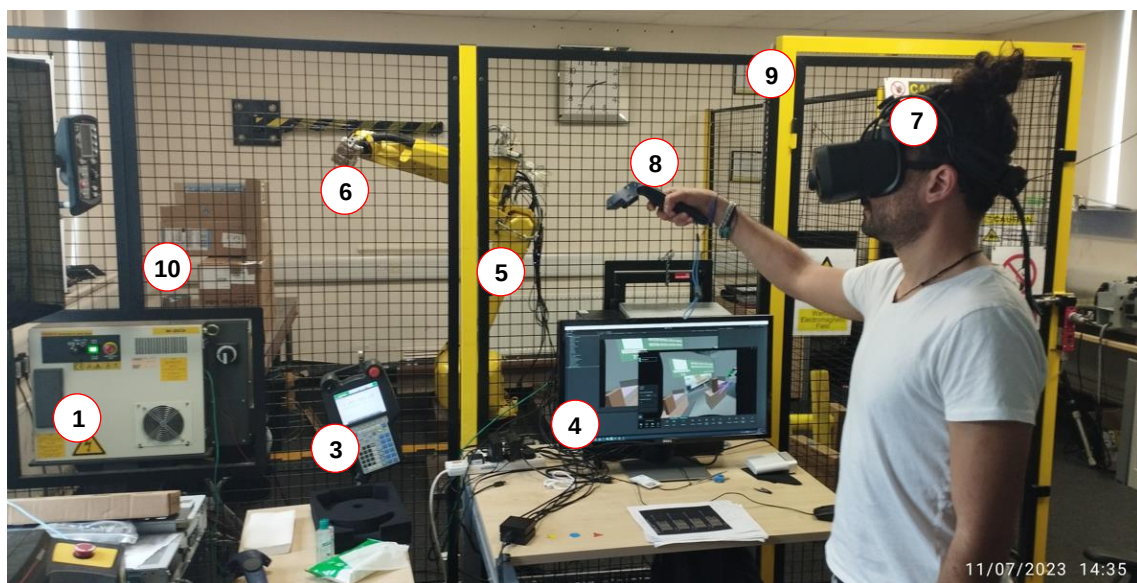


Figure 3-4 Participant during the experiment 2

Cores, an NVIDIA Quadro RTX 8000, and 64GB DDR4 RAM. The Varjo XR-3 HMD was used (7) for the VR interface within a free-tracked area of approximately 3m x 3m. The headset has a resolution of 2880 x 2720 pixels per eye, a 90Hz refresh rate, and a 115-degree field of view. The Varjo XR-3 was used with two HTC VIVE 6DoF (degree of freedom) controllers (8). The modalities scene setup was placed in front of the robot (10) and made of cardboard boxes and foam pipes to prevent any damage to the experiment equipment in case the participant collides with the setup.

3.5.2 Participants

The study was conducted with twenty-eight participants. However, due to inconsistencies in baseline pupillometry measurements, data from four participants were excluded, resulting in a final sample of **twenty-four participants for the analysis with 19 males and 5 females**. Data is presented in terms of mean $\pm$ standard deviation. The demographic composition revealed an average age of **28.17 $\pm$ 6.94** years. On the experiential front, seventeen participants had prior experience with VR, while seven did not. Similarly, fifteen participants were familiar with robotics, contrasting with **9** who lacked such experience. When gauging confidence in harnessing new technologies, **15** participants expressed moderate confidence, **8** were very confident, and **1** participant expressed a lack of confidence.

The exclusion of participants based on their baseline pupil dilation measurements was a necessary step for statistical integrity and not a reflection of their inherent ability to engage with or benefit from telepresence technologies. The four participants were excluded because their pupil dilation measurements were higher than the modalities measured. This is important because pupillometry, which gauges cognitive workload based on pupil dilation, would be skewed by such outliers. Baseline measurements offer a reference point against which changes in pupil dilation are compared. If the baseline is anomalously high, it could lead to misleading interpretations about cognitive load. Thus, it is judicious to exclude such participants to maintain the integrity and reliability of the study's findings. The order of modalities selected for participants was systematically

arranged so that, by the study's conclusion, each of the four orders was experienced by six distinct participants. Higher baseline pupil dilation in pupillometry studies can be attributed to many factors. The ambient lighting in which the measurement is taken plays a significant role; brighter environments typically induce pupil constriction, while dimmer settings lead to dilation. Participants' emotional states, such as anxiety or excitement about the experiment, can also cause variations in baseline measurements. Prior cognitive engagement, fatigue, or the intake of certain medications or substances like caffeine can influence pupil size. Furthermore, participants' anticipation or specific expectations about the study might have impacted their initial measurements. It is crucial in pupillometry studies to consider these variables to ensure the accuracy and consistency of the baseline measurements.

3.5.3 Data collection

3.5.3.1 Cognitive workload

As for this experiment, Varjo's HMD API data expose raw eye-tracking data, including pupil size from both eyes. The mean pupil size time series was analysed, as pupil diameters of both eyes are highly correlated (Jackson and Sirois, 2009). Before analysing pupils, data is essential to detect and deal with noise appropriately (Kret and Sjak-Shie, 2019). Kret et al. described the process in four steps: preparing the raw data, filtering the raw data, processing the valid samples, and sectioning and analysing the data. During the measurement of pupil dilation, artefacts are often created by blinks or the onset of eyelid occlusion and are characterised by significant changes in pupil size.

The pipeline for preprocessing pupillometry data proposed by (Kret and Sjak-Shie, 2019) was applied as described in the following steps:

1. Preparing the raw data

Eye tracking calibration was realised at the beginning of the experiment with the Varjo XR-3 as the eye tracker to reliably track participants' pupils (Figure 3-5). The Varjo XR plugin for Unity 3D saved eye data directly from Unity 3D for each modality in a CSV file. The recording was stopped between each modality but not

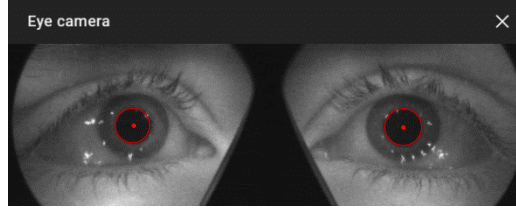


Figure 3-5 Pupils diameters measurement from the camera of the Varjo XR-3

each task of modalities. Data recorded consisted of the sample ID, the Timestamp of the recording, and the pupil size in millimetres for each eye. At this step within Unity 3D, data samples with either left or right eyes invalid were not saved in the CSV file as they do not require filtering.

2. Filtering the raw data

Using RStudio software, a function was created to perform comprehensive data processing on eye-tracking data from participants across the different modalities. The data for each participant and modalities were recorded in individual CSV files. The function is designed to read and process these files one at a time, ensuring that each data set is managed independently.

Firstly, the function checks if the recorded pupil diameters for both left and right eyes are within the valid range of 1.5 to 8. Any rows with values outside of this range are filtered out, helping to reduce the impact of any potential measurement errors or artefacts.

Next, the function calculates the dilation speed, which measures how quickly the pupil size changes. This is computed by dividing the absolute change in pupil diameter between consecutive samples by the corresponding change in time.

$$d'[i] = \max\left(\frac{|d[i] - d[i - 1]|}{t[i] - t[i - 1]}, \frac{|d[i + 1] - d[i]|}{t[i + 1] - t[i]}\right) \quad (3-1)$$

The function employs a method for outlier detection to minimise the impact of extreme values. This is based on the Median Absolute Deviation (MAD), a measure of data variability less sensitive to outliers than traditional methods like standard deviation. A constant scaling factor is applied to the MAD, and any sample with a dilation speed exceeding this scaled MAD is flagged as an outlier.

Following outlier detection, the function removes these flagged samples from the data. This helps ensure that extreme values do not skew the subsequent analysis.

3. Processing the valid sample

Now that the valid data for both eyes remain, the calculation of the mean pupil size time series is possible since both pupil diameters are correlated (Jackson and Sirois, 2009).

3.5.3.2 Scoring framework for participant's performance

The participant's performance was evaluated using a 5-point scoring framework (Figure 3-6) on five parameters: collision, shape/colour reporting, distance accuracy, photo quality, and task completion time. Each parameter was evaluated on a binary basis, where participants were awarded either one point for successful completion or 0 points for failure. The collision parameter is paramount, as ensuring safety in teleoperation tasks is critical. Thus, in the grading framework, any collision instantly results in a total score of 0, regardless of performance on the other criteria.

The binary nature of the grading system allowed for a clear distinction between successful and unsuccessful task completions, which simplified the grading process and made the results easier to analyse. The shape/colour recognition and distance accuracy parameters were crucial for assessing the participant's

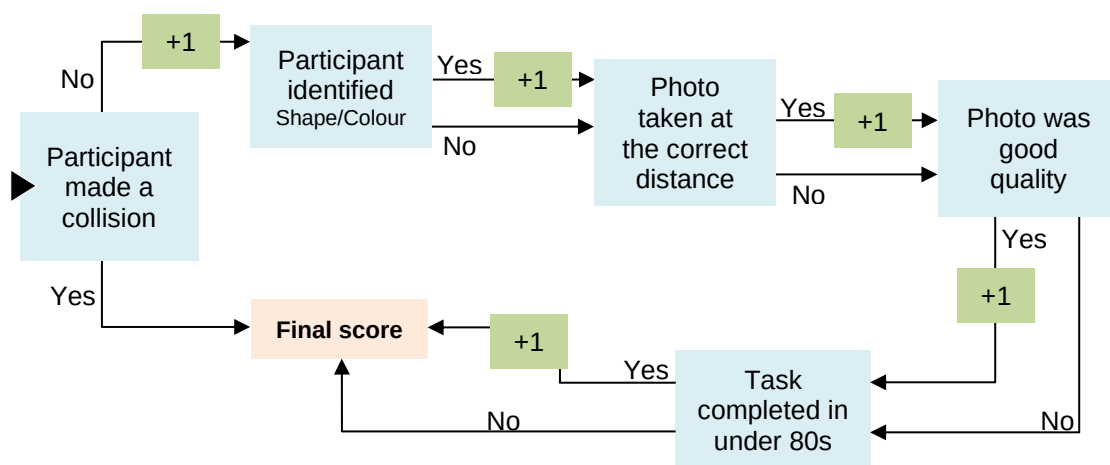


Figure 3-6 Scoring framework for participants performance

ability to correctly interpret and respond to the visual information of the different visualisation modalities provided. Photo quality was included to evaluate the participant's ability to correctly align the robot's camera, a key part of the visual inspection task. Finally, task completion time was included as a parameter to account for the efficiency of task completion, with a threshold set at 80 seconds based on preliminary pilot testing before the main experiment.

Each participant completed three tasks in each modality; the maximum total score across all three tasks for each modality is fifteen points. These total scores were converted into completion percentages to facilitate comparison across participants and modalities. This approach also normalised the scores, allowing for a fair comparison across participants who may have completed different parameters across the different tasks and modalities.

3.5.4 Procedure

First, the order of modalities was counterbalanced to prevent order effects, such as learning of fatigue, from confounding the results. Counterbalancing was employed to eliminate any effects due to the order in which the modalities were presented. The Latin Square design ensures that each modality appears in each position equally often, making the experiment more robust against order effects (Shadish, Cook and Campbell, 2002). Each group was ensured to have an equal number of participants. The modalities of data visualisation were displayed on the Varjo XR-3 (Varjo, 2023). They were developed with Unity 3D Engine using the Varjo XR Plugin for Unity¹ and the Azure Kinect Sensor SDK² for the RGBD cameras.

¹ <https://developer.varjo.com/docs/unity-xr-sdk/getting-started-with-varjo-xr-plugin-for-unity>

² <https://github.com/microsoft/Azure-Kinect-Sensor-SDK>

Automation was a crucial aspect of the experiment, ensuring consistency in how participants experienced the modalities. Using JSON files to control various parameters reduced potential human errors and inconsistencies in experiment setup and progression (Figure 3-7). The "Participant.json" file was in charge of loading participant ID, the experiment order of modalities (1: 2D>3D>AV>2D3D, 2: 3D>AV>2D3D>2D, 3: AV>2D3D>2D>3D, and 4: 2D3D>2D>3D>AV), and language (English or French). The "Questions.json" file was responsible for loading the cognitive workload questionnaire into the correct language. The "Settings.json" was responsible for the modalities, task flow, and display instructions.

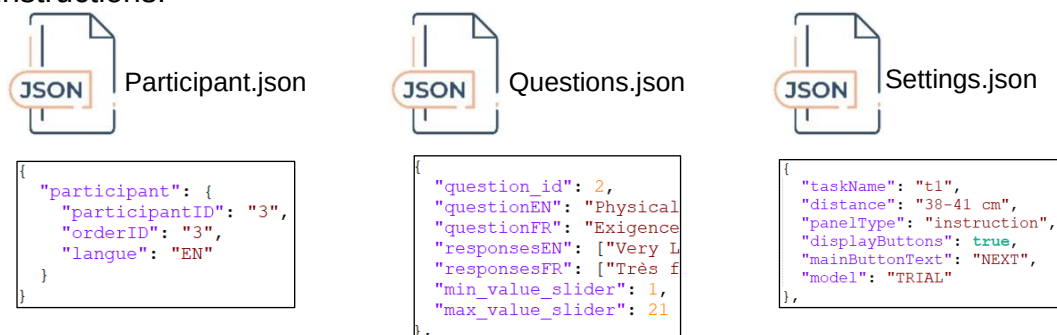


Figure 3-7 JSON configuration files in charge of the experiment flow

The recording of pupils' data was also started and stopped automatically. While most of the experiment was automated, eye calibration required manual initiation to ensure the eye-tracking system was accurately calibrated to each participant's eyes, ensuring dependable pupillometry data.

The experiment flow was carefully crafted to ensure participants were well-acquainted with the setup, understood their tasks, and could provide feedback without feeling rushed. The initial steps, including the welcome, pre-test questionnaire, and software presentation, set the stage, and ensured participants were comfortable and informed. They were handed the consent form detailing the experiment's objectives, methods, and potential risks. This form adhered to the Cranfield University Research Ethics System guidelines (8.3.3Appendix A). The baseline measures captured participants' resting state pupil diameters, providing a reference point against which task-induced cognitive workload could be compared. The allocation of time for each step was based on preliminary trials,

ensuring that participants had adequate time to engage with each modality, practice robot manipulation, and provide feedback without feeling fatigued. The procedure ensured participants' consistency, reliability, and comfort. By employing a counterbalanced design, automating the experiment flow, and carefully allocating time for each step, the aim was to gather accurate data on the cognitive workload associated with different data visualisation modalities. The summary flow of the experiment is exposed below in Table 3-3. The total expected duration for each participant was approximately **41 minutes**, making it a short yet insightful engagement as participants were not paid.

Table 3-3 Flow of the experiment (blue steps are repeated four times)

	Steps	Activities	Duration
	1	Welcome & informed consent signature	2 min
	2	Pre-test questionnaire	2 min
	3	Software and tasks presentation	4 min
	4	Eyes calibration	1 min
	5	Baseline measures (resting state)	1 min
	6	Robot manipulation practice	5 min
 Pupils' diameter + Video Recording	7	Tasks conducted on the modality (done four times)	3 min (x4)
	8	Workload questionnaire (done four times)	3 min (x4)
	9	Debriefing open questions and comments	2 min
The total planned duration of experiment 2: 41 min			

3.5.5 Data analysis

This experiment's data analysis was orchestrated to understand participants' interactions with different visualisation modalities in VR-based industrial robot

teleoperation. The analytical approach spanned both quantitative and qualitative realms. Quantitatively, descriptive statistics were generated for the NASA-TLX, elucidating overarching trends and intricate item-specific insights. Performance indicators, notably task completion time and error rates, were scrutinised using their mean values and standard deviations. A deeper layer of insight was introduced with repeated measures of ANOVA, shedding light on potential interplays between the visualisation modality and the order of their presentation to participants.

3.6 Ethical considerations

The research for this PhD project adhered to Cranfield University's ethical practice (Appendix A). All participants involved in the studies were provided with comprehensive information about the research project, including its purpose, procedures, potential risks, the rights as participants, and the ways their data would be collected, stored, and used. The participants were given sufficient time to read and understand the information before providing informed consent. Consent was obtained before participation, and participants were made aware that they could withdraw their consent at any point without any negative consequences.

The privacy and confidentiality of all participants were strictly protected throughout the research project. All collected data were anonymised and stored securely, with access granted only to the research team members. Any personal details were omitted or pseudonymised in any presentations, reports, or publications used in the research.

4 Enabling remote operations: A comprehensive framework for telepresence capability

4.1 Design Principles for Telepresence Capability Framework

- The framework is designed to integrate with various hardware and software components seamlessly, ensuring broad compatibility and adaptability. It aims to accommodate different robotic systems, sensory inputs, and user interaction devices without requiring substantial modifications. This principle supports the system's future expansion and integration with emerging technologies, fostering a versatile and sustainable telepresence ecosystem.
- At the heart of the framework is the commitment to real-time communication, which is crucial for the fidelity and responsiveness of telepresence experiences. Employing Data Distribution Service (DDS) ensures that data exchange and control commands between the remote operator and the robot are timely and dependable, minimising latency and enhancing the precision of remote operations.
- Recognising the diverse requirements of different telepresence applications, the framework incorporates flexible Quality of Service (QoS) settings. These settings allow for tailored communication protocols that can balance the trade-offs between speed, reliability, and resource consumption according to the specific needs of each application scenario.
- The framework prioritises intuitive and accessible user interfaces and interaction mechanisms, catering to users with varying degrees of technical expertise. By focusing on user experience, the system aims to lower the barrier to entry for operating complex robotic systems and to enhance the effectiveness of human-robot interactions.
- Designed with scalability in mind, the framework allows for the incremental addition of functionalities, robots, or sensors without disrupting the existing system architecture. This modular approach enables users to customise and expand their telepresence setups as needed, promoting scalability and long-term applicability.

4.2 System design

Telexistence-based systems (Figure 4-1) involve multiple interacting elements: an industrial maintenance robot and machinery space, a virtual environment or digital twin, a network platform, and an immersive interface and inputs. These elements are interconnected by data flows and control mechanisms, illustrating how humans interact with and guide industrial machines through virtual interfaces.

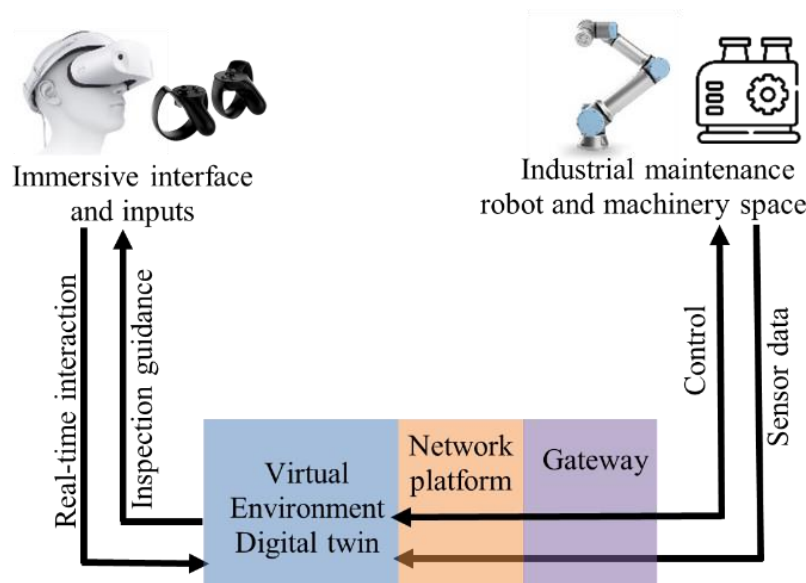


Figure 4-1 Elements to consider as part of a telexistence system

Elements to consider as part of a telexistence system:

- **Industrial maintenance robot and machinery space:** This is the physical environment where the robots and machinery exist and operate. It includes the machinery, robots, and the entire space, including its state, condition, and activities.
- **Sensor Data:** Data is continuously captured from this space using sensors equipped on the robots and throughout the space. This can include, for example, visual data from cameras, temperature data from thermometers, vibration data from accelerometers, and so on. This data is transmitted to the gateway, forming the basis of the "digital twin" or virtual representation of the space.

- **Gateway:** This is the interface where the sensor data from the physical environment is collected and transferred to the network platform. The gateway can also perform initial processing and analysis of the data.
- **Network platform:** This element receives, processes, and transmits the data from the gateway, making it available to the virtual environment.
- **Virtual environment:** This is a real-time representation of the physical environment (the robot and machinery space). It is created and updated based on the sensor data, allowing users to understand the current state and conditions of the environment without being physically present.
- **Inspection Guidance:** Once the virtual environment is created, it guides the inspection process. This guidance is transmitted to the immersive interface and inputs. This may include identifying points of interest, highlighting areas that need attention, or providing instructions for specific maintenance tasks.
- **Immersive Interface and Inputs:** This is where the human operators interact with the system. It provides an immersive experience, allowing the operator to "exist" and interact with the virtual environment. Based on the inspection guidance, the operator can provide inputs that control the robot and machinery.
- **Real-Time Interaction:** The operator's inputs are transmitted to the virtual environment. This allows the operator to manipulate the virtual environment in real-time, affecting the control of the industrial maintenance robot and machinery in the actual physical space.

Figure 4-1 thus illustrates the bi-directional communication and interaction between the physical space and the human operator through a virtual interface.

4.3 System architecture based on DDS

Based on the literature and gathered information described above in this Chapter, a system was proposed to control an industrial robotic arm. Figure 4-2 represents the traditional Data Distribution Service (DDS) Infrastructure, the basis of the proposed framework for controlling an industrial robotic arm. This framework is segmented into three main groups for clarity and understanding:

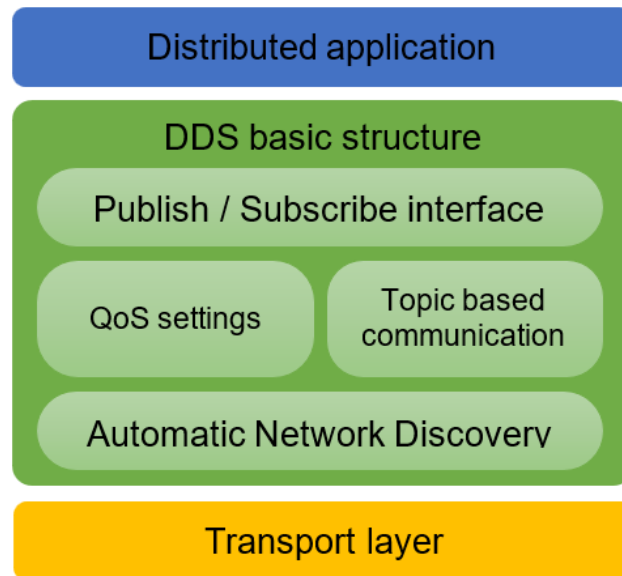


Figure 4-2 DDS Infrastructure

1. **Transport layer:** The infrastructure's foundation constitutes the primary data communication infrastructure. The robot arm is connected via Ethernet to a computer running the gateway application. The computer and robot have defined static local IP addresses, allowing uninterrupted and dependable data communication.
2. **DDS basic structure:** Built upon the transport layer, this represents the core components of DDS, which are indispensable for data distribution in distributed applications:
 - **Automatic discovery network architecture:** This aspect of DDS allows for automatically detecting and connecting nodes, minimising the need for manual network configuration.
 - **Topic-based anonymous communication:** Communication in DDS is based on 'Topics'. A 'Topic' is a channel for a specific data type, which applications can publish or subscribe to. It allows for anonymous communication, as publishers and subscribers do not need to know each other directly.
 - **Quality of Service (QoS) settings:** These parameters define how reliable, urgent, or resource-intensive the data communication should be. It provides a way to control the behaviour of data communication to meet the specific application's needs.

- **Publish/Subscribe interface:** In this interface, nodes publish data under specific 'Topics'; other nodes subscribe to those 'Topics' to receive the data.
3. **Distributed application:** This represents the application layer where the actual robotic arm control application resides. Here, the DDS infrastructure facilitates the transmission of necessary data, which is designed to be reusable across different functions. These data structures are modelled after ROS data types for potential future interoperability with ROS applications in future work.

The proposed framework described in Figure 4-3 involves two remotely situated computers connected through the same network, each containing an XML file outlining the chosen DDS system's QoS profile. This profile includes vital elements such as the 'initial peers list' (hostnames of the two machines), which is instrumental during the participant phase of the discovery process. Additional attributes of the QoS profile include the naming of participants ("Operator" and "Fanuc") and defining the size constraints for the buffer.

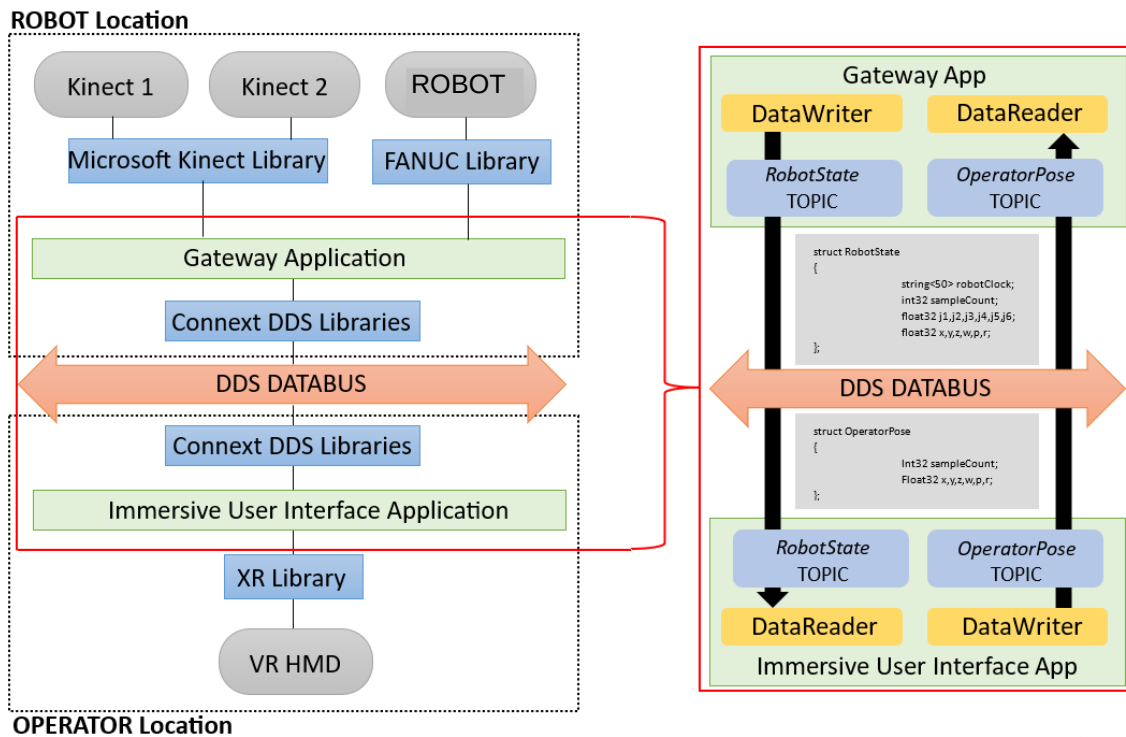


Figure 4-3 Architecture and components overview of the proposed framework

The QoS profile employed is based on the BEST_EFFORT_RELIABILITY_QOS, which is particularly suited for frequent updates of sensor values due to its efficiency and low resource intensity. This approach delivers the most recent value of a topic from DataWriters to DataReaders without accounting for prior samples sent. However, it is essential to note that this method does not guarantee data delivery and may be susceptible to losses arising from connectivity issues or limitations within the physical transport layer.

The setup also integrates two Azure Kinect cameras placed strategically around the robot. One camera focuses on the table in front of the robot to generate the point cloud, while the other provides a broader view of the operation area by focusing on the robot itself. The two cameras operate under different configurations and are connected to the computer on the robot side.

The UI Figure 4-4 can be initiated from any location. While the current framework shows promise, future work must be deployed across different robots. Preserving the current structure of the framework is feasible for all FANUC robots, requiring only the import and setup of new robot CAD files in Unity 3D. However, the gateway application must be redesigned with relevant libraries to extend this framework to robots of different brands. Furthermore, it is essential to note that

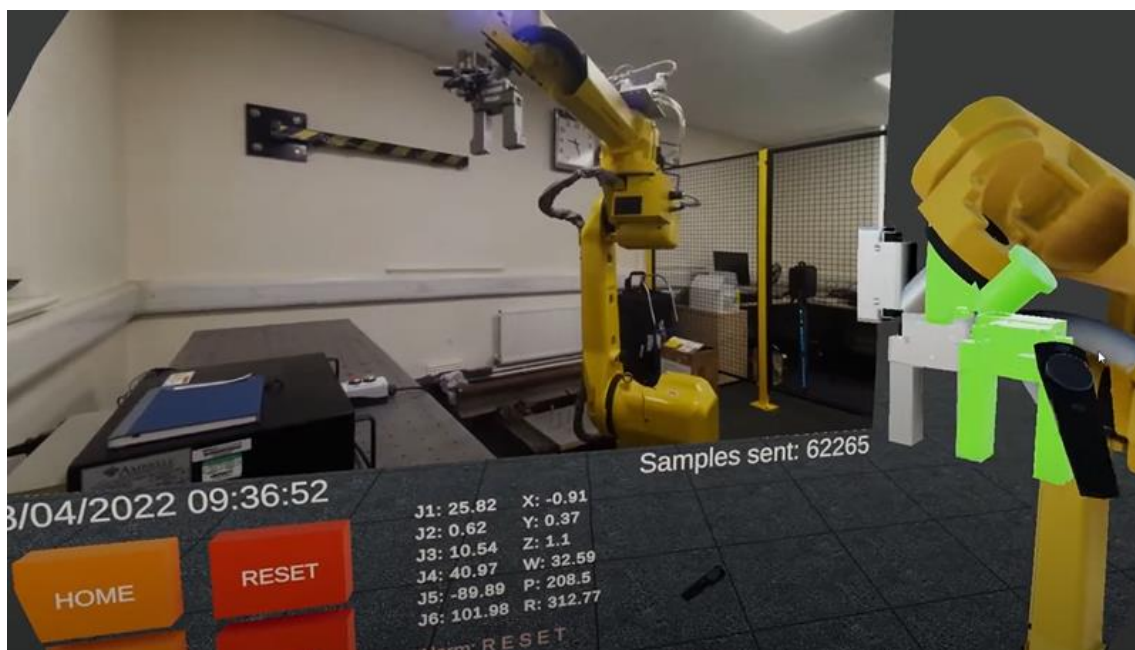


Figure 4-4 Proposed user interface

FANUC SDK requires the installation of a specific robot server program. This program allows access to the robot controller and facilitates communication using TCP/IP with the robot and the gateway application via a COM interface.

Table 4-1 provides an overview of the various data structures used in the DDS-based framework to facilitate communication between the FANUC robot and the remote operator. The data structures represent distinct types of data samples transmitted in both directions, each comprising specific members with defined roles.

Table 4-1 Data structures description of samples transmitted

	Data Structure Name	Data Structure Members
From the FANUC robot to the remote operator	Robot_State	- Clock (string /robot internal timestamp) - Sample (int/samples count) - J1, J2, J3, J4, J5, J6 (floats/joints angles in degree) - X, Y, Z (floats/robot world position) - W, P, R (floats/robot world rotation)
	Robot_Point_Cloud	- Clock (string/robot internal timestamp) - Sample (int/samples count) - Sequence Memory (floats/ sequence containing points position and Colour data)
	Robot_Image	- Clock (string/robot internal timestamp) - Sample (int/samples count) - Memory (bytes/sequence containing MPEG image data)
	Robot_Alarm	- Clock (string/robot internal timestamp) - Sample (int/samples count) - Message (string/alarm messages)
	Robot_Reachability_State	- Clock (string/robot internal timestamp) - Sample (int/samples count) - IsReachable (bool/state regarding reachability for a position by the robot)
From the remote operator to the FANUC robot	Operator_Request	- Clock (string/operator computer timestamp) - Sample (int/samples count) - Buttons (enumeration of buttons, RESET, ABORT, HOME, PATH)
	Operator_Teleop_Target	- Clock (string/operator computer timestamp) - Sample (int/samples count) - X, Y, Z (floats/operator target position) - W, P, R (floats/operator target rotation)

	Operator_Path_Points	- Clock (string/operator computer timestamp) - Sample (int/samples count) - ID (int/modified pose ID) - IsUpdating (boolean/sample is for updating point values) - IsDelete (boolean/sample is for deleting pose) - X, Y, Z (floats/pose position values) - W, P, R (floats/pose rotation values)
--	----------------------	---

Each data structure is designed to carry specific information pertinent to the robot's operation, making the communication between the robot and the operator efficient, organised, and transparent.

Appendix E details the technical foundation and implementation specifics of the Fanuc Remote Control Framework. The framework facilitates remote operations of a Fanuc industrial robot, leveraging DDS for real-time data exchange and operational command transmission between the robot and the remote operator.

4.4 Immersive interface design

Using an industrial robotic arm typically requires experience. Recognising this, developing a telexistence rig was initiated with a focus on simplicity, especially for untrained users. To achieve this, specific functionalities were designed to be user-friendly and intuitive:

- Interact with the immersive UI.
- Interact with the robot's digital twin.
- Jog³ the robot.
- Register a new user pose.
- Switch between robots' programs.
- Abort the current robot motion.

Research on interaction techniques and metaphors has been crucial for developing immersive virtual environments, as presented by Bowman (Bowman,

³ Jogging the industrial robot is the term used to describe the act of manually moving the robot via the user interface.

1999) and Poupyrev (Poupyrev et al., 1998). Bowman proposed interaction guidelines depending on the task and distance from the user. Poupyrev compared remote selection techniques, including ray cast and non-linear mapping of the real hand position.

The interaction in the 3D environment is conducted as shown in Figure 4-5; a controller button is dedicated to the teleport function (1), another for a pointer (2), and one for direct manipulation of the 3D cursor (3). Teleportation was chosen as a mode of movement in the VR environment due to its minimal tendency to induce VR sickness and contribute to higher performance (Prithul, Adhanom and Folmer, 2021). Remote button selection has been done using the basic raycasting technique. As for manipulating the robot, the primary way of programming a robot without offline simulation software is to jog the robot using

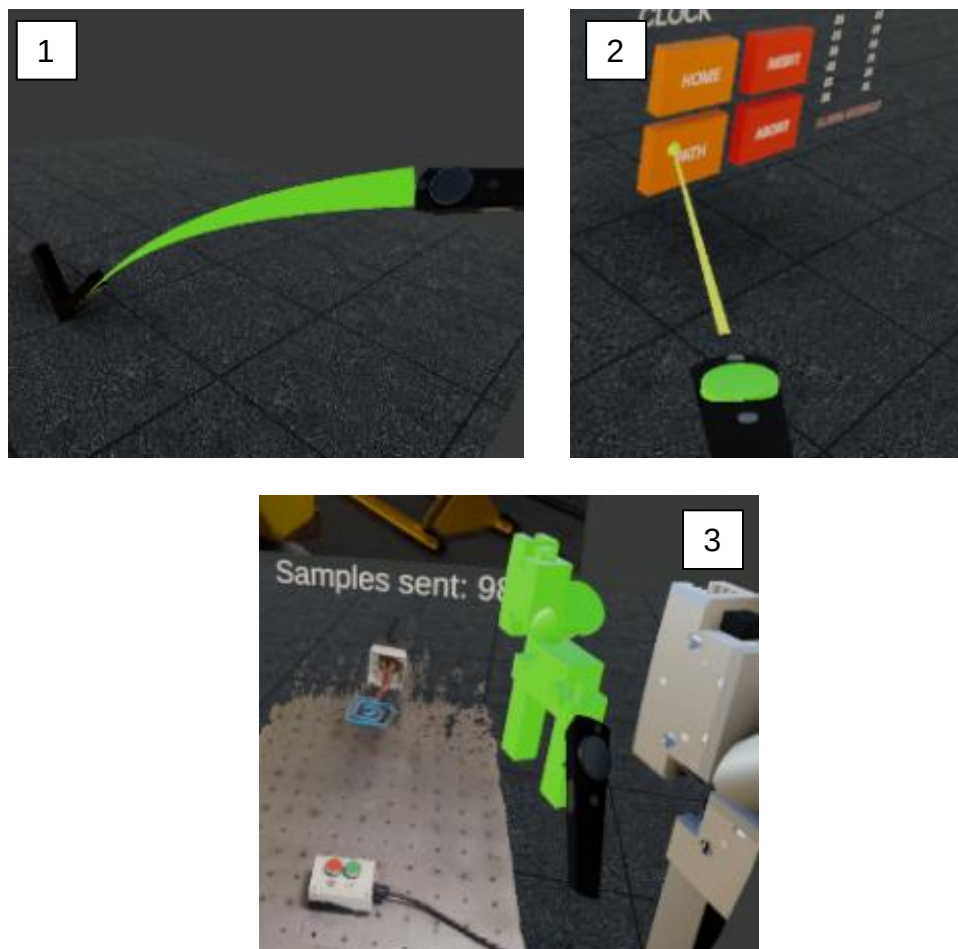


Figure 4-5 Interaction methods in the 3D environment for moving around (1), interact with the UI (2), and interact with the robot (3)

the teach pendant device (a handheld control device often used in industrial robotics to program and control robotic machines directly and manually). Within the immersive virtual environment, a 3D cursor will aim to replace the teach pendant device. Hand metaphors are often used for direct manipulation in a VR environment (Wagner, Stuerzlinger and Nedel, 2021) and are sometimes combined with hand-tracking technologies (Han et al., 2020). The decision was made to use virtual controller models exclusively, as they enable users unfamiliar with VR to understand quickly which buttons they are touching and activating.

The key components and functionalities of the interface design are illustrated in Figure 4-6 and described below:

- First, Figure 4-6 (1) the 3D buttons for the following features:
 - The HOME program button asks the robot to return to its default pose.

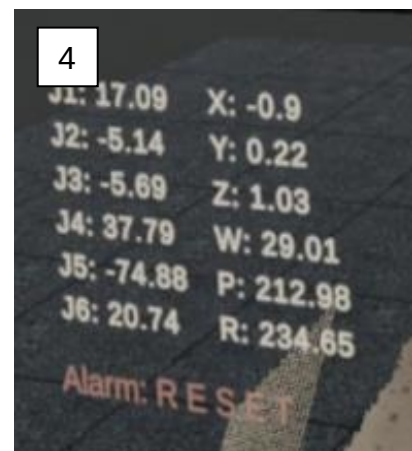


Figure 4-6 Components of the interface: UI of the framework, 3D buttons (1), 2D video feed viewer (2), point cloud (3), and robot pose values (4)

- The “PATH” button to manually teleoperate the robot.
- The “RESET” button is used to reset the current robot controller unit fault.
- The “ABORT” button aborts the current robot controller unit task.
- Figure 4-6 (2) a 2D video feed viewer of the first Kinect with the remote robot's current ping in milliseconds. The number of samples received from the robot is on the bottom right of the viewer.
- Figure 4-6 (3) a 3D visualisation of the 3D data received from the second Kinect. Additional Unity packages were used, including a High-Definition renders pipeline and Visual Effect Graph, a node-based visual logic graph creator. The Visual Effect graph created for this project takes a mesh generated from the received Kinect data by a unity C# script as an input, using vertices of this mesh to output particles in a cube shape. Position and Colour are updated twenty times per second (20 samples per second). Iterations were conducted to improve the performance of point cloud data streaming but at the expense of quality. Several iterations were conducted to optimise point-cloud processing speed for VR environments, as detailed in
- Table 4-2. The initial stage involved managing raw data from the Kinect sensor, which included both the Kinect mesh and a PNG

Table 4-2 Iterations done for optimising the point-cloud processing speed

Description	Point-cloud size per sample	Average samples per second received in Unity 3D
Raw Kinect mesh + PNG texture	~ 13 MB	0,5
Manual Decimation: deletion of 1 point out of two + colours and max distance filtering	~ 4 MB	14
Decimation + byte data compression API	~ 0,8 MB	20

texture, resulting in a substantial point-cloud size of approximately 13 MB per sample. This large size limited the system to processing only about 0.5 samples per second in Unity 3D. A manual decimation technique was applied to improve this, which involved selectively removing every other point, filtering out points based on colours (solid black), and excluding points beyond a 2-meter range (Figure 4-7). This method effectively reduced the point-cloud size to around 4 MB per sample, consequently increasing the processing speed to fourteen samples per second. The final optimisation integrated byte data compression using the LZ4⁴ algorithm alongside the previous manual decimation. This combination further decreased the size of each data sample to approximately 0.8 MB, enabling the system to process around twenty samples per second.

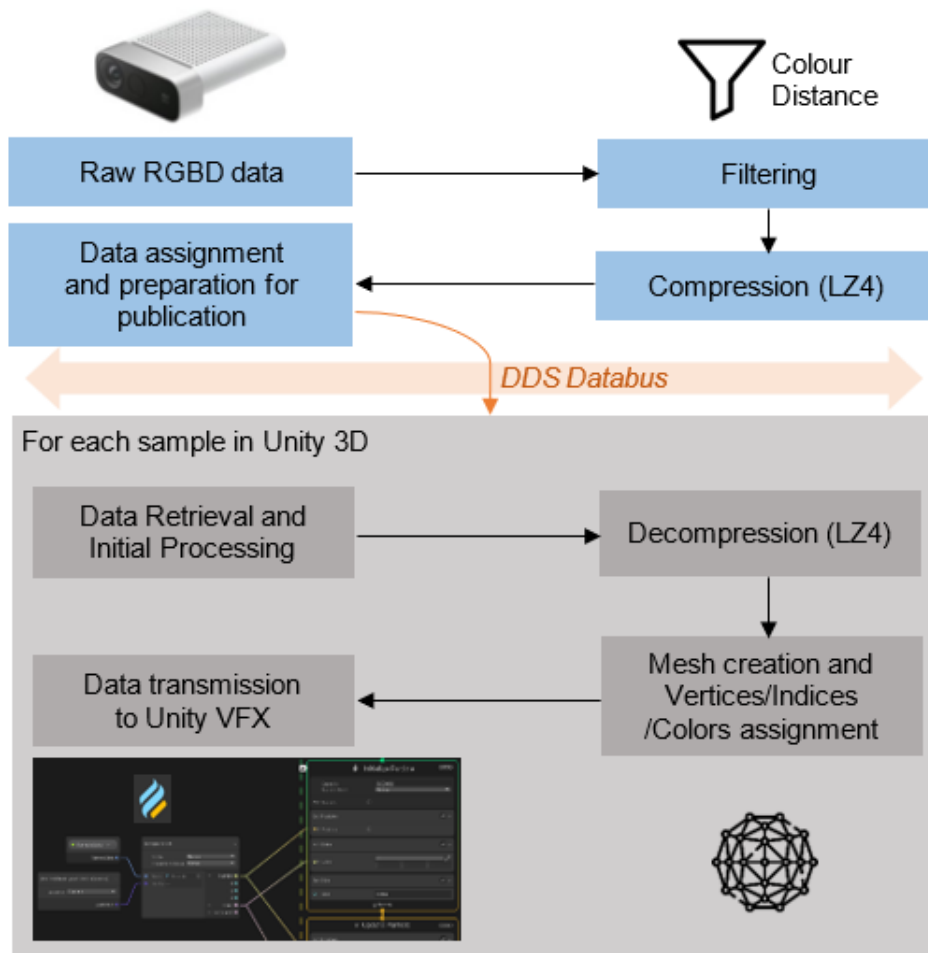


Figure 4-7 Pipeline for transmitting the point-cloud data

⁴ <https://github.com/MiloszKrajewski/K4os.Compression.LZ4>

- Figure 4-6 (4) angles value of the six joints of the robot in degrees, world pose in millimetres and degree of the robot, and the last alarm message is also displayed to the user. It is helpful if someone tries to access the robot fence; the robot will send a “fence open” alert message while stopping the robot program and motion.

4.5 Use case application: Fanuc industrial robot

4.5.1 Control of the physical robot

To demonstrate the proposed framework, a FANUC M-20iA was used (Figure 4-8). It is controlled by an R-30iB mate cabinet and equipped with a pneumatic gripper. Fences and Safety limits surrounding the robot have been set to prevent it from colliding with the environment.

The table below describes the robot specification and motion range.



Figure 4-8 Fanuc M-20iA (left) and R-30iB mate cabinet (right)

Table 4-3 FANUC M-20iA specification and motion range

Specification		Motion Range	
Axes	6	J1	$\pm 370^\circ$
Payload	20 kg	J2	$\pm 260^\circ$
H-Reach	1811 mm	J3	$\pm 458^\circ$
Repeatability	± 0.08 mm	J4	$\pm 400^\circ$
Mass	250 kg	J5	$\pm 380^\circ$
Structure	Articulated	J6	$\pm 900^\circ$
Mounting type	Floor		

4.5.2 Digital representation of the robot

The virtual representation is part of the digital twin concept, which denotes a system encompassing three primary subsystems: a tangible asset in the physical space, an intangible virtual representation of the asset in the digital space, and data and information links between physical and virtual spaces (Grieves, 2005). The concept of a digital twin is increasingly popular in robotics and automation, and it can have profound implications when integrated with game development engines such as Unity 3D. The importance of structuring scene elements (Figure 4-9) within a hierarchical framework cannot be overstated. Such a structure enables the effective management of received and sent state information afterwards. Each robot axis joint (J1-6) is assigned a parent-child relationship and updated thanks to the FANUC controller unit.

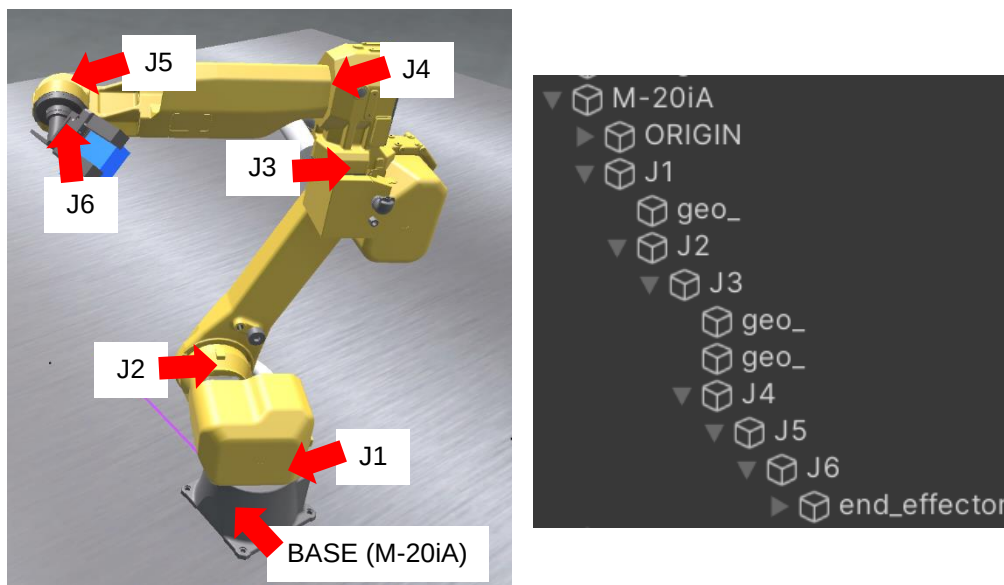


Figure 4-9 Hierarchy robot elements setup in the software Unity 3D

4.5.3 Coupling the physical and digital twins

The selection of Unity 3D as the platform for developing the digital twin (Figure 4-10) of a robotic system is underpinned by several compelling factors, distinguishing it from alternative frameworks like Unreal Engine. Unity 3D based on C# language is celebrated for its robust cross-platform support and versatility. It is the most popular game engine amongst robotics researchers and a leader in XR development (Mizuchi and Inamura, 2017). Unity 3D also has its own hub for

robotics simulation based on ROS⁵. The extensive documentation and an active community surrounding Unity provide an invaluable resource pool, facilitating problem-solving and development efficiency. Furthermore, Unity's user-friendly interface and development tools enable rapid prototyping and iteration, significantly accelerating development. Unity can seamlessly integrate with external systems and C# libraries, including communication protocol frameworks like RTI DDS. This amalgamation of features—ranging from interoperability, advanced simulation, and community support to graphical excellence and ease of use—solidifies Unity 3D's position as a particularly suitable platform for the intricate and nuanced development of digital twins in robotic applications (Moshayedi, Kolahdooz and Liao, 2022; Santos Pessoa de Melo et al., 2019).

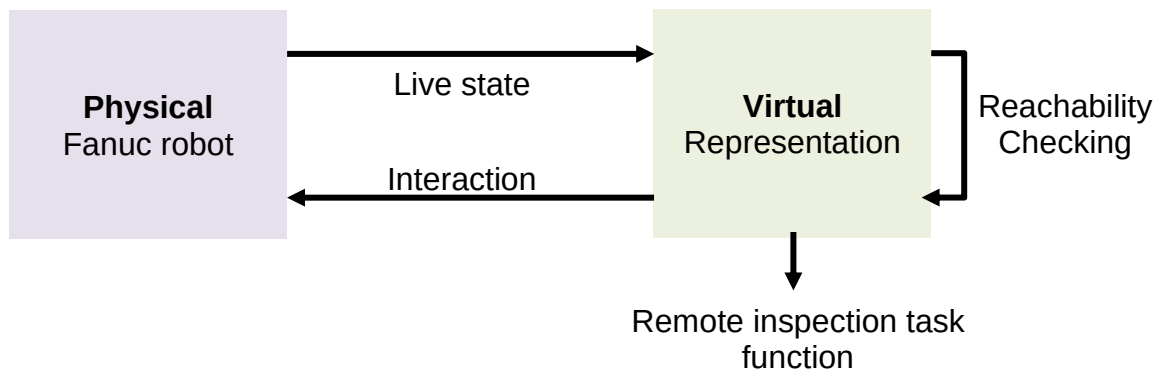


Figure 4-10 Digital twin for remote inspection task

The real-world robot is a tangible asset in physical space. In contrast, the intangible virtual representation in the digital space is the robot's model created in Unity 3D. Integrating the robot's virtual representation within Unity 3D requires careful setup.

Indeed, the Unity 3D and FANUC coordinate systems are different, as shown in Figure 4-11. The FANUC robot control systems use a system known as WPRXYZ to specify the position and orientation of the robot's end effector (the tool at the end of the robotic arm). It also refers to the Tool Center Point (TCP).

⁵ <https://github.com/Unity-Technologies/Unity-Robotics-Hub>

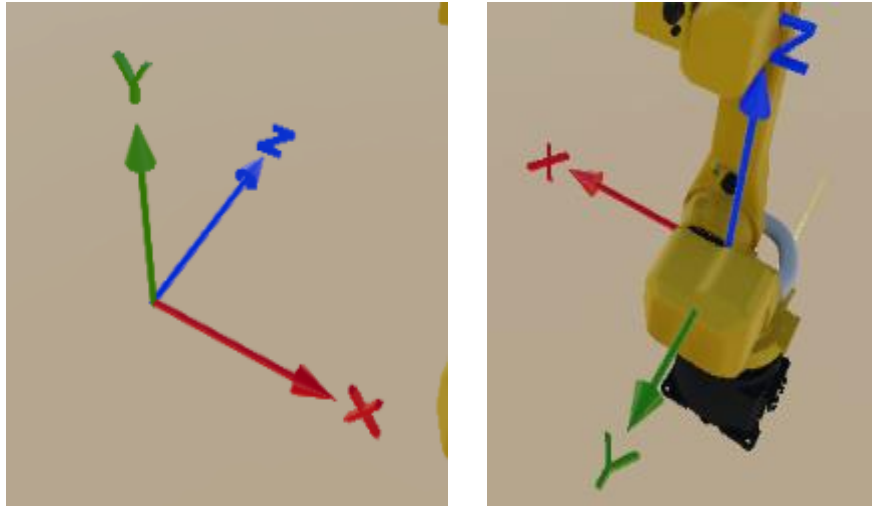


Figure 4-11 Coordinate systems of Unity 3D (left) and Fanuc (right)

WPR stands for Wrist, Pitch, and Roll. These are three of the rotational degrees of freedom that define the orientation of the robot's end effector. In other words, they describe how the end effector is rotated about the X, Y, and Z axes, respectively. These are typically represented as Euler angles or a similar rotation representation.

XYZ refers to the Cartesian coordinates in three-dimensional space. These define the position of the end effector in terms of its distance along the X, Y, and Z axes from a reference point, usually the robot's base.

FANUC WPRXYZ is a six-dimensional vector that provides a complete description of the position and orientation of a robot's end effector in space. The first three dimensions (XYZ) provide the position, and the last three dimensions (WPR) provide the orientation. The units for XYZ are typically in millimetres, while the units for WPR are typically in degrees.

4.5.3.1 From the FANUC robot to Unity 3D

In the Unity 3D application, the transform (position and rotation) of a `GameObject`⁶ represents the robot's virtual end effector. First, for the rotation,

⁶ In the Unity 3D game engine, a `GameObject` is the most fundamental object in a scene. Everything in a scene in Unity 3D is a `GameObject`, such as characters, lights, cameras, special effects, props, and more. It is the base class for all entities in Unity 3D.

converting FANUC WPR (Wrist, Pitch, Roll) angles to a Quaternion is done with a method created to convert the FANUC WPR into a quaternion and detailed next. Quaternions are a mathematical construct that is particularly useful in providing a convenient mathematical description of rotations in three-dimensional space and are used in Unity 3D game engines. They are widely used in computer graphics, robotics, navigation, and physics for their computational efficiency and to avoid specific problems that can arise with other rotation methods, such as gimbal locks.

Given the Wrist (W), Pitch (P), and Roll (R) angles in degrees, these angles are first converted to radians:

$$\begin{aligned} W_{rad} &= W \times \frac{\pi}{180} \\ P_{rad} &= P \times \frac{\pi}{180} \\ R_{rad} &= R \times \frac{\pi}{180} \end{aligned} \tag{4-1}$$

Then, the components of the quaternion (q_x, q_y, q_z, q_w) are calculated using the specified formulas by roboception (roboception, 2021):

$$\begin{aligned} q_x &= \cos\left(\frac{R_{rad}}{2}\right) \times \cos\left(\frac{P_{rad}}{2}\right) \times \sin\left(\frac{W_{rad}}{2}\right) - \sin\left(\frac{R_{rad}}{2}\right) \times \sin\left(\frac{P_{rad}}{2}\right) \times \cos\left(\frac{W_{rad}}{2}\right) \\ q_y &= \cos\left(\frac{R_{rad}}{2}\right) \times \sin\left(\frac{P_{rad}}{2}\right) \times \cos\left(\frac{W_{rad}}{2}\right) + \sin\left(\frac{R_{rad}}{2}\right) \times \cos\left(\frac{P_{rad}}{2}\right) \times \sin\left(\frac{W_{rad}}{2}\right) \\ q_z &= \sin\left(\frac{R_{rad}}{2}\right) \times \cos\left(\frac{P_{rad}}{2}\right) \times \cos\left(\frac{W_{rad}}{2}\right) - \cos\left(\frac{R_{rad}}{2}\right) \times \sin\left(\frac{P_{rad}}{2}\right) \times \sin\left(\frac{W_{rad}}{2}\right) \\ q_w &= \cos\left(\frac{R_{rad}}{2}\right) \times \cos\left(\frac{P_{rad}}{2}\right) \times \cos\left(\frac{W_{rad}}{2}\right) + \sin\left(\frac{R_{rad}}{2}\right) \times \sin\left(\frac{P_{rad}}{2}\right) \times \sin\left(\frac{W_{rad}}{2}\right) \end{aligned} \tag{4-2}$$

The quaternion is represented as $q = (q_x, q_y, q_z, q_w)$. In Unity, the quaternion is converted into Euler Angles to apply it to the rotation of GameObjects effectively. These Euler angles are represented as a Vector3, where X, Y, and Z correspond to the rotation of their respective axes.

$$GameObject_{localEulerAngles} = \begin{pmatrix} X = EulerAngles.x \\ Y = -EulerAngles.y \\ Z = -EulerAngles.z \end{pmatrix} \tag{4-3}$$

Secondly, the FANUC XYZ values are divided by one thousand to convert them from millimetres to meters, the unit of measurement used in Unity 3D. The X position values are also negated to account for the difference in the coordinate systems used by FANUC and Unity 3D.

$$GameObject_{localPosition} = \begin{pmatrix} X = -\frac{fanuc_x}{1000} \\ Y = \frac{fanuc_y}{1000} \\ Z = \frac{fanuc_z}{1000} \end{pmatrix} \quad (4-4)$$

4.5.3.2 From Unity 3D to the FANUC robot

To obtain the Fanuc position and rotation representation from the Unity 3D application, the Unity 3D GameObject quaternion is first converted to serve as the controller for updating the FANUC real robot. Here, q_w q_x q_y q_z are the quaternion components, W P R represent the Wrist, Pitch, and Roll angles, respectively.

$$\begin{aligned} W &= \arctan\left(\frac{2(q_w q_x + q_y q_z)}{1 - 2(q_x^2 + q_y^2)}\right) \times \left(\frac{180}{\pi}\right) \\ P &= \arcsin\left(2(q_w q_y - q_z q_x)\right) \times \left(\frac{180}{\pi}\right) \\ R &= \arctan\left(\frac{2(q_w q_z + q_x q_y)}{1 - 2(q_y^2 + q_z^2)}\right) \times \left(\frac{180}{\pi}\right) \end{aligned} \quad (4-5)$$

The GameObject position values are multiplied by one thousand to convert them from meters to millimetres, which is the unit of measurement used by FANUC. The X position values are also negated to account for the difference in the coordinate systems used by FANUC and Unity.

$$Fanuc_{Position} = \begin{pmatrix} X = -GameObject_x \times 1000 \\ Y = GameObject_y \times 1000 \\ Z = GameObject_z \times 1000 \end{pmatrix} \quad (4-6)$$

4.6 Conceptual design refinement of the telexistence-based framework

Figure 4-12 describes the elements implemented in the telexistence framework at the vessel and maintainer levels. The DDS protocol allows communication between both. The DDS protocol facilitates communication and coordination between the various components of a telexistence system, such as the remote robot, the VR interface, and the digital twin. For example, the DDS protocol publishes sensor data thanks to the robot's gateway and subscribes to the maintainer's commands, allowing real-time, bi-directional communication between the remote and maintainers. The virtual representation of the robot for telexistence provides a more realistic and immersive experience for the maintainer controlling the remote robot (Lopez Pulgarin et al., 2022). It can simulate and analyse its behaviour and performance (Rosen et al., 2015). In the context of telexistence for remote maintenance, the virtual representation of the remote robot can be created and used to provide real-time feedback and data on the movements and position of the robot regarding the machinery spaces

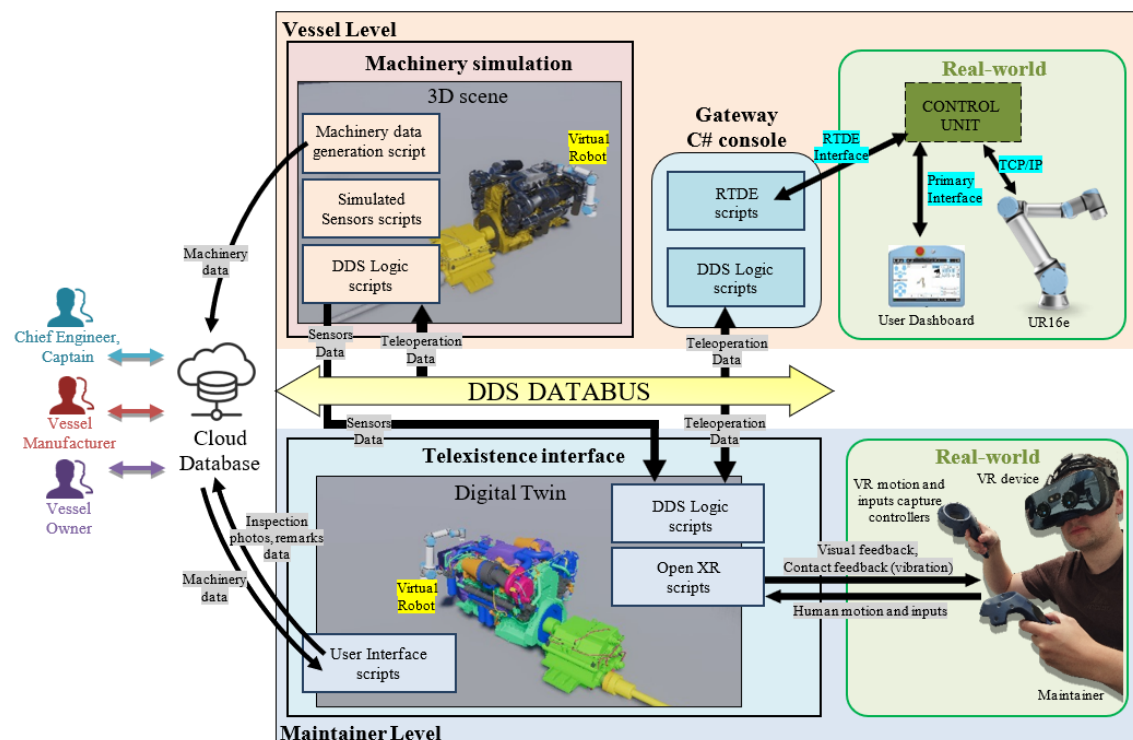


Figure 4-12 High-level system architecture of the refined framework and data flow overview

environment. This feedback can be displayed to the maintainer controlling the robot, allowing them to understand the remote environment better (Tan et al., 2021).

Additionally, the virtual robot helps to simulate potential scenarios, allowing the maintainer to evaluate different actions and strategies without physically affecting the remote robot. This simulation setting benefits training and testing and makes more informed decisions about operating the robot remotely. In the framework Figure 4-12, the marine engineer uses the machine space of the digital twin vessel and the digital twin robot to operate the latter. They take advantage of the virtual representation of the asset to be inspected by combining the RGB camera sensors and the available data. In a vision of the future of a system based on teleexistence CBM, data can be made available to the chief engineer and captain, the vessel manufacturer and the vessel owner through a cloud-based database and web interface. The teleexistence interface displays the data in an immersive and interactive 3D environment for the maintainer to report and store maintenance procedures, such as inspection photos, so the relevant stakeholders can access them.

4.7 Latency and time-delay discussion

Latency and time delay were used as metrics to analyse performance. Time delay in robotics refers to the interval between issuing a command and its execution, encompassing aspects like signal processing and physical distance, which can impact the system's real-time performance. Latency focuses on the delay in transmitting data over a network or communication channel. It is affected by network congestion, bandwidth limitations, and processing delays at intermediate nodes. Both time delay and latency can influence the performance of robotic systems. Still, they differ in scope: time delay refers to internal delays within the system, while latency refers to data transmission over a network.

To evaluate the latency between the two DDS participants, the C# console application, and the Unity 3D application, one acts as a publisher and the other as a subscriber. The publisher will send data, and the subscriber will receive and echo it back to measure the latency. The same as the round-trip time (RTT) test

between a client and a server. The evaluation was performed under controlled conditions within the university's local network infrastructure, where the server and client machines were connected. This setting was chosen to mimic a real-life usage scenario while keeping extraneous network variables to a minimum.

A preliminary 'ping' test was conducted between the two machines to obtain a baseline for the network latency. The ping command works by sending an Internet Control Message Protocol (ICMP) Echo Request to a specified interface on the network and waiting for a reply. The time taken from sending the command until receiving the network's response gives an approximate measure of the latency. In our tests, the round-trip times in milliseconds (ms) averaged eight ms, with a maximum recorded latency of 9 ms. This baseline latency represents the best-case scenario, i.e., the minimum network-induced latency when no other processing or operational delays are considered.

Secondly, during the test within the 2 DDS participants with a test message sent (simple string message), the average latency of **68.68 ms** was measured, with a minimum latency of **41.99 ms** and a maximum latency of **177.99 ms**. In exploring the Fanuc Tool Center Point (TCP) and its integration with Unity 3D software for VR teleoperation, our analysis reveals significant insights into the discrepancies between intended and actual robot paths. Through a detailed comparison, illustrated in 8.3.3Appendix F, we examine the deviations in both the position and rotation of the TCP during a VR manipulation sequence illustrated in Figure 4-13. It compares a randomly generated path in VR (visualised in red) against the actual movement path of the Fanuc TCP (shown in blue).

The differences highlighted in graphs in 8.3.3Appendix F, in terms of position (in millimetres) and rotation (in degrees), underscore the challenge of aligning virtual commands with physical robot movements. The observed discrepancies, with rotations varying from **21.93° to -23.54°** and positions from **-209.65 mm to 228.60 mm**, can be attributed to a multitude of factors, including latency, mechanical delays, and the inherent inaccuracies in the VR's tracking system. These variations are especially pronounced during rapid or sharp operator movements, indicating a direct correlation between the operator's control dynamics and the

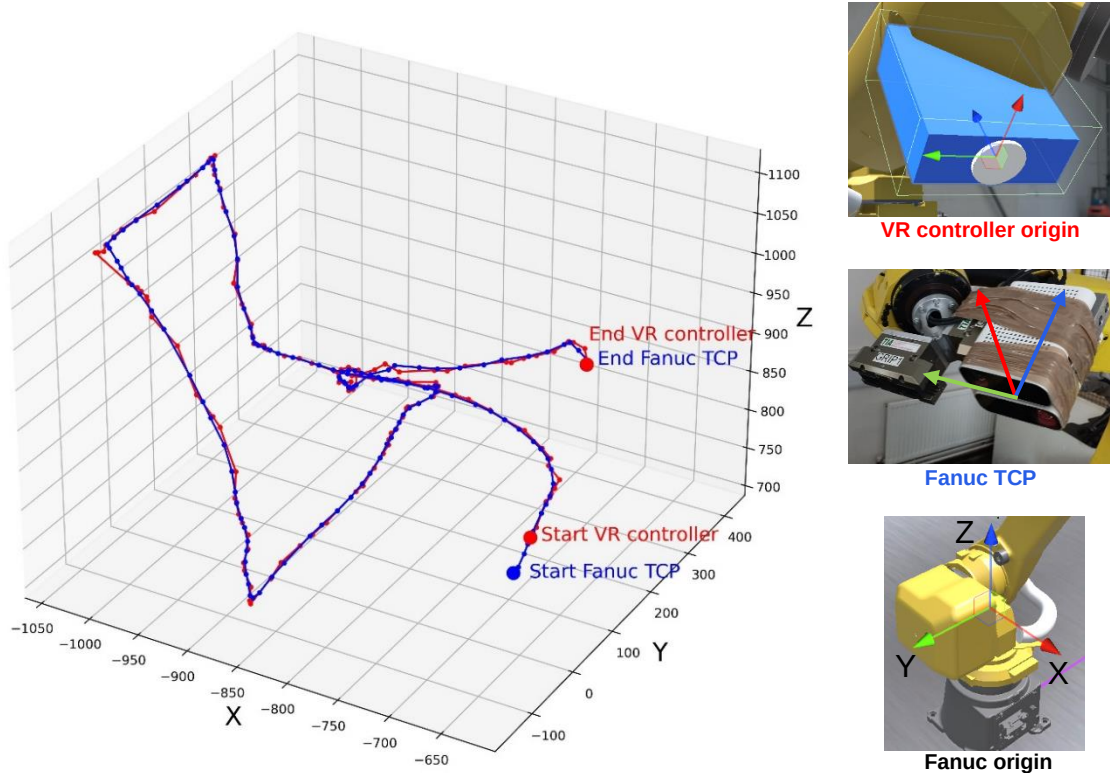


Figure 4-13 Robot trajectory recording: 133 data samples

robot's ability to replicate these movements accurately. The priority for Human-Computer Interaction (HCI) in VR systems is to ensure a natural and intuitive control environment, mimicking real-time responsiveness to maintain the operator's engagement and control proficiency (Murphy et al., 2010)

This analysis emphasises the importance of effectively managing latency within telepresence systems to minimise such discrepancies. While adjustments to the Fanuc robot's CNT (Cornering Numerical Control) parameter, enhancements in network infrastructure, and code optimisation can mitigate time delays and latency, integrating solutions within the teleexistence interface that can adeptly manage these challenges becomes imperative. The CNT parameter in Fanuc robots is a crucial setting that influences the robot's movement behaviour, particularly around corners and during trajectory changes. This parameter adjusts the balance between the robot's speed and precision. A higher CNT setting allows the robot to move through its programmed path with less deceleration at corners or changes in direction, prioritising speed, and fluidity of movement.

Conversely, a lower CNT value makes the robot slow down more at these points to achieve higher precision and tighter adherence to the exact path.

The findings from our analysis illustrate the intricate balance required between speed and accuracy in robotic operations, especially under varying operator control behaviours. The system demonstrates a higher degree of accuracy for meticulous, slow operations, as it affords the control system sufficient time to match the operator's intent precisely. Conversely, prioritising a constant speed during fast-paced, dynamic operations helps maintain fluid and responsive robot movements, albeit with a compromise on absolute positional accuracy. Such insights are pivotal for developing more adaptive and resilient teleoperation systems, ensuring they can accommodate various operational demands and operator preferences. Accuracy becomes the focus in use cases like teleoperation in medicine due to the critical nature of surgical procedures and patient safety (Korte et al., 2014; Murphy et al., 2010).

Studies suggest that delays beyond one hundred milliseconds can disrupt the seamless nature of real-time interactions, negatively impacting user satisfaction and efficiency in task execution (Card et al., 1991; Miller, 1968). The sensitivity to latency varies among different applications, with precision-critical tasks like remote telesurgery exhibiting a low tolerance for delay, contrasting with less critical applications such as remote visual inspections. Integrating autonomy and artificial intelligence (AI) presents a promising avenue to alleviate the effects of latency, particularly in scenarios where real-time decision-making is crucial. An example is the autonomy implemented in Mars rovers, designed to navigate, and avoid obstacles independently despite the substantial communication delay between Earth and Mars. This autonomy enables the rover to continue operations effectively, even when commands from Earth are delayed. Incorporating such AI-driven autonomy into telepresence systems can significantly enhance their resilience to latency, ensuring that operations remain smooth and efficient, regardless of the inherent delays in communication. This approach addresses the direct challenges posed by latency and leverages technology to adaptively

overcome these hurdles, thereby expanding the capabilities and applications of telepresence systems.

4.8 Limitations

Notably, the Robot Operating System (ROS), commonly used in robotics projects, particularly for research, was not employed in this project. However, ROS2 has introduced the DDS protocol, which tends to be adopted by industries (Mayoral-Vilches et al., 2020). When developing an immersive interface for robotics uses, difficulties still need to be solved with the interoperability of commercial and open-source tools. The development is complex; for instance, the FANUC SDK depends on the operating system. The company does not provide official ROS packages; however, unofficial packages use Karel for communication. Garg et al. (Garg, Kuts and Anbarjafari, 2021) have also recently developed a digital twin application of a FANUC robot using Karel for communication. Most research involving the teleoperation of industrial robotics like FANUC is limited to simulation. The framework has been evaluated with an actual physical industrial robot.

Gaining insights into the overall functionality and performance of the system interface was challenging due to the difficulty in directly comparing the proposed system with other advanced systems. This difficulty arose from variations in system setups, like differences in robot platforms, number of variables, transmission delays, and types of visual aids, which are not easily replicable across different systems. The focus was primarily on confirming the feasibility of the suggested systems and interface solutions by implementing and using the interface in a real-world setting. Despite thoroughly examining system performance in VR teleoperation, the present study has several inherent limitations. Firstly, the consistency and smoothness of operator movements affect both the latency and time delay measurements. The uncontrolled nature of these movements could introduce variability and uncertainty in the data, making it difficult to ascertain the full impact of the system's design on the results. Secondly, the limitations of the VR tracking system, such as inaccuracies, could have influenced the measurements of position and rotation discrepancies. These

inaccuracies may exaggerate the perceived time delay and latency, skewing the results. Thirdly, the tests were conducted under controlled conditions within a local network infrastructure, limiting the potential influence of external network variables. While this approach allows for a focused analysis, it may only partially replicate the conditions found in real-world applications where network variability, congestion, and latency could impact performance. Finally, the chosen setting of the FANUC robot's CNT parameter, which prioritises speed constancy over accuracy, inherently introduces inaccuracy to the system's operation. Although this trade-off was justified for the specific use case in this study, it may limit the applicability of the findings to other use cases where a different balance of speed and accuracy may be required.

5 Experiment 1: Evaluating remote visual inspection interfaces

5.1 Introduction

Navies worldwide are increasingly exploring and developing surface vessels designed for minimal or zero crew operation. Unmanned surface vessels (USVs) have become indispensable in modern naval operations due to the hazardous, repetitive, and asset-based tasks traditionally performed by personnel during missions. The deployment of unmanned vessels necessitates re-evaluating maintenance strategies and equipping marine engineers with innovative tools for remotely monitoring and maintaining UVS. Preventive maintenance, such as Condition-Based Maintenance (CBM), is crucial for these vessels. CBM relies on various sensors for data acquisition, processing, assessment, diagnosis, and decision-making. Given the absence of onboard personnel in unmanned vessels, CBM becomes even more essential, demanding enhanced remote capabilities. However, remote Human-Machine Interfaces (HMIs) often need to provide more representations of the operating environment, hindering the detection of visual anomalies and the execution of complex maintenance tasks.

This Chapter compares and evaluates the impact of different maintenance interface modalities, specifically between a non-immersive (2D interface) and an immersive (VR interface) approach, developed to aid remote visual inspection. It aims to determine whether the interface modality influences performance metrics and human factors such as presence, workload, and usability. These interfaces will be assessed through mission-like tasks, gamification, and real-time feedback from a physical robotic arm in a hybrid simulated setting.

The primary aim behind this experiment is manifold, stretching from the evaluation of interface modalities to validating and enhancing system architectures. Specifically, this study aims to evaluate the effectiveness of VR interfaces by doing a comparative analysis with a 2D interface. User experience metrics assess nuances in user experience between the two modalities, focusing on human factors and performance metrics.

5.2 Tasks and interfaces design

5.2.1 Gamified task creation

Gamification, or the application of game design elements and principles in non-gaming contexts, has emerged as a promising approach to enhance engagement, motivation, and learning outcomes in various domains (Treiblmaier, Putz and Lowry, 2018). Gamification elements include points, challenges, leaderboards, and performance feedback that can be incorporated to encourage healthy competition and promote a sense of achievement (Toda et al., 2019). In the study, a mission was designed to require participants to take four inspection photos of various engine parts, as illustrated Figure 5-1. While applying gamification to maintenance tasks might seem non-traditional, its utility is grounded in extensive research on human behaviour and motivation (Toda et al., 2019). Gamified experiences increase engagement, heighten attention to detail, and elevate the user experience (Hicks et al., 2019). The integration of game elements into maintenance tasks aligns with the objectives of this study, which seek to provide effective teleexistence capabilities. The mission-like design, which tasks participants with capturing photos at specific distances and comparing views, stimulates active participation and emulates real-world scenarios where meticulous observation and documentation are essential. By contrasting and comparing participant interactions and outcomes between VR and 2D interface tasks, inferences can be drawn about each interface's efficacy and user experience. The gamified approach is a methodological tool, enabling us to assess the interface's performance in a context that mirrors the attentiveness and precision required in real-world maintenance operations.

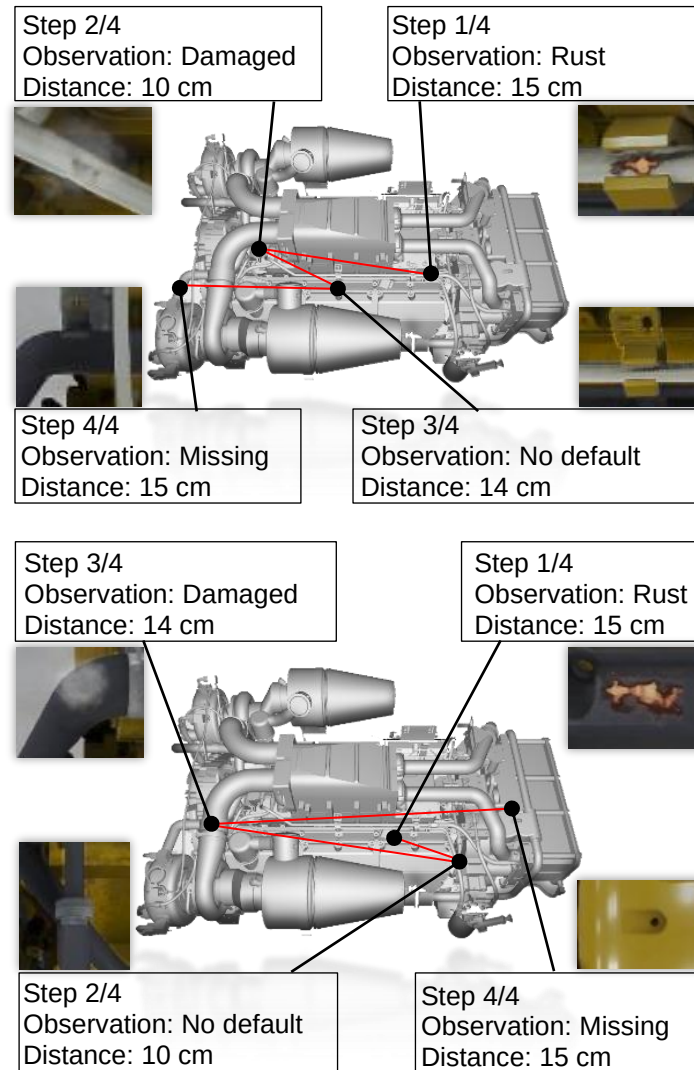
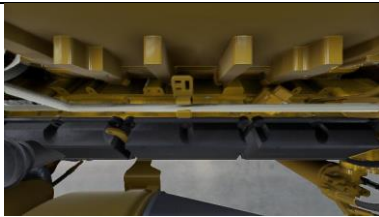
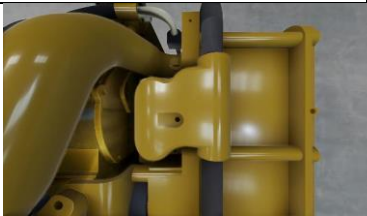


Figure 5-1 Description of the two designed scenarios for each Interface

The experiment required participants to complete a mission-like task involving inspecting four distinct engine parts, demanding careful navigation, identification, and documentation in VR and 2D settings. Participants navigated a simulated environment, mimicking a maintenance scenario in an unmanned vessel. The process involved identifying specific engine parts for inspection, manoeuvring a robot to these locations, and documenting the conditions observed. Any defects, such as missing or damaged parts or signs of wear like rust marks, were to be reported. This process was replicated across both interface modalities to ensure a comprehensive comparison. Each photo must be captured between the camera and the engine at a specific distance. Before taking the photo, participants had to

compare the camera view simulating real space with the engine's digital mock-up and indicate their observations. These observations can be (1) no defaults to report, (2) missing parts on the engine, (3) damaged parts on the engine, or (4) rust marks. This approach introduces an element of gamification to the user study by engaging participants in a mission-like task. The goal of capturing photos at specific distances and comparing views adds a challenge that encourages active participation and promotes a sense of accomplishment when completed. Table 5-1 presents a side-by-side comparison of the inspection photos captured by participants using both the VR and 2D interfaces.

Table 5-1 Example of participant photos taken during VR and 2D Interface tasks

VR Interface photos		2D Interface photos	
Step 1 Rust		Step 1 Rust	
Step 2 Nothing Wrong		Step 2 Damaged Element	
Step 3 Damaged Element		Step 3 Nothing Wrong	
Step 4 Missing Element		Step 4 Missing element	

The eight different areas for inspection were divided into two scenarios, each containing four areas to inspect. One scenario was designated for the 2D

interface and the other for the VR interface. The total distance between the four zones within each scenario was the same. Participants needed to visit four zones sequentially with the 2D and VR scenarios. This approach maintains a balanced comparison between the two interface modalities in the study. Participants were given the following instructions: first, they must identify the area to be inspected; second, they must manipulate the robot and navigate through the scene; and third, they must report any defects or issues found during the inspection.

5.2.2 Real-time feedback and kinematic control of collaborative robots

A gateway program, developed in C# and integrating the RTDE (Real-Time Data Exchange) interface, provides the data for the virtual representation of the robot joint states for the marine engineer performing the remote inspection.

Real-time feedback on the robot's range state was provided to the participant as a helpful tool. When the robot was close to being out of its operating range, a message was displayed on the screen for the 2D interface, or the colour of the robot shadow model was modified for the VR interface. The same occurred when the robot was out of range, with a different message and colour to alert the participant that the robot could not reach the area and was no longer moving for safety reasons. Figure 5-2 shows a flowchart of the proposed process part on real-time feedback.

The framework uses the Denavit-Hartenberg (D-H) method to calculate the inverse kinematics solutions directly inside Unity 3D. The D-H representation specifies parameters: the angle between two links, the offset between links, the

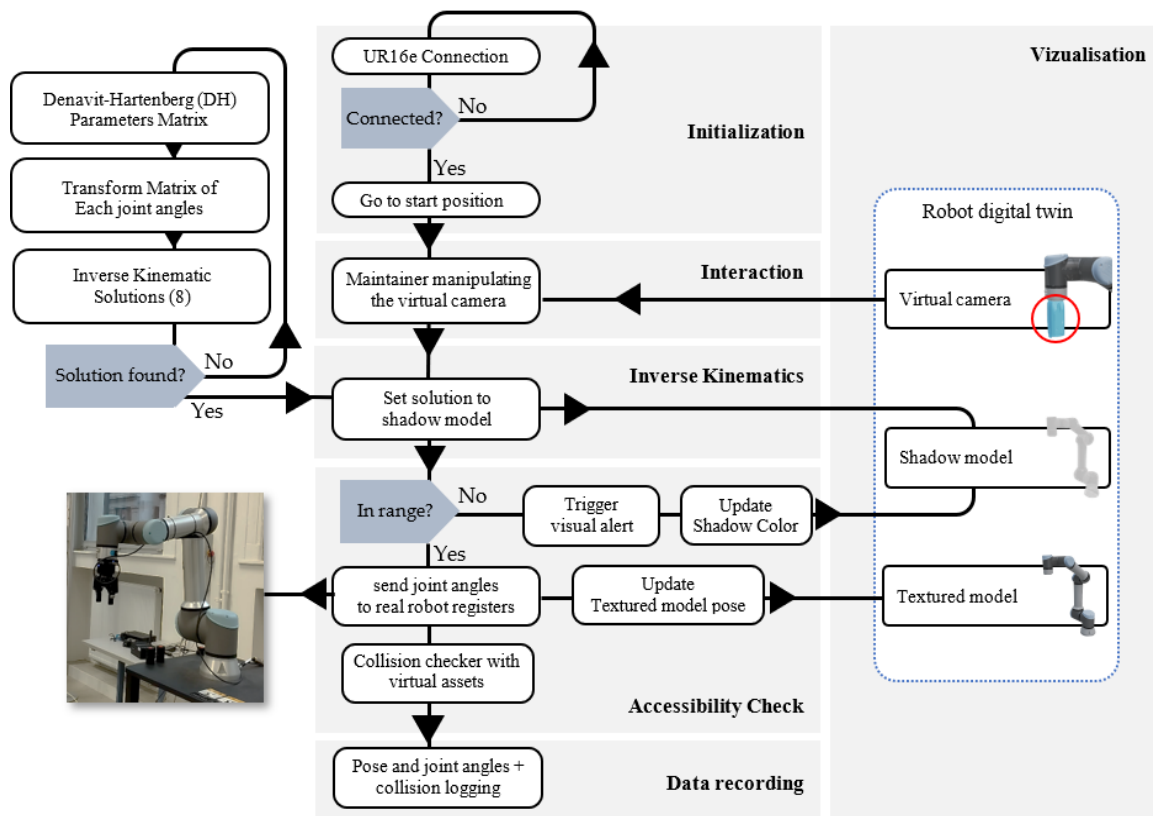


Figure 5-2 Flowchart of the data for real-time feedback in the proposed process

length of the link along the common normal and the twist angle between axes. A video presenting the result is available⁷.

D-H parameters are commonly used in robotics to describe the geometry of a robot arm in a standard way. This parametrisation simplifies the kinematic and dynamic analysis process, allowing for calculations related to positioning, velocity, acceleration, and the forces involved (Universal Robots, 2022). The D-H parameters of the UR16e used during the experiment are shown below in Table 5-2.

⁷https://cord.cranfield.ac.uk/articles/media/Immersive_robotics_experiments_-_Telexistence_PhD_videos_/24190620?file=42445554

Table 5-2 D-H parameters for the UR16e robot (Universal Robots, 2022)

UR16e					
Kinematics	Theta [rad]	Alpha [m]	d [m]	a [rad]	Dynamics
Joint 1	0	0	0.1807	$\pi/2$	Link 1
Joint 2	0	-0.4784	0	0	Link 2
Joint 3	0	-0.36	0	0	Link 3
Joint 4	0	0	0.17415	$\pi/2$	Link 4
Joint 5	0	0	0.11985	$-\pi/2$	Link 5
Joint 6	0	0	0.11655	0	Link 6

- Theta is the rotation angle along the common normal (z-axis).
- Alpha is the angle of twist (around the x-axis).
- d is the distance along the previous z to the common normal.
- a is the length of the common normal.

To perform the Inverse Kinematics (IK) in Figure 5-3, calculations are necessary for articulating the UR16e in response to a given target position and orientation in Unity 3D.

The initiation point of this loop is forward kinematics. This computational technique translates joint parameters, such as joint angles and link lengths, into the corresponding position and orientation of the robot's end-effector in Cartesian coordinates. This predictive process provides critical insights into the spatial

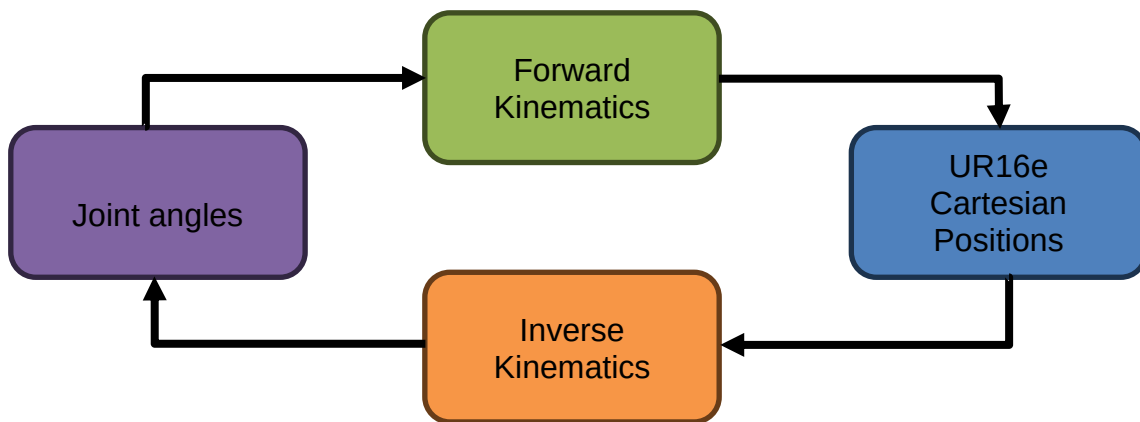


Figure 5-3 Forward and Inverse kinematics representation

position that the robot's end-effector would occupy based on the current configuration of its joints.

The outcome of the inverse kinematics process is a distinct set of joint angles for the UR16e. These angles represent the exact positioning of each joint within the robotic manipulator to ensure that its end-effector reaches the desired Cartesian coordinates. This process, forming a continuous cycle between forward and inverse kinematics, is a crucial mechanism in the operation and control of robotic manipulators. The author has integrated the process with Unity 3D without any external library. The source code available on GitHub is explained in 8.3.3Appendix D.

5.2.3 Interfaces designs

5.2.3.1 2D interface

Requirements:

- **Data Acquisition and Display:** The interface must acquire and display real-time data from multiple sensors and cameras. The videos should not require mental rotation; the same camera image should be presented on one specified panel.
- **Remote Control and Navigation:** Enable remote operation of a robotic arm to inspect engine parts, including manoeuvrability, in all necessary directions to capture inspection photos from specified distances.
- **Real-Time Feedback:** Provide instant feedback on the robotic arm's range state, alerting the operator when the arm is close to or out of its operational range.
- **Anomaly Detection:** This feature supports identifying and reporting anomalies such as missing parts, damaged components, or signs of wear, such as rust marks.

The 2D interface comprises four distinct panels:

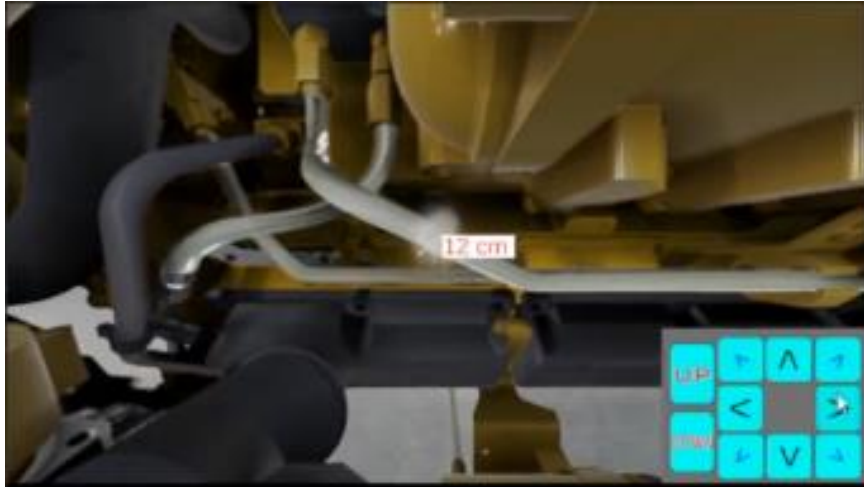


Figure 5-4 Top-left panel

- The top-left panel (Figure 5-4) displays the end effector inspection camera view. It provides a direct feed from the end effector's inspection camera, offering a close-up view of the engine parts under inspection. On this panel are buttons to zoom and move the camera to closely examine areas of interest. The distance between the middle camera and the environment is also displayed in cm.

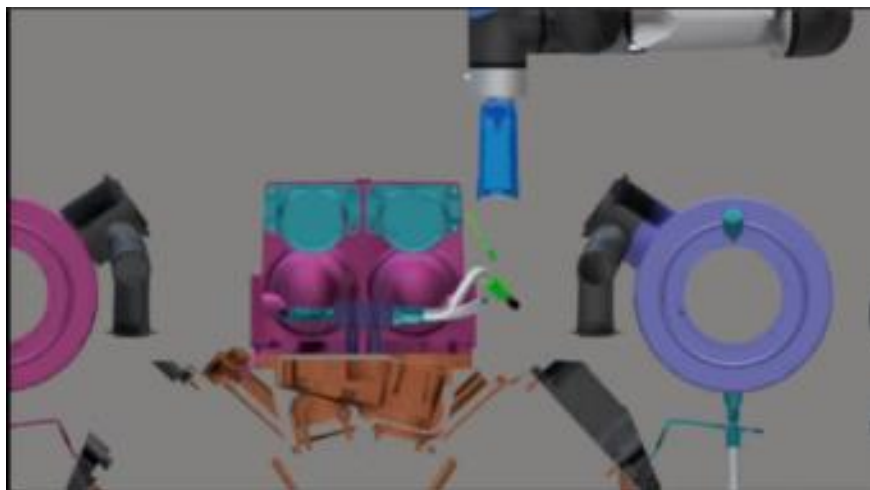


Figure 5-5 Top-right panel

- The top-right panel (Figure 5-5) displays a sectional view from the camera's position, allowing the operator to understand the spatial

relationship between the camera and the engine parts. There are no interactive elements in this panel.



Figure 5-6 Bottom-right panel

- The bottom-right (Figure 5-6) panel features a top orthographic view of the vessel's engine and surrounding area, providing context and aiding navigation. Two buttons allow the user to translate the robotic arm left and right.

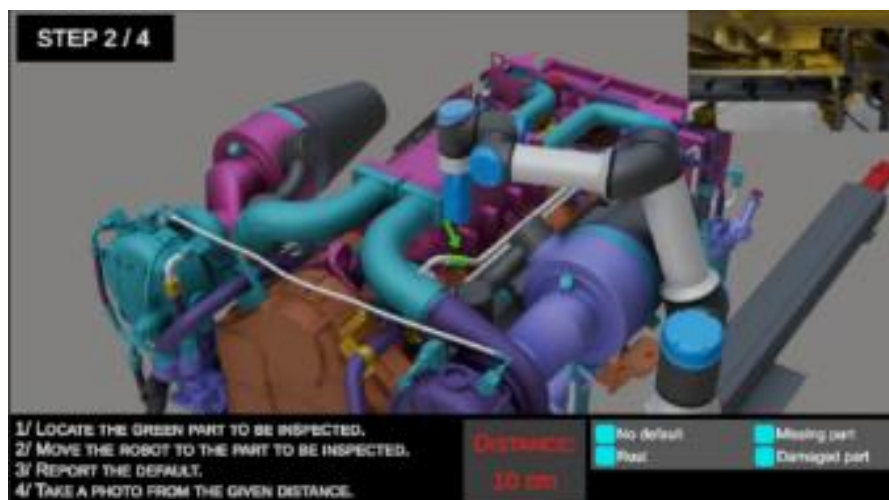


Figure 5-7 Bottom-left panel

- The bottom-left (Figure 5-7) panel integrates a 3D model of the engine and inspection area with interactive inspection instructions and reporting mechanisms.

The design choices were made for the following reasons. The 2D interface offers comprehensive spatial awareness by providing multiple views, allowing the operator to understand the object's context without changing the view frequently (Wickens et al., 2015). When transitioning from one camera view to another, operators need to reassess the robot's positioning within the new view, often causing a pause in operations (Kamezaki et al., 2016). Designing seamless video transitions poses a challenge, as it requires a deep understanding of both the task and the operator's needs. To address this, our approach entirely avoids the need for video switching. By maintaining a consistent arrangement of cameras, their functions, and the perspectives they offer, our system ensures clarity and quick comprehension of the surrounding environment.

Mouse and keyboard interaction is widely used and intuitive for most users, minimising the learning curve. Shneiderman et al. discussed the advantages of the mouse as a pointing device and its widespread adoption due to its intuitive nature and ease of use (Shneiderman et al., 1997). Given the interface's stationary nature and longer interaction durations, a seated position minimises user fatigue.

5.2.3.2 VR interface

Requirements:

- Immersive 3D Environment Creation: Ability to generate a realistic and detailed 3D virtual environment that mimics the real-world setting of the unmanned surface vessel's engine and surrounding areas for inspection.
- Real-time Interaction and Control: Support for real-time manipulation of virtual objects, including a digital representation of the robot, using VR controllers to emulate hands-on interaction.
- Dynamic Robot Operation Simulation: Simulate the operation of the robotic arm for inspection purposes, including movement, orientation, and interaction with engine parts, with feedback on accuracy and effectiveness.

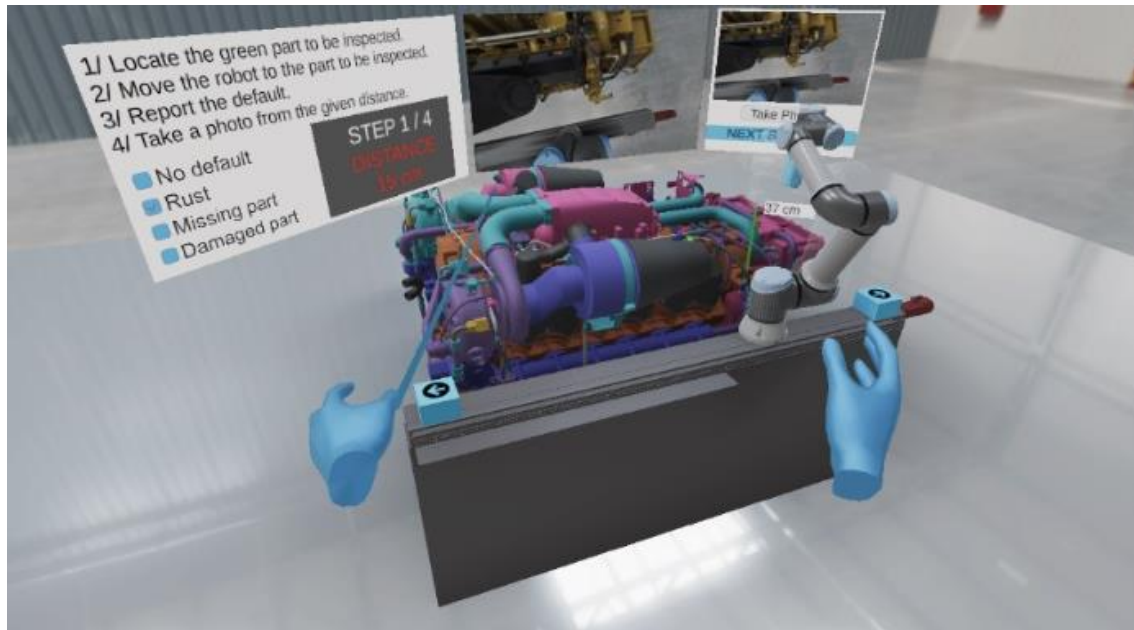


Figure 5-8 Screenshot of the VR Interface

The VR interface Figure 5-8 consists of a virtual environment featuring the digital mock-up of the engine in the middle of the scene and the digital representation of the robot. This setup is designed to immerse the operator fully in the maintenance task, simulating a physical presence alongside the engine and robot.

Three panels are situated in front of the engine:

- **Left Panel:** This panel serves as an information dashboard, providing the operator with instructions and reporting tools. Its location ensures that the operator can quickly glance and gather required information without straying too far from the primary view of the engine. The proximity of the information dashboard to the engine model allows the operator to seamlessly transition between reviewing instructions and engaging with the engine or robot, minimising cognitive load, and enhancing task efficiency.
- **Central Panel:** Acting as the eyes of the operation, this panel delivers a simulated real-space camera view. This means the operator can view the engine's real-time state and the robot's actions in the same visual plane, enhancing accuracy and coordination. This direct view fosters a deeper

connection between the operator and the task, improving spatial awareness and coordination. The central positioning ensures it is the focal point of the experience, prioritising the inspection task's visual aspects.

- **Right Panel:** This provides additional reporting instruments as a supplementary tool. Its positioning ensures it is within easy reach but does not overshadow the central view. Its placement to the side ensures that it supports, rather than interrupts, the immersive experience. Operators can quickly access these tools when needed, ensuring a streamlined workflow.

The VR interface is designed to enhance user engagement and efficiency by leveraging immersive technology to simulate a physical presence and intuitive interaction within a maintenance task environment. This design choice is justified by the potential for increased accuracy, faster learning curves, and overall improved user satisfaction when performing complex tasks such as remote visual inspection and maintenance of unmanned surface vessels. Below is a description of the VR vs. 2D interface parameters difference:

- *Input System & Participant Position:* VR allows for a more natural interaction through VR controllers, mimicking real-life movements (e.g., grabbing, clicking) in a standing position, which can enhance the sense of presence and immersion compared to the seated, mouse-based interactions of the 2D interface.
- *Interaction Method & Area Identification:* Using virtual hands and a laser pointer in VR versus cursors in 2D creates a more intuitive and engaging interaction model. Areas for inspection are indicated similarly in both interfaces but are more immersive in VR, enhancing the operator's ability to identify and focus on specific tasks.
- *Scene Navigation & Robot Operation:* VR excels with its free walk-around capability, offering a significant advantage over the 2D interface's fixed view. This freedom enhances the operator's spatial understanding and ability to inspect from multiple angles. Robot operation in VR mimics real-life control, providing a more intuitive experience than the abstracted camera controls in 2D.

- *Camera Distance & Default Reporting:* The VR interface positions the camera view next to the robot's end effector and follows the user's head orientation, creating a more coherent and integrated viewing experience that closely aligns with the operator's natural sightline. Reporting mechanisms in VR use a laser pointer for selection, which aligns more with the immersive experience than the mouse clicks required in 2D.
- *Camera Distance & Default Reporting:* The VR interface positions the camera view next to the robot's end effector and follows the user's head orientation, creating a more coherent and integrated viewing experience that closely aligns with the operator's natural sightline. Reporting mechanisms in VR use a laser pointer for selection, which aligns more with the immersive experience than the mouse clicks required in 2D.

5.3 Hypotheses

This Chapter attempts to evaluate several hypotheses and the associated null hypotheses:

H1: The sense of presence experienced by users during remote inspection tasks is higher when using an immersive VR interface than a traditional 2D interface.

H01: There is no difference in the sense of presence experienced by users when using an immersive VR interface versus a traditional 2D interface for remote visual inspection tasks.

H2: Users' workload is lower when performing remote visual inspection tasks via an immersive VR interface compared to a traditional 2D interface.

H02: There is no difference in users' workload when performing remote visual inspection tasks via an immersive VR interface compared to a traditional 2D interface.

H3: The VR interface will show higher usability scores than the traditional 2D interface for remote visual inspection tasks.

H03: There is no difference in usability scores between the VR and traditional 2D interfaces for remote visual inspection tasks.

H4: Task completion time and accuracy are improved using an immersive VR interface compared to a traditional 2D interface.

H04: Task completion time and accuracy are not different between an immersive VR interface and a traditional 2D interface.

H5: Users with prior VR experience exhibit better performance outcomes when using an immersive VR interface for remote visual inspection tasks than users without such experience.

H05: There is no difference in performance outcomes between users with prior VR experience and those without when using an immersive VR interface for remote visual inspection tasks.

5.4 Results

Participants were split into two groups, one starting with the VR interface and the other with the 2D interface. The experiment duration ranged from a minimum of 24 minutes to a maximum of 50 minutes. Nevertheless, on average, it took approximately **41 minutes** per participant.

The following three sections describe the subjective data analysis regarding the presence, cognitive workload, and usability.

5.4.1 Presence

As recommended by iGroup, the items were transformed on a seven-point scale (0-6), with the higher mean scores indicating a greater sense of presence. The results of the descriptive statistics calculated in Table 5-3 provide helpful information about the subjective experiences reported by the study participants in both 2D and VR conditions.

Table 5-3 Descriptive statistics of IPQ questionnaire results

Factors	Interface Modality	Mean Score	Min value	Max value
General Presence (GP)	2D	2.21 ± 1.72	0.00	5.00
	VR	5.71 ± 0.55	4.00	6.00
Spatial Presence (SP)	2D	2.68 ± 0.63	1.80	3.60
	VR	3.73 ± 0.82	2.40	4.80
Involvement (INV)	2D	2.33 ± 0.87	0.75	3.75
	VR	3.71 ± 0.71	3.00	4.50
Experienced Realism (REAL)	2D	2.4 ± 0.69	1.00	3.50
	VR	3.13 ± 0.73	2.00	4.25

Overall, the VR condition scored higher on every factor than the 2D condition. This could suggest that participants experienced a greater sense of presence, spatial presence, and involvement and experienced realism in the VR condition

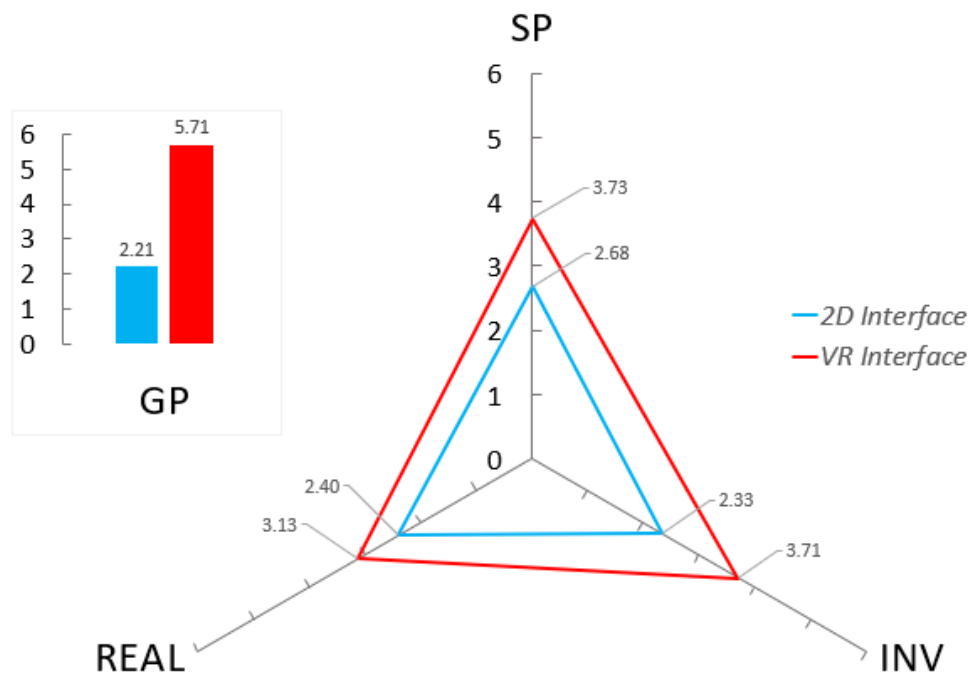


Figure 5-9 Presence profile of the 2D and VR interface regardless of the order effect

compared to the 2D condition. The General Presence score for VR is higher than the 2D counterpart, and the standard deviation in VR is also relatively low. This suggests that participants consistently experienced a strong sense of presence in VR. Figure 5-9 illustrates the presence profile of the 2D and VR interfaces regardless of the order effect.

5.4.1.1 Orders Effect on Presence

Figure 5-10 compares IPQ factor scores across two distinct interface modalities: 2D and VR. In the first order (Order 1), participants experienced the 2D interface before the VR modality, while in the second order (Order 2), the VR experience preceded the 2D interface. The bars represent the average scores, with error bars indicating the standard deviation.

The following observations are made:

- **Standard deviation:** The standard deviation bars for the 2D interface Modality are larger overall than those for the VR Interface Modality. This suggests more variability in the responses for the 2D interface in General Presence and Involvement factors. In the 2D and VR interface modalities,

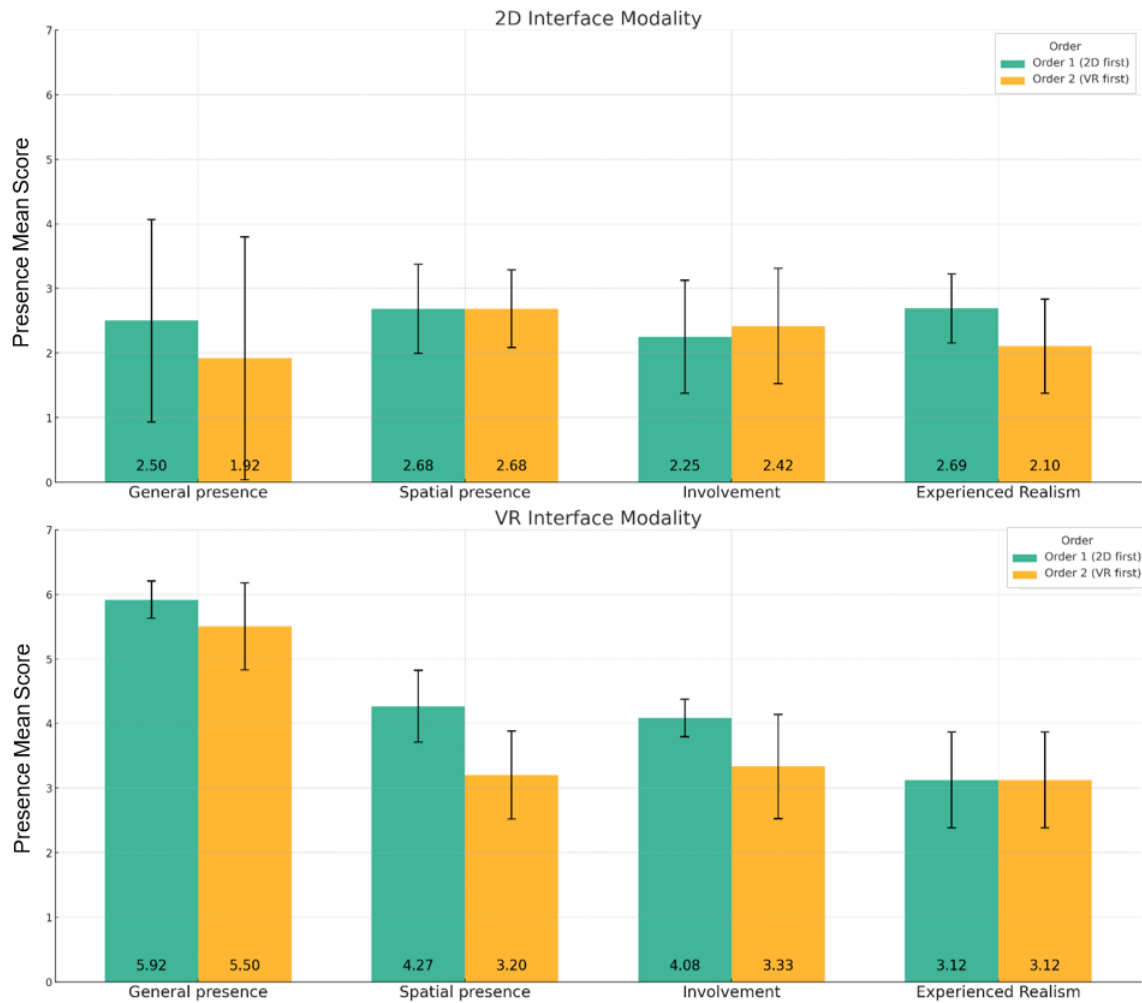


Figure 5-10 Comparison of IPQ factors across 2D and VR modalities based on testing order

the standard deviation appears similar across Spatial Presence and Experienced Realism factors, indicating that the spread of the scores around the mean is consistent.

- **Higher general presence in VR:** Regardless of the testing order, participants consistently felt a stronger sense of presence in the VR modality compared to the 2D interface. This could be due to VR's immersive nature, which may create a more engaging and lifelike experience.
- **Testing order matters:** The order in which 2D and VR were evaluated affected the scores, though the impact varied. For instance, in the 2D

modality, General Presence, Involvement, and Experienced Realism scores were higher when 2D was evaluated after VR.

- **Consistent spatial presence in 2D:** Interestingly, the Spatial Presence scores for the 2D modality remained consistent regardless of the testing order. This might suggest that the spatial aspects of the 2D interface are not influenced by prior VR exposure.
- **Reduced spatial presence in VR when evaluated first:** The VR modality witnessed a notable decrease in Spatial Presence when the first interface was evaluated. This could indicate that experiencing a 2D interface first might set certain spatial expectations, which are then exceeded in the VR experience.
- **Involvement in VR:** Participants felt more involved in the VR modality, especially when VR was tested after the 2D interface (Figure 5-10, VR interface modality – order 1).
- **Experienced realism is comparable across orders:** For both 2D and VR modalities, the Experienced Realism scores were consistent irrespective of the testing order.

To understand the influence of interface modality, a Two-way ANOVA was conducted, considering both the type of modality (2D vs. VR) and the order of exposure. Preliminary tests, including Levene's test, confirmed that the data met the assumptions for ANOVA. There was a statistical main effect of modality on the sense of presence $F(1,44) = 84.86, p < 0.001$, indicating a pronounced difference in the sense of presence between the 2D and VR modalities, also affected by the order in which participants were exposed to the modalities. $F(1,44) = 7.85, p = 0.0075$.

An interaction was observed between modality and order, $F(1,44) = 8.15, p = 0.0065$, suggesting that the influence of one factor was contingent on the level of the other. The findings support the hypothesis that an immersive VR interface elevates the sense of presence in remote inspection tasks compared to a traditional 2D interface. Moreover, the sequence of exposure to these modalities plays a pivotal role in shaping this experience.

5.4.2 Cognitive workload

Figure 5-11 shows the results of the NASA-TLX questionnaire for the subjective evaluation of the cognitive workload perceived by the participants during the experiment. It provides three dimensions related to the demands imposed on the participant (Mental, Physical, and Temporal demands) and three to the participant's interaction toward the task (Performance, Effort, and Frustration) on a 21-gradation scale.

Both modalities displayed similar mental demand scores, with the 2D interface registering a mean of approximately **11.33 ± 4.11** compared to VR's mean of **10.25 ± 4.10**. This similarity in mental demand scores could be attributed to several factors. For instance, the tasks performed using both interfaces might have required similar levels of cognitive processing. Also, despite VR being typically more immersive, the specific design of the 2D interface might have been engaging enough to require a similar level of mental investment from the users. The quality of immersion and engagement can impact mental workload.

However, the physical demand revealed a pronounced difference: the VR interface had a mean score of **7.50 ± 5.22**, higher than the 2D interface's mean of **2.38 ± 1.50**. This could be attributed to the inherent nature of VR technology, which often involves more physical interaction, like head and hand movements,

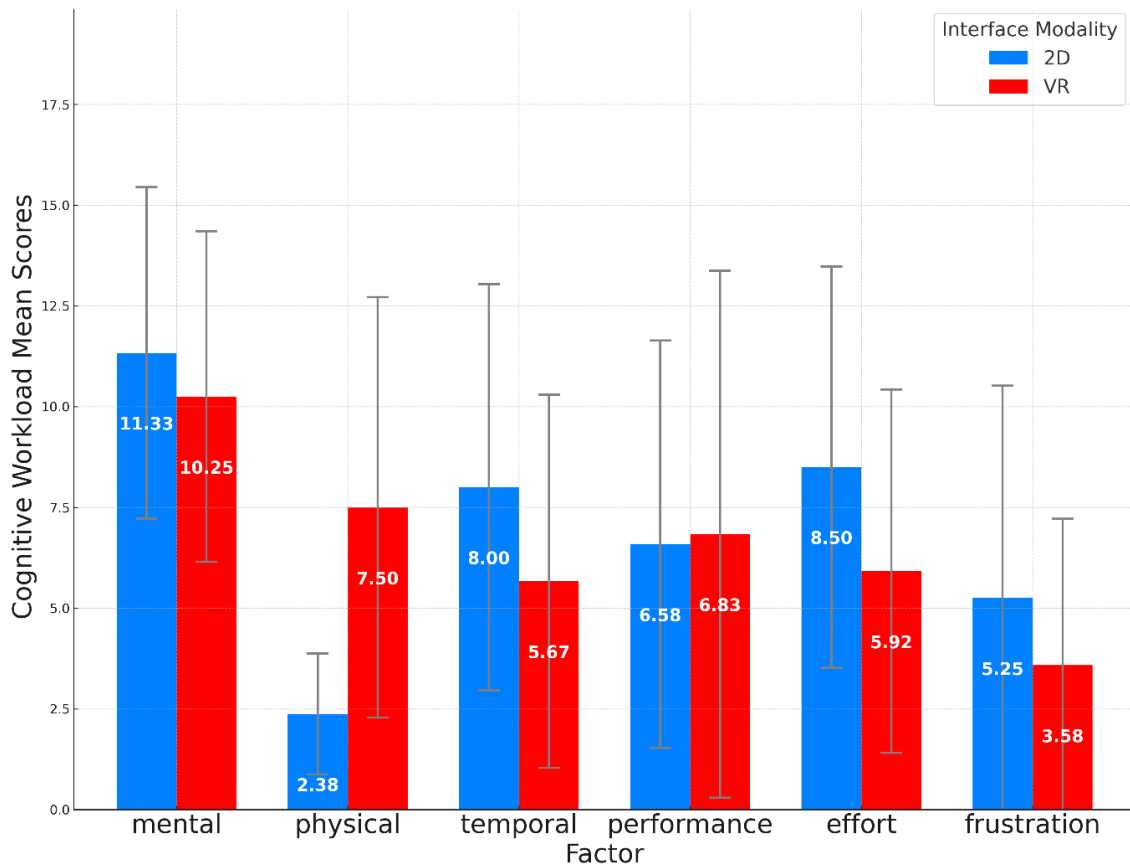


Figure 5-11 Comparison of cognitive workload between 2D and VR interfaces

to navigate and interact within the virtual environment. In contrast, 2D interfaces typically require less bodily movement, primarily involving hand-eye coordination with a mouse or keyboard.

Temporal demand was also slightly elevated in the 2D interface with a mean of **8.00 ± 5.04** compared to VR's **5.67 ± 4.63**. This might suggest that tasks in the 2D environment require more time-sensitive responses or quicker decision-making. By offering more immersive and intuitive experiences, VR environments might have allowed users to operate more relaxedly.

Performance perceptions were comparable across both interfaces, with VR at **6.83 ± 6.54** and 2D at **6.58 ± 5.06**. This parity might indicate that users felt similarly effective in accomplishing tasks in both interfaces despite the differences in physical and temporal demands.

The effort scores indicated that participants found the 2D interface more demanding, with a mean score of **8.50 ± 4.98**, as opposed to VR's **5.92 ± 4.51**.

This could suggest that, despite the higher physical demand for VR, users might have found the 2D interface more mentally or conceptually challenging, requiring more cognitive effort to achieve the same results.

Lastly, frustration levels were slightly higher in the 2D interface, with a mean of **5.25 ± 5.27**, compared to VRs of **3.58 ± 3.63**. This could be related to the higher effort and temporal demand in the 2D interface, potentially leading to more user frustration.

Overall, the findings highlight the distinct cognitive challenges associated with each interface, emphasising the pronounced physical demand of VR and the slightly elevated mental and temporal demands of the 2D interface.

5.4.2.1 Orders effect on cognitive workload

Figure 5-12 presents the perceived cognitive workload of participants when navigating two distinct interface modalities, 2D and VR, emphasising the influence of testing order. A striking observation emerges when considering the physical demand within the VR environment for those introduced to the 2D interface first (Order 1). Here, the mean score is **9.50 ± 6.67**, suggesting a substantial physical strain. This pronounced demand could imply that an initial 2D interface experience amplifies the recognition of physical exertion in the subsequent immersive VR context. On the other side, starting with the VR interface (Order 2) curtails the physical demand score for VR to **5.50 ± 1.93**. However, this order accentuates the effort and frustration scores for the 2D interface to **10.83 ± 3.86** and **7.17 ± 5.62**, respectively. These changes could indicate that the VR interface defines a distinct set of expectations that intensify cognitive effort and potential frustration when juxtaposed with the 2D experience.

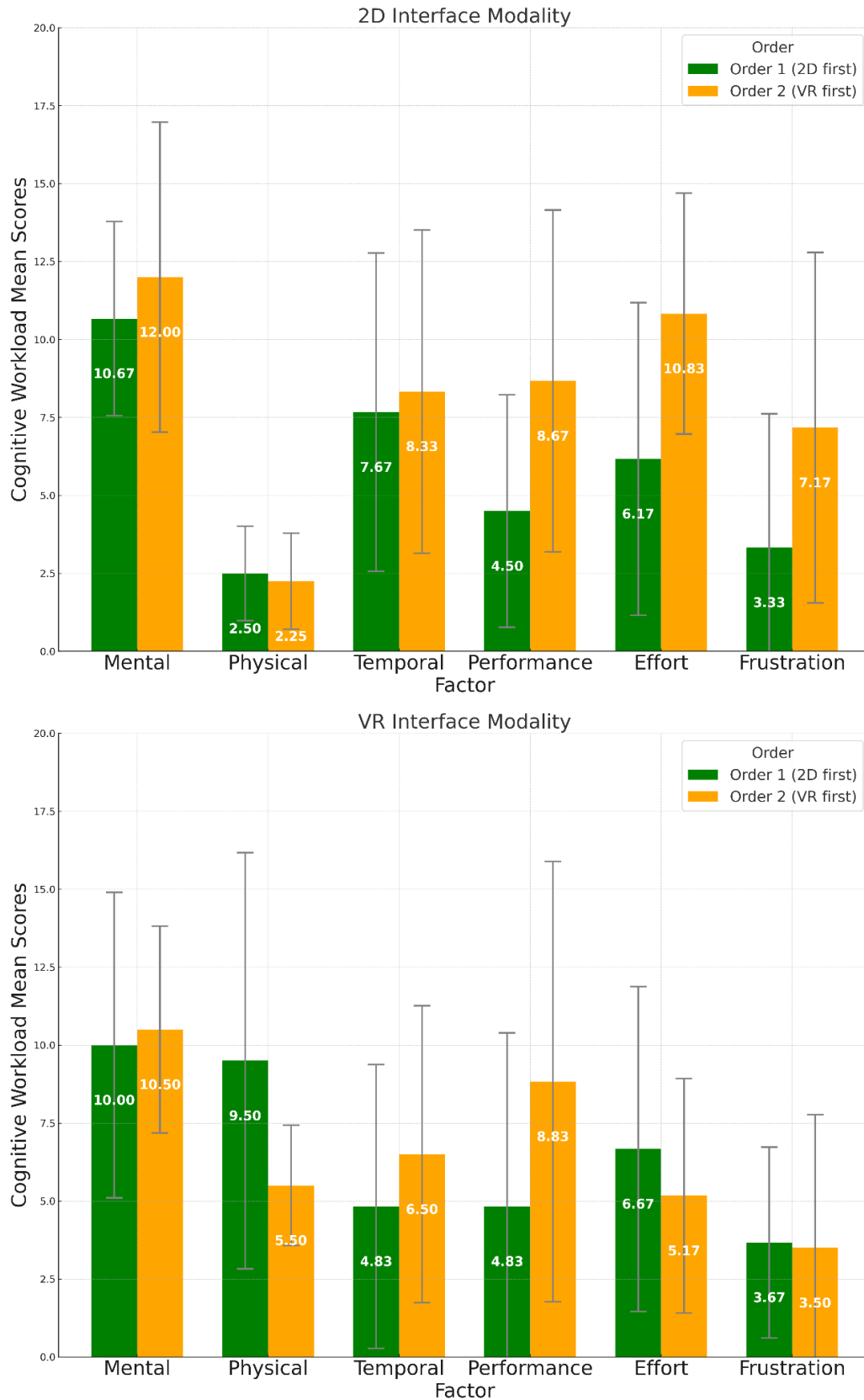


Figure 5-12 Impact of testing order on cognitive workload assessment in 2D and VR

Given that the dataset involves repeated measurement, where the same participants interacted with both the 2D and VR interfaces, a paired t-test was employed for the analysis. A paired t-test can determine if a difference exists between two related groups. For the study, these groups represent the distinct orders in which participants encountered the interfaces. The sequence of exposure can potentially shape a participant's perceptions and expectations. Therefore, by juxtaposing scores from the two orders across each NASA-TLX dimension, the aim was to identify any influence the exposure sequence might have on participants' perceptions of cognitive workload.

Table 5-4 Paired T-test Results Comparing Workload Perceptions Based on Testing Order for 2D and VR Interfaces

Dimension	t-statistic (2D)	p-value (2D)	t-statistic (VR)	p-value (VR)
Mental	-0.771	0.457	-0.561	0.586
Physical	0.447	0.643	1.985	0.073
Temporal	-0.274	0.789	-0.718	0.488
Performance	-3.837	0.00276	-1.515	0.158
Effort	-2.938	0.0132	0.861	0.408
Frustration	-1.829	0.0946	0.140	0.891

The table presents the paired t-test that compares participants' cognitive workload perceptions between the two testing orders: Order 1 (2D first, then VR) and Order 2 (VR first, then 2D).

For the 2D interface:

- The Performance dimension produced results of $t(11) = -3.837, p = 0.00276$, indicating a difference based on the testing order.
- Similarly, the Effort dimension returned $t(11) = -2.948, p = 0.0132$, highlighting another variance based on exposure order.

- These values suggest that when introduced to the VR interface first, participants had altered perceptions of their performance and effort when navigating the 2D interface.

For the VR interface:

- Although no dimension showcased a clear-cut difference equal to the $\alpha = 0.05$ threshold, the Physical Demand dimension is worth noting. It edges to with $t(11) = 2.9645, p = 0.0129$. This result suggests an emerging trend that previous interactions with the 2D interface may influence the perception of physical exertion in VR.

The order in which participants are introduced to interfaces shapes their perception of cognitive workload. This effect is particularly pronounced in the context of the 2D environment.

This effect is known as an order effect and is a common phenomenon in experimental design, where the experience of one condition can influence the response to subsequent conditions (Strack, 1992). Participants may become more familiar with the type of tasks or the experimental setup after using the first interface. This familiarity might reduce the perceived workload when they use the second interface because they have adapted to the environment and the task requirements. Experiencing one interface first sets a baseline of expectations and experiences against which the second interface is judged. If the first interface is less immersive and requires less physical interaction (like a 2D interface), the physical demands of a VR interface seem more pronounced when used second. Skills or strategies learned in the first condition could transfer to the second, making the tasks seem more accessible or manageable, thus reducing perceived workload. Alternatively, if the first interface is more complex or taxing, it may make the second interface feel more difficult due to fatigue or cognitive depletion. The first interface might create an 'anchor' for expectations. If the VR interface is used first, its immersive nature might make the 2D interface seem less engaging, negatively affecting the participants' perceived performance and effort when they switch to the less immersive interface.

5.4.3 Usability

The SUS is a simple, ten-item scale giving a global view of subjective usability assessments and expectations from the system. It was developed by John Brooke at Digital Equipment Corporation in the UK in 1986. The SUS consists of 10 questions from 0 to 5 (being the highest), the first, third, fifth, seventh, and ninth of which are positively worded and the second, fourth, sixth, eighth, and tenth of which are negatively worded. A score of less than 60 indicates poor usability, while an average score of more than 80 indicates good usability.

On average, as shown in Figure 5-13, participants rated the VR interface with a score of **79.27 ± 12.74**, approaching the SUS benchmark of 80 for “relatively good usability.” This suggests that most participants perceived the VR interface as highly usable and intuitive. The standard deviation suggests that participant’s ratings were consistent, clustering around a high usability score. The 2D interface achieved a mean score of **69.69 ± 19.24**. The 2D interface has a more significant standard deviation and suggests a broader range of experiences among participants, indicating less agreement on usability. While the mean score is above the SUS midpoint of 50, indicating a positive user experience, it is noteworthy that it falls below the 80 benchmarks.

5.4.3.1 Orders Effect on Usability

The sequence in which participants encountered the two modalities, 2D and VR, seems to have imparted a nuanced effect on their usability perceptions (Figure 5-13). Participants who first interacted with the 2D interface before transitioning to the VR modality (Order 1) registered a high mean score of **78.13 ± 19.24** for the 2D interface. In contrast, those who started with the VR interface and subsequently evaluated the 2D interface (Order 2) gave a noticeably lower score of **61.25 ± 19.24** for the 2D modality. This differential suggests that initial exposure to the immersive and novel VR experience might set a heightened usability expectation, making the subsequent 2D experience feel less intuitive or engaging.

The order effect was less pronounced for the VR modality, with scores of **79.79 ± 12.74** for Order 1 and **78.75 ± 12.74** for Order 2. This relative consistency

implies that the VR experience's perceived usability remained stable, irrespective of prior exposure to the 2D interface.

For participants who first interfaced with the 2D modality and then the VR (Order 1), the paired t-test of $t(11) = 0.3363, p = 0.7430$. This indicates that the difference in usability scores between the 2D and VR modalities was not statistically significant when 2D was evaluated first.

Conversely, for participants who began their experience with the VR interface followed by the 2D modality (Order 2), the t-test returned $t(11) = 2.9645, p = 0.0129$. This p-value, being less than the conventional threshold of 0.05, suggests a statistical difference in usability scores between the two modalities under this sequence.

The implications of these findings are multifaceted. Initial exposure to the VR interface can set heightened usability expectations, potentially influencing subsequent evaluations of the 2D interface. Conversely, beginning with the 2D interface, the perception of the VR modality is impacted to a different degree.

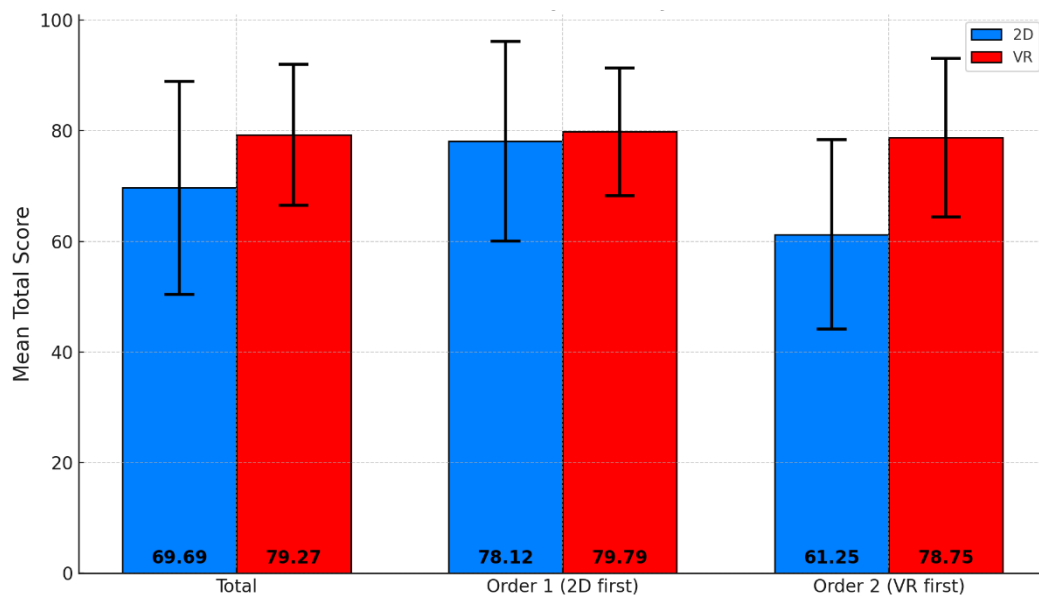


Figure 5-13 Comparative usability scores: 2D vs. VR interfaces

5.4.4 Performance

5.4.4.1 Completion Time

On average, participants took approximately **7 minutes and 29 seconds** (448.78 seconds) to complete a task using the 2D interface. In contrast, with the VR interface, the average completion time was **6 minutes and 28 seconds** (387.64 seconds). This suggests a time increase of about **16%** when using the 2D interface compared to the VR platform. Delving into individual performances revealed variability. For the 2D interface, completion times ranged from a swift **3 minutes and 44 seconds** to a lengthier **12 minutes and 56 seconds**, with a median time of **7 minutes and 51 seconds**.

In contrast, the VR interface saw participants completing the task in times ranging from **3 minutes and 41 seconds** to **10 minutes and 39 seconds**, with a median of **6 minutes and 33 seconds**. The variability in performance, as indicated by the standard deviation, was approximately **2 minutes and 27 seconds** for the 2D interface and **1 minute and 53 seconds** for the VR platform. The initial observations suggest that the VR interface might offer a more efficient user experience than its 2D counterpart. However, further statistical evaluation is essential to establish the significance of this difference.

Figure 5-14 compares completion times based on the order and interface. The boxes represent the interquartile range (IQR), capturing the middle 50% of the data. The red lines within the boxes show the median completion times. The whiskers (lines extending from the boxes) represent the range where most values fall, excluding potential outliers. Any points outside the whiskers are potential outliers. **When the 2D interface was evaluated first, there was a slight increase in completion times for the VR interface. However, the medians are close. When the VR interface was evaluated first, there was a more noticeable difference in the median completion times between the two interfaces, with the 2D interface taking longer.** The observed differences in task completion times between the 2D and VR interfaces and the influence of the exposure order can be attributed to various factors. **VR's immersive nature facilitates a more intuitive interaction and engagement, potentially leading**

to more efficient task completion. In contrast, the more conventional 2D interface might impose a more significant cognitive load due to less natural interaction modalities, mainly when used after the VR interface, thus extending task duration. Additionally, the novelty of the VR experience could enhance user motivation and concentration, contributing to faster completion times. The order effect suggests that the first interface sets a baseline for user expectations and adaptation, which carries over and influences the performance of the subsequent interface, highlighting the interplay between user interface design and human cognition.

A two-way repeated measures ANOVA was employed to study the effect of interface type and the order in which they were evaluated. The analysis yielded an F-value of **6.7417** (with a corresponding p-value of **0.0248**) for the difference between the 2D and VR interfaces, indicating a statistical variation in completion times between them. The order's standalone effect on completion times was not statistically significant, with an F-value of **0.8019** and a p-value of **0.3897**. Notably, the interaction between the interface type and the testing order emerged as significant, with an F-value of **7.4925** and a p-value of **0.0193**. This suggests that while participants took different amounts of time on the two interfaces, the

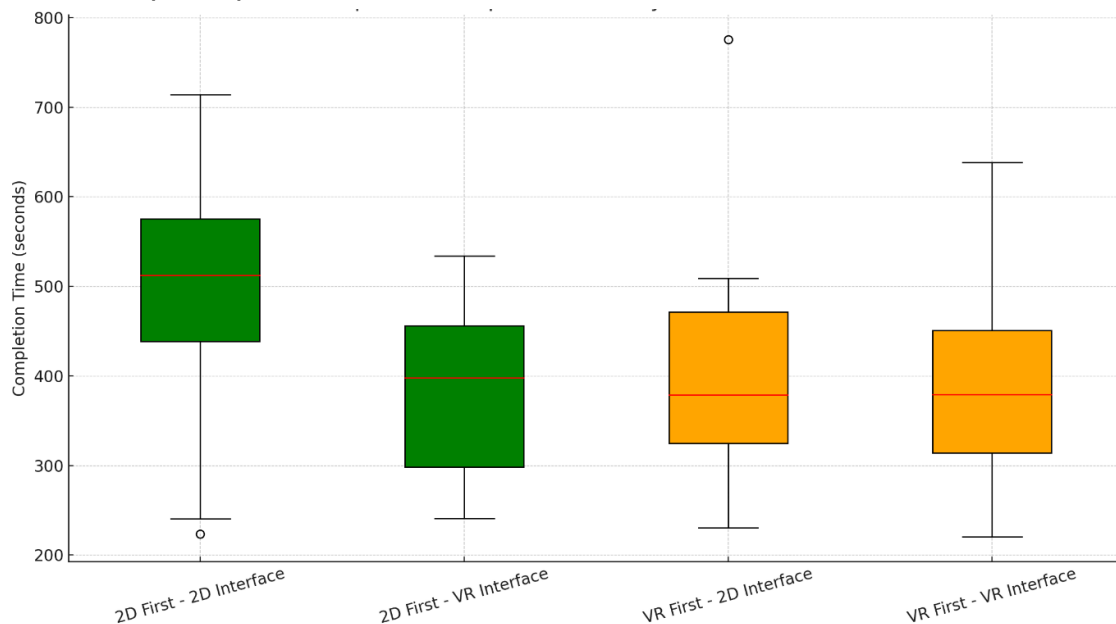


Figure 5-14 Boxplots of Completion Times by Order and Interface

magnitude of this difference was influenced by the testing sequence, precisely whether they first encountered the 2D or VR interface.

5.4.4.2 Collision Rate

The analysis of participants' performance across two distinct interfaces assessed the number of errors, specifically instances where participants collided with or hit the virtual engine. For the traditional 2D interface, participants registered an average of **2.0 ± 1.142** errors, with values ranging from a minimum of none to a maximum of 4 errors. In contrast, the VR interface saw a more proficient performance, with participants making an average of **1.125 ± 0.947** errors. The errors in the VR setting spanned from a minimum of none to a maximum of 3.

Table 5-5 provides a recap of the descriptive statistics for the number of errors made by participants, segmented by the order in which the interfaces were evaluated:

Table 5-5 Descriptive Statistics of Errors Based on Interface and Testing Order

Order	Interface	Average Errors	Standard Deviation	Median	Min Errors	Max Errors
2D First (Order 1)	2D	2.083	0.900	2.0	0	3
2D First (Order 1)	VR	1.083	1.083	1.0	0	3
VR First (Order 2)	2D	1.917	1.379	1.5	0	4
VR First (Order 2)	VR	1.167	0.835	1.0	0	2

Upon examining the effect of the order on performance, patterns become evident. For participants who interacted with the 2D interface first (Order 1), they averaged **2.083 ± 0.900** errors on the 2D interface and **1.083 ± 1.083** errors on the VR interface. In contrast, participants first introduced to the VR environment (Order 2) recorded an average of **1.917 ± 1.379** errors on the 2D platform and **1.167 ± 0.835** errors in the VR setting. This analysis suggests that, regardless of the order, participants made slightly more errors on the 2D interface compared to the VR platform. However, the difference in error rates seems influenced by the

sequence in which the interfaces were presented. A difference was observed in the number of errors between the 2D and VR interfaces ($F = 11.1517$, $p = 0.0066$). This indicates that **the type of interface used by participants influenced the errors they made**. The sequence in which the interfaces were evaluated did not impact the number of errors ($F = 0.0099$, $p = 0.9227$). This suggests **whether participants used the 2D or VR interface first did not affect their overall error rates**. There was no interaction between the interface type and testing order regarding influencing errors ($F = 0.2558$, $p = 0.6230$). This means that the effect of one factor (interface type) on errors did not depend on the level of the other factor (order).

To finish, while the choice of interface (2D vs. VR) played a notable role in the number of errors participants made, the order in which they were introduced to these interfaces did not. Moreover, the interface's effect on errors remained consistent irrespective of the testing sequence. Participants made 44% fewer errors in the VR environment compared to the traditional 2D setting, emphasising the VR interface's relative advantage in reducing errors.

5.4.4.3 Effect of VR experience of participants on performance

The analysis aimed to evaluate the hypothesis that users with prior VR experience demonstrate superior performance outcomes in immersive VR interface tasks related to remote visual inspection compared to those without such experience. To investigate this, an Ordinary Least Squares (OLS) regression was employed, analysing the completion times of tasks within a VR setting against the binary variable of VR experience (having experience vs. none). The results revealed that VR experience corresponds to a marginal decrease in task completion time by approximately **11.09 seconds**. However, this difference was not statistically significant (**p-value = 0.818**), indicating insufficient evidence to assert that prior VR experience enhances performance in the given tasks.

The data on the number of collisions suggested a trend towards improved precision among those with VR experience. While the average collision count for

participants without prior VR experience was 1.40 ± 0.97 , those with prior experience demonstrated a slightly lower average of 1.28 ± 0.89 collisions.

The statistical analysis, employing an OLS regression model, calculated a t-statistic that approached significance, with a p-value marginally above the conventional threshold (p-value = 0.057). This outcome implies a subtle, though not unequivocally significant, difference in collision rates based on prior VR exposure.

5.4.5 Participants feedback during the experiment

These comments from the participants highlight several important aspects and differences between the 2D and VR interfaces for manipulating and navigating the 3D environment:

Positive aspects of the VR interface:

1. Participants found the immersive environment of VR intriguing and enjoyable. *Mentioned by 18 out of 24 participants (75%).*
2. The 3D immersive view was reported to make it easier to see the pieces and visualise anomalies. *Mentioned by 3 out of 24 participants (12.5%).*
3. Participants felt more present in the virtual world, leading to more fluid control and quicker navigation. *Mentioned by 5 out of 24 participants (20.83%).*
4. VR was reported to be more intuitive and realistic, leading to more efficient completion of tasks. *Mentioned by 2 out of 24 participants (8.33%).*
5. Some participants appreciated the fun and engaging aspect of VR. *Mentioned by 10 out of 24 participants (41.67%).*

Negative aspects of the VR interface:

1. Some participants reported blurry visuals at times. *Mentioned by 3 out of 24 participants (12.5%).*
2. A few participants pointed out that the hand can sometimes obstruct depth perception in VR. *Mentioned by 3 out of 24 participants (12.5%).*

3. Complaints about the heaviness and discomfort of the headset were raised. *Mentioned by 5 out of 24 participants (20.83%).*
4. Some participants wanted to play a video game rather than an application that controls a physical robot for inspection. *Mentioned by 2 out of 24 participants (8.33%).*
5. One participant commented: "The VR headset caused some motion sickness after prolonged use. VR could limit the amount of time I can work on a task." *Mentioned by 1 out of 24 participants (4.17%).*

Positive aspects of the 2D interface:

1. One participant commented: "In tasks that required precise, fine-tuned input, I found the mouse and keyboard more dependable than a VR controller; maybe a mix of both interfaces could be smart in some situations." *Mentioned by 1 out of 24 participants (4.17%).*
2. Allows the user to step back and view the environment from a distance rather than being fully immersed. *Mentioned by 4 out of 24 participants (16.67%).*
3. The 2D view can help avoid collisions, as stated by one participant. *Mentioned by 1 out of 24 participants (4.17%).*
4. One participant highlighted that the learning curve for the 2D interface was less steep, making it potentially more inclusive for people with varying technology knowledge levels. *Mentioned by 1 out of 24 participants (4.17%).*

Negative aspects of the 2D interface:

1. One participant commented, "I felt more detached from the task in the 2D interface. It felt more like manipulating an object in a software program rather than interacting with a real robot." *Mentioned by 1 out of 24 participants (4.17%).*
2. Participants reported exerting more effort and mental focus when using the 2D interface. *Mentioned by 6 out of 24 participants (25%).*
3. There were reports of difficulty visualising anomalies' locations with a 2D screen. *Mentioned by 4 out of 24 participants (16.67%).*

4. Some participants found the 2D interface less intuitive and less effective for quickly completing tasks. *Mentioned by 6 out of 24 participants (25%).*

Participants lean towards VR due to its intuitive and immersive nature, even though it may have certain physical limitations. The 2D interface was considered less intuitive and more mentally demanding, though it still needed to be considered functional. A suggestion was made for a more advanced training period to improve the VR experience before starting the task and the possibility of having some features of the 2D interface, such as cutaway views, available in VR.

5.5 Summary

This Chapter evaluated and compared non-immersive (2D) and immersive (VR) interfaces for remote visual inspection in unmanned surface vessels, emphasising the significance of human factors and performance metrics to improve telepresence capabilities for marine engineering maintenance tasks.

The primary aim of the experiment was to evaluate the effectiveness of both interface modalities in terms of human factors and performance metrics. Upon data analysis, the findings provided clear insights into the hypotheses posed. The primary hypothesis, H1, postulated that users would experience a heightened sense of presence during remote inspection tasks when using an immersive VR interface compared to a traditional 2D interface. The results from the Two-way ANOVA supported this hypothesis, revealing a statistically significant main effect of modality on the sense of presence $F(1, 44) = 84.86, p < 0.001$. This implies that the VR interface offers a more profound sense of presence than its 2D counterpart. Consequently, the null hypothesis, H01, which proposed no difference in the sense of presence between the two modalities, was decisively rejected. **In summation, the immersive nature of VR provides a superior sense of presence in remote visual inspection tasks relative to traditional 2D interfaces.**

Examining the cognitive workload data, the findings shed light on the hypotheses. The primary hypothesis, H2, posited that users would experience a reduced

workload when executing remote visual inspection tasks via an immersive VR interface, in contrast to a traditional 2D interface. The results from the paired t-tests, particularly for the Performance and Effort dimensions, support this hypothesis. Specifically, the **2D interface recorded higher effort scores** with a mean of **8.50 ± 4.98** , compared to VRs of **5.92 ± 4.51** . Additionally, the Performance perceptions were comparable across both interfaces but indicated a variance based on the testing order, notably when participants were introduced to VR first. Given these differences in specific dimensions of the NASA-TLX between the two modalities, the null hypothesis, H02, which asserted that there would be no discernible difference in users' cognitive workload between the two interfaces, is challenged. However, it is worth noting that while some dimensions showcased differences, others did not, indicating that **the workload experience is multifaceted and not solely reliant on the interface modality**.

Participants, on average, rated the VR interface closer to the “relatively good usability” benchmark than the 2D interface, lending support to H3. However, the sequence in which the interfaces were presented influenced their perceptions. Beginning with the VR interface, elevated usability expectations were set, causing a subsequent 2D experience to appear less favourable. While **the initial exposure to the 2D interface showed no significant usability score difference between the two interfaces, starting with VR revealed a statistical advantage in favour of the VR interface**. Thus, while H3 gains partial validation based on sequence, the null hypothesis **H03 holds when the 2D interface is introduced first but is refuted when the sequence begins with VR**. This means that there is another effect of learning on the results.

When comparing average completion times, tasks conducted via the VR interface were completed approximately 16% faster than those on the 2D platform. Statistical analysis further confirmed a difference in completion times between the two interfaces. Interestingly, the order in which the interfaces were evaluated affected the magnitude of this difference. **Regarding errors, participants registered 44% fewer collisions in the VR environment than in the 2D setting**. While the interface choice influenced error rates, the testing order did

not, and the interface's effect on errors was consistent regardless of the testing sequence.

The Pearson correlation coefficient measures the strength and direction of a linear relationship between two variables. It was used to understand the relationship between the number of errors (collisions) participants made and the time they took to complete the task. It showed a moderate positive relationship for the 2D interface ($r = 0.412$), indicating that as errors increased, so did completion time. In contrast, the VR interface displayed a weaker positive relationship between these two variables ($r = 0.283$). This suggests that while there is some association between more errors and longer completion times, it is less pronounced in the VR environment. Considering the findings, the hypothesis (H4) that **users perform better in terms of time and accuracy using a VR interface than a 2D interface is supported**. The null hypothesis (H04), suggesting no difference in performance between the interfaces, is rejected.

However, it is important to contextualise these findings within the scope of the tasks performed. The similarity of the tasks across modalities suggests that any learning effect might not necessarily translate to a genuine preference for one interface over the other; instead, it reflects an adaptation to the task irrespective of the interface used. The efficiency gains and reduced error rates in VR may stem from the interface and increased spatial awareness and engagement, which aligns more closely with natural human perception and interaction. **Therefore, while the benefits of VR are evident, they are not solely attributable to a novelty effect or a learning curve but also to the inherent qualities of the interface that resonate with innate human cognitive and perceptual capabilities.** This nuanced understanding is vital for interpreting the results beyond surface-level preferences, pointing to the intrinsic advantages of VR in facilitating more natural and effective interaction paradigms for distant visual inspection tasks.

The study evaluated if prior VR experience improves performance in immersive VR tasks, using OLS regression to analyse task completion times and collisions. Results showed a non-significant 11.09-second reduction in completion time for

those with VR experience ($p=0.818$) and a slight but not significant decrease in collisions ($p=0.057$), suggesting minimal impact of VR experience on task efficiency and precision. Consequently, the null hypothesis, H_{05} , which posited no difference in performance outcomes between users with prior VR experience and those without, cannot be firmly rejected based on this data. This implies that factors other than prior VR experience might play a more significant role in determining performance outcomes in these remote visual inspection tasks. **Thus, while prior VR experience may offer some advantages, such as potentially quicker adaptation to VR interfaces, these benefits are not strongly pronounced in enhancing task efficiency or accuracy within the scope of this study.**

Overall, this Chapter has presented a comprehensive evaluation of 2D and VR interfaces for remote visual inspection in unmanned surface vessels, showcasing the profound advantages of VR interfaces regarding presence, workload, usability, and performance. **The insights derived not only confirm the superiority of the immersive VR environment over traditional 2D interfaces in specific domains but also emphasise the need to consider sequential effects in human-computer interaction studies. It would be worth investigating how this superiority was measured with specific user interfaces and the impact of user interfaces.**

6 Experiment 2: Impact of data visualisation modalities on cognitive workload in VR-based industrial robot teleoperation

6.1 Introduction

The previous Chapter demonstrated the advantages of immersive VR interfaces for remote visual inspection, including improved user spatial awareness and reduced cognitive workload. However, it also highlighted the impact of data visualisation on cognitive workload and task performance, indicating potential areas for further exploration.

This Chapter investigates the implementation and evaluation of four data visualisation modalities designed to represent the spatial information of a remote robot operating space. These modalities include (1) a 2D live camera feed, (2) a 3D live point cloud, (3) a combination of a 2D live camera feed and 3D live point cloud, and (4) Augmented Virtuality with OpenCV markers.

A user study was conducted to investigate the cognitive workload associated with different modalities and how the order of testing influenced the cognitive demand and effectiveness of the display methods. Pupillometry, a potentially helpful metric for studies involving different visualisation methods in VR for remote visual inspection, was employed to measure cognitive workload.

VR interfaces enhance the sense of user presence in both real and virtual environments, enabling untrained users to operate robotic systems intuitively. When using VR interfaces for teleoperation, the digital representation of the robotic system typically offers a visual and interactive depiction of the robot and its immediate surroundings. However, more than this representation is required to provide a comprehensive understanding of the spatial layout and the full extent of the remote environment where the robot operates. This limitation in spatial awareness arises because the digital representation often focuses on the robot and its direct interactions, potentially neglecting the broader spatial context.

The primary aim is to evaluate and compare the cognitive workload associated with four distinct data visualisation modalities: 2D live camera feed (named 2D),

3D live point cloud (named 3D), a combination of 2D live camera feed and 3D live point cloud (named 2D3D) and augmented virtuality with 2D live camera feed and 3D live point cloud (named AV). This will be achieved through a user study involving tasks that require participants to interact with a remote environment via a VR interface. Cognitive workload will be measured using pupillometry and subjective workload assessments.

6.2 Experiment and modalities setups

6.2.1.1 Robot Safety Constraints

The participants manipulated an industrial robotics arm with appropriate safety constraints and limits in place to prevent any wrong operations that could damage the equipment and unwanted robot rotation motion due to singularities. Safety constraints have been applied at Unity 3D software and robot controller levels. Robot singularities originate from the robot's inverse kinematics. When a robot is positioned at a singularity, infinite ways exist to achieve the same position. If the optimal solution is not selected as the robot controller does, the robot's joints could be commanded to move impossibly, causing issues such as infinite velocity. Manufacturers often program robots to halt entirely if a joint is commanded to move too fast to prevent damage caused by singularities. However, this is not an elegant solution, and contemporary manufacturers are refining their singularity avoidance strategies. For instance, they program the maximum speed for each joint, so when a joint is commanded "to infinity", it reduces the tip's velocity instead of a complete halt. Programmers often resort to advanced techniques to prevent singularities. Research has led to various solutions, some of which are starting to be implemented in industrial robots (Rebouças Filho et al., 2019). Robots with more axes offer more potential for singularities, but at the same time, extra axes can provide alternative positions to reach the same point, reducing the effect of singularities. Robot technicians also have methods to prevent singularities, such as adding small angles to the tooling to reduce the likelihood of moving into a singularity or relocating the task to a part of the workspace devoid of singularities. Understanding singularities and

developing appropriate countermeasures are crucial steps in advancing the field of robotics.

To mitigate the associated risks, strategic constraints were applied to the end-effector rotation capabilities of the Fanuc robot. This limitation is designed to minimise the likelihood of encountering singularities - points of instability that could compromise the robot's operational integrity and unwanted hazardous rotation by the participant leading the collision with the experiment scene. In addition, a designated safety zone was delineated (Figure 6-1), thereby reducing the area susceptible to singularities. This precaution not only further safeguards against these problematic points but also secures the robot from being manipulated by individuals unfamiliar with its operation, thereby ensuring overall safety.

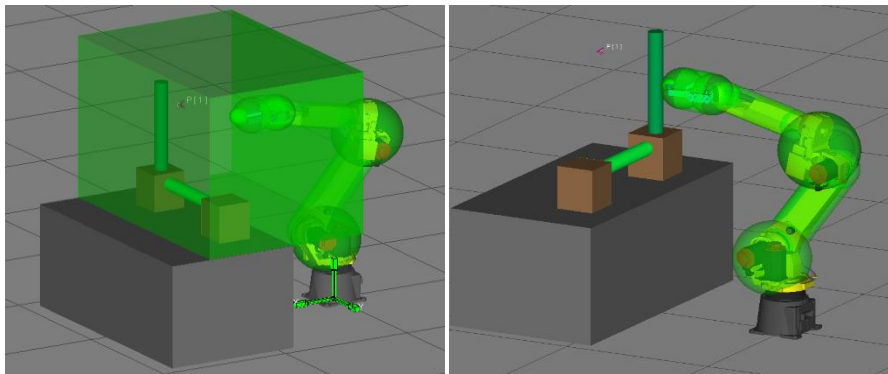


Figure 6-1 Dual Check Safety (DCS) areas (in green)

Figure 6-2 details the actions taken for the safe operation of the robot. The 3D cursor refers to the virtual representation of the Kinect camera used for manipulation. If one of the step answers is "No", the pose of the Fanuc robot is not updated. When the robot cannot reach the operator's desired pose, the virtual representation of the Kinect becomes red, the VR controller's vibration is turned on, and a virtual box is instantiated, representing the last valid pose to help the operator retrieve the control of the robot.

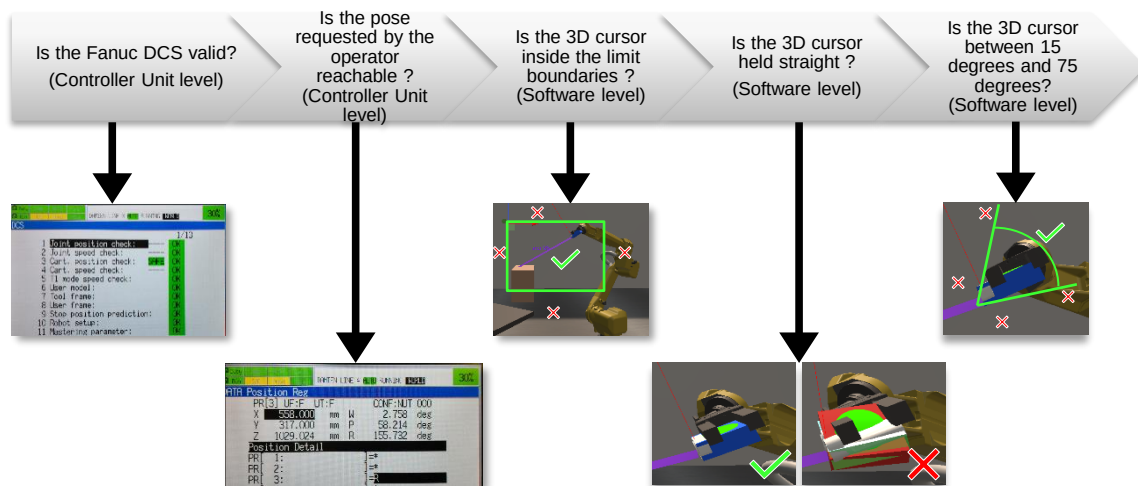


Figure 6-2 Actions taken for the safe operation of the robot

6.2.2 Modalities setups

depicts the layout of the experimental environments, illustrating the positions of the boxes, foam pipes, and task markers. Each task commences from a consistent start point, with the robot returning to this position after completing each task. The tasks have been designed to span a range of use cases. Task 1 (Figure 6-3, t1) directly faces the robot, visible from the start position without requiring a camera or robot complex rotation. On the other hand, task 2 (Figure 6-3, t2) is visible from the start position; successful completion demands the camera's rotation to face the participant's ground. Task 3 (Figure 6-3, t3) requires the participant to navigate the camera and robot, as the shapes to be inspected on the left side are not initially visible from the robot's starting position. This variety of scenarios allows us to examine the participant's cognitive workload and operational capabilities under varying degrees of challenge and, therefore, robustly evaluate the effectiveness of the four data visualisation modalities.

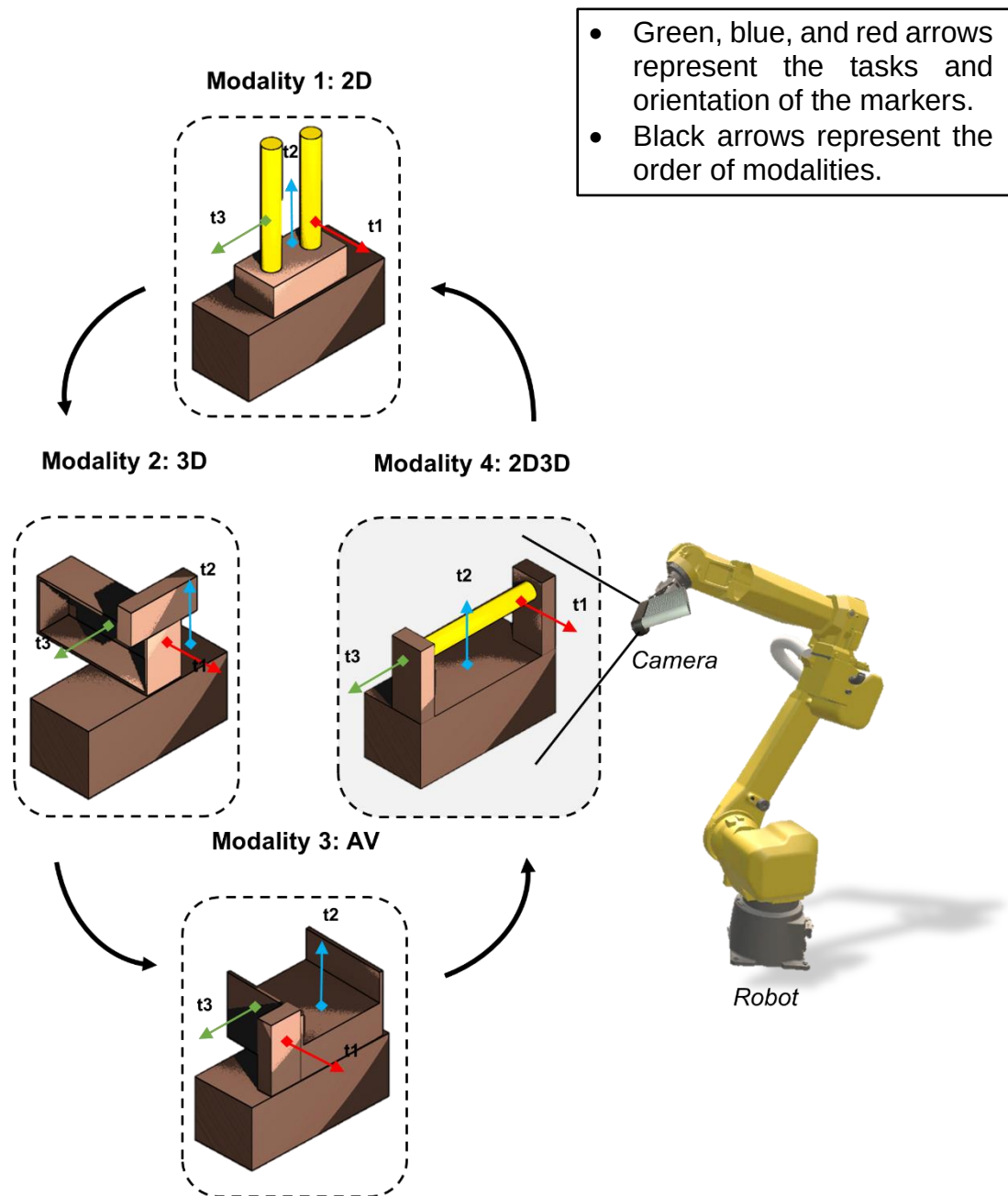


Figure 6-3 Modalities setup and positions of areas to inspect (t1, t2, t3)

6.3 Hypotheses

This Chapter attempts to test several hypotheses and the associated null hypotheses.

These hypotheses aim to determine if a specific modality (or combination of modalities) provides a lighter cognitive load, leading to higher performance scores.

H1: The modality type affects the perceived workload and the physiological cognitive load.

H01: The type of modality does not affect the perceived workload or the physiological cognitive load.

H2: The type of visualisation modality and the order in which they are presented influence the task completion times of participants.

H02: Neither the type of visualisation modality nor the order in which they are presented influences participants' task completion times.

H3: The type of visualisation modality and the sequence of their presentation influence participants' performance scores.

H03: The type of visualisation modality and the sequence of their presentation do not affect participants' performance scores.

6.4 Results

For all four modalities, the user will manipulate the robot in VR, allowing for natural and easy trajectory planning without extensive programming or training.

1. **Modality 2D:** In this modality, users relied solely on the live 2D camera feed from the Azure Kinect sensor to perform the inspection task. This modality relies on the video feed captured by an Azure Kinect sensor. It gives the user a standard two-dimensional view of the remote environment, like watching a video. Users perform inspection or operational tasks by viewing this live feed. This modality might be more straightforward and intuitive for those accustomed to regular video feeds

but could lack depth perception and spatial awareness provided by 3D representations.

2. **Modality 3D:** In this modality, users are given a real-time 3D point cloud visualisation of the remote environment generated from the Azure Kinect's depth data. This modality generates a 3D point cloud from the depth data the Azure Kinect sensor captured. A point cloud is a set of data points in space, representing the external surface of objects in the environment. This modality provides users with a more spatially accurate representation of the remote environment, enhancing depth perception and understanding of the spatial layout.
3. **Modality 2D & 3D:** In this modality, users are provided with both modality 2D and modality 3D. Users can switch between the 2D live camera feed and the 3D point cloud or compare them. This combination allows for a more comprehensive understanding of the environment. This modality can benefit tasks where spatial depth and detailed surface features are essential. The ability to switch or compare views helps in identifying crucial details in both 2D and 3D spaces.
4. **Modality Augmented Virtuality (AV):** Integrates live 2D camera feed with minimalist 3D geometries. This modality overlays simple 3D geometrical shapes (obtained from OpenCV markers) onto the live 2D camera feed. OpenCV markers detect and track objects or locations in the environment. It provides a hybrid view combining real-time video with 3D spatial cues, which can be especially beneficial in tasks requiring precise interaction with objects in the environment. The 3D geometries can provide a basic but effective representation of spatial relationships, aiding navigation, and object manipulation.

Figure 6-4 shows screen captures during the experiment showing the modalities described above. Additionally, a screen capture of the questionnaire is presented, and cognitive workload and usability questions are displayed in VR.

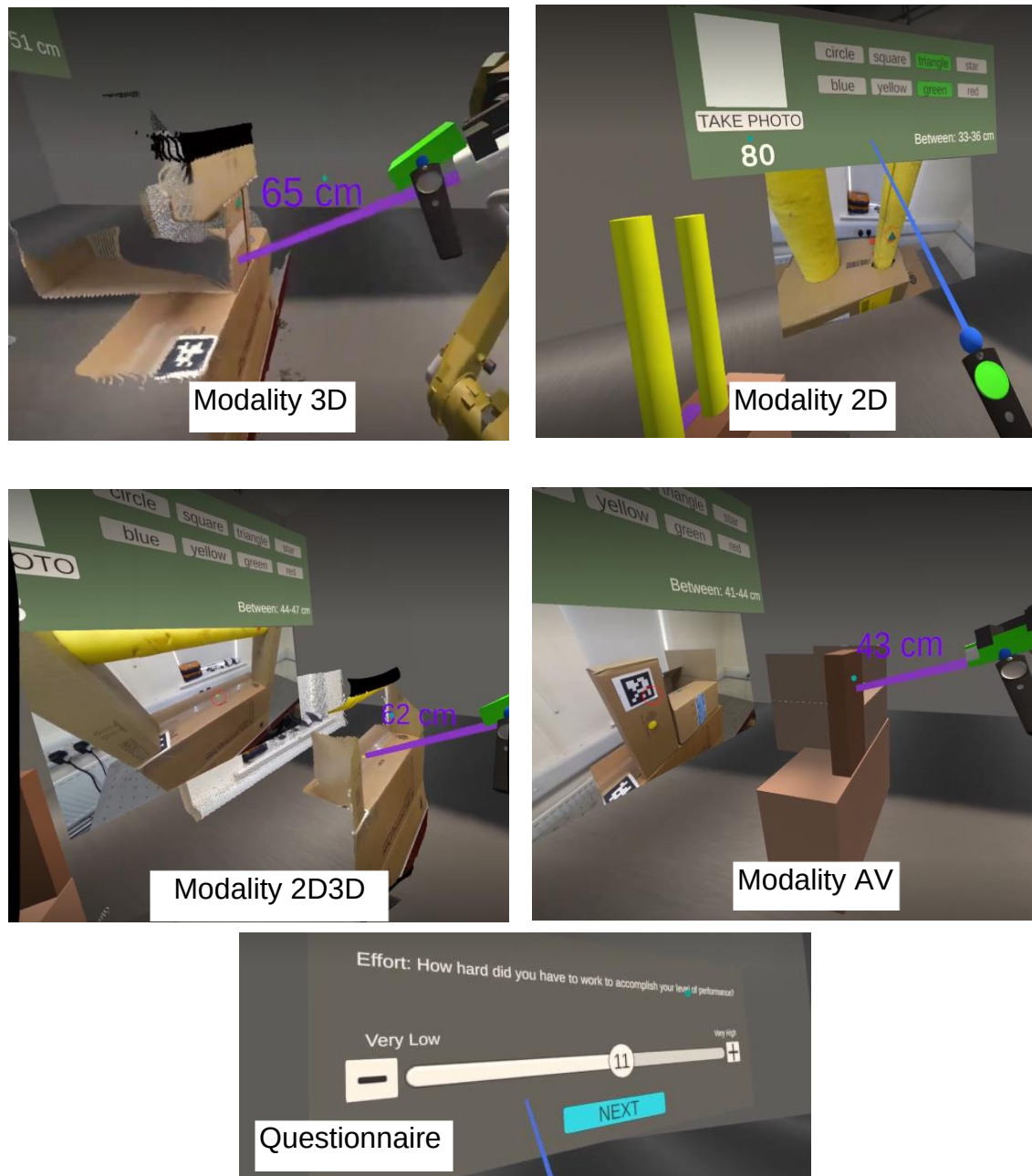


Figure 6-4 Screen capture of the modalities during the experiment

6.4.1 Cognitive workload

6.4.1.1 Questionnaire

Across the four modalities (2D, 3D, 2D3D, AV), participants' perceived workloads were similar for each NASA-TLX dimension, as illustrated in Figure 6-5. The 2D modality showed the highest mental demand, while the 3D modality indicated the highest temporal demand. In contrast, the 2D3D modality was perceived as less

mentally and temporally demanding than the others. Overall performance perceptions were consistent across modalities. As indicated by error bars, the variability in responses suggests that participants' experiences varied within each modality, particularly noticeable in the AV modality's temporal dimension. While there are nuances in the perceived workload for each modality, no single modality was consistently seen as the most or least demanding across all dimensions.

The analysis of variance was used to determine if the order in which participants experienced different modalities affected their NASA-TLX scores. The ANOVA compares the variance between different groups (in this case, order groups) to the variance within those groups. In the AV modality, it was identified that the order in which tasks were presented had an influence on participants' perceptions of Mental Demand ($p = 0.0054$), Performance ($p = 0.0339$), and effort ($p = 0.0128$). Similarly, within the 2D modality, the order markedly affected the Frustration dimension ($p = 0.0448$). Additionally, in the 2D3D modality, the Temporal Demand scores were notably affected by the order ($p = 0.0071$).

However, the order's influence did not reach statistical significance for other modalities and dimensions. This suggests that while presentation order shaped

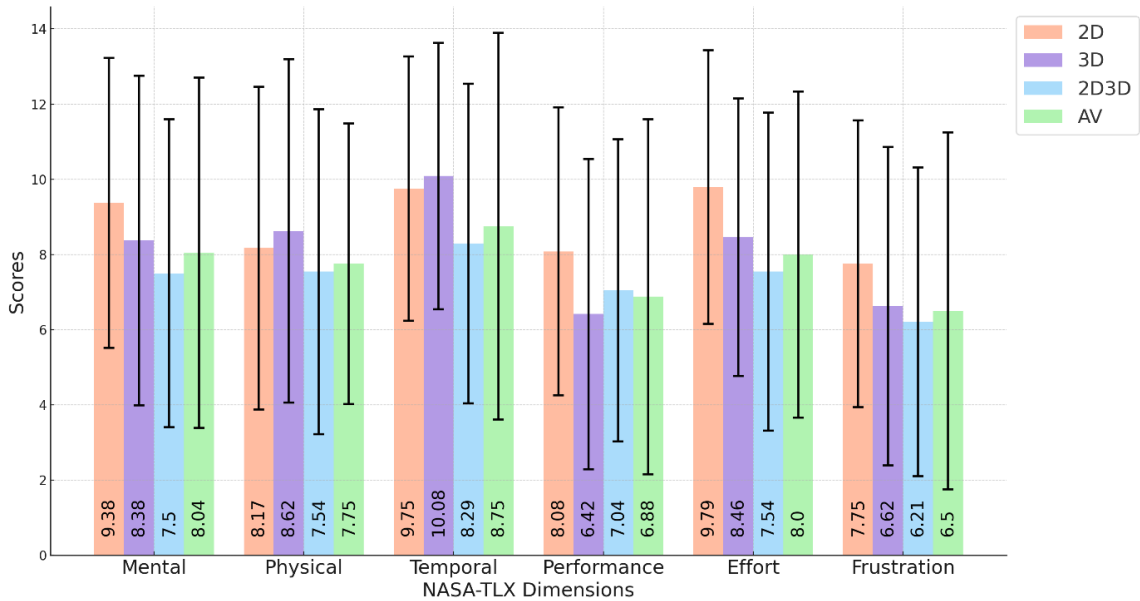


Figure 6-5 NASA-TLX workload assessment across modalities

perceptions for specific dimensions in certain modalities, it might not be a universal factor across all modalities and dimensions.

6.6.1.2 Pupil Dilation

The graph Figure 6-6 shows the differential impact of modalities on pupil dilation, which is used to assess cognitive load. The 3D modality manifests the most pronounced average increase in pupil dilation at approximately $7.43 \pm 3.49\%$. This sizable change suggests that the 3D modality might exert the heaviest cognitive demands on participants. Contrastingly, the 2D modality delineates a more modest average change of about $3.83 \pm 3.22\%$. The Augmented Virtuality (AV) modality charts an intermediate course with an average change hovering around $5.27 \pm 3.30\%$. Lastly, the 2D3D modality echoes the 2D modality in its average change, registering at $3.98 \pm 3.36\%$.

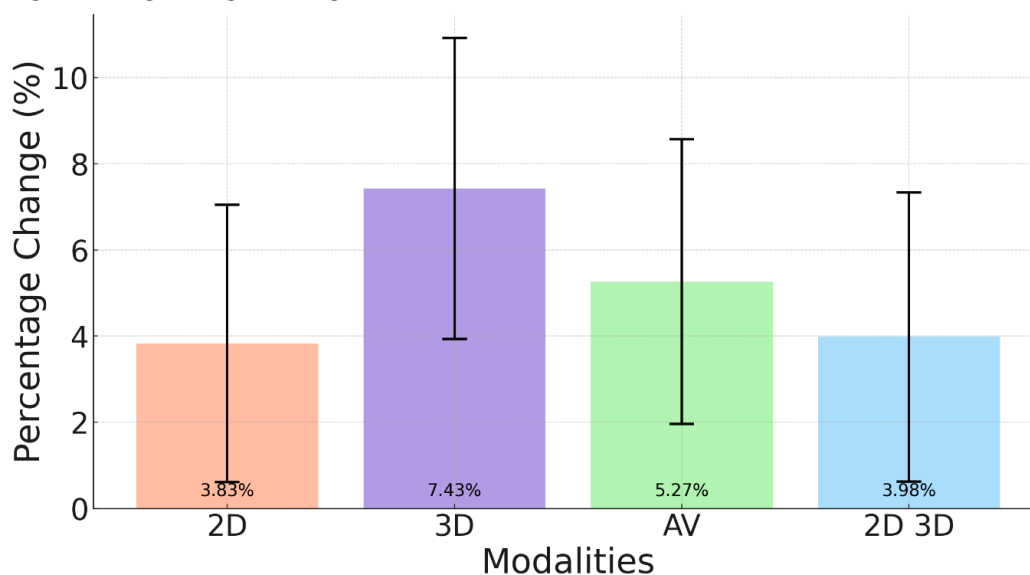


Figure 6-6 Change in pupil dilation from baseline across modalities

An Analysis of Variance was conducted to gauge the influence of presentation order. The results yielded varied p-values: **0.8793** for the 2D modality, **0.2101** for the 3D, **0.2081** for AV, and **0.8369** for the 2D3D modality. The elevated p-values for the 2D and 2D3D modalities suggest a minimal impact of presentation order on pupil dilation changes. Although the 3D and AV modalities exhibited reduced p-values, they remained above the conventional threshold of 0.05, implying that the differences based on order might not be statistically compelling.

6.4.2 Performance

6.4.2.1 time

In the box plots Figure 6-7, the distribution of task completion times across the four visualisation modalities, 2D, 3D, 2D3D, and AV, is presented. At the core of the box plot is a rectangular box that encapsulates the data's interquartile range (IQR). The bottom and top of this box denote the first (Q1) and third quartiles (Q3), respectively, representing the 25th and 75th percentiles. Thus, the height of the box captures the range within which the central 50% of the values lie. A crucial feature within the box is a horizontal line, which signifies the median (Q2) of the dataset—the point at which 50% of the data lies below and 50% lies above. The position of the median within the box can provide insights into the skewness of the data. If the median is closer to Q1, it indicates a skew towards the lower end of the dataset, and conversely, if it is nearer to Q3, it suggests a skew towards the upper end. The individual points plotted outside the whiskers are typically considered outliers. They represent data points that fall outside the typical range of the bulk of the data.

The 2D modality exhibits the highest median time, suggesting that participants typically took longer in this setting. The 3D and 2D3D modalities display similar

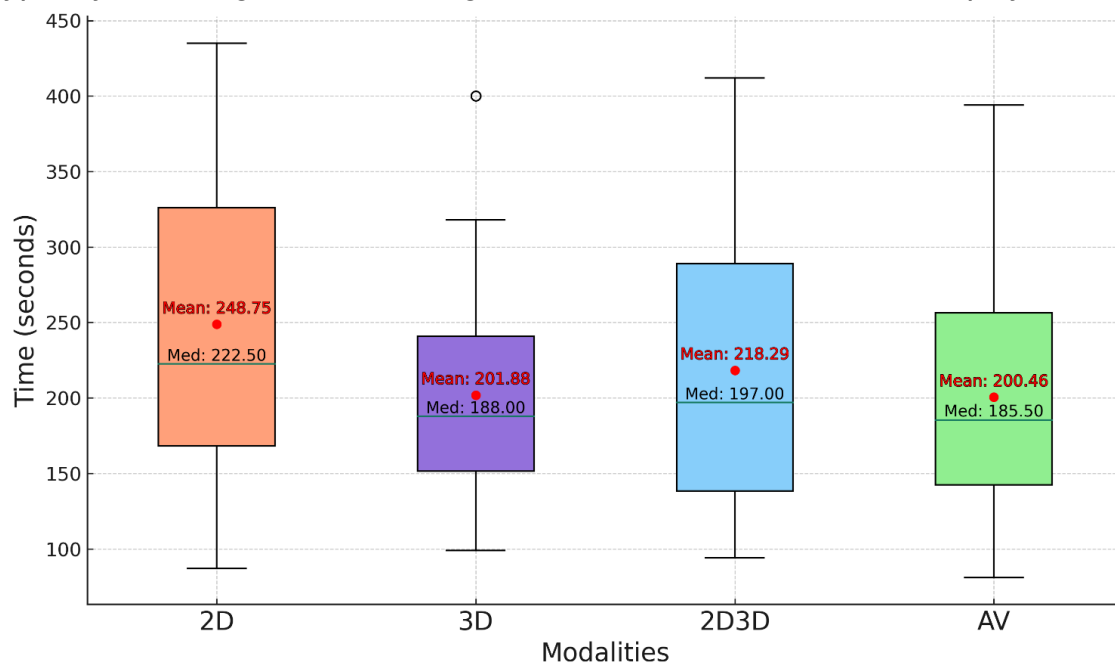


Figure 6-7 Box plot of task completion times by modality

median times, with 2D3D having slightly more variability in its data distribution. The AV modality has the lowest median time, indicating that participants, on average, were quickest in this setting. The mean time values, represented by red dots, further emphasise these observations. The 2D modality's mean is the highest, while AV's is the lowest.

In assessing the influence of modality and presentation order on completion time, a two-way ANOVA was conducted and descriptive Figure 6-8. The results indicated a main effect of presentation order on completion time, $F(3,80) = 11.52, p < 0.0001$, suggesting that **the order in which the modalities were introduced to participants played a notable role in determining how quickly they completed tasks**. In contrast, when examining the modalities themselves, irrespective of order, no differences in completion times were detected across the modalities, $F(3,80) = 2.13, p = 0.102$. This suggests that, on average, no modality inherently took more or less time than the others. This could be due to a balance between the complexity and information richness each modality provided and the cognitive load they imposed. While advanced modalities like augmented virtuality (AV) and the 3D live point cloud might have offered more spatial information, they potentially increased cognitive workload, counterbalancing their intuitive advantages. Conversely, simpler modalities like

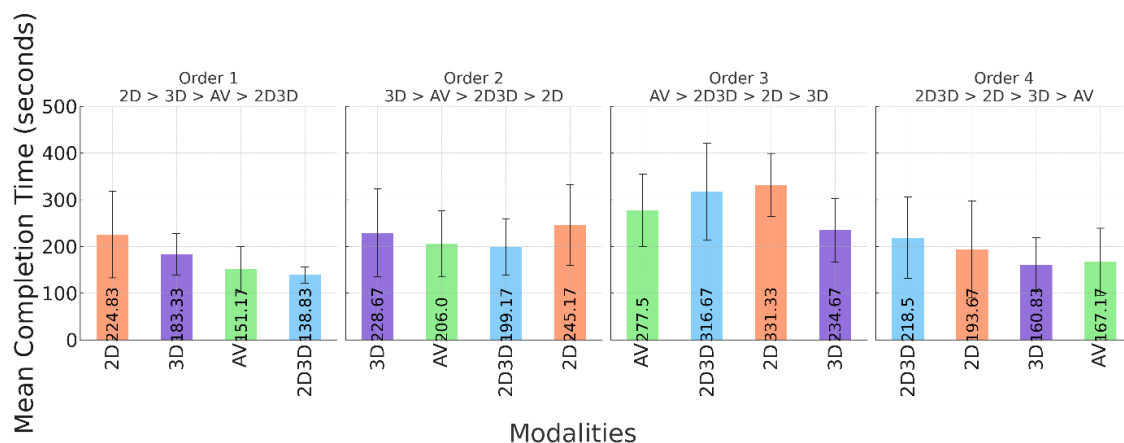


Figure 6-8 Mean completion time of modalities by order

the 2D feed, though less immersive, might have been easier to process cognitively, leading to similar completion times across modalities. Lastly, the interaction between modality and presentation order was not significant,

$F(9, 80) = 0.877, p = 0.5493$. This implies that the effect of a specific modality on completion time remained consistent regardless of the order in which it was presented to the participants. In summary, while the order of presentation had a clear impact on task completion times, the type of modality used and its interaction with order did not significantly alter these outcomes.

Moreover, the 2D modality consistently has longer completion times than the other modalities, except for Order 4. This could indicate that the 2D visualisation is inherently more challenging or less intuitive for participants than the other modalities. The 3D modality has stable completion times across different orders. While it is quicker than 2D, it tends to be slower than the AV and 2D3D modalities. This might indicate that the 3D visualisation strikes a balance between complexity and intuitiveness.

6.4.2.2 Other performance indicators and scoring framework

The AV modality recorded the highest number of collisions in the experiment data on collisions across the visualisation modalities, with approximately **5.56%** of the trials resulting in a collision. In contrast, the 2D and 2D3D modalities had about **1.39%** of their trials ending in collisions, while the 3D modality observed no collisions.

In the analysis of distance accuracy, the 2D modality emerged as the most compliant, with approximately **84.72%** of the trials respecting the distance threshold. The 3D and AV modalities followed closely, with **79.17%** and **80.56%** of their trials adhering to the distance guidelines. The 2D3D modality, though the least compliant among the four, still maintained a notable **76.39%** adherence rate.

Among the modalities, the 2D3D visualisation had the highest rate of well-taken photos, with almost **97.22%** of its trials achieving this mark. The 3D and AV modalities had approximately **93.06%** success rates, while the 2D modality was slightly higher with **94.44%**. Figure 6-9 shows examples of photos; correct photos are the ones with at least half of the marker in the red circle.

Participants showed high accuracy in reporting shape and colour across all modalities. The 3D visualisation modality had a perfect score, with all participants



Figure 6-9 Examples of incorrect and correct photos from participants

reporting accurately in every trial. The 2D3D modality followed closely, with a **98.61%** success rate, while both 2D and AV modalities achieved a commendable **97.22%** accuracy rate.

In evaluating participants' scores across various visualisation modalities, each modality displayed high average scores out of 15 points as presented in section 6.4.2.2. The 3D modality achieved the highest average score of **13.46**, followed closely by 2D3D, with an average of **13.33**. Both the 2D and AV modalities tied with an average score of **13.00**. Results are shown in Figure 6-11. Despite these high averages, it is notable that the minimum scores recorded for the 2D3D, and AV modalities were **8**, slightly lower than the 2D and 3D modalities, which had a minimum of **10**.

Several patterns emerge when examining the influence of modality presentation order on participants' scores. The sequence of modalities plays a crucial role in participants' performance, as illustrated in Figure 6-10.

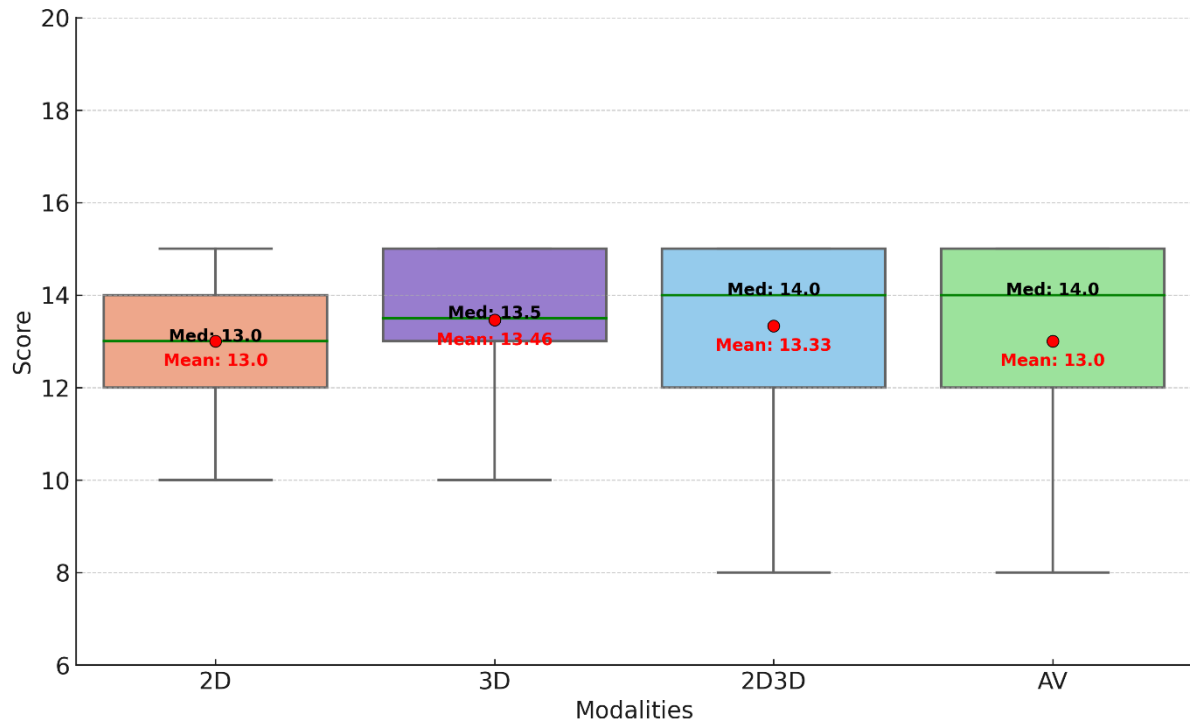


Figure 6-11 Distribution of scores (from the scoring framework introduced in section 3.5.3.2) across visualisation modalities

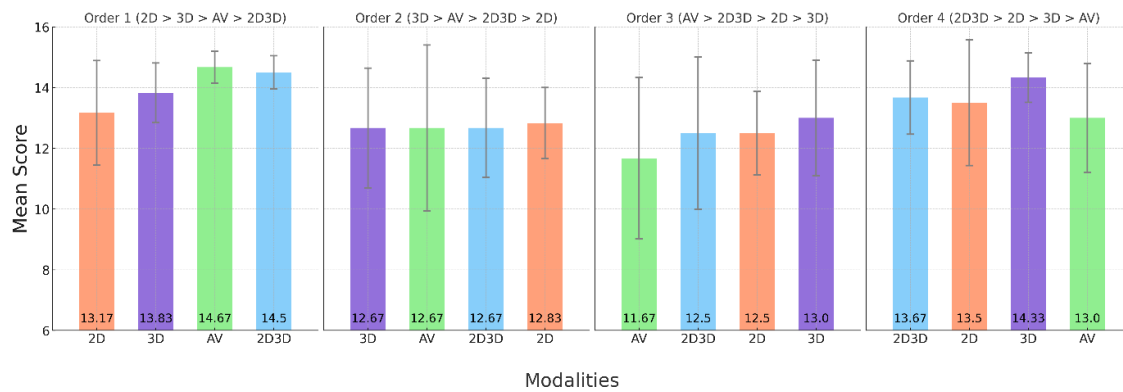


Figure 6-10 Comparison of mean scores across modalities by presentation order

- For Order 1 (2D > 3D > AV > 2D3D), the AV modality, which was presented third, achieved the highest average score of **14.67 ± 0.52**. This suggests that participants may have become more adept or comfortable with the tasks when they encountered the AV modality. The 2D3D modality, presented last, also scored highly, with an average of **14.50 ± 0.55**, further supporting this hypothesis.

- In Order 2 (3D > AV > 2D3D > 2D), there was a noticeable drop in scores, with all modalities clustered closely around the **12.67–12.83** range but with varying levels of variability. Despite being presented second, the AV modality shared the same average score of **12.67** with 3D and 2D3D modalities but had a higher standard deviation (**2.73**), indicating a broader spread in participant scores.
- For Order 3 (AV > 2D3D > 2D > 3D), the 3D modality, presented last, secured the highest score of **13.00 ± 1.90**. The AV modality, introduced first in this sequence, had the lowest average score of **11.67 ± 2.66**, suggesting that introducing a complex modality first might pose initial challenges for participants.
- Lastly, Order 4 (2D3D > 2D > 3D > AV) saw the 3D modality presented third, garnering the top score of **14.33 ± 0.82**. The sequence began with the 2D3D modality with a commendable average of **13.67 ± 1.21**, potentially setting a positive trajectory for the following modalities.

The two-way ANOVA analysis assessed the effects of visualisation modality and presentation order on participants' performance scores. The results revealed that neither the type of visualisation modality (2D, 3D, AV, 2D3D) nor the sequence of their presentation statistically influenced the scores. Specifically, the modality factor had an **F-value of 0.406 (p = 0.749)**, the order factor had an **F-value of 0.880 (p = 0.351)**, and the interaction between the two factors had an **F-value of 1.154 (p = 0.332)**. This implies that participants' performance was consistent regardless of the visualisation type or the order in which they experienced these visualisations. The lack of a significant interaction further indicates that the effect of presentation order remains consistent across different modalities.

6.5 Summary

At first, perceived workloads across the four modalities, 2D, 3D, 2D3D, and AV, appear homogenous for each NASA-TLX dimension. The data on pupil dilation shows a clear differential impact of modalities on cognitive load. The 3D modality, with a pronounced increase in pupil dilation of approximately $7.43 \pm 3.49\%$, underscores its potential to impose the most substantial cognitive

demands on participants. This starkly contrasts the 2D modality, which recorded a modest average change of about $3.83 \pm 3.22\%$. The AV and 2D3D modalities also displayed notable average changes, at $5.27 \pm 3.30\%$ and $3.98 \pm 3.36\%$, respectively, hinting at varying levels of cognitive engagement. These physiological findings provide tangible evidence supporting Hypothesis H1, suggesting that the type of modality does indeed affect cognitive load. Given the pronounced differences in pupil dilation across modalities, particularly between the 3D and 2D modalities, substantial support was found to refute the null hypothesis H01. However, the role of presentation order adds complexity to the interpretation. The ANOVA results aimed to decipher the influence of presentation sequence on pupil dilation. Despite observing some variability in p-values across modalities, none of the values breached the conventional threshold of 0.05. **This implies that while the type of modality impacts cognitive load, the order in which these modalities are presented might not play as pivotal a role in shaping the physiological response.**

About Hypothesis H2, the analyses have provided substantial insights, including differences in median task completion times across visualisation modalities. Specifically, the 2D modality necessitated longer completion times compared to others, hinting at unique challenges or reduced intuitiveness associated with it. Conversely, the AV modality emerged as the most efficient. This differentiation in modalities suggests an inherent effect on task completion times. The results of two-way ANOVA validated that the order in which modalities were introduced to participants played a pivotal role in how quickly they completed tasks. These analyses show that the order of presentation is consistent with the assertions of hypothesis H2. However, when considering the individual impact of each viewing modality, without the influence of presentation order, the data tend to confirm null hypothesis H02. **This suggests that while the order of presentation plays a crucial role in task completion times, individual modalities show uniform performance for varied orders when evaluated independently.**

A hypothesis (H3) was posited regarding the potential impact of visualisation modality types and their presentation order on participants' performance scores,

suggesting that both the type of visualisation modality and the sequence of their presentation influence participants' performance outcomes. The null hypothesis (H03) posited no effects from either factor. The descriptive analysis of the performance data, thanks to the scoring framework, postulated that both factors would have a discernible effect. However, the evidence culled from the two-way ANOVA paints a different picture. Despite observable variations in individual modality outcomes, statistically, neither the visualisation modality ($F=0.406$, $p=0.749$) nor its presentation sequence ($F=0.880$, $p=0.351$) influenced performance scores. Moreover, the interaction term further cemented this, yielding an F-value of 1.154 with a p-value of 0.332. This statistical uniformity across different modalities and their sequences refutes the original hypothesis, leading to an acceptance of the null hypothesis. While participants may exhibit preferences or slight variations in performance across modalities, these differences are not statistically significant, indicating that **the visualisation modalities' inherent properties and presentation sequence do not fundamentally alter user performance in this context.**

In summary, this exploration into modality type, presentation order, and their effects on perceived workload, cognitive load, and task performance provides valuable insights into the multifaceted interactions between visualisation modalities and user experiences. Understanding these nuanced interactions will be pivotal in designing more effective, intuitive, and user-centric telepresence systems. **Future studies could investigate how different visual modalities interact with individual user characteristics, such as spatial abilities, prior experience with VR systems, and personal preferences.** This would provide a more personalised understanding of the user experience in telepresence systems. **Additionally, research could explore the integration of adaptive interfaces that dynamically adjust the complexity and type of visual information based on real-time assessments of user cognitive load and task performance.** This could lead to more intuitive and user-friendly systems that cater to the needs of a broader range of users, thereby enhancing overall efficiency and effectiveness. Investigating the long-term effects of using different modalities on skill acquisition, retention, and transfer in teleoperated tasks would

also be valuable. Such research would contribute to developing more advanced teleexistence systems and a deeper understanding of human-computer interaction in immersive environments.

7 Discussion

7.1 Principles for assessing and implementing remote inspection with telexistence capability

This thesis has contributed to human-machine interaction within immersive environments by determining several principles. It supports developing and deploying telexistence capabilities in a remote inspection context. These principles are instrumental for developing, assessing, and implementing telexistence capabilities, ensuring they meet the specific demands of remote inspection and maintenance tasks. Most current research focuses on evaluating human performance on VR-based teleoperation frameworks with little or no consideration of human factors of cognitive workload imposed on the user.

This research has methodologically progressed through a series of distinct yet interconnected stages. Initially, it involved developing and proposing a novel framework with telexistence capabilities specifically designed to operate an accurate industrial robotic arm. This framework represents a significant advancement in the field of human-robot interaction, offering a more immersive and intuitive means of controlling industrial machinery remotely. Subsequently, the research analysed this proposed telexistence interface against a conventional 2D interface in a remote visual inspection context. This comparison, grounded in a comprehensive user study, focused on evaluating human factors. Finally, the study delved into the impact of different data visualisation modalities on cognitive workload. This aspect was rigorously explored through another user study to understand how various visualisation strategies can affect the operator's mental load, efficiency, and overall interaction experience with the telexistence system. The principles of development of telexistence for remote inspection are as follows:

- **Framework Development for Seamless Teleoperation:** Develop frameworks integrating telexistence capabilities with industrial robotic arms using scalable, real-time interaction protocols like DDS. This scalability allows for quick deployment of the architecture on different

robotic systems. This includes creating immersive interfaces for safe and effective human-robot interaction.

- **Evaluating Human Factors in Interface Design:** Assess human factors, including sense of presence, cognitive workload, usability, and performance metrics, when comparing different interface modalities (e.g., 2D vs. VR) for remote inspections. When choosing the display and interaction devices, it is important to conduct a human-centred design assessment of their suitability, as outlined by Stone (Stone, 2012).
- **Cognitive Workload and Data Visualisation Modalities:** Consider the impact of different data visualisation modalities on cognitive workload and task performance, understanding how various modalities influence user experience and efficiency.
- **Adaptation and Customization for Specific Applications:** Adapt and customise telexistence systems for specific applications, considering different industrial sectors' unique challenges and needs, such as marine engineering.
- **Safety and Efficiency in System Design:** Design telexistence systems focusing on safety and operational efficiency, prioritising features that minimise operational errors and optimise task performance.
- **Incorporation of Learning and Adaptation Effects:** Acknowledge and incorporate learning and adaptation effects in user training and interface design, recognising how familiarity with different interfaces influences user performance.
- **Comprehensive Evaluation and Iterative Improvement:** Continuously evaluate and iteratively improve telexistence systems based on user feedback, technological advancements, and changing operational requirements.
- **Exploration of Adaptive Interfaces:** Investigate the development of adaptive interfaces that adjust complexity and type of visual information based on real-time assessments of user cognitive load and task performance, leading to more intuitive and user-friendly systems.

7.2 Key actors of the telexistence industry

Telexistence and its application in various industries have gained significant momentum, with several key players driving innovation in this space. Focusing on companies like Telexistence Inc. (<https://tx-inc.com/en/home/>), Extend Robotics (<https://www.extendrobotics.com/>), and Sarcos Robotics (<https://www.sarcos.com/>), we can gain insights into the current landscape and focus areas in the field of telexistence.

- **Telexistence Inc.:** This company is at the forefront of developing remote-controlled robots designed to replicate human actions and presence in distant locations. Their focus is on creating robots that can operate in environments that are either hazardous or challenging for humans. Telexistence Inc. emphasises enhancing the capabilities of these robots to interact with the physical world like humans, thus expanding the potential for remote operation in sectors like retail, manufacturing, and disaster response.

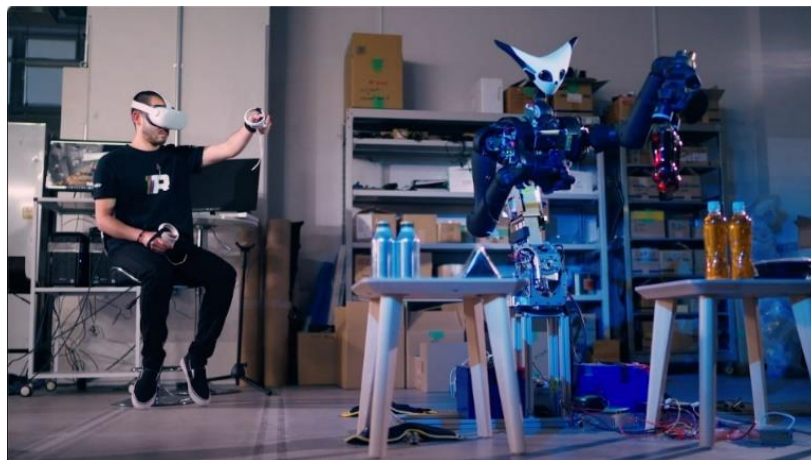


Figure 7-1 Telexistence's model-T robot (June 2021)

- **Extend Robotics:** Extend Robotics takes a unique approach by focusing on robotics and VR to create an immersive teleoperation experience. Their technology enables operators to control robotic arms from a distance with high precision and intuitiveness, effectively extending human capabilities into remote environments. This has applications in areas such as healthcare, where delicate operations can be performed remotely, and in industries where precise manipulation from a distance is crucial.



Figure 7-2 Extend Robotics low latency VR teleoperation demonstration between the UK and Turkey (October 2023)

- **Sarcos Robotics:** Specialising in robotic exoskeletons and teleoperated robots, Sarcos Robotics is leading the way in enhancing human strength and endurance through robotics. Their exoskeletons are designed to augment human capabilities, allowing for safer and more efficient execution of tasks that would otherwise be physically demanding or dangerous. Additionally, their teleoperated robots are being developed for industrial, military, and public safety applications, where they can perform tasks in environments inaccessible or unsafe for humans.



Figure 7-3 Sarcos Guardian XT demo (May 2022)

These companies contribute significantly to advancing telexistence technologies, focusing on unique aspects like human-like robot interaction, immersive teleoperation, and human capability augmentation. Their innovations are shaping the future of various industries, from healthcare and manufacturing to public safety and defence, by enabling remote operation and collaboration in ways that were previously not possible.

7.3 Enhancing safety and efficiency through AI and machine vision

Integrating AI and Machine Vision into remote operations signifies a significant leap towards enhancing operational safety and efficiency across various sectors, including maritime operations and manufacturing. This transformation in the concept of telexistence, which blends telepresence with autonomous functionality, can change the way unmanned vessels are inspected and maintained. This not only significantly elevates safety by mitigating the need for human presence in hazardous environments but also ensures continuity and efficiency by minimizing operational downtime. In broader applications, the fusion of AI and machine vision streamlines monitoring and inspection processes, ensuring quality control and predictive maintenance. This predictive approach allows for anticipating equipment failures before they occur, safeguarding against unexpected downtimes and maintenance costs. Moreover, the real-time monitoring of work environments helps identify unsafe behaviours or conditions, facilitate immediate corrective actions, and thus reduce the risk of accidents.

Beyond safety, these technologies drive operational efficiency by optimizing production processes and logistics operations. They enable the analysis of vast amounts of data to refine production methodologies, enhance throughput, and streamline supply chain management. Furthermore, when integrated with immersive technologies for training, they offer simulated real-world scenarios for remote workers, enhancing skill sets and preparedness for onsite tasks. Machine vision, combined with AI, also extends real-time support and guidance to field workers, improving problem-solving efficiency and reducing errors. However, implementing these technologies requires a robust infrastructure, effective data management, and ongoing skill development. Ensuring data privacy and system security remains paramount. Despite the challenges of integration, cost, and technology adoption, the long-term benefits of improved operational performance, cost reduction, and enhanced worker safety provide a compelling case for adopting AI and machine vision in remote operations.

7.4 Future steps of telexistence in maintenance

Figure 7-4 illustrates a potential concept that uses a collaborative robot (cobot) equipped with a camera, projector, smartphone, and VR interface. This setup allows the cobot to share images and 3D captures of the workspace with the remote expert. The expert can then manipulate the cobot to project guidance cues directly into the workspace, effectively improving their involvement in the task and allowing the on-site worker to focus on other critical tasks. Such a setup not only enriches the quality of remote assistance but also grants experts unprecedented freedom and flexibility in their work. They can seamlessly interact with and guide on-site personnel or even directly manipulate environments previously constrained by the need for physical presence. The potential of

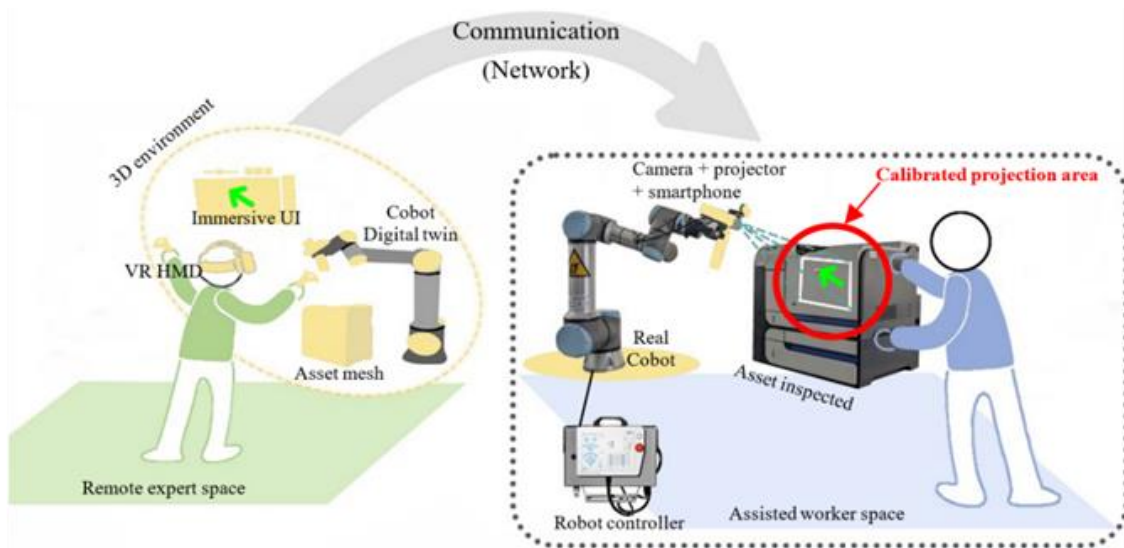


Figure 7-4 Example of expert-worker remote collaboration with telepresence capability

telepresence in transforming remote expert collaboration is immense, offering a more dynamic, efficient, and effective approach to complex maintenance tasks and operational challenges.

Incorporating viewpoint management into telepresence significantly advances remote operational capabilities and aligns with the foundational principles established in my PhD research. This integration is essential for overcoming the current limitations of static or restricted viewpoints in telepresence systems, particularly pertinent in complex industrial environments where spatial awareness and interaction are crucial. The central idea is to enhance the remote operator's ability to interact efficiently with the robotic environment through dynamic viewpoint control (Livatino et al., 2021; Thomason et al., 2017).

Envision a telepresence system where industrial robots are equipped with multiple cameras, providing diverse perspectives like egocentric, exocentric, and bird's-eye views. These cameras, integrated with a VR interface, would allow the operator to switch seamlessly between various viewpoints, offering a comprehensive understanding of the robot's surroundings and tasks. This approach is especially beneficial for intricate operations such as machinery inspection, repair, or assembly in industrial settings.

Future research in this area could focus on developing adaptive viewpoint systems with algorithms that dynamically adjust viewpoints according to task requirements, environmental constraints, and operator preferences. An ergonomic analysis of these viewpoint management systems is also crucial, investigating how different viewpoints impact operator ergonomics, explicitly aiming to reduce fatigue and improve task accuracy. Moreover, integrating MR technologies could further enhance viewpoint management by overlaying essential information onto the operator's view, augmenting the telepresence capability of the system.

8 Conclusions

8.1 Addressing the aim

In pursuit of the research aim, a focus is placed on the repeatability of the process for designing and deploying telexistence capabilities. Repeatability, in this context, refers to the ability to consistently replicate the development and implementation of telexistence systems across different remote maintenance scenarios without the need for significant redesign or reconfiguration. This aspect is vital for several reasons:

- **Standardization:** Developing a repeatable process contributes to the standardization of telexistence technologies in industrial maintenance. Standardisation facilitates the adoption of these technologies by providing clear guidelines and protocols that can be quickly followed, reducing the barrier to entry for organisations and individuals.
- **Scalability:** A repeatable process ensures scalability, allowing telexistence systems to be deployed in a variety of maintenance contexts, ranging from small-scale operations to large, complex environments. Scalability is crucial for the widespread adoption and effectiveness of telexistence technologies.
- **Quality Assurance:** Repeatability inherently includes mechanisms for quality assurance, ensuring that each deployment of the telexistence capability meets predefined standards of performance and user experience. This is essential for maintaining the reliability and safety of remote maintenance operations.
- **Efficiency and Cost-effectiveness:** Organisations can achieve greater efficiency and cost-effectiveness by standardising the design and deployment process. Repeatability reduces the time and resources required to tailor telexistence solutions to specific maintenance tasks, enabling a more streamlined and economical approach.
- **Continuous Improvement:** A focus on repeatability does not preclude innovation; instead, it provides a structured framework for continuous improvement. As new technologies and methodologies emerge, they can

be integrated into the repeatable process, enhancing the capability and applicability of telepresence solutions over time.

To achieve repeatability, the research proposed a modular and component-based design based on DDS, where telepresence systems are constructed using a set of reusable data structures. This approach allowed for the customisation of the system to fit specific maintenance needs while maintaining a consistent and repeatable core structure and functionality reused with the two conducted experiments.

8.2 Addressing the objectives

In addressing the initial objective of this PhD thesis, a thorough investigation into the constraints associated with traditional teleoperation systems was conducted through a literature review. This review unveiled the primary issues stemming from the labour-intensive requirements of conventional inspection techniques and the significant shortfall in providing comprehensive sensory feedback within existing teleoperation frameworks. These identified challenges are found to notably undermine the efficacy, safety, and cognitive workload of operators involved in remote inspections. Moreover, the compilation of the literature's design criteria and best practices marks a significant stride toward developing advanced teleoperation interfaces and control systems. These envisioned systems are anticipated to offer extensive environmental feedback and bolster efficient decision-making processes. By tackling the limitations previously identified, this research contributes vitally to the evolution of more effective, secure, and user-centric teleoperation strategies.

The second objective of this PhD thesis has been the development of a framework designed to facilitate the scalable and repeatable creation of telepresence capabilities. Aimed at enhancing human operators' sense of presence in remote maintenance tasks, this framework prioritizes the improvement of collaboration and interaction within virtual environments. It specifically seeks to overcome the inefficiencies and safety concerns associated with current human-robot interfaces, focusing on hazardous maintenance activities. A critical aspect of this objective was to investigate if a telepresence

framework could effectively install a sense of presence in a virtual setting and enable straightforward management of an industrial robotic arm's digital representation. Special attention was given to evaluating the ease of interaction between operators and the robotic arm, highlighting the role of communication protocols and system architecture in ensuring fluid and intuitive control. Central to the framework is its adaptability and compatibility with a wide range of hardware and software designed to accommodate various robotic systems and user interfaces without necessitating significant modifications. This adaptability lays the groundwork for future technological integrations, contributing to a dynamic and sustainable telepresence ecosystem.

The investigation of the third objective within this thesis entailed a meticulous evaluation of immersive (VR) and non-immersive (2D) interface modalities to determine their efficacy in conducting remote visual inspection tasks. Using a human-centred research approach, this segment focused on evaluating critical aspects such as presence, workload, usability, and performance metrics. The aim was to identify which interface modality significantly elevates the user experience and facilitates telepresence capabilities within remote maintenance environments. A key strategy employed in this investigation was the integration of gamification within task design to bolster participant engagement and motivation. Participants were tasked with conducting inspection activities on various engine components, employing VR and 2D interfaces. This study sought to illuminate the interfaces' effectiveness in remote visual inspection scenarios by closely monitoring performance metrics and evaluating human factors like presence, workload, and usability within a simulated maintenance context. The participant pool comprised 24 volunteers from Grenoble INP, spanning a diverse range of experiences with VR and robotics.

In addressing the fourth objective of this dissertation, a comprehensive investigation was conducted to explore the effects of different data visualisation modalities on the cognitive workload of maintainers during remote maintenance tasks. By integrating various visualisation techniques—including 2D feeds, 3D point clouds, augmented reality, and mixed modalities—this study assessed

their impact on the operator's telepresence capability, focusing on task performance and the sense of presence. The primary goal was to identify which visualisation modalities could optimise cognitive workload without compromising the efficiency and safety of remote maintenance tasks. The research methodology encompassed an experimental setup at the Centre for Digital Engineering and Manufacturing of Cranfield University, using advanced technologies such as the Fanuc M-20iA robot and Varjo XR-3 HMD. This setup facilitated a realistic simulation environment for conducting the experiments. A total of 28 participants were initially involved, with the final analysis focusing on 24 participants after excluding data based on baseline pupillometry measurements to ensure the accuracy of cognitive workload assessments. Data collection methods included pupillometry and employing Varjo's HMD API for precise pupil size tracking. This approach, combined with a detailed preprocessing pipeline, ensured the reliability of the cognitive workload data. Additionally, a 5-point scoring framework evaluated participants' performance across various parameters, precisely measuring task effectiveness under different visualisation modalities.

8.3 Answering the research questions

The research described within this thesis aimed to develop a repeatable process for designing and deploying telepresence capability that enhances the sense of presence in remote maintenance contexts. The primary outcome of the studies undertaken was the formulation of evidence-based recommendations for implementing immersive technologies to control industrial robots intended for remote inspection tasks. The two experiments comparing the 2D / VR interfaces and the evaluation of data visualisation modalities were performed primarily to measure human factors regarding the sense of presence, cognitive load, and usability. Performance was also measured when interacting with the various proposed interface designs under varying environmental conditions. This was achieved through a human-centred technology assessment, the proposed system architecture, experiments with participants, and the use of VR in hybrid

simulation settings. The three sub-research questions were defined at the beginning of the study.

8.3.1 Research question 1

The first question was **RQ1: Can a scalable telexistence-based framework provide seamless interaction with an industrial robotic arm?** To address this question, Chapter 4, reinforced by insights from Chapter 2's literature review on VR and telexistence, establishes a foundational understanding of the elements integral to a telexistence system. This includes everything from the actual environment in which the robot and machinery operate to the immersive interfaces that facilitate human-machine interactions. A pivotal factor in achieving this seamless interaction is bi-directional communication, ensuring a continuous and real-time connection between the virtual representation and its physical counterpart. Chapter 4 delves into the importance of such communication, highlighting the critical role of the Data Distribution Service protocol. Selected for its aptitude in sharing sensor data and guaranteeing real-time interactions, this protocol eases seamless interactions.

Furthermore, Chapter 4 illustrates how real-time digital representation can be leveraged to create an environment conducive to telexistence. This digital representation grants operators immediate visual feedback and empowers them with interactive capabilities, facilitating telexistence capability. The framework underwent rigorous validation using a physical FANUC industrial robot, thus proving its practical applicability. Performance metrics, key indicators of the system's efficiency and efficacy, were subsequently evaluated to cement the answer to RQ1 further. Interacting seamlessly with industrial robotic arms via a telexistence framework significantly enhances operational efficiency. In complex and dynamic industrial environments, the agility and precision of such interactions can lead to considerable improvements in task completion times and quality of outcomes. This is particularly crucial in tasks that require fine manipulation or are in hard-to-reach areas. Telexistence systems enable operators to perform tasks remotely, a significant advantage in hazardous environments or situations where human presence could be dangerous. By

allowing operators to control robotic arms from a safe distance, these systems minimise the risk of accidents and injuries, promoting a safer workplace.

8.3.2 Research question 2

The second question was **RQ2: How does the interface choice impact teleexistence capability during the remote visual inspection?** This was addressed through a detailed comparative study of non-immersive (2D) and immersive (VR) interfaces. This investigation revealed user experience and task performance distinctions based on the interface modality. The immersive VR interface consistently outperformed the 2D interface in several key aspects. Firstly, it provided a higher sense of presence, which is crucial for effective teleexistence. Users reported feeling more engaged and 'present' in the virtual environment, enhancing their connection with the remote task. This heightened presence, facilitated by the immersive qualities of VR, is fundamental for tasks requiring precise and careful inspection, as is often the case in remote maintenance scenarios. Secondly, the study indicated differences in cognitive workload between the interfaces. While the VR interface demanded more physical engagement, it did not increase mental or temporal demands compared to the 2D interface. This is noteworthy, as it suggests that the immersive nature of VR does not necessarily translate to increased cognitive strain, a critical factor for tasks that require sustained attention and precision. Usability was another key factor, and the VR interface showed a distinct advantage. Despite some physical discomfort reported by users, the overall usability of the VR system was rated closer to the 'relatively good usability' benchmark. This indicates that VR's intuitive and engaging nature can enhance user experience, making complex tasks more manageable and less intimidating. Performance metrics further underscored the benefits of the VR interface. Task completion times were notably shorter, and error rates were lower in the VR environment. This suggests that VR's intuitive design and immersive experience can lead to more efficient and accurate task execution, which is essential in high-stakes maintenance tasks. Participants' feedback echoed these findings, highlighting the immersive and intuitive nature of the VR interface as a key advantage. While the VR system had

some drawbacks, such as physical discomfort and the potential for cognitive strain over extended periods, its benefits in engagement, control, and task efficiency were clear. Finally, the choice of interface impacts telexistence capability during remote visual inspection. The immersive VR interface offers a more engaging, efficient, and accurate experience than traditional 2D interfaces, making it a superior choice for remote maintenance tasks.

8.3.3 Research question 3

The third question was **RQ3: How does data visualisation impact the telexistence capability of the maintainer at a cognitive workload level?** This was explored through an experiment involving four distinct data visualisation modalities in VR-based industrial robot teleoperation. The investigation focused on assessing the cognitive workload associated with these modalities and understanding their influence on task performance. The study demonstrated that different data visualisation modalities have varying impacts on cognitive workload. The findings revealed that the 3D live point cloud modality imposed the highest cognitive demand on participants, as indicated by the increase in pupil dilation. This suggests that while 3D visualisation provides rich spatial information, it also requires more cognitive processing. On the other hand, the combined 2D and 3D modality and the Augmented Reality (AR) modality presented a more balanced cognitive load, pointing toward the potential benefits of multi-modal data visualisation in reducing cognitive strain. Furthermore, task completion times varied across modalities, with the 2D live camera feed requiring longer completion times, indicating potential challenges or reduced intuitiveness in this singular modality. Conversely, the AR modality emerged as the most efficient in terms of completion time, suggesting an advantageous blend of visual elements for quicker task execution. In terms of performance scores, while each modality displayed high average scores, the analysis revealed that neither the type of visualisation modality nor the sequence of their presentation influenced these scores. This finding is particularly noteworthy as it suggests that, despite the inherent differences in cognitive workload and task completion times, the end performance of participants remained consistently high across all modalities. This

research provides valuable insights into the cognitive aspects of telexistence systems. It highlights the importance of considering how data visualisation techniques affect the cognitive workload of maintainers. The study underlines that while specific modalities may introduce higher cognitive demands, this does not necessarily translate to reduced performance. The nuanced understanding of these interactions between visualisation modalities, cognitive workload, and performance has implications for future telexistence systems, especially in ensuring they are both practical and cognitively manageable for users. In summary, the findings from this experiment address RQ3 by studying the relationship between data visualisation modalities and cognitive workload in telexistence systems.

References

Alqahtani, A.S., Daghestani, L.F. and Ibrahim, L.F. (2017) 'Semi-Immersive Virtual Reality for Improving the Mental Rotation Skill for Engineering Students: an Experimental Study', *Journal of Computer Engineering & Information Technology*, 2017(04) SciTechnol Available at: 10.4172/2324-9307.1000180 (Accessed: 14 May 2023).

Al-Sada, M., Jiang, K., Ranade, S., Kalkattawi, M. and Nakajima, T. (2020) 'HapticSnakes: multi-haptic feedback wearable robots for immersive virtual reality', *Virtual Reality*, 24(2), pp. 191–209.

Anon (n.d.) *The Factory of the Future: How AI, XR and Real-Time 3D Are Shaping Manufacturing* This whitepaper is sponsored by Unity. 2 THE FACTORY OF THE FUTURE: HOW AI, XR AND REAL-TIME 3D ARE SHAPING MANUFACTURING.

Avgousti, S., Christoforou, E.G., Panayides, A.S., Voskarides, S., Novales, C., Nouaille, L., Pattichis, C.S. and Vieyres, P. (2016) Medical telerobotic systems: Current status and future trends *BioMedical Engineering Online*. BioMed Central Ltd., Available at: 10.1186/s12938-016-0217-7 (Accessed: 7 April 2021).

Ballantyne, G.H. and Moll, F. (2003) 'The da Vinci telerobotic surgical system: the virtual operative field and telepresence surgery', *Surgical Clinics of North America*, 83(6), pp. 1293–1304.

Beck, L., Wolter, M., Mungard, N., Kuhlen, T. and Sturm, W. (2007) 'Combining Virtual Reality and Functional Magnetic Resonance Imaging (fMRI): Problems and Solutions', in *HCI and Usability for Medicine and Health Care*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 335–348.

Benmansour, N., Labayrade, R., Aubert, D., Glaser, S. and Gruyer, D. (2008) 'A model driven 3D lane detection system using stereovision', *2008 10th International Conference on Control, Automation, Robotics and Vision*. IEEE, pp. 1277–1282.

Benotsmane, R., Kovács, G. and Dudás, L. (2019) 'Economic, Social Impacts and Operation of Smart Factories in Industry 4.0 Focusing on Simulation and Artificial Intelligence of Collaborating Robots', *Social Sciences*, 8(5), p. 143.

Bienstock, C.C., Mentzer, J.T. and Bird, M.M. (1997) 'Measuring physical distribution service quality', *Journal of the Academy of Marketing Science*, 25(1), pp. 31–44.

Biocca, F. (2006) 'The Cyborg's Dilemma: Progressive Embodiment in Virtual Environments [1]', *Journal of Computer-Mediated Communication*, 3(2) Oxford University Press (OUP), pp. 0–0. Available at: 10.1111/J.1083-6101.1997.TB00070.X/4080399 (Accessed: 8 July 2023).

de Boer, T.A.B., de Winter, J.C.F. and Eisma, Y.B. (2023) 'Augmented reality-based telepresence in a robotic manipulation task: An experimental evaluation', *IET Collaborative Intelligent Manufacturing*, 5(4)

de Boissieu, F., Godin, C., Guilhamat, B., David, D., Serviere, C. and Baudois, D. (2009) 'Tactile texture recognition with a 3-axial force MEMS integrated artificial finger', *Robotics: Science and Systems V. Robotics: Science and Systems Foundation*.

Bowman, D.A. (1999) *Interaction techniques for common tasks in immersive virtual environments: design, evaluation, and application*. Georgia Institute of Technology.

Bric, J.D., Lumbard, D.C., Frelich, M.J. and Gould, J.C. (2016) Current state of virtual reality simulation in robotic surgery training: a review *Surgical Endoscopy*. Springer New York LLC, Available at: 10.1007/s00464-015-4517-y (Accessed: 28 April 2021).

Brooke, J. (1996) 'SUS: A "Quick and Dirty" Usability Scale', *Usability Evaluation In Industry*, CRC Press, pp. 207–212. Available at: 10.1201/9781498710411-35 (Accessed: 20 March 2023).

Buhler, H., Misztal, S. and Schild, J. (2018) 'Reducing VR Sickness Through Peripheral Visual Effects', *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, pp. 517–9.

Caldas, O.I., Aviles, O.F. and Rodriguez-Guerrero, C. (2020) 'Effects of Presence and Challenge Variations on Emotional Engagement in Immersive Virtual Environments', *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(5), pp. 1109–1116.

Card, S., K., R., G. G. and Mackinlay, J.D. (1991) 'The information visualizer: An information workspace', *ACM CHI'91*. New Orleans, LA, pp. 181–188.

Cruz-Neira, C., Sandin, D.J., DeFanti, T.A., Kenyon, R. V. and Hart, J.C. (1992) 'The CAVE: audio visual experience automatic virtual environment', *Communications of the ACM*, 35(6) Association for Computing Machinery, Inc., pp. 64–73. Available at: <https://go.gale.com/ps/i.do?p=AONE&sw=w&issn=00010782&v=2.1&it=r&id=G ALE%7CA12353475&sid=googleScholar&linkaccess=fulltext> (Accessed: 11 June 2023).

Diemer, J., Alpers, G.W., Peperkorn, H.M., Shiban, Y. and Mählberger, A. (2015) 'The impact of perception and presence on emotional reactions: a review of research in virtual reality', *Frontiers in Psychology*, 6

Draper, J. V., Kaber, D.B. and Usher, J.M. (1998) 'Telepresence', *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40(3), pp. 354–375.

Dwivedi, Y.K., Hughes, L., Baabdullah, A.M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M.M., Dennehy, D., Metri, B., Buhalis, D., Cheung, C.M.K., Conboy, K., Doyle, R., Dubey, R., Dutot, V., Felix, R., Goyal, D.P., Gustafsson, A., Hinsch, C., Jebabli, I., Janssen, M., Kim, Y.-G., Kim, J., Koos, S., Kreps, D., Kshetri, N., Kumar, V., Ooi, K.-B., Papagiannidis, S., Pappas, I.O., Polyviou, A., Park, S.-M., Pandey, N., Queiroz, M.M., Raman, R., Rauschnabel, P.A., Shirish, A., Sigala, M., Spanaki, K., Wei-Han Tan, G., Tiwari, M.K., Viglia, G. and Wamba, S.F. (2022) 'Metaverse beyond the hype: Multidisciplinary perspectives on emerging

challenges, opportunities, and agenda for research, practice and policy', *International Journal of Information Management*, 66 Pergamon, p. 102542. Available at: 10.1016/J.IJINFOMGT.2022.102542 (Accessed: 26 July 2022).

Eckstein, M.K., Guerra-Carrillo, B., Miller Singley, A.T. and Bunge, S.A. (2017) 'Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development?', *Developmental Cognitive Neuroscience*, 25, pp. 69–91.

Egashira, A., Horikawa, Y., Hayashi, T., Kawamura, A. and Kurazume, R. (2021) 'Near-future perception system: Previewed Reality', *Advanced Robotics*, 35(1), pp. 19–30.

Faul, F., Erdfelder, E., Lang, A.-G. and Buchner, A. (2007) 'G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences', *Behavior Research Methods*, 39(2), pp. 175–191.

Finn, A. and Scheduling, S. (2010) *Developments and Challenges for Autonomous Unmanned Vehicles*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Foley, J.D. (1995) *Computer Graphics: Principles and Practice in C (2nd Edition)*. Germany: Addison-Wesley.

Fong, T., Rochlis Zumbado, J., Currie, N., Mishkin, A. and Akin, D.L. (2013) 'Space Telerobotics', *Reviews of Human Factors and Ergonomics*, 9(1) SAGE PublicationsSage CA: Los Angeles, CA, pp. 6–56. Available at: 10.1177/1557234X13510679 (Accessed: 7 April 2021).

Fontaine, G. (1992) 'The Experience of a Sense of Presence in Intercultural and International Encounters', *Presence: Teleoperators and Virtual Environments*, 1(4), pp. 482–490.

Garg, G., Kuts, V. and Anbarjafari, G. (2021) 'Digital Twin for FANUC Robots: Industrial Robot Programming and Simulation Using Virtual Reality', *Sustainability 2021, Vol. 13, Page 10336*, 13(18) Multidisciplinary Digital Publishing Institute, p. 10336. Available at: 10.3390/SU131810336 (Accessed: 30 March 2022).

- Gaschler, A., Springer, M., Rickert, M. and Knoll, A. (2014) 'Intuitive robot tasks with augmented reality and virtual obstacles', *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 6026–6031.
- Gobbetti, E. and Scateni, R. (1998) '*Virtual reality: past, present, and future*' Available at: <http://www.p-arch.it/handle/11050/682> (Accessed: 14 May 2023).
- Grassini, S., Laumann, K., Thorp, S. and Topranin, V. de M. (2021) 'Using electrophysiological measures to evaluate the sense of presence in immersive virtual environments: An event-related potential study', *Brain and Behavior*, 11(8)
- Gray, S., Chevalier, R., Kotfis, D., Caimano, B., Chaney, K., Rubin, A., Fregene, K. and Danko, T. (2017) 'An Architecture for Human-Guided Autonomy: Team TROOPER at the DARPA Robotics Challenge Finals', *Journal of Field Robotics*, 34(5), pp. 852–873.
- Grieves, M. and Vickers, J. (2017) 'Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems', *Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches*, Springer, Cham, pp. 85–113. Available at: 10.1007/978-3-319-38756-7_4 (Accessed: 31 May 2022).
- Grieves, M.W. (2005) 'Product lifecycle management: the new paradigm for enterprises', *International Journal of Product Development*, 2(1/2), p. 71.
- Hagiwara, H. (2018) 'Trends in HRC and their effects on human operators' sense of presence in manufacturing settings', *Proceedings of the Human Factors and Ergonomics Society*, Vol.2.
- Han, S., Liu, B., Cabezas, R., Twigg, C.D., Zhang, P., Petkau, J., Yu, T.-H., Tai, C.-J., Akbay, M., Wang, Z., Nitzan, A., Dong, G., Ye, Y., Tao, L., Wan, C. and Wang, R. (2020) 'MEgATrack', *ACM Transactions on Graphics*, 39(4)
- Harris, D., Wilson, M. and Vine, S. (2020) 'Development and validation of a simulation workload measure: the simulation task load index (SIM-TLX)', *Virtual Reality*, 24(4)

Hayward, V., Astley, O.R., Cruz-Hernandez, M., Grant, D., and Robles-De-La-Torre, G. (2004) 'Haptic interfaces and devices', *Sensor Review*, 24(1), pp. 16–29.

Heeter, C. (1992) 'Being There: The Subjective Experience of Presence', *Presence: Teleoperators and Virtual Environments*, 1(2), pp. 262–271.

Heilig, M.L. (1962) 'Sensorama Simulator', *U.S. Patent No. 3,050,870*

Held, R.M. and Durlach, N.I. (1992) 'Telepresence', *Presence: Teleoperators and Virtual Environments*, 1(1), pp. 109–112.

Hicks, K., Gerling, K., Richardson, G., Pike, T., Burman, O. and Dickinson, P. (2019) 'Understanding the effects of gamification and juiciness on players', *IEEE Conference on Computational Intelligence and Games, CIG*, 2019-August IEEE Computer Society Available at: 10.1109/CIG.2019.8848105 (Accessed: 25 September 2023).

Hoeg, E.R., Ruder, K. V., Nilsson, N.C., Nordahl, R. and Serafin, S. (2017) 'An exploration of input conditions for virtual teleportation', *2017 IEEE Virtual Reality (VR)*. IEEE, pp. 341–342.

iGroup (2006) *igroup presence questionnaire (IPQ) overview* | *igroup.org – project consortium*. Available at: <http://www.igroup.org/pq/ipq/index.php> (Accessed: 4 July 2023).

Inoue, Y., Saraiji, M.Y., Kato, F. and Tachi, S. (2019) 'Enhancing bodily expression and communication capacity of telepresence robot with augmented reality', *21st International Symposium on Measurement and Control in Robotics.*, pp. 49–52.

Isabet, B., Pino, M., Lewis, M., Benveniste, S. and Rigaud, A.S. (2021) 'Social Telepresence Robots: A Narrative Review of Experiments Involving Older Adults before and during the COVID-19 Pandemic', *International Journal of Environmental Research and Public Health* 2021, Vol. 18, Page 3597, 18(7) Multidisciplinary Digital Publishing Institute, p. 3597. Available at: 10.3390/IJERPH18073597 (Accessed: 25 July 2022).

Jackson, I. and Sirois, S. (2009) 'Infant cognition: going full factorial with pupil dilation', *Developmental Science*, 12(4), pp. 670–679.

Jang, I., Carrasco, J., Weightman, A. and Lennox, B. (2019) 'Intuitive bare-hand teleoperation of a robotic manipulator using virtual reality and leap motion', *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer Verlag, Vol.11650 LNAI, pp. 283–294.

Jarrahi, M.H. (2018) 'Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making', *Business Horizons*, 61(4), pp. 577–586.

Jaselskis, E., Sankar, A., Yousif, A., Clark, B. and Chinta, V. (2015) 'Using Telepresence for Real-Time Monitoring of Construction Operations', *Journal of Management in Engineering*, 31(1)

Jason Jerald (2015) *Human-Centered Design for Virtual Reality*. Morgan & Claypool.

Johns, M., Mok, B., Sirkin, D., Gowda, N., Smith, C., Talamonti, W. and Ju, W. (2016) 'Exploring shared control in automated driving', *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, pp. 91–98.

Kamezaki, M., Yang, J., Iwata, H. and Sugano, S. (2016) 'Visibility Enhancement using Autonomous Multicamera Controls with Situational Role Assignment for Teleoperated Work Machines', *Journal of Field Robotics*, 33(6), pp. 802–824.

Kandemir, C. and Celik, M. (2021) 'Determining the error producing conditions in marine engineering maintenance and operations through HFACS-MMO', *Reliability Engineering & System Safety*, 206 Available at: 10.1016/j.ress.2020.107308 (Accessed: 2 December 2022).

Kang, W., Kapitanova, K. and Son, S.H. (2012) 'RDDS: A Real-Time Data Distribution Service for Cyber-Physical Systems', *IEEE Transactions on Industrial Informatics*, 8(2), pp. 393–405.

Kiryu, T. and So, R.H. (2007) 'Sensation of presence and cybersickness in applications of virtual reality for advanced rehabilitation', *Journal of NeuroEngineering and Rehabilitation*, 4(1), p. 34.

Kofman, J., Xianghai Wu, Luu, T.J. and Verma, S. (2005) 'Teleoperation of a robot manipulator using a vision-based human-robot interface', *IEEE Transactions on Industrial Electronics*, 52(5), pp. 1206–1219.

Koh, E. (2010) 'Conferencing room for telepresence with remote participants', *Proceedings of the 16th ACM International Conference on Supporting Group Work, GROUP'10*, , pp. 309–310. Available at: 10.1145/1880071.1880124 (Accessed: 25 July 2022).

Kohn, S., Blank, A., Puljiz, D., Zenkel, L., Bieber, O., Hein, B. and Franke, J. (2018) 'Towards a Real-Time Environment Reconstruction for VR-Based Teleoperation Through Model Segmentation', *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 1–9.

Korte, C., Sudhakaran Nair, S., Nistor, V., Low, T.P., Doarn, C.R. and Schaffner, G. (2014) 'Determining the Threshold of Time-Delay for Teleoperation Accuracy and Efficiency in Relation to Telesurgery', *Telemedicine and e-Health*, 20(12), pp. 1078–1086.

Kret, M.E. and Sjak-Shie, E.E. (2019) 'Preprocessing pupil size data: Guidelines and code', *Behavior Research Methods*, 51(3), pp. 1336–1342.

Leeper, A., Hsiao, K., Chu, E., and Salisbury, J.K. (2014) 'Using Near-Field Stereo Vision for Robotic Grasping in Cluttered Environments', in , pp. 253–267.

LI, Y., HUANG, J., TIAN, F., WANG, H.A. and DAI, G.Z. (2019) 'Gesture interaction in virtual reality', *Virtual Reality & Intelligent Hardware*, 1(1) Elsevier, pp. 84–112. Available at: 10.3724/SP.J.2096-5796.2018.0006 (Accessed: 27 July 2022).

Livatino, S., Guastella, D.C., Muscato, G., Rinaldi, V., Cantelli, L., Melita, C.D., Caniglia, A., Mazza, R. and Padula, G. (2021) 'Intuitive Robot Teleoperation

Through Multi-Sensor Informed Mixed Reality Visual Aids', *IEEE Access*, 9, pp. 25795–25808.

Lo, C.K., Chen, C.H. and Zhong, R.Y. (2021) 'A review of digital twin in product design and development', *Advanced Engineering Informatics*, 48 Elsevier Ltd Available at: 10.1016/J.AEI.2021.101297 (Accessed: 9 March 2022).

Lopez Pulgarin, E.J., Niu, H., Herrmann, G. and Carrasco, J. (2022) 'Implementing and Assessing a Remote Teleoperation Setup with a Digital Twin Using Cloud Networking', *Towards Autonomous Robotic Systems*. Springer Science and Business Media Deutschland GmbH, Vol.13546 LNAI, pp. 238–250. Available at: 10.1007/978-3-031-15908-4_19/COVER (Accessed: 9 January 2023).

Lorusso, M.L., Travellini, S., Giorgetti, M., Negrini, P., Reni, G. and Biffi, E. (2020) 'Semi-Immersive Virtual Reality as a Tool to Improve Cognitive and Social Abilities in Preschool Children', *Applied Sciences*, 10(8), p. 2948.

Ma, R. and Kaber, D.B. (2006) 'Presence, workload and performance effects of synthetic environment design factors', *International Journal of Human-Computer Studies*, 64(6), pp. 541–552.

Maciaś, M., Dąbrowski, A., Fraś, J., Karczewski, M., Puchalski, S., Tabaka, S. and Jaroszek, P. (2020) 'Measuring Performance in Robotic Teleoperation Tasks with Virtual Reality Headgear', *Advances in Intelligent Systems and Computing*., Vol.920.

Makhal, A., Thomas, F., and Gracia, A.P. (2017) 'Grasping Unknown Objects in Clutter by Superquadric Representation'

Malińska, M., Zużewicz, K., Bugajska, J. and Grabowski, A. (2015) 'Heart rate variability (HRV) during virtual reality immersion', *International Journal of Occupational Safety and Ergonomics*, 21(1), pp. 47–54.

Marino, D.L., Grandio, J., Wickramasinghe, C.S., Schroeder, K., Bourne, K., Filippas, A. V. and Manic, M. (2020) 'AI Augmentation for Trustworthy AI:

Augmented Robot Teleoperation', *2020 13th International Conference on Human System Interaction (HSI)*. IEEE, pp. 155–161.

Marvel, J.A., Falco, J. and Marstio, I. (2015) 'Characterizing Task-Based Human–Robot Collaboration Safety in Manufacturing', *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 45(2), pp. 260–275.

Mayoral-Vilches, V., Pinzger, M., Rass, S., Dieber, B. and Gil-Uriarte, E. (2020) 'Can ROS be used securely in industry? Red teaming ROS-Industrial' Available at: <https://arxiv.org/abs/2009.08211v1> (Accessed: 2 July 2023).

Mihelj, M., Novak, D. and Beguš, S. (2014) *Virtual Reality Technology and Applications*. Dordrecht: Springer Netherlands.

Mihelj, M. and Podobnik, J. (2012) *Haptics for Virtual Reality and Teleoperation*. Dordrecht: Springer Netherlands.

Miller, R.B. (1968) 'Response time in man-computer conversational transactions', *AFIPS Fall Joint Computer Conference.*, pp. 267–277.

Milleville-Pennel, I. and Charron, C. (2015) 'Do mental workload and presence experienced when driving a real car predispose drivers to simulator sickness? An exploratory study,' *Accident Analysis & Prevention*, 74, pp. 192–202.

Minsky, M. (1980) TELEPRESENCE, *OMNI magazine*,

Mizuchi, Y. and Inamura, T. (2017) 'Cloud-based multimodal human-robot interaction simulator utilizing ROS and unity frameworks', *2017 IEEE/SICE International Symposium on System Integration (SII)*. IEEE, pp. 948–955.

Moshayedi, A.J., Kolahdooz, A. and Liao, L. (2022) *Unity in Embedded System Design and Robotics: A Step-by-Step Guide*. Available at: <https://www.routledge.com/Unity-in-Embedded-System-Design-and-Robotics-A-Step-by-Step-Guide/Moshayedi-Kolahdooz-Liao/p/book/9781032205588> (Accessed: 5 March 2024).

Murphy, R.R., Nomura, T., Billard, A. and Burke, J.L. (2010) 'Human–Robot Interaction', *IEEE Robotics & Automation Magazine*, 17(2), pp. 85–89.

Mütterlein, J. (2018) 'The Three Pillars of Virtual Reality? Investigating the Roles of Immersion, Presence, and Interactivity,' *Proceedings of the 51st Hawaii International Conference on System Sciences*. Hawaii International Conference on System Sciences. Available at: 10.24251/HICSS.2018.174 (Accessed: 28 April 2021).

N. H. P. R. Group et al. (1987) *Task load index (nasa-tlx) v1. 0 computerised version*.

Naceri, A., Mazzanti, D., Bimbo, J., Tefera, Y.T., Prattichizzo, D., Caldwell, D.G., Mattos, L.S. and Deshpande, N. (2021) 'The Vicarios Virtual Reality Interface for Remote Robotic Teleoperation: Teleporting for Intuitive Tele-manipulation', *Journal of Intelligent and Robotic Systems: Theory and Applications*, 101(4)

Nass, C., Steuer, J. and Siminoff, E. (1994) 'Computer are social actors', *Conference on Human Factors in Computing Systems - Proceedings.*, p. 204.

Nikolaidis, S., Zhu, Y.X., Hsu, D. and Srinivasa, S. (2017) '*Human-Robot Mutual Adaptation in Shared Autonomy*'

Nostadt, N., Abbink, D.A., Christ, O. and Beckerle, P. (2020) Embodiment, presence, and their intersections *ACM Transactions on Human-Robot Interaction*.

Nuzzi, C., Ghidini, S., Pagani, R., Pasinetti, S., Coffetti, G. and Sansoni, G. (2020) 'Hands-Free: a robot augmented reality teleoperation system', *2020 17th International Conference on Ubiquitous Robots (UR)*. IEEE, pp. 617–624.

Okura, F., Ueda, Y., Sato, T. and Yokoya, N. (2014) '[Paper] Free-viewpoint Mobile Robot Teleoperation Interface Using View-dependent Geometry and Texture', *ITE Transactions on Media Technology and Applications*, 2(1), pp. 82–93.

Olaverri-Monreal, C., Krizek, G.Ch., Michaeler, F., Lorenz, R. and Pichler, M. (2019) 'Collaborative approach for a safe driving distance using stereoscopic image processing', *Future Generation Computer Systems*, 95, pp. 880–889.

Omarali, B., Denoun, B., Althoefer, K., Jamone, L., Valle, M. and Farkhatdinov, I. (2020) 'Virtual Reality based Telerobotics Framework with Depth Cameras', *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, pp. 1217–1222.

Oyama, E., Tokoi, K., Suzuki, R., Nakamura, S., Shiroma, N., Watanabe, N., Agah, A., Okada, H. and Omori, T. (2021) 'Augmented reality and mixed reality behavior navigation system for telepresence remote assistance', *Advanced Robotics*, 35(20), pp. 1223–1241.

Pakanen, M., Alavesä, P., van Berckel, N., Koskela, T. and Ojala, T. (2022) "Nice to see you virtually": Thoughtful design and evaluation of virtual avatar of the other user in AR and VR based telepresence systems', *Entertainment Computing*, 40, p. 100457.

Park, C. and Jang, K. (2019) 'Investigation of Visual Self-Representation for a Walking-in-Place Navigation System in Virtual Reality', *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, pp. 1114–1115.

Petroski, K. (2018) 'Virtual realities', in *Fiction and the Languages of Law*. New York, NY: Routledge, 2018. | Series: Law, language, and communication: Routledge, pp. 121–150.

Piolain, G., Geffard, F., Coudray, A., Garrec, P., Thro, J.F. and Perrot, Y. (2010) 'Dedicated and standard industrial robots used as force-feedback telepresence remote devices at the AREVA recycling plant', *2010 1st International Conference on Applied Robotics for the Power Industry, CARPI 2010*. Available at: 10.1109/CARPI.2010.5624418 (Accessed: 7 April 2021).

Poupyrev, I., Ichikawa, T., Weghorst, S. and Billinghurst, M. (1998) 'Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques', *Computer Graphics Forum*, 17(3), pp. 41–52.

Pour Rahimian, F., Seyedzadeh, S., Oliver, S., Rodriguez, S. and Dawood, N. (2020) 'On-demand monitoring of construction projects through a game-like

hybrid application of BIM and machine learning', *Automation in Construction*, 110, p. 103012.

Power, S.E. (2016) 'Relative and Absolute Temporal Presence', in *Philosophy and Psychology of Time*. Cham: Springer International Publishing, pp. 69–100.

Prithul, A., Adhanom, I.B. and Folmer, E. (2021) 'Teleportation in Virtual Reality; A Mini-Review', *Frontiers in Virtual Reality*, 2

Qu, J., Zhang, Y. and Bu, L. (2023) 'Functional near-infrared spectroscopy in the assessment of rehabilitation efficacy of virtual reality products for people with cognitive disorders', *International Journal of Industrial Ergonomics*, 97, p. 103500.

Quintero, C.P., Li, S., Pan, M.K., Chan, W.P., Machiel Van der Loos, H.F. and Croft, E. (2018) 'Robot Programming Through Augmented Trajectories in Augmented Reality', *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 1838–1844.

Ramacciotti, M., Milazzo, M., Leoni, F., Roccella, S. and Stefanini, C. (2016) 'A novel shared control algorithm for industrial robots', *International Journal of Advanced Robotic Systems*, 13(6), p. 172988141668270.

Ramezani, M., Brandao, M., Casseau, B., Havoutis, I. and Fallon, M. (2020) 'Legged Robots for Autonomous Inspection and Monitoring of Offshore Assets', *Day 1 Mon, May 04, 2020*. OTC.

Rauschnabel, P.A., Felix, R., Hinsch, C., Shahab, H. and Alt, F. (2022) 'What is XR? Towards a Framework for Augmented and Virtual Reality,' *Computers in Human Behavior*, 133 Pergamon, p. 107289. Available at: 10.1016/J.CHB.2022.107289 (Accessed: 26 July 2022).

Rebouças Filho, P.P., P. da Silva, S.P., Praxedes, V.N., Hemanth, J. and de Albuquerque, V.H.C. (2019) 'Control of singularity trajectory tracking for robotic manipulator by genetic algorithms', *Journal of Computational Science*, 30, pp. 55–64.

Reiner, M. and Hecht, D. (2009) 'Behavioral Indications of Object-Presence in Haptic Virtual Environments', *CyberPsychology & Behavior*, 12(2), pp. 183–186.

roboception (2021) *FANUC XYZ-WPR format*. Available at: https://doc.rc-visard.com/v21.07/en/pose_format_fanuc.html?highlight=fanuc (Accessed: 22 November 2023).

Roldán, J., Peña-Tapia, E., Martín-Barrio, A., Olivares-Méndez, M., Del Cerro, J. and Barrientos, A. (2017) 'Multi-Robot Interfaces and Operator Situational Awareness: Study of the Impact of Immersion and Prediction', *Sensors*, 17(8), p. 1720.

Rosen, R., Von Wichert, G., Lo, G. and Bettenhausen, K.D. (2015) 'About The Importance of Autonomy and Digital Twins for the Future of Manufacturing', *IFAC-PapersOnLine*, 48(3) Elsevier, pp. 567–572. Available at: 10.1016/J.IFACOL.2015.06.141 (Accessed: 9 January 2023).

Rusu, R.B. and Cousins, S. (2011) '3D is here: Point Cloud Library (PCL)', *2011 IEEE International Conference on Robotics and Automation*. IEEE, pp. 1–4.

Saakes, D., Choudhary, V., Sakamoto, D., Inami, M. and Lgarashi, T. (2013) 'A teleoperating interface for ground vehicles using autonomous flying cameras', *2013 23rd International Conference on Artificial Reality and Telexistence (ICAT)*. IEEE, pp. 13–19.

El Saddik, A. (2018) 'Digital Twins: The Convergence of Multimedia Technologies', *IEEE Multimedia*, 25(2) IEEE Computer Society, pp. 87–92. Available at: 10.1109/MMUL.2018.023121167 (Accessed: 31 May 2022).

Sanchez-Vives, M. V. and Slater, M. (2005) 'From presence to consciousness through virtual reality', *Nature Reviews Neuroscience*, 6(4), pp. 332–339.

Santos Pessoa de Melo, M., Gomes da Silva Neto, J., Jorge Lima da Silva, P., Natario Teixeira, J.M.X. and Teichrieb, V. (2019) 'Analysis and Comparison of Robotics 3D Simulators', *2019 21st Symposium on Virtual and Augmented Reality (SVR)*. IEEE, pp. 242–251.

Schaik, P. Van, Turnbull, T., Wersch, A. Van and Drummond, S. (2004) 'Presence Within a Mixed Reality Environment', *CyberPsychology & Behavior*, 7(5), pp. 540–552.

Schroeder, R. (2006) 'Being There Together and the Future of Connected Presence', *Presence: Teleoperators and Virtual Environments*, 15(4), pp. 438–454.

Schultze, U. (2010) 'Embodiment and presence in virtual worlds: a review', *Journal of Information Technology*, 25(4), pp. 434–449.

Schwind, V., Knierim, P., Haas, N. and Henze, N. (2019) 'Using Presence Questionnaires in Virtual Reality', *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM, pp. 1–12.

Sean, B., Paul, U., Noah, S. and Kavita, B. (2015) 'Material Recognition in the Wild with the Materials in Context Database', *Computer Vision and Pattern Recognition*

Senft, E., Hagenow, M., Welsh, K., Radwin, R., Zinn, M., Gleicher, M. and Mutlu, B. (2021) 'Task-Level Authoring for Remote Robot Teleoperation', *Frontiers in Robotics and AI*, 8

Serebrenniy, V. V., Lapin, D. V. and Mokaeva, A.A. (2019) 'Concept of multifunctional modular mobile robotic complex for marine vessels maintenance', , p. 190003.

Shadish, W.R., Cook, T.D. and Campbell, D.T. (2002) *Experimental and quasi-experimental designs for generalized causal inference*. Houghton, Mifflin, and Company.

Sherman, W.R., and Craig, A.B. (2003) *Understanding Virtual Reality: Interface, Application, and Design*. Elsevier Inc. Available at: 10.1162/105474603322391668 (Accessed: 21 July 2022).

Shin, M., Lee, S., Song, S.W. and Chung, D. (2021) 'Enhancement of perceived body ownership in virtual reality-based teleoperation may backfire in the execution of high-risk tasks', *Computers in Human Behavior*, 115

Shishehgar, M., Kerr, D. and Blake, J. (2019) 'The effectiveness of various robotic technologies in assisting older adults', *Health informatics journal*, 25(3) Health Informatics J, pp. 892–918. Available at: 10.1177/1460458217729729 (Accessed: 25 July 2022).

Shneiderman, B., Plaisant, C., Cohen, M. and Jacobs, S. (1997) *Designing the user interface: strategies for effective human-computer interaction*. 5th edn. Addison-Wesley Longman Publishing Co. Available at: Must specify DOI or URL if using the accessed field (Accessed: 25 September 2023).

Slater, M. (2003) 'A note on presence terminology',

Slater, M. and Sanchez-Vives, M. V. (2014) 'Transcending the Self in Immersive Virtual Reality', *Computer*, 47(7), pp. 24–30.

Slater, M. and Wilbur, S. (1997) 'A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments', *Presence: Teleoperators and Virtual Environments*, 6(6), pp. 603–616.

Statista (2023) *VR headset unit sales worldwide 2024* | Statista. Available at: <https://www.statista.com/statistics/677096/vr-headsets-worldwide/> (Accessed: 11 June 2023).

Steuer, J. (1992) 'Defining Virtual Reality: Dimensions Determining Telepresence', *Journal of Communication*, 42(4), pp. 73–93.

Stone, R.J. (2012) *Human Factors Guidance for Designers of Interactive 3D and Games-Based Training Systems* *Human Factors Integration Defence Technology Centre Publication*.

Stotko, P., Krumpen, S., Schwarz, M., Lenz, C., Behnke, S., Klein, R. and Weinmann, M. (2019) 'A VR System for Immersive Teleoperation and Live

Exploration with a Mobile Robot', *IEEE International Conference on Intelligent Robots and Systems*, Institute of Electrical and Electronics Engineers Inc., pp. 3630–3637. Available at: 10.1109/IROS40897.2019.8968598 (Accessed: 6 April 2022).

Strack, F. (1992) “Order Effects” in Survey Research: Activation and Information Functions of Preceding Questions’, in *Context Effects in Social and Psychological Research*. New York, NY: Springer New York, pp. 23–34.

Stuart, Rory. (1996) ‘*The design of virtual environments*’, McGraw-Hill, p. 274. Available at: Must specify DOI or URL if using the accessed field (Accessed: 27 July 2022).

Sutherland (1968) ‘*HEAD-MOUNTED THREE DIMENSIONAL DISPLAY*’, 33(pt 1), pp. 757–764. Available at: 10.1145/1476589.1476686 (Accessed: 26 July 2022).

Tachi, S. (2015a) ‘*Telexistence*’, in , pp. 229–259.

Tachi, S. (2015b) ‘*Telexistence*’, *Telexistence*, WORLD SCIENTIFIC Available at: 10.1142/9248 (Accessed: 25 July 2022).

Tachi, S. (2015c) ‘*Telexistence*’, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8844 Springer, Cham, pp. 229–259. Available at: 10.1007/978-3-319-17043-5_13 (Accessed: 7 January 2022).

Tan, Y., Niu, C., Tian, H. and Zhang, J. (2021) ‘A Digital Twin Based Design of the Semi-physical Marine Engine Room Simulator for Remote Maintenance Assistance’, *Image and Signal Processing*. Institute of Electrical and Electronics Engineers Inc., pp. 137–141. Available at: 10.1109/ICVISIP54630.2021.00032 (Accessed: 8 January 2023).

Thomason, J., Ratsamee, P., Kiyokawa, K., Kriangkamol, P., Orlosky, J., Mashita, T., Uranishi, Y. and Takemura, H. (2017) ‘Adaptive View Management for Drone Teleoperation in Complex 3D Structures’, *Proceedings of the 22nd*

International Conference on Intelligent User Interfaces. New York, NY, USA: ACM, pp. 419–426.

Toda, A.M., Klock, A.C.T., Oliveira, W., Palomino, P.T., Rodrigues, L., Shi, L., Bittencourt, I., Gasparini, I., Isotani, S. and Cristea, A.I. (2019) ‘Analysing gamification elements in educational environments using an existing Gamification taxonomy’, *Smart Learning Environments* 2019 6:1, 6(1) SpringerOpen, pp. 1–14. Available at: 10.1186/S40561-019-0106-1 (Accessed: 25 September 2023).

Treiblmaier, H., Putz, L.-M. and Lowry, P. (2018) ‘Setting a definition, context, and theory-based research agenda for the gamification of non-gaming applications.’, in *AIS Trans. Human-Comput. Interactions.* , pp. 129–163.

Tremmel, C., Herff, C., Sato, T., Rechowicz, K., Yamani, Y. and Krusienski, D.J. (2019) ‘Estimating Cognitive Workload in an Interactive Virtual Reality Environment Using EEG’, *Frontiers in Human Neuroscience*, 13

Tsui, K.M., Desai, M., Yanco, H.A. and Uhlik, C. (2011a) ‘Exploring use cases for telepresence robots’, *Proceedings of the 6th international conference on Human-robot interaction*. New York, NY, USA: ACM, pp. 11–18.

Tsui, K.M., Desai, M., Yanco, H.A. and Uhlik, C. (2011b) *Exploring Use Cases for Telepresence Robots*. Available at: Must specify DOI or URL if using the accessed field (Accessed: 28 April 2021).

Universal Robots (2022) *Universal Robots - DH Parameters for calculations of kinematics and dynamics*. Available at: <https://www.universal-robots.com/articles/ur/application-installation/dh-parameters-for-calculations-of-kinematics-and-dynamics/> (Accessed: 25 September 2023).

van 't Wout, M., Spofford, C.M., Unger, W.S., Sevin, E.B. and Shea, M.T. (2017) ‘Skin Conductance Reactivity to Standardized Virtual Reality Combat Scenes in Veterans with PTSD’, *Applied Psychophysiology and Biofeedback*, 42(3), pp. 209–221.

Varjo (2023) *Varjo XR-3 - The industry's highest resolution mixed reality headset* | Varjo. Available at: <https://varjo.com/products/xr-3/> (Accessed: 15 July 2023).

Villani, V., Pini, F., Leali, F. and Secchi, C. (2018) 'Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications', *Mechatronics*, 55

Wagner, J., Stuerzlinger, W. and Nedel, L. (2021) 'Comparing and Combining Virtual Hand and Virtual Ray Pointer Interactions for Data Manipulation in Immersive Analytics', *IEEE Transactions on Visualization and Computer Graphics*, 27(5), pp. 2513–2523.

Walker, M.E., Hedayati, H. and Szafir, D. (2019) 'Robot Teleoperation with Augmented Reality Virtual Surrogates', *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, pp. 202–210.

Wang, L., Li, Q., Lam, J., Wang, Z. and Zhang, Z. (2022) 'Intent inference in shared-control teleoperation system in consideration of user behavior', *Complex & Intelligent Systems*, 8(4), pp. 2971–2981.

Wang, Z., Reed, I. and Fey, A.M. (2018) 'Toward Intuitive Teleoperation in Surgery: Human-Centric Evaluation of Teleoperation Algorithms for Robotic Needle Steering', *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 5799–5806.

Wei, D., Huang, B. and Li, Q. (2021) 'Multi-View Merging for Robot Teleoperation With Virtual Reality', *IEEE Robotics and Automation Letters*, 6(4), pp. 8537–8544.

Whitney, D., Rosen, E., Ullman, D., Phillips, E. and Tellex, S. (2018a) 'ROS Reality: A Virtual Reality Framework Using Consumer-Grade Hardware for ROS-Enabled Robots', *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 1–9.

Whitney, D., Rosen, E., Ullman, D., Phillips, E. and Tellex, S. (2018b) 'ROS Reality: A Virtual Reality Framework Using Consumer-Grade Hardware for ROS-Enabled Robots', *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 1–9.

Wickens, C.D., Hollands, J.G., Banbury, S. and Parasuraman, R. (2015) 'Engineering Psychology and Human Performance', *Engineering Psychology and Human Performance*, Psychology Press Available at: 10.4324/9781315665177 (Accessed: 25 September 2023).

Witmer, B.G. and Singer, M.J. (1998) 'Measuring presence in virtual environments: A presence questionnaire', *Presence: Teleoperators and Virtual Environments*, 7(3) MIT Press Journals, pp. 225–240. Available at: 10.1162/105474698565686 (Accessed: 27 July 2022).

Wonsick, M. and Padir, T. (2020) 'A Systematic Review of Virtual Reality Interfaces for Controlling and Interacting with Robots', *Applied Sciences*, 10(24), p. 9051.

Wu, H., Houben, G.-J. and De Bra, P. (2000) *Supporting user adaptation in adaptive hypermedia applications* Technische Universiteit Eindhoven, Available at: <https://research.tue.nl/en/publications/supporting-user-adaptation-in-adaptive-hypermedia-applications> (Accessed: 14 May 2023).

www.gov.uk (2022) *Dstl grasps telexistence potential to reduce risk to personnel* - GOV.UK. Available at: <https://www.gov.uk/government/news/dstl-grasps-telexistence-potential-to-reduce-risk-to-personnel#full-publication-update-history> (Accessed: 18 May 2023).

Xi, B., Wang, S., Ye, X., Cai, Y., Lu, T. and Wang, R. (2019) 'A robotic shared control teleoperation method based on learning from demonstrations', *International Journal of Advanced Robotic Systems*, 16(4), p. 172988141985742.

Xiang, Z., Zhang, Q., Zhu, W., Zhang, Z. and Zhang, Y.-Q. (2004) 'Peer-to-Peer Based Multimedia Distribution Service', *IEEE Transactions on Multimedia*, 6(2), pp. 343–355.

Yanco, H.A. and Drury, J. (2004) 'Classifying human-robot interaction: An updated taxonomy', *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, 3, pp. 2841–2846. Available at: 10.1109/ICSMC.2004.1400763 (Accessed: 27 July 2022).

Zhou, T., Zhu, Q. and Du, J. (2020) 'Intuitive robot teleoperation for civil engineering operations with virtual reality and deep learning scene reconstruction', *Advanced Engineering Informatics*, 46, p. 101170.

Appendices

Appendix A Ethical approval letter



14 March 2023

Dear Mr Mazeas ,

Reference: CURES/17958/2023

Project ID: 20805

Title: Telexistence capability in remote visual machinery inspection: Evaluation in virtual and real settings

Thank you for your application to the Cranfield University Research Ethics System (CURES).

We are pleased to inform you your CURES application, reference CURES/17958/2023 has been reviewed. You may now proceed with the research activities you have sought approval for.

If you have any queries, please contact CURES Support.

We wish you every success with your project.

Regards,

CURES Team

Appendix B Expert 1 questionnaire responses

1. Number of years of experience in the field of marine engineering/maintenance?

31. BEng in Ship Science (Southampton University's naval architecture degree) and MSc in Marine Engineering, also at Southampton, graduating in 1997.

2. Past/Current responsibilities?

Current and recent experience (at UNCLASSIFIED): Marine engineer in Defence Science & Technology Laboratory (Dstl), part of the United Kingdom Ministry of Defence. I primarily work on ship and small craft concept designs and technology studies into related equipment. Studies into implementing modern and advanced maintenance and repair techniques and technologies for Royal Navy (RN) surface ships and implementation in other environments. Input to various RN policy and technical initiatives in related areas.

2012-2016: UK Exchange Marine Engineer/Naval Architect – US Navy Surface Warfare Centre Carderock (one of the USN's naval architecture research facilities, in Maryland, US)

Elsewhere, I analysed various naval warfare areas, program shaping and researched customers for maritime and air logistics research (including maintenance & repair issues), sonar engineer and trials officer, and in the early 90s, engine room watchkeeper and motorman on a civilian merchant vessel.

3. What does a typical day on the job look like for you as a marine engineer in charge of maintenance?

Never having overseen maintenance at sea, I will try to reconstruct from what I saw my Chief Engineer at sea do and what I have observed from Royal Navy engineers.

The marine engineer in charge of maintenance is also usually the one leading the department, so dealing with personnel issues, setting up rosters, understanding policy implications from higher up the chain, dealing with suppliers and so on.

Work with logged machinery performance data and observations from watchkeepers at sea to understand imminent failures or systems needing further investigation. Balance this against scheduled maintenance tasks (generated by running hours or the system's life), etc. Prioritise work schedules. Understand the effect of system outages (either caused by taking systems offline for maintenance or where they cannot be fixed at sea and need to either head to port or a dry-dock) and how this will impact the performance and capability of the ship as a whole; will the ship be able to carry out its mission given current or likely system failures?

Additionally, the Chief Engineer will conduct their walk-around of the engine room and machinery space to maintain situational awareness and maintain standards.

4. What everyday maintenance tasks do you perform in the machine room?

Daily tasks include engine room watchkeeping (every 4 hours at sea on my ship), where you inspect everything in the engine room and machinery spaces and log all data values if not recorded automatically.

Beyond that, keeping everything clean and tidy (the absence of which usually correlates with a fire or health & safety risk), tightening up weeping pipe joints, replacing blown gaskets, cleaning fuel and seawater filters (a daily task, can be more in adverse environments) and other general maintenance.

Conducting/assisting with scheduled or directed maintenance tasks as needed.

5. How do you diagnose and troubleshoot problems with the engines and other machinery in the engine marine room?

Often from experience in your engine room and spotting that something has changed (sound, smell, the temperature of a surface, fluid accumulating where there should not be fluid, etc). This raises concerns and often leads to investigation by the watchkeeper or other, more experienced engineers.

6. What tools and equipment do you use in your work? How do you ensure that these tools are properly maintained and calibrated?

On my ship, hand and power tools, welders, etc. Light machine tools (lathes, bench pillar drills etc). Common sense maintenance of tools (clean up and wipe down to dry after use, etc.) ensures longer life and a lower likelihood of tool failure.

For measurement, micrometres, tape measures, and ultrasonic probes for plate thickness. I do not recall ever seeing a calibration certificate; I suspect the RN take this far more seriously.

7. What safety precautions do you take when working in the engine marine room?

Firstly, always know who is in the engine room. Most ships have a board with several hooks in different areas, and each person has a disc with a distinct number. On entering the engine room, you move your disc to the correct area to show you are in the engine room. In this way, officers, rescue teams, etc., can see who is in the engine room briefly.

Secondly, the buddy system. If you are working in confined spaces, tanks, etc, always have someone else aware of what you are doing and able to take remedial action if you get into trouble.

Appropriate PPE, including fire-resistant overalls, safety boots, hard hats, etc.

Appropriate training and certification. This includes having properly trained fire and personnel rescue teams and confidence in working with breathing apparatus, fire safety equipment, etc.

8. What training and experience are required for a marine engineer to be proficient in the maintenance of the engine marine room?

Appropriate classroom and workshop training onshore before going to sea to gain the appropriate basic skills, followed by mentoring from more experienced personnel while carrying out their roles and more advanced study.

9. What are some challenges that you face in your job, and how do you overcome them?

When designing a ship, everything is a trade-off, not least constrained by a limited budget. In my current roles, I would say that the main challenge is the preference for operational capabilities over design for better MBTF and improved maintenance and repair.

10. Are there any new technologies or practices you see adopting to maintain marine engines and machinery in the future?

For me, the next frontier is the development of subsystem Digital Twins at a sufficient level to allow both warnings of imminent failure or issues and conduct 'what if' games to see the effect on system performance and failure of proposed operations, operating regimes, and profiles, etc.

11. What are the most important factors to consider when designing a user interface for maintenance on a marine vessel?

When needed, give the user the correct information in a format they can use. Firstly, users will take a while to gain confidence in the system, so it must be as helpful and correct as possible before being rolled out to sea. Also, as the amount of data produced by naval vessel systems goes up dramatically, the data needs to be processed and presented to the user without drowning them.

12. How do maintenance needs and requirements differ across different types of marine vessels?

Naval vs non-naval vessels: If something goes very wrong on a civilian vessel, the priority is to ensure the safety of personnel by evacuation if necessary. On a naval vessel, the priority is to continue carrying out the mission, so more risk may be taken to achieve that.

Also, different ship types will have different systems, which may need different approaches, tools, etc.

13. How do you envision the user interface being used in practice? What kind of information and functionality should it provide?

For example, underway management. The provision of operational guidance to a platform while underway, in peacetime and wartime conditions, via awareness of expected versus actual Hull, Mechanical and Electrical (HM&E) system performance, likelihood of what system failures with different operating profiles, etc.

14. Are there any existing user interfaces for maintenance on marine vessels that you think are particularly well-designed? What elements of these interfaces could be incorporated into the design of a new user interface?

Unit Maintenance Management System (UMMS). UMMS is a web-based computerised MMS based on the principles of RCM. UMMS administers the management of ships, submarines, RFA, and shore environments.

Assets and their associated details are displayed on the asset hierarchy screen. The Asset Tree contains all a ship's assets arranged in a logical hierarchy displayed in a graphical form. In a ship system, the highest level of the hierarchy is the ship itself; the next level down includes all the ship's major systems. UMMS is significantly different from previous computer-based maintenance systems used in the Royal Navy in that scheduling most jobs is done in the shore environment and, through replication, the ship's staff and shore authorities can view the maintenance state of the platform. In conjunction with the ship's staff, COMs, and appropriate SCA/PA for non-SSS vessels are responsible for compiling work packages for those platforms fitted with UMMS. A web-based application, UMMS is widely available in the shore environment where Design Authorities can view maintenance tasks, overdue maintenance, and defects.

Due to classification constraints, there is not much more I can go into here beyond what is covered in this conference paper:

https://theforge.defence.gov.au/sites/default/files/reliability_centred_maintenance.pdf

Appendix C Expert 2 questionnaire responses

1. *Number of years of experience in the field of marine engineering/maintenance?*

37

2. *Past/Current responsibilities?*

Royal Navy Marine Engineer Officer so sea jobs as a Deputy Marine Engineer Officer of a Type 42 Destroyer, Propulsion Section Head and Senior Watchkeeper of our previous generation of aircraft carriers, and Marine Engineer Officer (Chief Engineer) of a Type 22 Frigate and Type 45 Destroyer. In addition to these sea-going roles several roles in direct support of shops (directing maintenance & repair) and defining support policy in the HQ.

3. *What does a typical day on the job look like for you as a marine engineer in charge of maintenance?*

Maintenance planning is/was conducted by the senior ratings in the Marine Engineering Department, led by the Warrant Officer and then Group and Section Heads. Much of this day-to-day work was left to them, with my input being to concede overdue maintenance items that could not be carried out and arrange with the Command for the standing down of machinery so that maintenance could be carried out. Daily, I would be involved in defect management, assessing the impact, reporting to the internal and external command chain, and arranging and managing defect rectification. In terms of maintenance, I would be looking ahead to the following maintenance period, be it a fleet period or a more in-depth Upkeep, using my knowledge of the material state of the vessel to produce a work package.

4. *What everyday maintenance tasks do you perform in the machine room?*

As mentioned above, I did not do any hands-on maintenance! However, my team would do routine watchkeeping when I started; this required physical recording of local readings by watchkeepers doing rounds to be presented to the watch leader on watch and daily to the Marine Engineer Officer. They would also monitor and replenish fluid levels, etc. Machinery sections would have daily servicing tasks, usually simple fluid levels and oiling and greasing routines. The planned maintenance system determined other servicing. Additionally, there was condition monitoring, usually taking vibration readings to be analysed both onboard and ashore.

5. *How do you diagnose and troubleshoot problems with the engines and other machinery in the engine marine room?*

Typical technique: Watchkeeping senses – sound, sight, smell, touch etc, Running parameters, Condition Monitoring, Vibration analysis – applicable to most rotating machinery, Machine specific techniques, Gas turbine debris monitoring, Diesel oil analysis – some testing done onboard, and spectrographic analysis done ashore.

6. *What tools and equipment do you use in your work? How do you ensure that these tools are properly maintained and calibrated?*

Normal hand tools and a Common Range of Electrical Test Equipment. Some specific support equipment is held to match what is expected of the ship's staff or because it would be too bulky to ship out for a repair when deployed. As an example, the Type 21/22/42 and Type 23 frigates could all conduct propulsion gas turbine changes deployed but relied on a change team of naval personnel to fly out with the spare gas turbine to do the change. All the heavy change rails, etc, were held onboard to minimise the air transport requirement.

7. *What safety precautions do you take when working in the engine marine room?*

PPE, Fire retardant overalls and underwear, Ear defenders – with built-in receive comms to the machinery space broadcast, Nonslip and steel toe capped safety boots, Man Below Routine adhered to, i.e. the Ship Control Centre knew you were in the machinery spaces. You would check in if you moved between spaces, e.g., for rounds purposes. Aware of the hazards, e.g. work going on, defects, hot surfaces, slipper decks and ladders and extra precautions if the ship was manoeuvring or in a high sea state.

8. *What training and experience are required for a marine engineer to be proficient in the maintenance of the engine marine room?*

Royal Navy practice across all ranks and rates is to have a period of shore training consolidated onboard using the 'Task Book' process of activities that must be conducted and signed off. There are then onboard examinations, both written and oral, for watchkeeping qualifications and professional advancement.

9. *What challenges do you face in your job, and how do you overcome them?*

The fundamental issue is understanding the design intent on a system basis, so you know what the marine engineering system is supposed to be capable of. This needs to be better understood in general. Thinking on a system basis so that you can effectively guide fault finding. Finally, understand the material state so you can direct maintenance/repair where needed.

10. *Are there any new technologies or practices you see adopting to maintain marine engines and machinery in the future?*

- Providing technical information at the point of work.
- Condition monitoring that works and allows interventions to prevent failure instead of telling you why something has failed.

11. *What are the most important factors to consider when designing a user interface for maintenance on a marine vessel?*

- How the interface performs in an emergency, e.g., how alarms are presented, prioritised, etc. Current systems present alarms and warnings as the system interrogates them. This results in a mass of alarms, so you cannot see the wood for the trees and alarms in an order that does not reflect how the defect occurred, e.g. for a cascade failure.
- The issue of parameters continually being in alarm or warning must be addressed. Constant resetting by the operator leads to poor watchkeeping habits; inhibiting the channel is a little better. The root cause needs to be determined, but that is not easy if you have 200 channels in alarm.

12. How do maintenance needs and requirements differ across different types of marine vessels?

- My experience is in warships, although I have some experience in auxiliaries. For a warship, I would say the onboard maintainer's time is split between:
 1. Maintenance
 2. Defect Repair
 3. Watchkeeping
 4. Whole-ship evolutions

This is an even split, although the watchkeeping burden is less than it was.

- For a more merchant ship configuration, they spend much more time percentagewise on maintenance, but because they have simpler systems and fewer people, they do less maintenance.

13. How do you envision the user interface being used in practice? What kind of information and functionality should it provide?

From what Ian tells me, we are attempting to augment reality. So, we will remove the need to do mundane tasks, and in fairness, a lot of this has gone already, so there should be no need for manual recording of running parameters at the machine, for example. So how could we attempt to enhance what the operator/maintainer/supervisor is presented with, and it could/should be different depending on the role:

Possible examples:

- Overlay running parameters on a view of a machine; do some analysis, e.g., where this pump is on its pressure/flow curve and what its performance trend is.
- Present known defects and allow new ones to be added, e.g. leaks.
- Material state information – more for maintainer/supervisor, e.g. latest CM results, when last run for standby machinery.

14. Are there any existing user interfaces for maintenance on marine vessels that you think are particularly well-designed? What elements of these interfaces could be incorporated into the design of a new user interface?

Modern Platform Management systems can present data supported by underlying graphics. Still, I miss the overall view you could get from a control panel, even if it presented much less data. There are few attempts to use large displays, but most use standard monitor sizes.

Appendix D Inverse Kinematics for UR16e Robot

GitHub repository link: https://github.com/mazeasdamien/sample_IK_UR16e

Result video:

https://cord.cranfield.ac.uk/articles/media/Immersive_robotics_experiments_-_Teleexistence_PhD_videos_/24190620?file=42445548

The Update() C# method for calculating and setting inverse kinematic solution is explained below:

```
// Update is called once per frame
void Update()
{
    // Calculate the transformation matrix for the IK
    Matrix4x4 transform_matrix = GetTransformMatrix(ik);

    // Reflect the matrix along the Y-axis
    Matrix4x4 mt = Matrix4x4.identity;
    mt.m11 = -1;
    Matrix4x4 mt_inverse = mt.inverse;
    Matrix4x4 result = mt * transform_matrix * mt_inverse;

    // Compute the inverse kinematics solutions
    double[,] solutions = Inverse_kinematic_solutions(result);
    IK_Solutions.Clear();
    IK_Solutions = DisplaySolutions(solutions);

    // Set the robot arm joints based on the selected solution
    ApplyJointSolution(IK_Solutions, solutions, solutionID, robot);
    goodSolution.Clear();
    goodSolution.Add(solutions[0, 5]);
    goodSolution.Add(solutions[1, 5]);
    goodSolution.Add(solutions[2, 5]);
    goodSolution.Add(solutions[3, 5]);
    goodSolution.Add(solutions[4, 5]);
    goodSolution.Add(solutions[5, 5]);
}
```

1. **Transform matrix calculation** represents the desired position and orientation of the end-effector in the workspace.
2. **Reflection operation:** A reflection operation is conducted along the Y-axis. This is achieved by multiplying the transformation matrix by a matrix 'mt' and its inverse. The matrix 'mt' is an identity matrix, creating a mirror effect along the Y-axis. The inverse of 'mt' is then calculated, and these

three matrices are multiplied to yield the resulting transformation matrix. This process is often used when the robotic system's orientation is mirrored from the standard configuration.

3. **Inverse kinematics calculation:** The function created by the author to calculate the inverse kinematics is then called using the resulting transformation matrix as input based on the Universal Robots calculation example (Universal Robots, 2022). This function computes joint configurations that allow the robot's end-effector to reach the desired position and orientation. The set of solutions is returned as a 2D array values.
4. **Solution selection and application:** The results of the inverse kinematics calculation are then displayed, visualising the various solutions for the user. Then, a function applies the selected solution to the robotic arm. This function sets the joint angles of the robotic arm to the angles specified by the selected solution, effectively moving the arm to the desired position and orientation.

This solution is sent to the cobot controller unit using the RTDE interface over a standard TCP/IP connection at 125Hz. Real-time data can be read and written from the UR16e registers input/output. Thanks to URScript, a high-level programming language that controls the cobot at the script level, access to the registers chosen for storing the correct joint angles solution is possible. The SERVOJ() command structure uses the registered solutions joint angles values and re-requests the UR16e to perform the movement; here is an example:

SERVOJ(Q, T=0.002, LOOKAHEAD_TIME=0.1, GAIN=300)

- Q = [0.0,1.57, -1.57,0,0,3.14] joint angles in radians representing base, shoulder, elbow, wrist1, wrist2, and wrist3 rotations.
- T = 0.008 time where the command is controlling the robot. The function is blocking for time t [S].
- LOOKAHEAD_TIME = 0.1 time [S], range [0.03,0.2] smoothens the trajectory with this lookahead time.
- GAIN = 300 proportional gains for following target position, range [100,2000]

Appendix E Fanuc remote control framework

Repository of the unity application:

https://github.com/mazeasdamien/p2_fanuc_cranfield/tree/main/Assets/DDS

Quality of Service file:

https://github.com/mazeasdamien/p2_fanuc_cranfield/blob/main/TelexistenceRig.xml

Repository link for the gateway application (running in parallel of the unity application): <https://github.com/mazeasdamien/TelexistenceRig>

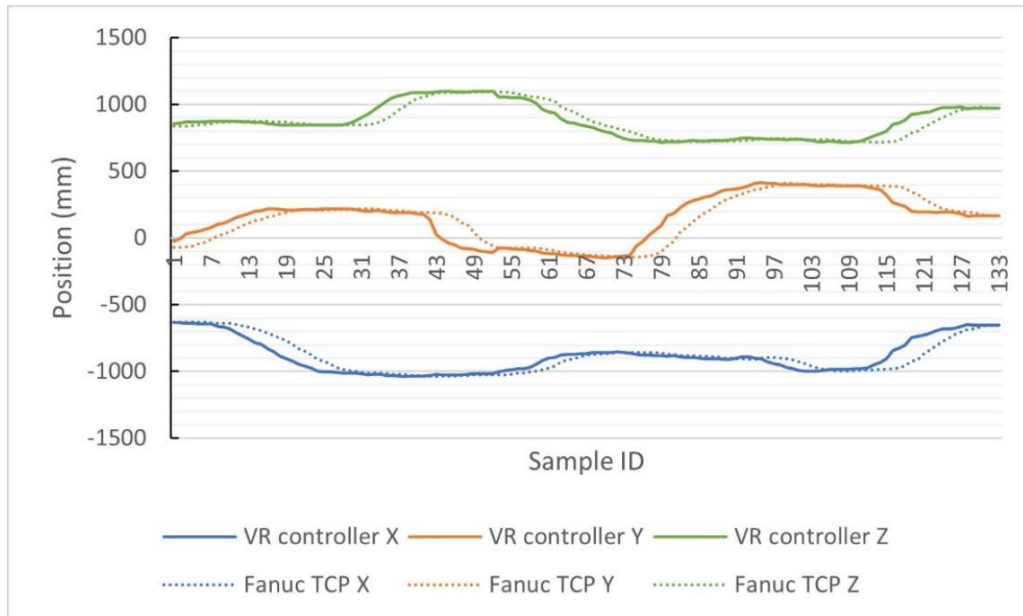
Here is a brief overview of the files in the "DDS" directory, each script has comments included for further explanation:

- **DDSHandler.cs:** This script is responsible for initialising and configuring DDS components such as the domain participant, subscribers, and publishers within a Unity environment. It provides methods for setting up DataReaders and DataWriters for dynamic data types associated with specific topics. The script ensures proper disposal of DDS resources when the Unity application is quitting. The Quality of Service (QoS) settings are loaded from external files for flexibility and customisation.
- **FanucManager.cs:** This script sets the target angles for each joint of the Fanuc robotics system based on the Fanuc joint angles received from DDS. It then updates the ArticulationBody joint drives with the new target angles. This script bridges the DDS communication with the Fanuc robotics system in Unity, updating the Unity scene's transformation and joint angles based on the received data. The CreateQuaternionFromFanucWPR method is used to convert Fanuc Wrist Pitch Roll angles to a Unity-compatible quaternion for accurate orientation representation in the Unity scene.
- **MovetoManager.cs:** This script is a Unity component responsible for handling DDS communication related to the "Moveto" action. It utilises a DragObject component to track object dragging. On initialisation, it sets up a DDS DataWriter for a specific topic ("Moveto_Topic") with a dynamic data structure containing a boolean member ("Released"). During each

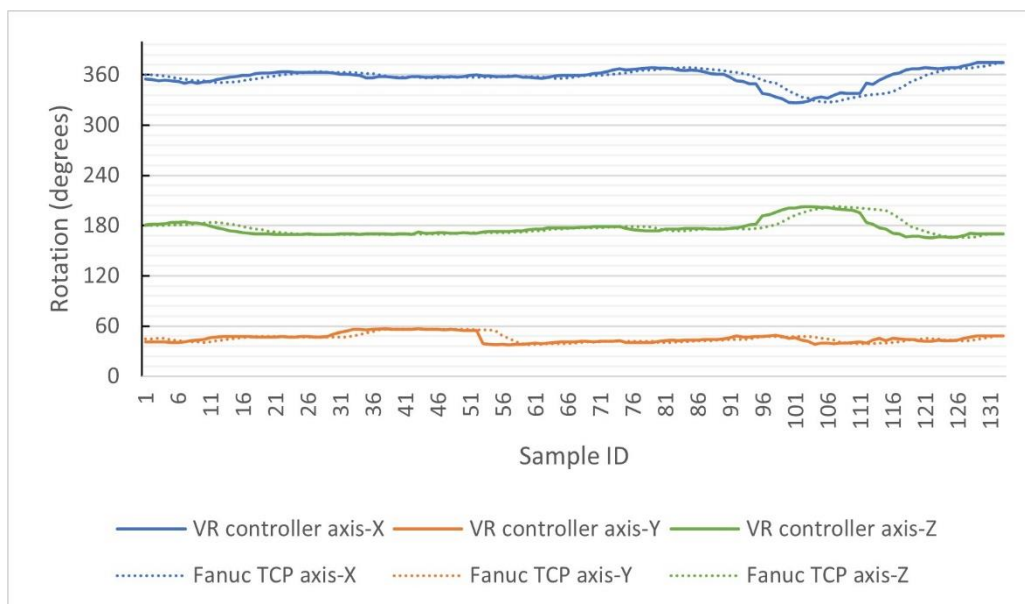
frame, it checks if the object has been released, and if true, it logs a message and sends a DDS sample with the boolean value set to true.

- **PointCloudManager.cs:** This script is a Unity component responsible for handling DDS communication and updating a point cloud visualisation using the Visual Effect Graph. It sets up a DDS DataReader for a specific topic ("PointCloud_Topic") with a dynamic data structure containing a sequence of float values. It processes incoming DDS data during each frame and updates the point cloud visualization using the Visual Effect Graph.
- **ReachableManager.cs:** This script is a Unity component responsible for handling DDS communication related to reachability information. It sets up a DDS DataReader for a specific topic ("Reachability_Topic") with a dynamic data structure containing a boolean member ("isReachable"). During each frame, it processes incoming DDS data and updates the visualization (material color) based on the reachability information. If the object is reachable, the color is set to green; otherwise, it is set to red.
- **TeleopManager.cs:** This script is a Unity component responsible for handling DDS communication related to teleoperation. It sets up a DDS DataWriter for a specific topic ("OperatorNewPose_Topic") with a dynamic data structure containing six float members representing position and rotation. During each frame, it checks if there is a change in the position or rotation of the associated GameObject and, if so, updates the DDS sample and sends it to the DDS topic. The rotation information is converted from a Unity Quaternion to Fanuc World Position Representation (WPR). This script enables the teleoperation of a virtual object in Unity by communicating its pose information through DDS.

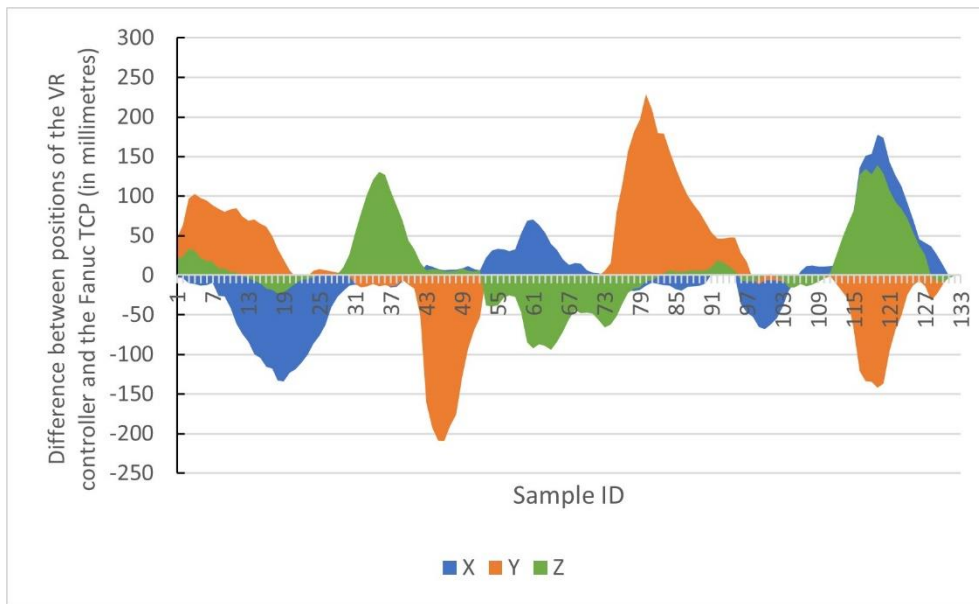
Appendix F Analysis of the Fanuc trajectory sample



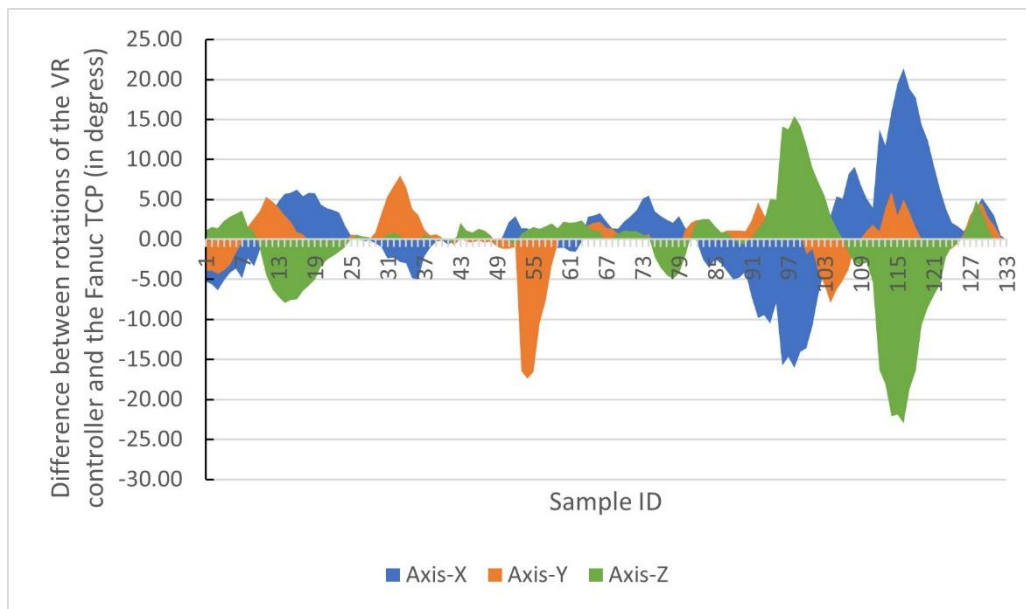
Range of position movement for the VR controller and the Fanuc TCP



Range of rotation movement for the VR controller and the Fanuc TCP



Difference between the rotation values of the VR controller and the actual Fanuc TCP



Difference between the rotation values of the VR controller and the actual Fanuc TCP