UNIVERSITY OF BIRMINGHAM

DOCTORAL THESIS

Human Factors Issues in Telerobotic Decommissioning of Legacy Nuclear Facilities

Author:
Mohammed TALHA

Supervisor:
Prof. Rustam STOLKIN
Prof. Robert Stone
Prof. Martin Freer
Mr. Wayne Ingamells

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Engineering

in the

Department of Physics & Astronomy

July 2018

Abstract

This thesis investigates the problems of enabling human workers to control remote robots, to achieve decommissioning of contaminated nuclear facilities, which are hazardous for human workers to enter.

The mainstream robotics literature predominantly reports novel mechanisms and novel control algorithms. In contrast, this thesis proposes experimental methodologies for objectively evaluating the performance of both a robot and its remote human operator, when challenged with carrying out industrially relevant remote manipulation tasks.

Initial experiments use a variety of metrics to evaluate the performance of human test-subjects. Results show that: conventional telemanipulation is extremely slow and difficult; metrics for usability of such technology can be conflicting and hard to interpret; aptitude for telemanipulation varies significantly between individuals; however such aptitude may be rendered predictable by using simple spatial awareness tests. Additional experiments suggest that autonomous robotics methods (e.g. vision-guided grasping) can significantly assist the operator.

A novel approach to telemanipulation is proposed, in which an "orbital camera" enables the human operator to select arbitrary views of the scene, with the robot's motions transformed into the orbital view coordinate frame. This approach is useful for overcoming the severe depth perception problems of conventional fixed camera views.

Finally, a novel computer vision algorithm is proposed for target tracking. Such an algorithm could be used to enable an unmanned aerial vehicle (UAV) to fixate on part of the workspace, e.g. a manipulated object, to provide the proposed orbital camera view.

Publications

Publications arising from this thesis:

Talha, M., Ghalamzan, E. A. M., Takahashi, C., Kuo, J., Ingamells, W., & Stolkin, R. (2016, October). Towards robotic decommissioning of legacy nuclear plant: Results of human-factors experiments with tele-robotic manipulation, and a discussion of challenges and approaches for decommissioning. In Safety, Security, and Rescue Robotics (SSRR), 2016 IEEE International Symposium on (pp. 166-173). IEEE.

Marturi, N., Rastegarpanah, A., Takahashi, C., Adjigble, M., Stolkin, R., Zurek, S., Kopicki, M., Talha, M., Kuo, J.A. and Bekiroglu, Y. (2016, December). Towards advanced robotic manipulation for nuclear decommissioning: a pilot study on tele-operation and autonomy. In Robotics and Automation for Humanitarian Applications (RAHA), 2016 International Conference on (pp. 1-8). IEEE.

Talha, M., & Stolkin, R. (2014). Particle filter tracking of camouflaged targets by adaptive fusion of thermal and visible spectra camera data. IEEE Sensors Journal, 14(1), 159-166.

Acknowledgements

Firstly, I would like to express my sincerest gratitude to my supervisor Prof. Rustam Stolkin for his endless support and guidance during my EngD through thick and thin. Words cannot express how much I appreciate the motivation he gave and patience he showed to overcome obstacles while giving me the freedom to find my own way. I have utmost respect for his selflessness to help not just his students but those struggling with other supervisors. There is not a better supervisor in the world.

I would also like to thank my co-supervisors. Prof. Robert Stone for his expertise in human factors and experiment design without which I would be lost and Prof. Martin Freer for his accommodating nature and bureaucratic support. I also wish to express sincere thanks to my industrial co-supervisor Wayne Ingamells. Wayne provided the initial inspiration that led to this project being funded. He provided great insight into the practicalities of nuclear decommissioning problems and working practices. Wayne also generously provided mentoring for the nuclear engineering modules which I had to undertake as part of this research, as well as advice on my papers and parts of this thesis.

Many thanks to Amir Ghalamzan and Chie Takahashi for giving me the push I needed to get over my first hurdle; your help was invaluable. I would also like to extend my thanks to the rest of the members of the Extreme Robotics Lab especially Maxime and Henry for their technical help.

Special mentions to the following: Peter Brewer and Daniela Ionescu for dealing with admin issues and making life for everyone in the robotics group a lot easier. Michael Mistry who is no longer at this university but was instrumental to understanding the complex world of robotics. Also to the folks at Coppelia Robotics GmbH for creating the best robotics simulator, V-rep, as well as providing prompt support all free of charge.

I am thankful for the Engineering and Physical Sciences Research Council and our industrial partners NNL for their financial support. Thanks to Lorraine McDermott for her guidance and special thanks to Dr. Jeffrey Kuo, one of the industrial co-supervisors of this project, who tragically became ill and passed away during the project. Dr. Kuo was a pioneer in the field of nuclear robotics, and we have benefited greatly from his expertise and generosity of spirit.

I am specially thankful and grateful to my family for their emotional support throughout this rollercoaster project without which this would not have been possible.

In loving memory of
my beloved mother
1960 - 2016

Contents

List of Abbreviations

Nuclear terms

ALARP	**A**s **L**ow **A**s **R**easonably **P**racticable
HLW	**H**igh **L**evel **W**aste
LLW	**L**ow **L**evel **W**aste
VLLW	**V**ery **L**ow **L**evel **W**aste
PBO	**P**arent **B**ody **O**rganisation
POCO	**P**ost-**O**perational **C**lean-**O**ut
NDA	**N**uclear **D**ecommissioning **A**uthority
SLC	**S**ite **L**icence **C**ompanies
UKAEA	**U**nited **K**ingdom **A**tomic **E**nergy **A**uthority

Robotics & HRI terms

AAAI	**A**merican **A**ssociation of **A**rtificial **I**ntelligence
ANOVA	**A**nalysis **O**f **V**ariance
API	**A**pplication **P**rogramming **I**nterface
CAD	**C**omputer **A**ided **D**esign
DARPA	**D**efense **A**dvanced **R**esearch **P**rojects **A**gency
DOF	**D**egree(s) **O**f **F**reedom
DRC	**D**ARPA **R**obotics **C**hallenge
EE	**E**nd **E**ffector
EIH	**E**ye **I**n **H**and
FOV	**F**ield **O**f **V**iew
GUI	**G**raphic **U**ser **I**nterface
HCI	**H**uman **C**omputer **I**nteraction

HMD	**H**ead **M**ounted **D**isplay
HMI	**H**uman **M**achine **I**nterface
HRI	**H**uman **R**obot **I**nteraction
IK	**I**nverse **K**inematics
IR	**I**nfra **R**ed
LOA	**L**evel **O**f **A**utomation
MSM	**M**aster **S**lave **M**anipulator
NASA-TLX	**N**ational **A**eronautics and **S**pace **A**dministration **T**ask **L**oad **I**ndex
PTU	**P**an-**T**ilt **U**nit
ROS	**R**obot **O**perating **S**ystem
ROV	**R**emotely **O**perated **V**ehicle
RSE	**R**obotics **S**imulation **E**nvironments
SA	**S**ituational **A**wareness
SAGAT	**S**ituational **A**wareness **G**lobal **A**ssessment **T**echinque
SAR	**S**earch **A**nd **R**escue
SLAM	**S**imultaneous **L**ocalisation **A**nd **M**apping
SUS	**S**ystem **U**sability **S**cale
TTC	**T**ime **T**o **C**ompletion
UAV	**U**nmanned **A**ir **V**ehicle
USAR	**U**rban **S**earch **A**nd **R**escue
VR	**V**irtual **R**eality

Chapter 1

Introduction

1.1 Overview

This thesis investigates the problems of enabling human workers to control remote robots, for decommissioning of contaminated nuclear facilities, which may be prohibitively hazardous for human workers to enter. In particular, the focus of this thesis is on the issues of human-robot interaction during remote manipulation tasks. The thesis begins with a discussion of the nuclear decommissioning challenges.

1.2 Background

The United Kingdom began nuclear operations at the Sellafield site in 1947, with the first nuclear reactor becoming active in 1950 [1]. The first reactors were built for weapons production (especially plutonium) in response to the emergence of the cold war arms race era, with the first Soviet atomic bomb test in 1949. However, the UK also pioneered the peaceful use of nuclear energy, with the world's first industrial scale civil nuclear power station in 1956 [1].

The UK has thus been engaging in nuclear activity for some 70 years, and this has resulted in a very large legacy waste remediation problem. In 2004, the Nuclear Decommissioning Authority (NDA) [2] was formed as part of the Energy Act 2004 with the core aim of decommissioning legacy nuclear facilities “safely, securely, cost-effectively”. The UK is estimated to contain 4.9million tonnes of legacy waste [3], approximately 75% at the Sellafield site (largest nuclear site owned by the NDA [4]) in Cumbria. The UK nuclear legacy is regarded as the largest and most complex environmental remediation project in western Europe, expected to take more than 100 years at costs in the order of £90-200billion [3]. It is expected that a significant part of this work will have to be done by using remotely operated robots, because some materials and facilities are too hazardous for humans. Additionally, demolition work currently done by humans wearing air-fed plastic suits is hazardous, and extremely uncomfortable for workers. Furthermore, the protective suits become secondary nuclear waste, which can greatly exceed the original volume of materials being decommissioned. Without significant advances in robotics, it is estimated that one million suited human entries into hazardous environments will be needed.

The NDA believes that the condition of the historic 1950s-built facilities at Sellafield is deteriorating due to poor maintenance, and they are no longer fit to hold waste [2]. Therefore novel methods such as robotics, which might achieve faster decommissioning, are increasing in priority. It is estimated that the cost to decommission all NDA sites is estimated at £125 billion of which 75% is attributed to Sellafield .

A motivating example for this project is the decommissioning of contaminated legacy glove-boxes. The UK Atomic Energy Authority (UKAEA) was formed in 1954 to manage the country’s nuclear weapons research and nuclear reactor programmes [5]. As part of these programmes, the UKAEA used numerous glove-boxes in laboratories and facilities to protect personnel from contamination when conducting research with materials such as plutonium, which forms a fine dust that can be toxic as well as emitting alpha radiation.

To decommission these legacy facilities, they have to be decontaminated, demolished, and all materials must be size-reduced and manipulated into long-term storage containers. However, some of these facilities have been non-operational for many decades. They were constructed and operated with different regulations and with little thought for future decommissioning [6]. As a consequence, there is significant uncertainty about some facilities and their contents, e.g. blueprints and inventories.

This uncertainty causes significant challenges for planning of decommissioning interventions. Uncertainty is especially challenging for the deployment of robotic systems, which have predominantly been used in highly structured and precisely known environments in the manufacturing industry. This is why manufacturing has undergone a robotics revolution since the 1980s, while the nuclear industry remains comparatively un-roboticised.

1.3 Nuclear Decommissioning

Decommissioning can be considered to be a series of actions taken on a facility when it reaches the end of its useful life to ensure the safety of workers, public and environment [7, p. 41]. This can vary between closing the facility and completely dismantling it to the end state, at which point it can be used for other purposes. Bayliss and Langley [7] outlines the decommissioning process to be as follows:

- Shut-down transition phase (defueling & post-operational clear-out known as POCO).
- Preparation for safe enclosure.
- Safe enclosure period.
- Final dismantling.

There are two strategies for decommissioning:

- Continuous- Starts as soon as a facility becomes non-operational.
- Deferred- Radiation in a facility is allowed to decay in a safe enclosure until radioactivity reaches a level safe enough to allow operation.

The method of decommissioning depends on the type of facility. Reactors are immediately defueled and safely enclosed until radiation levels become safe; plutonium facilities require immediate decommissioning because, if left for long periods, alpha-emitting contaminants may decay to form more dangerous gamma-emitting substances; hot cells can safely contain radioactivity so decommissioning can be postponed; waste treatment and storage facilities are treated similarly to reactors but the timing of their dismantling is discretionary.

The NDA contracts this work to Site Licence Companies (SLCs) which operate on NDA-owned nuclear sites. The SLCs are managed by Parent Body Organisationsco (PBOs) selected through bidding processes run by the NDA. This helps to provide SLCs with additional resources and expertise. The NDA also funds research and development programmes to develop new technologies to tackle the challenges it faces. This project is an example of NDA-funded research.

The UK government now has extensive regulations which must be followed by nuclear site operators to ensure safety and security of workers and the public and the protection of the environment. The NDA is responsible for monitoring the operators to make sure they meet all the requirements. In cases such as these, where the risk is intolerable [2], the NDA focuses on hazard and risk reduction for determining how to proceed. A hazard is defined as the“potential for harm arising from an intrinsic property or ability of something to cause detriment“ and risk is the “chance that someone, or something that is valued, will be adversely affected by the hazard" [2, p. 14]. Risk can be described on 3 levels:

- Intolerable- Risk level is unacceptable except in extraordinary circumstances; urgent action is needed to reduce risk to tolerable.
- Tolerable- Risk is As Low As Reasonably Possible (ALARP); further reduction is not cost-efficient.
- Broadly Acceptable- Low risk; focus on task completion.

Since some aging UK facilities pose an intolerable risk to safety and the environment, it is imperative they are decommissioned as soon as reasonably practicable through continuous decommissioning. The timing of decommissioning is determined on a case-by-case basis through evaluation of risk as well as other factors such as safety, security, environmental impact, funding, resources and waste management NDA2011.

Another essential aspect of decommissioning is waste management. Before decommissioning can begin, a safe and effective method of waste management must be identified [8]. This consists of waste minimisation strategies, re-use and recycling, waste treatment, packaging, storage, transport and disposal [2, p. 39]. Waste is classified based on the level of radioactivity present in it [7]:

- Very Low Level Waste (VLLW)- can be disposed of with ordinary refuse
- Low Level Waste (LLW)- waste from routine/ decommissioning operations with small concentrations of radioactivity present
- Intermediate Level Waste (ILW)- waste containing high concentrations of radioactivity requiring remote handling but no heat generation
- High Level Waste (HLW) - waste such as spent nuclear fuel with the most radioactivity. At this level, consideration needs to be given to the

significant amount of heat generated from the concentration of radioactivity present.

The type of disposal required for each waste type depends on the period of radioactivity [7] as shown in table 1.1. The current method of storage is shown in table 1.2.

Waste classification	Short-lived	Long-lived
LLW	Shallow disposal	Deep disposal
ILW	Shallow disposal	Deep disposal
HLW	Not applicable	Deep disposal

TABLE 1.1: Waste disposal classification based on period of radioactivity.

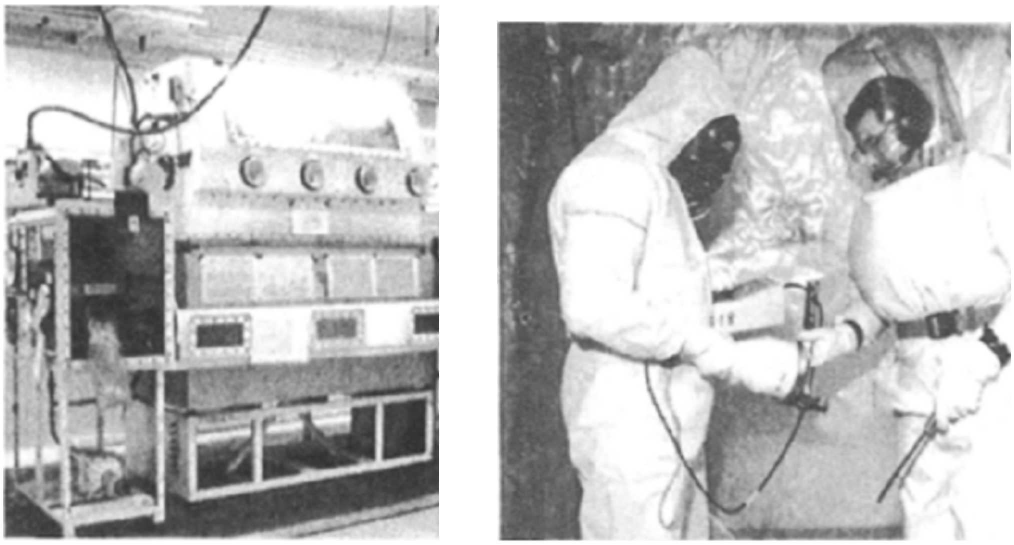
Waste classification	Short-lived	Long-lived
LLW	Landfill sites	On-site storage
ILW	On-site storage	On-site storage
HLW	On-site storage	On-site storage

TABLE 1.2: Current method of waste storage.

ILW is immobilised in cement before storage pending availability of disposal sites and HLW is processed and stored as liquid in air-cooled storage pending availability of disposal sites [9].

1.3.1 Case Study: Glovebox Decommissioning

Gloveboxes (fig. 1.1(a)) are sealed, transparent containers made of Perspex reinforced with steel or plywood with built-in glove attachments that allow users to work with objects in containment. The atmosphere inside a glovebox is controlled at a lower pressure than the external environment to contain any radioactive leaks. Legacy facilities primarily used these to handle plutonium



(a) (b)

FIGURE 1.1: (a): Example of a glovebox; (b): Workers in pressurised suits during decommissioning. *Source: [7]*

and uranium. Plutonium is a man-made element derived from the decay of uranium [10]. As plutonium decays, it emits alpha radiation. Alpha radiation is easily contained by gloveboxes because it lacks penetration. However, if the plutonium isotope Pu-241 is left to decay, it becomes americium which is a gamma-emitter. Gamma radiation is dangerous because it can penetrate even steel [7].

Throughout the UK, there are numerous sites with tens, or even hundreds of labs, each containing many glove boxes. Typically, these labs have been closed off and abandoned at the end of their working lives without decommissioning. They now contain a mixture of alpha and gamma emitting contaminants because they were not decommissioned using the continuous strategy. Due to prolonged exposure to radiation, the older glove boxes are themselves beginning to decay. For example, their Perspex windows have become brittle and can fracture and collapse on minimal disturbances, risking severe contamination of the room surrounding the glove box, severely compounding the decommissioning problem.

Dismantling these facilities currently requires personnel to wear pressurized suits (fig. 1.1(b)) which are expensive and restrictive. Additionally, work can only be carried out for short durations due to personnel reaching their gamma radiation dosage allowance. Upon exiting the contaminated area, personnel must be decontaminated and their suits stored as ILW in addition to the waste generated from the decommissioning. Decontamination is a lengthy procedure requiring additional personnel in protective clothing generating additional LLW. Moreover, since workers use sharp tools for their work, there is a risk of these suits becoming compromised. Ingestion of alpha emitters such as plutonium through inhalation can have serious long-term health implications.

The current method of decommissioning is clearly not an efficient or cost-effective method. It requires excessive time, personnel being exposed to hazards and generates additional collateral waste. With the rejection of an underground disposal facility in Cumbria [11] on the only potential site, space for waste storage is at a premium. Therefore, using a robot controlled from a safe remote location is a more economical, efficient and safe alternative for decommissioning [12].

1.4 Research Context

Current methods for decommissioning of alpha-contaminated facilities predominantly rely on human workers. These facilities are put in sealed containment, with human entries carried out wearing pressurized air-fed suits for protection. Typically a worker can only carry out two hours of work per day inside an airfed suit. The suits, and multiple layers of gloves, severely inhibit the ability to carry out manual tasks, making decommissioning a time consuming and expensive operation. Working with sharp power tools whilst inside the hazardous environment presents a risk of an accident that will compromise the suit. Upon exiting the contaminated area, personnel must be decontaminated, requiring additional personnel (often themselves suited). All such protective clothing must then be disposed of as contaminated waste. This is known as “secondary waste“. Typically secondary waste volume can be more than ten times the volume of the original material being decommissioned. Since long-term storage of all such waste is extremely expensive, the secondary waste associated with human interventions poses a serious problem.

The alternative, as proposed in this thesis, is to use a teleoperated robot to assist with decommissioning. The robot may be left inside the facilities for extended periods of time. It can be operated by rotating shifts of workers, in safe zones. This means that decommissioning work can progress continuously, and removes the need for most of the human entries and corresponding hazard and secondary waste. This approach is consistent with the UK nuclear industry’s drive towards “safer, faster, cheaper“ solutions.

While a variety of remote manipulation devices have been used on nuclear sites for decades, evidence of their performance within the industry is largely anecdotal. Meanwhile the mainstream academic robotics literature predominantly focuses on the design of novel robotic mechanisms and novel control algorithms. In contrast, this thesis is a contribution towards formal methods of objectively exploring and evaluating how well humans can control remote robots, to carry out realistic and relevant tasks.

Traditional telemanipulation using push-buttons with joint control is slow and tedious. Recent rapid advances in technological areas such as gaming, data processing, mobile phones and artificial intelligence has led to the release of mature technological devices (e.g. Kinect, drones, HTC Vive) that are affordable and widely available and made powerful techniques (e.g. deep learning, simultaneous localisation and mapping (SLAM), image processing) more feasible. The use of this technology has allowed robotic systems to sense environments in real-time and perform complex tasks. This thesis explores the application of similar technologies in the nuclear context to improve the human machine interface and evaluates them with human factors issues in mind through a principled empirical approach.

1.5 Thesis Structure

The remainder of this thesis is arranged as follows:

Chapter 2 provides a literature review. This thesis is interdisciplinary, combining elements of experimental design and analysis from psychology and human factors fields, with robotics, virtual reality, and also nuclear engineering problems. Therefore, a wide variety of literature needs to be discussed. This includes sections on telerobotics, experiment design and human-machine interfaces.

Chapter 3 presents the design and technical development of a testbed in a virtual reality simulation environment. This enables a human test-subject to control a virtual remote robot, while viewing the scene using virtual camera views. The testbed forms the basis for human-factors testing.

Chapter 4 presents experiments to objectively evaluate the performance of humans using remote robot arms to perform complex manipulations, and makes several observations from the analysed test data.

Chapter 5 presents an experiment to show how the use of advanced autonomous robotics methods (vision-guided grasping) as assistive tools for the human operator, i.e. "human-supervised autonomy", can significantly outperform conventional direct teleoperation for pick-and-place tasks.

Chapter 6 presents a novel approach to teleoperation and situational awareness. This approach involves the use of an "orbital camera", which enables the operator to view the scene from an arbitrary viewpoint. A control method is proposed, in which the robot's motion is referenced from the coordinate frame of the orbital camera. This enables the robot to be controlled relative to the orbital camera view. Experiments are presented to evaluate this approach in comparison to conventional approaches with fixed camera views.

A robust implementation of the orbital camera from chapter 6 would require computer vision methods to enable features such as automated gaze fixation and end-effector tracking as well as reducing camera instability due to movement. This would also help to relieve the workload on the operator by reducing the amount of error control required for positioning the camera. Chapter 7 presents a novel computer vision algorithm for tracking moving targets for use with the orbital camera. Such a tracking algorithm could be used to enable a UAV to fixate its gaze on a user-designated region of the manipulator's work-space. This would enable the use of a UAV to provide the orbital camera view in a real application such as decommissioning a glovebox facility with a mobile-manipulator robot.

Chapter 8 provides a summary of the doctoral research work with concluding remarks and suggestions for future extensions to this work.

Chapter 2

Related Work

This project aims to evaluate technologies for use in a teleoperated robot system for remotely dismantling contaminated radioactive labs in the nuclear industry. Such a complex problem requires the use of highly interdisciplinary robotic and sensing technologies as well as comprehensive human factors and human robot interaction studies.

2.1 Teleoperation

Over the last 50 years, teleoperation has been employed in various ways, ranging from space robots, remotely operated underwater vehicles, surgical robots to unmanned military robots in the air and on the ground. However, it was the nuclear industry which inspired the initial research into teleoperation. The earliest application of teleoperated manipulators was handling nuclear materials in "hot cells" by operators situated outside the cell using Master-Slave Manipulators (MSMs) as demonstrated by Goertz in 1949 [13].

2.1.1 Definitions

The term teleoperation and its sister terms are often subject to interpretations. To avoid ambiguity, this thesis makes use of the definitions and concepts provided by the seminal work of Thomas B. Sheridan. He [13, p.487] defines

teleoperation as *"the extension of a person's sensing and manipulation capability to a remote location"* which refers to extending the capability of a human operator to manipulate objects remotely. When teleoperation involves interaction at the remote site, it is commonly referred to as **telemanipulation**. Teleoperation encompasses **telerobotics** which is defined as *"a form of teleoperation in which a human operator acts as a supervisor… (using 2-way communication with a computer) … while the subordinate telerobot executes the task based on information received from the human operator plus its own artificial sensing and intelligence"*[13, p.488]. In generic terms, telerobotics is robotics with human-in-the-loop control. In the context of this thesis, a robot, as defined in the oxford dictionary [14], is a machine that can be programmed by a computer to carry out a series of complex actions. An autopilot system on an aircraft can be considered a robot as it carries out complex maneuvers but a washing machine is only a machine as it only performs a few preprogrammed operations.

There is a slight distinction between the terms telerobot and teleoperator. A **teleoperator** can be considered to be any machine that extends the capabilities of a human to a remote location requiring direct and continuous control from the human operator whereas a **telerobot** is an advanced teleoperator that has autonomous capabilities requiring only supervisory control from the human operator [15]. **Supervisory control** in this context refers to the high-level monitoring and direction of a semi-autonomous system by the human operator. Teleoperation can be combined with **telepresence** which involves providing sufficient stimuli to the human operator so that they feel physically present in the remote environment.

A teleoperation system can be thought of as two subsystems: a master (input device/ human operator) at the local site and a slave (telerobot/ teleoperator), at the remote site, that communicate in the form of signals (e.g. positions, velocities or forces), fig. 2.1. In this connected system, the teleoperator executes the commands of the human operator based on the feedback

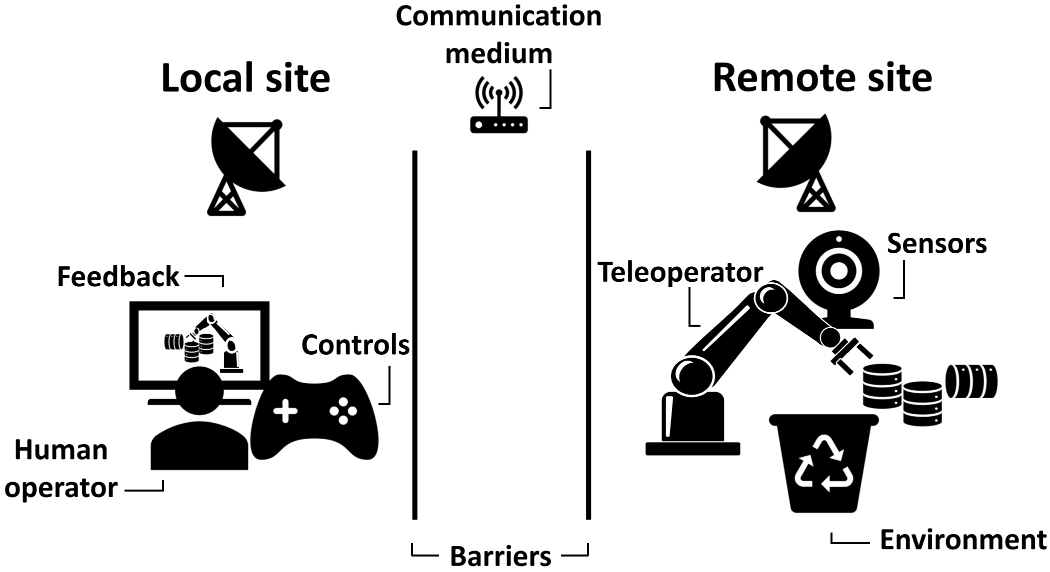


FIGURE 2.1: A typical teleoperation system.

(visual, auditory, haptic etc.) from the remote sensors. Such a system requires the integration of many areas of robotics including force and motion control, sensing as well as haptic and tactile feedback. If the slave system has force/ torque sensors, it can transmit back reaction forces from the remote environment to the master which in turn renders the forces to the operator providing touch feedback [16]. This is known as bilateral control.

The wide variety of teleoperation systems can be classified into three classes based on their control as follows [17]:

1. Closed loop control (direct teleoperation)- the operator directly controls the actuators of the teleoperator with real-time feedback.

2. Coordinated teleoperation- similar to 1 but the teleoperator now possesses internal control loops for basic tasks such as controlling speed. There is no autonomy.

3. Supervisory control- the operator mainly monitors and gives high-level commands. The teleoperator is highly autonomous and can perform most tasks without operator intervention.

In most cases, direct teleoperation on its own is not a very effective method of operation as it encumbers the operator and leads to cognitive overloading. To improve on this, research in telerobotics is aimed at improving the autonomy of the robot through teleautonomy and improving the immersion of the operator through telepresence with the assumption that both will lead to better overall performance of the system.

2.2 Teleautonomy

Before telerobotics, robots were pre-programmed to carry out tasks as on a factory assembly line. With the introduction of wireless technologies and human-machine interfaces, robots can now be directly controlled by human operators over long distances allowing them to be used in unstructured environments. In such dynamic complex environments with inherent uncertainty, telerobots have yet to prove that they can break free from the control of the human operator. This limits the effectiveness of a telerobot to the capabilities of the human operator receiving sensory feedback from the remote environment. If there is interference in the feedback such as time delay due to long distances or radiation affecting sensors, this can significantly cripple the performance of the telerobotic system. To avoid these issues, the long-term goal of telerobotics is to make robots increasingly autonomous. This can reduce the workload and fatigue of the operator and also augment their capabilities by increasing situational awareness and freeing time for multi-tasking.

Fully autonomous systems are not currently achievable due to the complexity and impracticality of modeling every aspect of the environment with all its uncertainties. To be fully autonomous, a robot would need to fully replicate critical aspects of human abilities such as sensing, planning, decision making, navigation, manipulation, and learning. The next best option is to integrate the best of both; the intelligence of the human operator with the physicality of the robot. The human operator monitors the robot and gives high-level task goals acting as a supervisor while the telerobot performs low-level and repetitive tasks using its internal control loop. This is essentially supervisory control [13].

Although supervisory control suggests the presence of autonomy in a telerobotic system, it is difficult to determine the degree to which the telerobot is autonomous. To that end, Sheridan [18] is credited with developing the spectrum of autonomy known as level of autonomy (LOA) for human-machine systems. LOA consists of 10-points ranging from direct teleoperation, with gradual increase in decision making autonomy to fully autonomous, 2.1. Its limitation to a single dimension was acknowledged and was later modified [19], to include other aspects of a human-machine system that could be automated such as sensing and perception.

1.	Human does it all
2.	Robot offers alternatives
3.	Robot narrows alternatives down to a few
4.	Robot suggests a recommended alternative
5.	Robot executes alternative if human approves
6.	Robot executes alternative; human can veto
7.	Robot executes alternative and informs human
8.	Robot executes selected alternative and informs human only if asked
9.	Robot executes selected alternative and informs human only if it decides to
10.	Robot acts entirely autonomously

TABLE 2.1: Levels of automation adapted from [18]

Sensory Processing	Perception/ Working Memory	Decision Making	Response Selection

TABLE 2.2: Four-stage model of human information processing including aspects other than decision making that can have LOA applied to from table 2.1.

However, teleautonomy is not without its drawbacks. Research has shown that even when a system can be fully automated, it may not be desirable as in the event of an error, this leads to operators failing to recover effectively [20]–[22]. This is due to the operator becoming complacent leading to skill degradation and inability to perform tasks manually when automation fails. Another factor is the lack of communication between the silent automated system and operator leaving the operator in an uninformed state. Research in teleautonomy is focused on developing an understanding of the human-robot system to efficiently allocate appropriate tasks to both human and robot [23], [24].

2.3 Telepresence

In the context of telerobotics, the term “telepresence“ was first coined by Marvin Minksy in 1980 to describe a future where humans could feel the sense of being teleported to a remote work location using technology [25]. There are multiple interpretations of telepresence but all refer to the sense of being transported to a space through the use of technology [26]. Telepresence allows the human operator to feel as if present at the remote environment, making the telerobot transparent through replication of sufficient human senses from the remote environment. Ideally, all human senses would be replicated but this can vary depending on the task and requirements. Telepresence can be extended to also include actions by the operator (movement, speech etc.) being conveyed from the local environment.

Visual feedback using a monitor to display a camera feed is a start towards achieving telepresence but is not enough on its own; as Sheridan [27] stated the greater the number of senses engaged, the better the sense of telepresence achieved. The typical way to induce telepresence is to use technology to replicate the senses [17]. The human eye is responsible for approximately

90% of sensory input to the brain [17]. Since vision accounts for the majority of sensory input, it follows that providing cutting-edge visual immersion is critical to instilling telepresence. The human vision system is a complex and versatile system. It allows for the world to be perceived in 3D with full colour, high resolution, great contrast and a wide field of view. These features are restricted when performing tasks in teleoperation which leads to reduced performance compared to direct operation [13]. The use of head mounted displays (HMD) in conjunction with cameras mounted on pan/tilt units (PTU) has been prevalent in the goal to achieve visual telepresence [28]–[30]. The display in the HMD can also be stereoscopic to provide depth information if the remote camera system allows. Additionally, the visual feedback can be augmented to improve operator performance [31]. Previous work [32], [33] to incorporate this presented a virtual 3D world to an operator constructed in real-time through SLAM and other mapping data. It proved that that there was a reduced workload on the operator which led to fewer errors and an increased sense of control.

Hearing is the second sense which dominates the sensory channels. Sound can be transmitted from the remote environment to provide useful feedback to the operator e.g. information about the volume and pitch of a sound during a cutting operation. Although audio can be beneficial for telepresence, it can also weaken the experience if the audio is of poor quality unlike vision or if the synchronicity is off as identified by Lessiter [34]. Lessiter conducted a study which showed that the sense of telepresence varies as a function of the perception of audio. Kiselev et al. [35] evaluated stereo vs. mono sound feedback from varying angles in a remote environment through an experiment and showed that participants were more successful at localising sound using stereo. Additionally, the method of audio delivery is also important as research shows that users prefer headphones to speakers since they offer better realism [36]. Le Groux [37] explored the use of sonification, the use of non-speech sounds for visualisation, as an additional method of providing immersion as well as information to the user.

Since vision and sound are already heavily occupied in terms of perception, touch sensing is an alternative to conveying information effectively to the operator during a task. This can be achieved through haptic feedback [38]. Haptic displays generate skin-based as well as proprioceptive (body position, orientation and movement) feedback while tactile is a type of haptic feedback that uses pressure to stimulate the skin [39].

Research into haptic feedback has been going on for decades but only recently have haptic devices reached a level where realistic feedback could be achieved at an affordable price and practical size [40], [41]. The Touch (previously Phantom), from 3D Systems [42], is a 6DOF haptic feedback device with industry leading performance characteristics. Alternatively, Force Dimension [43] provides a range of devices from affordable 3-DOF solutions to sophisticated 6-DOF devices that cover the range of motion of the human hand and provide gravity compensation.

The combination of haptic and force feedback can improve teleoperation as shown by Sarakoglou et al. [44] in experiments comparing performance with force and haptic feedback combined against force feedback alone. Subjects performed a simple path-following exercise from a remote location using an Omega7 force feedback device and KUKA 7-DOF arm. The results indicated an improvement in trajectory and lower contact forces when using both haptic and force feedback.

Research from UCLA [45] utilised a tactile feedback system for the *da Vinci* surgical robot system demonstrating that operations without tactile feedback involve the use of excess force. This implies the use of tactile information in precise and delicate tasks such as tissue handling can be crucial for preventing mistakes. Enayati, De Momi and Ferrigno [46] reviewed the use of haptics in surgical robotics across a number of studies. These included knot tying, suture maneuvers, dissection and needle insertion with benefits ranging from improved performance times and reduced errors to reduced forces exerted by operators and improved accuracy in tissue characterisation. They also note that although studies show advantages in using haptic feedback during surgery more clear evidence is needed to convince surgeons of the value of robot surgical systems. This may need to be provided in the form of a cost-benefit analysis.

2.3.1 Applications

Teleoperated robots have become essential tools for many areas where it is necessary to work in hazardous environments. Space exploration necessitates the use of teleoperated robots to reduce cost and risk to astronauts' safety [16]. The first ever use of a remotely operated robot arm in space was by the DLR German Aerospace Research Centre [47]. It was successfully able to execute a set of tasks including connecting & disconnecting plugs, assembling structures and capturing a free-floating object over a long distance using a predictive display. Current more notable applications include the NASA Voyager probe, Mars Rovers and the Hubble telescope.

Underwater robots were first used by the US Navy for deep sea salvage operations in the 1960s [48, p. 129]. From the 1970s, remotely operated vehicles (ROVs) started to be employed for underwater applications to reduce costs and risk from the use of human divers [13]. The initial proponents of ROVs were offshore oil and gas industries who utilised them for inspection and monitoring of pipelines [13]. The use of ROVs is now ubiquitous with research surveys, seabed mapping, military missions, cabling and repair operations in deep waters. Revenues for ROV manufacturers worldwide are estimated to be several hundred million dollars [48], [49].

Other areas include military applications with the use of a variety of drones and ground vehicles which are remotely controlled by Global Positioning Satellite (GPS). Unmanned Air Vehicles (UAVs) or "drones" can be used for reconnaissance and combat whilst ground vehicles are predominantly used for bomb disposal [50]. The nature of the hazardous tasks in the forest & mining industry has led to the development of robots such as the Work Partner [51] designed for heavy duty outdoor tasks. It is able to learn tasks and maintain a close relationship with human operators. Robots in the mining industry are used mainly for inspection and rescue operations in the event of mine collapse.

Aside from applications to hazardous environments, another important application of telerobotics is in robot assisted surgery. One goal of telerobotics in medicine is to provide access to medical expertise over long distances saving time and money. The daVinci is a prominent dexterous telerobot system that can be used for a variety of surgeries [52]. It is considered to be one of the most advanced master-slave systems in medical telerobotics providing dexterous control and immersion through high-quality stereo visualisation [53]. The benefits over conventional surgical methods include reduced blood loss, post-surgery pain and quicker recovery for the patient although surgical times are increased [53]. However, as the daVinci system is specialised for short-distances, it faces the challenge of time-delay between the operator input and robot execution for transatlantic distances. It also lacks autonomy relying on the surgeon to control all actions through direct teleoperation.

Telerobots are also being used to provide telepresence. Teleconferencing robots allow a new level of physical presence in remote locations through the use of a videoconferencing-capable device mounted on a mobile base [54]. These have seen use in meeting rooms, classrooms and events such as conferences and concerts.

2.4 Human robot interaction

Human Robot Interaction (HRI) is an interdisciplinary field spanning research and methodologies from engineering disciplines such as robotics and computer science to social sciences disciplines such as psychology and human factors among others. The aim of HRI is to foster natural interaction between humans and robots. However, this is complicated as "robot" is a continually changing concept. The wide variety of robots, among the equally broad range of industries, with constantly evolving characteristics and behaviours require examining from both the human and robot perspective to reveal the intricacies of interaction [55]. As robots start to leave the factory and seep into many aspects of society, this higher and more complex level of interaction between humans and robots requires its own theoretical framework which helps to develop and evaluate the effectiveness of robotic systems [56].

HRI is a relatively young field and only recently has work started on researching HRI in the context of autonomous robotics [57], [58]. However, much of this has focused on improving the robot's perception, cognition and action capabilities during interaction with the ultimate aim of creating a personal service robot with the human aspect of the interaction often taking second place [58]–[61]. HRI is different from the widely studied Human Computer Interaction (HCI) as HRI involves physically instantiated systems that interact with their environment and operate in the real-world [62], [63].

Yanco and Drury [63] classified the HRI of a system as being dependent on team composition, decision support for the user, spatial location and required interaction. The level of human interaction required is inversely related to the LOA of the teleoperation system with constant interaction at low LOAs reducing to supervisory control at high LOAs. Goodrich and Schultz [56] classified the overall interaction between humans and robots as either remote (e.g. field robots) or proximate (e.g. assistants, social robots).

Multiple classifications of interaction in HRI are given in the literature [56], [62]–[64] with similar themes of spatial location, level of automation (LOA) of the system, team structure and task criticality. In this thesis, the focus is on real-time remote interaction using direct teleoperation where the human and robot are not co-located.

2.4.1 Application of HRI to telerobotics

In the recent paper on the status of HRI, Sheridan [24] proposed the following four areas of application for HRI: supervisory control of telerobots for routine tasks (e.g. as employed in Amazon warehouse), robots for nonroutine tasks in hazardous/ inaccessible environments, automated vehicles with humans as passengers (e.g. autopilot systems in aeroplanes) and robots for social interaction in all forms of social environments such as healthcare, offices and education. This thesis is concerned with HRI as applied to the remote control of telerobots/ teleoperators where areas other than nuclear have made promising advances. Space and "search and rescue" (SAR) have been two key areas for HRI in this regard [56]. Research by NASA into long-distance teleoperation with significant time delays has seen success in the form of the semi-autonomous telerobot Mars Rover [65] that is carrying out scientific investigations via supervisory control. SAR robotics have been used to provide rapid response to environmental disasters where it may be dangerous or physically impossible for humans to go [66]. Example of robots used include mobile snake-like robots that are designed to fit through tight spaces in search of victims or structure instability. In medical robotics, the *da Vinci* Surgical System, medically approved in 2000, allows surgeons to carry out complex operations while seated at an ergonomically designed console.

2.4.2 HRI Research Trends

HRI research is prevalent across a variety of fields as robots start to see widespread usage. As robots become more advanced and start to share the same space as humans, safety of the human both physically and mentally has become a concern [67]. As well as safety, there are ethical concerns about replacing humans with robots in the workplace, for socialising with children and the elderly; giving robots authority such as exercising judgment to harm or even kill [68]–[71]. Robots are increasingly becoming smaller and more affordable allowing research to expand into swarm robotics where a distributed architecture is advantageous [72]. Applications include reconnaissance, environemntal monitoring and exploration.

This thesis is concerned with telerobotics in hazardous environments, where the dream of fully autonomous robots is still far away. Semi-autonomous robots with humans acting in a supervisory role look to be in use for the forseeable future [22]. The use of human factors research in their design and development is vital to make them practical for their intended purpose.

2.5 Human factors

According to the International Ergonomics Association (IEA) [73], human factors is:

> the scientific discipline concerned with the understanding of interactions among humans and other elements of a system, and the profession that applies theory, principles, data and methods to design in order to optimize human well-being and overall system performance.

Human factors methodology has been neglected surprisingly often in HRI research, due to a focus on developing novel interface technology, rather than empirical studies to validate theories and provide scientific conclusions about the usefulness of such systems [24]. Over 90% of errors contributing to accidents are human errors, because machines have become more reliable, leaving opportunities for humans to slip-up [74]. Technological advances have led to the introduction of increasingly complex products demanding human factors research and design [24], [41]. Although technological tools are becoming increasingly advanced, they are at times limited by their usability. For example, home entertainment systems continue to get increasingly complicated with the introduction of Smart TVs, HD PVRs, sound systems, various assortments of adapters etc. each with their own set of convoluted menus and controls which can create major usability problems. This is the consequence of systems created by experts, arguably without accommodation for a variety of users, especially from a holistic perspective.

Human factors studies aim to maximise the performance of systems by bringing together knowledge of human cognitive, physical and social characteristics [75]. While other fields in HRI aim to improve interaction between human robots through advanced perception capabilities (computer science) or studying the ethical and trust issues of companionship robots for vulnerable children and adults (humanities), the field of human factors in the context of robotics is concerned with designing robots to achieve symbiosis with humans. The aim is to maximise system performance through the application of behavioural and biological sciences in the design of a human-robot system [76]. To receive the most value from human factors methods in terms of time and money, they are best applied in the early stages of system design [75], [77].

To achieve an optimum or "symbiotic" relationship between humans and robots, not only do machines need to be designed to fulfill a certain task; the design of the robot needs to be unified with the end user as the overall performance of the system is limited by the performance capabilities of the user [78]. Researchers typically lack field knowledge when proposing systems and design systems according to their own beliefs without consulting end users [79]. This leads to sytems that are not suitable for purpose where end users reject the technology as Nagatani et al. [79] found. Their robots were used by engineers responding to the Fukushima Daiichi nuclear plant disaster. The engineers did not trust the autonomous capabilities of the robot, neither shared or full. They also did not need the robot's ability to reconstruct the environment using laser range scanners as they had it memorised. Proposals to use multiple robots were also rejected by the engineers as it would expose more personnel to radiation.

Wickens et al. [77] view the study of human factors as a cycle. This cycle consists of using knowledge of human physiological and cognitive capabilities to identify problems in the human-machine interface and correcting them through applying a solution using one of the following five methods:

- *Equipment design*, changing the nature of the interaction between humans and physical equipment
- *Task design*, changing the tasks assigned to humans
- *Environmental design*, changing the physical environment where humans work
- *Training*, improving the teaching and practice of skills required for the job
- *Selection*, improving the recruitment criteria to optimise system performance.

It is relatively difficult to adapt the task and environment in nuclear facilities due to the strict regulations. However, the design and evaluation of teleoperated robots, the focus area for this thesis, is an attainable goal as well as the recruitment and training practices for operators.

Adams [80] proposes how the *"vast pool"* of existing human factors research from other fields such as complex human-machine systems (e.g. air traffic control, nuclear power plants etc.) can be leveraged for telerobotics. This would entail the use of user-centered design to centre the focus of system design and development on the user. Consideration also needs to be given to understanding the mechanisms that drive human-decision making, maintenance of attention span, workload and situational awareness; and factors contributing to human error.

The complexity of dynamic systems such as telerobots incurs significant strain on human cognitive capabilities. Human factors aims to understand and predict human performance using constructs from the fields of cognitive engineering and psychology [81]. There is a vast amount of empirical research highlighting the effects of these constructs on human performance. These constructs include, among others, the stress and mental workload on the operator when using the system, operator vigilance, the operator's situational awareness of the remote environment, trust in automation and system usability [56], [81].

The impact of these factors on human performance in telerobotic systems in nuclear decommissioning needs to be clarified to satisfy industry that such systems are safe and reliable. In this thesis, as well as overall task performance, emphasis is given to mental workload and situational awareness.

2.5.1 Mental Workload

There is not a single accepted definition of mental workload but it commonly refers to the amount of effort required to complete a task [81]–[83]. It is a relatively old construct that has evolved from being the subject of theoretical research trying to define and measure it, to research exploring its practical application [83]. It is different to performance in that poor performance could be attributed to a badly designed interface while a workload problem stems from requiring operators to hold large amounts of information in memory while performing actions [81].

Measuring mental workload helps to identify areas where the mental load on the operator is high which may be a contributing factor to degraded performance. It can have huge impact in the design of human-machine systems and has previously been used to evaluate aircraft pilot teams which resulted in downsizing of the team from three to two members [81].

A popular notion in human factors research is function allocation [81]; the distribution of tasks between human and robot. Usual function allocation policies involve allocating the robot all tasks within its capabilities with the human responsible for the rest [20], however this is not optimal. Human performance measures such as workload are important for optimally distributing tasks between the teleoperated system [80].

According to Dixon and Wickens [84], it is generally accepted that an operator's workload is highest during direct teleoperation and reduces as a function of the autonomy of a system. However, this reduction is dependent on the reliability of the automation. Previous research has shown that higher degrees of human involvement in tasks leads to higher subjective workload although it improves operator situational awareness [20]. Context acquisition, switching between tasks, can also incur workload on the operator where memory recall is required [84]. Prewett et al. [85] conducted an in-depth review of studies investigating methods to reduce operator workload. They investigated many variables for robot systems such as display design, automated functions and intelligence frameworks. They found that workload could be reduced by using optimal visual displays, multimodal feedback and reliable automation resulting in improved operator performance. However, they cautioned that results are subject to task requirements and the conclusions are inconclusive due to the small samples used in studies.

2.5.2 Situational Awareness

The seminal work of Endsley [86] presents the widely accepted notion of situational awareness (SA) as the "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and projection of their status in the near future". This implies that SA is not merely about perceiving the environment but building upon the perception to gain an understanding of the environment as proposed in her model of SA with three levels of SA:

1. Level 1 SA: perception of the status of environment and its elements
2. Level 2 SA: understanding the significance of those elements with respect to operator goals to form a comprehensive picture
3. Level 3 SA: ability to project the actions of elements into the near future

An operator's ability to acquire and maintain SA depends on the operator, system and task environment. Endsley lists individual factors that contribute to SA as innate ability, training and experience. Other various factors include task complexity, system and interface design; workload and stress on the operator. She notes that although SA can also be limited by an operator's ability to allocate attention and memory resources, this can be alleviated through forming mental models in the long-term through training and experience.

Generally, SA is linked to performance such that an operator's ability to acquire SA in a given context directly affects their overall task performance [86]. SA shows its importance in large complex systems such as aircraft, air traffic control, and large system operations such as nuclear power plants where it is critical that operators have up-to-date knowledge of the current situation [86]. Assessing the level of SA in such systems helps to identify areas of improvement (e.g. interface design) allowing designers to improve operator performance and prevent errors.

Acquiring SA in a control room of a complex system such as a nuclear plant is analogous to telemanipulation [80]. In both cases, an operator is situated remotely and monitors the remote environment through limited sensory information. The level of SA which the operator is able to develop is limited by the quality of the information presented to them.

Poor perception of the remote environment in primitive teleoperated systems using narrow field of view (FOV) cameras coupled with the lack of depth information compromises SA when operators need to estimate distances to targets or size of objects [39]. If the interface presenting the information to the operator is not optimally designed and overloads the operator with information, this leads to cognitive tunneling where some parts of the interface are favoured over others and vital information can be missed [86]. Moreover, high levels of automation (LOA) can also be deterimental to SA as research shows that systems with high LOA significantly decrease SA leading to increased time-to-recovery from failure states of the automated system [20]. This is due to complacency on the part of the operator as discussed below.

Other constructs include trust in automation, automation may not always be used as intended by designers and its use depends on operator's confidence in its capability [81]. High levels of trust is known to lead to complacency and eventual degradation of operator's skills [20], [81], [84]. Stress influences how operators interpret information and can affect performance (negatively or positively) and it can come from physical (noise, temperature, lighting, weather) or psychological (fear, anxiety, uncertainty, workload, time pressure) sources [77].

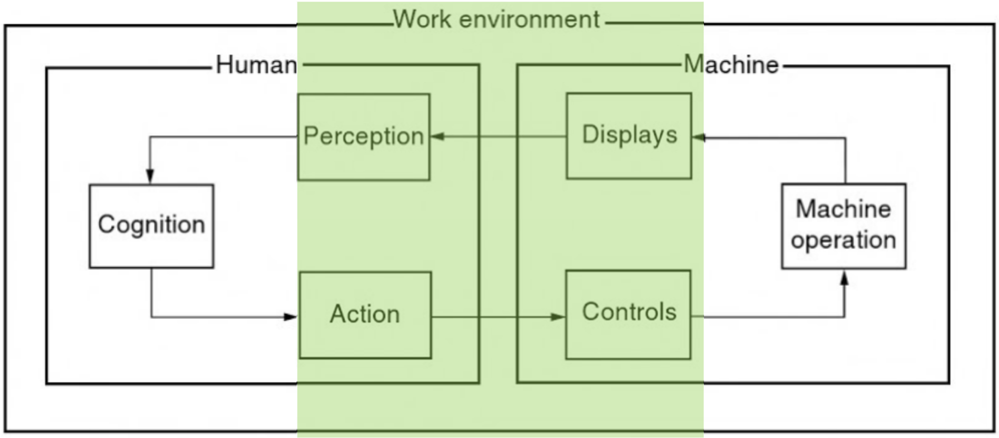


FIGURE 2.2: The human and machine subsystems in a human-machine system. Highlighted is the domain of the human factors specialist. *Adapted from [75]*

2.5.3 Individual factors

Research has shown that individual cognitive differences between operators can have an effect on performance as identified by Chen and Barnes [22]. They identified attention, spatial ability and gaming experience as being able to impact overall performance. Attention is important in human performance as operating a telerobot requires multitasking which requires being able to continuously shift attention allocation to gain SA [86]. Spatial ability has been found to be significant in various areas such as navigation, visual search tasks and overall task performance with higher spatial ability linked to better performance [22], [87]. Finally, frequent gaming has been shown to affect brain functionality leading to positive changes in areas of cognition known to affect task performance in telerobotics [22].

2.5.4 Human-Machine Interface

Proctor and Zandt [75] present a model of the domains of engineers, human performance expert and human factors specialist, fig. 2.2. While an engineer is concerned with the design of the machine and the human performance expert is concerned with human capabilities, the human factors specialist evaluates the interface between the two subsystems, the *human-machine interface (HMI)*. When the machine is a robot, this is referred to as the human-robot interface. In addition, the work environment of the human-machine system is an external factor that needs consideration as it also influences the performance of the system. A minimal HMI for a typical direct teleoperation system consists of monitors for visual feedback of the remote environment and an input controller (e.g. joystick) for sending commands from the operator to the teleoperator.

HMI Design

Human-machine interface experts were surveyed by Steinfeld for insight into interfaces in autonomous and semi-autonomous robots [88]. They unanimously recommended designing the human-machine interface in parallel with the robot system allowing the facilitation of a structure for human-robot communication suitable for non-expert users. They also encouraged the use of displays for providing situational awareness for understanding remote environments but cautioned that it must be kept relevant. Other suggestions included the implementation of safe modes for controlled failure in safety-critical tasks, multiple control methods for command input and status information to keep the operator updated.

The U.S. Nuclear Regulatory Commission reviews human factors engineering in its nuclear sites and provides a detailed document with guidelines addressing the physical and functional characteristics of human-system interfaces in nuclear power plants [89]. Although not designed to cover robot systems, elements such as displays, controls and interface interaction are equally applicable in this domain.

Interface design is important as it plays a role in determining operator SA acquisition capabilities. Endsley [86] gives several guidelines for designing interfaces to maximise SA gain focusing on minimising the effort required by the operator to acquire information. These include, among others, displaying information that needs processing and integration directly rather than relying on the operator to calculate them mentally, present information from the operator's perspective rather than the robot's and using filtering to display only the relevant information in any given context to avoid tunneling operator information.

The guidelines reviewed here form the foundations in the development of teleoperated systems in this thesis.

2.5.5 Human Factors Issues with HMIs

Chen, Haas and Barnes [39] present a very detailed, though somewhat dated (2007), review of teleoperated and semi-autonomous robot interfaces and associated human performance issues classified as either remote perception and remote manipulation (navigation and manipulation). They found that remote perception and manipulation was affected by various factors such as FOV, multiple cameras, viewpoints, lack of depth perception, time delays and motion leading to numerous human factors issues such as degraded performance, increased workload, reduced situational awareness, cognitive tunneling, motion sickness etc. They also gave many potential solutions including, but not limited to, use of wide FOVs, multimodal interfaces (e.g. audio alerts, haptic feedback etc.), stereoscopic displays for improved depth perception and predictive displays to counter feedback delays.

The current trend of moving towards semi-autonomous systems for their benefits in reducing mental workload of operators and increasing their situational awareness (SA) has raised several unique human performance issues

Factor	Effects	Suggested solution
Field of View (FOV)	Erroneous speed & distance judgments; peripheral vision loss; degraded remote driving	· Increase FOV (e.g., Perspective Folding); changeable FOV can be considered · *Caveats:* perceived speed increases & motion sickness
Orientation & Attitude of the Robot	Orientation in the environment; North-up vs. Track-up map; mismatch between actual & perceived attitude of robot; unawareness of (polymorphic) robot's inclination & shape	· Map - Track-up map for navigation; North-up map for tasks involving integration of spatial relations in the environment (e.g., recon, surveying, planning tasks). · Gravity referenced view (GRV) - better awareness of robot's attitude, better route selection, and faster completion of route. · Polymorphic views - operator less likely to tip the robot or have it caught on objects.
Multiple Cameras	Attention switching; change blindness; perceptual registration	Auditory alerts; multimodal solutions & visual momentum techniques
Camera Viewpoint & Frame of Reference (FOR)	Egocentric - cognitive tunneling; exocentric - loss of immediacy & true ground view; integration of info from different FORs may be challenging for operator; saliency effect	Dual mode & inserts of other views (e.g., Sensory Ego-Sphere); peripheral cues for egocentric
Depth Perception	Underestimation of distance & size; degraded navigation, driving, & telemanipulation	· Stereoscopic Displays (SDs) - improved depth perception, obstacle avoidance, arm manipulation, important for difficult terrain & remote arm manipulation; inter-camera distance should be less than inter-ocular distance. · *Caveats:* limited use - benefits mainly for difficult tasks; may induce motion sickness & perceived stress; hyper-stereo SD have multiple negative effects
Video Image/ Frame Rate	Degraded motion perception & spatial orientation; degraded target identification & latency	· Min frame rate: 10HZ · Augmented reality/ Synthetic overlay · SDs (see above)
Time Delays	Task dependent: negative effects range from 170 ms to over 1 sec; degraded driving, tracking, & telemanipulation; over-actuation when delay is variable; robot-to-operator delay more detrimental than the other direction; motion sickness; degraded telepresence	· Minimum -170 ms for driving like tasks; other minimums task dependent · Predictive displays (e.g., Ecological Display) - navigation faster & more accurate. · *Caveats:* disturbances in remote environment may make prediction model unreliable.
Motion	Degradation on accuracy & latency; sometimes severe motion sickness	Multimodal user interfaces; tailor interface to vibratory & motion effects, possible medical remediation

FIGURE 2.3: Factors affecting remote perception and manipulation with potential solutions. *Source: [39]*

[23], [84]. These include increased complacency from the operator leading to failure to detect errors, skill degradation from lack of skill use due to automation, mistrust in the robot's automation, poor communication between human and robot during decision making and task switching leading to reduced SA and increased mental workload [22], [23], [90].

Multimodal interfaces utilising multiple channels for operator feedback rather than overloading the visual channel to improve immersion promise great potential and a review of these is included in [39]. However, these solutions are task-dependent requiring usability testing and research into their effective integration into systems is still ongoing [91], [92].

2.5.6 HMI in Field Robotics

Robot competitions have provided valuable opportunities to evaluate HRI in state-of-the-art robot systems where multiple robots aim to complete the same tasks with different HMIs. Yanco and Drury [93] studied the American Association of Artificial Intelligence (AAAI) Rescue Robots competitions for three years from 2002-2004. From their analysis, they presented the following guidelines for improving HRI in urban search and rescue (USAR) scenarios:

- Utilize a single monitor for the interface.
- Avoid small video windows on the interface.
- Avoid window occlusion.
- Use one robot to view another when more than one robot is available.
- Design for the intended user, not the developer.

When the earthquake and tsunami struck the Fukushima Daiichi nuclear power plant in 2011, robots were not ready to assist in the disaster response. The Defense Advanced Research Projects Agency (DARPA) initiated the DARPA Robotics Challenge (DRC). The DRC was designed to advance the state-of-the-art in USAR robotics to respond to future disasters in a timely manner [94]. A number of mobility and manipulation anthropomorphic tasks were designed to test the abilities of telerobotic systems. Yanco et al. [94] also studied this competition under the scope of HRI and presented the following guidelines:

- Increase sensor fusion;
- Decrease the number of operators;
- Decrease the amount of operator input needed to control the robot;
- Don’t separate the robot into legs and arms;

- Plan for low bandwidth; and
- Design for the intended users.

The guidelines from both competitions hint at the lack of consideration given to human factors in the design of the HMIs. Yanco et al. noted that most of the interaction was "not very far from teleoperation" as operators heavily micromanaged the robots. It should be additionally noted that the robots were controlled by the developers in both competitions as both competitions had complex interfaces leading to the recommendation of designing for the intended users.

2.5.7 Human Factors Evaluation

Human factors evaluation of HRI is almost always empirical [76]. It is based on the measurement of relevant task-specific variables. Performance evaluation can be difficult due to the variability of the results and diverse range of human-robot applications [95]. Successful interaction in the system and the degree to which the end goal is achieved will vary depending on the operators' attitudes and previous experiences with technology. Sound and valid studies necessitates the need for rigorous experiment design using methods from fields such as Psychology and human factors which many current studies lack [85], [96], [97].

The traditional method of performance evaluation is to measure the "time-to-completion" of a task or the overall human-robot system performance (e.g. number of errors) [95]. Applying such objective measures helps to improve the system performance, but does not discriminate or break down which aspects of the human-robot collaboration were more or less effective. Alternative approaches focus on subjective evaluation through interviews and surveys [98]. [85] presents a comprehensive review of studies aimed at reducing operator workload and their shortcomings indicating that most results still remain inconclusive due to differences in experiment design and methodology. To overcome these issues, Bethel and Murphy [96] present a detailed account of designing experiments in HRI with specific recommendations for the experiment design and execution. They suggest the biggest improvements in HRI studies to come from using larger sample sizes and multiple methods of evaluation such as task performance metrics, observations, self-assessments and interviews.

Quantitative Measures

A number of papers [99]–[101] define performance metrics for human-robot collaboration systems with the aim to be able to improve and compare different collaboration schemes with each other. They suggest and define a number of useful metrics, including "task effectiveness" (how well the task was performed), "neglect tolerance" (how effectiveness declines without human intervention), "robot attention demand" (how much time the human must spend interacting with the robot), "interaction effort" and various other quantitative metrics, that assess an interface based on the attention required.

As well as task-specific measures, the human-robot interactions can also be evaluated. Steinfeld et al. [95] developed an extensive framework for evaluating human-robot collaboration. It focuses on task-oriented mobile robots, with HRI metrics split over five categories: navigation, perception, management, manipulation and social. These metrics cover quantitative as well as subjective measures.

Subjective Measures

There are a number of well-established techniques to measure subjective factors such as mental workload, situational awareness and system usability. Cain [102] conducted an in-depth review of workload measurements and categorised them into three categories: performance measures, subjective rating and psychophsyiological. Performance measures can either consist of the primary task measures or the use of a secondary task solely for measuring workload on the operator. Subjective rating measures involve participants rating their own perceived workload on a numerical or categorical scale. Finally psychophsyiological measures are more intrusive but also the most objective. These measure physiological features such as heart rate, neuron activity etc. and can be difficult to administer. However, the technology behind these measures is improving so this method may become more attractive in the future [102].

The most widely used method for workload measurement is the National Aeronautics and Space Administration Task Load Index (NASA-TLX) [82], [103]. In this questionnaire, participants subjectively rate the workload on six different sub-scales [82]:

- Mental Demand: how mentally demanding was the task,
- Physical Demand: how physically demanding the task was,
- Temporal Demand: how hurried or rushed the pace of the task was,
- Effort: how hard the participant has to work to accomplish the level of performance,
- Frustration Level: how insecure, disorganised, irritated, stressed, or annoyed the participant was, and
- Overall Performance: the participant's view of how successful in accomplishing the task the participant was.

The raw scores are within a 100 points range (very low: 0 to very high: 100). Additionally, the NASA-TLX employs 15 paired comparisons between the subscales to account for between-rater variability: differences in workload definition between the participants and also differences in the sources of workload between the tasks. A variant developed by Byers, Bittner and Hill in 1989 ignores the pairwise comparisons and in subsequent testing found no significant differences to the original NASA-TLX [104]. The questionnaire can be administered after each trial or at the end of the experiment.

There are many techniques for performing SA analysis but the most popular is SAGAT [104]. The SAGAT technique was formally introduced by Endsley in 1988 [21] and has since gained widespread use in the research community after demonstrating its empirical validity [20], [80], [105]. In the SAGAT assessment, queries are constructed based on task goals using SA requirements analysis. Then during task execution, the freeze probe technique is used to freeze the task at pre-determined random points to administer the queries and calculate the SA score.

As the focus in this thesis is on analysing the human-machine interface (HMI), the HMI needs a measure for evaluation. Interface analysis techniques can be used to assess interfaces across various aspects such as usability, user satisfaction, error, layout, labelling and the controls and displays used [104]. The System Usability Scale (SUS) [106] questionnaire is a simple ten-item Likert scale designed to give an overview of the subjective assessment of usability across different systems. The score is calculated out of 100 using the procedure mentioned in [106]. According to reviews of studies using SUS as a usability measure [107], [108], a score of 68 is considered average. As well as being a reliable and valid measure of usability, it has recently been found to also measure learnability, the ease at which users can learn to use the system [109].

Human factors issues play a key role in the design and evaluation of the telerobot HMI. The interface for controlling the robot in part determines the workload on the operator, the stress encountered during a task, when fatigue comes into play and the overall performance during execution. The principled evaluation of the HMI, its effects on human cognition and the overall performance *of the combined human and telerobotic system* using human factors methods is the goal of this thesis.

2.6 State-of-the-art Robotic Applications in the Nuclear Industry

Due to the extremely hazardous nature of legacy nuclear environments, in many cases, the only means of achieving decommissioning is through the use of automation, robotics and "remote engineering" in order to reduce the dose exposure of workers [110]. As well as decommissioning, nuclear fission and fusion life cycles for new build reactors will require significant use of robotics and automation [111]. Other areas of robotic applications include inspection, plant maintenance and waste disposal [112].

The level of technology associated with the use of telerobotics within the nuclear industry has remained stagnant compared to other industries. Remote manipulation systems range from mechanical MSMs to teleoperated wheeled/tracked robots [7]. Rather than investing in robotics research, the nuclear industry has previously chosen to apply off-the-shelf technology from other industries, such as defence and off-shore, modified to serve a single specific task [113]. This has limited the industry to the technology currently available, which is suboptimal to requirements, and is a significant factor for the limited number of "real robots" have been deployed in the nuclear industry.

A number of teleoperation issues have been identified that, in terms of the decommissioning of nuclear plants, make the technology slow to develop and expensive to use [113]. For example, it may take a team of several highly trained operators to operate a single teleoperated robot [114]. Also, due to the complexity of working in an unstructured and hazardous environment and the inadequacies of the HMI, teleoperation is extremely fatiguing for the operators because of the cognitive load on the operator to carry out each movement of the slave manipulator while observing via a CCTV system [113]. What follows is a discussion on the current applications of robotics in the various areas of the nuclear life cycle.

2.6.1 Inspection

Recent telerobotic developments for the nuclear industry have focused on inspection and maintenance tasks. Nuclear regulations regarding the inspection of reactor vessels have necessitated the industry to employ specialist nuclear engineering companies to find economical robotic solutions for these challenging operations [112]. These systems have seen mixed success. Noteworthy solutions include the OCRobotics' SAFIRE robot used for pipework inspection in CANada Deuterium Uranium (CANDU) reactors. SAFIRE comprises a mobile platform with a"snake-arm" mounted on top. The arm contains a camera at the tip to provide a visual inspection of the remote scene. Another solution is the WesDyne's SUPREEM system, a submersible platform with two manipulators used to inspect the structural integrity of underwater vessels [112].

Sellafield Ltd., which manages decommissioning at the Sellafield plant in the UK, have begun testing a UAV system for 3D imaging with radiation overlays with positive early results and have been using underwater remotely operated vehicles (ROV) for inspection as well as simple manipulation tasks [115].

2.6.2 Remote Handling

Many nuclear industry operations take place in "hot cells", which are shielded rooms, where nuclear materials can be safely handled remotely. Such handling is predominantly supported using mechanically coupled MSMs. Most cells that contain electro-mechanical systems for handling fuels and wastes (as opposed to process vessels and pipe work) are serviced by overhead cranes and may have the additional benefits of gantry-mounted telemanipulators. The control systems for these devices offer very little sensory feedback to the operator and there is a high reliance on the operator viewing the activity via CCTV or through very thick shielded windows. Such operations are consequently very slow and tedious [116].

Remote intervention activities are generally carried out using bespoke systems developed specifically for the task, and consequently are expensive both in terms of cost and timescale, unreliable, and require highly skilled operators. For example, Sellafield Ltd. have installed a Silo Emptying Plant robot, weighing 400 tonnes, to retrieve radioactive materials from a swarf storage silo [117]. Additionally, Wayne Ingamells (industrial supervisor of this doctorate, and Nuclear Decommissioning Specialist, 3i Technology) worked on a problem where a specially commissioned robot arm was used to reach over a shielding wall and achieve precision cutting, cleaning and resealing of deteriorating pipework and leaking concrete wall of a legacy storage pond under high radiation dose conditions [113], [115]. However, this was an expensive bespoke system and required many months of modifications and the equivalent of 80,000-hours of training and test runs in a mock-up facility.

2.6.3 Decommissioning

As expressed before, the use of robotics in decommissioning is to limit the exposure of personnel to radiation, reduce secondary waste and improve productivity. The case for robotics is even stronger where continuous decommissioning strategies are employed due to the level of radiation, making it infeasbile for personnel to carry out work [112]. However, most telerobotic systems for decommissioning remain directly teleoperated [112], [113]. One of the first teleoperated systems used was the NEATER robot [118], a radiation-hardened industrial manipulator intended to replace previous bulky and expensive bespoke manipulators for remote operations.

Currently available off the shelf mobile demolition machines, such as the Brokk are heavily used but little is done to meet the requirements of the operators and the control systems offer little more than basic switch-box control and a CCTV feedback loop [110], [112]. Furthermore, Brokk machines are missing basic sensors such as rotation encoders at joints that enable inverse kinematics, so operators must control each degree of freedom individually by hand using levers, i.e. Cartesian control with joysticks is typically unavailable on such machines.

In a small number of cases, modern industrial robot arms, with gradually more modern control methods, are starting to be adopted by the industry. New work at UK National Nuclear Laboratory, commissioned by Sellafield Ltd, is currently experimenting with bespoke systems for joystick teleoperation (with CCTV camera views) of large KUKA KR500 robots. However, baseline industry standard for controlling such robots still remains limited to pushing buttons on an industrial teach pendant to control translations/rotations along/about each Cartesian axis. [paragraph adapted from [119]]

2.7 Discussion

The meltdown of the Fukushima Daiichi Nuclear Power Plant in 2011 was a wakeup call for the robotics community. The subsequent radiation release necessitated the use of mobile robots for the disaster response effort, yet none were available to take up the challenge [79]. This failure to respond forced robotics companies such as Honda, famous for its ASIMO humanoid robot, to shift focus from the entertainment sector to the more practical search and rescue robotics, and inspired competitions such as the DARPA Robotics Challenge to fast-track developments in search and rescue robotics [120].

There are now a variety of innovative telerobotic systems for monitoring and inspection, but there is a clear lack of robots for remote manipulation. In addition to this, current bespoke systems are bulky, expensive and difficult to maintain.

Although telerobotic systems which make use of modern technological advances have been proposed in the literature, they are either limited to functional demonstrations, lack rigorous human-subject testing or are limited to the artificial laboratory environment [96], [121]–[124]. Yet there is a great need for such systems as demonstrated by Uematsu, Kashiro and Tobita [12]. They used a combination of human personnel and robots to dismantle gloveboxes. An "arm-type robot and three manipulators" were used to carry out 20% of the dismantling tasks such as removing bolts and cutting panels. They showed that compared to conventional dismantling, radiation dosage reduced by 57%, secondary waste generation by 76% and personnel costs by 65%.

Historically, the application of telerobotic systems for nuclear decommissioning has met with opposition due to insufficient evidence of their performance and reliability [125]. However, as shown by the introduction of new systems in various areas of the nuclear life cycle, the opposition to robotics is starting to curtail with specialist nuclear engineering companies often leading the advances. This presents a unique opportunity for academia to create an impact by operationalising its novel research and validating its utility through principled human factors analysis and human-subject experiments. This thesis is intended as a step forward in this direction.

Chapter 3

Testbed Development

The aim of this thesis is to systematically evaluate telerobotics through human subject experiments. The pursuit of this goal requires a platform that can be used for rigorous and repeatable testing of existing systems as well as new developments; this is known as a testbed. This chapter presents the development of the testbed to be used as the foundation for these experiments.

3.1 Task requirements

In order to make the evaluation of the telerobotic system relevant, it is helpful to define a set of scenarios that the final system is likely to encounter.

Each nuclear facility is unique in its structure and contents because it served a specific purpose so it is difficult to generalise their decommissioning process since their structure and inventory differ. However, the process of decommissioning can be split into phases. These are:

1. Inspection- Determine structural integrity, inventory and spread of radioactive contamination

2. Removal of non-structural items- Clear facility of non-structural equipment and contents into waste storage

3. Removal of structures- Dismantle pipework, ventilation and gloveboxes and place into storage

4. Demolition- Demolish building and ground remediation

This outline can be used to generate the following scenarios for a telerobotic system:

Scenario 1: Facility inspection The robot is used to perform an inspection of the facility to determine its structural integrity, radiation in the environment and contents.

Scenario 2: Clearing the facility The robot is used to remove any radioactive sources to minimise the radiation and dismantle non-structural items in the facility.

Scenario 3: Cutting pipework and other structures The robot is used to disconnect the facility from the pipe work and ventilation systems and systematically dismantle the rest of the facility with all waste put into storage containers.

3.1.1 End Users

With a particular focus on Sellafield, it is envisaged that there are two groups of end users. The first group consist of current personnel working on manual decommissioning who may be retrained to work with the telerobotic system as well as current workers involved in the use of MSMs for decommissioning in "hot cells". The second group consist of new recruits who will eventually have to take over as the current batch of personnel retire. These novice recruits will require training on the use of the telerobotic system.

The average worker currently working in low-level decommissioning position fits the following profile [113]: caucasian male, middle-aged, college-level education or below and live in the surrounding rural countryside, mainly from Cumbria. They are likely resistant to change in their work, strict about their 9-5 hours, not go above and beyond their job requirements and generally have a laid-back attitude towards work. It is predicted that future workers will also fit around this profile.

Other than the workers who currently work with MSMs, it is assumed that these workers will have little to no experience with similar systems and will require extensive training. They will also have their own preconceptions of how the robot should work. Therefore, it is essential that their feedback is utilised during the design and development of the telerobotic system so it facilitates the end-user cognitive processes.

3.2 Testbed

The purpose of the testbed was to replicate the current state of the art of robotics in the nuclear industry. This would then serve as a benchmark for further iterative development and testing through controlled experiments and user feedback.

The testbed needed to be a typical teleoperation system consisting of a remote and local environment as described in section 2.1 and representative of systems in current use in the nuclear industry as shown in the video in [126]. The robot arm manipulator is situated in the remote environment along with cameras for visual feedback. The human operator remains out of line of sight of the robot in the local environment. The local environment consists of the operator in front of a workstation which displays the remote environment and sends input commands to the robot through an input device.

The functionality required from the testbed was as follows:

- Display the robot in the remote environment to the operator using basic webcams
- Feed input commands from the input device (i.e. mouse & keyboard) to the robot

- Allow the operator to interact with the remote environment.

3.3 Platform

The platform forms the foundation on which the testbed is to be developed and then further improved upon. In HRI experiments, the robot system can either be a real robot or replicated in a simulation.

3.3.1 Simulated vs. Real-World Robots

Before the testbed could be developed, a decision needed to be made on whether to use real robots or a simulation environment. In making this decision, a number of factors need to be considered.

Development

At the time this project commenced (2011), using industrial manipulators for research was extremely laborious and lacking support. As these robot arms were designed for use on factory floors using preprogrammed instructions, there was a lack of documentation and after-sales support (e.g. software, drivers etc.) to re-purpose them for research. This barrier was nearly impossible to overcome except for those with close ties to the robot manufacturers themselves.

In contrast, using a simulation environment is much easier as it removes the need for expertise in interfacing with the propriety programmable logic controllers (PLC) of the robots. Unlike with physical robots, most simulation environments allow for plug and play functionality with robot models with some even incorporating a library of 3-D models of robots, components and environments. They allow for rapid prototyping development and some even offer deployment of code to real robots.

Robustness

As well as being difficult to work with, industrial manipulators are extremely fragile making them unsuitable for human-robot interaction. As the robots do not come with compliance control built-in, until it is implemented, they are susceptible to irrepairable damage if they come into contact with the environment as well as causing harm to the environment. This is problematic as safe manipulation is the objective of any telemanipulation system, especially in a nuclear reactor. Moreover, industrial manipulators must function under the constraints set by the manufacturer in the PLC such as real-time control, speed limits etc. Violation of the constraints can cause the robot to halt and require lengthy resetting procedures making it unsuitable for use in experiments during the early stages of development. Using simulation can avoid these issues completely.

Accessibility

A major issue in research laboratories can be managing access to the physical robot as priority is often given to senior researchers or other students if the robot is not owned by the supervisor. Also, robots often have significant periods of downtime when firmware is being upgraded or the robot has to be sent to a different country for repairs.

Simulation removes the need for access to a physical robot saving considerable time. Although simulators can be missing models of a particular robot, unless the robot is really niche or developed in-house, a model is usually available either through the robotics community or from the robot manufacturers themselves.

Experiment Control

For human-robot interaction experiments, simulation offers stricter controls of the experiment environment minimising confounding variables such as lighting conditions, emergency stoppages due to loss of control, collision damage etc. This improves the validity of results as well as repeatability of experiments. It also offers the perfect environment for introducing novice users to remotely operating robots through training without the risks involved in using a physical robot.

Simulation Fidelity

The major drawback of using simulation is fidelity. Although simulation can replicate most essential aspects of an environment, other aspects such as physics are yet incomplete and can be prone to erratic physics behaviour at random. Additionally, the highly safety-oriented nuclear industry is not likely to accept a system that cannot be shown to be deterministic. For this reason, a simulation can be supplemented by deployment onto real robots during the later stages of development for testing and evaluation.

Safety

Using industrial manipulators in research and human subject experiments requires extensive risk assessment and safety assessments. Originally intended for the fast-paced industrial environment, they are able to cause significant harm and/or loss of life if they come into contact with humans or the environment during operation. This does not include the damage caused to the robot itself leading to costly repairs if not irrepairable damage.

Simulation environments are also known to cause simulator sickness but this is a significantly lower risk as it is not common in screen-based simulators. For this reason, simulation can be much safer for development and human-subject experiments with novice users who are likely to lose control of the robot.

In summary, simulation is a safer and efficient method during the early stages of development which can be easily extended to physical robots for testing and evaluation. In addition, simulation can continued to be used with physical systems for training and review purposes.

3.3.2 Review of Robot Simulation Environments

Robot controllers have a specific set of requirements that separates them from other software with the chief requirement being real-time control. Numerous robotics simulation environments (RSE) have been developed to facilitate these needs. In order to select the most appropriate RSE for this project, a review of the most complete and popular packages was conducted. The following requirements were used to limit the number of RSEs to a manageable amount during the review.

Firstly, the RSE had to be general purpose as opposed to being limited to specific hardware (e.g. vehicular robots) or purpose (e.g. OpenRAVE- motion planning algorithms). In addition, the RSE had to be well-documented with an active developer community so that it was continuously being developed and any problems could be resolved in a timely manner. It also had to support human-robot interaction, be it through native support for interfaces and input devices or through plugins. An easily extendable and modular architecture allows for users to quickly add specific missing components as needed. A scene editor which allows 3D meshes to be imported as well as interacted with during simulations is vital for constructing realistic interactive environments . Many RSEs are limited with regard to human-robot interaction as they are oriented towards development of autonomous robots and

	MSRS 4	Gazebo 1.2	V-rep 2.6.7	Webots 7	USARSim
Platform	Windows	Multi	Multi	Multi	Windows, Linux
Main language	C# .NET	C++	Lua	C	C++
Licensing	Free	Open source	Free for edu use	Paid	Paid
Setup difficulty	Difficult	Difficult	Easy	Unknown	Medium
Scene editor	Yes, clunky UI	Yes, but limited	Yes, very polished	Yes, but limited	Yes
Rendering engine	DirectX 3D	OpenGL	OpenGL	Propriety	Unreal engine
Physics engine	PhysX	ODE	ODE, Bullet	ODE	PhysX
Extendibility through plugins etc.	Medium	High, ROS support	High, ROS support	High, ROS support	High, ROS support
Supports robot manipulators	Yes, limited models	Yes, limited models	Yes, wide range	Yes, limited models	Yes, but none built-in
Model library	Large but messy	Fairly limited	Large, organised	Large	Large
Documentation & Tutorials	3/5	2/5	5/5	4/5	2/5
Support & active community	2/5	5/5, through ROS	3/5	5/5	1/5
Integrated algorithms for robotics	No	No	Yes	No	No
Real-time support	No	Through plugin	Yes	Yes	No

TABLE 3.1: Comparison of robotics simulation environments (RSE). *MSRS = Microsoft Robotics Studio*

/or algorithms. Finally, the simulation environment needed to be in 3-D utilising physics engines for high-fidelity so that users felt a sense of immersion. Table 3.1 shows a summary of the popular simulator environments and their characteristics at the time of review (Dec,2012).

V-rep and Webots were both suited to the task being high-fidelity simulators with excellent documentation and developer support. V-rep won out due to being free for use as well as having short start-up time. Although Gazebo has lived up to its potential of being a successful RSE widely supported by the ROS community, at the time it was very difficult to use with bugs and obscure dependencies. Additionally, missing documentation and need to learn a complex framework (ROS) potentially prolonging start-up time meant that it was overshadowed by V-rep.

3.3.3 V-rep Simulator

The Virtual Robot Experimentation Platform (V-rep), fig. 3.1 is publicised as a "versatile and scalable robot simulation framework" [127]. It combines a high-fidelity simulation environment with built-in robotics functionality and a multitude of ways to program controllers via its distributed architecture conveniently in one file. The benefits of using V-rep for this project include quick setup using the extensive documentation, developer support and tutorials; rapid environment construction using the drag'n'drop scene editor, support for a wide variety of robots, sensors and input devices for use in human-robot interaction in the built-in model library and pre-installed modules covering common robotic techniques such as inverse kinematics and collision detection to allow easy handling of robot logic.

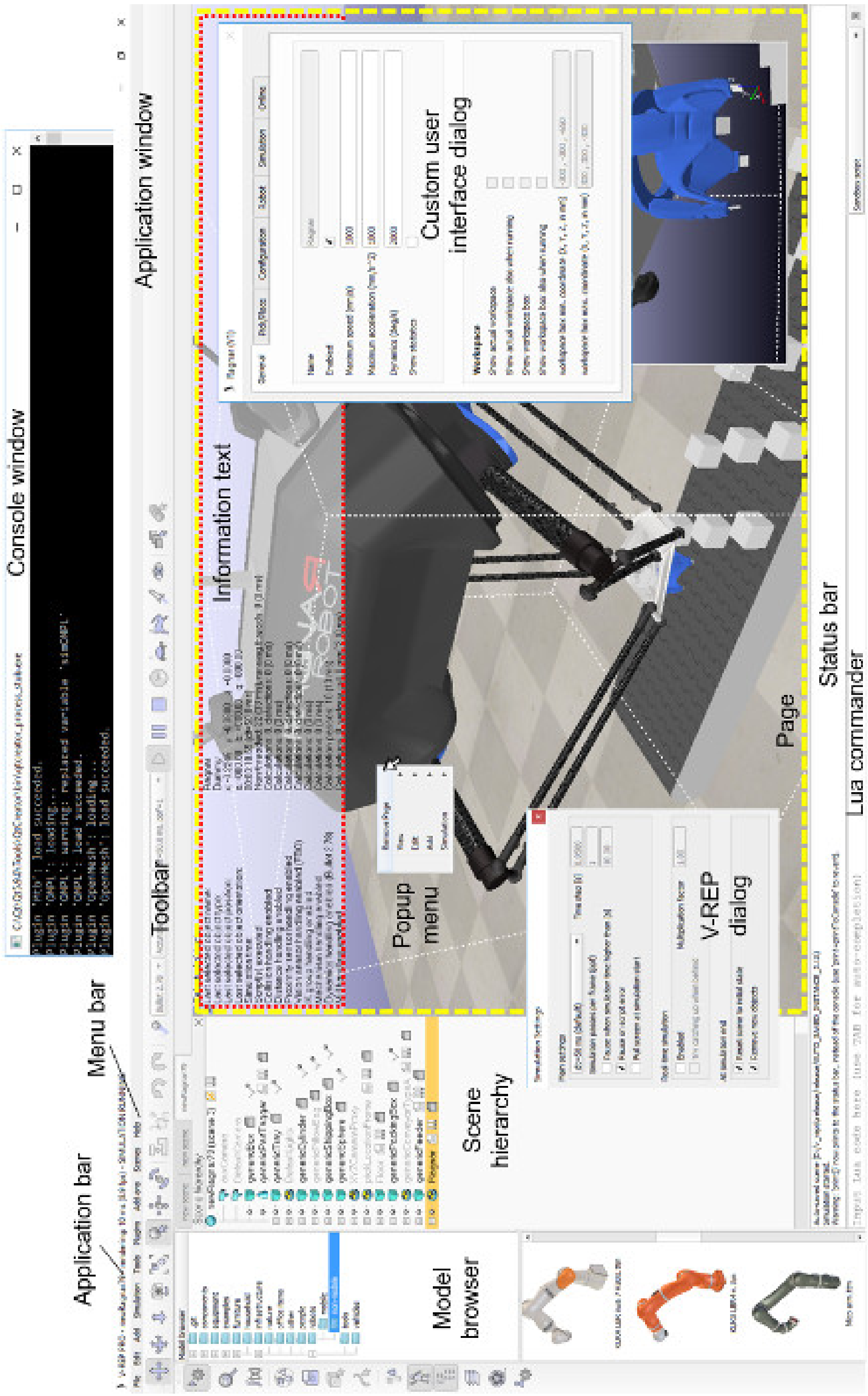


FIGURE 3.1: V-rep user interface. *Source: [128]*

V-rep has the ability to interface with numerous input and output devices. These include common devices such as mouse & keyboard, joystick and cameras as well as haptic devices, Kinect cameras and transceivers. In addition, V-rep offers modules to quickly perform a variety of robotic functions without having to interface with external libraries. Among these are the inverse kinematics module to convert from cartesian coordinates to joint values, the collision detection module to track and visualise object collisions and a motion planning module to plan motion with respect to constraints. The versatility allows for swift development of the robotic system so that the focus can be on evaluating the human-robot interface.

3.4 Architecture

The V-rep application programming interface (API) provides numerous ways of programming control of the robot simulation resulting in considerable flexibility and extensibility (fig. 3.2). Embedded scripts within V-rep can be combined with plugins, remote clients using sockets and ROS to allow control of the simulation from outside V-rep (e.g. from real robots) and/ or extend the simulation with additional features not included with V-rep (e.g. stereoscopic HMDs such as Oculus) as required.

The initial architecture for the benchmark testbed is depicted in fig. 3.3. Its simple requirements meant that it was encompassed within V-rep as all functionality required for the initial testbed was provided within. The remote environment was constructed using the scene editor and displayed within V-rep through simulated cameras. Control of the robot was programmed in an embedded script attached to the objects in the scene.

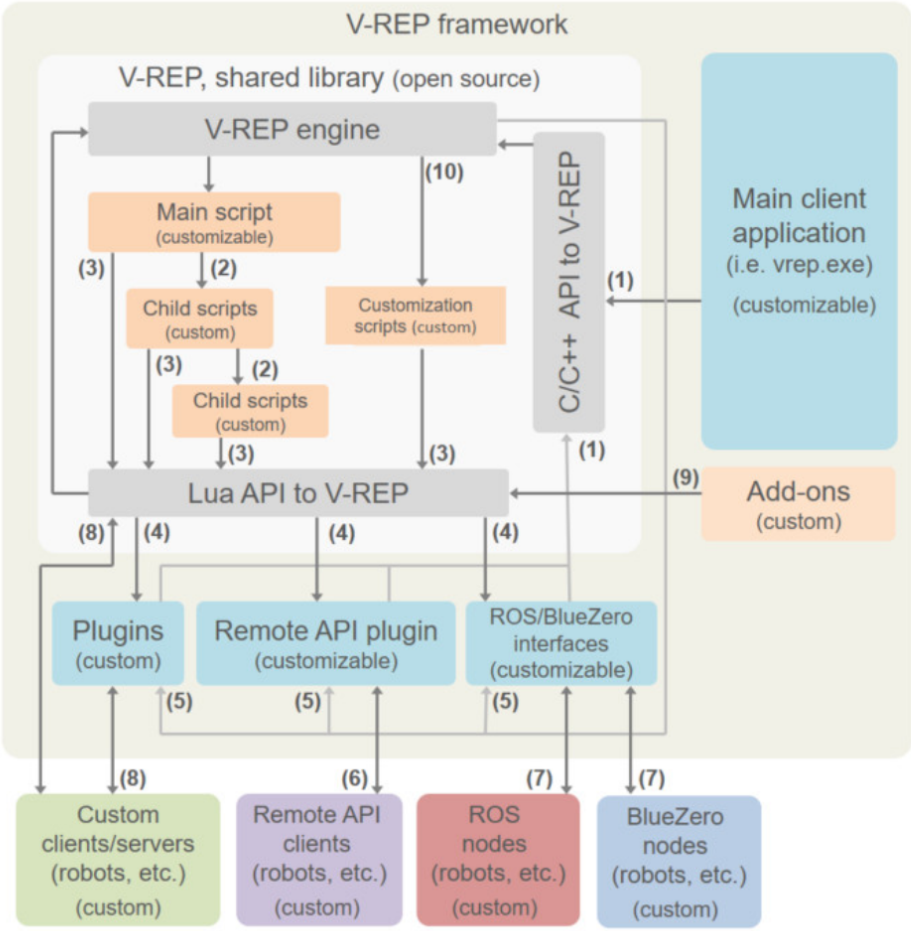


FIGURE 3.2: V-rep architecture. *Source: [128]*

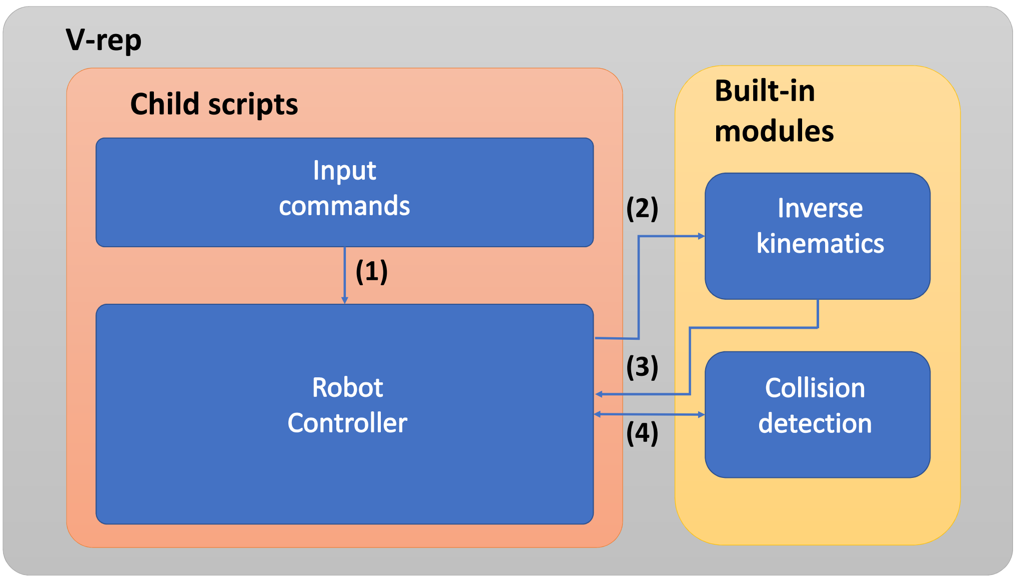


FIGURE 3.3: Testbed architecture. Input commands are converted to joint positions using inverse kinematics (IK) which are checked for collisions by the built-in collision detection module.

Chapter 4

Performance Evaluation of a Conventional Teleoperation Interface

This chapter is a reformatted copy of the following publication: *Talha, M., Ghalamzan, E. A. M., Takahashi, C., Kuo, J., Ingamells, W., & Stolkin, R. (2016, October). Towards robotic decommissioning of legacy nuclear plant: Results of human-factors experiments with tele-robotic manipulation, and a discussion of challenges and approaches for decommissioning. In Safety, Security, and Rescue Robotics (SSR-R), 2016 IEEE International Symposium on (pp. 166-173). IEEE.* [119]

This chapter presents a pilot study to evaluate a replica of current primitive telerobotic systems in the nuclear environments. The Baxter robot is used as a teleoperator in conjunction with V-rep to remotely complete a block stacking task and task performance is evaluated based on a number of performance and subjective metrics. The results and feedback from users are analysed and recommendations on improvements for potential systems in this sector are offered.

4.1 Related work

There are a variety of innovative telerobotic systems for monitoring and inspection but there is a clear lack of robots for remote manipulation. Current telerobotic arm systems consist of an industrial manipulator controlled using joysticks or push buttons via the use of CCTV cameras placed around the task area. Such systems can lead to cognitive overload of operators who have to micromanage the robot whilst also performing a potentially dangerous task, under conditions of severely limited situational awareness. One illustrative example might be trying to mentally merge images from multiple cameras to create a mental scene model, to determine how the robot will move; painstakingly moving each joint of a robot individually with a non-intuitive controller; using this system to weld a leaking pipe without any haptic feedback and only a 2D monitor for scene monitoring. To fully understand the underlying problems with such primitive systems, a testbed was built where current system performance could be evaluated through principled experimentation.

An experiment was conducted to evaluate the extent to which humans can control such baseline control interfaces to perform complex remote manipulation tasks using modern robots. Contributions from this chapter highlight shortcomings of such teleoperated manipulation interfaces and suggest how human factors principles can be applied to develop more sophisticated telerobotic manipulation systems in the future.

4.2 Experiment

Although the testbed was intended to be developed purely in simulation as mentioned in section 3.3, the introduction of the Baxter robot [129][1] changed

[1]Full specifications of Baxter can be found at: http://www.rethinkrobotics.com/baxter/tech-specs/

the landscape for human-robot interaction research. Originally intended for industrial factories, the Baxter robot quickly gained popularity in the research community thanks to its integration with ROS. Unlike the rigid industrial robots, the Baxter robot was designed from conception with human-robot collaboration in mind significantly removing/ reducing the issues present in other industrial robots. It allows humans to work in close proximity to the robot removing the need for safety cages as well as offering a more programmer friendly interface for communicating with the robot and development support.

The Baxter robot is an optimal solution for human-robot interaction experiments as it inherently supports backdrivability. This unique feature means the robot is able to accommodate interaction with the environment as collisions no longer result in damage to the robot. It consists of two 7-DOF arms each with a payload of 2.2kg and force sensors on each joint. In addition, there is a camera in each hand with a third mounted on the head. This combination means the robot is well-situated to providing the remote interface for a telerobotic system as well as a more realistic reconstruction of real-life scenarios.

This experiment sought to evaluate the performance of human participants, using a control interface similar to the industry baseline (push buttons to control axial motions) in teleoperating a robot to execute fundamental remote manipulation tasks. Each participant was tasked with grasping and stacking five wooden cubes to assemble a tower. This activity incorporates elements of popular pick-and-place tasks, but the precision and delicacy needed to stack the blocks successfully entails precise and complex manipulations which better represent those needed for complex decommissioning tasks. The task involves primitive robotic movements, including translation and rotation of the robot's end-effector.

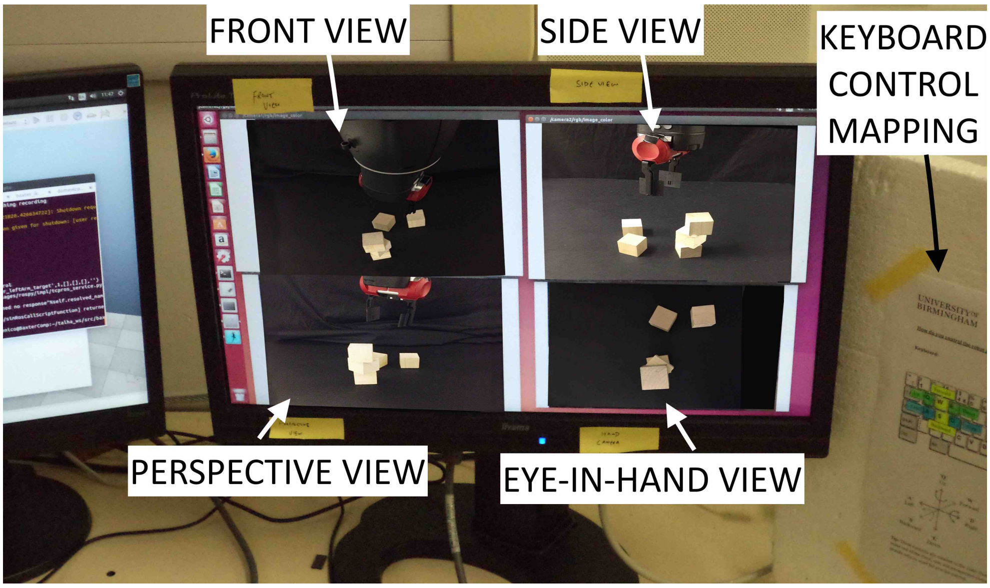
4.2.1 System setup

The experimental setup comprised a master control interface (computer and monitor with robot controls and camera feeds) local to the human operator, which is referred to as the "local workspace", and a slave robot used to perform manipulative tasks in a location remote from the human operator, referred to as the "remote workspace". The remote workspace comprised the robot along with the blocks situated within a frame on a bench top. Fig. 4.1 shows both workspaces.

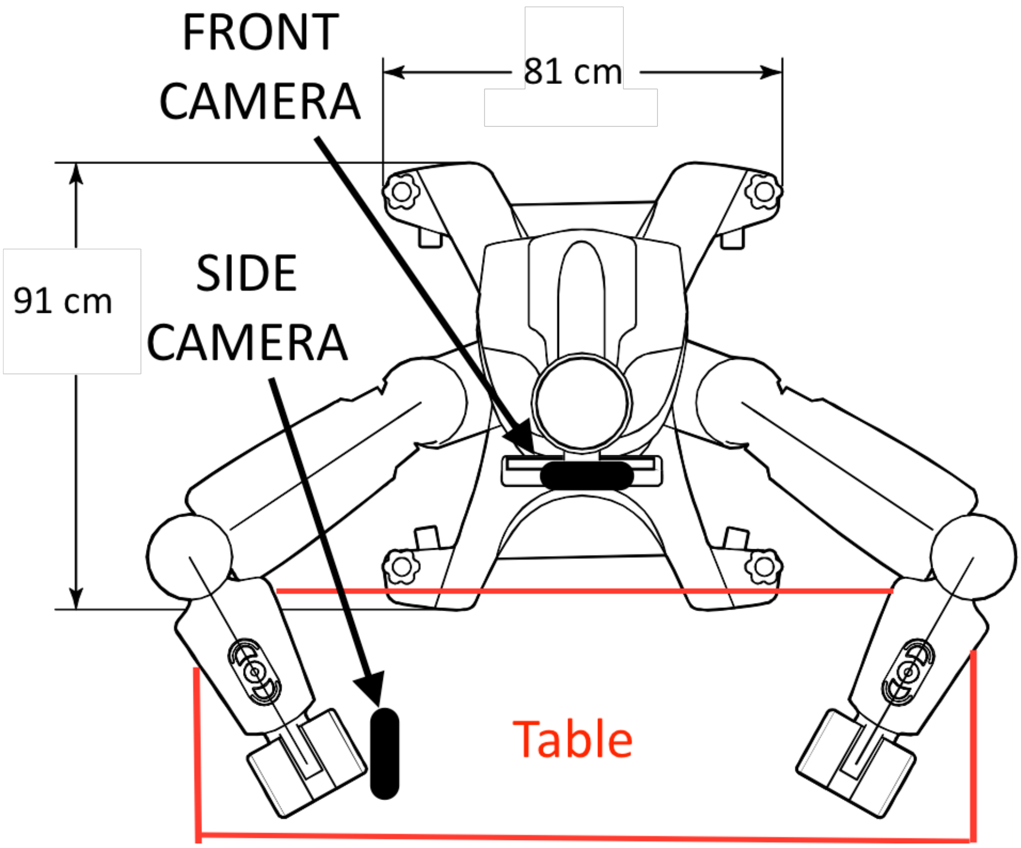
For this experiment, only Baxter's left arm was used, with the gripper width adjusted to a range of 34-68mm to accommodate the 40mm wooden cubes. Similar to many real nuclear deployments, participants had no direct line of sight to the remote workspace, and relied on multiple CCTV camera views, fig. 4.1(a), for situational awareness. Two cameras provided orthogonal views of the remote workspace: first camera attached to Baxter's torso, facing forwards and angled downwards toward the table to provide a front facing view; second camera placed orthogonally on the table to provide a side view. A third camera was fixed above the side view camera, providing an overall view of the scene from a more distant viewing position; and Baxter's left wrist camera provided an eye-in-hand view which moved with human-controlled end-effector motions.

The slave robot was controlled from the local workspace by the human operator pushing buttons on a keyboard. Three pairs of keys were used to control Cartesian X, Y, Z motion relative to the front camera. An additional key could be used to toggle the function of the first three key-pairs between translation along these Cartesian axes, or rotation of the robot's end-effector about the respective axes.

The robot was connected via the lab network to a standard PC using Robot Operating System (ROS). The Baxter SDK was used through ROS to



(a)



(b)

FIGURE 4.1: Experimental setup: (a) master control interface (local to the human) comprising a monitor showing four different camera views and a paper explaining keyboard controls; (b) view from the top of slave robot and remote workspace.

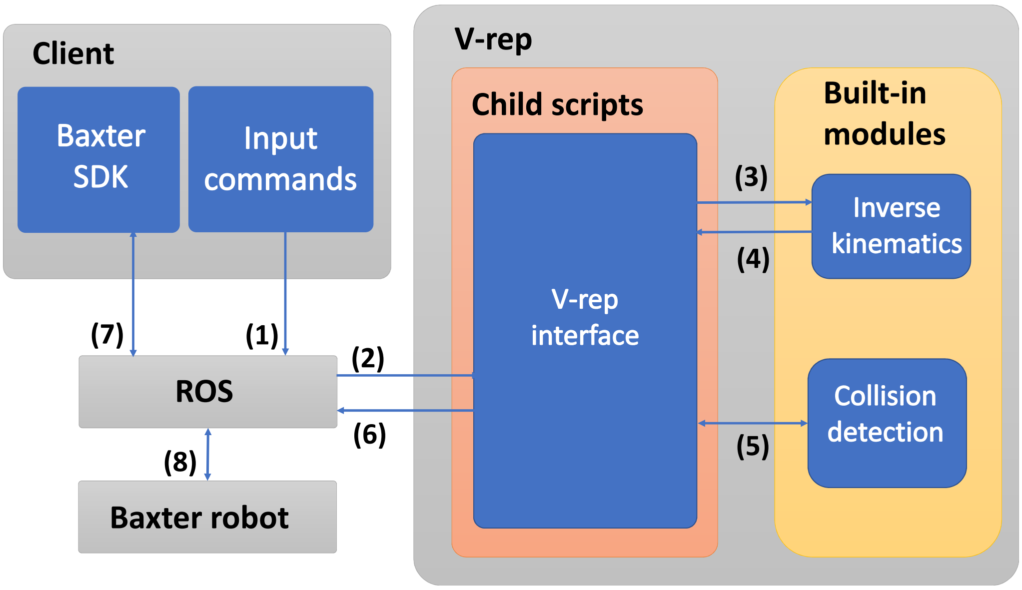


FIGURE 4.2: Pilot experiment architecture. Extension of the testbed architecture by using ROS as middleware for communicating between joint positions generated by V-rep's IK and sending positions to the physical Baxter robot.

communicate with the robot. V-rep was used to calculate inverse kinematics for the robot arm. The overall software architecture can be seen in fig. 4.2.

4.2.2 Block Stacking Task

The block stacking task was chosen because it includes the primitive movements of the robotic arm relevant to the nuclear glovebox context. While being easy to perform and administer for the first experiment, it is deep enough to allow for a human factors analysis and set up the foundation for future experiments involving more complex tasks.

The blocks were situated to ensure they were within the reachable workspace of the robot arm. All five 40mm wooden cubes were randomly distributed within the reachable workspace of the robot at the beginning of each experimental trial. A trial was considered successful if the tower was able to stand freely after placement of the final block and removal of the robot gripper. Each trial commenced once the participant sent the first movement command to the slave robot. Each trial ended once the participant pressed a termination key. Participants could decide to start a new trial, either after finishing a successful tower, or whenever they felt unable to continue controlling the slave robot. In most cases, unsuccessful trials occurred when the robot ended up in awkward configurations which were unintuitive to control from the provided camera views.

Dependent variables, adapted from [95], [99] were measured during the task: Time to Completion (TTC) is time for completion of a trial, from initial robot movement until the finished tower; Number of Collisions (NOC) includes collisions with the table as well as unintentional collisions with blocks; Number of Dropped objects (NOD); and Number of Critical Errors (NCE) (any incidents where the robot must be emergency stopped by the experimenter).

4.2.3 Participants

A group of 16 (4 female, 12 male) participants volunteered from the University of Birmingham student and staff population. Participants' aged in the range 18-50. Each participant was compensated for their time with a £10 Amazon voucher.

4.2.4 Method

Prior to experiments, each participant completed a pre-screening questionnaire comprising 9 questions, covering background information such as age, illness/disabilities and previous experience with 3D games, applications and simulators. Participants were then given a five minutes assessment to determine the extent of their spatial abilities, using images selected from Google images using the "spatial awareness tool" query. Two questions assessed 2D rotation, two assembly and four 3D rotation/folding. A final question consisted of matching 2D rotations between two groups of 25 images.

Participants were briefed by reading a standard instruction sheet, and receiving verbal instructions on the robot controls. Participants were told they could choose to terminate a trial, if they became stuck in an awkward robot configuration or other frustrating situation, after which the robot was reset to its home position. To stabilise performance, participants were given 10 minutes of training time prior to the experiment, to explore the controls and practice the block-stacking task.

Following the 10 minute training session, each participant was allowed 45 minutes to complete as many trials of the block-stacking task as possible, up to a maximum of five trial attempts. Participants were informed that they could take breaks but would not be given extra time. After each trial, the participants knocked down any stacked blocks to randomise their position before starting the next trial.

At the end of their 45 minute experiment, participants were asked to complete the computerised version of the NASA Task Load Index (NASA-TLX) assessment. Participants were also asked to answer a system usability scale (SUS) questionnaire, separately for both the visual display and for the keyboard control system. Additionally, participants were asked to provide comments on the systems they had used. Overall, the experiment took approximately one hour for each participant.

4.3 Results

Out of 16 participants, two participants were considered as anomalies and removed from further analysis: participant 3 completed four successful trials but stated in the post-experiment questionnaire that he intentionally ignored the task objectives; participant 5 withdrew from the experiment after two

TABLE 4.1: Collected data from block stacking task, $\bar{x} = 289, \sigma = 76$. The following abbreviations are used in the table: Participant No. (PN), Trials Completed (TLC), Time to completion (TTC), Number of Collisions (NOC), Number of Dropped Objects (NOD) and Spatial Awareness Score (SAS)

PN	TLC	TTC [s]		NOC		NOD		SAS
		Avg.	σ	Avg.	σ	Avg.	σ	
1	5	233	58	2.0	1.7	0.8	0.7	68.0
2	5	190	33	0.2	0.4	0.0	0.0	70.4
4	5	286	24	0.4	0.5	0.0	0.0	58.4
6	5	445	241	4.6	2.2	3.0	1.9	52.0
7	5	281	51	1.8	1.7	0.6	0.5	65.6
8*	4	260	126	2.3	0.8	0.8	0.8	80.0
9	5	301	46	0.6	0.5	1.0	1.3	64.0
10	5	322	141	2.8	1.6	1.2	1.9	52.0
11	5	245	18	2.0	0.6	0.0	0.0	88.0
12*	3	418	139	2.3	1.7	0.7	0.5	57.0
13	5	164	28	0.6	0.8	0.2	0.4	65.0
14	5	278	85	0.6	0.5	0.0	0.0	87.0
15	5	309	15	1.2	1.0	0.8	0.7	53.0
16*	3	309	55	0.0	0.0	0.3	0.5	84.0

trials, likely because they had not understood the instructions due to poor English language ability. These two participant's data were eliminated from the analyses.

Table 4.1 shows performance metrics: TTC, NOC, NOD and Spatial Awareness Score (SAS) for each participant. Note that no critical errors (situations requiring the experimenter to intervene using e-stop button) occurred during any trials, so this zero statistic is omitted in Table 4.1. Three participants (marked by an asterisk in the table) successfully completed three trials within 45 minutes, but did not complete all five trials because they terminated some trials before completion due to frustration or driving the robot into awkward configurations from which they could not recover. On average participants took approximately 5 minutes to complete a trial.

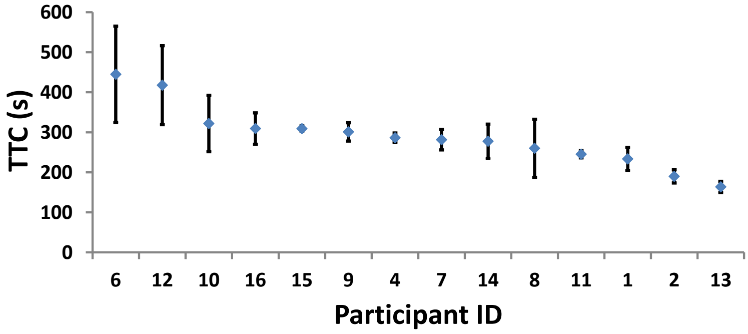
Fig. 4.3(a) plots average TTC of successful trials for each participant, sorted by size. Fig. 4.3(b) shows how TTC (averaged over all participants) improves with each successive trial. A repeated-measure analysis of variance

(ANOVA) was conducted in order to test the mean differences on the task performance (here, the TTC) between trials. In cases where the assumption of sphericity of the ANOVA was violated (i.e. the variances of the differences were not equal across trials), ANOVA with the Greenhouse-Geisser correction was employed. Results suggest significant differences over the course of the trials $F(1.927, 25.049) = 5.549,\ p < 0.05$. A result is considered to be significant if its p value is less than 0.05, i.e. there is less than 5% chance that observed results occurred merely by chance. F value is a ratio of the difference in means of each trial.

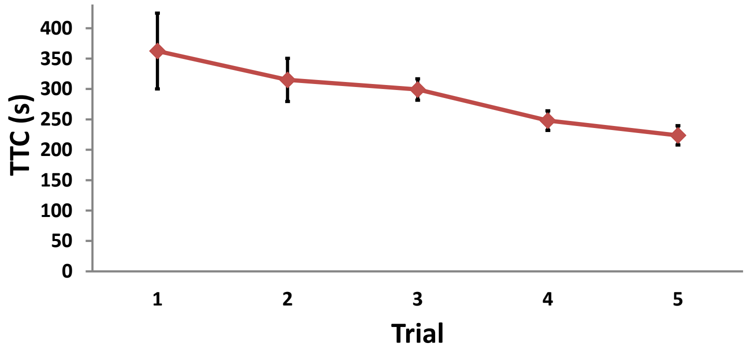
Fig. 4.3(b) shows TTC improving steadily from first to fifth trials, suggesting a learning effect during the repetitive trials. Additionally, standard deviation decreases with each trial, suggesting that the difference between high and low aptitude participants diminishes with training. In addition, a Pearson correlation coefficient was computed to assess the relationship between the two variables which were averaged across the completed trials. There were positive correlations between TTC & NOC ($r(14) = 0.62, p < 0.05$) and TTC & NOD ($r(14) = 0.69, p < 0.05$). Note that although there were incompleted trials for three participants, overall task performance consistently improved during successive trials.

Correlation between task performance and individual spatial ability scores was also evaluated, Table 4.1. A Pearson correlation coefficient was computed; there was a negative correlation between TTC and the SAS ($r(14) = -0.45,\ p = 0.10$). It also had a negative effect on the NOD ($r(14) = -0.56,\ p < 0.05$); however, no significance for NOC ($r(14) = -0.39,\ p = 0.165$).

There is high variation in six NASA-TLX sub-scales across different participants. Fig.4.4(a) shows the participants' subjective responses averaged across 14 participants for six sub-scales whereas Fig.4.4(b) presents the weighted value of the NASA-TLX. The collective weighting given to each factor via the pairwise comparison indicates relatively high mental demand in this



(a)



(b)

FIGURE 4.3: (a): Average TTC of each participant over all completed trials; (b): average participant time for each of 5 successive trials. The error bar indicates the standard error of mean (SEM).

block stacking task (Fig. 4.4(b)). Counterintuitively, participants with a higher spatial awareness score did not experience a significantly lower workload on the NASA-TLX scores, either raw or weighted scores. Based on additional post-experiment participant statements it can be concluded that, although the task involved a considerable mental effort, most participants tended to believe that they were "good" at teleoperating the robot.

The relationship between task performance and demographic factors was also examined. Fig. 4.5 shows the TTC trends compared between two groups: participants with minimal or no $3D$ games/applications experience ($n = 8$) and those with substantial experience ($n = 6$). Mixed ANOVA analysis was conducted to evaluate differences between these two groups across the five trials and the result suggests significant differences between groups:

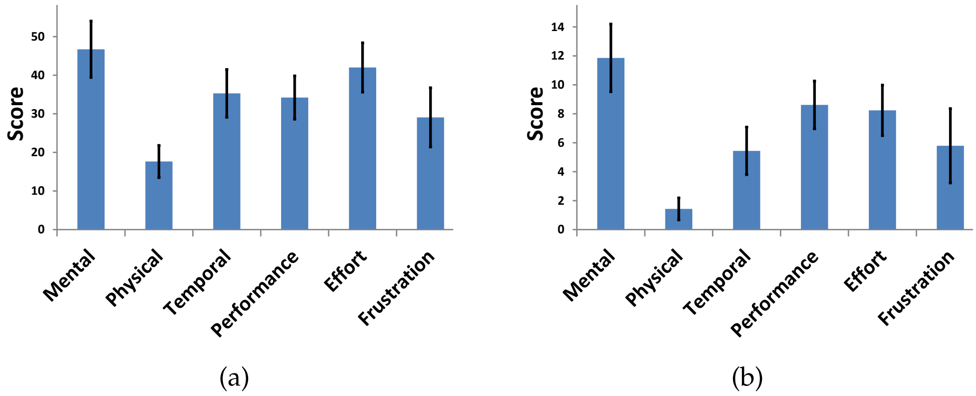


FIGURE 4.4: (a): raw scores, (b): weighted scores, the NASA-TLX for six subscales, averaged across the participants (n=14). The errorbar indicates the SEM.

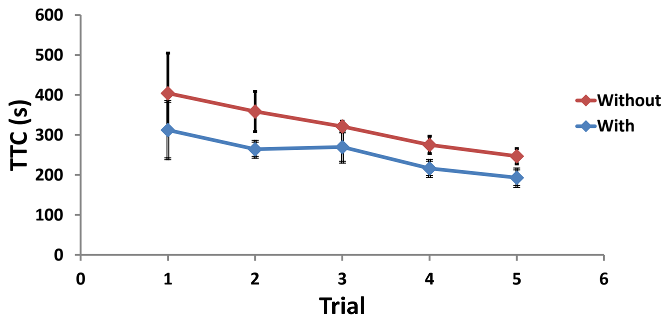
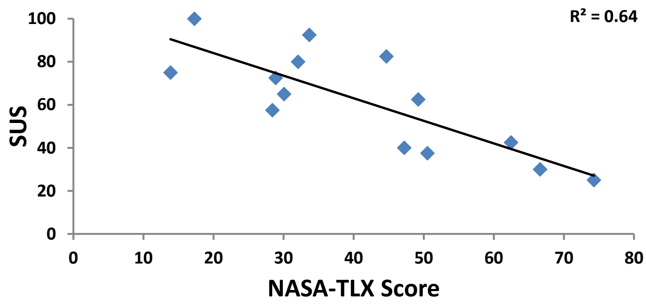


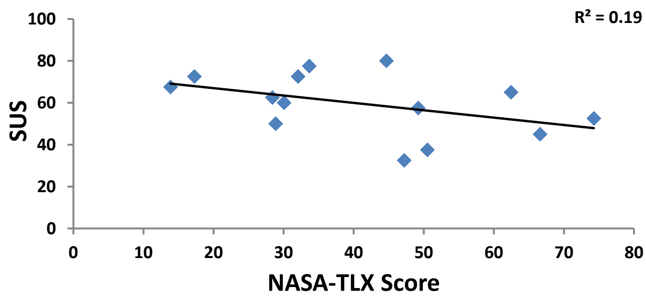
FIGURE 4.5: Previous experience with games or 3D applications (n=6) vs without (n=8). The errorbar indicates the SEM

$F(1, 12) = 5.286$, $p = 0.04$. Performance comparison between gender, however, did not show any significant differences. It should be noted that the sample size used for gender analysis is relatively small, so this is not conclusive.

System usability scores alone do not provide a clear picture as to what participants thought of each system element: (($n = 14$), camera feed display: ($\bar{x} = 61.6$, $\sigma = 23.7$); keyboard controller: ($\bar{x} = 59.5$, $\sigma = 14.6$)). Pearson Correlation analyses were conducted to explore whether any relationships exists between these characteristics. The results suggest that there were no significant correlations between the SAS and the SUS: ($r(14) = 0.16$, $p = 0.60$) for the camera feed display and ($r(14) = 0.19$, $p = 0.52$) for the keyboard. In Fig. 4.6, a comparison of workload (NASA-TLX weighted scores)



(a)



(b)

FIGURE 4.6: Comparison of workload experienced vs usability rating for each participant with camera feed display: (a); and keyboard control system (b).

with SUS scores reveals a significant negative correlation ($r(14) = -0.81,\ p < 0.05$) for the camera feed display system (Fig. 4.6(a)) and a weaker and non-significant negative correlation for the keyboard (Fig. 4.6(b)) ($r(14) = -0.44$, $p = 0.12$).

4.4 Discussion

This experiment was designed to investigate limitations of current baseline telerobotic systems used for remote manipulation in the nuclear industry, and to analyse relationships between overall system performance and human factors.

4.4.1 Task Performance

TTCs measured in this study evidenced that the use of baseline telemanipulation control systems is very slow and very difficult for novice operators, requiring high mental workload. Such systems would require extensive training and practice to achieve success in real nuclear deployments, and deployments would still be slow. This has serious implications for: i) training a new generation of workers to replace the current highly skilled workforce who are close to retirement (mean age 55); ii) handling the enormous quantities and types of waste with sufficient throughput speed. The high variance of the TTC suggests that the experimental population had significantly varying aptitude for remote manipulation skills, and this may have practical implications for selecting new operators for training.

The experiment highlighted the significance of adequate initial training to eliminate confounding factors of learning effects. The results suggest that 10 minutes of pre-experiment training was not sufficient, since a significant learning improvement is apparent throughout successive trials, Fig. 4.3. Although similar training periods have been shown to work well for experiments with 2-DOF robot vehicles [130], the spatial cognition problems of controlling a 7-DOF robot arm are more complex. For future experiments, structured training is recommended, customised to the task context, rather than free-form practice periods with minimal guidance. During training, subgoals can be used to assess task comprehension and control proficiency. Such subgoals should be demonstrated by participants in competency tests, post-training and pre-experiment, to ensure a common minimum competence between participants.

4.4.2 Workload & SUS

Measuring workload proved to be difficult, due to high variance in subjective responses from participants. However fig. 4.4 evidences mental demand contributed greatly to the workload, in particular for the weighted scores. Mental workload arises from having to resolve 2D camera images, from multiple views, to reconstruct 3D mental models of the scene, while deciding robot input commands with respect to a single eye-in-hand view. The lack of robot configuration feedback in the current baseline GUI also contributes to effort and frustration experienced by participants.

In the experiment the NASA-TLX scores seemed to be sensitive to individual interpretations of each subscale. For example, some participants gave a high rating to temporal demand even though there were no time limits other than the overall 45 minutes allowed (while each trial required only around 5 minutes to complete). Upon further investigation, responses from the participants indicated that many felt rushed due to self-imposing time constraints as they wanted to minimise the time taken as much as possible. Due to the current experimental settings, there were possible reasons that their mindset became competitive. Firstly, participants were told in advance their TTC data would be recorded, and this may have affected the way they performed during the task. Secondly, the controlled laboratory experiment does not replicate the seriousness of a high-consequence nuclear environment, so that confounding factors can be introduced if participants view the experiment as a competition. These observations emphasise the importance of taking into account how workload metrics can be affected by the operator's mindset when using telerobotic systems; and the importance of designing experiments that better reflect the end-usage scenario. Further research is required to create realistic settings to control the influence of the environment on participant responses.

Despite high variability in the SUS scores, significant correlation was found between SUS and NASA-TLX scores for the multi-camera view GUI, but little correlation for usability of the keyboard push-button control system. The usability ratings of the GUI (scene visualisation) likely reflect the mental demand and effort, assessed by the NASA-TLX, where participants needed to reconstruct 3D mental models of the remote workspace from 2D images. In contrast, no reasonable explanation was found for the lack of usability correlation for the keyboard controls. Future experiments comparing different types of controller (keyboard, joystick, haptic devices etc) in conducting the same block stacking are likely to provide functional insights into the relationship between the controller and the workload. Addition of other sensory cues (e.g. sound transferred from a microphone on the robot) might also cause changes to be observed in SUS and workload scores.

4.4.3 Demographics and Spatial Awareness

Participants had a range of previous experience with 3D applications and games, assessed by the pre-screening questionnaire. Fig. 4.5, suggests that such experience is advantageous for teleoperation, as these participants consistently performed better, likely due to transfer of skills developed in such experiences.

Participants with better spatial abilities demonstrated superior task performance, suggesting that ability to mentally manipulate 2D & 3D images is useful for teleoperation, in agreement with other recent work [131]. The findings suggest how training and recruitment processes could be improved by pre-screening of participants according to such traits. Applicants identified through pre-screening with significant previous 3D games or applications experience could be allocated to a fast track training route, reducing cost and improving efficiency. Spatial ability tests could also be used to identify candidates for teleoperation training.

4.4.4 Limitations

This experiment was an early foray into the study of human factors for telerobotic decommissioning, and as such there are many limitations and questions that remain. The following text reports the general observations of task performance during this experiment, which suggest important lessons for future research directions and methodology.

Firstly, controlling end-effector orientation proved mentally overwhelming for participants, with the majority choosing to ignore it. This was likely due to cognitive overloading as participants could be observed using trial and error to check how the robot responded to orientation control. Incorporating autonomous methods for controlling end-effector orientation could significantly assist human operators. Secondly, participants consistently misjudged distances, with frequent unsuccessful grasping attempts highlighting the importance of depth cues. Improving depth perception is therefore a crucial research issue for future telerobotic systems. Thirdly, some participants had preferences for different keyboard mappings, and involuntarily pressed keys from their preferential mapping on multiple occasions. These simple yet informative observations need to be considered in the future design of control interfaces. In addition, note that in this initial experiment, while participants could only view the remote workspace via camera views, the (curtained off) robot was nearby, enabling participants to hear some audio cues (motor noises and contacts).

Extensive human factors improvements are needed for telerobotic systems to enable nuclear decommissioning. Participant comments, obtained by the SUS questionnaire, recommend improving situational awareness, for example employing a 'visual indicator' to show the movement direction when controlling the robot arm. Participants also found the baseline push-button controls unintuitive, and difficult to relate to the 3D motions of the robot. Participants suggested that automated assistance would be beneficial, in particular a function to automatically reset the robot from awkward configurations. Other recommendations are to provide robot configuration indicators to the operator.

No common standard yet exists for selecting metrics for teleoperation [132]. Additional metrics might measure: participant "thinking time" and "idle time" spent without input; frequency of key presses, to highlight possible frustrating aspects; and distance covered by robot motions, as a measure of efficiency. It is important that appropriate metrics fully encompassing the effects of both robot and operator are identified and measured. Future experiments should expand the number of metrics being measured to cover a greater breath of human factor issues.

4.5 Conclusions

This experiment has presented a rigorous human-factors evaluation of humans using industry-baseline technology (push-button control of a robot arm via multiple camera views) to perform remote manipulation tasks. The results suggest severe limitations of such basic direct tele-operation interfaces, especially considering the urgency of rapidly training a new generation of nuclear workers to replace the aging (mean age 55) workforce of very highly skilled operators.

Static camera views with poor depth perception inhibit performance and make trajectory planning mentally challenging for human operators. Non-autonomous teleoperation is extremely demanding especially when the robot is driven into complex poses, causing frustration and increasing operator fatigue. Rudimentary teleoperation interfaces will not enable the complex manipulations needed to decommission legacy plant infrastructure such as gloveboxes, and cannot provide the increased throughput rates that are needed to decommission very large quantities of legacy nuclear waste within socially acceptable periods of time.

The results also suggest that, while human subjects can be a vital source of feedback in human factors development, subjective assessments by human subjects require careful administration. Standard subjective assessment methods may sometimes be inappropriate and need to be adjusted depending on the context and participants.

Chapter 5

Direct vs. Semi-autonomous Teleoperation

This chapter is a reformatted copy of the following publication: *Marturi, N., Rastegarpanah, A., Takahashi, C., Adjigble, M., Stolkin, R., Zurek, S., Kopicki, M., Talha, M., Kuo, J.A. and Bekiroglu, Y. (2016, December). Towards advanced robotic manipulation for nuclear decommissioning: a pilot study on tele-operation and autonomy. In Robotics and Automation for Humanitarian Applications (RAHA), 2016 International Conference on (pp. 1-8). IEEE.* [133]

This chapter presents work from a follow-up experiment using an industrial robot arm to perform telemanipulation. This was made possible through collaborative effort from members of the Extreme Robotics Lab in the Metallurgy and Materials department of the University of Birmingham including expert industrial robot arm programmers. The author's contribution to this collaborative effort consisted of the principled experiment design to evaluate human factors issues while colleagues developed the semi-autonomous capabilities.

The purpose of the experiment was to demonstrate a proof-of-concept in using semi-autonomy for nuclear decommissioning applications. Manual teleoperation was compared against semi-autonomous teleoperation with supervisory control to evaluate the effect on task performance during basic grasping and manipulation tasks. The proposed benefits of using semi-autonomy in such tasks include efficient task performance, lower cognitive demand on the operator and reliability of the overall system.

5.1 Related Work

Computer vision has been used heavily in industrial processes based on object characteristics (e.g. size, shape, colour, pose etc.) to automate repetitive tasks using automated machinery or robots in structured environments [134], [135]. However, visual-servoing techniques have failed to penetrate as deeply in the nuclear industry partly due to the industry's highly conservative nature and partly due to the unstructured environment. Examples of applications where computer vision methods have been used in nuclear facilities are [136]–[139]. [136] present a vision-based solution for identification and classification of nuclear waste items. [137] proposed a vision-based method for controlling the trajectory of inspection robots for nuclear reactors and techniques for estimating radiation levels and nondestructive testing of piping systems can be found in [138] and [139], respectively.

These techniques however do not cover the area of nuclear waste sort and segregation where research is still limited to the laboratory environment [140]. In an effort to resolve this, a semi-autonomous framework was developed for assisting the remote operator with autonomously grasping and stacking objects. A model-based object tracker [141], [142] is used to track and estimate object poses, which are passed to a robot trajectory planner to execute autonomous grasping. The object is then placed in the desired location under human-supervision.

5.2 Semi-autonomy Features

The semi-autonomy features used in this experiment consist of the model-based tracker and visual servoing for grasping.

5.2.1 Model-based Tracker[1]

A model-based object tracker is used to track each object. Objects are represented using 3-D CAD models whose poses are estimated and tracked in real-time through monocular cameras. The objective of the tracker is to obtain a camera pose for which the projected CAD model fits best with the 2D image contours of the object in the scene. This process involves estimating a rigid transformation between the virtual camera frame and the tracked object frame. The literature shows that previous uses of such methods for tracking objects in complex environments under intensity and occlusion constraints have been successful [143], [144].

5.2.2 Visual Servoing for Grasping[1]

A pose-based visual controller is used to take the pose features obtained from tracking to navigate the robot to the grasping locations of each object. This is achieved by minimising the error between the current and desired poses using homogeneous transformations. After grasping, the robot follows a predefined path to place the object at the desired location.

5.3 Experiment

The experiment was designed to study various factors affecting performance of manual vs. vision-guided semi-autonomous teleoperation. This section details the developed experimental setup along with the tasks used for the

[1]Feature implemented by colleagues. See chapter introduction.

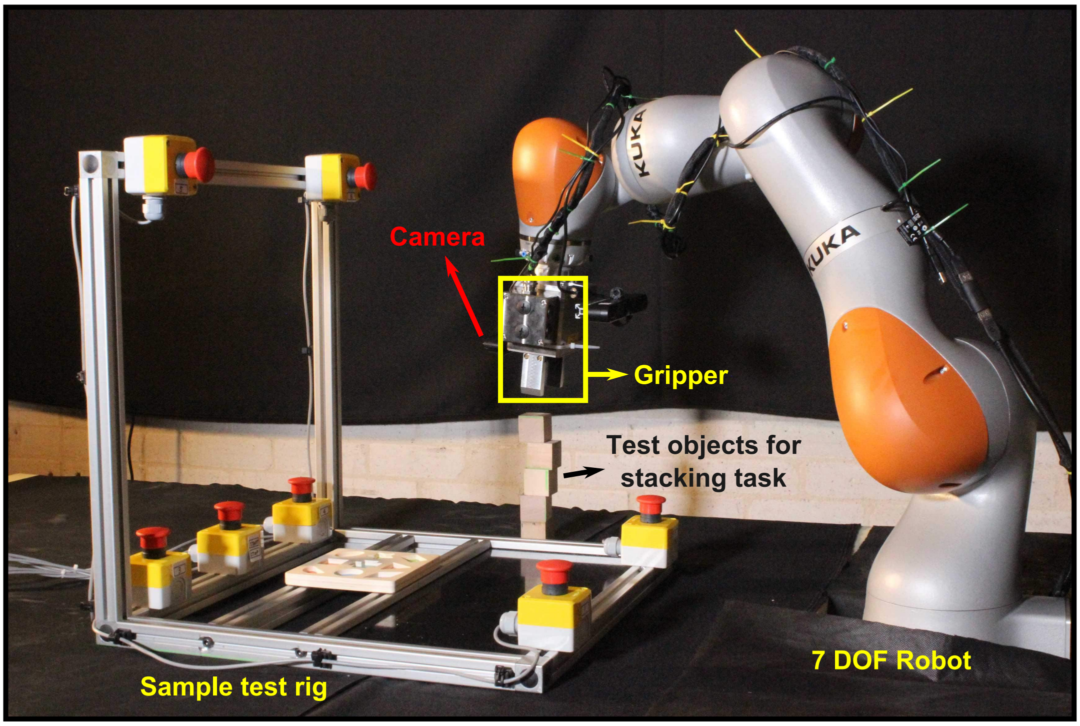


FIGURE 5.1: Experimental set-up showing various components used in this work. The test rig with seven buttons, used for the point-to-point dexterity task, can be seen on the left.

experiment. Each task was performed as a separate study using different participants on different days.

5.3.1 System

The experimental set-up used to perform both the tele-operated as well as the semi-autonomy trials are shown in fig. 5.1. It consisted of a 7-DOF lightweight robot arm: KUKA LBR iiwa 14 R820, a Schunk PG plus-70 2-finger parallel jaw gripper with a stroke size of 68 mm and a commercial Logitech c920 USB camera. The gripper was attached to the tool centre point of the robot and the camera was mounted on top of the gripper whose optical axis coincides with the z-axis of the gripper frame. The vision-guided control architecture for semi-automatic tests was implemented in C++ and is executed from a remote computer running Windows (8 GB RAM, 2.3 GHz Intel core i7 CPU). The communication between the PC and the robot controller has been made using UDP and the gripper control commands are transmitted over a serial port.

5.3.2 Tasks

Point-to-point Dexterity Task

The initial task of point-to-point robot navigation was designed to evaluate the performance of human subjects who were required to perform the task by using the camera-provided views of the remote workspace. The task required the participants to move the robot's end effector from button-to-button on the test-rig shown in fig. 5.1 in a specific order. To successfully complete this task, participants needed to change both the position and orientation of the end effector to be able to reach the buttons.

Block Stacking

The second task involved stacking wooden blocks as in the previous experiment in section 4.2. Criteria for success was the finished block tower had to be able to stand freely.

5.3.3 Method

Point-to-point Dexterity Task

6 participants from the students and staff at the University of Birmingham voluntarily participated in this experiment. All had normal or corrected to normal vision, and had no known neurological motor deficits (self-reported). Each participant was given a 10 minutes training session followed by the live trials. In the training session, they were provided with an instruction set of the task and asked to familiarize with the robot control. After training, they

were asked to complete the live task three times without any time limit. They were allowed to have a break between trials as required.

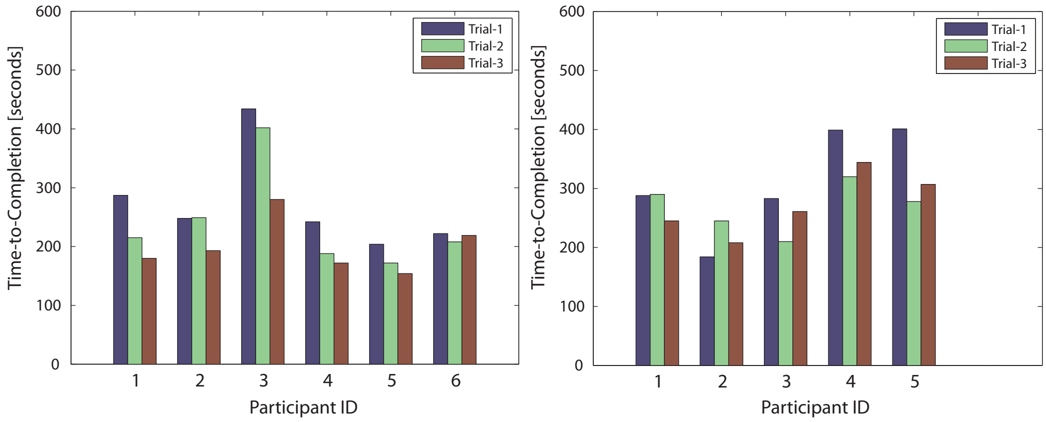
At the end of the session, participants were asked to answer the computerized version of the NASA Task Load Index (NASA-TLX) assessment as well as system usability scale (SUS) questionnaires. In this study, the computer version of the NASA-TLX was used and only the raw scores were analysed [145], [146]. In addition, the task performance was measured based on several metrics: time taken to complete the task, number of collisions with the environment, the number of missed buttons and the number of unacceptable critical incidents such as failure due to incorrect input causing the robot to be halted with the emergency button.

Block Stacking Task

5 new participants were recruited for this task. Each participant had 10 minutes training to get themselves familiar with the robot controls. Similar to the previous experiment, participants tele-operated the robot using keyboard functions. The blocks were situated such that they were within reach of the robot and also sufficiently spaced apart from each other such that the gripper fingers do not collide with neighbouring blocks.

After training, each participant was asked to perform the task three times and during each trial, similarly to the previous task, the time taken to complete the task, the number of collisions with the environment, the number of dropped objects, and the number of unacceptable critical incidents were recorded.

For comparison, the task was repeated using semi-autonomous teleoperation with human supervision. Initially, the robot was moved to a predefined home position, which is fixed throughout the experiment. Trackers are initialised through human input in the form of initial poses for each object. The telerobot then performs the rest of the task autonomously to stack the blocks.



(a) point-to-point, time (b) stacking, time

FIGURE 5.2: Time spent to complete a) the point-to-point dexterity task and b) the block stacking task for all participants. For each participant task completion time in each trial is shown.

5.4 Results

5.4.1 Results for the Point-to-point Dexterity Task

The performance of all 6 participants during the point-to-point dexterity task was evaluated based on the criteria listed in section 5.3.3. Although some participants failed to press the buttons a number of times during their training session, no such instances were observed during the main trials. As shown in Fig. 5.2(a), the averaged time to completion (TTC) across 6 participants was 273 seconds in the first trial reduced to 200 seconds in the third trial. The repeated measure ANOVA (Analysis of Variance) was used to evaluate whether the performance significantly changed or not across three trials. Mauchly's test indicated that the assumption of sphericity for ANOVA is satisfied: $\chi^2(2) = 2.17$, $p = 0.34$. The ANOVA test shows that there was a significant learning effect across the three trials: $F(2, 10) = 8.74$, $p =< 0.01$. These results suggest that the task performance constantly improved throughout the trials.

As fig. 5.3 shows, the mental demands are significantly high compared

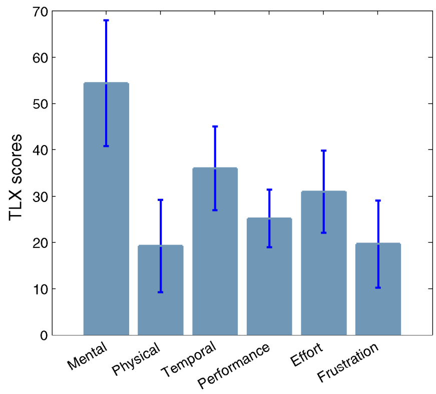


FIGURE 5.3: NASA Task Load Index scores for the point-to-point dexterity task, averaged across 6 participants.

TABLE 5.1: Statistical analysis of teleoperated block stacking including task completion time (Π), no. of collisions (α), no. of dropped objects (β), no. of critical incidents (γ)

Participant ID	Π		α		β		γ	
	Av.	Std.	Av.	Std.	Av.	Std.	Av.	Std.
1	274.3	25.4	3.6	3.2	0.6	0.5	0.6	0.5
2	212.3	30.7	0.3	0.5	0.3	0.5	0.0	0.0
3	251.3	37.4	2.6	3.0	0.6	0.5	0.3	0.5
4	354.3	40.5	2.0	3.4	1.6	0.5	0.0	0.0
5	328.6	64.2	1.6	2.8	0.3	0.5	0.0	0.0

Av. = Average over 3 trials, Std. = Standard deviation.

with other sub-scales due to participants having to construct the 3D perception of the remote workspace through the 2D images of the live camera-feeds while at the same time teleoperating the robot arm resulting in high cognitive load. On the contrary, the physical demands and frustration were relatively low compared with other sub-scales, suggesting that teleoperation could help to reduce fatigue for such repetitive tasks.

5.4.2 Results for the Block Stacking Task

Fig. 5.2(b) shows the TTC of all 5 participants for the block stacking task. The average TTC across 5 participants was 311 seconds in the first trial and then slightly reduced to 273 seconds in the third trial. Table 5.1 shows the descriptive statistics averaged over all trials. Similarly to the previous task, the repeated measure ANOVA was used on the TTC data. Mauchly's test indicated that the sphericity assumption was satisfied: $\chi^2(2) = 2.17$, $p = 0.26$. In contrast to the point-to-point dexterity task, the results show that there was no significant effect across the three trials: $F(2, 8) = 1.80$, $p = 0.23$. This could be due to the fact that during the block stacking task, there was more room for error through dropped blocks, misaligned blocks leading to tower collapse etc. meaning that further practice was needed for learning to occur. This is supported by the fact that overall, the number of critical incidents (collisions and drops) were relatively high compared with the point-to-point dexterity task.

Participants encountered difficulties in constructing aligned and stable block towers. This seemed to have a large negative impact on the task performance, which possibly was one of the main reasons that the learning effect disappeared across three trials. Overall, this task was more difficult to achieve compared to the previous one and required more trials for participants to excel.

5.4.3 Results for Semi-Autonomous Block Stacking

The task was executed 11 times using the semi-autonomous system and the corresponding time for each trial is shown in fig. 5.4. On average, the semi-autonomous system took 88 seconds (in robot T1 mode) to complete one trial, over three times faster than the teleoperation experiments. Due to the constant lighting and clutter-less environment, no task failures or collisions were observed during all 11 trials.

5.5 Discussion

Similarly to the previous experiment (section 4.2), participants found it difficult to directly teleoperate the robot due to the high cognitive demand of

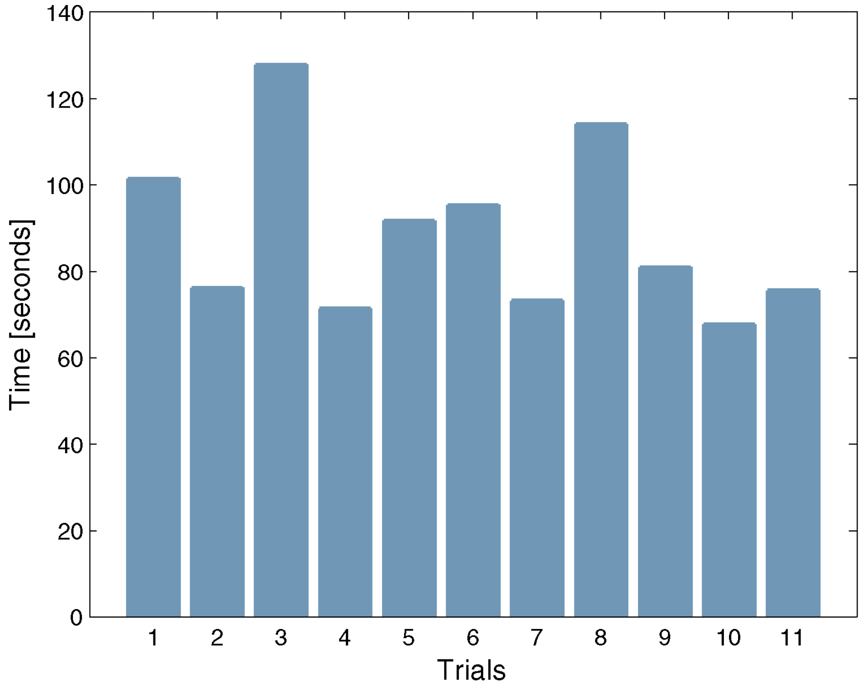


FIGURE 5.4: Overall time taken for semi-automated block stacking during 11 runs. On average the stacking of 5 blocks took 88 seconds, more than three times faster than teleoperation mode.

the teleoperated system. The use of push-button controls made it awkward to handle the teleoperator as it was not clear how the controls mapped to the teleoperator movement. This combination of controller and lack of depth perception led to errors and irrecoverable configurations frustrating participants. This highlights again how primitive teleoperated systems currently used for nuclear deployment are unfeasible for handling the large quantities of waste within the time-frame set by government.

The developed semi-autonomous system has been demonstrated to be a more efficient and reliable method for teleoperation. As well as reducing the TTC by a factor of 3, errors due to incorrect input from the operator were completely removed reducing the chances of hazards occuring in a real scenario such as a safety-oriented nuclear facility. It also reduced the workload of the human operator freeing time that can be allocated to other tasks such as controlling other robots. This proof-of-concept has shown the feasibility of using semi-autonomy for teleoperation to make sort and segregation tasks in nuclear facilities "safer, faster and cheaper".

5.6 Conclusions

An initial proof-of-concept has been presented and demonstrated for semi-autonomous manipulation using visual servoing. Performance of human test-subjects using both systems was evaluated empirically and with regard to human factors issues. The results show how performance improves with training, and suggest how training requirements scale with task complexity. The results also demonstrate how the incorporation of autonomous robot control methods can reduce workload for human operators, while improving task completion time, repeatibility and precision. However, the appropriate level of autonomy for increasingly complex teleoperation tasks in the safety-oriented nuclear industry remains to be investigated.

Chapter 6

Orbital Camera Control

In remote teleoperation, such as in hazardous nuclear facilities, the loss of a direct view of the environment reduces situational awareness for the operator, negatively affecting overall task performance. Even with modern technological advancements, it is still not possible to replicate human eyes using cameras. 2D cameras make it difficult to ascertain distances between objects due to lack of depth information and suffer from a reduced field of view (FOV). Using multiple cameras to expand coverage exerts mental workload on the operator. It makes it challenging to control a robot due to confusion with frames of reference for control input. Even simple transformations can incur high cognitive load due to misalignment between camera display and robot movement [147]. In typical scenarios during nuclear teleoperation, camera angles do not conform to the human eyes i.e. rather than a line of sight view as seen through the human eyes; the operator is presented with an arbitrarily transformed view. The majority of human perception takes place through vision so therefore to reduce the workload on the operator and improve performance, it is important to remove or minimise the effects of such misalignment. As a step towards achieving this goal, this chapter presents an “orbital camera“ concept for a shared-control approach to camera views. The camera remains fixated on the end-effector of the robot arm and can be dynamically controlled by the operator to obtain a desired camera view.

6.1 Related Work

Common techniques in the literature tackling these problems can be split into three branches: display configuration, viewpoints & reference frames and camera control. Display configuration deals with the methods used to display the remote environment to the operator. Viewpoint is concerned with the best positioning for cameras to maximise operator situational awareness whilst camera control explores efficient interfaces for positioning the cameras in the remote environment during teleoperation.

6.1.1 Display configuration

Early systems were homogenous in the use of monitors (single/ multiple) for visual feedback [148]–[151]. Video streams from static cameras are directly displayed on monitors with multiple streams either split on a single screen or across multiple screens. The operator controls the robot based on a world frame of reference and has to mentally map control input to robot motion for each camera view.

Research has shown that this method induces a higher workload and can lead to the keyhole effect where the user focuses on a single stream and neglects the other views for a prolonged period of time [93], [152], [153]. As a result, the reduced situational awareness can have repercussions on safety in a nuclear environment. Conversely, using a single camera is insufficient due to obstructions and coverage issues particularly during telemanipulation tasks where bulky robot arms are used [154]. A study on teleoperated interfaces in a competition environment demonstrated that operators use video as the main source of feedback and can frequently neglect feedback from other sensors [93].

A middle ground approach was presented by Nielsen et al. [155] for a mobile teleoperated robot that fused data from multiple streams. An augmented reality interface integrated a map of the environment, robot pose, camera pose and video feedback to provide a 3D view on a single display. The results from their user studies showed that this approach improved performance, reduced cognitive load and raised situational awareness of operators. This method works well for mobile navigation and search scenarios where a single view is usually sufficient to carry out all tasks but telemanipulation entails complex tasks that cannot be performed using a single view especially without stereoscopy.

Head mounted displays (HMDs) have seen a recent rise in use for teleoperation visual feedback. Combined with virtual reality, they are particularly suited to viewing environments created by depth cameras using point clouds [156]. Theofilis et al. [157] used HMDs to increase user immersion by presenting stereo panoramic views. Other cases of using HMDs to provide stereoscopic feedback have been implemented in [156]. Numerous others have used HMDs in mobile teleoperation with either a single static camera or pan tilt unit (PTU) to provide visual feedback [158]–[160]. However there is comparatively little research on the use of HMDs to teleoperate robot arms. One such use is by Adamides et al. [161] who compared HMDs to standard display monitors in a telemanipulation task using a mobile robot and found that HMDs increased the workload perceived by the users.

6.1.2 Viewpoints & Reference Frames

One of the first studies conducted on visual feedback for remote operators was by NASA [148]. They compared 3 different camera views for a remote manipulation task: line of sight, top-down & eye-in-hand. The line of sight view consists of a camera placed along the sight axis of the operator to replicate a natural view. The findings revealed that the optimum viewing angle is highly dependent on the task; no single camera position serves well for general work. The recommendation from this study was to use a line of sight camera view augmented by other views, preferably an eye in hand camera view. Smith and Stuart [162] studied the effects of spatially displaced camera views on operator performance and found that the normal camera view offered the best performance when direct view of the remote environment was not available.

McCormick et al. [150] conducted a study on 3 frames of reference (egocentric- views from the robot's perspective, exocentric- views from an external viewpoint & tethered- camera attached to the robot using a restraint) in a remote navigation task and found that egocentric views improved navigation performance near targets whilst exocentric views improved search and global judgement subtasks.

A number of studies have shown that minimising the number of cameras to the most effective views (usually two) and displaying them on a single monitor leads to the best operator performance as operators do not have to mentally combine images from multiple orthogonal cameras (e.g. cameras in the x,y,z axes) [93], [153], [163]. Egocentric views are best for manipulation tasks whilst exocentric and tethered views allow for greater situational awareness [150], [164]. According to Rahnamaei Sirouspour [165], a good viewpoint covers critical task-specific points. For navigation, these consist of robot body, destination and obstructions and additionally the end effector for manipulation.

Teleoperated mobile navigation works well with exocentric views as this allows for a global scene understanding, improves search capabilities and reduces workload for the operator [166], [167]. A wide field of view from a line of sight perspective, which includes the robot body and its general surroundings, significantly improves the situational awareness of the operator [153], [168].

A line of sight view is not sufficient during telemanipulation as it can become obstructed due to the arm configuration or may not provide adequate coverage [154]. For teleoperated arms, the best views have shown to be the global and end effector view [161], [163]. The global view is best during large scale movement, scene understanding and search tasks whilst the end effector view allows for precise movements required during manipulation and interaction. Adamides et al. [161] evaluated a number of camera views for a teleoperated agricultural robotic sprayer which showed that users experienced an increased sense of presence with a combined exocentric and end effector view, compared with an egocentric view alone.

6.1.3 Camera control

To improve operator perception, the camera can be controlled separately from the robot. This technique has shown to improve operator performance in a navigation task during mobile teleoperation [169], [170]. However, controlling a camera and a robot at the same time can be challenging due to the secondary mental workload [171]. In some cases, a second operator is required just for controlling the camera/s [93], [172].

An early study [149] recommended the control input to be camera-oriented where the reference frame is located at the robot end effector. Brooks, McKee and Schenker [172] were one of the first to propose the use of automated camera control. They introduced the Visual Acts theory which uses the geometric, kinematic models and geometric description of the task to reason about the task being performed and select the optimal camera viewpoint in a particular context.

Recent years have seen the introduction of novel interfaces for controlling cameras as well as algorithms to automate camera control. Conventional joystick, head gaze tracking and autonomous tracking were compared in [173] for a camera mounted on a PTU. They found significant performance advantages and user preference for using the head gaze tracking. Saakes et al. [174] used a drone to provide an autonomous tethered aerial view of a teleoperated mobile robot during a navigation task. However, this view was unstable due to the vibrations from the movement of the drone causing users to become disoriented and there was no control provided for users to override the current view. Automated control was compared against manual and shown to improve performance for a gantry mounted camera system [165]. A role assignment method was developed in [175] to autonomously control camera feeds displayed to the user. This approach holds promise but needs further improvements as frequent switching during task execution led to cognitive overload for users. Other approaches include multi vantage systems with the use of a second robot vehicle mounted with a camera to autonomously follow the primary robot [167] and a robot arm mounted with a camera on the end effector controlled manually using forward kinematics [176]. Head and gaze tracking were used in [152] and the gesture based LeapMotion device in [160].

Most previous research related to visual feedback is concentrated around teleoperated mobile robots. However, telemanipulation using anthromorphic robot arms is more complex and presents unique challenges.

HMDs, as mentioned above, have not yet been proven to be suitable for prolonged use in telerobotics. Although they offer improvements in 3D perception capabilities for operators [177], they still suffer from a number of issues such as simulator sickness and high computational demand and require evaluation during prolonged periods of usage by an individual [157], [159]. Additionally, using the orientation of a HMD to control the camera view in a telemanipulation task is not appropriate as the range of motion of an operator's head is likely to not cover the requirements of such a task.

The use of novel interfaces such as gaze and head tracking or devices like the LeapMotion to control cameras are not ergonomically viable. Gaze and eye tracking require users to hold the gaze/ head in awkward positions relative to the display [173]. Prolonged use of such devices in unnatural positions can lead to eye/ neck strain and fatigue. Additionally, such devices remain relatively unreliable with high variation in tracking performance [171]. Automated control of cameras does not always offer the best view and can create instability. Such methods can lead to conflicts between operator and robot diminishing trust and therefore reducing system efficiency.

6.1.4 Novelty of Contributions

the approach to telemanipulation visual feedback involves the use of a camera mounted on a drone or a secondary robot arm with two degrees of freedom control. This is referred to as an "orbital camera". Inverse kinematic control of the drone or arm in Cartesian space in 2 dof would reduce the mental workload of the operator unlike other approaches that require monitoring multiple camera feeds [161] or have complex high dof manual controls [165], [176]. Such a camera offers a high degree of flexibility to the operator in acquiring optimal views during teleoperation tasks such as navigation and manipulation.

This will be coupled with an eye-in-hand (EIH) camera mounted on the end effector to provide an egocentric view for telemanipulation tasks. It was hypothesised that only these two camera views are needed for optimal visual feedback for an operator during telemanipulation with robot arms. Unlike many in the literature, an evaluation of the system is also presented in a user study consisting of individuals from the University of Birmingham. Based on a thorough search of the literature, such a system has not been proposed before in the context of nuclear teleoperation nor experimentally evaluated against industry standard systems.

6.2 Visual Feedback Methods

Two visual feedback methods were compared in this experiment through simulation and are outlined below.

6.2.1 Static Cameras

In most teleoperation environments in the nuclear industry, the standard method of visual feedback is the use of multiple static cameras placed in the remote environment. These are positioned to provide maximum coverage to the operator. The camera stream is then displayed to the operator across multiple monitors in no particular order at the control station.

A similar method was chosen to serve as the basis for comparison in this experiment. The implementation consists of 4 static orthogonally placed camera views: top-down, line of sight, left and right. Additionally, there is an eye-in-hand (EIH) camera mounted on the end effector which provides an egocentric dynamic view as the robot arm moves. The line of sight and EIH views are scaled to a resolution of 800x600 whilst the rest of the cameras are set to 400x300. It should be noted that the camera positioning in this setup is idealised as the perfect orthogonal positioning of cameras in such a manner is rare in real-world scenarios. Lack of access and obstructions mean that operators often have to deal with suboptimal camera positions not aligned with coordinate axes.

6.2.2 Orbital Camera

The concept of the orbital camera consists of an operator-controlled camera mounted on a 6 dof platform. This could either be a camera on the end-effector of another robot arm manipulator or attached to a UAV. The camera remains fixated on the end effector of the teleoperator, in this case a robot arm. This fixation point is used as the pivot point for operator-controlled

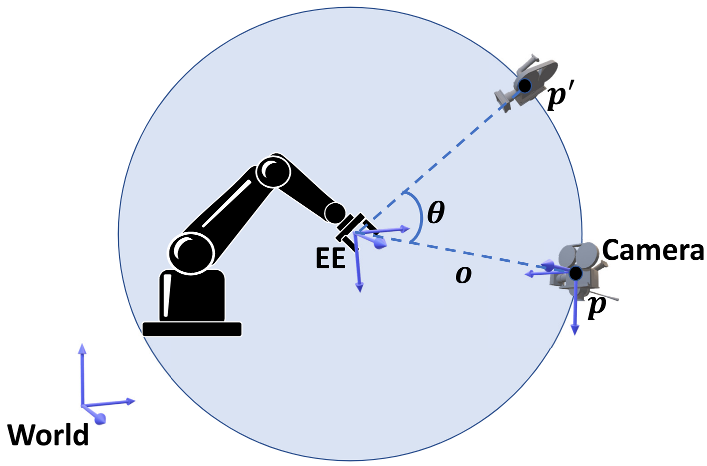


FIGURE 6.1: Representation of a camera at multiple points in its spherical orbit.

camera movement in the azimuth and elevation directions. This allows for a reduced 2-DOF control as the roll axis of the camera is redundant for tele-operation. However, the orbital camera concept allows for zoom towards or away from the centre of the sphere so in total the camera has 3-DOF. This is inspired by video games that utilise the 3rd person perspective to give the view of the player-controlled character within the surrounding environment improving awareness.

Implementation

Consider a camera located at a point p from the origin of the *end-effector (EE)* coordinate frame pointing towards the EE, fig. 6.1. The workspace of the camera forms a hollow sphere with the EE as the centre and current camera position as the radius. Assuming the camera location relative to the robot is known through calibration, the camera can be made to orbit its centre based on operator input using vector rotation.

The position of the camera relative to the EE is given by the vector $\mathbf{p}$. The goal is to rotate p by θ around the axis perpendicular to both p and $p\prime$ to get the vector $p\prime$.

Quaternions can be used to efficiently represent this rotation. A quaternion is defined as $q = q_0 + q_1 i + q_2 j + q_3 k$ where $q0, q1, q2, q3$ are scalars and i, j, k are imaginary components. The angle-axis ($\theta\mathbf{v}$) representation of the required rotation can be converted to quaternion as follows:

$$
\begin{aligned}
q_0 &= \cos\frac{\theta}{2} \\
q_1 &= \mathbf{v_x}\sin\frac{\theta}{2} \\
q_2 &= \mathbf{v_y}\sin\frac{\theta}{2} \\
q_3 &= \mathbf{v_z}\sin\frac{\theta}{2}
\end{aligned}
$$

The resulting quaternion q_{rot} holds the information on the amount of rotation around the chosen axis to be applied to the camera. This can be multiplied by the null quaternion q_{null} to obtain the transformation from point p to $p\prime$:

$$\mathbf{q_{tf}} = \mathbf{q_{rot}}\mathbf{q_{null}} \tag{6.1}$$

The transformation now needs to be applied to the camera. This can be achieved by using the quaternion $\mathbf{q_{tf}}$ to generate a computationally efficient homogeneous transformation matrix:

$$\mathbf{T} = \begin{bmatrix} 1 - 2q_2^2 - 2q_3^2 & 2q_1q_2 - 2q_0q_3 & 2q_1q_3 + 2q_0q_2 & 0 \\ 2q_1q_2 + 2q_0q_3 & 1 - 2q_1^2 - 2q_3^2 & 2q_2q_3 + 2q_0q_1 & 0 \\ 2q_1q_3 - 2q_0q_2 & 2q_2q_3 - 2q_0q_1 & 1 - 2q_1^2 - 2q_2^2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{6.2}$$

The transformation matrix $\mathbf{T}$ can then be multiplied by the original camera vector $\mathbf{p}$ to give the new position vector of the camera $\mathbf{p\prime}$:

$$\mathbf{p\prime} = \mathbf{Tp} \tag{6.3}$$

Camera zoom can be achieved by scaling the unit vector $\hat{\mathbf{p}}$ by an operator-controlled zoom factor d:

$$\mathbf{p} = d\hat{\mathbf{p}} \tag{6.4}$$

The robot EE can be controlled by the operator in the camera reference frame using coordinate transformations. Assuming the transformation from the EE to the world frame $\mathbf{T_{EE}^{world}}$ (using inverse kinematics (IK)) and the transformation from the camera to the world frame $\mathbf{T_{world}^{cam}}$ are known, the pose (a 4x4 homogeneous matrix) of the EE $\mathbf{x_{EE}^{cam}}$ with respect to the camera is given by the following transformation:

$$\mathbf{x_{EE}^{cam}} = \mathbf{T^{-1}{}_{world}^{cam}}\,\mathbf{x_{EE}^{world}} \tag{6.5}$$

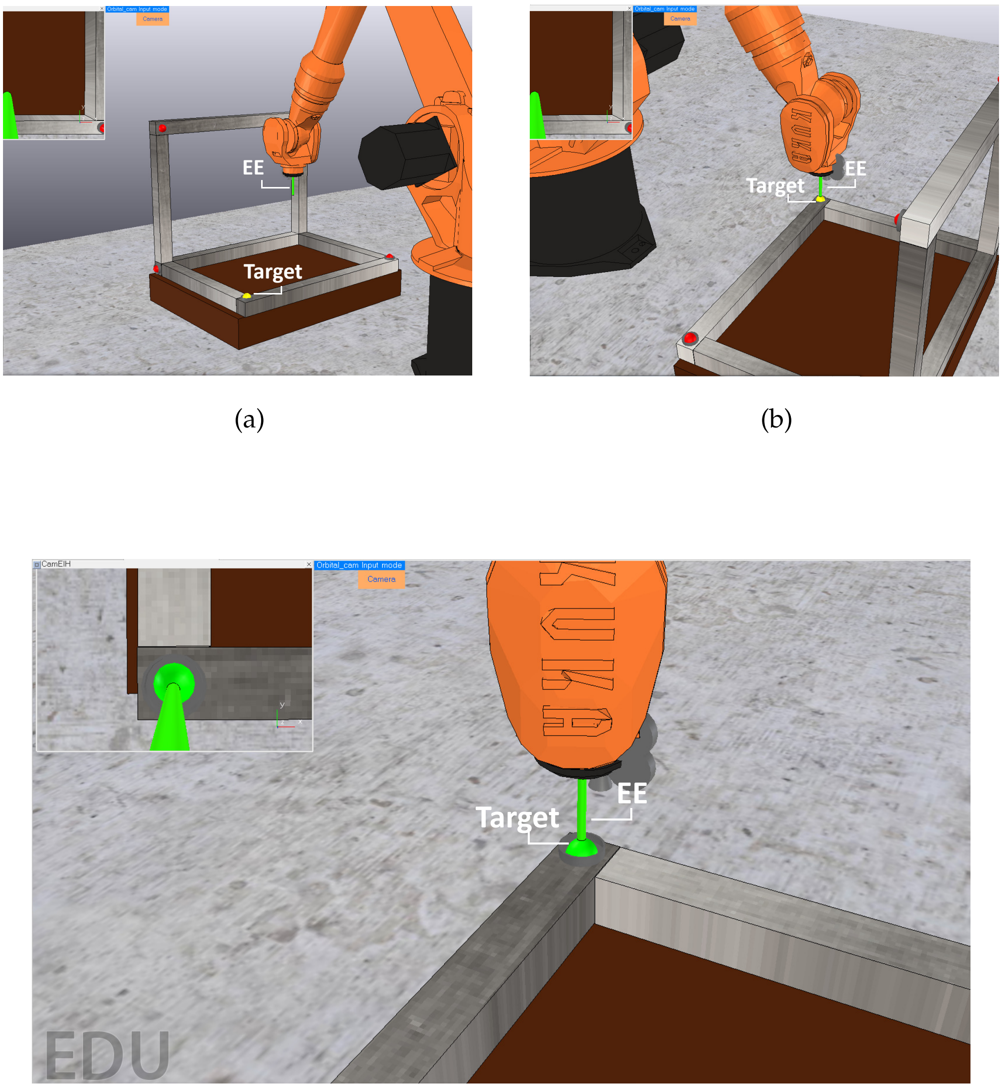
The pose $\mathbf{x_{EE}^{cam}}$ can then be updated based on the operator input with the new position and orientation in cartesian axes. Although the position along the cartesian axes is updated with a relatively simple increment/ decrement operation, the equivalent cannot be achieved with an Euler angle representation of orientation. The axis-angle representation of the change in orientation converted to quaternion can be used with eq. 6.1-6.3 to increment/ decrement orientation. The new pose $\mathbf{x\_des_{EE}^{cam}}$ can be set as the new target pose for the robot EE and solved using IK.

Task Space Reduction

A major benefit of using the orbital camera view to control a telemanipulator is the ability to reduce the DOF of positioning the robot in task space from 3-DOF to 1-DOF, fig.6.2. Given a target position (or intermediate position if target position is obstructed) for the robot EE in the orbital camera view, fig.6.2(a), the camera can be positioned to align the EE with the target, fig.6.2(b). Subsequently, to reach the target position, the operator only has to move forward on a single axis until the target position has been reached, fig.6.2(c). This reduces the mental workload of the operator by avoiding the need to align robot movement using multiple cameras/ monitors as well as reducing the risk of hazards by minimising the time and space the robot is in motion.

For the purposes of this experiment, the orbital camera is simulated using the V-rep simulation environment. However, in a real robot deployment, the orbital camera can be implemented in three ways. It can be mounted on the end effector of a secondary arm or a mini-drone in the remote environment depending on the environment constraints. Through calibration of the secondary arm or the drone, the camera can then be controlled using inverse kinematics in the Cartesian axis. In addition, the orbital camera in its current virtual implementation can be used with a real robot and a virtual reconstruction of the environment to improve the situational awareness of the operator.

In this implementation, it is assumed that collision avoidance is implemented in the camera platform. There is numerous literature dealing with real-time collision avoidance for robot arms [178], [179] and UAVs [180], [181]. The latest commercial drones now also come with built-in collision avoidance technologies. Alternatively, drones can be combined with roll cages [182] in non-sensitive environments where small collisions can be tolerated.



(a) (b)

(c)

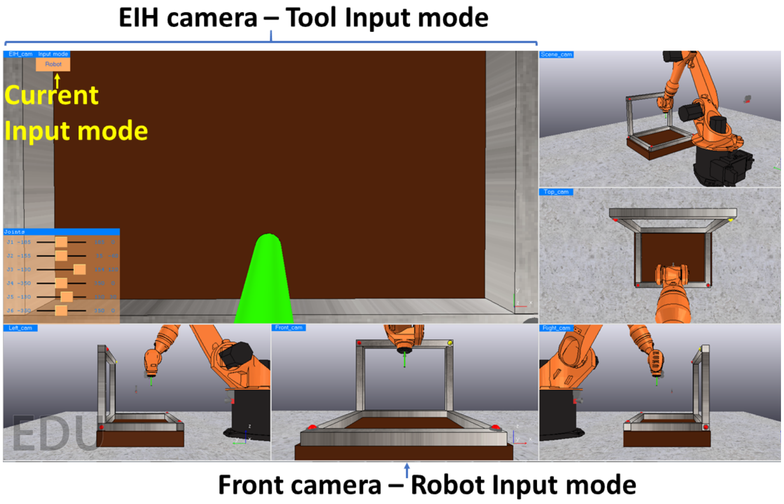
FIGURE 6.2: Demonstration of aligning robot EE with target to reduce task complexity when using the orbital camera. The target is brought into the camera view (a); the camera is rotated to align the target and EE (b); the robot EE is moved forward until the target is successfully (target changes colour to green) reached ((c).

6.3 Experiment

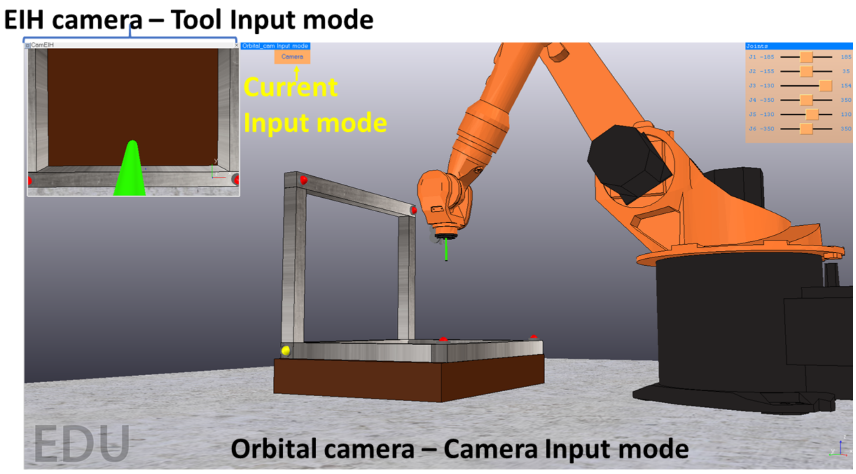
This study was designed to evaluate the performance of participants in a simple remote push button task using two modes of camera feedback as the independent variables, static mode and dynamic mode. The static mode consists of the six camera views as described in 6.2.1 while the dynamic mode is comprised of the orbital camera view and the EIH camera view, fig. 6.3. The task consisted of participants manoeuvring a virtual robot to sequentially press six buttons using the robot end effector on a frame presented in a virtual environment. The previous experiment, (section 4.2), highlighted that participants mostly avoided changing the orientation of the robot as much as possible due to it increasing control complexity. Therefore, two of the six buttons were deliberately placed on the side of the frame to force participants to change the robot's orientation and is also the primary reason for the removal of the block stacking task in section 4.2. This task represents the fundamental point-to-point motion required of a teleoperated robot arm to precisely control both position and orientation of the EE to reach a desired location. The task was chosen so that the two camera configurations could be evaluated on basic motion alone with further more complicated tasks to be used in the future.

6.3.1 System

The experimental setup consisted of a typical teleoperation master-slave interface. The master interface consisted of a monitor for visual feedback and a gamepad controller for input. This formed the "local workspace". An Xbox 360 wired controller was used by the participants for controlling the arm and orbital camera. Movement was implemented using Cartesian control as in section 4.2. The control mapping is shown in fig. 6.5(a). For each camera mode, user input could be switched between two reference frames. In static



(a)



(b)

FIGURE 6.3: Setup of task with push button frame in static mode: (a); and dynamic mode (b).

mode, the robot base was used as the main reference frame while the orbital camera was used for the dynamic mode. In both modes, control could be switched to/from the tool frame in the EIH camera.

The slave interface comprised a robot, the KUKA KR16 arm, along with cameras for visual feedback in a remote location which acted as the "remote workspace". The remote workspace was simulated in the V-rep simulation software. The simulated environment also contained the push button frame situated near the robot arm on top of a cuboid block with the tip of the robot

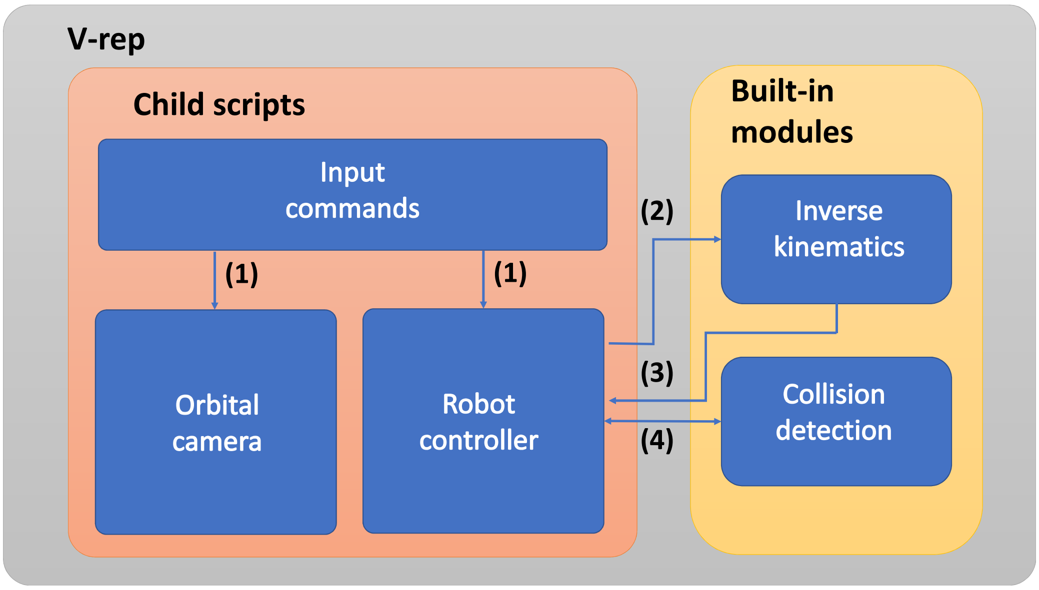


FIGURE 6.4: Software architecture. Extension of the testbed architecture adding an operator-controlled camera.

arm . The arm was configured to be controlled using the built-in V-rep inverse kinematics module. The experiment setup is shown in fig. 6.5(b).

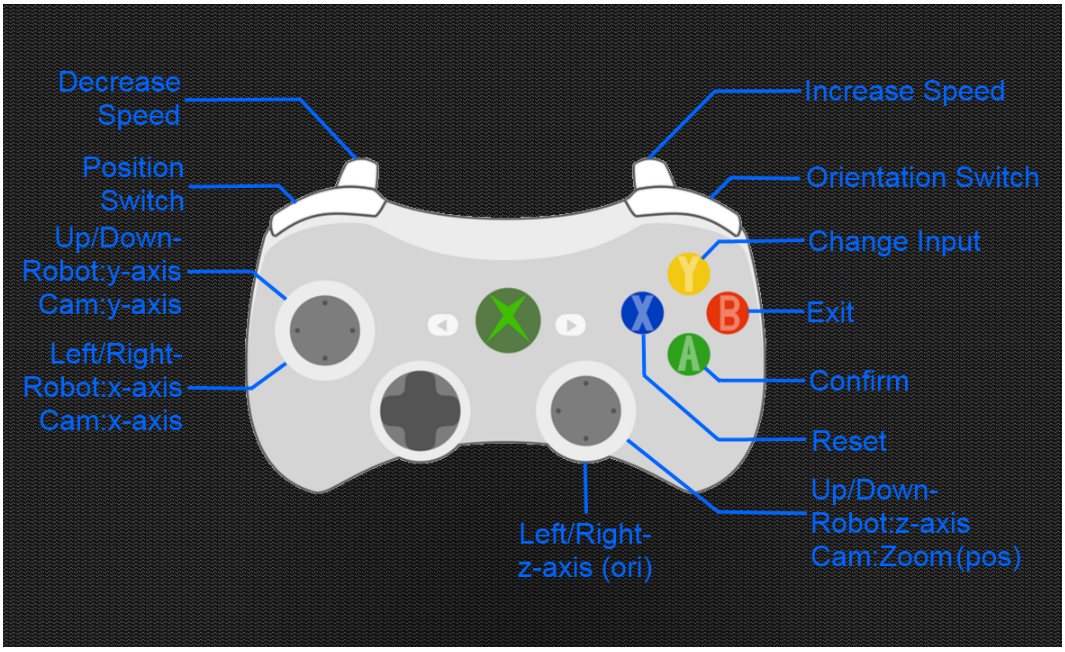
6.3.2 Participants

Participants consisted of volunteers from the general population of the University of Birmingham. Ads were placed around the university and on the Psychology school research participation system. No restrictions were placed on participation. A total of 9 participants were recruited with only 5 participants (1 male,4 female, ages: 18-25) able to finish the experiment. Participants were also compensated with a £10 Amazon gift voucher for their time.

6.3.3 Method

Prescreening

Before the experiment, background information about participants was collected including previous experience in 3D environments (e.g. controlling



(a)

(b)

FIGURE 6.5: Input mapping for the Xbox controller (a); and experiment setup (b).

RC planes, experience with 3D software/ games etc.), vision and disabilities as well as a spatial awareness test to gauge spatial abilities. All five participants were right-handed with no vision impairment or disabilities. Most had little or no previous experience with 3D video games with one participant reporting moderate gaming experience and none holding a driving licence.

They also completed a spatial abilities test using a computerised version [183] of the mental rotations test in [184]. This standardized test requires subjects to determine as quickly as possible whether pairs of drawings of three-dimensional cubes are identical but rotated or different. Participants were instructed to complete the test of 25 questions as quickly as possible although no time limit was set.

Spatial scores were weighted to account for completion speed and participants guessing answers. Although a time limit was not set, participants were expected to finish within five minutes giving a reasonable 12s for each question. The weighted score is calculated as follows:

- Raw score ($0 - 100$): 4 points for each correct answer
- Time bonus: 1 point for every $10s$ less than 300s, up to a maximum of 10 points
- Time penalty: -1 point for every $5s$ above 300s, with a maximum penalty of -30 points
- Score penalty: -2 points for each incorrect answer, with a maximum penalty of -40 points
- Weighted score ($-70 - 110$): raw score + time bonus/penalty + score penalty

Training

A standard instruction sheet was used to brief the participants on the experiment after which they were allowed to familiarise with the system in practice mode until they felt ready to proceed. The practice mode consisted of three

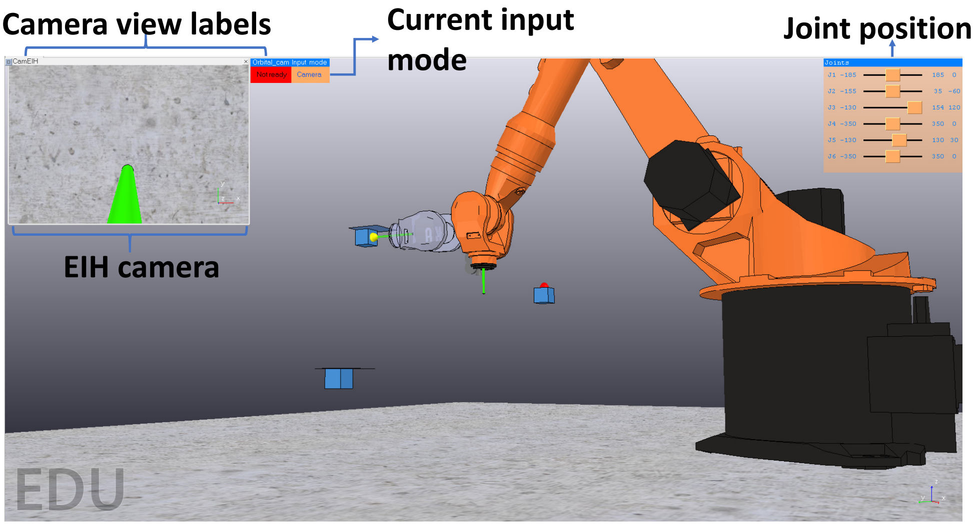


FIGURE 6.6: Training mode with three training buttons each chosen randomly to be pressed. Current button is highlighted yellow with a guide.

buttons suspended in space in different orientations which together necessitated the full use of the gamepad controls to press, fig. 6.6. The button to press was randomly chosen and the task could be repeated indefinitely. During this training period, participants were allowed to ask questions. A brief demo was given on how to move the robot to press the buttons, use of cameras in relation to robot movement and optimal use of controls. To achieve a similar level of competency between all participants, a performance target was set. Three buttons had to be pressed in under a minute before proceeding to the live task. Due to resource limitations, after 15 mins, if the participants had not reached the competency goal, a judgement was made by the experimenter to continue to the experiment or abandon it based on the participant's progress. If the participant made sufficient progress where they could press the training buttons using the robot arm but not fast enough to reach the competence goal, they were allowed to continue on to the actual task. Training was available before the live trials for each camera mode.

Trials

After the training session, participants completed three trials of the task using each camera mode with each trial randomising the button sequence. All participants were informed to minimise collisions and only reset the robot when absolute necessary. A trial started when the first input was sent by the user to the robot and ended successfully when all six buttons had been pressed. Counterbalancing was used to remove the effects of learning, fatigue and/or boredom from confounding the results by randomising the order in which participants used each of the camera modes. The various metrics recorded during the experiment are shown in table 6.1. Upon completion of the trials, for each camera mode, the participants were asked to complete a NASA-TLX questionnaire as well as a system usability scale (SUS) questionnaire.

Metric	Operational definition
Task completion time/s	Robot base (mode 1 ONLY)
Collisions	No. of unintentional collisions with environment (does not include buttons)
Ideal Path length/s	Straight line distance travelled between buttons
Actual Path Length/m	Actual distance covered by the arm
Camera	Time spent controlling orbital camera (mode 2 ONLY)
Input modes time/s	Time spent in each input mode (e.g.robot,camera)
Control modes time/s	Time spent controlling the arm using orientation and position
Speed modes time/s	Time spent in each speed mode (i.e. slow, normal, fast: for position and orientation input)
Robot idle time/s	Time spent thinking (no user input) or using orbital camera
Arm resets	No. of times arm reset to starting position after getting stuck
Arm position	Arm pose at approx. 50Hz
Camera position	orbital camera pose at approx. 50Hz

TABLE 6.1: Experiment metrics and their definitions

6.4 Results

The experiment had to be prematurely ended for modifications due to the influence of confounding factors from unforeseen sources (see section 6.5). This section briefly presents the results from the five participants who completed this pilot experiment. In addition, participant 9 only performed one trial of the task in static mode.

The statistical analysis consisted of **two-tailed paired samples t-tests** to test the sample means for significant difference. Two-tailed tests so that a positive or negative effect could be found between the two camera modes and the paired samples version of the t-test was used since the samples in the experiment was generated using a within-subject design. **Pearson's r** was used when the strength of a linear correlation between two variables needed to be determined. The Pearson correlation coefficient r can take values ranging between -1 to 1 with values close to the extrema indicating strong association between two variables. The effect size of this measure can also be determined by r^2 which gives a measure of the extent to which one variable affects the other in percentage.

6.4.1 Performance Metrics

Tables 6.2 and 6.3 show a summary of the dispersion measures for each participant on each camera mode, static ($n = 5, \bar{x} = 264, \sigma = 72$) and dynamic ($n = 5, \bar{x} = 376, \sigma = 165$).

ID	Spatial Score		TTC/s		Environment Collisions		Path ratio	Arm resets		Performance Score
	Raw	Weighted	$\bar{x}$	σ	$\bar{x}$	σ		$\bar{x}$	σ	
2	24	-4.00	187	61	2.33	0.58	2.47	1.33	1.53	250
4	56	44.00	190	47	3.67	3.21	4.03	1.33	1.15	266
6	76	58.95	190	24	11.33	7.64	2.68	**0.67**	0.58	323
7	**84**	**77.90**	**118**	24	2.33	0.58	2.85	0.33	0.58	**151**
9	68	62.00	280	0	**2.00**	0.00	**2.22**	1.00	0.00	330

TABLE 6.2: Results for static camera mode with average score $\bar{x} = 264, \sigma = 72$. Path ratio is ratio of the robot's actual path to ideal path. Underlined values are best in column. Participant 9 performed ONLY 1 trial. *Note: Lower performance score is better*

ID	Spatial Score		TTC/s		Environment Collisions		Camera/s		Path ratio	Arm resets		Performance Score
	Raw	Weighted	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	%		$\bar{x}$	σ	
2	24	-4.00	224	29	**1.33**	2.31	28.67	12.80	2.76	1.00	0.00	267.33
4	56	44.00	299	30	29.33	16.17	**41.00**	13.71	4.44	1.00	0.00	622.33
6	76	58.95	**151**	44	3.00	2.65	*14.33*	9.51	2.38	0.67	0.58	**200.67**
7	**84**	**77.90**	232	34	7.00	5.57	23.00	9.93	4.17	1.33	1.15	341.67
9	68	62.00	383	34	5.33	8.39	39.33	10.28	**2.51**	**0.33**	0.58	446.00

TABLE 6.3: Results for dynamic camera mode with average score $\bar{x} = 376, \sigma = 165$. Underlined values are best in column. *Note: Lower performance score is better*

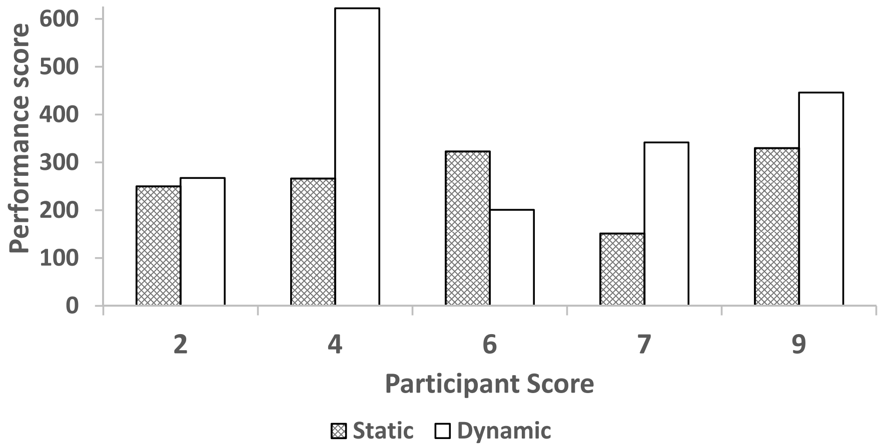


FIGURE 6.7: Performance scores for all participants for both camera modes, static: $\bar{x} = 264, \sigma = 72$, dynamic: $\bar{x} = 376, \sigma = 165$. Lower is better.

Performance score is parameterised by *TTC, collisions and arm resets* as $Score = TTC + (10 \times collisions) + (30 \times armresets)$. The penalties were calculated based on the literature but extra emphasis was given to *arm resets* due to it being a critical incident to have to reset the robot in a hazardous environment due to loss of control. Figure 6.7 plots the performance score of all participants for both camera modes. Participant 4's spike in performance score for the dynamic mode is mainly due to the *number of collisions*. The number of collisions were recorded programatically where the collision variable was incremented by 1 for every second the robot was in a state of collision. The act of thinking for a prolonged period of time ($>$ 10s) about how to move the robot to exit a collision state resulted in the participant's final score being doubled in this case.

A two-tailed paired samples t-test was conducted to compare the scores between camera modes. A significant difference was not found $t(4) = 1.38$, $p = 0.24$. This seems to suggest that the orbital camera does not offer a significant improvement over the static cameras and based on the large standard deviation, there is high variation in performance among the participants.

Figure 6.8 shows the TTC (base performance score) of all participants. A paired samples t-test did not reveal a significant difference $t(4) = 2.20, p =$

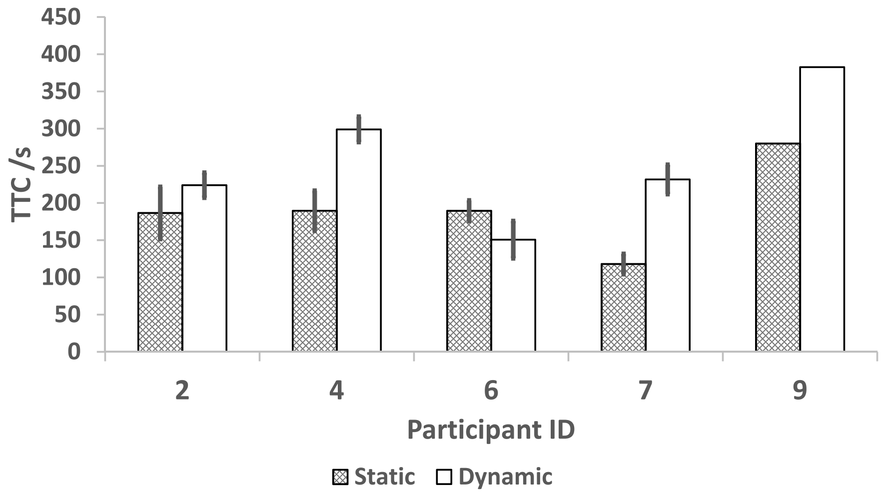


FIGURE 6.8: Time to Completion (TTC) for all participants for both camera modes, static: $\bar{x} = 193s, \sigma = 39s$, dynamic: $\bar{x} = 258s, \sigma = 34s$.

0.09 between static ($\bar{x} = 193s, \sigma = 39s$) and dynamic ($\bar{x} = 258s, \sigma = 34s$) camera modes.

Further analysis of metrics was deemed superfluous due to the limited sample size and subsequently limited statistical power. However, experimenter observations and fig.6.9 revealed key information on user input mode usage. From the three input modes: robot (control based on the robot base frame), camera (control based on the orbital camera frame) and tool (control based on the robot wrist frame), participants tended to favour the tool mode input over robot or camera with some completely ignoring the camera/ robot input modes for the tool mode. Although tool mode was intended for use when precision was required, participants also frequently used it to navigate between the buttons on the frame.

6.4.2 Subjective Metrics

This section covers the subjective metrics measured to supplement the performance metrics. NASA-TLX was used to measure the participants' subjective workload and the SUS was used to measure the subjective usability rating for each camera mode.

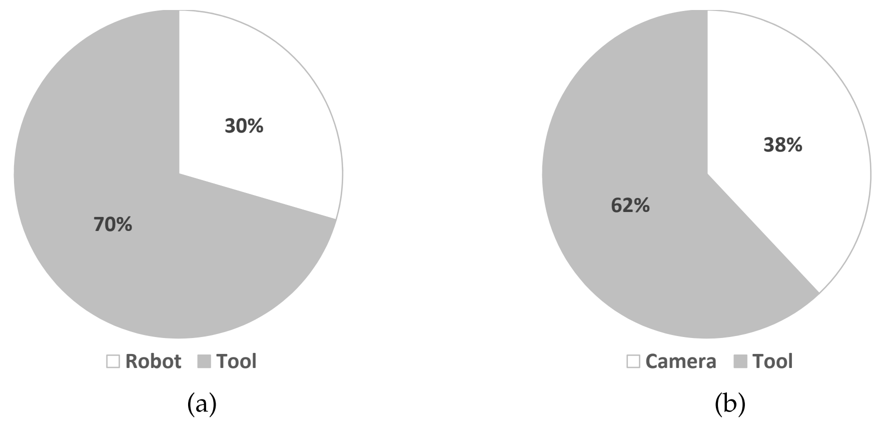


FIGURE 6.9: Comparison of input mode usage for each participant with static cameras: (a); and dynamic camera (b).

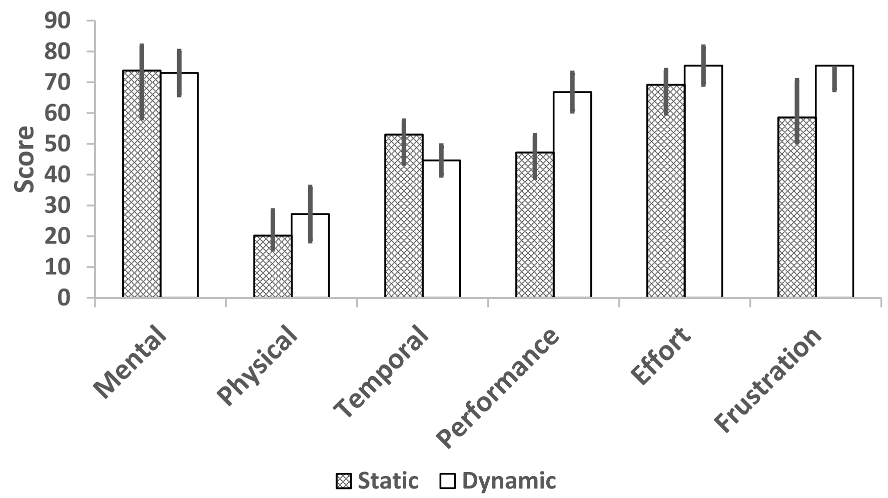


FIGURE 6.10: Average NASA-TLX scores across all components for both camera modes, static: $\bar{x} = 53, \sigma = 7$, dynamic: $\bar{x} = 58, \sigma = 10$. Max score: 100

NASA-TLX did not show any significant differences with similar scores for each camera mode, fig.6.10, static ($\bar{x} = 53, \sigma = 7$) and dynamic ($\bar{x} = 58, \sigma = 10$). Neither camera mode offered benefits in terms of lower workload for the participants with both modes being equally mentally challenging and effort inducing.

SUS showed significant difference $t(5) = 3.00, p = 0.01$ between camera modes, fig. 6.11, static ($\bar{x} = 63, \sigma = 9$) and dynamic ($\bar{x} = 46, \sigma = 20$). Static mode was rated as having better usability. *Note: The df for this t-test was five due to using the SUS results from a participant who attempted both camera modes*

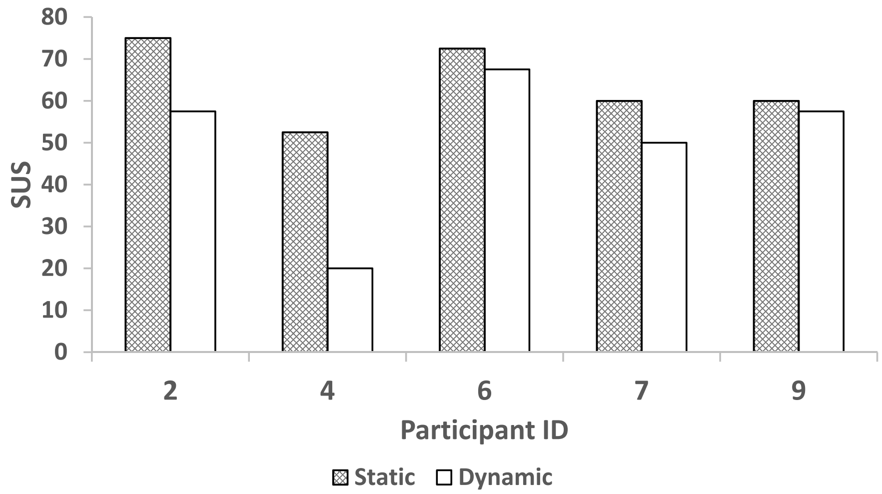


FIGURE 6.11: Average SUS for all participants for both camera modes, static: $\bar{x} = 63, \sigma = 9$, dynamic: $\bar{x} = 46, \sigma = 20$. Max score: 100

but did not successfully complete the trials.

6.4.3 Demographics

Based on the responses to the background questionnaire, participants were given a score for gaming experience out of 125 by multiplying the values of three Likert scales ranging from 1 to 5. In the previous experiment (section 4.2) , mixed ANOVA analysis found a significant difference between gaming experience and task performance. Examining this relationship again using Pearson's correlation found a significant negative correlation with performance scores in the dynamic mode ($r(3) = -0.62, p = 0.05, r^2 = 38\%$)(see fig. 6.12) but not with the static mode ($r(3) = -0.25, p > 0.05, r^2 = 6\%$). Other demographic factors such as age, gender, right/left-handedness could not be investigated due to insufficient data.

6.5 Discussion

The experiment had to be terminated early due to unaccounted confounding factors affecting the experiment validity. Participant skill levels remained

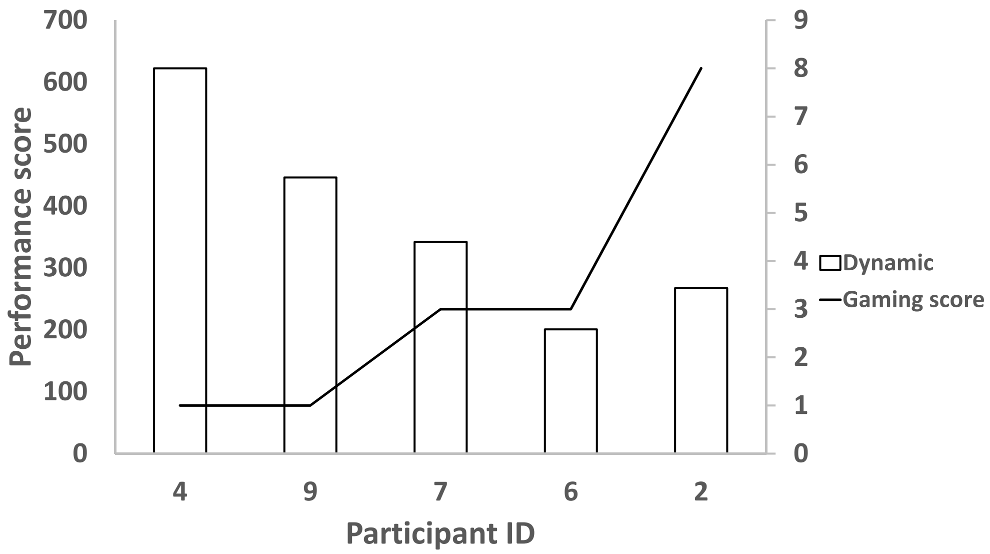


FIGURE 6.12: Graph depicting relationship between gaming score and task performance in dynamic mode. Pearson's r showed a strong correlation: $r(5) = -0.62, p = 0.05, r^2 = 38\%$

chaotic despite the use of a training mode with a demo given. This resulted in highly volatile data making statistical examination of the effects of the two camera modes unreliable.

Improper training meant that participants were not competent enough to control the telerobotic system effectively. Symptoms of ineffective training included controls such as orientation being ignored, inefficient use of cameras and over-reliance on tool mode input and EIH camera view.

Although background information was collected from participants, there was no prescreening. However, the experiment could have benefited from screening participants for gaming experience. A correlation between gaming experience and performance, see 4.4.3, was confirmed again in this experiment albeit for only the dynamic mode. This could be due to the transfer of skills gained from experience of traditional third-person cameras in video games that the orbital camera is based on.

For the training demo, participants were given all the instructions and demonstration before starting training. This information overload meant that instructions not retained were ignored diminishing the benefit of training. This was evident with participants finding it difficult to grasp the controls resulting in controls such as orientation being ignored as it was not completely necessary to press the training buttons.

There was also the need to balance the amount of training time and total experiment time. Only one participant managed to pass training achieving the goal of pressing three buttons in under a minute and on average participants spent 13 mins on training over both camera modes with 15 mins allocated to training. This method of training was clearly not sufficient.

Participants overwhelmingly used the EIH camera view coupled with the tool mode to operate the robot. For the static mode, this meant the other views were not needed except for finding the next target location. However, the orbital camera in the dynamic mode had to be positioned into similar positions as those readily provided in the static mode resulting in longer completion times. This also led to the ineffective use of the orbital camera. Its use was mostly relegated to finding the button to press and using the tool mode with the EIH camera to press the button.

This endorses the view in the literature that egocentric views are preferred to exocentric views for navigation. However, for telemanipulators performing manipulation using the end-effector, the EIH camera view will not always be available and may be blocked when an object is grasped at the tool. This prevented a sound evaluation of the camera modes, as the main benefit of the orbital camera mode was to offer alternative dynamic views in the event that the EIH camera is not available.

6.6 Revised Experiment

The previous short "pilot" experiment revealed key areas for improvement so that a valid evaluation of the two camera modes could be performed. This section covers the modifications made and presents the results from the experiment.

The above experiment was modified so that for half of the time the eye-in-hand camera view was partially hidden by a plate at the end effector simulating an object being held by the robot arm. This simulated a more realistic scenario and made it more difficult to rely on the EIH camera view all of the time and forced participants to use other camera views.

6.6.1 Modifications to Method

The experimental method was revised to improve metric definitions, prescreen participants, improve training procedures and gain relevant user feedback as discussed here.

Prescreening

In the previous experiments, results showed that gaming experience and task performance were correlated. To improve the quality of the results in future experiments and in the interest of saving time by avoiding failed experiment runs, participants were prescreened for gaming experience.

The experiment recruitment advertisement was covert and did not mention any restrictions placed on participant selection to avoid priming effects. Interested participants were asked to fill out a prescreening questionnaire which included questions on gaming experience using Likert scales among other demographic questions. Participants with no gaming experience were then excluded from the experiment.

Participants were given a score based on their gaming experience. A *gaming score* metric was developed to quantify gaming experience so that its effects if any, could be analysed from the results. A literature search did not reveal any previous work on such a metric with the popular option being to categorise participants into low, medium and high bins based on response. The *gaming score* metric is somewhat subjective but provides a rough estimate of participants' relative gaming experience. The scoring system used was as follows:

1. 0, 0.5, 1, 3 or 5 points for gaming frequency
2. 0 - 5 points for experience within the last year
3. 0 - 5 points for experience with gamepad controllers within the last year
4. 0 - 5 points for experience with 3D simulators within the last year
5. 9 points for having a full driving licence
6. Gaming score was then calculated as: $[1] \times ([2] + [3] + [4]) + 5$ to give a total score out of 84

The sum of the experiences within the last year was multiplied with gaming frequency so that those who played games more often had relatively higher scores compared to those who played irregularly or had stopped playing. The reasoning for this stems from human nature where it is common knowledge that memory retention or skill level diminishes if not continuously practiced. However, this reduction is not linear and hence the use of a non-linear scale for gaming frequency. Points were also awarded for being able to drive a car as this requires similar cognitive skills to moving around in games on a 2D plane using a controller. 9 points were given to signify its equivalence in gaming experience as someone who plays games occasionally.

Training Improvements

Despite the inclusion of a demo in the training mode, participants still found it difficult to grasp key concepts and robot controls. Therefore, further improvements were required for the training mode.

The demo was broken down into a step-by-step tutorial rather than a full demo at the beginning of the training mode. Participants were gradually introduced to the controls and concepts with practice in between to reduce information overload.

One of the big issues was the lack of training in orienting the robot. To fix this, covers were added near the buttons to prevent them being pressed from certain directions. This forced participants to practise changing the robot orientation.

The competency goal of pressing three buttons in under a minute for training was also too high and unrealistic with most participants finding it difficult to achieve. This was reduced to two buttons in under a minute.

Redefined Collision Metric

Data collected in the previous experiment did not give an accurate picture of robot collisions with the environment highlighting the effect of imprecisely operationalising variables. While some participants quickly moved to safety after a collision, others required more time to think about how to recover from the collision.

As it is more interesting to find how many collisions occured, the collision metric was redefined. Rather than incrementing collision count every second the robot was in a state of collision, it was incremented only when the robot moved whilst in a state of collision.

Post-experiment interview

Subjective metrics often fail to capture the full picture and can be open to interpretation both from the experimenter's and participant's point of view. Therefore, a post-experiment interview method was also employed to gather the views and thoughts of participants. A brief interview was conducted after the experiment to ascertain which camera mode if any, the participants preferred and what they thought of each camera mode.

6.6.2 Results

This section presents the results from the revised experiment. T-tests were conducted on a number of metrics to assess for significant differences and Pearson's correlation was used to discover association between relevant metrics.

Performance Metrics

Table 6.4 shows a summary of the dispersion measures for each participant on each camera mode, static ($n = 16, \bar{x} = 368, \sigma = 91$) and dynamic ($n = 16, \bar{x} = 415, \sigma = 92$). At first glance, it seems overall performance using the static mode was better with negligible differences among the other metrics.

T-tests

To determine whether the differences between the two camera modes were significant, t-tests were used. A summary of the results from two-tailed paired sample t-tests on various metrics is given in Table 6.5.

Camera Mode	TTC /s		Environment Collisions		Path ratio		Arm resets		Performance Score		Camera Control /s		
	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	%
Static	**315**	**67**	3.92	2.98	3.11	0.63	0.67	0.89	**368**	**91**			
Dynamic	371	86	**3.42**	**2.36**	**3.04**	**0.61**	**0.51**	**0.88**	415	92	79	29	31

TABLE 6.4: Results for both camera modes in the push button task ($n = 16$). Underlined values are best in column. Path ratio is ratio of the robot's actual path to ideal path. *Note: Lower performance score is better*

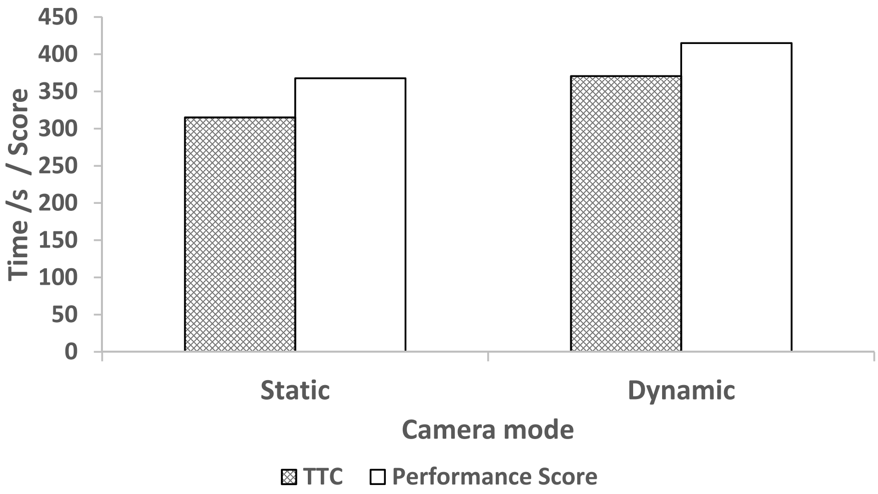


FIGURE 6.13: TTC vs Performance Score for both camera modes

No statistical significance was found for performance scores ($t(15) = -1.35, p = 0.20$), suggesting task performance between the two camera modes was similar, see fig.6.13. Similarly, *TTC* showed a non-significant difference between the two camera modes, static ($\bar{x} = 315s, \sigma = 67s$) and dynamic ($\bar{x} = 371s, \sigma = 86s$); $t(15) = -1.91, p = 0.07$ suggesting similar task completion times.

TTC was also investigated for learning effects due to ordering. Participants were split into two groups, *static-dynamic* and *dynamic-static*, depending on which camera mode they completed first. The mean TTCs for each group are shown in fig. 6.14. There was no significant difference across the two groups for either camera mode, static ($t(7) = 1.74, p = 0.13$) and dynamic ($t(7) = 2.10, p = 0.07$) although the t-score for the dynamic mode is so close to the t-critical value $t - crit = 2.36$ such that a slightly larger sample is likely to show a significant learning effect. This is supported by the t-tests within the two groups, static-dynamic ($t(7) = 0.49, p = 0.64$) and dynamic-static ($t(7) = -3.75, p < 0.01$) which indicates that there was a significant learning effect for participants who completed the dynamic mode followed by the static mode, see fig. 6.14.

Metric	t-value	p-value
Performance Score	-1.35	0.19
TTC /s	-1.91	0.07
Collisions	0.71	0.49
Path ratio	0.17	0.87
Input time /s	**3.05**	**0.01**
Position time /s	**3.52**	**0.00**
Orientation time /s	1.67	0.12
Idle time /s	**-4.65**	**0.00**
Idle time w/o Camera /s	-0.76	0.46

TABLE 6.5: Summary of t-values from t-tests between the static and dynamic camera modes with *df = 15, α = 0.05*. Bold values represent significant difference.

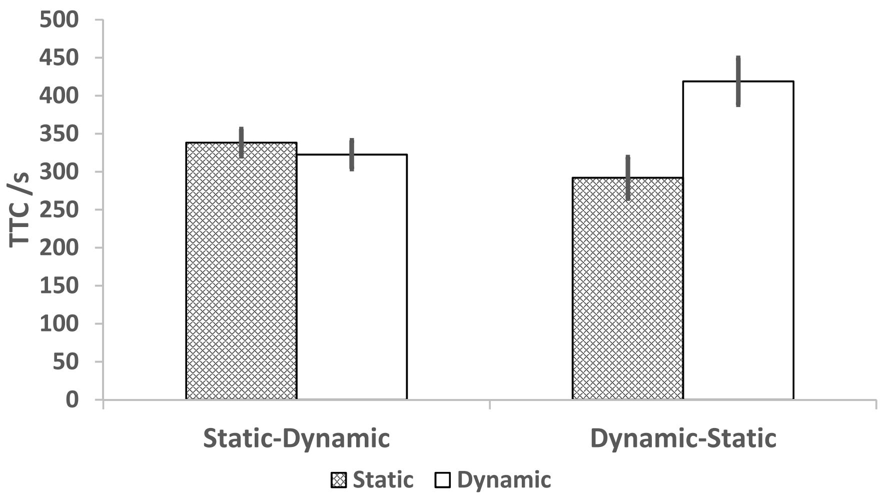


FIGURE 6.14: The effect of condition order on TTC.

The effect of the camera mode on the number of *collisions* with the environment was also not significant, $t(15) = 0.71, p = 0.2$ although the dynamic mode had minimally fewer collisions ($\bar{x} = 3.42, \sigma = 2.36$) compared to static mode ($\bar{x} = 3.92, \sigma = 2.98s$).

Path ratio compares the ratio between the actual path length traversed by the robot and straight line path length between buttons for each mode, static ($\bar{x} = 3.11, \sigma = 0.63$) and dynamic ($\bar{x} = 3.04, \sigma = 0.61$). As the ratios were very similar, the t-test also showed no disparity, $t(15) = 0.28, p = 0.78$.

Time spent on user input from the controller to the robot was also measured for each camera mode. Total time spent on input (static: $\bar{x} = 191s, \sigma = 50s$, dynamic: $\bar{x} = 157s, \sigma = 36s$) showed the dynamic mode required significantly less input, $t(15) = 3.05, p = 0.01$, suggesting it was easier to control the robot in this mode. However, the overall TTC for the dynamic mode was higher due to also having to control the camera as discussed in the next paragraph.

Time spent controlling the robot's position, orientation and idle time were also measured. Position time (static: $\bar{x} = 129s, \sigma = 35s$, dynamic: $\bar{x} = 103s, \sigma = 24s$) revealed a significant difference, $t(15) = 3.52, p < 0.01$. This implies it was easier to position the robot in the dynamic mode. However there was no effect, $t(15) = 1.67, p = 0.12$, on the orientation time (static: $\bar{x} = 61s, \sigma = 17s$, dynamic: $\bar{x} = 52s, \sigma = 16s$). Idle time was shown to be affected ($t(15) = -4.65, p < 0.01$) by the camera mode (static: $\bar{x} = 126s, \sigma = 28s$, dynamic: $\bar{x} = 215s, \sigma = 66s$) however removing the time spent controlling the orbital camera from the idle time in dynamic mode ($\bar{x} = 137s, \sigma = 51s$) showed no significance, $t(15) = -0.76, p = 0.46$. This suggests a significant amount of time during the task is spent on controlling the orbital camera while thinking time remains the same.

To summarise the above two paragraphs, the dynamic mode allowed for quicker positioning of the robot. However, the significant time required to think and operate the orbital camera nullified this benefit.

Pearson's correlation

Pearson's correlation was used to assess the relationship strength between a number of metrics across the two camera modes and within each camera mode. Where a statistically significant result is found, the power of the result is also reported. A generally accepted value of power is 0.8. When the power is greater than this threshold, the chance of missing a significant difference is less than 20%. A summary of the results is given in section 6.6.

Cross comparison between camera modes of *performance score* and *TTC* showed no significant correlation, $r(14) = -0.18, p > 0.05, r^2 = 3\%$) and ($r(14) = -0.08, p > 0.05, r^2 = 0\%$) respectively, between the camera modes suggesting performance on one did not affect the other. However, a significant yet moderate positive correlation was found for *collisions*, $r(14) = 0.47, p = 0.05, power = 0.45, r^2 = 22\%$, indicating that neither camera mode helped to reduce collisions.

There was a medium strength correlation between time spent using the *robot input* mode and *idle time*, $r(14) = 0.52, p = 0.05, power = 0.55, r^2 = 27\%$. This is likely due to participants finding it difficult to map the input based on multiple camera views.

In dynamic mode, *performance score* and *TTC* both corresponded positively to time spent on *dynamic camera* control, $r(14) = 0.51, p = 0.05, power = 0.53, r^2 = 26\%$ and $r(14) = 0.63, p = 0.05, power = 0.76, r^2 = 40\%$, suggesting inefficient use of the orbital camera. This is corroborated by the lack of a connection between the previous two metrics with *%camera* use, $r(14) = -0.12, p > 0.05, r^2 = 1\%$ and $r(14) = 0.00, p > 0.05, r^2 = 0\%$, implying that participants that spent a higher percentage of time using the orbital camera did not necessarily take longer to complete the task.

Mode	Metric		r	p	Power	r^2 /%
Static vs Dynamic						
	Performance Score		-0.18	>0.05	<0.8	3
	TTC		-0.08	>0.05	<0.8	0
	Collisions		**0.47**	**0.05**	**0.45**	**22**
Static						
	Robot Input vs	**Idle Time**	**0.52**	**0.05**	**0.55**	**27**
		Orientation Input	**0.76**	**0.05**	**>0.8**	**58**
	Tool Input vs Slow Positioning		**0.61**	**0.05**	**0.73**	**37**
Dynamic						
	Camera Control vs	**Performance Score**	**0.51**	**0.05**	**0.53**	**26**
		TTC	**0.63**	**0.05**	**0.76**	**40**
	%Camera vs	Performance Score	-0.12	>0.05	<0.8	1
		TTC	0.00	>0.05	<0.8	0
	Path Length vs	**Performance Score**	**0.59**	**0.05**	**0.69**	**35**
		TTC	**0.50**	**0.05**	**0.51**	**25**
		Collisions	**0.50**	**0.05**	**0.51**	**25**
	Path ratio vs Performance Score		**0.55**	**0.05**	**0.61**	**30**
	Camera Input vs Orientation Input		**0.83**	**0.05**	**>0.8**	**69**
	Tool Input vs Slow positioning		**0.56**	**0.05**	**0.62**	**31**

TABLE 6.6: Summary of Pearson's r across and within both camera modes. Bold values are statistically significant.

Performance score and *TTC* also showed a significant relationship with the robot's *path length* in the dynamic mode pointing to inefficient control of the robot, $r(14) = 0.59, p = 0.05, power = 0.69, r^2 = 35\%$ and $r(14) = 0.50, p = 0.05, power = 0.51, r^2 = 25\%$ however only the *performance score* correlated significantly with the *path ratio*, $r(14) = 0.55, p = 0.05, power = 0.61, r^2 = 30\%$. Moreover, *path length* had an effect on *collisions*, $r(14) = 0.50, p = 0.05, power = 0.51, r^2 = 25\%$, indicating collisions increased proportionally to distance travelled.

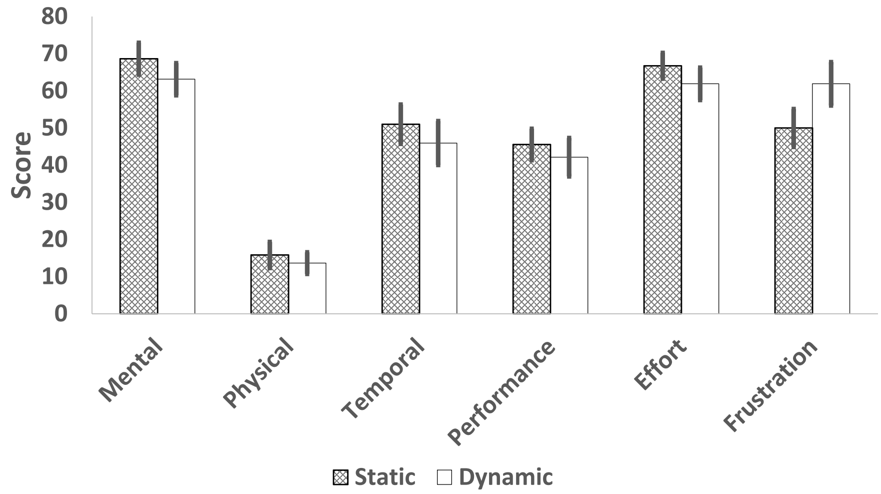
For both camera modes, *orientation* input to the robot had a notable connection to the main *input mode*; robot for static mode ($r(14) = 0.76, p = 0.05, power > 0.8, r^2 = 58\%$) and orbital camera for dynamic mode ($r(14) = 0.83, p = 0.05, power > 0.8, r^2 = 69\%$). This suggests it was easier to orientate the robot using the exocentric views rather than the egocentric views.

Tool input mode was used for precise movement of the robot as shown by its moderate correlation with *slow speed* positioning during both static and dynamic mode, $r(14) = 0.61, p = 0.05, power = 0.73, r^2 = 37\%$ and $r(14) = 0.56, p = 0.05, power = 0.62, r^2 = 31\%$ respectively.

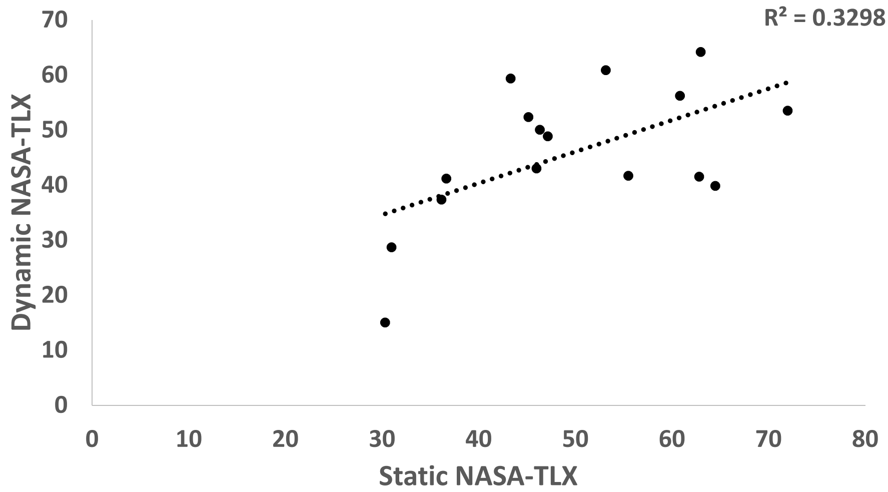
Subjective Metrics

Subjective analysis was used to complement the performance analysis in the previous section. The measures used were the NASA-TLX to measure perceived workload and the SUS to measure system usability.

The NASA-TLX ratings across the six subscales were similar for both camera modes (fig. 6.15(a)). The average workload (out of 100) was as follows: *static:* $\bar{x} = 50, \sigma = 13$; *dynamic:* $\bar{x} = 46, \sigma = 13$. Although the workload was slightly lower for the dynamic mode, no statistical significance was found, $t(15) = 1.29, p > 0.05$. NASA-TLX scores for both modes were significantly correlated, $r(14) = 0.57, p = 0.05, power = 0.65, r^2 = 33\%$ (fig. 6.15(b)), suggesting participants rated the workload across both camera modes equally.



(a)



(b)

FIGURE 6.15: (a): Average NASA-TLX subscale ratings (out of 100) for both camera modes; (b): Relationship between static and dynamic NASA-TLX scores.

SUS also gave a similar result between the two camera modes, *static:* $\bar{x} = 65, \sigma = 17$; *dynamic:* $\bar{x} = 68, \sigma = 14$; $t(15) = -0.68, p > 0.05$. Although the results were similar, the dynamic mode score of 68 is equal to the average score of according to analysis of normative data [107], [108]. Unlike for NASA-TLX ratings, there was no correlation in SUS scores between camera modes, $r(14) = 0.47, p > 0.05, power = 0.45, r^2 = 29\%$.

The SUS ratings were compared to the NASA-TLX scores (fig. 6.16) to determine a relationship between perceived workload and usability rating. For the static mode, workload had a substantial connection to usability, $r(14) =$

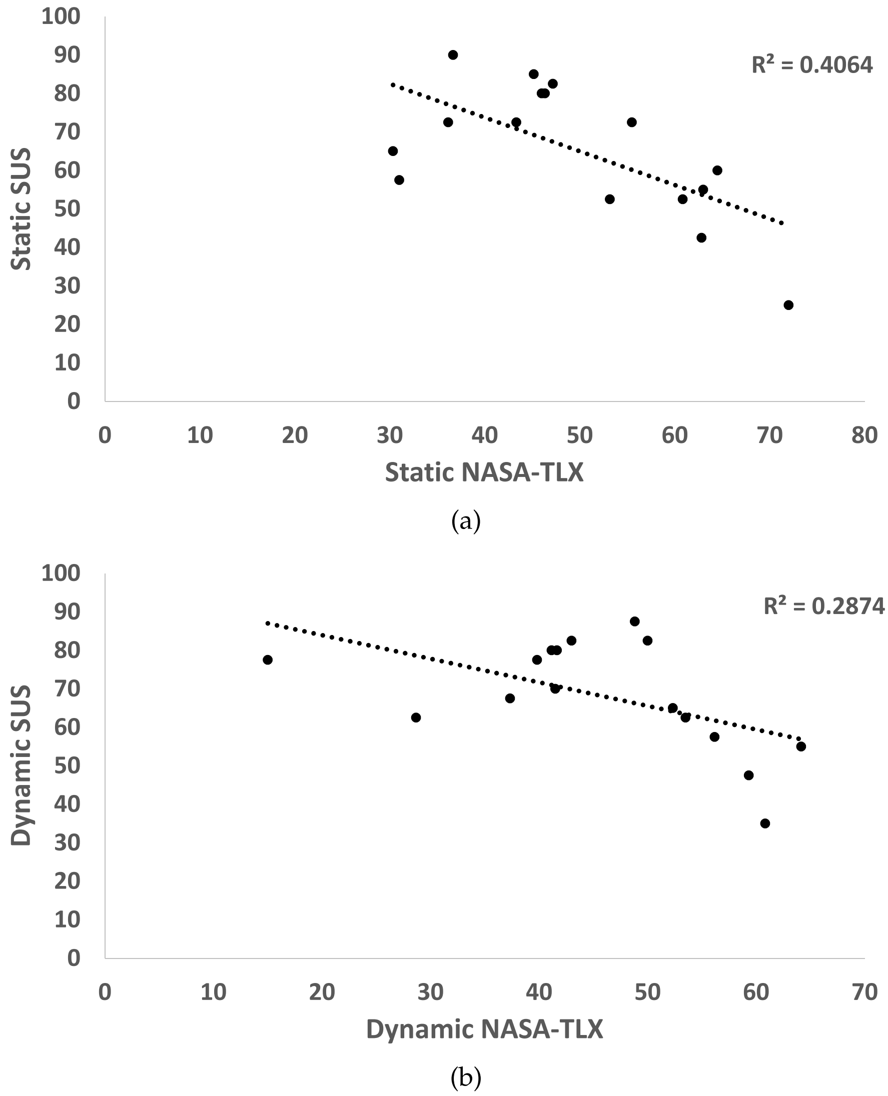


FIGURE 6.16: Perceived workload vs usability for both camera modes (a): static: $r(14) = -0.64, p = 0.05, power = 0.78, r^2 = 41\%$; (b): dynamic: $r(14) = -0.54, p = 0.05, power = 0.59, r^2 = 29\%$.

$-0.64, p = 0.05, power = 0.78, r^2 = 41\%$, indicating usability rating was negatively affected by the amount of workload experienced as shown in fig. 6.16(a). There was a weaker correlation with borderline statistical significance for the dynamic mode which showed greater variation (fig. 6.16(b)), $r(14) = -0.54, p = 0.05, power = 0.59, r^2 = 29\%$.

Analysis of subjective metrics against performance score for each camera mode did not reveal any significant correlations, see table 6.7.

The data collected from the post-experiment interviews showed that there

Metric		r	p	Power	r^2 /%
NASA-TLX	Static	-0.04	>0.05	<0.8	0
	Dynamic	-0.29	>0.05	<0.8	8
SUS	Static	-0.35	>0.05	<0.8	12
	Dynamic	-0.01	>0.05	<0.8	0

TABLE 6.7: Correlation between performance score and subjective metrics: NASA-TLX and SUS .

was a 9 : 6 split in favour of the dynamic mode with one participant abstaining. Due to the unequal sample sizes and variance between the two groups, *static_pref* and *dynamic_pref*, two-sample t-tests assuming unequal variance were used to analyse the means.

While there was no significant difference between static *performance scores* (*static_pref*: $\bar{x} = 334, \sigma = 49$; *dynamic_pref*: $\bar{x} = 401, \sigma = 103$) in either group, $t(12) = -1.69, p > 0.05$, the dynamic_pref group had a notably better dynamic *performance score* (*static_pref*: $\bar{x} = 491, \sigma = 67$; *dynamic_pref*: $\bar{x} = 373, \sigma = 79$), $t(12) = 3.09, p < 0.01$.

Reviewing the data showed that all other metrics being equal, this is likely due to the significantly lower *idle time* during the dynamic mode for those in the *dynamic_pref* group, (*static_pref*: $\bar{x} = 185, \sigma = 34$; *dynamic_pref*: $\bar{x} = 108, \sigma = 38$), $t(12) = 4.13, p < 0.01$. This shows that those who preferred the dynamic mode performed better mainly due to needing less thinking time with a reduction in TTC by 109s shortening the thinking time by 9% to 33% of TTC.

Unsurprisingly, this group also gave the dynamic mode a higher SUS score and a lower NASA-TLX workload rating. Although the SUS score was marginally not statistically significant (*static_pref*: $\bar{x} = 58, \sigma = 18$; *dynamic_pref*: $\bar{x} = 74, \sigma = 8$), $t(6) = -2.06, p > 0.05$, the NASA-TLX workload rating was very significant (*static_pref*: $\bar{x} = 57, \sigma = 6$; *dynamic_pref*: $\bar{x} = 39, \sigma = 11$), $t(12) = 4.00, p < 0.01$.

Ordering effects were also investigated for both NASA-TLX ratings and

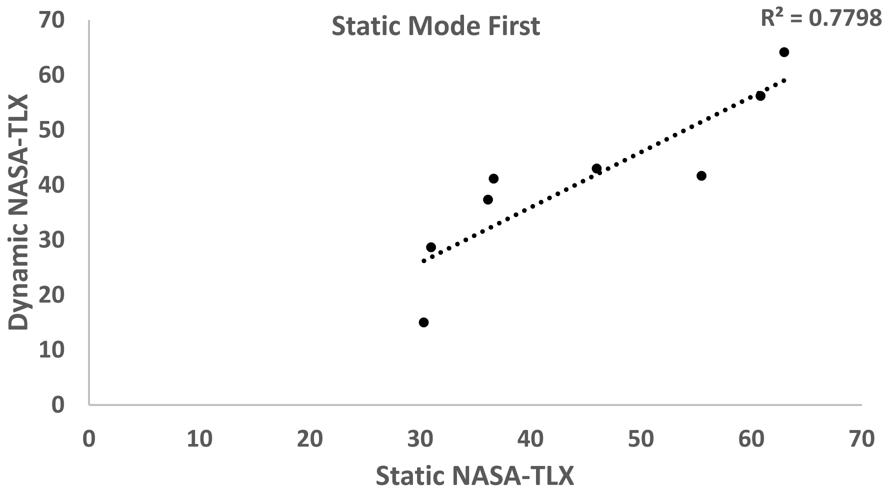
SUS scores to ascertain whether they were affected by the order in which participants used the camera modes. As before, participants were split into two groups based on which camera mode they used first and their NASA-TLX and SUS scores were analysed, table 6.8.

Order	**NASA-TLX**				**SUS**			
	Static		**Dynamic**		**Static**		**Dynamic**	
	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ
Static- Dynamic	45	13	**<u>41</u>**	**<u>15</u>**	68	13	**<u>70</u>**	**<u>11</u>**
Dynamic - Static	54	11	51	8	63	21	66	18

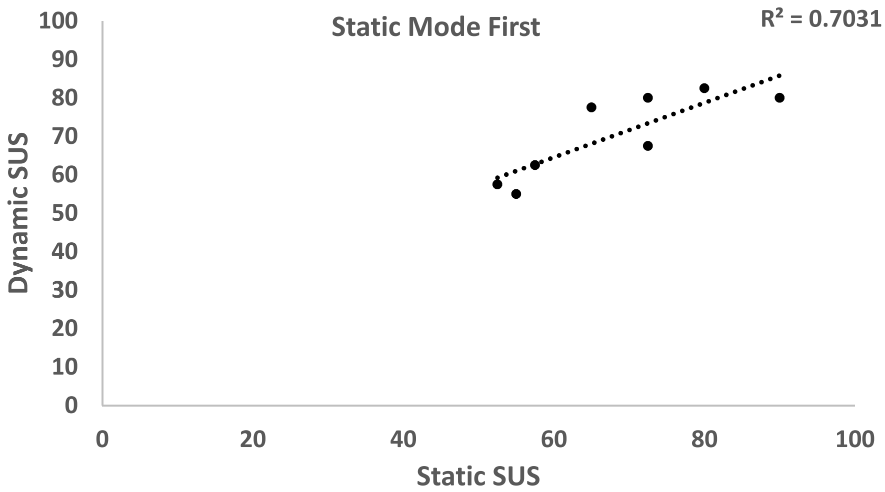
TABLE 6.8: Results of subjective metrics based on order of condition administered. Bold and underlined values are best for each metric.

As table 6.8 shows, when the dynamic mode was used second, participants experienced lower workload and found it user-friendly which also corresponded with the previous finding where 6 out of the 9 participants who preferred the dynamic mode over the static mode used the dynamic mode second. However, this was not statistically different to the other permutations nor were there any statistically significant differences between any of the permutations. Interestingly, participants using the dynamic mode first did not rate the workload and usability lower than the static mode even though they performed significantly worse on the dynamic mode as shown in fig. 6.14

Participants who used the static mode first had strongly correlated NASA-TLX ratings for the camera modes, fig. 6.17(a), $r(6) = 0.88, p = 0.05, power > 0.8, r^2 = 78\%$. Additionally, SUS scores were also strongly correlated (fig. 6.17(b)), $r(6) = 0.84, p = 0.05, power = 0.78, r^2 = 70\%$. This was not the case for participants using the dynamic mode first, NASA-TLX: $r(6) = -0.43, p > 0.05, r^2 = 18\%$; SUS: $r(6) = 0.32, p > 0.05, r^2 = 10\%$. This suggests the effect of using the static mode first was that participants did not find much of a difference between the two camera conditions. .



(a)



(b)

FIGURE 6.17: Relationship between ratings for subjective metrics for participants who used static mode first, (a): NASA-TLX, $r(6) = 0.88, p = 0.05, power > 0.8, r^2 = 78\%$; (b): SUS, $r(6) = 0.84, p = 0.05, power = 0.78, r^2 = 70\%$.

Metric		r	p	Power	r^2 /%
Performance	Static	-0.14	>0.05	<0.8	2
	Dynamic	-0.41	>0.05	<0.8	17
% Camera	Dynamic	0.26	>0.05	<0.8	7
NASA-TLX	Static	-0.15	>0.05	<0.8	2
	Dynamic	-0.32	>0.05	<0.8	10
SUS	Static	0.14	>0.05	<0.8	2
	Dynamic	0.14	>0.05	<0.8	2

TABLE 6.9: Correlation of metrics with gaming score, $n = 16, df = 14$

Demographic factors

In order to build a profile of the participants and allow for generalisations, demographic factors were also examined. In the prescreening questionnaire, data on age, sex, dominant hand and gaming experience was collected.

Based on the responses to the questionnaire, a total of 21 (19 males, 2 females) participants were recruited. Due to the homogeneity of some factors, further examination was not possible. All 16 participants who successfully completed the experiment were male. All were aged between 18-25 years except one who was 26-35 and all were right-handed except one who was left-handed.

There was a mix of responses for previous gaming experience and a *gaming score* statistic was added, $n = 16, \bar{x} = 39, \sigma = 22$. There were four questions on gaming experience covering playing frequency, experience in the last year with games, game controllers and 3D simulators/ virtual reality in terms of hours spent per week on a Likert scale.

As table 6.9 shows, there was little to no significant correlation between gaming score and performance score indicating gaming experience didn't have an effect on performance. Orbital camera usage was expected to be linked to gaming due to its gaming origins but this was not the case. Workload and usability likewise remained unaffected by previous experience with similar environments.

6.6.3 Discussion

Performance Metrics

The results alone do not paint a clear picture as to which camera mode was better. Although performance-wise the static mode was quicker to use giving better performance, there was high variation between the results so the difference wasn't significant. Apart from TTC, other metrics such as collisions, path length and arm resets did not reveal a notable change. The high variation and slow TTC in results emphasises the difficulty involved in controlling a robot arm and the need for extensive training. A 1 - 1.5 hour experiment does not allow sufficient time in taking novices with no prior experience in robotics control to competent remote operators. For future experiments, a recommendation would be for complex systems such as this to be evaluated through a pilot experiment to determine if a multi-session experiment structure would be more appropriate.

Subjective Metrics

The subjective metrics also showed similar results between the camera modes. As expected, participants found the task mentally challenging requiring high effort and causing stress. Despite there being no time limit set on the task, participants felt that they were under pressure to complete the task quickly yet did manage to complete it to a satisfactory level. This is similar to the previous experiment (section 4.2) where participants viewed the experiment as a personal challenge and ignored the serious connotations of a nuclear environment. This also has a negative effect on the performance metrics such as environment collisions and arm resets which are sacrificed in favour of achieving a faster TTC. For future experiments to be meaningful and allow for generalisations to the population, more work needs to be done to increase participant immersion through the use of scenarios and immersive simulations.

Although the consensus slightly favoured the dynamic mode, it is not readily apparent why this was since both performance and subjective metrics are so closely matched for both camera modes. Preference of a particular mode did not necessarily always entail that performance in that mode was better. This was supported by the lack of correlation in the results between performance score and subjective metrics. This further supports the importance of using the post-experiment interview technique, as the other metrics failed to capture this information.

No explanation has been found as yet for the apparent affinity to the dynamic mode of the participants who preferred it. Since all metrics remained the same other than idle time, it may be concluded that their ability to interpret the visual information from the orbital camera in a timely manner greatly improved their TTC and subsequently performance score.

Demographic factors

Results of the spatial test were unfortunately unusable due to unsolvable questions introducing noise in the answers. The particular computerised version of the spatial test [183] was not fit for use; a pen and paper test would be more suitable although more time consuming.

Questions in the prescreening questionnaire were also misunderstood. The recent rise of virtual reality and 3-D TVs has made the definition of the phrase "3-D video games" ambiguous. 3-D video games can now either mean games that utilise perspective projection to display 3-D graphics or games that make use of stereoscopic technologies to create a sense of depth. The aim was to distinguish 3-D video games from 2-D video games which do not require camera control or movement in 3-D space. It was not anticipated that participants would respond based on this latter definition when asking about "previous 3D video games experience".

As participant recruitment was screened based on gaming experience, the participant pool consisted of mainly young males. Female participants and some male participants who were recruited but did not manage to complete the experiment revealed that they mainly played mobile games (which are mostly not true 3-D) or 3-D games that fell outside of the view of the stereotypical games (i.e FPS, 3^{rd} person shooters, action-adventure) such as Minecraft or 2.5-D games such as MOBAs and strategy games.

These shortcomings in the questionnaire made the *gaming score* metric largely unreliable and as can be seen from the results, did not lead to any significant findings. This highlights the importance of removing ambiguity from subjective analysis to avoid wasted effort and resources.

Although steps were taken to reduce the effect of priming by not stating any restrictions on participants when advertising the experiment, some participants may have been inadvertently affected. As the prescreening questionnaire focused on questions about gaming experience, coupled with the experiment details, it may have influenced the participants to think gaming experience was related to the outcome of the experiment. To reduce the effect of priming, prescreening questionnaires should shuffle the questions so that they are not bunched together as well as releasing experiment details as and when needed to keep speculation of the experiment purpose to a minimum.

Observations

Controller effects

From observing the experiment, participants found it confusing to use an Xbox controller for 6-DOF movement. As evident by the long idle times ($>$

2$mins$) in both camera modes, a sizable amount of time was spent on thinking about mapping the movement of the robot to the controller. This was emphasised when the robot end-effector (EE) was oriented off-axis. Moreover, there were numerous errors where the participant intended the robot to go in a certain direction but ended up moving in the opposite direction due to a mismatch in the mapping between controller and robot. This was especially evident with the orientation controls.

Therefore, it can be concluded that the controller was a major confounding factor in participant performance, workload experienced and usability rating. In the author's opinion, based on the experiment data and observations, a joystick controller does not adequately replicate the 6-DOF of a robot arm. A controller with a closer kinematic or $1:1$ mapping to a robot arm such as an exoskeleton controller [185] or a 6-DOF haptic device such as the Haption Virtuoso [186] would be much better suited for this purpose. This would drastically reduce the thinking time required for the task by removing the need to mentally pre-map how a joystick would translate to robot movement. However, such controllers need to be empirically validated for their suitability to the nuclear environment as they are still in the early stage of development [185], [187].

Training

Although the training was improved in the revised experiment with gradual introduction of controls and constraints to force the utilisation of all controls, participants in many cases still did not manage to achieve the intended target of pressing two buttons in under a minute in the 15 mins of allocated practice time. This meant that some participants who did not reach a stable plateau in skill level, showed effects of learning during the live task.

In the interest of time and to prevent other factors such as boredom and fatigue from affecting the results, it was decided to either discontinue the experiment if the participant was unable to press a button in a reasonable time or continue to the actual experiment if the participant had reasonable aptitude for the task. This approach inevitably introduces variation in the results so it would be apt in future experiments to carefully balance sufficient training with resources such as time, compensation available and physiological factors such as engagement, fatigue and skill retention. If the need for training exceeds 30 mins, it may be better to schedule multiple one hour sessions at different times which would require management of skill retention.

Orbital camera

There was a split in participants when it came to the orbital camera. High variance in the results for the dynamic mode indicates that training was still insufficient. Initially, participants found the orbital camera difficult to use but preferred it with the structured training and subsequent practice. In the post-experiment interview, participants mentioned the panning and zoom capability of the orbital camera offered more flexibility when the eye-in-hand (EIH) camera view was blocked as well as making it easier to correct positioning errors. This was supported by the results which showed faster positioning of the robot in the dynamic mode. Reduced input times are beneficial as this reduces the risk of accidents due to the robot in a safety-critical environment such as a nuclear power plant. In the post-experiment interview, participants also mentioned that the camera compensated for the lack of depth perception through user-controlled saccades of the camera.

However, other participants were overwhelmed by having to control the robot as well as the camera with results showing prolonged camera usage leading to high completion times. Unfortunately, due to the structure of the task (the EIH view being blocked randomly), participants did not get enough chances to use the task space reduction technique to simplify the robot movement. Interestingly however, the results showed that some participants were efficient in their use of the orbital camera such that prolonged usage did not result in high completion times. Participants found it difficult to keep up with the dynamic control reference frame in the orbital camera view and neglected to use the camera even when the robot was blocking the line of sight to the button. From the video footage, this was exacerbated when at non-orthogonal viewing angles which increase the mental workload required to map controls from the controller to the robot. Some participants expressed that there was too much freedom in the camera controls and would prefer an option to toggle restriction on certain axes. Others, who were used to controlling cameras in video games, would prefer to have camera control on a single separate joystick in the right hand rather than split over two joysticks on a gamepad.

Static camera

As with the dynamic mode, there was also a split in preference over the static cameras. Some participants found having multiple camera views at orthogonal angles useful and preferred the static control of the robot they offered although it was slightly more stressful than the dynamic mode. It should be noted that the orthogonal camera views were very idealistic and in a real scenario are unlikely to be afforded to the remote operator and as noted previously, non-orthogonal angles are detrimental for mapping controls.

Other participants found it confusing to create a mental model of the multiple views with results showing participants spending considerable time idle whilst in the robot input mode. Predictably, there were issues with lack of depth and without any zoom functionality, there were frequent alignment issues when close to a button with the EIH camera view blocked. This is similar to the issues encountered with current systems [93], [152] which were overcome by the use of the orbital camera.

Input mode switching

The use of multiple input modes (tool, robot/ camera) during the task (standard in the industry) led to cognitive overload as all participants forgot to switch between modes at the appropriate times often causing collisions and loss of control of the robot. There were prolonged periods of time where the participants did not realise why the robot was not moving the way they had pictured it would since they were concentrating on the wrong camera view. This was despite there being a pop-up notification to confirm a switch as well as an overlay on the edge of the GUI to constantly show the current mode.

In both camera modes, a number of participants refused to switch away from tool mode control even when the EIH view was blocked. Tool mode was used to control the robot arm while "guesstimating" the input mapping from the other available view/s. The finding supports the previous literature that humans are most comfortable with an egocentric view [150], [164].

Keyhole effect

In addition to the aforementioned issues, during the course of the experiment, participants could be observed to regularly fixate on a single camera viewpoint neglecting vital information from other views. In many cases this led to collisions and alignment issues between the robot and buttons. An avenue of interest in the future would be to investigate operator gaze during remote perception using eye-tracking methods to further understand operator visual behaviour and develop an automated system to automate switching between camera views similarly to [188].

GUI

The use of sliders to display information on joint limits was ineffective as it was too mentally taxing to connect each slider to a dynamic joint configuration. Participants opted to ignore the sliders and only used them to identify if they had a reached a joint limit when the robot became unresponsive to input commands. Consequently, participants often hit the joint limits and became stuck resulting in poor performance. It may help to add colour indicators on joints to inform operators when joints are close to limits [78].

Further work is needed on information overlays to effectively disseminate contextual information in a timely manner to avoid critical incidents such as collisions and loss of control.

6.7 Conclusions

This experiment presented a systematic empirical analysis of the potential for camera systems in a telerobotic system using robot arms in contrast to the literature which focuses on developing novel systems while neglecting user feedback and human factors analysis. A sound experimental design was used to conduct a principled evaluation using statistical methods and human factors research to support the findings.

Even though this experiment failed to show a clear difference between the camera modes, the results show a potential for the orbital camera to improve telerobotic control. The EIH camera view is the preferred method of manipulation when using a robot arm. However, as shown in this experiment, in the event where the EIH camera is blocked by tools or objects at the end-effector, a camera with the flexibility to pan and zoom is invaluable for compensating the loss of close control and correcting errors in positioning. It also removes the workload of having to mentally piece together information from multiple views. In a safety-critical environment such as a nuclear power plant where there is a low tolerance for errors, dynamic camera views could be vital.

This has been an initial foray into developing an alternative camera system and requires further development to reach maximum effectiveness. Further improvement in training for baseline competence and a better kinematically similar controller are the key for achieving this goal. There is an initial learning curve to overcome but this can be facilitated with more training time. Other improvements to camera control can include a free mode switch for those confident and proficient enough for total freedom of control as well as semi-automated assistance such as snapping to orthogonal angles when in close proximity and automated movement when the camera is obstructed.

This experiment has also highlighted areas of improvement and further development for experiment design such as improvements to training methods, removing areas of ambiguity in subjective analysis, reducing priming effects as well as addressing participant mindsets. A part-way to getting participants to adhere to task requirements such as safety rather than viewing the experiment as a personal challenge or just taking part for the compensation is to split the compensation over best pre-test results, best performance and participation. This will incentivise participants to perform their best over all aspects of the experiment according to the criteria set by the experimenters.

Chapter 7

Visual Tracking

This chapter is a reformatted copy of the following publication: *Talha, M., & Stolkin, R. (2014). Particle filter tracking of camouflaged targets by adaptive fusion of thermal and visible spectra camera data. IEEE Sensors Journal, 14(1), 159-166.* [189]

In the previous chapter, a novel camera system for remote perception was proposed that could be deployed through a UAV or a secondary robot. Since current industry practice requires a team of several people simply to pan/tilt/zoom cameras to assist the robot operator, partially automating such camera control can significantly enhance the efficiency and capabilities of robot operators.

To enable a real robotic implementation of the orbital camera, with automatic gaze fixation and tracking of the end-effector (or other elements of a scene, such as an object being manipulated), it is likely that computer vision methods for automatic target tracking will need to be used. Abi-Farraj et al. [190] show how vision-based tracking can be used to enable an arm-mounted camera to automatically move to, and maintain, advantageous views of an object being manipulated by a tele-operated second arm and gripper. However, this work relied on a very simple box-shaped target object, which could be tracked using a relatively simple vision algorithm, based on fitting a wire-frame model to observed edges of the box. The wrist of the gripper-arm was also tracked, but again a simple vision algorithm was used to track a specially designed calibration target which was mounted to the robot's wrist. The tracking algorithm could not track objects of arbitrary shape and appearance.

In general, nuclear waste comprises and enormous diversity of objects and materials, and even the appearance of the robot's end-effector is variable, since typically the robot will use an automatic tool-changer to switch between a variety of different end-effector tooling (cameras, grippers, shears, scrapers etc). It is therefore desirable to develop an "anything tracker", i.e. a computer vision algorithm which can track objects of completely arbitrary appearance (including deformable targets such as rubber gloves, contaminated plastic suits, hoses etc. which are common in nuclear waste). Furthermore, the algorithm should be robust against objects which move against cluttered or "camouflaged" background scenes (e.g. a waste object which moves past a heap of other waste material).

Moreover, as the camera is dynamic with continuous movement, this can create instability [174], due to forces acting on the camera platform, in the visual feedback reducing its effectiveness.

To alleviate these problems, visual tracking can be used to assist the operator with gaze fixation through visual servoing to keep the focus point (e.g. robot end-effector) in the centre of the image. Additionally, rather than solely relying on traditional colour-based tracking, the visual tracking can be made more robust by employing bi-modal visual information. This combines a conventional visible colour spectrum camera with a secondary radiation source for intelligent tracking based on maximal information. Since the secondary radiation can be any radiation in the electromagnetic spectrum, this opens up the possibility of tracking areas of interest with gamma radiation in the nuclear environment.

In this chapter, the data fusion problem of combining image data from a colour camera and a deep infra-red thermal imaging camera is examined enabling target tracking that is more robust than would be possible with either imaging modality alone. Subsequently, the results of testing the algorithm on a number of videos to determine its effectiveness are presented.

7.1 Related Work

Until recently, there was comparatively little literature on the use of thermal cameras emerging from the robotics and computer vision communities, presumably because these devices were extremely expensive and did not feature in the standard equipment of a robotics or vision laboratory. In recent years, un-cooled thermal sensor chips have become more widely available at much more affordable prices, and with far higher resolution than previously possible, so that there seems to be a steadily increasing interest in their use to facilitate robotic vision tasks, especially for human-robot interaction, e.g. [191], [192].

A large part of the computer vision literature which discusses the fusion of thermal and visible data, is focused on face recognition, e.g. [193], [194], since thermal images solve some of the difficulties of this problem by providing relatively illumination invariant face images. Some approaches to surveillance-type tracking of moving targets include [195], and [196]. In [193], a human pedestrian is initially segmented from high resolution visible spectrum images using background subtraction. The segmented image is then modified by combining it with opinions from pixels of a low resolution thermal image, expressing the fusion problem probabilistically as an Extended-Markov Random Field, [195], [197]. [198] formulates a statistical background model for each pixel, similar to the well known surveillance-tracking formulation of [199], but including thermal parameters and colour parameters in a single, high dimensional distribution for each pixel. Both these methods are based on background subtraction with stationary cameras. In contrast, the present chapter will focus on tracking objects viewed from an arbitrarily moving camera.

Treptow et al., [191], are interested in enabling a mobile robot to detect, track and interact with humans. They use a thermal camera to detect and localize a human (with an elliptical contour model of the head and shoulders), and then direct a visible spectrum pan-tilt camera to track the human's face. This work is valid and useful, however it is not really a fusion of thermal and visible information, instead being a hand-over from one modality to the other. In contrast, O'Conaire et al., [200] propose what is really a "product of experts" type technique, [201], with one expert for thermal data and other experts for visible spectrum data. They use separate individual trackers for thermal brightness in an infra red camera, and luminosity, edge orientation and other visible light features from corresponding pixels in a visible spectrum camera. For any candidate target location in an image, each of these trackers computes a similarity measure, between that tracker's target model and the pixels of the candidate image region. The opinions of each of these trackers are then combined as a product of these "beliefs". Cielniak et al. [192] use a thermal camera to detect humans. A bounding contour from the thermal detection is then used to learn a colour model from relevant pixels in a visible spectrum camera. Probabilities of a particular target can then be assessed as a product of the beliefs of separate colour-based and thermal-based detection schemes.

Many of the above methods perform some kind of detection or tracking separately in each modality, and then combine the opinions of these trackers through some kind of voting scheme or product of experts. However, predominantly such fusion is not actually informed by information about the relative discriminating power of each modality. This can cause difficulties – consider the case where one modality indicates a strong detection and one indicates a weak detection. A conventional combination method may result in a medium strength combined opinion – but if the target is highly camouflaged in one or other of the modalities, then this result may be incorrect. Depending on which modality is "wrong", that modality's opinion should be largely eliminated, and output a combined opinion that is either truly strong or truly weak. This chapter presents an approach to achieve this by means of rapidly and continuously re-learning local background models in each imaging modality, and using these background models to assess the current discriminating power of each modality.

This chapter makes several novel contributions. Firstly, the important colour-based particle filter of [196] is extended to make additional use of thermal imaging data. Secondly, a method of optimally fusing data from both thermal and visible spectra cameras when evaluating each particle is provided. Thirdly, the data fusion is shown to be adaptive, weighting the data-fusion process in favour of whichever imaging modality happens to be most discriminating for each particle in each video frame. This adaptation is enabled by a process of continuously relearning local background models for each particle in each image frame of each imaging modality. When backgrounds are similar to the foreground being tracked in one modality, the data fusion process weights in favour of information from the other imaging modality if it is more discriminatory.

7.2 Tracking Algorithm

This section describes the algorithm used for tracking. A particle filter combining colour and thermal information is used to keep track of multiple hypotheses with weighting used to adapt the combined information to scene changes through continuous background re-learning.

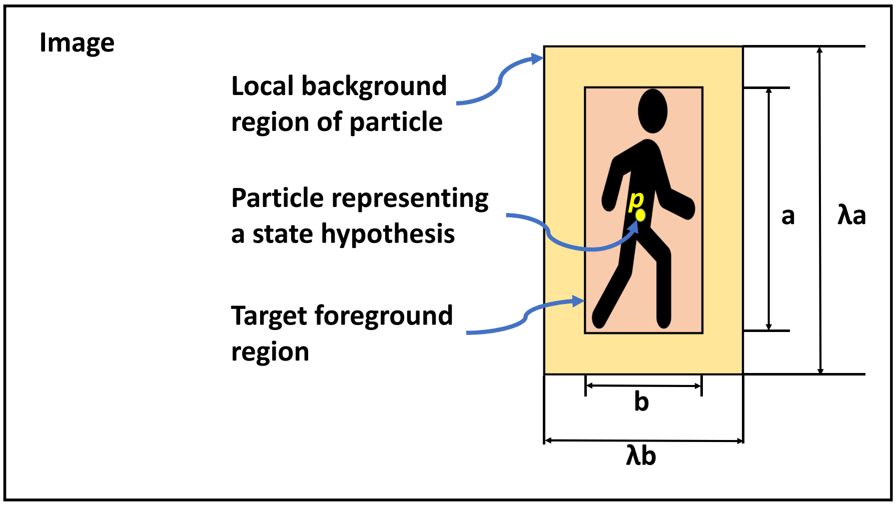


FIGURE 7.1: Each particle p represents a target state hypothesis indicating a candidate position, foreground region and local background region inclusive of the foreground region.

7.2.1 Particle Filter

At any time, t, the state s_t of the tracked target is represented by a distribution, approximated by a weighted set of I particles, $p_0, ...p_i, ...p_I$, with weights $w_0, ...w_i, ...w_I$ which are normalized so that:

$$\sum_{i=1}^{I} w_i = 1 \tag{7.1}$$

Each particle encodes a hypothesis target position, x_i, as well as size and shape information. For simplicity a tracked person is represented as a vertical rectangle of height and width parameters a and b respectively – so each particle, p_i can maintain it's own size and shape hypothesis, a_i and b_i, as well as its position hypothesis, x_i, see fig. 7.1.

In the usual manner, at each time step, these particles are re-sampled according to their weights, propagated according to a motion model, and then re-weighted according to new sensory information (a new pair of thermal and colour frames) at the next time step. A simple zeroth order diffusion motion model was used (although more complex predictive motion models may be used [196], [202], [203]). A random walk is applied to the size and shape parameters of each particle, as well as to its position, so that a cloud of particles can maintain multiple hypotheses of the target's size and shape at any time, i.e. shape and size are handled efficiently as part of state propagation (similar to [196]), without the need for extra computations as in [204].

Nummiaro et al., [196], first showed how to evaluate particle weights, w_i, using simple colour histogram models of tracked targets. This chapter demonstrates how to additionally make use of thermal information from an IR camera, as well as colour information from a conventional camera. Additionally, the data fusion process is continuously adapted, so that the process for evaluating particle weights takes into account that one or other of the imaging modalities may be more or less discriminatory at different times.

7.2.2 Particle Weighting

A visible spectrum colour camera is used, for which each pixel yields a colour, to represent each frame by a 3D colour vector RGB. To incorporate a degree of robustness against changes in illumination changes, normalisation is used to give a 2D normalised colour vector $\bar{c} = \{r, g\}$ where

$$r = \frac{R}{R+G+B} \qquad g = \frac{G}{R+G+B} \tag{7.2}$$

Since $r + g + b$ must sum to unity, all of the colour information for a pixel is now contained in just two variables, $\bar{c} = \{r, g\}$, and a convenient target model for the colour modality is a 2D colour histogram, H_{t_Col}. Similarly, 1D pixel intensities (where intensity varies monotonically with temperature), derived from the thermal imaging camera, yield a thermal target model in the form of a 1D histogram, H_{t_Therm}.

A well known method, first proposed by Ennesser [205], and later popularized by [204] and [196], for evaluating how well the local region of image around a particle matches a target model, is to compare a target histogram against a histogram for the local region using the Bhattacharyya coefficient, [206], which provides a simple measure of similarity between two histograms. For two histograms, p and q, each with m bins:

$$p = \{p^{(u)}\}_{u=1..m} \qquad \text{and} \qquad q = \{q^{(u)}\}_{u=1..m} \tag{7.3}$$

the Bhattacharyya coefficient is defined as:

$$B[p,q] = \sum_{u=1}^{m} \sqrt{p^{(u)} q^{(u)}} \tag{7.4}$$

If both histograms are identical, $B = 1$, and maximally dissimilar histograms yield $B = 0$.

Nummiaro et al., [196], use a Bhattacharyya coefficient, denoted by $B^{i}_{tf\_Col}$, between the target colour histogram, H_{t_Col}, and the local foreground colour histogram, $H^{i}_{f\_Col}$, to assign a fitness weight w_i to particle p_i. Additionally, a second measure of particle fitness can also be generated, derived from the Bhattacharyya coefficient, $B^{i}_{tf\_Therm}$, between the target thermal histogram, H_{t_Therm}, and the local foreground thermal histogram, $H^{i}_{f\_Therm}$.

More robust tracking should result from weighting particles using an appropriate combination of these coefficients for each modality, than might be obtained by assigning weights using one or other alone. However, it is not obvious how the coefficients should be combined. For example, simply taking an arithmetic mean of the thermal and colour Bhattacharyya coefficients, and then substituting this instead of the colour coefficient in the Nummiaro method, then at times when one imaging modality is more discriminatory than the other, performance would be worse than with the good modality alone. Similarly, simply multiplying the coefficients as a product of experts

[201], does not necessarily improve robustness – a modality that is performing poorly may spoil a meaningful coefficient from a modality that is performing well.

In general, some optimally weighted combination of the coefficients will be best. Therefore, particles are weighted using:

$$B^i_{tf\_Fused} = \alpha B^i_{tf\_Col} + (1-\alpha) B^i_{tf\_Therm} \tag{7.5}$$

where is a weighting factor (varying between 0 and 1) which is continually relearned at each frame, to make the data fusion process adaptive to changing scene conditions. The method of calculating α by using continuous model re-learning is explained in the next section.

Note, with appropriate normalization it is possible to assign $B^i_{tf\_Fused}$ directly as the particle weight w_i, and this results in a reasonably effective tracker. However, to make particle weights handle rather more like true probabilities, and arguably providing more robustness, weights are evaluated as:

$$w_i = \frac{1}{\sqrt{2\pi\sigma}} \exp\{-\frac{(1 - B^i_{tf\_Fused})}{2\sigma^2}\} \tag{7.6}$$

7.2.3 Adaptive Data Fusion

This section discusses how best to re-compute the parameter α of equation 7.5. Intuitively, if the target is very similar to its background in the thermal modality (e.g. a hot person walking past a hot background), then the thermal Bhattacharyya coefficient should not be trusted by weighting in favour of colour information, using a high value for α. Conversely, when the target is heavily camouflaged in the colour domain (e.g. a person wearing red and walking past a red background), the colour Bhattacharyya coefficient should not be trusted by weighting in favour of thermal information, using a low value for α.

Since the above discussion hinges on characterizing the background conditions around candidate target locations, some new terms need to be introduced for handling these. Fig. 7.1 shows how, as well as considering a candidate foreground region around each particle, a local background region can also be defined, comprising a border strip that surrounds the foreground region, the border region being maintained at a size, some multiple λ times the size of the foreground region (see fig. 7.1). At each new frame, thermal and colour histograms can be learned for this background region surrounding each of the particles. For the i^th particle, these are denoted by $H^i_{b\_Therm}$ and $H^i_{b\_Col}$ respectively. Bhattacharyya's coefficient can now be used to evaluate how similar the target model is to the local background in each imaging modality. The Bhattacharyya coefficients between target and local background are denoted as $B^i_{tb\_Col}$ for the colour modality and $B^i_{tb\_Therm}$ for the thermal modality. The larger these coefficients, the less confidence there is in the discriminating power of each respective imaging modality. The data fusion adaptation term α can now be meaningfully re-computed for every particle in every new video frame as:

$$\alpha = \frac{B^i_{tb\_Therm}}{B^i_{tb\_Therm} + B^i_{tb\_Col}} \tag{7.7}$$

Substituting into equation 5 gives:

$$B^i_{tf\_Fused} = \frac{B^i_{tb\_Therm} B^i_{tf\_Col} + B^i_{tb\_Col} B^i_{tf\_Therm}}{B^i_{tb\_Therm} + B^i_{tb\_Col}} \tag{7.8}$$

When $B^i_{tb\_Therm}$ is large and $B^i_{tb\_Col}$ is small, the target is very similar to the background (i.e. a camouflage situation) in the thermal modality – hence equation 7 produces a large α which weights heavily in favour of colour information, to the exclusion of thermal information, in equation 5. The converse results for a large $B^i_{tb\_Col}$ and a small $B^i_{tb\_Therm}$.

The algorithm can now instantly relearn local background conditions for every particle, at every new video frame, and instantly adapt to pay more attention to data from whichever imaging modality is most discriminating.

7.3 Experiment

This section examines the performance of the algorithm on a number of videos.

7.3.1 Experiment Setup

The adaptive thermo-visual particle filter tracker was implemented using a FLIR Photon 320 thermal imaging camera with 320×240 pixel resolution, mounted side by side with a standard, research quality, colour firewire camera.

For initial proof of principle work, the pixel correspondence problem between the two cameras was solved with an offline calibration, which assumes a simple affine mapping from one image plane to the other, similar to the approach taken in [200] r, g data respectively. Pixel intensity ranges of 0-255 are divided up into 10 bins and 10×10 bins respectively for the two histograms.

7.3.2 Results

7.3.3 Adaptive Data Fusion

This section is designed to provide a clear and simple demonstration of how the tracking algorithm automatically adapts to changing background conditions in each imaging modality.

Fig. 7.2 shows an example of tracking a target which may be either highly camouflaged or highly discriminated in each modality at different times. The target is a rectangular piece of red card which has been heated up using a hair

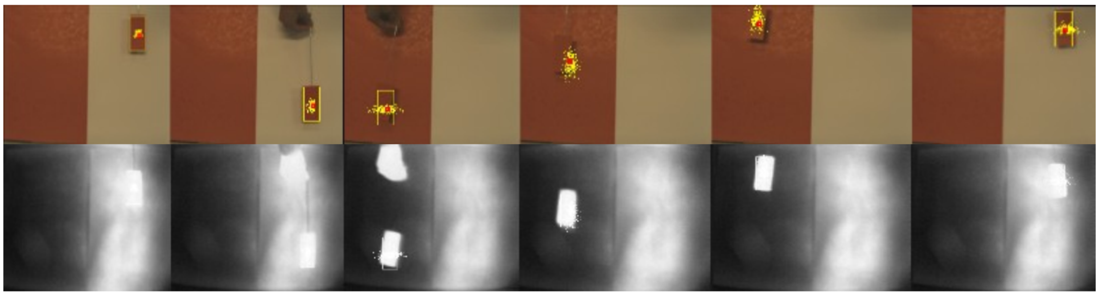


FIGURE 7.2: A simple target (card rectangle heated with hair dryer) is both red and hot. It is tracked while moving against a cold red background (where thermal camera is most discriminating) and a hot white background (where the colour camera is most discriminating). Red square denotes the maximum likelihood particle with foreground region in the yellow rectangle, and particle centres are denoted by yellow dots.

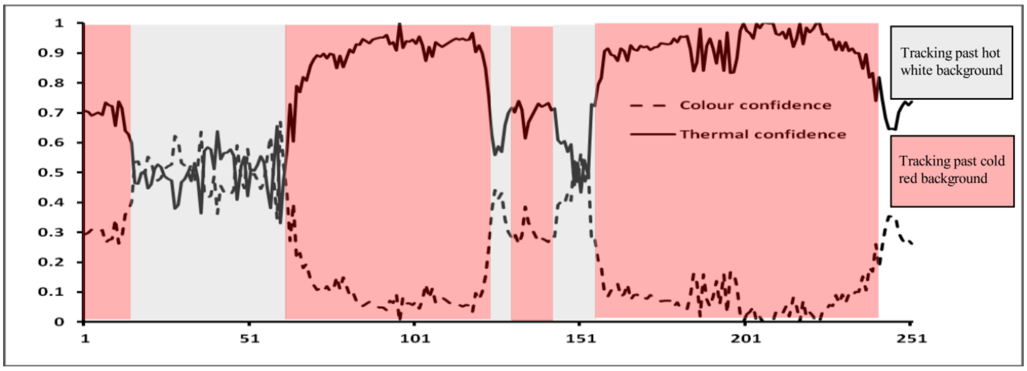


FIGURE 7.3: This graph shows the variation of the continuously relearned confidence values (α in the colour modality and 1- α in thermal modality), as the target of fig. 7.2 (hot red) moves past cold red backgrounds and hot white backgrounds. Note that there is an apparent steady offset in confidence between the two modalities (thermal always greater than colour) – this is because, in this particular example, it so happens that the hot target camouflaged against hot background is better discriminated in the thermal camera than the red target camouflaged against red background is discriminated in the colour camera.

dryer. It is tracked across changing background scenery, comprising first a red background which is cold (where the colour modality is camouflaged and the thermal modality is discriminating), and later a white background which is hot (where the thermal modality is camouflaged but the colour modality is discriminating).

Fig. 7.3 shows how the terms α (confidence in colour modality) and $(1-\alpha)$ (confidence in thermal modality) automatically vary throughout the tracking sequence of figure 7.2. It can be seen that, when a red target moves past red background, the algorithm automatically learns to lower confidence in the colour modality and increase confidence in the thermal modality. Conversely, when the hot target is viewed against a hot background, the algorithm adapts by reducing confidence in the thermal modality, while raising confidence in the colour modality. A curious feature of this experiment is an apparent steady offset in confidence between the two modalities – the simple explanation is that, in this particular example, the hot target camouflaged against hot background appears consistently better discriminated in the thermal camera than the red target camouflaged against red background in the colour camera.

7.3.4 Tracking Under Severe Camouflage

FIGURE 7.4: Extreme camouflage in the colour modality – a person wearing red, walks past a red background, while being occluded by a red chair. The algorithm automatically lowers confidence in the colour modality and raises confidence in the thermal modality, to successfully track.

FIGURE 7.5: Extreme camouflage in the thermal modality – a hot target person, walks past other distracting hot people, and is completely occluded by another hot person who walks between the target and the camera. The algorithm automatically lowers confidence in the thermal modality and raises confidence in the colour modality, to successfully continue tracking throughout the sequence.

FIGURE 7.6: An extremely challenging test sequence where a person is tracked while moving against backgrounds for which he is strongly camouflaged in both the thermal and colour modalities. The person is hot and moves past background distractors including other hot people. The person also wears red clothing but in the middle of the sequence he sits down on a red chair against a red background and covers his legs with a red coat. The algorithm adaptively weights in favour of whichever imaging modality is most discriminating in each phase of the tracking – it successfully tracks throughout the sequence.

FIGURE 7.7: Tracking with servoed camera motion, past extreme occlusion. In the middle of this sequence (4th frame) the target person crouches down behind a wooden fence, where he is almost completely occluded. The algorithm tracks correctly, with the exception of overly fixating on the target's legs rather than torso – perhaps because this person's white shirt is almost completely camouflaged (especially in normalised colour space) against the white background wall. This sequence was tracked and recorded in real-time, while using the algorithm to visually servo a pan-tilt motorised camera rig.

The algorithm was tested on a variety of video sequences, of many hundreds of frames each, with and without visual servoing. Some salient and instructive examples are shown here (figures 7.4-7.7), which are designed to illustrate how the data fusion strategy has successfully enabled:

- Automatic increase in confidence in one feature modality, to assist when the other modality is unable to discriminate.
- Rapid and automatic adaptation, by continuously adjusting the confidence in each modality, to cope with video sequences where both modalities are unable to discriminate at various moments.

Some of the presented test sequences may, at first glance, look somewhat artificial and contrived. However, it should be noted that they are indeed contrived - so as to be far more challenging (in specific and informative ways) than, e.g., randomly recorded street scenes.

Figure 7.4 shows a video sequence where a person wearing red clothes, walks past a red background while being partially occluded at times by a red chair. This sequence is designed to show how the algorithm successfully adapts in favour of thermal information, when colour information is unable to discriminate due to severe colour camouflage.

Figure 7.5 shows a video sequence where a hot person walks past other hot people, while at one time being completely occluded by another hot person who passes between the target person and the cameras. This sequence is designed to show how the algorithm successfully adapts in favour of colour information, when thermal information is unable to discriminate due to severe thermal camouflage.

Note that, at first glance, the tracking problems of figures 7.4 and 7.5 might appear trivial – "of course the colour camouflage sequence will be successfully tracked, because the algorithm has thermal information too" and conversely for the thermal camouflage sequence. However, in both these cases, tracking is only successful because the algorithm has re-learned its background models, used them to detect which modality is best, and weighted the data fusion process in favour of this modality – that procedure is non-obvious, and (as discussed in section II. B) a naive data fusion approach (such as averaging, producting or voting) would often fail in such cases.

Figure 7.6 shows a more challenging sequence, in which the target often becomes camouflaged in both imaging modalities. In this case a hot person, wearing red, moves past red background regions while also moving past other distracting hot people. In the middle of the sequence, the red target person sits down on a red chair while holding a red coat over his legs, before getting up and walking past other thermally distracting people again. This sequence is designed to show how the algorithm can rapidly adapt, from one frame to the next, performing smart data fusion by extracting whichever combination of modalities is best, under rapidly changing circumstances.

Figure 7.7 shows a scene which is more conventional in appearance, but which features a period of almost complete occlusion. The target person walks across the scene, crouches down behind a barrier where he is almost completely concealed, before standing up and walking back across the scene again.

7.4 Conclusions

This chapter presented a method for converting a fundamental tracking technique (particle filtering using colour histograms) to perform continuously adaptive data fusion with colour and deep-IR thermal cameras.

This new algorithm exhibits several useful properties:

- Local background models are rapidly re-learned at every frame, for every particle, and for both imaging modalities.
- These background models are used to continually re-assess how much confidence should be assumed for each modality, when computing particle weights.
- If the target becomes camouflaged in one or other modality, this is automatically detected, and the data fusion process is weighted in favour of data from the other modality.
- Where the target may be partially camouflaged in both modalities, a best compromise blending of the data from both modalities is found, by adaptively re-computing confidences in each modality.
- The adaptation is extremely fast, because background models are completely re-learned from scratch for every image. This enables the technique to track difficult sequences, where the background scene rapidly changes.
- Rapidly changing backgrounds often occur in situations where the camera itself is moving. Hence this rapid adaptation technique is especially

useful for visual servoing of a pan-tilt camera rig, or for tracking from cameras mounted on a moving robot.

- Although the technique has here been demonstrated for merely fusing simple thermal and colour pixel intensities, it can readily be extended to fusing an arbitrary number of different pixel features (e.g. edginess or texture plus colour etc.) from one or arbitrarily many imaging modalities. Future work will experiment with using this method to incorporate and fuse various additional kinds of features to improve tracking robustness.

The tracking algorithm can track objects of arbitrary shape and appearance. The tracker can be initialised by the operator simply designating the object to track with a bounding box in the first frame of an image sequence, from which the tracking algorithm learns a statistical model of the target's appearance. The tracker copes with deformable target objects, and is also highly robust to objects which move against cluttered or camouflage backgrounds. This is achieved by exploiting an adaptive background model, which continually relearns the background at each successive frame. Once the background model has been learned, the tracker then weights its attention on parts of the target (foreground) model that are most distinct from the current background scene.

In nuclear environments, it is possible that radiation imaging can be used in addition to conventional colour cameras. This might help discriminate e.g. a highly radioactive target object, against a background that had low radioactivity but shared common colours with the target. Radiation imaging is still however a relatively immature technology, especially real-time imaging and it is difficult to perform experiments with radioactive materials.

Nevertheless, many radioactive waste materials are also physically warm, information which can be captured using a deep infra-red (IR) thermal imaging camera. Therefore, the data from the thermal IR camera is adaptively combined with data from a conventional colour camera. In principle, this tracking algorithm can be easily applied to combining conventional cameras with any kind of additional imaging modality (e.g. radiation imaging, multi-spectral cameras, acoustic imaging devices for underwater work etc.) helping to reduce the workload of the operator through automated gaze fixation of user-selected areas and keeping track of areas with radioactive radiation.

Chapter 8

Conclusions and Future Work

The aim of this thesis has been to investigate human factors issues in using telerobotics for the decommissioning of nuclear facilities. Current decommissioning practices are far too slow requiring human personnel to carry out work in air-fed suits which is exhausting, risks exposure to hazards and generates secondary waste. This presents a unique opportunity for introducing telerobotics to make decommissioning "faster, cheaper and safer". Nuclear facilities present a significant challenge to the integration of telerobotics not only because of the need for caution in handling radioactive materials to avoid accidents but also due to the mistrust in the use of robotic technology in an industry that is highly safety-oriented.

In contrast, technological advancements in areas such as computer processors, AI, wireless communication and HMI capabilities etc. have allowed telerobotics to be increasingly used in other hazardous sectors such as space, underwater and energy. This presents an opportune moment for penetrating the nuclear industry with telerobotics as research from other sectors can be bootstrapped to quickly develop sophisticated solutions to problems in the nuclear domain. Although the lure of novel sophisticated telerobotics technology is tempting and many systems have been proposed as the solution in the literature, their application to industry remains elusive. Either the system has been engineered by expert researchers in isolation without regard for the end-user and/ or tasks requirements resulting in a complex system only usable by the researchers themselves (e.g. DARPA challenge robots), or the system has been left untested at the prototype stage.

This thesis demonstrates how methods from fields such as human factors and pscyhology can be applied through empirical studies from the early stage of the design process to identify and resolve problems that occur during HRI in an attempt towards achieving the goal of developing a telerobotic system that emphasises synergy between man and machine.

A testbed was developed to replicate the current industry-baseline technology. This was evaluated using a principled experiment design to determine the capabilities of current technology in use in the industry as well as its limitations. The results illustrated the limited effectiveness of direct teleoperation systems. Primitive HMIs with lack of depth perception and kinematic feedback make it cognitively challenging to carry out telemanipulation tasks efficiently and require extensive training. It also revealed how subjective metrics for determining usability can be difficult to interpret. The results showed that aptitude for telemanipulation varies significantly between individuals but standardised tests may help to select prospective operators. The study also revealed gaps in human-subject experiment methodology related to the appropriate training of novice participants to achieve minimum competence and reduce the effect of confounding factors such as learning effects as well as potential useful metrics to be measured in future experiments.

As a first improvement to direct teleoperation, a proof-of-concept semi-autonomous telerobotic system with supervisory control was developed which showed how automating simple navigation tasks can help to drastically improve system efficiency as well as freeing the operator to carry out other tasks.

Another human factors aspect that can now be improved with modern technology is visual awareness of the remote environment. Removing the human operator from direct contact with the environment severely restricts their perception abilities by restricting their field of view to the cameras in use. To this end, an orbital camera was developed that could be dynamically controlled by the operator to view the remote scene from an arbitrary viewpoint and to control the robot based on the camera reference frame.

Although the comparison with traditional camera views did not show significant improvement gains due to confounding factors, the orbital camera showed potential. It was the preferred method of viewing for the majority of participants who found it improved awareness of the scene and was invaluable for precise control of the robot.

A major confounding factor was the gamepad controller. The experiment identified the urgent need for anthropomorphic controllers in telemanipulation for kinematic feedback. Participants found it mentally challenging to map the input from the gamepad controller to robot movement which often resulted in errors and movement in unintended directions. Another factor that contributed to the system performance was the use of multiple input modes. Participants frequently forgot to switch between input modes when switching between different camera views (e.g. participants forgot to switch from tool mode to camera mode when shifting control from the EIH camera view to the orbital camera view) resulting in loss of control. This work has been instrumental in illustrating such strongly interlinked challenges of perception and control of high degree of freedom teleoperators. Additionally, the study also highlighted further areas of improvement for experiment design such as reviewing questionnaires for ambiguous phrases, reducing priming effects and addressing participant mindset to improve engagement.

Finally, a novel visual tracking method was presented that uses a particle filter to track targets in multiple modalities while continuously adapting to the target background in each modality. It is made robust against camouflage in one modality by adjusting the weighting in favour of the alternative modality to successfully keep track of the target. These unique features are verified through experiments. The flexibility to fuse multiple different features from one or multiple modalities allows for robust tracking and can be readily used for gaze fixation with the previously mentioned orbital camera or for keeping track of radiation contaminated areas using thermal information or other radiation sensors.

8.1 Thesis Contributions

This thesis provides the following contributions to knowledge:

- **Development of an experimental methodology for evaluating the performance of teleoperated systems *together with their human operators* for remote manipulation tasks**- A systematic and rigorous experimental design taking into account human factors, controlled experimental conditions, confounding variables and standardised training, to obtain repeatable results on which statistical analysis can lead to inferences being made. Previous robotics literature rarely incorporates experimental design for evaluation of the performance of the combined human-robot system. Where such experiments have been attempted, they tend to be somewhat adhoc, often with insufficient details provided to enable repetition and comparison by other researchers. In particular, there has often been inadequate control for a variety of confounding factors, such as gaming experience and baseline competencies of the human participants.

- **Development of a VR testbed environment for *exactly repeatable* evaluations of interface technologies/techniques for use in telerobotics**- A VR simulation testbed as described in Ch.3 was developed, for empirically evaluating system performance, under highly repeatable conditions, to determine the effects of new technologies and/or techniques.

The testbed is capable of: realistically simulating a variety of remote manipulation tasks, with a variety of input devices and slave robots; simulating real-time CCTV video views of the remote manipulation tasks to provide realistic situational awareness (SA) to the robot operator; automatically recording and computing a number of performance metrics during task execution; recording a time series of all human control inputs and robot movements during experimental trials; as well as taking screen-shots and recording video footage, from multiple camera views, allowing comprehensive analysis and evaluation. By recording the trajectories of the robot and other entities, as well as all operator inputs, these can be combined with a playback feature. As well as assisting with experimental evaluations, this could potentially also be used as a valuable operator training tool, and assist with operator performance reviews.

- **Development of a novel approach to situational awareness and robot control, comprising an orbital camera for dynamic views, and control of the robot with respect to the coordinate frame of such dynamic views**- A thorough search of the literature has shown that an operator-controlled orbital camera concept has not been proposed in the literature before. The camera allows operators to obtain dynamic views of the remote environment which significantly help to overcome the severe depth perception problems of conventional camera views. In addition, a camera-referenced control system has been developed, in which the robot's control signals are transformed into the coordinate frame of the orbital camera. This enables intuitive control of robot motions with respect to the current viewing direction.

- **Evaluation of the performance of human operators when using dynamic vs. conventional camera views**- Principled and repeatable performance evaluation experiments were undertaken in Ch. 6, to measure the relative benefits of the proposed orbital camera, versus performance using conventional static camera views.

- **Adaptive visual target tracking algorithm which adaptively combines data from multiple imaging modalities**- A visual tracking method that fuses data from multiple modalities with optimal weighting for tracking targets that move with respect to the camera. The algorithm has been demonstrated with deep infra-red and visible spectra cameras, but could in principle be extended to other kinds of sensing modalities. Such tracking can be used to enable a moving platform, such a UAV drone, to keep its camera fixated on a region of interest in the scene. It could therefore could be used to help implement the orbital camera system proposed in this thesis.

- **Data sets including videos and pictures**- Data set covering performance metrics, subjective metrics and questionnaire responses as well as video footage of experiments. In addition, data capturing robot poses and camera poses for the orbital camera which can be utilised for analysing input command patterns from users.

8.2 Future Work

This thesis has presented methods for human factors testing of telerobotic systems. The logical step forward is to use the evaluation methods developed in this thesis in conjunction with the iterative design methodology to incrementally develop a telerobotic system taking into consideration input from end users. In this thesis, the focus was on establishing cause-and-effect relationships so the users consisted of the population of the University of Birmingham and the experiments took place in a laboratory. However, laboratory research is artificial and does not always resemble real-world performance. For development at a higher technology readiness level, it would be prudent to analyse and profile current personnel and tasks involved in the decommissioning of nuclear facilities to better inform the design process. Moreover, using the same personnel to conduct human factors testing in a realistic mock-up facility during the development stage would offer the chance for specialised feedback from domain experts that would not be available from the general public.

Additionally, the human factors evaluation can be further expanded to include methods such as SA assessment to measure operator perception, analysis of factors that induce stress, cognitive task analysis to assess the cognitive demands imposed on operators and human error prediction to identify potential areas for operator error and implement recovery procedures.

As well as incrementally developing the telerobotic system, the experiment design methodology should also be continually developed in conjunction with the system. Over the course of the experiments in this thesis, many confounding factors were revealed and the methodology revised and refined illustrating the complexity of the overall task. Of special importance is the need to balance training with participants' psychological state of mind. Observations of participants in the experiments showed frustration and boredom at times; particularly during training when participants could not grasp the robot controls. Training novice users for telemanipulation entails prolonged periods of practice; especially if primitive interface methods such as game controllers or joysticks are used. Further research is needed to determine whether training sessions need to be split over multiple sessions to avoid the effects of boredom and fatigue creeping into the results; a factor in the decision to exclude SA measurements from the analysis in addition to the tasks being too simple to allow for comprehensive SA analysis. A sound experiment methodology will lead to validity in the results allowing inferences to be made leading to generalisations beyond the scope of the specific experiment.

Furthermore, the novel technologies proposed in this thesis need more development to prove their viability for real-world scenarios. The current baseline teleoperated systems are clearly not practical and require more intuitive interfaces as well as some level of autonomy. The semi-autonomous system demonstrated in Ch.5 requires pre-built 3-D CAD models of objects to grasp. However, many legacy nuclear facilities do not have inventory logs to show what is currently stored in them. Further work is needed to extend this approach using machine learning and computer vision techniques to automatically extract object models in a scene under human supervision. Although the majority of participants expressed preference for the orbital camera, it could still benefit from some operator-assistance capabilities to reduce operator workload and frustration. Examples of such capabilities include snapping to orthogonal views when in close proximity and automated movement when obstructed by environment or robot. Moreover, some participants expressed preference to toggle between fixation of the camera on the robot end-effector and free mode where the camera could be moved around without being tethered to the robot. This would allow complete freedom to the operator to control the camera view as well as a more complete awareness of the remote environment.

Aside from this, for enhanced telepresence, the iterative development of the telerobotic system through human-subject testing should aim to incorporate as much state-of-the-art multi-modal feedback methods from the literature as appropriate, considering user and task requirements. Examples of these include bilateral feedback using haptic devices for sensing the remote environment through touch, auditory feedback using microphones and visual feedback through head-mounted displays connected to cameras mounted on pan/tilt units for immersion.

Bibliography

[1] *Nuclear provision: The cost of cleaning up britain's historic nuclear sites*, 2017. [Online]. Available: https://www.gov.uk/government/publications/nuclear-provision-explaining-the-cost-of-cleaning-up-britains-nuclear-legacy/nuclear-provision-explaining-the-cost-of-cleaning-up-britains-nuclear-legacy.

[2] NDA, "NDA Strategy", Tech. Rep. April, 2011, p. 5. [Online]. Available: http://www.nda.gov.uk/documents/upload/NDA-Strategy-Effective-from-April-2011-full-colour-version.pdf.

[3] NDA, "NDA Strategy- Effective from April 2016", Tech. Rep., 2016. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/518669/Nuclear_Decommissioning_Authority_Strategy_effective_from_April_2016.pdf.

[4] NDA, *Sellafield – the most complex nuclear site in Europe*, 2009. [Online]. Available: http://www.nda.gov.uk/stakeholders/newsletter/sellafield.cfm (visited on 05/18/2013).

[5] UKAEA, *History*. [Online]. Available: http://www.uk-atomic-energy.org.uk/about/history.html (visited on 05/18/2013).

[6] NDA, *The Nuclear Legacy*. [Online]. Available: http://www.nda.gov.uk/aboutus/the-nuclear-legacy.cfm (visited on 05/18/2013).

[7] C. Bayliss and K. Langley, *Nuclear decommissioning, waste management, and environmental site remediation*. Elsevier, 2003.

[8] UK, "The Decommissioning of the UK Nuclear Industry's Facilities", Tech. Rep. February, 2004. [Online]. Available: http://webarchive.nationalarchives.gov.uk/+/http:/www.berr.gov.uk/files/file30124.pdf.

[9] NDA, *How is radioactive waste managed now?*, 2013. [Online]. Available: http://www.nda.gov.uk/ukinventory/waste/waste-managed-now.cfm (visited on 05/20/2013).

[10] EPA US, *Plutonium*, 2013. [Online]. Available: http://www.epa.gov/radiation/radionuclides/plutonium.html (visited on 05/18/2013).

[11] M. Wainwright, *Cumbria rejects underground nuclear storage dump*, 2013. [Online]. Available: http://www.guardian.co.uk/environment/2013/jan/30/cumbria-rejects-underground-nuclear-storage.

[12] S. Uematsu, K. Kashiro, and N. Tobita, "Dismantling of gloveboxes for mox fuel fabrication by a glovebox dismantling facility", in *Proc. Waste Management Conference, Tucson, Arizona, USA*, 2002.

[13] T. B. Sheridan, "Telerobotics", *Automatica*, vol. 25, no. 4, pp. 487–507, 1989.

[14] Oxford, *Robot | definition of robot in english by oxford dictionaries*. [Online]. Available: https://en.oxforddictionaries.com/definition/robot.

[15] T. B. Sheridan, *Telerobotics, automation, and human supervisory control*. MIT press, 1992.

[16] P. F. Hokayem and M. W. Spong, "Bilateral teleoperation: An historical survey", *Automatica*, vol. 42, no. 12, pp. 2035–2057, 2006.

[17] S. Lichiardopol, *A survey on teleoperation*, English, ser. DCT rapporten. Technische Universiteit Eindhoven, 2007, DCT 2007.155.

[18] T. B. Sheridan and W. L. Verplank, "Human and computer control of undersea teleoperators", MASSACHUSETTS INST OF TECH CAMBRIDGE MAN-MACHINE SYSTEMS LAB, Tech. Rep., 1978.

[19] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "A model for types and levels of human interaction with automation", *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans*, vol. 30, no. 3, pp. 286–297, 2000.

[20] D. B. Kaber, E. Onal, and M. R. Endsley, "Design of automation for telerobots and the effect on performance, operator situation awareness, and subjective workload", *Human factors and ergonomics in manufacturing & service industries*, vol. 10, no. 4, pp. 409–430, 2000.

[21] M. R. Endsley, "Situation awareness global assessment technique (sagat)", in *Aerospace and Electronics Conference, 1988. NAECON 1988., Proceedings of the IEEE 1988 National*, IEEE, 1988, pp. 789–795.

[22] J. Y. Chen and M. J. Barnes, "Human–agent teaming for multirobot control: A review of human factors issues", *IEEE Transactions on Human-Machine Systems*, vol. 44, no. 1, pp. 13–29, 2014.

[23] J. Y. Chen, M. J. Barnes, and M. Harper-Sciarini, "Supervisory control of multiple robots: Human-performance issues and user-interface design", *IEEE transactions on systems, man and cybernetics, part C: Applications and reviews*, vol. 41, no. 4, pp. 435–454, 2011.

[24] T. B. Sheridan, "Human–robot interaction: Status and challenges", *Human factors*, vol. 58, no. 4, pp. 525–532, 2016.

[25] I. Rae, G. Venolia, J. C. Tang, and D. Molnar, "A framework for understanding and designing telepresence", in *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, ACM, 2015, pp. 1552–1566.

[26] K. M. Lee, "Presence, explicated", *Communication theory*, vol. 14, no. 1, pp. 27–50, 2004.

[27] T. B. Sheridan, "Musings on telepresence and virtual presence", *Presence: Teleoperators & Virtual Environments*, vol. 1, no. 1, pp. 120–126, 1992.

[28] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino, "Augmented reality: A class of displays on the reality-virtuality continuum", in *Telemanipulator and telepresence technologies*, International Society for Optics and Photonics, vol. 2351, 1995, pp. 282–293.

[29] S. Kratz and F. R. Ferriera, "Immersed remotely: Evaluating the use of head mounted devices for remote collaboration in robotic telepresence", in *Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on*, IEEE, 2016, pp. 638–645.

[30] J. P. Whitney, T. Chen, J. Mars, and J. K. Hodgins, "A hybrid hydrostatic transmission and human-safe haptic telepresence robot", in *Robotics and Automation (ICRA), 2016 IEEE International Conference on*, IEEE, 2016, pp. 690–695.

[31] G. Kipper and J. Rampolla, *Augmented reality: An emerging technologies guide to ar*. Elsevier, 2012.

[32] D. J. Bruemmer, D. A. Few, R. L. Boring, J. L. Marble, M. C. Walton, and C. W. Nielsen, "Shared understanding for collaborative control", *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 35, no. 4, pp. 494–504, 2005.

[33] D. J. Bruemmer, D. A. Few, M. C. Walton, R. L. Boring, J. L. Marble, C. W. Nielsen, and J. Garner, "Turn off the television!: Real-world robotic exploration experiments with a virtual 3-d display", in *System Sciences, 2005. HICSS'05. Proceedings of the 38th Annual Hawaii International Conference on*, IEEE, 2005, 296a–296a.

[34] J. Lessiter, J. Freeman, and J Davidoff, "Really hear? the effects of audio quality on presence", in *Proceedings of the Fourth Annual International Workshop on Presence*, 2001, pp. 288–324.

[35] A. Kiselev, M. Scherlund, A. Kristoffersson, N. Efremova, and A. Loutfi, "Auditory immersion with stereo sound in a mobile robotic telepresence system", in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts*, ACM, 2015, pp. 55–56.

[36] C. C. Bracken, G. Pettey, T. Guha, and B. E. Rubenking, "Sounding out small screens and telepresence", *Journal of Media Psychology*, 2010.

[37] S. Le Groux, J. Manzolli, and P. Verschure, "Interactive sonification of the spatial behavior of human and synthetic characters in a mixed-reality environment", in *Proceedings of the 10th Annual International Workshop on Presence*, 2007, pp. 27–34.

[38] S. Brewster, *Haptic human-computer interaction: First international workshop, glasgow, uk, august 31-september 1, 2000, proceedings*. Springer Science & Business Media, 2001, vol. 1.

[39] J. Y. Chen, E. C. Haas, and M. J. Barnes, "Human performance issues and user interface design for teleoperated robots", *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1231–1245, 2007.

[40] R. J. Stone, "Haptic feedback: A brief history from telepresence to virtual reality", in *Haptic Human-Computer Interaction*, Springer, 2001, pp. 1–16.

[41] R. J. Stone, "Human factors guidelines for interactive 3d and games-based training systems design",

[42] D. Systems, *Touch*. [Online]. Available: `https://uk.3dsystems.com/haptics-devices/touch/specifications`.

[43] F. Dimension. [Online]. Available: `http://www.forcedimension.com/products`.

[44] I. Sarakoglou, N. V. Garcia-Hernandez, N. G. Tsagarakis, and D. G. Caldwell, "A high performance tactile feedback display and its integration in teleoperation.",

[45] C.-H. King, M. O. Culjat, M. L. Franco, C. E. Lewis, E. P. Dutson, W. S. Grundfest, and J. W. Bisley, "Tactile feedback induces reduced grasping force in robot-assisted surgery", *IEEE transactions on haptics*, vol. 2, no. 2, pp. 103–110, 2009.

[46] N. Enayati, E. De Momi, and G. Ferrigno, "Haptics in robot-assisted surgery: Challenges and benefits", *IEEE reviews in biomedical engineering*, vol. 9, pp. 49–65, 2016.

[47] G. Hirzinger, B. Brunner, J. Dietrich, and J. Heindl, "Rotex-the first remotely controlled robot in space", in *Robotics and Automation, 1994. Proceedings., 1994 IEEE International Conference on*, IEEE, 1994, pp. 2604–2611.

[48] X. Chen, X. Chen, Y. Chen, and J. Chase, *Mobile robots: State of the art in land, sea, air, collaborative missions*. InTech, 2009.

[49] L. L. Whitcomb, "Underwater robotics: Out of the research laboratory and into the field", in *Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on*, IEEE, vol. 1, 2000, pp. 709–716.

[50] S. J. Zaloga, *Unmanned aerial vehicles: Robotic air warfare 1917–2007*. Bloomsbury Publishing, 2011.

[51] I. Leppänen, S. Salmi, and A. Halme, "Workpartner–hut-automations new hybrid walking machine".

[52] G. S. Guthart and J. K. Salisbury, "The intuitive/sup tm/telesurgery system: Overview and application", in *Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on*, IEEE, vol. 1, 2000, pp. 618–621.

[53] S. Avgousti, E. G. Christoforou, A. S. Panayides, S. Voskarides, C. Novales, L. Nouaille, C. S. Pattichis, and P. Vieyres, "Medical telerobotic systems: Current status and future trends", *Biomedical engineering online*, vol. 15, no. 1, p. 96, 2016.

[54] D. Robotics, *Double*. [Online]. Available: `https://www.doublerobotics.com/`.

[55] K. Dautenhahn, "Human robot interaction", in *The encyclopedia of human-computer interaction*, The Interaction Design Foundation, 2012, ch. 38.

[56] M. A. Goodrich, A. C. Schultz, *et al.*, "Human–robot interaction: A survey", *Foundations and Trends® in Human–Computer Interaction*, vol. 1, no. 3, pp. 203–275, 2007.

[57] S. Kiesler and P. Hinds, "Introduction to this special issue on human-robot interaction", *Human-computer interaction*, vol. 19, no. 1, pp. 1–8, 2004.

[58] P. Tsarouchi, S. Makris, and G. Chryssolouris, "Human–robot interaction review and challenges on task planning and programming", *International Journal of Computer Integrated Manufacturing*, vol. 29, no. 8, pp. 916–931, 2016.

[59] D. Goren-Bar, "Designing model-based intelligent dialogue systems", in *Information modeling in the new millennium*, IGI Global, 2001, pp. 268–284.

[60] N. Roy, G. Baltus, D. Fox, F. Gemperle, J. Goetz, T. Hirsch, D. Margaritis, M. Montemerlo, J. Pineau, J. Schulte, *et al.*, "Towards personal service robots for the elderly", in *Workshop on Interactive Robots and Entertainment (WIRE 2000)*, vol. 25, 2000, p. 184.

[61] H. A. Yanco, J. L. Drury, and J. Scholtz, "Beyond usability evaluation: Analysis of human-robot interaction at a major robotics competition", *Human-Computer Interaction*, vol. 19, no. 1, pp. 117–149, 2004.

[62] J. Scholtz, "Theory and evaluation of human robot interactions", in *System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on*, IEEE, 2003, 10–pp.

[63] H. A. Yanco and J. L. Drury, "A taxonomy for human-robot interaction".

[64] H. A. Yanco and J. Drury, "Classifying human-robot interaction: An updated taxonomy", in *Systems, man and cybernetics, 2004 IEEE International Conference on*, IEEE, vol. 3, 2004, pp. 2841–2846.

[65] *Mars science laboratory/curiosity*. [Online]. Available: `https://www.jpl.nasa.gov/news/fact_sheets/mars-science-laboratory.pdf`.

[66] R. R. Murphy, *Disaster robotics*. MIT press, 2014.

[67] M. Vasic and A. Billard, "Safety issues in human-robot interactions", in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, IEEE, 2013, pp. 197–204.

[68] H. Yan, M. H. Ang, and A. N. Poo, "A survey on perception methods for human–robot interaction in social robots", *International Journal of Social Robotics*, vol. 6, no. 1, pp. 85–119, 2014.

[69] J.-J. Cabibihan, H. Javed, M. Ang, and S. M. Aljunied, "Why robots? a survey on the roles and benefits of social robots in the therapy of

children with autism", *International journal of social robotics*, vol. 5, no. 4, pp. 593–618, 2013.

[70] J. Złotowski, D. Proudfoot, K. Yogeeswaran, and C. Bartneck, "Anthropomorphism: Opportunities and challenges in human–robot interaction", *International Journal of Social Robotics*, vol. 7, no. 3, pp. 347–360, 2015.

[71] T. S. Tadele, T. de Vries, and S. Stramigioli, "The safety of domestic robotics: A survey of various safety-related publications", *IEEE robotics & automation magazine*, vol. 21, no. 3, pp. 134–142, 2014.

[72] M. Brambilla, E. Ferrante, M. Birattari, and M. Dorigo, "Swarm robotics: A review from the swarm engineering perspective", *Swarm Intelligence*, vol. 7, no. 1, pp. 1–41, 2013.

[73] *Definition and domains of ergonomics | iea website*. [Online]. Available: `https://www.iea.cc/whats/index.html`.

[74] J. Reason, "Understanding adverse events: Human factors.", *BMJ Quality & Safety*, vol. 4, no. 2, pp. 80–89, 1995.

[75] R. W. Proctor and T. V. Zandt, *Human factors in simple and complex systems, second edition*. CRC Press, 2008.

[76] T. B. Sheridan, *Humans and automation: System design and research issues.* Human Factors and Ergonomics Society, 2002.

[77] C. D. Wickens, S. E. Gordon, Y. Liu, and J. Lee, "An introduction to human factors engineering", 1998.

[78] M. Johnson, B. Shrewsbury, S. Bertrand, T. Wu, D. Duran, M. Floyd, P. Abeles, D. Stephen, N. Mertins, A. Lesman, *et al.*, "Team ihmc's lessons learned from the darpa robotics challenge trials", *Journal of Field Robotics*, vol. 32, no. 2, pp. 192–208, 2015.

[79] K. Nagatani, S. Kiribayashi, Y. Okada, K. Otake, K. Yoshida, S. Tadokoro, T. Nishimura, T. Yoshida, E. Koyanagi, M. Fukushima, *et al.*, "Emergency response to the nuclear accident at the fukushima daiichi nuclear power plants using mobile rescue robots", *Journal of Field Robotics*, vol. 30, no. 1, pp. 44–63, 2013.

[80] J. A. Adams, "Critical considerations for human-robot interface development", in *Proceedings of 2002 AAAI Fall Symposium*, 2002, pp. 1–8.

[81] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs", *Journal of Cognitive Engineering and Decision Making*, vol. 2, no. 2, pp. 140–160, 2008.

[82] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research", in *Advances in psychology*, vol. 52, Elsevier, 1988, pp. 139–183.

[83] M. S. Young, K. A. Brookhuis, C. D. Wickens, and P. A. Hancock, "State of science: Mental workload in ergonomics", *Ergonomics*, vol. 58, no. 1, pp. 1–17, 2015.

[84] J. Y. Chen, E. C. Haas, K. Pillalamarri, and C. N. Jacobson, "Human-robot interface: Issues in operator performance, interface design, and technologies", ARMY RESEARCH LAB ABERDEEN PROVING GROUND MD, Tech. Rep., 2006.

[85] M. S. Prewett, R. C. Johnson, K. N. Saboe, L. R. Elliott, and M. D. Coovert, "Managing workload in human–robot interaction: A review of empirical studies", *Computers in Human Behavior*, vol. 26, no. 5, pp. 840–856, 2010.

[86] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems", *Human factors*, vol. 37, no. 1, pp. 32–64, 1995.

[87] C. E. Lathan and M. Tracey, "The effects of operator spatial perception and sensory feedback on human-robot teleoperation performance", *Presence: Teleoperators & virtual environments*, vol. 11, no. 4, pp. 368–377, 2002.

[88] A. Steinfeld, "Interface lessons for fully and semi-autonomous mobile robots", in *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on*, IEEE, vol. 3, 2004, pp. 2752–2757.

[89] J. M. O'Hara, W. S. Brown, P. M. Lewis, and J. J. Persensky, *Human-system interface design review guidelines (nureg-0700, revision 2)*, 2002. [Online]. Available: `https://www.nrc.gov/reading-rm/doc-collections/nuregs/staff/sr0700/`.

[90] L. Laursen, *Robot to human: "trust me"*, 2013. [Online]. Available: `https://spectrum.ieee.org/robotics/artificial-intelligence/robot-to-human-trust-me`.

[91] J. J. LaViola Jr, S. Buchanan, and C. Pittman, "Multimodal input for perceptual user interfaces", *Interactive Displays: Natural Human-Interface Technologies*, pp. 285–312, 2014.

[92] M. Turk, "Multimodal interaction: A review", *Pattern Recognition Letters*, vol. 36, pp. 189–195, 2014.

[93] H. A. Yanco and J. L. Drury, "Rescuing interfaces: A multi-year study of human-robot interaction at the aaai robot rescue competition", *Autonomous Robots*, vol. 22, no. 4, pp. 333–352, 2007.

[94] H. A. Yanco, A. Norton, W. Ober, D. Shane, A. Skinner, and J. Vice, "Analysis of human-robot interaction at the darpa robotics challenge trials", *Journal of Field Robotics*, vol. 32, no. 3, pp. 420–444, 2015.

[95] A. Steinfeld, T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, and M. Goodrich, "Common metrics for human-robot interaction", in *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, ACM, 2006, pp. 33–40.

[96] C. L. Bethel and R. R. Murphy, "Review of human studies methods in hri and recommendations", *International Journal of Social Robotics*, vol. 2, no. 4, pp. 347–359, 2010.

[97] G. Charalambous, S. Fletcher, and P. Webb, "Development of a human factors roadmap for the successful implementation of industrial human-robot collaboration", in *Advances in Ergonomics of Manufacturing: Managing the Enterprise of the Future*, Springer, 2016, pp. 195–206.

[98] J. E. Young, J. Sung, A. Voida, E. Sharlin, T. Igarashi, H. I. Christensen, and R. E. Grinter, "Evaluating human-robot interaction", *International Journal of Social Robotics*, vol. 3, no. 1, pp. 53–67, 2011.

[99] D. R. Olsen and M. A. Goodrich, "Metrics for evaluating human-robot interactions", in *Proceedings of PERMIS*, vol. 2003, 2003, p. 4.

[100] T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, and A. Steinfeld, "Common metrics for human-robot interaction", in *IEEE 2004 International Conference on Intelligent Robots and Systems, Sendai, Japan*, 2004.

[101] R. Murphy and D. Schreckenghost, "Survey of metrics for human-robot interaction", in *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*, IEEE Press, 2013, pp. 197–198.

[102] B. Cain, "A review of the mental workload literature", Defence Research And Development Toronto (Canada), Tech. Rep., 2007.

[103] S. G. Hart, "Nasa-task load index (nasa-tlx); 20 years later", in *Proceedings of the human factors and ergonomics society annual meeting*, Sage Publications Sage CA: Los Angeles, CA, vol. 50, 2006, pp. 904–908.

[104] C. Baber, D. P. Jenkins, G. H. Walker, L. A. Rafferty, P. M. Salmon, and N. A. Stanton, *Human factors methods: A practical guide for engineering and design*. Ashgate Publishing, Ltd., 2013.

[105] M. R. Endsley, "Direct measurement of situation awareness: Validity and use of sagat", *Situation awareness analysis and measurement*, vol. 10, p. 14, 2000.

[106] J. Brooke *et al.*, "Sus-a quick and dirty usability scale", *Usability evaluation in industry*, vol. 189, no. 194, pp. 4–7, 1996.

[107] A. Bangor, P. T. Kortum, and J. T. Miller, "An empirical evaluation of the system usability scale", *Intl. Journal of Human–Computer Interaction*, vol. 24, no. 6, pp. 574–594, 2008.

[108] J. Sauro, *A practical guide to the system usability scale: Background, benchmarks & best practices*. Measuring Usability LLC Denver, CO, 2011.

[109] J. R. Lewis and J. Sauro, "The factor structure of the system usability scale", in *International conference on human centered design*, Springer, 2009, pp. 94–103.

[110] D. W. Seward and M. J. Bakari, "The use of robotics and automation in nuclear decommissioning", in *22nd International Symposium on Automation and Robotics in Construction ISARC*, 2005, pp. 11–14.

[111] John, *Robotics and autonomous systems*, 2014. [Online]. Available: `http://nuclearmatters.co.uk/2014/07/robotics-and-autonomous-systems/`.

[112] R. Bogue, "Robots in the nuclear industry: A review of technologies and applications", *Industrial Robot: An International Journal*, vol. 38, no. 2, pp. 113–118, 2011.

[113] W. Ingamells, *Decommissioning specialist*, Private Communication, 2012.

[114] J. L. Burke and R. R. Murphy, “The safe human-robot ratio”, in *Human-robot interactions in future military operations*, CRC Press, 2016, pp. 51–70.

[115] Sellafield Ltd., “Introducing our new recruits”, *Sellafield Magazine*, pp. 46–53, 2015. [Online]. Available: `http://www.sellafieldsites.com/wp-content/uploads/2015/08/Sell-Magazine{\_}Q1.pdf`.

[116] K. Kim, B.-S. Park, H.-D. Kim, S. Hassan, and J. Yoon, “Robotic virtual manipulations of a nuclear hot-cell digital mock-up system”, *Assembly Automation*, vol. 31, no. 1, pp. 17–28, 2011.

[117] Sellafield Ltd., *Energy minister praises decommissioning progress at Sellafield [Press release]*, 2015. [Online]. Available: `http://www.sellafieldsites.com/press/energy-minister-praises-decommissioning-progress-at-sellafield-on-first-visit-to-site/`.

[118] “A telerobot for the nuclear industry”, *Industrial Robot: An International Journal*, vol. 17, no. 1, pp. 42–44, 1990.

[119] M. Talha, E. Ghalamzan, C Takahashi, J Kuo, W Ingamells, and R. Stolkin, “Towards robotic decommissioning of legacy nuclear plant: Results of human-factors experiments with tele-robotic manipulation, and a discussion of challenges and approaches for decommissioning”, in *2016 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, IEEE, 2016, pp. 166–173.

[120] T. Hornigold, *How fukushima changed japanese robotics and woke up the industry*, 2018. [Online]. Available: `https://singularityhub.com/2018/04/25/how-fukushima-changed-japanese-robotics-and-woke-up-the-industry/`.

[121] K. Qian, A. Song, J. Bao, and H. Zhang, “Small teleoperated robot for nuclear radiation and chemical leak detection”, *International Journal of*

Advanced Robotic Systems, vol. 9, 2012, ISSN: 17298806. DOI: 10.5772/50720.

[122] J. Cornell'a, D. Zerbato, L. Giona, P. Fiorini, and V. Sequeira, "Dynamics simulation for the training of teleoperated retrieval of spent nuclear fuel", *2012 IEEE International Conference on Robotics and Automation*, pp. 5012–5017, 2012, ISSN: 1050-4729. DOI: 10.1109/ICRA.2012.6224886. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6224886.

[123] Y. A. Anas, T. KAWAMURA, and H. YAMADA, "Operational evaluation of a construction robot tele-operation with force feedback : Evaluation of tele-robotic force perception using auditory feedback", *Transactions of The Japan Hydraulics & Pneumatics Society*, vol. 43, no. 1, pp. 8–15, 2012, ISSN: 18803121. DOI: 10.5739/jfps.43.8. [Online]. Available: http://ci.nii.ac.jp/naid/10030309918/en/.

[124] A. H. Herdocia, A. Shademan, and M. Jägersand, "Unimodal asymmetric interface for teleoperation of mobile manipulators: A user study", in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, IEEE, 2012, pp. 5214–5219.

[125] J. De Geeter, M. Decréton, and E. Colon, "The challenges of telerobotics in a nuclear environment", *Robotics and Autonomous Systems*, vol. 28, no. 1, pp. 5–17, 1999.

[126] NNL, *Robotics*. [Online]. Available: https://www.nnl.co.uk/innovation-science-and-technology/collaborations/universities/robotics/.

[127] E. Rohmer, S. P. N. Singh, and M. Freese, "V-rep: A versatile and scalable robot simulation framework", in *Proc. of The International Conference on Intelligent Robots and Systems (IROS)*, 2013.

[128] C. Robotics, *V-rep user manual*. [Online]. Available: http://www.coppeliarobotics.com/helpFiles/index.html (visited on 06/29/2018).

[129] Rethink Robotics, *Baxter robot*. [Online]. Available: http://www.rethinkrobotics.com/baxter/ (visited on 02/17/2016).

[130] M. Chiou, R. Stolkin, G. Bieksaite, N. Hawes, K. L. Shapiro, and T. Harrison, "Experimental analysis of a variable autonomy framework for controlling a remotely operating mobile robot", *IEEE IROS*, 2016.

[131] X. Su, J. Pan, and M. Grinter, "Improving Construction Equipment Operation Safety from a Human-centered Perspective", *Procedia Engineering*, vol. 118, pp. 290–295, 2015, ISSN: 18777058. DOI: 10.1016/j.proeng.2015.08.429. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1877705815020846.

[132] P. Cong Dung, P. Huynh Nhat Trinh, and F. Pål J., "Evaluation of Subjective and Objective Performance Metrics for Haptically Controlled Robotic Systems", *Modeling, Identification and Controll*, vol. 35, no. 3, pp. 147–157, 2014, ISSN: 0332-7353. DOI: 10.4173/mic.2014.3.2. [Online]. Available: http://www.mic-journal.no/ABS/MIC-2014-3-2.asp.

[133] N. Marturi, A. Rastegarpanah, C. Takahashi, M. Adjigble, R. Stolkin, S. Zurek, M. Kopicki, M. Talha, J. A. Kuo, and Y. Bekiroglu, "Towards advanced robotic manipulation for nuclear decommissioning: A pilot study on tele-operation and autonomy", in *2016 International Conference on Robotics and Automation for Humanitarian Applications (RAHA)*, IEEE, 2016, pp. 1–8.

[134] J. L. Sanz, *Advances in machine vision*. Springer Science & Business Media, 2012.

[135] O. Tahri, H. Araujo, F. Chaumette, and Y. Mezouar, "Robust image-based visual servoing using invariant visual information", *Robotics and Autonomous Systems*, vol. 61, no. 12, pp. 1588–1600, 2013.

[136] A. Shaukat, Y. Gao, J. A. Kuo, B. A. Bowen, and P. E. Mort, "Visual classification of waste material for nuclear decommissioning", *Robotics and Autonomous Systems*, vol. 75, pp. 365–378, 2016.

[137] J.-Y. Park, B.-H. Cho, and J.-K. Lee, "Trajectory-tracking control of underwater inspection robot for nuclear reactor internals using time delay control", *Nuclear engineering and design*, vol. 239, no. 11, pp. 2543–2550, 2009.

[138] C.-H. Baek, S. J. An, H.-i. Kim, S.-W. Kwak, and Y. H. Chung, "Development of a pinhole gamma camera for environmental monitoring", *Radiation Measurements*, vol. 59, pp. 114–118, 2013.

[139] A. Ettemeyer, "New techniques for nondestructive testing of piping systems in nuclear power plants", *Nuclear engineering and design*, vol. 139, no. 2, pp. 183–185, 1993.

[140] C. Kapoor and D. Tesar, "Integrated teleoperation and automation for nuclear facility cleanup", *Industrial Robot: An International Journal*, vol. 33, no. 6, pp. 469–484, 2006.

[141] E. Marchand, F. Spindler, and F. Chaumette, "Visp for visual servoing: A generic software platform with a wide class of robot control skills", *IEEE Robotics and Automation Society*, vol. 12, no. 4, pp. 40–52, 2005.

[142] V. Ortenzi, N. Marturi, R. Stolkin, J. A. Kuo, and M. Mistry, "Vision-guided state estimation and control of robotic manipulators which lack proprioceptive sensors", in-press.

[143] S. Dembélé, N. Piat, N. Marturi, and B. Tamadatze, "Gluing free assembly of an advanced 3d structure using visual servoing", in *Micromechanics and Microsystems Europe 2012*, 2011.

[144] A. Comport, E. Marchand, M. Pressigout, F. Chaumette, *et al.*, "Real-time markerless tracking for augmented reality: The virtual visual servoing framework", *Visualization and Computer Graphics, IEEE Transactions on*, vol. 12, no. 4, pp. 615–628, 2006.

[145] S. G. Hart, "Nasa-task load index (nasa-tlx); 20 years later", in *Proceedings of the human factors and ergonomics society annual meeting*, Sage Publications, vol. 50, 2006, pp. 904–908.

[146] N. H.P. R. Group *et al.*, "Task load index (nasa-tlx) v1. 0 computerised version", *NASA Ames Research Centre*, 1987.

[147] B. DeJong, J. Colgate, and M. Peshkin, "Mental transformations in human-robot interaction", in *Mixed Reality and Human-Robot Interaction*, Springer, 2011, pp. 35–51.

[148] J. Mackro, "Study of design and control of remote manipulators. part 4: Experiments in video camera positioning with regard to remote manipulation", 1973.

[149] M. A. Stuart, M. K. Manahan, J. M. Bierschwale, C. E. Sampaio, and A. J. Legendre, "Adaptive strategies of remote systems operators exposed to perturbed camera-viewing conditions", in *Proceedings of the Human Factors Society Annual Meeting*, SAGE Publications Sage CA: Los Angeles, CA, vol. 35, 1991, pp. 151–155.

[150] E. P. McCormick, C. D. Wickens, R. Banks, and M. Yeh, "Frame of reference effects on scientific visualization subtasks", *Human Factors*, vol. 40, no. 3, pp. 443–451, 1998.

[151] S. N. Plishka, "An experimental investigation of direct and indirect viewing of a remote manipulation", 2002.

[152] G. Doisy, A. Ronen, and Y. Edan, "Comparison of three different techniques for camera and motion control of a teleoperated robot", *Applied ergonomics*, vol. 58, pp. 527–534, 2017.

[153] A. Nishiyama, M. Moteki, K. Fujino, and T. Hashimoto, "Research on the comparison of operator viewpoints between manned and remote control operation in unmanned constuction systems", in *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, Vilnius Gediminas Technical University, Department of Construction Economics & Property, vol. 30, 2013, p. 1.

[154] D. Clark, "The effect of camera placement and display configuration on a remote manipulation task", *Theses - Daytona Beach*, 2006.

[155] C. W. Nielsen, M. A. Goodrich, and R. W. Ricks, "Ecological interfaces for improving mobile robot teleoperation", *IEEE Transactions on Robotics*, vol. 23, no. 5, pp. 927–941, 2007.

[156] C. Pérez Mejías, *Design of a telepresence interfacefor direct teleoperation of robots: The synergy between virtual reality and freelook control*, 2016.

[157] K. Theofilis, J. Orlosky, Y. Nagai, and K. Kiyokawa, "Panoramic view reconstruction for stereoscopic teleoperation of a humanoid robot", in *Humanoid Robots (Humanoids), 2016 IEEE-RAS 16th International Conference on*, IEEE, 2016, pp. 242–248.

[158] F. Okura, Y. Ueda, T. Sato, and N. Yokoya, "Free-viewpoint mobile robot teleoperation interface using view-dependent geometry and texture", *ITE Transactions on Media Technology and Applications*, vol. 2, no. 1, pp. 82–93, 2014.

[159] K. Krückel, F. Nolden, A. Ferrein, and I. Scholl, "Intuitive visual teleoperation for ugvs using free-look augmented reality displays", in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, IEEE, 2015, pp. 4412–4417.

[160] J. C. García, B. Patrão, L. Almeida, J. Pérez, P. Menezes, J. Dias, and P. J. Sanz, "A natural interface for remote operation of underwater robots", *IEEE computer graphics and applications*, no. 1, pp. 34–43, 2017.

[161] G. Adamides, C. Katsanos, Y. Parmet, G. Christou, M. Xenos, T. Hadzilacos, and Y. Edan, "Hri usability evaluation of interaction modes for a teleoperated agricultural robotic sprayer", *Applied ergonomics*, vol. 62, pp. 237–246, 2017.

[162] R. L. Smith and M. A. Stuart, "The effects of spatially displaced visual feedback on remote manipulator performance", in *Proceedings of the Human Factors Society Annual Meeting*, SAGE Publications Sage CA: Los Angeles, CA, vol. 33, 1989, pp. 1430–1434.

[163] Y. Fu, W. Chen, S. Li, and Z. Jiao, "Evaluating operator performance in teleoperated manipulator system factored by camera configurations", in *International Conference on Engineering Psychology and Cognitive Ergonomics*, Springer, 2015, pp. 327–334.

[164] P. Lamb and D. Owen, "Human performance in space telerobotic manipulation", in *Proceedings of the ACM symposium on Virtual reality software and technology*, ACM, 2005, pp. 31–37.

[165] S. Rahnamaei and S. Sirouspour, "Automatic viewpoint planning in teleoperation of a mobile robot", *Journal of Intelligent & Robotic Systems*, vol. 76, no. 3-4, pp. 443–460, 2014.

[166] B. Keyes, R. Casey, H. A. Yanco, B. A. Maxwell, and Y. Georgiev, "Camera placement and multi-camera fusion for remote robot operation", in *Proceedings of the IEEE international workshop on safety, security and rescue robotics*, National Institute of Standards and Technology Gaithersburg, MD, 2006, pp. 22–24.

[167] S. Maeyama, T. Okuno, and K. Watanabe, "View point decision algorithm for an autonomous robot to provide support images in the operability of a teleoperated robot", *SICE Journal of Control, Measurement, and System Integration*, vol. 9, no. 1, pp. 33–41, 2016.

[168] N. Shiroma, N. Sato, Y.-h. Chiu, and F. Matsuno, "Study on effective camera images for mobile robot teleoperation", in *Robot and Human Interactive Communication, 2004. ROMAN 2004. 13th IEEE International Workshop on*, IEEE, 2004, pp. 107–112.

[169] S. Hughes, J. Manojlovich, M. Lewis, and J. Gennari, "Camera control and decoupled motion for teleoperation", in *Systems, Man and Cybernetics, 2003. IEEE International Conference on*, IEEE, vol. 2, 2003, pp. 1339–1344.

[170] S. B. Hughes and M. Lewis, "Task-driven camera operations for robotic exploration", *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 35, no. 4, pp. 513–522, 2005.

[171] A. Pandya, L. A. Reisner, B. King, N. Lucas, A. Composto, M. Klein, and R. D. Ellis, "A review of camera viewpoint automation in robotic and laparoscopic surgery", *Robotics*, vol. 3, no. 3, pp. 310–329, 2014.

[172] B. G. Brooks, G. T. McKee, and P. S. Schenker, "The visual acts model for automated camera placement during teleoperation", in *ICRA*, 2001, pp. 1019–1024.

[173] D. Zhu, T. Gedeon, and K. Taylor, "Exploring camera viewpoint control models for a multi-tasking setting in teleoperation", in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2011, pp. 53–62.

[174] D. Saakes, V. Choudhary, D. Sakamoto, M. Inami, and T. Lgarashi, "A teleoperating interface for ground vehicles using autonomous flying

cameras", in *2013 23rd International Conference on Artificial Reality and Telexistence (ICAT)*, IEEE, 2013, pp. 13–19.

[175] M. Kamezaki, J. Yang, H. Iwata, and S. Sugano, "Visibility enhancement using autonomous multicamera controls with situational role assignment for teleoperated work machines", *Journal of Field Robotics*, vol. 33, no. 6, pp. 802–824, 2016.

[176] M. Wilde, Z. K. Chua, and A. Fleischner, "Effects of multivantage point systems on the teleoperation of spacecraft docking", *IEEE Transactions on Human-Machine Systems*, vol. 44, no. 2, pp. 200–210, 2014.

[177] H. Martins, I. Oakley, and R. Ventura, "Design and evaluation of a head-mounted display for immersive 3d teleoperation of field robots", *Robotica*, vol. 33, no. 10, pp. 2166–2185, 2015.

[178] O. Khatib, "Real-time obstacle avoidance for manipulators and mobile robots", in *Autonomous robot vehicles*, Springer, 1986, pp. 396–404.

[179] F. Flacco, T. Kröger, A. De Luca, and O. Khatib, "A depth space approach to human-robot collision avoidance", in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, IEEE, 2012, pp. 338–345.

[180] J. Israelsen, M. Beall, D. Bareiss, D. Stuart, E. Keeney, and J. van den Berg, "Automatic collision avoidance for manually tele-operated unmanned aerial vehicles", in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, IEEE, 2014, pp. 6638–6643.

[181] A. Budiyanto, A. Cahyadi, T. B. Adji, and O. Wahyunggoro, "Uav obstacle avoidance using potential field under dynamic environment", in *Control, Electronics, Renewable Energy and Communications (ICCEREC), 2015 International Conference on*, IEEE, 2015, pp. 187–192.

[182] E. Ackerman, *The secret to small drone obstacle avoidance is to just crash into stuff*, 2016. [Online]. Available: `https://spectrum.ieee.org/automaton/robotics/drones/the-secret-to-small-drone-obstacle-avoidance-is-to-just-crash-into-stuff`.

[183] Imagine3D, *Imagine3d download*, 2013. [Online]. Available: `https://sourceforge.net/projects/imagine3d/` (visited on 08/20/2017).

[184] R. N. Shepard and J. Metzler, "Mental rotation of three-dimensional objects", *Science*, vol. 171, no. 3972, pp. 701–703, 1971.

[185] J. Rebelo, T. Sednaoui, E. B. den Exter, T. Krueger, and A. Schiele, "Bilateral robot teleoperation: A wearable arm exoskeleton featuring an intuitive user interface", *IEEE Robotics & Automation Magazine*, vol. 21, no. 4, pp. 62–69, 2014.

[186] Haption, *Virtuose^{TM} 6d*. [Online]. Available: `https://www.haption.com/en/products-en.html`.

[187] I. Jo, Y. Park, and J. Bae, "A teleoperation system with an exoskeleton interface", in *Advanced Intelligent Mechatronics (AIM), 2013 IEEE/ASME International Conference on*, IEEE, 2013, pp. 1649–1654.

[188] C. Häggström, M. Englund, and O. Lindroos, "Examining the gaze behaviors of harvester operators: An eye-tracking study", *International journal of forest engineering*, vol. 26, no. 2, pp. 96–113, 2015.

[189] M. Talha and R. Stolkin, "Particle filter tracking of camouflaged targets by adaptive fusion of thermal and visible spectra camera data", *IEEE Sensors Journal*, vol. 14, no. 1, pp. 159–166, 2014.

[190] F. Abi-Farraj, N. Pedemonte, and P. R. Giordano, "A visual-based shared control architecture for remote telemanipulation", in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 2016, pp. 4266–4273.

[191] A. Treptow, G. Cielniak, and T. Duckett, "Active people recognition using thermal and grey images on a mobile security robot", in *Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on*, IEEE, 2005, pp. 2103–2108.

[192] G. Cielniak, T. Duckett, and A. J. Lilienthal, "Data association and occlusion handling for vision-based people tracking by mobile robots", *Robotics and Autonomous Systems*, vol. 58, no. 5, pp. 435–443, 2010.

[193] M. A. Hanif and U. Ali, "Optimized visual and thermal image fusion for efficient face recognition.", in *FUSION*, 2006, pp. 1–6.

[194] J. Heo, S. G. Kong, B. R. Abidi, and M. A. Abidi, "Fusion of visual and thermal signatures with eyeglass removal for robust face recognition", in *Computer Vision and Pattern Recognition Workshop, 2004. CVPRW'04. Conference on*, IEEE, 2004, pp. 122–122.

[195] R Stolkin, M Hodgetts, A Greig, and J Gilby, "Extended markov random fields for predictive image segmentation", in *Advances In Pattern Recognition*, World Scientific, 2007, pp. 208–214.

[196] K. Nummiaro, E. Koller-Meier, and L. Van Gool, "An adaptive color-based particle filter", *Image and vision computing*, vol. 21, no. 1, pp. 99–110, 2003.

[197] R. Stolkin, A. Greig, M. Hodgetts, and J. Gilby, "An em/e-mrf algorithm for adaptive model based tracking in extremely poor visibility", *Image and Vision Computing*, vol. 26, no. 4, pp. 480–495, 2008.

[198] C. O'Conaire, N. E. O'Connor, E. Cooke, and A. F. Smeaton, "Comparison of fusion methods for thermo-visual surveillance tracking", in *Information Fusion, 2006 9th International Conference on*, IEEE, 2006, pp. 1–7.

[199] W. E. L. Grimson, C. Stauffer, R. Romano, and L. Lee, "Using adaptive tracking to classify and monitor activities in a site", in *Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference on*, IEEE, 1998, pp. 22–29.

[200] C. Ó. Conaire, N. E. O'Connor, and A. Smeaton, "Thermo-visual feature fusion for object tracking using multiple spatiogram trackers", *Machine Vision and Applications*, vol. 19, no. 5-6, pp. 483–494, 2008.

[201] G. E. Hinton, "Products of experts", 1999.

[202] D. J. Duff, T. Mörwald, R. Stolkin, and J. Wyatt, "Physical simulation for monocular 3d model based tracking", in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, IEEE, 2011, pp. 5218–5225.

[203] T. Mörwald, M. Kopicki, R. Stolkin, J. Wyatt, S. Zurek, M. Zillich, and M. Vincze, "Predicting the unobservable visual 3d tracking with a probabilistic motion model", in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, IEEE, 2011, pp. 1849–1855.

[204] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking", *IEEE Transactions on pattern analysis and machine intelligence*, vol. 25, no. 5, pp. 564–577, 2003.

[205] F. Ennesser and G. Medioni, "Finding waldo, or focus of attention using local color information", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 8, pp. 805–809, 1995.

[206] A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions", *Bull. Calcutta Math. Soc.*, vol. 35, pp. 99–109, 1943.