CHALMERS
UNIVERSITY OF TECHNOLOGY

Control interface for teleoperated construction inspection robot

Usability focused design of interface for enabling remote inspections of construction sites

Master's Thesis in Industrial Design Engineering

Fredrik Longnell & Filip Sperr

DEPARTMENT OF INDUSTRIAL & MATERIALS SCIENCE
Division Design & Human factors

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2022
www.chalmers.se

Master of Science Thesis

Control interface for teleoperated construction inspection robot

Usability focused design of interface for enabling remote inspections of construction sites

Fredrik Longnell & Filip Sperr

DEPARTMENT OF INDUSTRIAL & MATERIALS SCIENCE
Division Design & Human factors

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2022

Control interface for teleoperated construction inspection robot
Usability focused design of interface for enabling remote inspections of construction sites
FREDRIK. LONGNELL & FILIP. SPERR

Department of Industrial & Materials Science
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone: +46 (0)31 772 + 1000

Cover:
The robot and interface side by side. By Fredrik Longnell and Filip Sperr

Print: Repro Service Chalmers
Gothenburg, Sweden 2022

Preface

This report is based on a master’s thesis carried out at the Industrial and Materials Science department of Chalmers University of Technology in the spring term of 2022. The project's scope was 30 ECTS and was carried out by two students in collaboration with the companies Merphi and Hines.

Acknowledgements

There are a number of individuals whose contributions we would like to acknowledge. Firstly, we would like to thank Mehrdad Farimani from Merphi who provided us with continuous feedback and was a source of great ideas throughout the project. We are also grateful towards Merphi at large for their contributions to the project and for taking this thesis on. From Hines we would like to thank Reza Farimani, who provided us with the background to the project as well as valuable input throughout.

We would also like to thank the construction experts from both the US and Sweden who volunteered to participate in the interviews, without whom the project would not have been possible. We are also grateful towards all the participants who helped evaluate the interface design in the usability test.

From Chalmers we would like to thank our examiner Bijan Aryana, who provided support and encouragement continuously throughout the project.

Glossary

Teleoperation - the control of a device or machine remotely.

BIM - Building Information Modeling.

Unity - software for creation of computer games.

Figma - software for creation of digital interface prototypes.

LIDAR - Light Detection and Ranging.

CAD - Computer Aided Design.

Hoist - a construction device that typically uses a pulley system to raise objects upward.

Fisheye lens - a very wide-angle lens with a field of vision covering up to 180°, the scale being reduced towards the edges.

Tracks - a system of vehicle propulsion used in tracked vehicles, running on a continuous band of treads or track plates driven by two or more wheels.

Abstract

This project was conducted to investigate the area of teleoperated robotic inspections at construction sites, as well as to design a prototype of the inspection robot's control interface with a focus on usability. The interface, a web based application, allows the user to access and control the driving, tool usage and various other features of the robot remotely.

The investigation into the area was conducted through two primary studies - one user study consisting of interviews with construction experts and one literature study focusing on relevant subareas of research. The data from these studies was analyzed using a modified thematic analysis approach, which yielded a number of themes and findings that formed the basis of the following design work. Throughout the design phase, concepts were evaluated using both Unity and Figma based prototypes by implementing these into usability tests aimed at identifying problems and opportunities for improvements.

From the research phase, a number of interesting findings were identified. Aspects such as the chaotic nature of construction sites, allocation of user attention and situation awareness proved especially challenging in ensuring satisfactory use of the interface. Additionally, enabling the entire spectrum of construction inspections, both from very general to specific ones, dictated much of the design work. The users also placed emphasis on certain overarching principles being adhered to, such as efficiency and safety. Through a set of iterations, a final design proposal was created, covering the primary use cases of the robot. The focus was on separating the advanced features from the basic ones, enhancing the explicitness of the controls and minimizing the operators' cognitive workload.

Keywords: Teleoperation, construction, inspection, robotics, interface, usability

Table of contents

1 Introduction

This chapter outlines the project's background, along with its delimitations, aim and research questions.

1.1 Background

The role of robotics has become more and more prevalent in industries across the world in recent decades (Statista, 2022). In factories and storage facilities alike, autonomous robots carry out repetitive tasks in a cheaper, safer and more reliable way than a force of humans ever could. In search and rescue missions, drones, robotic rovers and quadrupeds aid rescue personnel in their critical activities to an increasing degree (Delmerico, 2019). And this is not even mentioning the large changes occuring in the automotive industry, where autonomous cars are becoming a more viable option each year and are expected to replace manually operated cars to a large extent in the coming decades.

Advancements in the technologies and manufacturing of such robotic solutions mean that there are greater opportunities now than ever before in introducing new products and services onto markets - especially those that traditionally have not been willing to adopt new technologies. One such trade is the construction industry which, though it has made some progress in terms of digitizing and introducing robotic solutions, has been slow in comparison to many other industries (Schultz, 2020). Much of the work on construction sites is carried out with the same tried and tested tools and processes that have been used for decades, and inclusions of advanced robotics and machinery is often limited to narrow use cases.

A project was launched by Hines, an American real estate investment, development and management firm, to investigate the possibility of implementing a teleoperated robotic solution for alleviating the work associated with inspections on construction sites. The main purpose of this solution would be to enable remote inspections, with users accessing the product through the web. To handle the design of the robot, the Swedish consultancy firm Merphi was hired and given the task of developing the product. The authors of this report have assisted Merphi in this task by investigating the requirements of and designing an interface for controlling the inspection robot, with a focus on usability.

Hines believe that a teleoperated robotic solution could greatly aid the monitoring and inspection segment, which is a critical part of ensuring that work on-site progresses according to schedule and with sufficient quality. Currently, those conducting these activities often have to spend a large part of their day traveling to and from sites, which increases costs and makes their work less satisfactory. Some inspections may also be inconvenient or dangerous to perform, or may simply be monotonous and prone to human error. A teleoperated robot, controlled remotely and equipped with the tools required to perform various inspections could alleviate many of these issues and increase efficiency, reliability and safety on site.

The ability to effectively teleoperate the robot in a satisfying and intuitive way is crucial in enabling the work to be performed remotely. To achieve this, an interface for controlling and accessing the features of the robot is needed, and is what this thesis will focus on. The interface is the user’s access point to all the benefits the robot holds, and facilitates all remote interactions.

1.2 Aim

The aim of this project is to design a control interface for a teleoperated inspection robot, intended for construction work, that enables inspectors to perform their work remotely in an effective, efficient and satisfactory way.

1.3 Research questions

1. What needs and problems are associated with robot-enabled remote inspections on construction sites?

2. What are the requirements for designing a teleoperated robot interface for construction inspection?

3. What are the requirements for designing a teleoperated robot interface with high usability?

1.4 Delimitations

- The thesis will be based around an existing robot design, which will not be subject to change. The robot will be assumed to be capable of performing the things it is specified to do. The provided technical specifications as well as descriptions of the various softwares and intended work flows will act as the basis for the development of the interface.

- There will be no actual programming of the interface included in the project. As there is not enough time for making a fully functional interface, in addition to the authors not having extensive experience in programming, it is not included. The work will instead be presented using digital wireframe-based prototypes. These may then be used as a guide when creating the real software.

- Due to the lack of a physical robot, the driving can not be accurately evaluated and can therefore not be designed for with any specificity. The project will therefore not include input delay or other aspects of the driving which can only be experienced in combination with an actual robot.

- The interface will be designed with a web application structure in mind, and the intended usage will be imagined to be via a laptop or stationary computer. Although there may be benefits to providing users with alternate ways of accessing the interface, such as via their phones or ipads, computers provide good options for interaction (mouse and keyboard) and also provide the largest screen size for providing visual information to the user. The large size difference between phone- and computer screens also means that adapting a design from one medium to the other would require substantial work, which would take time away from other areas of the project.

2

Project Context

This chapter presents the relevant contextual information surrounding the project. Important elements such as the robot’s specifications, the target users and the nature of site visits are explained.

2.1 Robot description

The project is based around an existing design of a construction inspection robot, see figure 1, internally referred to as “Droidio", which is being developed by Merphi alongside the interface. Although there is no physical prototype at the time of writing this report, many of the specifications are already determined and act as a starting point for much of the project work. In this section, the relevant specifications are presented along with their implications for the interface and project at large. Due to confidentiality reasons, the exact measurements as well as the interior of the robot is not presented in this report. Information presented in this section is based on discussions with Merphi.

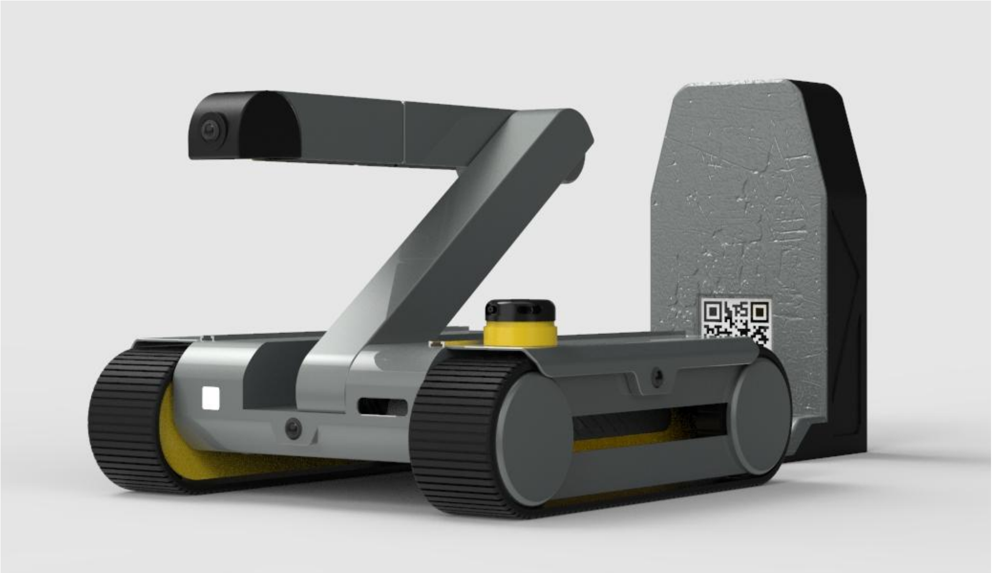


Figure 1: *The construction inspection robot design that the interface was based on.*

The robot consists of a rectangular body, roughly half a meter long and wide, propelled by two sets of tracks on either side of it. The tracks can accelerate independently of each other, providing the robot with a means of changing speed as well as turning. An arm is attached to the body, which can fold out to reach a height of around two meters. At the end of the arm, a full HD camera (1920x1080p) is located, which provides the user with visuals and enables them to perform inspections. In addition, there is an operational camera located at the front of the body, which is of lower quality and solely intended as a tool for the robot’s internal software. This software allows the robot to interpret the environment in front of it and react to obstacles.

On each side of the robot body, there is a fisheye camera, as can be noted in figure 2. Together, the feeds from these four cameras provide the ability to create a simulated top-down view. The robot also contains two 2-D lidars positioned on top of the body in opposite corners. These are able to provide sensory data in a plane around the robot, which allows for obstacle detection as well as basic 2-D mapping of the surrounding environment. Critically, a 2-D lidar is not able to reconstruct three dimensional environments, which prevents the robot from completely mapping its surroundings.

Figure 2: *Overhead view of Droidio.*

The robot uses a charging station to recharge its battery. This station will be placed on site and then function as a base of operations for the robot. The robot is able to autonomously dock and undock from the station using its operational camera coupled with a QR code attached to the front of the station. The camera will pick up the QR code in its field of view, at which point software is used to align and dock with the station. The station will likely also include a wifi hotspot and an improved ability to transfer data to and from the robot.

To allow for customization of the robot for any given use case, there are attachment points on both the body and arm of the robot. On the body, standard railings are used, which are typical for industrial robots and compatible with many existing tools. On the arm, the intention is for the entire front of the arm, which includes the camera, to be removable. This way, a separate front arm module containing additional tools can be swapped in to adapt the robot to different needs.

The robot also contains a speaker and a microphone for enabling communication with on-site personnel, as well as additional feedback from the environment in the form of audio. Critically, due to the nature of large construction sites, the robot must be able to travel between different floors within buildings. Although the robot is capable of maneuvering up stairs to some extent, the intended primary way of traveling between floors is via an elevator or construction hoist. In these cases, the communicatory capabilities of the robot is what enables users to instruct on-site personnel of what floor they wish to go to.

To allow for autonomous behaviors, the robot is outfitted with a range of different softwares. Firstly, to enable the creation of a map of the construction site, BIM data of the finished building is converted to a usable navigational map using existing methods. This BIM data is a digital, three dimensional representation of a building as it will be when construction is finished, which is present in practically all modern large scale building projects. Once the map is converted, algorithms for generating and selecting the most efficient path from one destination to another are used to enable autonomous pathfinding.

Because the BIM-data does not accurately reflect the current state of the construction site, but rather the intended final state, there will be cases where it does not line up with reality. Due to this, it is inevitable that the robot will encounter unforeseen obstacles as it conducts autonomous routes. As the robot contains a 2-D lidar, there are some options for handling such obstacles automatically. The one used as the basis for this project is using the lidar to create a secondary map, highlighting only the edges of walls and objects which the robot has detected through the lidar. This approach yields a less detailed version of the map, but is more accurate than the BIM-map as it is continuously updated. Theoretically, it is possible to update the BIM-map based on the lidar sensor data and then, using some advanced algorithms, maneuver the robot around obstacles in an intelligent way. However, these solutions represent a significant challenge and are therefore not considered in this project.

There is also a challenge of being able to accurately know where the robot is at any given time. Since the robot is operated remotely, and often will be used indoors in multi storey buildings, the GPS which is included in the robot is not enough to know exactly where the robot is. This is due to the GPS having roughly one to two meters of uncertainty in terms of planar positioning, as well as being unable to determine which height the robot is on. To resolve the height issue, the product is intended to either contain a specialized sensor for determining its height at any given point, or have the hoists at each site be connected to the robot and thereby be able to provide it with data regarding which floor it is on. To increase the accuracy of the planar positioning, additional sensors will be placed on each floor. This will provide a much more accurate positioning than the robot GPS, though it increases cost and setup time.

2.2 Site context and Target users

In this section, the site context and target users are presented. These segments are based on information that was relayed by Hines and were used as a basis for development in this project, but do not necessarily apply to all construction inspection robots.

The robot is primarily intended for use at large construction sites, such as skyscrapers or industrial complexes. These sites often feature multi storey buildings, several structures on the same site as well as both indoors and outdoors environments. For this project, indoor environments were the primary focus, though outdoor usage was still seen as something the robot should be able to handle.

On most sites, there is plenty of machinery, personnel and materials which form an intricate web of activities and ever changing conditions. Some particularly important aspects to consider are the safety of the workers that the robot will operate alongside, coexisting with large vehicles and machinery and getting around the site efficiently through rough terrain and changing conditions. Unlike smaller sites, this context also often includes stairs, construction hoists and other complications, which the robot will need to be able to navigate to some degree.

The intended user group mainly focuses on managerial individuals who conduct inspections on construction sites, while also spending a significant amount of time at their offices. This group ranges from project owners with little to no knowledge of construction to specialized contractors with expert knowledge in their field, which leads to very different requirements for the product based on which user is being considered. Many users are also complete novices when it comes to using robots in their work today, which has to be accounted for.

In addition to these primary users, the idea is also for on-site personnel and other secondary users to be able to operate the robot if there is a need for it. Since there is no real equivalent to the product on the market today, it is not clear which use cases may have a need for such a robot. Keeping the usage broad and accessible to many types of users is therefore beneficial, even though the primary use specifically concerns remote inspections. For the interface, this means that decisions that exclude users are generally seen as less attractive than those that are accessible for all potential users.

2.3 Types of construction inspections

Since it is not obvious what a typical construction inspection may contain, some information regarding inspections and site visits at large are presented below. This information mainly stems from interviews that were conducted with various construction professionals.

Most identified visits could be divided into one of two categories - general or specific inspections. General inspections deal with broad awareness of the work on site, and can often be described as wanting to get an overview of how things are progressing. Checking that contractors are working when they should and that certain features of a building are progressing according to schedule are both examples of general inspections. Specific inspections are more precise, and typically require more advanced methods for collecting the desired data. Due to their more advanced nature, these inspections are often conducted by an expert. Doing quality control inspections, taking measurements or sensor readings or comparing a physical construction element to one found in the BIM-model are all examples of specific inspections.

In addition to what kind of data is being collected, there is also a separation between inspections that are conducted with and without communication with on-site personnel. Though some planned inspections may not feature much communication, reactive visits often deal with some kind of problem solving and require the inspector to discuss the problem with the workers on site who will be the ones to solve it. There is also a social component to site visits, as managers often want and need to spend time socializing with their colleagues on-site and building trust.

The actual content of an inspection varies greatly depending on which user, field and environment it takes place in. The purpose of the inspection also has a massive impact on what data and which activities are included and excluded. Generally, visuals are the most important sensory data in any given inspection. Though there are inspections that deal with audial and haptic data, as well as sensors and tools for measuring very specific data points, visuals are the most important and frequently used of all. In conclusion, inspections can range from general to specific, can be cooperative or individual and mainly rely on visual data.

3

Theoretical Framework

A theoretical framework was established and used for definitions, guidelines and direction for what problems and areas to investigate and address. This framework is presented below.

3.1 Situation awareness

Endsley and Jones (2003), define situation awareness as both perceiving and understanding your current situation and surroundings, and specifically their implications for the task which you are currently undertaking. An example could be how a chef keeps an eye on their boiling pots, or a driver checking how much fuel is left in their car. The authors go on to highlight how the nature of situation awareness can drastically change between different fields. This is due to the fact that the purpose of situation awareness is to ensure that an individual has the necessary information to make informed decisions. Since decisions in different fields can be extremely dissimilar, so can the information relevant for situation awareness.

Endsley and Jones describe how situation awareness can be divided into three levels; perception, understanding and projection. These levels represent the different aspects of situation awareness and how well an individual can acquire situational information, interpret it and make projections based on it. They further highlight that many different factors affect how well situation awareness can be established, relating both to the operator, for example stress, and to the environment, like how informative an interface is.

According to Endsley and Jones, situation awareness is critical for effective completion of most tasks, but it may be more difficult to acquire in some areas. One such area is teleoperation, where the user is not co-located with the situation and environment on which they have to base their decisions. Gathering the necessary information to make effective decisions thus becomes a lot more difficult and Endsley and Jones therefore highlight the importance of properly designed human-machine systems to overcome this challenge. Taking situation awareness into account when designing the interface in this project will thus be necessary in order to create an effective teleoperation solution.

Endsley and Jones provide a number of design guidelines to enhance situation awareness in a general sense but also sets specifically for teleoperation and automation. These were used as input for the creation of requirements, based on which the concepts were evaluated throughout the project.

3.2 Usability

As much of this project has centered around the challenge of enabling even untrained and inexperienced users to successfully use the product, ease of use has been a core principle. Therefore, usability has been a critical concept throughout the project, as it relates to how user-friendly something is (Jordan, 1998).

The definition of usability that has been used in this project is the ISO definition, which reads: “...the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use” (International Organization for Standardization [ISO], 2018). Here, the effectiveness concerns how completely and accurately the goal set by the user could be met.

Efficiency relates to how much resources were required to reach the goal, for example time or mental effort. Satisfaction concerns the subjective experience of the user during the completion of the task (ISO, 2018). Because the ISO definition takes the user, task and environment into account, the usability of a product is highly context dependent which is something that must be taken into account when using it as a theoretical basis.

Jordan (1998), further categorizes usability into five different types, which has served as more concrete starting points for discussions and analysis during this project. The five types are:

- Guessability: A product's usability when the user tries to complete a task for the first time.
- Learnability: A product's usability when the user tries to complete a task with some previous experience.
- Experienced user potential: A product's usability when an experienced user attempts to complete a task.
- System potential: The highest degree of usability a product could enable for a certain task.
- Re-Usability: A product's usability when significant time has passed since the user last tried to complete the task.

Beyond these five concepts, Jordan (1998) describes ten guidelines to achieve good usability. These guidelines have been used as reference throughout the project to ensure that the created concepts maintained a high level of usability and served as groundwork for the requirements and guidelines used in this project. The ten guidelines are:

- Consistency: Tasks of similar nature should possess similar ways of execution.
- Compatibility: The user's interaction with the product should be compatible with their understanding of how other products function.
- Consideration of user resources: The product should seek to always match the users available resources at any given time during interaction.
- Feedback: The product should give clear indications about which actions the user has performed and what they resulted in.
- Error prevention and recovery: The product should counteract user errors and ensure that they can be easily recovered from if they occur.
- User control: The user should have the greatest possible control over the products actions and states.
- Visual clarity: All information that is shown to the user should be clear and easy to understand.
- Prioritization of functionality and information: The functions and information which are of the greatest importance to the user should be the most accessible.
- Appropriate transfer of technology: The product should, where suitable, utilize available technology from other fields and contexts to support its usability.
- Explicitness: The product should clearly communicate its features and how to interact with them to the user.

3.3 Human-robot interaction

Human-robot interaction (HRI) is an area which relates both to human factors and to technical areas and can, according to Sheridan (2016), generally be divided into four different areas. The first type of HRI relates to robots carrying out routine tasks with the human overseeing the processes. In this situation, the robots can act mostly independently within their limited context and the human only directs the robot where necessary. The second type of HRI relates to teleoperated robots. Sheridan defines teleoperation as directly controlling something for the purpose of moving around and interacting with the environment, where the operator and that which is controlled are not co-located. Teleoperated robots are commonly used in hazardous environments where it would not be safe for a physical operator to be present, like deep underwater, in space or under collapsed buildings.

While teleoperation can be a great tool for allowing users to complete tasks remotely or work in hazardous environments, it is associated with a number of problems which are important to consider when designing a teleoperated robot and its interface (Chen et al., 2007). These problems, which affect both the operator's ability to grasp their surroundings and to take appropriate actions, stem from the fact that the operator is physically disconnected from the thing they are controlling and the environment they are trying to perceive (Chen et al., 2007).

The third type of HRI concerns fully automated vehicles, such as cars or trains. In these situations, the human acts only as a passenger, although sometimes with the need to be ready in case a situation occurs where they are required to assume control (Sheridan, 2016). This relates to one of the main problems of automation which is the fact that the operator easily can end up being “out of the loop”, meaning that they lack sufficient understanding of what the automated robot is doing. This leads to them not being able to effectively take over control when it is needed (Endsley and Jones, 2003).

The final type of HRI is the one where the social aspect of the interaction is the most important, for example robots used for teaching or assisting patients in hospitals. With these types of social robots, user acceptance can be a big challenge.

The robot which this project is based around is primarily a teleoperated robot and thus the challenges of remote perception and effective control will be crucial areas to address. The robot will also have some autonomous features and therefore this kind of HRI will also be of interest, and the out-of-the-loop problem, as well as other related problems, will be important to resolve.

4 Method

The methods used throughout the project are outlined in this chapter. Adaptations made to certain methods based on the project needs are also explained.

4.1 Competitor analysis

A competitor analysis was compiled in order to gain an understanding of what similar products currently exist and gather inspiration. Competitors were sorted based on how similar the purpose of their products were to Droidio, as well as in which industry they were active. Competitors whose products' purposes were closely aligned with that of Droidio and who were active in the construction industry were generally regarded as the most interesting ones, whereas those with a less clear connection to inspection robots and the construction industry were regarded as less relevant. A variety of aspects were considered when studying the competitors. These aspects included the product's stage of operation, its functions, its interaction with people as well as what interfaces and physical controllers were used to operate it. The factors were listed for each competitor in a sheet, after which their respective strengths and weaknesses, as well as general conclusions, were extracted. These insights were then used as a basis when designing the literature and user studies, as they highlighted interesting ideas and topics to be investigated.

4.2 Literature study

A literature study entails searching current scientific literature for information and findings which can support the project (Wikberg Nilsson et al., 2015). The information can relate both to general theory about a certain field, to be used as background for the project, and to niche studies which can help answer specific research questions. In this project, literature studies were used both to create a theoretical basis for HRI, usability and teleoperation, and to learn from similar studies of teleoperated robots conducted in the past. The results from this study, together with the results from the interviews, laid the foundation for the functional analysis, and highlighted which areas could be investigated further throughout the project.

Based on the background research, a number of categories of interest were created which became the basis for the initial search for literature. It was conducted using Google Scholar to quickly find scientific material. As material relating to the area of teleoperated robots within the construction industry was limited, studies concerning teleoperated robots in other areas were also investigated. Information which was deemed as particularly important and relevant for this project was marked down and these insights were then summarized and organized into a list which formed the main takeaways from the literature study.

4.3 Interviews

Interviewing is a versatile method which can be used for many different purposes in order to probe potential users for their thoughts and feelings relating to a certain subject or object (Wikberg Nilsson et al., 2015). The interviewer asks the interview subject a predefined set of questions to trigger reflection and notes down their answers. The questions can be structured, semi-structured or unstructured, depending on the level of openness the interviewer wants in the answers (Wikberg Nilsson et al., 2015).

The primary purpose of interviews in this project was to identify and understand the needs and preferences of the users as well as the context of use, in relation to remote inspections at construction sites. Based on the project brief and background research, a number of questions and areas of interest had been identified which laid the groundwork for the interview template. The full template can be found in appendix B. There was, however, much uncertainty regarding the user needs and the context, and therefore a semi-structured format was used for the questions. This provided the subject with the option of focusing on what they thought was especially relevant for the project and pursuing interesting ideas which emerged during the interviews. The users who were sought for the interviews were those with a connection to inspections and construction work, as these were the intended final users of the product. As the product was primarily intended for the U.S market, a number of American users were included in addition to Swedish users.

4.4 Thematic analysis

A thematic analysis is used to analyze qualitative data, for example interviews, to find patterns within it. It consists of six different steps which aim to generate relevant conclusions regarding a certain topic, based on your collected data (Braun and Clarke, 2012). The first step is to become familiar with the data you are to analyze by reading or listening to it. The second step is to go through the data and, based on the questions you want to answer, mark it down with different codes. The codes are generated during the analysis and can be very diverse, noting recurring thoughts, feelings and concepts which show up during the interview. The codes may be semantic or latent, meaning the interviewees comments can be marked objectively or be subjectively interpreted.

The third step is to, based on the codes which have been created, look for themes which can help answer the questions that are being investigated. The fourth step consists of evaluating the created themes to ensure that they are relevant and representative of the original data, or if they should be removed or reworked. The last two steps concern defining and describing each theme and summarizing the analysis in a report (Braun and Clarke, 2012).

A thematic analysis was used in this project, as it provided a structured way to analyze the large amount of qualitative data which had been gathered through the user interviews**.** The flexibility of the method also supported the authors in dealing with the uncertainty regarding which aspects from the interviews would be most important. The first step was conducted by transcribing the interviews and reading them, after which the coding began, resulting in a list of the codes along with their relevant quotes. The authors deemed that a good enough understanding of the interviews and users had been achieved to allow for a more latent approach to codes and themes, which could generate more interesting insights in a shorter time. Most of the themes were therefore formulated as more concrete takeaways. Step four was then performed and all the codes and themes were evaluated and reorganized. Instead of performing the last two steps, which would yield a more presentable result, the final takeaways were extracted and used to build out the functional analysis, as this was deemed more valuable to the project.

4.5 Brainstorming

Brainstorming is an ideation method used to boost creativity and facilitate the generation of large amounts of ideas. The core concept of brainstorming revolves around four main rules: no criticism of ideas, explore wild ideas, merge the ideas together to create new ones and focus on quantity over quality (Wikberg Nilsson et al., 2015). These rules create an open environment that enables more productive ideation sessions, which are then followed by discussions. Due to the simplicity and speed of this method, it was used frequently during the ideation phase to explore the solution space and generate discussions. The ideation concerned both the general layout of the interface and the interaction with specific features. It was often based around different use scenarios, with the theme for the many brainstorming sessions concerning what the optimal interface would look like for a specific scenario

4.6 Functional analysis

A functional analysis is a method used to divide the functions of a product into a hierarchical structure. It begins with specifying a main function, the ultimate object which the product aims to facilitate, by combining a verb and a noun, and then dividing this function into sub-functions (Cross, 2000). A car, for example, could have the main function of “enable transportation” and a sub-function could be “provide steering”. Another key feature of the functional analysis is that it only expresses what the product should do, have or be, but not *how* it should do these things (Cross, 2000). This makes the method a great tool during ideation as it can act as a reference for what should be ideated upon and which functions must be addressed.

In this project, the functional analysis acted as the primary synthesis of the research phase, where conclusions from the background studies, literature study and user interviews were used to define the primary functions of the robot and interface. The ideation phase was then based around this list. While the method, as described by Cross, consists of arranging the sub-functions in a hierarchical diagram, in this project a list structure was used. The functions were also prioritized based on four different levels of importance which further helped during ideation as it made it easier to prioritize different features and information in the layout. The functional analysis can be found in appendix C.

4.7 Weighted decision matrix

A weighted decision matrix is a method used to compare and evaluate multiple concepts. The matrix is constructed by listing the most relevant design criteria or requirements for a design and weighting each one based on how important they are. One of the concepts is then chosen as a reference and then each of the other concepts are compared with this one, fulfilling each of the listed criteria to a higher, lower or similar degree. Depending on the concepts’ performance in this regard, and the weight of the criteria, the concepts receive positive or negative points. At the end, the points are tallied for each concept, including the reference concept, and then their overall scores can be compared to see which concepts performed best (Cross, 2000).

This method was used towards the end of the ideation phase, after a number of concrete concepts had been generated. As each concept contained numerous features and had multiple pros and cons, the decision matrix was an appropriate tool to help organize the evaluation and make sure that every important aspect had been considered. While the final score did help inform the decision of which concept to proceed with, the method was also valuable as a facilitator for discussion.

4.8 Prototyping

Prototyping entails creating representations of a design solution in order to have something concrete to evaluate and explore (Wikberg Nilsson et al., 2015). Different types of prototypes, such as sketches, physical models and digital wireframes, can be used to evaluate many different aspects of a solution, such as usability, ergonomics or appearance. A prototype can be either high- or low-fidelity, meaning it can appear more or less refined. A high-fidelity prototype is suitable to create later on in the process as the details of the concept are more well defined and low-fidelity prototypes are more suitable early in the project as they do not require a lot of time or resources to create and are thus easy to iterate upon (Benyon, 2010).

In this project, prototypes were used primarily as a way to evaluate the final concepts. To evaluate the concepts' usability, it was crucial that the users could interact with them during evaluation. For this purpose, functional prototypes were created. Low fidelity was seen as appropriate, as the appearance and feel of the interface was not the main aspect to be evaluated, but rather its functionality and usability.

As driving was a core feature of the product's usage, the controls were an important aspect to evaluate. For this purpose, Unity was used to create a prototype with functional driving. Unity is a game development platform which can be used to create many different types of games (Unity technologies, 2022). A second complimentary prototype was also created using Figma, which is a software that allows for the creation of interactable digital interfaces (Figma, 2022). The two prototypes were used in tandem throughout the usability tests; the Unity prototype enabling control of the robot's movement, and the Figma prototype providing a functional interface with buttons and menus.

Simpler, low fidelity prototypes, like sketches and digital mockups, were also used throughout the ideation phase to visualize different ideas and facilitate discussions.

4.9 Usability Testing

A usability test revolves around letting users try to accomplish tasks using an interface in a manner which represents the real use of the product, while observing and measuring the interaction (Hanington & Martin, 2019). Through analysis of the users' interaction and their comments during the test, conclusions regarding the product's usability and what problems exist within the interface can be determined. Using scenarios with specific goals rather than direct instructions is also common for usability tests (Hanington & Martin, 2019). This creates more realistic situations for the user and limits the help they get in understanding the interface, which can increase the validity of the test.

Performing usability tests with users was a crucial step in evaluating the final interface concept and its two versions. The tests provided insights into which of the two concepts had better usability, and enabled validation of the effectiveness and efficiency of the overall layout and control. Identifying concrete problems which could be addressed for the next concept iteration was also a key focus of the tests. The tests, which were made up of both Unity- and Figma-segments, consisted of six scenarios with a set of tasks that the users had to complete utilizing both the movement controls and the interface interactions from Unity and Figma respectively. After each task had been completed, the users were asked how easy or difficult the task was to complete and at the end of the test, a set of interview questions were asked to summarize their impressions and thoughts.

To evaluate the effectiveness, efficiency and satisfaction of the concepts, both qualitative and quantitative measurements were taken. The qualitative data consisted of comments made during the tests, as the users were asked to think out loud, answers to interview questions which took place after the tests and observations of the user's interaction with the interface. These observations focused on what transpired on the screen and were also, along with the audio, recorded to strengthen the analysis. The interview questions mainly concerned how well the user understood the interface, if they felt like they had situation awareness whilst driving and what problems they experienced.

The quantitative data consisted of completion times and numbers of errors for each task, split between the Unity and Figma prototypes. Actions which were considered errors included clicking on the wrong function, forgetting to perform an action, driving in the wrong direction and colliding with objects. As the errors were somewhat loosely defined, some subjective interpretation of the driving related errors was required to fairly compare the two concepts. The full test procedure can be found in Appendix A.

4.10 KJ Analysis

A KJ-analysis entails organizing data by grouping items which correlate somehow until distinct groups have formed. It is normally performed by writing data items on physical notes and placing them on a surface to get a better overview (Project-management.com, 2022). KJ-analysis was used to group all observations and comments from the usability tests to easier spot patterns and draw conclusions for different features of the interface. In this instance, the analysis was done digitally to save time.

5 Process

This chapter details the project's process and the execution of the methods.

5.1 Overall process

The process was loosely divided into four different phases, which are illustrated in figure 3. The research phase consisted of gathering and analyzing data to create knowledge about the users' needs and problems and how to best design an interface with high usability. In the ideation phase, these insights were used to generate a large number of interface designs, after which the best performing concept was selected for further development. The evaluation phase consisted of prototyping and usability testing to evaluate the final concept and identify areas of improvement. Iteration led to these phases somewhat overlapping as the exploration and synthesis was performed cyclically.

The research phase, for example, contained some ideation to generate insights about the design challenges and what areas warranted further investigation. The ideation phase also contained significant research, which served to fill knowledge gaps that were discovered as the concepts were being created. Throughout the ideation, the various concepts were also continuously evaluated and new insights were generated for the next iterations of the concepts. The final phase consisted of summarizing the results from the usability tests, as well as the results from the previous phases, to create one final iteration of the interface with clear connections to the conclusions from every phase.

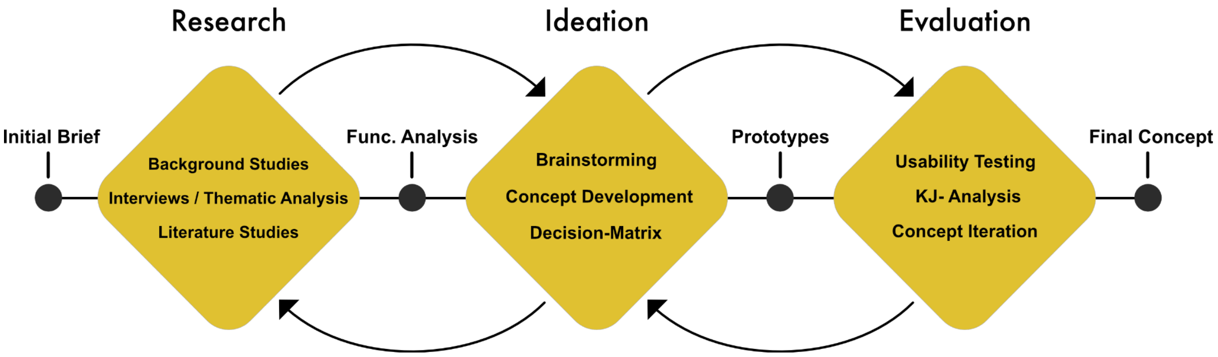


Figure 3: *Illustration of the project's phases.*

5.2 Research Phase

The first part of the project consisted of performing background studies in order to become familiar with the construction context, similar products on the market as well as the main challenges related to teleoperated robots. Benchmarking similar products and solutions was the most rigorous part of the research, but analyzes of potential features and the construction environment were also performed. The resulting findings formed the basis for the following literature- and interview studies.

The literature study then attempted to map out the opportunities and challenges associated with teleoperated robots as well as providing supporting theory regarding human-robot interaction, construction site technology and safety. The most interesting articles and their resulting takeaways were listed and grouped under different categories, such as "situation awareness", "information display" and "HRI". For each takeaway, supporting sources were noted down in order to ensure its validity. After this summary, the most important points were added to the functional analysis.

The interviews were the second main research activity, whose purpose was to create an understanding of the users and their work, as well as the context in which they operate. The end result consisted of a number of needs and problems which had a potential impact on the future design of the interface. Nine semi structured interviews were held in total, three of which were with American users and six with Swedish ones. All of the interviewees worked within the construction industry and were all involved in some kind of inspection work at their various sites. The first part of the interviews focused on the users' general work and the site context, as well as their thoughts on working remotely. The second part shifted the conversation towards their thoughts on inspection robots and what opportunities and possibilities they saw in terms of implementing remote inspections.

The thematic analysis was then conducted based on the transcribed interviews. After having generated all the codes and structured quotes, themes and takeaways, the relevant findings were added to the functional analysis.

5.3 Ideation Phase

An initial brainstorming session was conducted in parallel with the study analyses in order to start ideating on what the layout should look like and what features would be important to consider. Despite the analyses being unfinished, this ideation session generated some new insights and highlighted areas that required further investigation. It was clear at this point that the main features of the interface had to be decided on before more specific ideas could be generated. Thus, the final part of the research phase revolved around prioritizing the different functions in the functional analysis.

The main ideation sessions took place after the research phase was concluded and the results from the various analyses had been fully synthesized into the functional analysis. The main objective during these sessions was to ideate on the overall layout of the interface, such as the concept seen in figure 4, and was achieved by doing multiple rounds of quick sketches and discussions. Based on the results from the research phase, lists of user types and potential inspection scenarios were also created and these acted as additional support for the creation of new ideas. Layouts were, for example, created to match the needs of a specific type of user or site visit. The overall objective at this point was to create concepts of the layout which would be easy and efficient for the operators to use by providing visual clarity and by considering the users' mental resources.

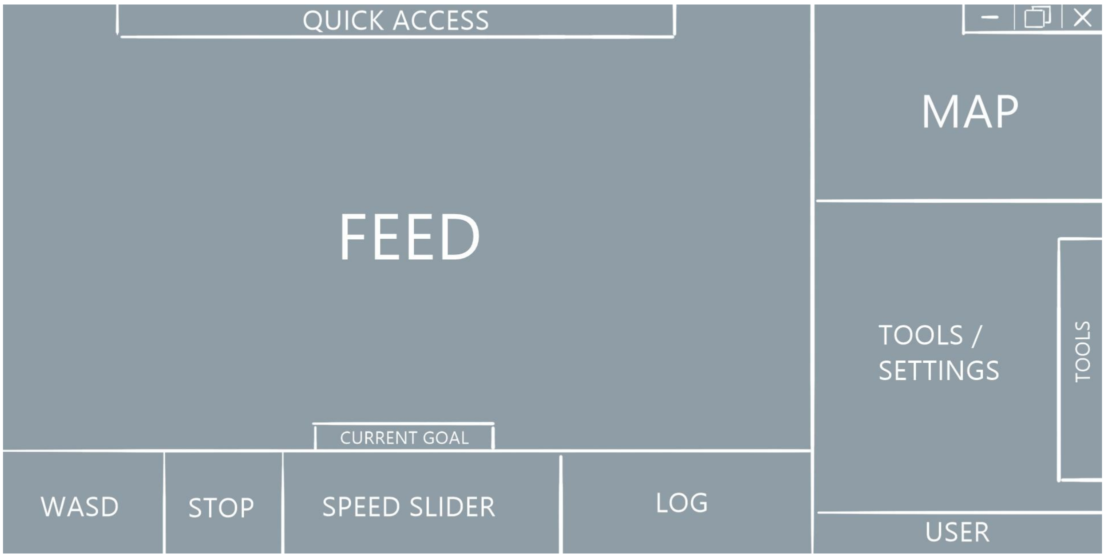
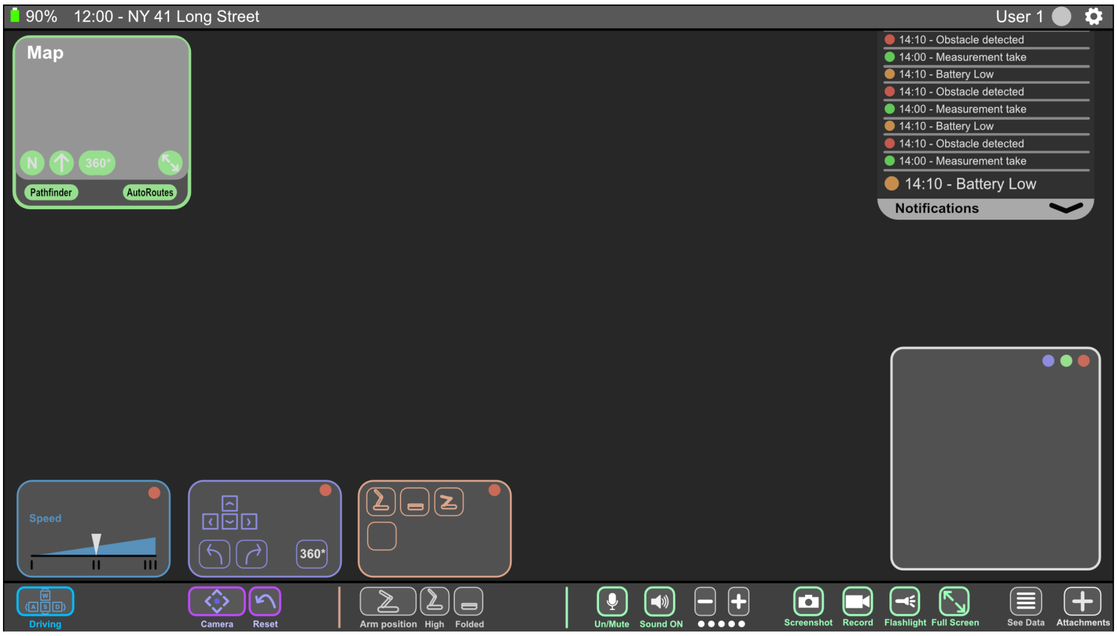


Figure 4: *Early topological layout.*

Once a large number of ideas had been generated and the main types of layouts had been identified, the ideas were combined into six distinct and more refined concepts, as the one seen in figure 5. These concepts were then evaluated using a weighted decision-matrix to identify their respective strengths and weaknesses and support comparisons, using the functional analysis' most central functionalities as a reference for the evaluation. Through these results and the discussion surrounding them, a final concept was chosen. This concept focused on separating the advanced features from the basic ones without limiting the user's options. This was achieved by having the basic features, such as driving, always available in a bar at the bottom and utilizing a togglable section which contained the more advanced features. While the main layout was determined, the designs of individual features were still not specified as there was not enough knowledge to support these decisions at this point. Thus, the final concept was divided into two separate versions containing different alternatives of some of the features. This enabled the different versions of the individual features to be evaluated in the usability test.



14:10 - Obstacle detected
14:00 - Measurement take
14:10 - Battery Low
14:10 - Obstacle detected
14:00 - Measurement take
14:10 - Battery Low
14:10 - Obstacle detected
14:00 - Measurement take
14:10 - Battery Low

Figure 5: *One of the six concepts used in the Decision matrix evaluation.*

5.4 Evaluation Phase

After the two versions of the final concepts had been created, the next objective was to conduct a usability test to evaluate their respective levels of usability and determine if any problems existed with their designs. To evaluate the usability, functional prototypes had to be created so that the users could actually interact with the interfaces. In order to evaluate the driving and control of the robot using the interface, the game development platform Unity was used to create a driving simulator for the robot, see figure 6. This allowed the users to control a 3-D model of the robot in real time and thus create a more realistic interaction. However, due to time limitations, all interface interaction could not be implemented in the Unity prototype and a Figma prototype was therefore created to complement it. Thus by utilizing the two prototypes together, the usability of both the driving and the general interface interaction could be evaluated.

Figure 6: *The Unity prototype used in the usability tests.*

The Unity segment included a simulation of a typical construction site environment, consisting of four zones in a square layout, see figure 7. Pathways to each of the different zones were present, though the width of these paths and amount of obstacles varied from zone to zone. Throughout the simulation, construction elements such as walls, pillars and blocks were positioned alongside models of people. A crude collision detection feature allowed for a simulated proximity sensor to communicate when any of these elements were collided with. The actual robot was a simple 3-D model which could go forwards, backwards, and rotate. The camera could also be controlled independently and, finally, the camera could be moved in a vertical and horizontal plane to simulate the arm movement of the robot. The tasks were structured in a way which required users to navigate to each of the different zones and complete tasks that mirrored ones that might take place in a real world scenario, such as looking for a particular construction element, taking photos or using the construction hoist. For many of the tasks, the users would first drive to the relevant zone in Unity and then swap to the Figma prototype for the specific interactions required for that task. The swaps were handled by the facilitator of the tests, using a simple “alt-tab” approach to switch between the two programs.

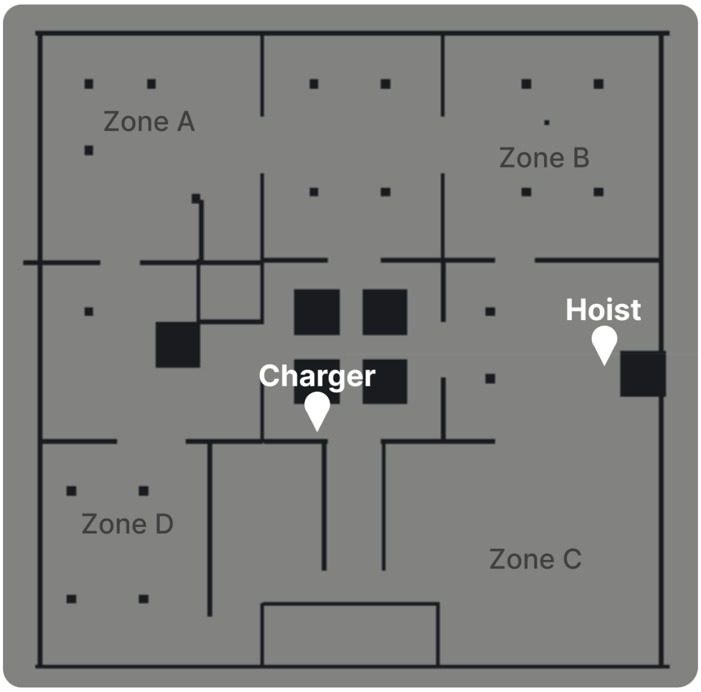


Figure 7: *Map of the Unity prototype's simulated environment.*

When the prototypes had been created, the usability test was performed with six users. One user was from the intended user group of construction professionals who regularly perform inspections on site. Of the other five, two were from the infrastructure sector and the remaining three did not work in the intended field or with the intended profession. As the core of the test related to general usability and was not directly related to the construction industry, this was seen as acceptable. The tests were conducted in home- or office environments as this was the intended use context for the final product, and the tasks attempted to combine both robot control and interface interaction to create more realistic scenarios.

The quantitative data was primarily used to evaluate the efficiency aspect of the concept's usability. Thus, the average completion time and number of error occurrences was summarized for each of the six tasks. The two concepts could then be compared to evaluate which of their respective parts had the best usability. The qualitative data was analyzed using a KJ analysis and the comments were labeled to indicate whether they related to concept 1 or 2, which made it easier to see similarities and differences between the two concepts.

5.5 Finalization

Based on the results from the usability test, decisions were made regarding which versions of the individual features had the best usability. For instance, the decisions were made to lock the camera panning to the robot and reduce the number of features to keep track of when driving. A number of problems with the interface were also identified during the test. The main problems related to lacking guessability and understandable feedback as well as demands on attention being too high. Rapid ideation and discussions were utilized to address these problems and enhance the final iteration in terms of explicitness, feedback, visual clarity and consideration of the users resources. When the final concept had been determined, it was then further iterated on to achieve a more uniform aesthetic, see figure 8. The usability tests did not indicate that the grouping of functions using colors were critical.

Thus, having a more uniform, clear and presentable appearance was deemed to be more important, especially considering that the remaining use of color still supported visual clarity and prioritization of functions and information.

Figure 8: *The finalized design.*

6 Results

This chapter presents the results of the user- and literature studies, as well as the usability tests.

6.1 Thematic analysis

The user study consisted of nine interviews with users from the construction industry and other relevant areas. Three of the users were American and the rest were Swedish. The interviews were then transcribed and a modified thematic analysis was conducted to find needs, problems and themes. The results of the thematic analysis are presented below.

6.1.1 Variation in user needs

Site visits can generally be divided into two categories depending on what kind of data is to be collected: general or specific visits. General data relates to basic and easily observable information, such as whether a certain task is currently being carried out on the site or if a particular part of the construction has been installed. Examples given during the interviews were checking if an oven had been installed or if the workers were performing the activity that was planned. General data visits are primarily carried out by those with more managerial roles in the project, for example project supervisors, investors and owners. These individuals were described as wanting the big picture of how the construction is proceeding and not being interested in the details. Specific data relates to that which cannot be seen at a glance. It concerns things like quality assurance and taking measurements, for example inspecting a small crack formation on a pillar or measuring the heat of a concrete section. These inspections are usually carried out by lower level managers, designers and specialized inspectors, though many actors may perform both general and specific inspections.

As general data relates to getting an overview of what is happening on the site and if things are proceeding as planned, it does not require as many tools or complicated procedures as specific inspections do. Rather, simply walking around the construction site and observing your surroundings is usually sufficient. Thus, this creates two distinct use cases for remote inspections with very different requirements in regards to the level of complexity needed in the interface, so as to match the level of complexity of the associated tasks. A teleoperation interface needs to support the more advanced use cases, without overcomplicating the basic ones, to successfully enable remote work at construction sites.

Another distinction which can be made between different types of visits is whether they are planned or reactive ones. Reactive visits are ones that the user did not anticipate, and usually relate to solving some kind of unforeseen problem in collaboration with on-site personnel. The planned visits on the other hand relate more to quality control and progress monitoring. The users generally have a specific goal in mind and can plan their visit based on what they are looking for. One user, for example, described a specific checklist which was used when a reinforcing bar was to be inspected. Thus, the users have a need for flexibility in order to perform both planned visits as well as reactive ones where speed is of greater importance. This need will also be present in remote inspections which means that a teleoperation interface must take this into account.

Beyond the direct driving controls of the robot, the operator will also have to be able to find their way around the site and know what to look for in order to complete their tasks. While some users, like on-site supervisors, might have extensive knowledge of the site, others, like contractors or investors, may not be as involved in the project. If these less involved individuals are to be able to perform their inspections remotely, they may need additional assistance in terms of navigation and understanding the site layout, especially since communicating and asking for directions will be more difficult due to not being co-located with on-site personnel.

Most of the users also pointed out that an interface meant to enable remote inspections had to be simple and easy to use, especially if it aims to be accessible to as many users as possible. Many users mentioned that they did not think that their colleagues would be able to use anything that was more advanced than the most basic computer software. This concern was primarily directed towards older users, as both young and old interviewees thought their older colleagues would be less capable of using the interface. Thus, ensuring that the teleoperation interface is simple and easy to use was determined to be a very important factor in enabling remote inspection. Furthermore, separating which features are meant for the most basic inspections, and which are meant for the more advanced inspections and experienced users, may be necessary to not overwhelm or deter less adept users.

6.1.2 Need for efficiency

Most users expressed that efficiency was highly valued in their work and that the overall efficiency in terms of time would have to increase for them to consider a teleoperated robot a viable option for performing their work. Several users also identified their own travel time to and from sites as having a negative impact on efficiency in their current situation. As one of the robot's primary objectives is to address this issue, this points towards a good match with the users' needs and problems. Furthermore, there might be value in highlighting just how much time or money a teleoperated robot is saving them. As users believed that cost and time efficiency would be the main hurdles to purchasing such a product and that its value to the project had to be clear, tracking these metrics could help communicate the product's value. The interface should also be quick to navigate through and enable the user to perform their work tasks in an efficient way.

There is, however, a part of the intended user group which already spend a majority of their time near the site, for example in field offices, and therefore might not benefit as much from being able to remotely access the site. Thus, for these users, controlling the robot has to be at least comparable in terms of time to performing the tasks themselves. For this to be the case, many of the users stated that automation was desirable, or even necessary, as they would otherwise not gain anything from using a robot. One example given was how an autonomous robot could save the workers time by fetching items. Another user suggested it could benefit larger sites where checking on something on the opposite end of the site could be time consuming. If a product is to be useful for a greater part of the user group, it may therefore be beneficial to include such automatic features.

6.1.3 Changing site conditions

One of the most prominent aspects of the construction site context is constant change, with one user describing them as being "alive". Not only are there always tools and material laying around and being moved, but the entire structure in which the workers operate is changing daily as the structure is being built or demolished. This means that entire sections of the site can appear or disappear between visits, and paths that were previously accessible may be blocked off due to recently constructed walls. This can cause even those who are familiar with the site to have trouble navigating around it. Thus, there is a need to stay updated on how the site has changed since your last visit, something which might prove to be even more important when the user is not there in person.

While the users had some ideas of how to address the need to stay updated, none of these ideas were particularly concrete, and the users thought it would be a difficult problem to solve. Thus, the most important takeaway is that a robotic solution meant to enable remote inspections needs to be adaptable and robust enough to handle changes in the environment. For instance, any automatic functions should optimally be able to resolve unexpected obstacles or otherwise smoothly transition control back to the user if their input is required. The interface should also assist the user in updating their mental model of the site layout in any way that it can.

6.1.4 Importance of functional communication

The reactive site visits, which usually relate to unforeseen problems, can be divided into those where the user is collecting data individually and those where the inspection is running in parallel to discussions with on site personnel. Many interviewees stated that workers on site would regularly ask for their input in solving a problem and then discuss the situation to find a solution. Because some users perform this type of inspection frequently, and almost all users described some type of communication during their visits, the need to communicate with on site personnel is clearly important. However, while many users thought a robot could support this type of functional and technical communication regarding problems on site, they were very skeptical towards the robot performing any social interactions. Aspects such as building team cohesion or ensuring that workers are following safety protocols were highlighted as examples where a robot would not be able to have the same effect as a human. Thus, this type of robot's communicatory functions should focus on facilitating effective two-way communication with on-site personnel to discuss concrete problems, rather than trying to replace the operator's social presence on site.

The users also mentioned that when trying to solve problems with on-site personnel, they often use some kind of mediating object. This could for example be the problem area itself where the user wants to highlight something in the environment, or the user might want to show a sketch or blueprint when explaining something. Thus, the user has a need to utilize mediating objects and interact with the physical environment. When the user is not physically present on site, supporting ways to perform these actions will therefore be important.

6.1.5 Problems with reaching inspection elements

Many of the elements which users may want to inspect are ones that are hard to reach. In particular, users described often encountering elements located in tight spaces or at high elevations. Users may in these cases have to use ladders or man lifts to reach some elevated areas and can sometimes not reach them at all, instead being forced to perform the inspection from a distance. Thus, the users have a need to inspect hard to reach areas, which could pose a problem for a robot that is lacking the capabilities of a human in reaching such places. The robot would need to enable users to conduct these types of inspection in other ways, for example through the use of camera zooming or extendable joints.

The users also indicated that the terrain on the construction sites could prove problematic for maneuvering a robot, due to factors such as uneven ground, debris, and various tools or material laying around. This could be a challenge for both the robot and the interface as the robot must be technically capable of driving through rough terrain and the interface must provide the user with sufficient situation awareness to maneuver around all the potential obstacles and debris.

Another challenge related to reaching the inspection elements relates to traversing between floors. This aspect is particularly important for high rise buildings where the user might need to visit several different floors for each inspection. Making sure that the robot can be transported between floors efficiently was highlighted as a big concern during the interviews, but users also pointed out that most larger construction sites have hoists which could be used for this purpose. However, during certain parts of the day, the hoist operator might not be present. Thus, the robot must be capable of using the hoist without an operator present, or incorporate an intuitive and simple way for the user to contact the operator or other workers on site to assist them in using the hoist. Furthermore, providing information about which floor the robot is on will also be an important feature, especially since users expressed how difficult it can be to differentiate between the numerous floors of a high rise building.

6.1.6 Need for situation awareness

Many of the users regularly work on large sites, both in terms of area and the height of the buildings in question, which can lead to some problems. Many of the users did, for example, describe how they sometimes could get lost on the site and not be able to reach their destination. Even users who were otherwise very familiar with the site and spent almost all their time there admitted that they sometimes felt a bit lost, especially in relation to their current floor. As global orientation is a problem even for those familiar with the site, it will be an ever greater challenge for those who have spent limited time there and on top of that must navigate through a teleoperated robot, which in itself diminishes the user's ability to locate themselves. Thus, the teleoperation interface must support the users' situation awareness in relation to their global position on the site so that they understand where they are and which way they should be going.

Almost all large construction sites today use some form of BIM. While some users stated that they did not use it in their daily work, those who did use it said it was a great help in understanding the site and the construction elements. The fact that BIM is so widely used is a great opportunity for the interface solution if it can be incorporated effectively. This would mean that detailed models of walls, electrical systems and more can be accessed by the user and support navigation, orientation and the inspection itself, filling the need of a visual representation of the site.

The users also stated that they desired local situation awareness if they were to drive a robot. They mentioned problems like understanding the scale of things in the environment, knowing what is behind the robot and understanding which direction you are facing. Perceiving scale and depth was described as particularly important for certain types of inspections where the placement of different components had to be observed. Thus, the interface should assist the user with their local situation awareness, both in terms of showing where the robot is in relation to the immediate surroundings and obstacles, but also through helping the user understand the scales and distances in their field of view.

6.1.7 Need for data handling

Many of the interviewees stated that they often used photos as a tool in their inspections. They took photos of objects of interest on-site, both to document their own work and to have a reference for discussing a problem with their colleagues. Beyond photos, they also expressed the benefit of recording and storing video as well as specific measurements from their visit. This data would also include meta data, such as what routes the robot took and how long it was active.
Together, this collected data would allow users to review previous visits and see what was going on at a certain location at a specific time and date, which is an idea that users expressed a lot of positivity towards.

However, they also noted that both photos and other forms of collected data need to be handled in an organized and efficient way so that their value can be realized. Thus, to enable remote inspections, the teleoperation interface should not only provide a way to collect the various types of data listed, but need to have an efficient way of handling this data. One possibility could be to simply provide the option to export the data to external, standardized locations in the cloud which some users are already working with.

6.1.8 Relative importance of senses

Most users expressed that the visual aspect of their inspections was the most important one, as opposed to other sensory data like audio or haptics. Some users even stated that visual data was the only type they collected. They also did not think that this would change if they performed the inspection remotely. The visual data concerns everything from getting an overview of site progress to closely observing a specific material for quality control. The importance of visual inspection means that the visual feedback and related controls must be prioritized in the interface to match this need.

There is also some need for audial feedback, both in general and specific scenarios. For instance, one user mentioned that they would like to know if there was work going on on the opposite side of a wall to support their situation awareness. Another user described a scenario where tapping on a rebar generated a sound which helped determine its quality and users also described more standardized sound inspections taking place to establish sound ratings. Some users also highlighted the need to adjust the sound levels as the sites are often very loud and sound levels can fluctuate dramatically. Thus the interface should have a way to record audio data as well as quickly adjusting the sound levels.

6.1.9 Safety

Safety was highlighted by the users as a crucial factor when trying to introduce new technologies into the construction industry. For a robot to be effectively implemented it is therefore crucial that consideration has been put into its operation to make sure that it does not become a safety hazard. For the interface, this means that it must anticipate various safety issues related to the robot's operation, such as driving near high elevation edges or in close proximity to people. By alerting the user of these risks, regardless of what the robot is capable of doing on its own, the risk of accidents will be minimized and users will be reassured of the robot’s awareness of these risky scenarios.

There was also some concern toward the other side of the spectrum of use in regards to the robots physical presence on site. Specifically, some users were concerned that workers on site might not see the robot and accidentally trip over it or run into it. Thus, ensuring that the robot's presence on site is made explicit to anyone in its vicinity will reduce the risk of accidents.

6.2 Literature study

The literature study consisted of identifying relevant literature on a number of areas related to the project's focus. The main findings from this study are presented below.

6.2.1 Errors and accidents in construction

Throughout the user study, it became clear that safety and avoiding accidents played a large role in the users' everyday work. Though the robot will not be particularly large or fast and will have sensors to prevent it from driving into objects or off of ledges, accidents could still occur due to the operation of the robot. It is therefore important to, despite the low perceived risk, investigate how accidents occur on construction sites and how to prevent user errors when operating the robot.

Mitropoulos et al. (2005) describes how human errors occur once the users are outside the "safe zone". Their accident causation model is based on the model created by Rasmussen et al. (1994), which illustrates the occurrence of errors by two primary zones in which the workers operate. Optimally, the worker should operate in the safe zone where they are in control and errors will not occur. However, the workers' desire to work efficiently, as well as external pressure to do so, pushes the user out of the safe zone as they take various kinds of shortcuts. Based on this model, Mitropoulos et al. suggest that since the workers will always try to find shortcuts, the key to accident avoidance is to allow them to take these shortcuts in a safe way. This can be achieved by making sure that the workers remain in control at all times, even if they try to take shortcuts, as well as establishing proper error prevention and recovery protocols (Mitropoulos et al., 2005).

For teleoperation, this means that the possible shortcuts the operator can take need to be considered and that measures need to be taken, through the interface, to ensure that they remain in control in these situations. For example, an operator might try to drive through a narrow gap instead of going around the area. To ensure that they remain in control of the situation, the interface could inform them of how close they are to the obstacle using sensor data, light up the path so that the user clearly sees the obstacle or use additional camera angles to give a better view of the robot's size compared to the gap. Employing error prevention and recovery protocols will also be important to enable safe and efficient operation. This could for example relate to automatic breaks if the robot is approaching an obstacle, or a quick reset for the robots joints if they were to become oddly positioned. Using the interface to provide increased information, awareness, as well as knowledge of how to handle errors, are all things which will allow the operator to remain in control and thus to work efficiently with minimized risk of errors and accidents.

6.2.2 Interaction and control

In their article on task-level authoring, Senft et al. (2021) seek to provide ways of alleviating the interaction between humans and robots. In the text, the authors describe a rule set for solving the challenges of having novices specify task plans for robots to complete autonomously. The rules are divided into four categories, the first two of which are especially relevant for this thesis. The first rule states that observation, planning and execution of tasks should be separated. The main point of doing this is to simplify the context in which each phase is conducted, so that novices will not have to keep each phase's specific considerations in mind at the same time.

It also provides an intuitive way of structuring feedback and verification, as the initial state of each phase is based on decisions that are made in prior phases. This means that separating the various tasks that can be performed in the interface into multiple stages and thus limiting how much information the user has to keep track of at any one point could be an effective way of taking their mental resources into consideration and ease the use of the product.

The second rule described by Senft et al. (2021), concerns the use of higher level controls. What this refers to is the exclusion of detailed or minor decisions in favor of action level controls. An example in the case of this thesis would be to specify for the robot to go and charge rather than specifying each substep required to reach the charging station, enter and commence charging. Following this rule will further alleviate the mental workload of using the teleoperation interface as well as increase efficiency.

6.2.3 Information display

The way information is displayed in teleoperation is a critical topic. Since the operator is not co-located with the robot, all sensory information that they would normally receive has to be conveyed through the controller interface, which is a much more narrow medium (Endsley and Jones, 2003). Thus, what information is displayed should complement the task at hand so that sufficient situation awareness can be established and the relevant features can be used.

One important rule, highlighted by Endsley and Jones (2003), is to successfully prioritize the right information to ensure that the operator always has sufficient information to make the required decisions. This could for example include showing a map so that the operator can determine which path to take to their destination. Endsley and Jones also discuss one of the main problems related to displaying information, which makes the above rule more difficult to implement, which is that concentrating too much information onto the interface can lead to cognitive overload for the operator. Nielsen et al. (2007) also highlight this problem and describe how integrating different parts of the interface could reduce the demands on attention and support situation awareness without overloading the operator's mental resources. In their study, the video feed and a three dimensional map were combined into a 3-D model which was found to better support teleoperation than a 2-D video feed with a separate map. This meant that the operator, instead of navigating using a first-person view, could see the robot from a third-person view, with a 3-D model of the robot inserted into a simulation of the environment either built from sensor data or preexisting models.

To design a teleoperation interface with high usability, the attention required by the operator needs to be minimized. At the same time, the operator must always have the necessary information to make decisions. Otherwise, the efficient use that is desired in construction cannot be realized.
To reconcile these factors, the interface could be adaptable to the current task, for example by utilizing multiple modes which can contain different information suited to each mode’s purpose. Another approach would be to integrate information as Nielsen et al. did. Using a 3-D model as the main feedback is one way to integrate the information, but other parts of the interface could potentially also be integrated to reduce the attention demands. Endsley and Jones (2003), for example, recommend presenting higher level information rather than lower level. By combining and interpreting different information in the interface, for example by showing a difference rather than two separate values, the operator is not required to perform the computation or split their attention unnecessarily.

One type of information which can be particularly important for teleoperation is direct feedback of movement or manipulation. This is due to the fact that the distance between the operator and the teleoperated robot causes some time delay in the control and feedback, which Chen et al. (2007) highlight as a significant problem for effective teleoperation. To counteract this problem, Chen et al. mention predictive displays, which can give operators direct feedback on their actions even if the real feedback from the robot, which is shown on the video feed, is delayed due to latency. This could for example be a virtual model of the robot which moves as soon as the operator executes the command. Through the use of these predictive displays, a clearer connection between the operator's input and the robot's response is created, something that Nielsen et al. (2003) mentions as a crucial factor in effective teleoperation and as a way to alleviate the operator's mental work.

Singh et al. (2013) also discuss this concept and highlight its usefulness for the telemanipulation of a robotic arm. In addition to placing the control input area in close proximity to the video feed to improve the connection, their study also indicates that using the tip of the robot's arm as a reference point for the manipulation enables intuitive control of the arm. This means that the user only has to consider how they want to move the end point of the arm, rather than moving each joint separately.

To ensure effective teleoperation, even when latency is present, the teleoperation interface should thus utilize some kind of predictor display where possible. This could be implemented both for the driving and the movement of joints and could be especially beneficial for novice users, who might be more unsure in their ability to drive and could use the additional feedback. Creating strong connections between user input and robot action in other places will also improve the usability of the interface, for example by placing the controls near the feedback. Finally, utilizing clear spatial mapping can support the explicitness of the robots controls, in particular when concerning the control of an arm.

6.2.4 General situation awareness

Situation awareness, as described in more detail in the theoretical background, relates to perceiving, understanding and predicting one's current situation (Endsley and Jones, 2003). For teleoperation, the operator's situation awareness relates both to the status of the robot and how the robot is situated in relation to the environment, and can include aspects such as orientation, depth and motion (Chen et al., 2007).

Chen et al. write that, while situation awareness is critical for effective teleoperation, it is often difficult to obtain due to the fact that the operator is not co-located with the robot. Endsley and Jones also highlight this problem, and detail how the sensory information which the operator would normally receive if present physically, now has to be communicated through a comparatively narrow interface.

Doing this successfully is difficult, as many problems emerge when trying to communicate the situation. Chen et al. describe a number of these problems, with limited FOV (field of view) being a primary one. The less the operator sees of the robot's surroundings, the more difficult it becomes to assess the situation and see obstacles. Chen et al. further details how the limited FOV has negative effects on the operator's ability to judge speed and distance, which greatly affects the operator’s ability to drive effectively.

While the use of additional cameras could address the problem, Chen et al. and Endsley and Jones all highlight the problem of visual overload and limited attention and suggest that using other modalities to support the visual feedback could be a better alternative.

As construction sites are very active environments, this places an even greater emphasis on situation awareness in order for the user to be able to operate safely and efficiently. Thus the problems related to situation awareness in teleoperation have to be properly addressed in the interface. Since novice users, whose attention will likely be in high demand while driving, make up a significant portion of the intended user group, the approach of using multiple modalities becomes even more appropriate.

To acquire situation awareness it is also important to have a sense of depth. Chen et al. describes how lacking such sense makes it more difficult to effectively teleoperate a robot since spatial dimensions will be miscalculated. When only a 2-D video feed is available however, as is often the case in teleoperation, assessing depth is challenging. Chen et al. suggest that stereoscoping displays could support depth sense, as they utilize a binocular feed, however they see multiple drawbacks, such as motion sickness. Mast et al. (2015), also highlight the importance of depth sense in teleoperation, and suggest that giving the operator other kinds of depth cues in the interface, for example haptic feedback, could serve the same purpose as the stereoscopic display and provide a sense of depth. Thus providing the operator with some form of depth indicator through the interface will be important to support their effective teleoperation, and this will be particularly important when manipulating the robot's arm. For robots with limited FOV this will also be crucial as it will be even more difficult for the operator to assess depth.

Another aspect of situation awareness which is relevant is what Endsley and Jones refers to as “global situation awareness”. This kind of awareness relates to understanding the situation on a higher level, like what the operator's current goal is and whether a new goal should be set. Optimally, the operator should always have direct access to this kind of information, even if they are focusing on a specific part of the software (Endsley and Jones 2003). This could for example be the robots current battery levels, which informs the operator of how much longer they can drive before they have to recharge. Informing the operator of the global situation awareness becomes even more important when using any autonomous features. If insufficient information is available to the operator while using autonomous features, there is a high risk of them being out of the loop, meaning that the operator does not understand what the robot is trying to do (Endsley and Jones 2003). This can lead to frustration and inefficient use, especially if the autonomous function fails and the operator has to figure out what has gone wrong.

Prioritizing critical information for global situation awareness in the interface will thus be important so that no matter what the operator is doing, they can easily determine the overall status of the robot and their current goal. Beyond battery level, this might include connectivity, current destination and global positioning. When autonomous features are active, the interface needs to clearly indicate both that this mode is active and what the current goal of the robot is, in order to keep the operator in the loop.

6.2.5 Local and global orientation

Chen et al. (2007), highlights orientation as one of the most important aspects of effective teleoperation. This includes both the global orientation, that is, where the robot is positioned in relation to the entirety of the operational area, and local orientation, which concerns the robot’s position in relation to nearby obstacles and paths. Endsley and Jones (2003), also discuss this issue and how it can often be difficult for the operator to successfully orient themselves due to being separated from the robot's location.

In the case of global orientation, Chen et al. highlight a map as an effective solution and details two ways of using a map in teleoperation. The first is a “north-up map”, which functions in such a way that north is always “up”, meaning that even if the user turns the robot, the map's orientation will remain the same. The strength of this type of map is that it better supports the operators planning of their route and global awareness.

The “track-up” map on the other hand, instead of being static, rotates as the user turns the robot, meaning that “up” on the map will instead of north be whichever direction the robot is facing, making this type more suitable for local orientation. Chen et al. further reference how research typically advises using the two map types together, in order to make use of their respective strengths. Endsley and Jones also give examples of how different kinds of overlaid information on the video feed could support the user's orientation, as an alternative to a map solution.

To support local orientation and allow the operator to effectively maneuver around obstacles without falling over, Chen et al. (2007) suggest that being informed about how the robot is tilted is important. They further explain that in environments where clear spatial references, such as planar surfaces or vertical walls, are missing, the tilt can be difficult to determine. In these cases, interface indicators of the robot's tilt could serve in place and reduce the risk of tipping.

This could be taken a step further by using a control interface based on a 3-D model, as those detailed by Nielsen et al. (2007), and Labonté et al. (2008). This type of interface not only includes the robot's local and global position, but also gives explicit feedback of the robot's tilt, as the operator sees a representation of the robot in third-person view. Both Nielsen et al. and Labonté et al. found that this type of interface is better suited for remote navigation than the 2-D, video-based interface type. It does however require accurate 3-D models and large amounts of sensor data to work.

These findings have multiple implications for this project. First of all, the importance of both global and local orientation for effective and efficient use of teleoperation, as well as the problem with acquiring them. As the robot will operate on construction sites, where rough terrain, complicated layouts and varied obstacles will all be common, these aspects become even more important. Information about local and global orientation must thus have high priority in the interface. Using a map might also be an effective way of achieving global orientation. Both types of maps that were discussed could be used, but since the site layouts might be complex and have many similar looking locations, switching between the two modes may confuse the users and thus decrease usability instead of increasing it.

Alternatively, overlaid indicators could be useful, as these keep the information more integrated compared to a separate map, which minimizes the operators attention workload, a previously established problem. Using a 3-D model as the basis for navigation has clear advantages, but due to the visual aspect being such an important part of construction inspections, this approach might still be a worse fit for this type of interface than one based around a video feed. Furthermore, the requirements for accurate 3-D models and robot sensors means that this approach will not be suitable for all solutions.

6.3 Usability test

A series of tests were conducted with the aim of evaluating the general performance and usability of the prototype. In addition, the tests were used to compare different approaches for various features in places where there was no one clear choice. These different approaches were allocated to one of two concepts, which resulted in these concepts becoming different from each other in certain ways, though they were still very similar in their overall appearance and functionality. In total, 6 tests were completed, three of them testing the first concepts and three testing the other. These tests' main results were qualitative findings of observations and comments made by the users in regards to the usability of the interface, but also quantitative measurements of, for instance, how long it took users to complete certain tasks and how many mistakes they made along the way.

6.3.1 Qualitative results

The qualitative findings were extracted from the KJ-analysis which was based on both observations made by the authors during the tests and comments and interview answers made by the users. The main takeaways from this analysis are listed below.

Compass

One of the most frequently recurring issues throughout the tests was misunderstanding the compass feature, see figure 9. Though it was intended as a proximity indicator and directional aid, and users also saw its usefulness in these regards, many also interpreted it as either an indicator of the robot's size in the environment or as the robot itself. Consequently, the users had difficulties in understanding what it was for, and often made incorrect interpretations that led to errors. For instance, those who thought that the visualization of the robot was the robot itself drove into obstacles and didn't understand why they were colliding with any of them. In their minds, the robot was not hitting anything, and so their entire usage was compromised by this one misunderstanding.

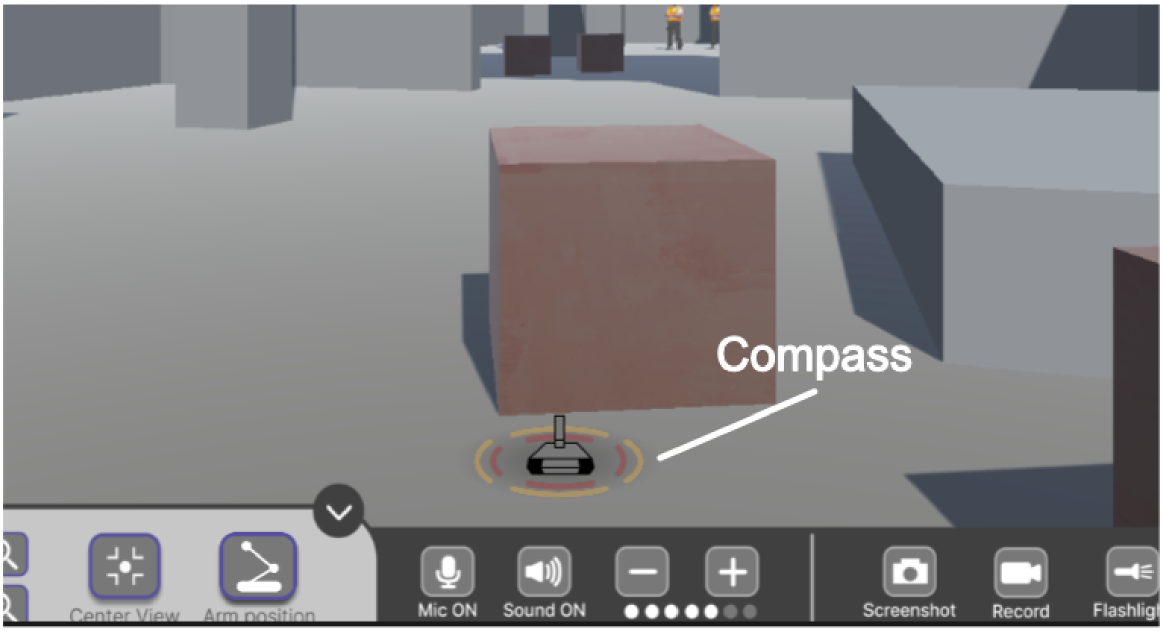


Figure 9: *The commonly misinterpreted compass feature in context.*

One aspect that allowed for this misunderstanding was clearly the similarly low fidelity of the prototype and the test environment. In real life, users would be less likely to assume that a simple model of the robot was the robot itself if the environment were more realistic. In the simulated environment however, it was less obvious that the low fidelity wasn't simply a version of the robot that matched the environment. Had there been a greater separation in the levels of fidelity between the compass and the environment, this issue might therefore have been minimized. Still, it is interesting that users did not reflect on the fact that they would have no way of viewing the robot in third person, despite the fact that they were shown a picture of the robot and had its basic usage explained to them prior to starting the test. In summary, if a compass feature like this one is used in a teleoperation interface, it should clearly communicate its purpose to the user to avoid the occurrence of this error.

Driving

Another problematic segment was the driving, which suffered from a number of features that were either difficult to use or to understand. Every user experienced some issues related to this in the beginning of the test, in which they would collide with obstacles, wouldn't know how to control the robot or didn't understand where to place their attention. Throughout the tests, it did become apparent that users improved significantly as they drove around the test environment, suggesting that many of the issues could be overcome given enough practice or a proper introduction.

One particularly confusing feature for many of the users was the use of the “WASD keys”, which were used for driving and turning the robot. Most users did not have prior experience with using these specific keys, as this convention is most commonly found in modern video games. Due to this, many of them tried to use the on screen arrow keys or camera controls as they were trying to figure out how to get the robot to move. Though the younger users did make the connection after some initial confusion, some of the older users were not able to get past this hurdle without help from the facilitators. Once they had understood which keys to use, the users had no trouble using them for controlling the robot. In terms of the design, this issue points to a lack of explicitness and adherence to the users' existing mental models of control interfaces. The results thus indicate that these aspects need to be addressed to design a teleoperation interface with good usability. However, this issue does not represent any critical fault in the interface design, as the initial introduction to the keys would be included in the first time tutorial, which would resolve this issue. Any error stemming from this misunderstanding would also be unlikely to cause any accidents, due to the fact that users would be unable to move the robot. Making the on-screen driving buttons interactable might also be a plausible solution, although this might lead to less efficient driving.

Management of attention

Many users expressed frustration and confusion regarding where they were supposed to place their attention throughout the test. Initially, users would generally look at the feed, as this is the most prominent feature on the screen and the one that makes the most intuitive sense. However, as the test progressed and tasks appeared which required the use of the top-down view and map tool, users would start to feel overwhelmed as they were forced to switch their attention between three different areas of the screen, as is illustrated in figure 10.

Generally, the three features worked as intended, with the top-down view being used for collision avoidance, the map for orientation and planning of routes and the feed for looking around and inspecting the environment. However, the test results showed that the users were not able to effectively use all three features simultaneously. Instead, when introduced to the top-down view, for instance, the users would often stop looking at the feed all together and only focus on the top-down view. Thus, there is a need for reducing the amount of elements to keep track of so that the user only has to split their attention between two items, while still receiving the necessary information. Furthermore, the users expressed that the map was of little use when minimized, as they did not see enough of the surroundings to effectively plan their paths. Zooming out the map or giving the users control of its zoom function in minimized mode could help resolve this.

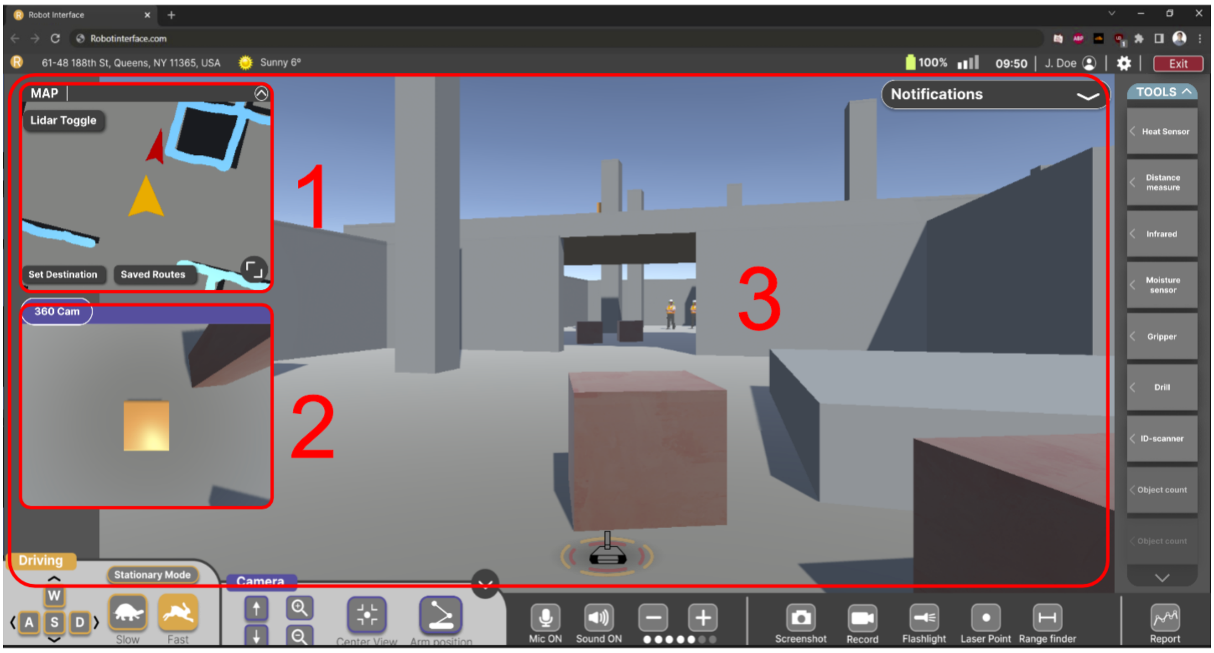


Figure 10: *Illustration of the three navigational elements which users struggled to pay attention to simultaneously.*

In conclusion, limiting the number of primary elements in the teleoperation interface that the user has to pay attention to when driving to two seems optimal for good usability. Solving this through integrating the elements could be a suitable solution, especially if the elements are frequently needed and thus cannot be hidden. For a minimap to have a positive effect, the test results also indicate that it needs to be sufficiently zoomed out to show a greater area around the robot. While the literature had already pointed out the need to limit the attention demands, the test gave further insight into how much the attention demands have to be minimized and which elements should be prioritized in this regard.

Camera and arm control

The two concepts which were evaluated in the tests contained different ways of being able to maneuver the camera, something which had not been thoroughly addressed in the identified literature. The first concept only allowed for panning up and down, which meant that the robot itself would have to turn in order to look at something to the sides. The second concept, however, allowed for both panning and tilting of the camera, which was intended to provide the user with more freedom in how they could inspect the environment.

The problem with this approach became apparent very quickly though, when users moved the camera to look around and then proceeded to start driving without re-centering the camera. This, in some cases, led to a complete collapse of the user's ability to control the robot, as the output from the screen no longer matched the inputs the users were giving. Most users did not understand why this problem occurred, and were generally not able to resolve the issue without help from the facilitators. The test prototype did contain a feature for centering the camera view, which was located in the center of the screen. However, this feature was not explicit enough, as no user was able to track it down and understand that it was a solution to their problem.

Finally, one of the last tasks of the test concerned moving the robot arm close to a wall in order to get an effective reading with an attachment placed next to the camera. Though users were generally able to find the arm controls and use them without issues, many commented on the fact that they would have had no idea how close they actually were to the wall in real life and that they subsequently would not have felt comfortable conducting the task. Thus, the interface must provide the user with information regarding how close they are to surrounding objects at any given point when using the arm, and reassure them that they can not hit these objects even if they use the controls recklessly.

6.3.2 Quantitative results

The tests were divided into 5 distinct tasks which, depending on whether the task contained driving or menuing, took place in one or both of the Figma and Unity prototypes. The tests were evaluated based on completion times and number of serious errors, including going severely in the wrong direction, hitting an object or trying to press a button that was not meant to be pressed. The quantitative measurements highlighted patterns relating to the users and the usability of the two concepts, the main takeaways of which are presented below.

The first takeaway was that younger users completed the tasks faster than older users. On average, the younger users performed the tasks 23% faster than the older users. The young users were all in their twenties and the older users were all over the age of forty. This supports the concerns the users had during the initial user study, where they stated their belief that their older colleagues would have more trouble operating the robot than younger ones. As some of the intended users will be older, this puts extra emphasis on the simplicity of the interface and ensuring that the user is sufficiently assisted when operating the robot.

The second main takeaway was that concept 1 was easier to drive, with users on average completing the driving tasks 27% faster for this concept. The average number of driving-related errors were also lower for concept 1, with users on average committing 59% fewer driving-errors than those who used concept 2. Though the small sample size does again limit the conclusiveness of the results, this finding does support the conclusion from the KJ-analysis that concept 2, due to the independent camera setup, is more difficult to use. The results thus indicate that to achieve high usability, a teleoperation interface should not allow the user to pan the camera independently of the robot's rotation, or at least not without significant measures taken to prevent errors stemming from this feature.

7

Design Description

The final design is presented in this chapter, along with explanations for the design decisions and how they meet the identified requirements.

7.1 Layout

In this section, the interface's features are introduced and their placement and proportional size is motivated. The entire layout in its intended normal mode can be seen in figure 11. Figure 12 shows the interface with some of the features expanded.

Figure 11: *The main page of the interface.*

For the layout design, a number of aspects were considered. Firstly, adherence to existing interfaces and web applications was an important factor. Although no exact equivalent to the product exists currently, many similar applications use principles and solutions that are relevant to any interface. Secondly, the flow from users' input to the interface's output was considered in order to ensure that the design would remain consistent with the users' mental models of where they expect information to appear. Lastly, visual clarity and appropriate prioritization of features were the primary considerations when determining the relative size and overlap of modules, both within modules and when considering the interface as a whole.

7.1.1 Feed

The camera feed from the robot's full HD camera (1920x1080p), positioned at the end of the arm, is the interface's primary way of providing the user with information about the robot's environment. It is also a key tool for conducting many of the tasks in the various inspection scenarios users are likely to find themselves in, such as looking at a particular feature of the environment, taking pictures or recording video.

Because of its central role in the usage of the product, the feed takes up a large part of the interface and is positioned in the center of the screen, following the guideline of prioritizing functions and information. This provides the user with the best possible conditions for viewing the environment, and also enforces the feed's importance and key role in the intended usage of the interface.

Its placement in the center of the screen adheres to the users' expectations based on their experience with driving other vehicles, and solidifies the fact that the robot is a vehicle with similar considerations and limitations as other vehicles. If the feed had been positioned in a corner or had been made proportionally smaller, the users may have interpreted the robot as more of a stationary tool, not one where driving is an integral part of the usage.

7.1.2 Compass

The compass, which acts as a directional aid as well as a proximity sensor, is located centrally at the bottom of the feed. Since both its directional arrow and the proximity data relate to the robot's positioning, placing it to the side would make the output less intuitively linked to the user's sense of their positioning. Its proximity to where the user will look during driving also reduces the need to split attention.

7.1.3 Controls

The controls for both the robot and camera movement are located in the bottom left of the screen. Though there is technically no need to have controls on the screen, as the user can and should use the keyboard to control the robot, the inclusion of on screen controls helps to inform the user of what kind of control is possible and which keys should be used. It also provides redundancy in case the user doesn't want to use the keyboard for whatever reason. The controls are grouped to facilitate consistency in the interface as the user will always know where to look for controls, whether they relate to driving the robot or moving the camera or arm. As these are the most important parts of the interface, besides the video feed, this section is also larger than the rest of the bottom bar and has a separate color. This supports the visual clarity of the interface as well as the prioritization of the critical functions.

Placing the WASD-keys in the bottom left of the screen also provides a natural flow from input to output, as the user's attention moves from their left hand which presses the keys through the bottom left controls panel and into the feed.

7.1.4 Quick access

The quick access features are located in the bottom of the screen, extending from the middle to the rightmost part of the screen. These features are ones that the user is likely to want to use regularly, and would therefore prefer to have readily available. The placement and scale of these features is intended to resemble those of online meeting platforms, since the purpose and type of features are similar and the users already have established mental models of how these kinds of panels function and what they look like.

Together with the controls, the quick access features form a bottom section which is confined to general features that should be actively used and which have a direct impact on the robot's status. In contrast, the upper panel contains features which relate to more indirect settings and passive status displays, while the sides of the screen are reserved for more specific features. Separating segments in this way provides an intuitive way of locating specific features and information.

7.1.5 Tools

The tools section, which contains controls and settings for attachments that may be added to the robot, is located on the right hand side of the screen. As these attachments are not part of the basic toolset and are not required for most tasks, they are separated from the more general features at the bottom of the screen. The tools are kept relatively small and neutrally colored, as they are not something the user should put attention on unless they require a specific tool.

7.1.6 Map

The map tool, containing a host of features related to providing the user with situation awareness, global positioning as well as automatic navigation, is located in the top left corner. Though there are many cases of maps being placed in other corners than the top left one, there is no clear standard choice for interfaces at large. The map's placement in combination with its relatively large size provides the user with an effective way of receiving information about the site and immediate surroundings, thereby increasing situation awareness and alleviating the planning and driving of routes.

7.1.7 Top-down view

The top-down view tool, consisting of a merged view of the robot's four fisheye cameras, is located below the map on the left side of the screen, due to its association with the map tool and navigation. Their close proximity means that it will be easier for the user to keep track of both tools simultaneously. The tool is intended to be minimized for much of the usage, which prevents its placement from interfering with the feed in most cases. It also lessens the attention load of the user, as the number of elements to keep track of on screen is reduced.

7.1.8 Notifications

The notification log is located in the top right of the screen. As with many applications, the notifications first appear on screen independently of the log itself. After the user has seen them, they're able to access all notifications in the log and go back to review their session. The notifications appear directly below the log. Though there is some merit in having notifications appear close to the feature that the message is associated with, such as having notifications regarding automation appear near the map, the association with and reinforcement of notifications being saved in the log is deemed more important.

7.1.9 Status indicators, account and settings

The remaining features, various status displays, settings and profile options, are located at the top of the screen. This generally adheres to most applications the users are likely to be experienced with, and places these less critical parts of the interface away from the more integral features.

Figure 12: *The main page with expanded top-down view and tools panel.*

7.2 Primary navigational feedback

The primary feedback that the operator will use for navigation is the video feed from the camera mounted on the end of the robot's arm. This is the main tool for viewing the remote environment and determining how to drive to successfully reach the destination without hitting obstacles. As the most important aspect of the users' work is related to visual inspections, the video feed will also play a critical role in most of the inspection use cases, which its size and placement reflects. The reason why a pure 2-D video feed was chosen over an integrated 3-D model, like those described in the literature study, was due to the importance of the visual feedback for inspections and due to the present sensors on the robot being insufficient to make proper use of a 3-D based interface.

With a 2-D video feed, the real visual feedback from the sites will be a lot more prominent and clear than it would be with feedback based on a 3-D model. Thus, the 2-D video feed is more aligned with the users' needs and follows the usability guidelines of appropriately prioritizing information. As the robot is only equipped with 2-D lidar, it would not be able to generate a 3-D model of sufficient fidelity to work as the basis for a 3-D model interface. Using a pre-built model would also not work, as the GPS on the robot would not be accurate enough to align the 3-D model of the interface with the real world site. Thus the hardware does not support the use of a 3-D model as the main form of navigational feedback.

To maximize the size of the video feed, some parts of the interface had to be placed slightly overlayed. While keeping the feed clear of overlapped elements would be optimal, this was deemed as an acceptable solution. The map, which normally occupies the most space on top of the feed, is positioned in the top left corner which is not a place where the user will normally need to look, as they will tend to keep focus on the center of the feed. Furthermore, when the map is fully expanded, the user is not supposed to drive and thus the increased obscurance is not a problem.

The compass, which is also positioned as an overlay to the feed, is mostly transparent and its proximity to the video feed and where the user is normally looking is meant to partly integrate these two features. This in turn reduces the demands on attention when driving.

7.3 Controls

Possibly the most critical part of the interface, in terms of usability, is the controls which are used to drive the robot as well as manipulating the arm and the camera, see figure 13. These three control setups will be presented below.

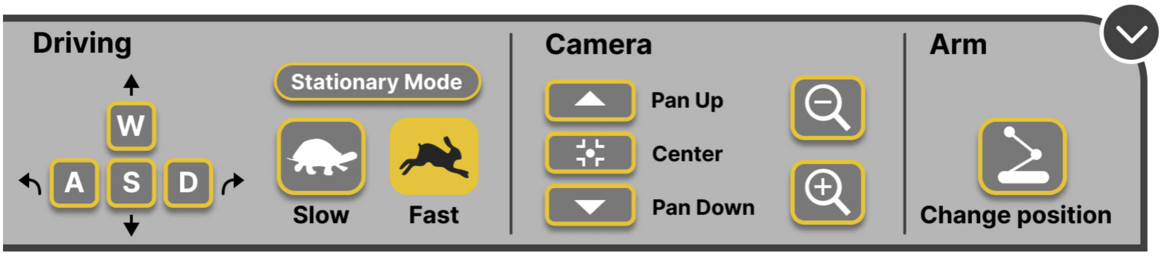


Figure 13: *Closeup of control panel in expanded mode.*

7.3.1 Driving controls

For driving the robot, a number of different hardware solutions were considered. Many options provide a greater sense of control than using a mouse and keyboard, such as having a physical joystick with which to maneuver the robot. Other options, such as drone controllers, possess additional potentially useful features, such as touch sensitive scroller wheels and joysticks for controlling the robot in multiple degrees of freedom. For the present application however, the decision was made to make mouse and keyboard controls the standard way of controlling the robot. The pros of not requiring users to have any additional material than their laptop, in combination with their familiarity of using mouse and keyboard, makes this approach preferable. For the future product, there is nothing limiting users or designers from adding additional ways of controlling the robot, should this prove to be beneficial.

The keys for operating the robot were chosen to be W, A, S and D, hereby referred to as the WASD-keys, which are commonly utilized in various web applications and video games for moving around. The reasoning behind this choice of keys mainly relates to wanting users to have their right hand free, so that they can use the mouse simultaneously and not have to keep switching back and forth. In order to get users to understand which keys to use, as well as the fact that they should use keys in the first place, buttons corresponding to the WASD-keys are present on the screen. The keys are arranged according to their placement on the keyboard, as well as being placed on the left of the screen - all of which aims to solidify the connection in the users' minds.

Next to the buttons are arrows which indicate how the robot will move or rotate when they are clicked which, combined with the "Driving" label, aims to increase explicitness. Though this introduces a risk of users confining themselves to only using the on screen buttons to drive the robot, this will be mitigated by including a short tutorial for first time users which presents the keyboard as the intended medium for controlling the robot.

Next to the WASD-keys, there are two buttons for controlling the speed of the robot, one for fast mode and one for slow mode. The buttons toggle on or off as the user presses either of them and lights up to indicate which mode is active at any given point. Though a slider would provide more detailed control, the simplicity of only having two modes alleviates the usage by eliminating uncertainty regarding which speed is appropriate. For the vast majority of use cases, there is now a clear choice between the two modes, with fast mode being suitable for going long distances and slow mode for tight spaces or detailed maneuvering. The icons featuring a hare and a tortoise also provide a symbolic connection to what the buttons are intended for, unlike the less explicit appearance of a typical slider. In the case of a typical inspection, the user will utilize fast mode to get to the area of the inspection, with the option of turning on slow mode for avoiding particularly tricky obstacles, and then use slow mode for getting close to inspection elements and for framing the camera feed.

In addition to being able to change the speed of the robot, there is also a button for engaging stationary mode. This feature acts much like a hand brake, with the intention of being activated whenever the user plans to stay still for a prolonged period of time or when they are concerned about accidentally pressing one of the control keys. Throughout the studies, it became clear that most inspections featured an element of particular interest which the users were likely to spend time analyzing and discussing with colleagues. In stationary mode, the user has access to all the tools that are relevant for static activity whilst also being certain that the robot won't move unexpectedly.

Once users have understood the controls and familiarized themselves with the panel, it can be minimized to take up less space on screen, see figure 14.

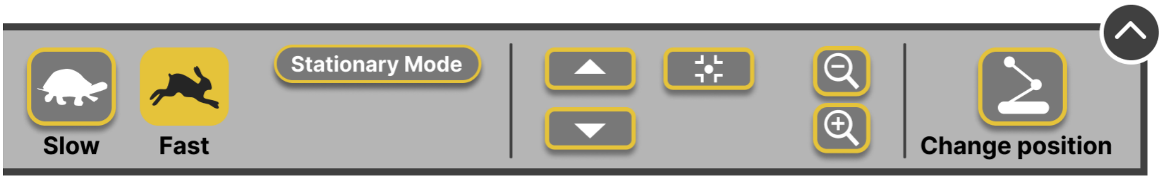


Figure 14: *Closeup of control panel in minimized mode.*

7.3.2 Camera controls

The camera is mounted at the end of the robot's arm and has only one degree of freedom, tilting. To pan, the user therefore needs to rotate the entire robot using the A and D keys. As was evident in the usability testing, allowing for panning of the camera independently of the robot causes a lot of confusion as users are unable to understand what is happening or recover from the situation. Simplifying the usage by excluding in-camera panning therefore removes a major hurdle whilst still allowing for panning through the robot's rotation. This significantly reduces the risk of errors and increases the effectiveness of the interface.

To tilt the camera, the arrow keys and a set of on screen buttons corresponding to the keyboard keys are used. The rationale behind placing keys on screen is similar to that of the WASD-keys, with the main purpose being to hint at which keys should be used for what control action. To zoom in or out, there are two buttons located next to the arrow keys in the control panel, which allows the users to inspect objects that are far away or difficult to reach.

Between the up and down keys on the screen, there is a button for centering the view of the camera. Though this feature would have played a much more central role if in-camera panning had been included, it still serves the purpose of giving the user a way to return to a standard position after they have titled the camera. The inclusion of this feature therefore removes the need for users to meticulously try to tilt the camera back into a forward facing position, and simplifies and elucidates the intended usage.

In addition to the keys and on screen buttons, the user is also able to drag the camera view with their mouse. This feature is mainly intended for stationary use, when the user might feel it is more intuitive to drag the view in order to look around instead of being forced to use the more mechanical motion of the arrow keys. Zooming can also be performed using the scroll wheel. Users may also be more familiar with such motions based on similar controls found in Google maps, for instance. Before being implemented into the actual product, the effects of delay caused by the distance between operator and robot would have to be examined, in order to determine if the dragging could function in a satisfying way.

7.3.3 Arm controls

The robot is equipped with an arm which can be manipulated in order to position the tools mounted at the wrist appropriately. While this function will be very important for some types of inspections where the required tool needs to be placed in close proximity to the object of interest, it will not be necessary for the majority of the inspections. This is due to the fact that most inspections rely on basic visual data, like simply observing an area or a section of a building, and thus the arm controls do not need to be highly prioritized in the interface. Furthermore, it became obvious during the usability test that manipulating the camera or arm often resulted in the user forgetting their new camera position and trying to drive as if they were in their normal position. Thus, it is better for the user to have a standardized setup when driving, which means that the manipulation of the arm should be limited to the situations where it is actually needed, and the controls can therefore be hidden by default.

The arm controls can be accessed by clicking the button with the same name in the control panel. By having all forms of physical robot control located in the same space, consistency is supported and the user will more likely associate the arm control with its location. When clicked, the arm controls appear above the compass, in the center of the feed, see figure 15. The position of the controls is meant to create a stronger connection between the users' input and the feedback, as they will see the physical result of their input at the same location as they click. This, however, means that the controls will somewhat obscure the highly important video feed. To address this issue, the controls are somewhat transparent and will fade even further when a directional button is pressed, letting the user see the feed clearly and know when an appropriate position has been reached.

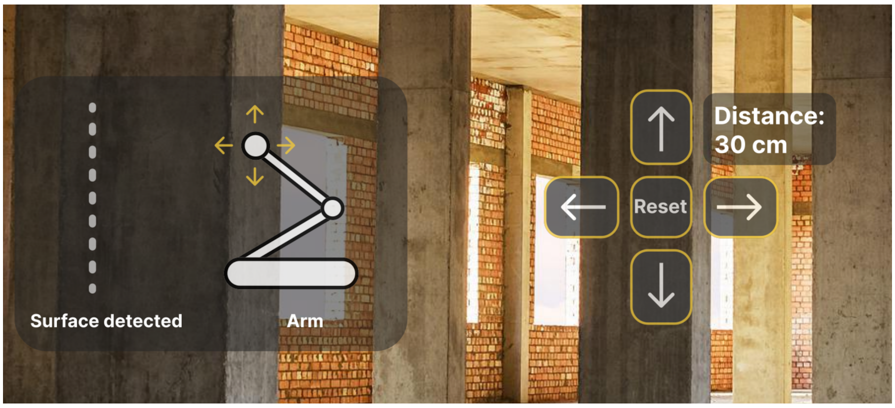


Figure 15: *The arm controls on top of the feed.*

The actual controls consist of four directional buttons representing the arms movement forwards, backwards, upwards and downward, as well as a button for resetting the arm’s position. A 2-D representation of the arm is located in a window to the side of the controls with directional arrows at the end point, indicating how the buttons change the position of this end point. This utilizes the same type of spatial mapping Singh et al. (2013) used in their study and it means that the user only has to think about how they want to move the arm’s end point. This control setup helps even inexperienced users to control the arm, as shown in the usability test. When the user presses one of the four available directional buttons, the arm begins to move in that direction. The user can see the arm moving forward as the feed changes and once they let go of the button, the movement stops. This way, the user remains in control during the operation and can precisely position the arm, something that was appreciated during the user tests. The reset button also supports the user's control and acts as error recovery if they were to position the arm incorrectly.

Since the camera which provides the main video feed is positioned at the end of the arm, the arm itself can not be viewed by this camera. This means that the only feedback the user will have is seeing things in the video feed move as the arm is moved. This might not always be sufficient feedback and therefore, the 2-D model of the arm positioned next to the arm control buttons will always match the pose of the real arm, as illustrated in figure 16. It will thus act as a predictor display, meaning that the user will get immediate feedback on their input, even in the presence of latency.

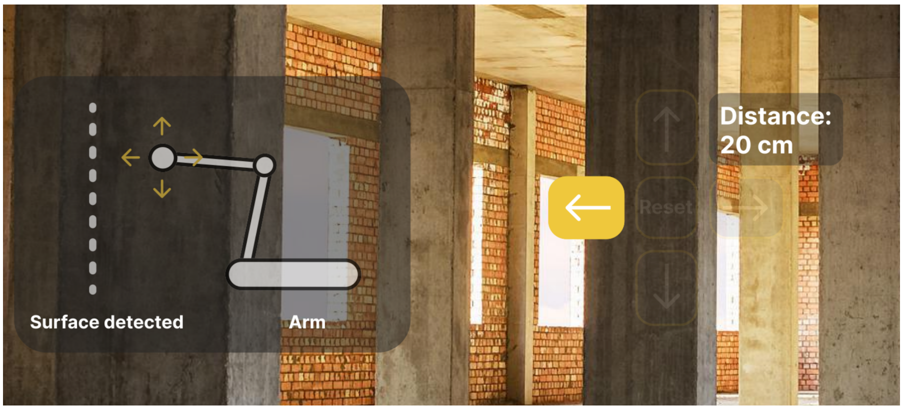


Figure 16: *The arm controls during arm movement.*

To support spatial awareness, specifically the sense of depth, a depth indicator will be shown next to the arm controls. This will enable the user to more accurately place the arm and also ensure that it can not be bumped into objects. Furthermore, a visual indication of the distance between the arm and any surface in front of it will also be displayed next to the 2-D model of the arm.

7.4 Quick access

Throughout inspections, there are a number of tools that are needed to complete various common tasks. These features are grouped into the quick access panel, which is readily available for users at all times, see figure 17. The panel functions both as an encouragement to use the tools that are there, whilst also discouraging the user from using the more specialized tools and advanced settings for their everyday tasks where they are for the most part not needed. Thus, the common use cases are made simpler and the commonly used functions are prioritized. To increase their explicitness, the buttons are outfitted with both names and icons.

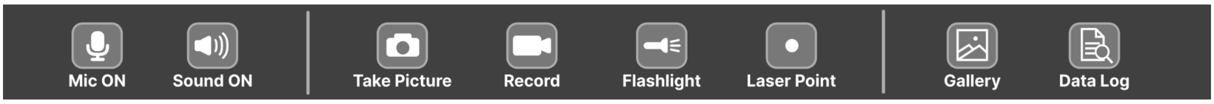


Figure 17: *The quick access panel.*

Going from left to right, the first quick access tools are mute buttons for the microphone and speaker of the robot. Since users need to communicate with on-site personnel at certain times as well as potentially wanting to use sound from the environment to get a sense of what is happening, the sound must be quick to unmute. This would also support quick communication with a hoist operator to allow the user to move between floors. In other cases, users may want to mute their microphone to not let sounds from their office reach the site, or mute the speakers to prevent distracting sounds from the site reaching them.

Next to the mute buttons are the buttons for taking pictures and for recording video. These features are critical parts of many inspections, as users often want to take pictures or video for sharing with colleagues or cataloging their visits. Next to these are the flashlight and laser point tools, which can be toggled on to provide better lighting or a laser pointing at whatever the camera is looking at, allowing users to interact with the environment and support effective communication. These features are mainly useful for taking pictures in better lighting or highlighting a specific point of interest to colleagues. However, the flashlight may also be used continuously whilst driving in case the environment is darkly lit.

To the rightmost of the panel are the buttons for accessing the gallery and data log tools. The gallery contains pictures and videos and is intended to be used as a tool for reviewing and comparing data from sessions. The data log contains measurements and metrics from attachments, the robot’s components and the user interface. This feature is not as critical as the other quick access tools throughout a session, but is included due to its usefulness at the end of sessions.

As became clear from the user study, having a way of managing or exporting your data is something the users expect from an interface such as this. The approach that was selected for the proposed design is to simply export all data from each session into an external software. Users will be notified upon exiting the interface that their data has been saved. For managing data during inspections, the data log will feature a basic structure for viewing and sorting entries.

7.5 Tools

The tools section on the right hand side of the screen is intended for attachments and their corresponding features, see figure 18. There are opportunities on both the arm and body of the robot to attach different kinds of sensors, tools and components that are suited to specific applications. This way, users can adapt the robot and its interface to suit their needs.

Figure 18: *Main page with expanded heat sensor tool.*

The tools section is separated from the controls and quick access tools at the bottom of the screen. Since the attachments are not part of the basic usage, general users should not have to use this section. Thus, by making it hidden by default, the basic and advanced features are clearly separated and the basic view does not become overly complicated. For specialized users or for companies that rely on very specific tools for their everyday operations, the tools section contains all the controls and settings for the relevant attachments.

After the tool section is opened, the individual tools and their names are shown in miniature and once pressed, a window appears, showing all the associated information with the clicked tool. This means that the relevant information needed for each task will be available when it is needed. Users then have the option of dragging the boxes onto the screen if they need to keep the window open for a prolonged period of time or change the order they appear in to better fit their needs.

7.6 User profile / settings / extra

As with most online applications, especially those requiring authentication to access, a profile feature is included in the interface. Since the way in which users will gain access to the interface is not yet determined, it is unclear how central this feature will be or what the requirements surrounding its functionality actually are. For the application of construction site inspections however, some conclusions can be drawn. Users are likely to have specific roles on site which determine their needs in the interface. Electricians, for example, may want certain attachments related to their work to be available on screen at all times. Providing these users with their own profiles with customizable settings is a way of alleviating their work and making their usage of the interface more satisfying.

Aside from profiles and those settings that relate to different users, there are also general settings for the interface at large. These settings concern options for controls, language and visibility that are not critical for everyday use, but which users may want to access at certain times.

7.7 Compass and top-down view

To support local situation awareness, both the top-down view and the compass feature is included in the interface. The top-down view uses cameras mounted on each side of the robot to create a live feed of what the robot and its surroundings would look like from above, as seen in figure 19. This view provides the user with a good understanding of their immediate surroundings which can be very useful when navigation in tight spaces, something which was observed in the usability tests. The top-down view will however not be active by default. Instead, the intention is for it to be opened and looked at during specific situations that require extended precise maneuvering around obstacles or narrow spaces. This is because much of the literature highlights the importance of limiting the number of things the user has to pay attention to during teleoperation, something which was also verified in the usability tests.

Figure 19: *Closeup of top-down view.*

To further address this issue, the compass feature will serve as a simplified version of the top-down view, integrating some of its information with the video feed. This way, the user only has to look at the feed during normal driving, using the overlaid compass tool to keep track of the robot's local surroundings through the proximity sensors, which should provide sufficient local awareness in most situations. The appearance of the compass is a simplified model of the robot surrounded by faded orange and yellow proximity indicators, as seen in figure 20. The simplified model is meant to disconnect the tool from the realistic environment in the video feed to avoid users thinking that the compass model is the real robot viewed from a third-person perspective, which was a recurring problem in the usability tests. Text underneath the compass also indicates that it is a proximity sensor and not the real robot. This appearance is inspired by the dashboard interface for reversing that is found in some modern cars, something the users might be familiar with. If an obstacle is approaching, the user can simply glance towards the compass and see if the indicators are lighting up and they will then know how close they are. The outer orange indicators will light up first and the inner red ones when the object is closer to the robot.

Figure 20: *The proximity sensor far from, close to and within touching distance of an obstacle.*

Thus, both the object's position relative to the robot and its proximity can be easily ascertained. The additional situation awareness gained from the compass is especially important when only one camera feed with limited FOV is available to the user. With a better understanding of the robot's surroundings, the user will gain greater control and be able to drive more safely even when taking shortcuts, for example when driving through a narrow passage.

Besides the proximity indicator, the compass has a secondary function which is the reason for its name; a compass arrow which points towards the user's current destination. Using the map, the user can set a destination which, when confirmed, makes an arrow appear on the compass which points towards this destination. As the robot is rotated, the arrow will remain pointed towards the destination, allowing the users to acquire global orientation. As this would otherwise require the user to observe the minimap, this feature partly integrates the minimap into the video feed, which eases the users attention workload by minimizing the number of features the users has to keep track of while driving, at least most of the time. Thus after making multiple turns in a row and being unsure of which way leads to their destination, the users can immediately identify their global orientation relative to their destination and choose the right direction.

7.8 Map

To support the users' global orientation and awareness, a map will always be present in the interface. Both literature findings and the users themselves pointed towards the usefulness of a map, especially for those who are less familiar with a site, and it will therefore remain present in a minimized mode regardless of what the user is doing. This ensures that global orientation can always be achieved.

The map, as seen in figure 21, shows the teleoperated robot as a blue arrow, indicating which direction the user is facing. Any hoists and other items of interest present on the site will be marked on the map so that the user can quickly find them. The current floor is also displayed at the top, as this is not something that users are always instinctively aware of. A button in the corner of the map window will expand the map, allowing the user to more clearly see the site's layout and plan their route.

Figure 21: *The minimized map-tool, with and without set destination.*

By clicking anywhere on the map, a red marker appears indicating that a new destination has been marked, after which this location name appears at the top of the map window next to “current destination:”. This gives additional confirmation to the user that the destination has been set correctly and that they now have this information present to remind them of their current goal, thereby supporting their global situation awareness. The colors of the user arrow and destination marker were chosen as these are commonly used in other map- and GPS softwares. Once the destination has been set, the red marker remains on the map and a preliminary path to the destination is generated on the minimap, giving users tips on how to reach the destination. While this path might not always be completely accurate due to the changing nature of the sites, it will still provide a general direction for the user.

The minimap also has buttons for zooming in and out on the site layout. This was a need identified in the usability test and will allow users to better plan their routes when in minimized mode if they encounter a complicated part of the layout. As the usability test did not generate any contradictory findings, the solution suggested by Chen et al. (2007) will be utilized, meaning that when the map is expanded it will be in north-up mode to facilitate better route planning, and in minimized mode it will be track-up to support direct navigation. An indicator of the northern direction will be present to connect the two views and help the user understand how the map is rotated when the map mode is changed. The expanded map can be seen in figure 22.

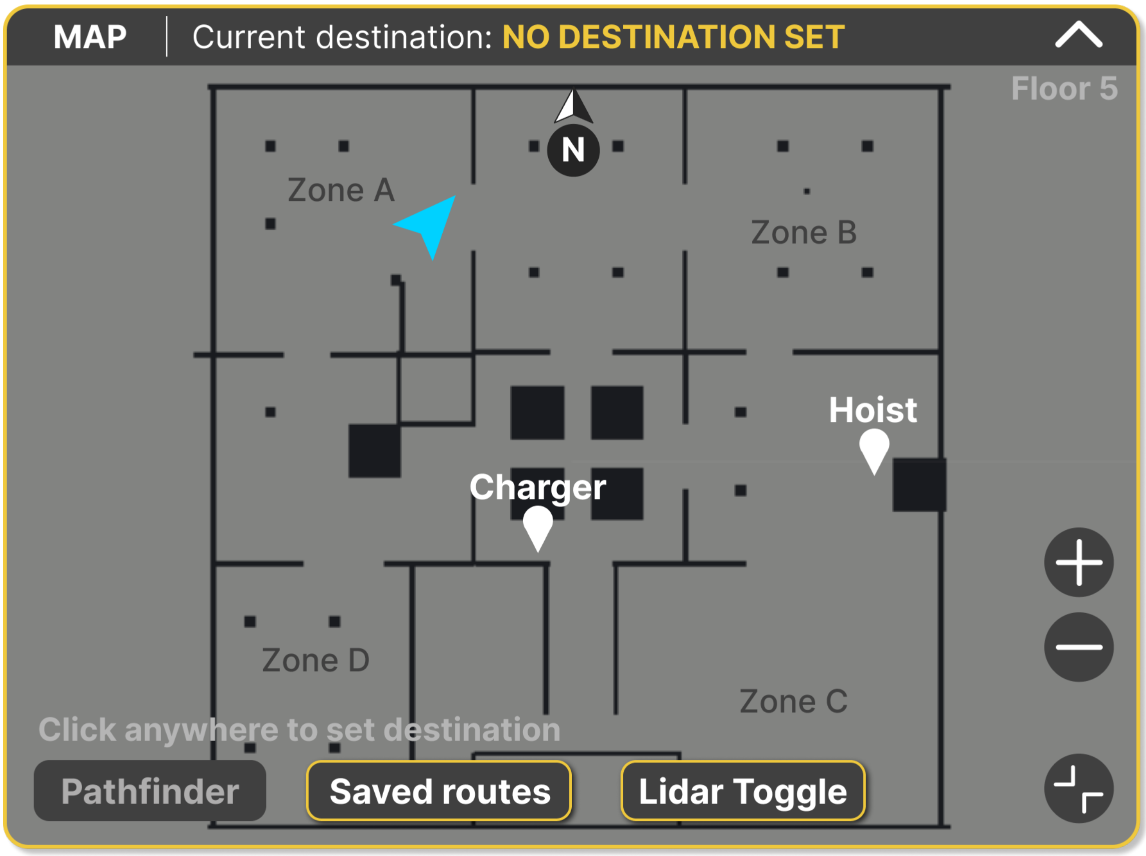


Figure 22: *The expanded map tool.*

To complement the BIM-based map, the robot's lidar system will be used to create a secondary map which can be toggled as an overlay on the BIM-based map. The lidar-based map will only contain outlines of walls and other obstacles which are sensed by the lidar sensors. Thus, if there are discrepancies between the BIM-based map and the real world, the lidar-based map will highlight this, allowing users to navigate more efficiently and avoid accidents. While this secondary map can be useful to increase accuracy, it will inevitably clutter the map module and might not be as necessary once a route has been planned, and can therefore be toggled on or off.

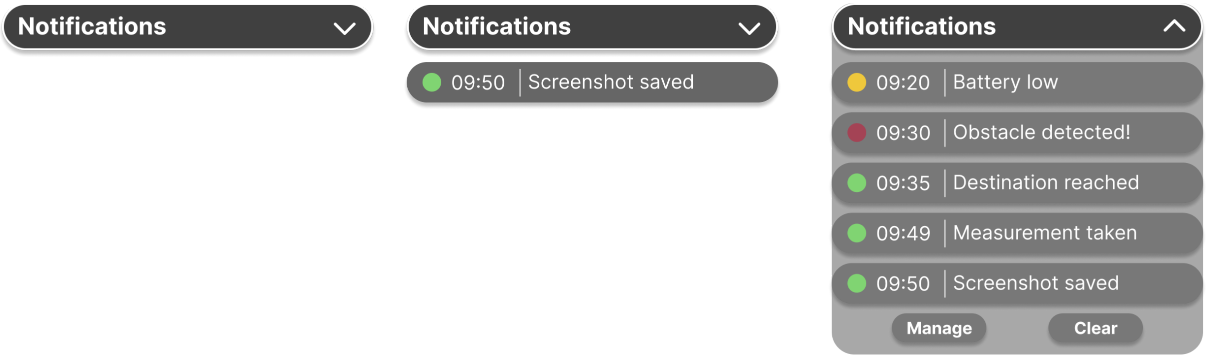
7.9 Status indicators

At the top of the interface is a bar containing basic information about the robot which is meant to inform the user of its current status and provide situation awareness on a higher level. This information will include the current time at the site where the robot is located, which could be useful when the robot is located in a different time zone, along with current battery- and connectivity levels. These features are meant to give users a general sense of the robot’s status which can help inform them of what they have time to do before the battery runs out or the site closes. However, they can also help the users understand why problems arise, for example by indicating a drop in connectivity which explains the increase in control time delay. Thus, the awareness of the robot’s status created by this information can both help the user plan their usage and resolve problems during their use, both of which will help maintain control throughout their usage.

7.10 Notifications

Notifications are a useful way of conveying relevant information about certain events to users in a standardized way. Though many features of the interface have their own ways of highlighting issues and providing feedback, the notification log is the primary source of feedback for the user and supports situation awareness, both of the robot's status and its actions. For all features in the interface, notifications may appear as a result of some important event or action taken within those features. For instance, notifications confirming that a picture has been taken, that the robot's battery level is low or that an obstacle has been detected in the robot's path may appear as their associated event occurs. The purpose of these notifications is, in other words, to provide the user with important information and to verify actions and events.

To adhere to established norms of interfaces, the notifications are gathered in a log which folds down from the top right corner of the screen. Users are able to access all notifications from their session along with time stamps and other information, as seen in figure 23. This allows the user to review their session and gives them an additional way of understanding the product and their usage of it. When clicked, notifications lead to whatever part of the interface they relate to. So for instance, a notification verifying that a picture has been taken leads the user to the corresponding picture in the gallery tool, and a notification about low battery levels leads to information and instructions regarding how to resolve the issue. This increases efficiency and aids the user's understanding of different features and the interface as a whole.



09:50 | Screenshot saved

09:20 | Battery low
09:30 | Obstacle detected!
09:35 | Destination reached
09:49 | Measurement taken
09:50 | Screenshot saved

Figure 23: *From left to right; The minimized notification log, the notification log with a pop-up notification and the fully expanded notification log..*

While all notifications will appear as described above, the most critical ones which require direct action from the user, will also appear as larger pop-ups in the center of the feed. This ensures that the users are made aware of information and feedback that is critical for the safety of the robot and people on site.

7.11 Automation

In addition to being able to drive the robot manually, users also have access to certain automatic features. These are separated into the pathfinder tool and the saved routes tool, both of which are located in the map module, and allow the user to utilize higher level commands to control where the robot goes. Due to their close association with the map and its features, along with the desire to not indicate to users that automation is the primary mode of operation, the modes are not featured as their own sections on the main page of the interface. The current extent of the autonomous feature requires the users to remain present and oversee the robot's actions in case a problem occurs or an obstacle is encountered. While this does not completely free up the robots transportation time for the users, it does alleviate the navigation task. In this present concept, there is no way of enforcing the user's presence in front of their computer, which would have to be considered for the future development of the interface.

When activated, a border will appear around the video feed, along with a panel at the top of the screen, as seen in figure 24. This way it is clearly indicated that autonomous mode has been engaged. The panel displays the destination of the autonomous mode which ensures that the user is always aware of the robot's current goal, which supports their global awareness and keeps them in the loop. Inside this panel, a large stop-button is also present to ensure that the user remains in control and can quickly stop the robot if a problem occurs. A visualization of the planned route is shown on the map to give more feedback about what the robot is doing.

Figure 24: *Interface when an automatic route is in progress.*

The pathfinder tool is a general feature for point to point autonomous navigation. The user can select a destination on the map and then have the robot navigate there autonomously. The way this is intended to work in reality is by converting BIM data into usable paths using existing software solutions. The robot can then figure out what the best route is and suggest it to the user, which means that planning and execution are separated and the user's cognitive load is thereby reduced. An example of this suggested route can be seen in figure 25.

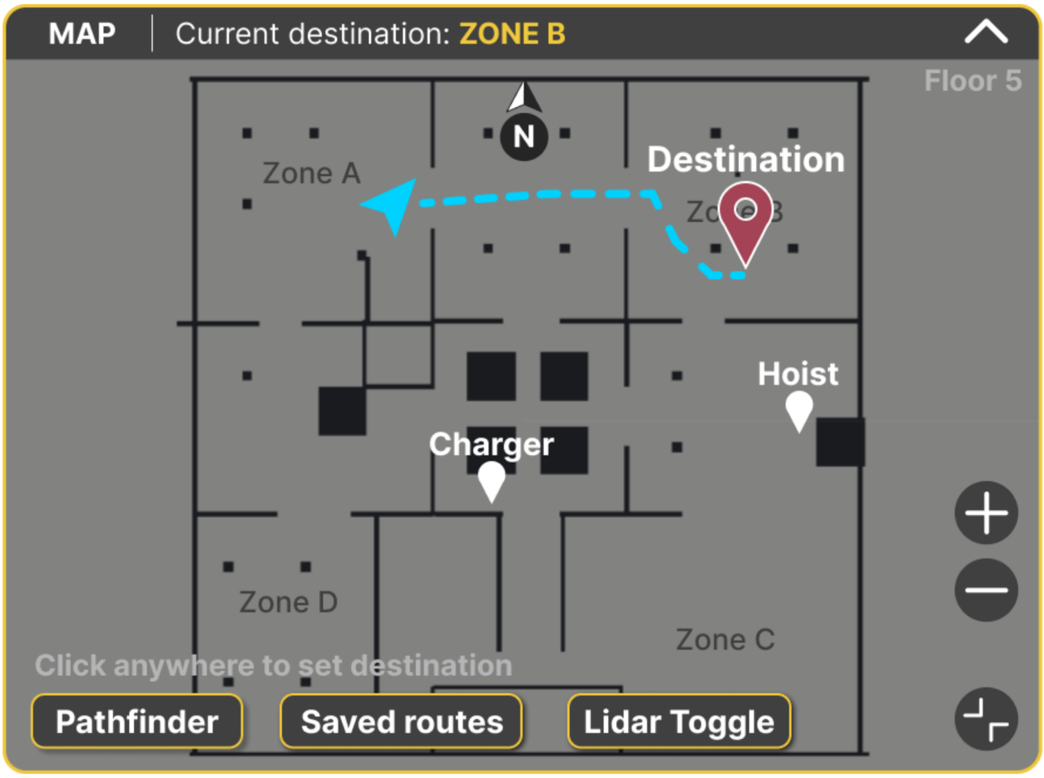

Current destination: ZONE B

Figure 25: *Expanded map with a destination set.*

The saved routes feature will function much like the pathfinder tool, but instead of requiring users to select a destination on the map, they can instead select a preexisting route from a list in the map tool, see figure 26. These routes are ones that the users themselves or others in their organization have driven manually and then saved as a route. For many inspections, this type of navigation is preferable, as it removes the need for users to drive manually and allows them to complete their tasks faster whilst being able to focus on inspecting the camera feed and using tools.


Current destination: NO DESTINATION SET

SAVED ROUTES

Safety round floor 5
Saved 3/4-2022 400m

To Hoist on floor 4
Saved 1/4-2022 200m

Security Round
Saved 5/4-2022 500m

Figure 26: *Expanded map with "Saved Routes" feature opened..*

For both autonomous features, there can be instances where the robot encounters obstacles along its routes or is interrupted by some other event. In this case, due to the difficulty associated with coding for autonomous handling of obstacles, the user must take manual control to maneuver around the obstacle. They may then restart the autonomous pathfinder to the same destination. The obstacle detection first informs the user of the problem via a notification shown in figure 27, and then requires confirmation that the user is aware of the obstacle before activating the controls. The user must then confirm again that the obstacle has been cleared before automatic pathing may be resumed, if the obstacle can be moved. From the usability tests, it became apparent that users needed the notifications and information they received to be explicit, so that they for instance didn't think that the robot would proceed to drive into the obstacle if they confirmed that they had seen the initial notification. Explanatory text of what will happen once the confirmation buttons are pressed will therefore accompany these notifications.


Automatic route in progress...
OBSTACLE DETECTED!

Obstacle detected!
The obstacle must be removed before autonomous mode can be resumed.

Press OK to close this notification

Figure 27: *Interface with notification alerting that an obstacle has been encountered.*

7.12 Color

The interface primarily uses three variations of colors: gray, light gray and yellow. The primary concern in regards to colors was ensuring visual clarity and making the purpose of features and segments of the interface explicit. For this purpose, a majority of the interface was kept gray, whilst interactable and control related design elements were assigned a color to distinguish them from their surroundings. Yellow was chosen primarily due to its association with the colors featured on the existing robot design, as well as being a common choice for construction applications. Gray was chosen due to its neutrality, which was intended to give the feed a greater presence on screen. For features where color association is particularly strong, such as red and green for negative and positive actions and feedback respectively, these colors were used in order to match the users' mental models of similar applications..

Since the feed takes up such a large space on the screen, the visuals that the feed displays can have a large effect on the surrounding modules. Considering lighting, particularly bright or dark conditions may lead to certain features being hard to see or distinguish from surrounding modules. Certain shapes and colors of the surroundings could also cause problems, as the interface design might blend into the feed and confuse users. Although no visual material has been created for this feature, the intention is to include a toggle for switching between light and dark modes, which is a common feature on many existing web applications. This solution provides the user with a tool for managing challenging visual conditions, but will likely also have to be complemented with additional design work for optimizing contrast and clarity of features in the future.

8 Discussion

In the following chapter, the process and results of the project are discussed. Ethical considerations are also reflected upon and future steps for development of the product are suggested.

8.1 Fulfillment of research questions

In this section, the research questions and the degree to which they have been answered is discussed.

8.1.1 Research question 1

The first research question, which revolved around understanding what needs and requirements were associated with a robot-enabled remote inspection, was answered to an acceptable degree. Through the use of interviews and thematic analysis, a good understanding was acquired regarding the needs that exist for various types of construction inspections as well as what possibilities and problems would appear if said inspections were to be carried out using a teleoperated robot. Users from multiple different professions within construction were interviewed, which not only provided insight into many types of use cases, but also enabled a better understanding of the construction context. The user study, together with the other studies and general learnings from the project, provided a robust set of needs and requirements which the design and intended usage could be based around.

Most of the needs that were relevant to the robot-enabled inspections originated from needs associated with regular, in-person inspections, which speaks to their validity. However, others were more related to the teleoperation aspect of the interface and are, due to the users lacking experience in this area, less certain and some problems and needs might also have been missed because of this. While the interviews attempted to provide the users with a context of the robot's use, to help them reflect on a possible robot-enabled inspection, this could only be achieved to a certain degree due to the limited time for the interviews. Thus, the depth and insight of the users' input was most certainly still limited by their inexperience with this concept, and future research in this area is likely to find additional takeaways.

In summary, while some problems and needs were likely missed due to this fact, the number of interviews and the context that was provided to the users did result in a good amount of needs and problems regarding robot-enabled inspections, which are also deemed to be the most fundamental and critical ones.

8.1.2 Research question 2

The second research question, which dealt with determining what needs and requirements the construction inspection context places on a teleoperation interface, was not fully answered. The input from the user interviews provided a basic understanding of the inspection tasks and the construction context. This understanding was then combined with the accumulated knowledge regarding general teleoperation interface design, which provided some idea of what requirements the specific context placed on the interface design. However, due to the usability test's limitations, these specific requirements could not be thoroughly validated. Firstly, only one participant in the usability test was from a construction background. Thus, the feedback suffered from a lack of expertise and diversity, which complicated the validation of the analysis. Secondly, there was no way of testing the interface in a real world context or validating that certain features would work in reality, primarily because no real functional prototype existed.

Though there was an attempt to simulate the environment through the Unity prototype used in the usability test, it was not detailed or advanced enough to accurately depict all the relevant aspects of using the product on a real construction site. For example, the prototype did not feature any moving people or machinery, which is a key feature of real construction sites.

The lack of real users and realistic environment in the usability tests thus meant that the established needs and requirements for the interface, which were specifically related to construction inspections, could not be fully validated. If the interface was to be used to complete a real inspection on a real site, previously undiscovered requirements might become evident and others might change. Because of this, while a theoretical set of needs and requirements were determined in relation to this research question, they have not been properly validated and the question has not been fully answered.

8.1.3 Research question 3

The third research question related to determining what requirements exist for a teleoperation interface with high usability. Through the literature study, a good basis for these requirements was created, relating to areas such as situation awareness and control feedback, and these insights were also verified through the usability test. Where possible, multiple sources were also used to verify the takeaways and provide a deeper understanding of the issues. The usability tests also expanded on these requirements by providing additional insights regarding the overall layout of the interface and the controls used for driving and joint manipulation. Insights regarding the general understandability of the control interface and the explicitness of the various features were also acquired through the tests. The fact that most participants were not part of the intended user group is not deemed to be a problem in regards to this research question, as their experience with this type of teleoperation interfaces was non-existent, which will also be the case for most of the intended users. Thus, the received feedback in regards to the interface's usability would most likely have been very similar if more of the participants had been from the intended user group.

One limitation in regards to this research question is that the conclusions from the usability test have not been possible to validate, as there was not enough time for a second test. However, the problems identified in the tests were mostly clear in how they impacted the usability of the interface and were able to be solved in ways that did not introduce new uncertainties. The fact that only six tests were performed may also have impacted the results, as there is a risk of certain user behaviors and flaws of the concepts not surfacing throughout the tests. However, as the tests were intended for broad evaluation of features and general usability, coupled with the length of the tests, the core problems are deemed to have been identified and the resulting requirements are still deemed to be valid. Thus, through the combination of the literature study and the usability test, a good understanding of the requirements for a teleoperation interface with high usability has been acquired and the research question has been answered to a satisfactory degree.

8.2 Impact of robot design

Since the project was based around an existing robot design, much of the flexibility and options for making changes to the hardware was lost. Though this was specified as a feature of the project from the start, it is still interesting to consider what impact the choice of robot design may have had on the interface design compared to a more generic, less rigorously specified design.

An important feature of many vehicles, which was identified early on in the project, is the ability for users to have a view of the rear and side of the vehicle. For many use cases, this allows users to more confidently conduct reversing maneuvers as well as have greater situation awareness of what is going on around them. For this project, however, the limitation of having only one available camera for the user meant that this feature was not possible to include in the interface. One might speculate on how differently the design may have ended up if a rear view window had been included early on.

The nature of having only one camera feed led to making decisions focused on increasing the user's situation awareness through other means, such as adding functionality to the map and including the top-down view tool. The layout of the features on screen was also heavily impacted by the number of features. Adding more windows that had to be available to the user may have led to a more crowded interface or one where the user would have to switch between different tabs containing different views.

Another impactful area of the robot's design was the set of sensors and softwares which enable various complex and autonomous behaviors. Because of the limitations of the robot, some behaviors were deemed to be too advanced to include or to be outside of the capabilities of the robot's set of features. Certain features which were considered to be added to the interface were therefore omitted. One such feature was the use of point cloud generation, which was deemed to be impossible given the lack of 3-D lidar capabilities on the robot. Such a feature could have provided the interface with the option of including a modeled version of the environment the robot is in, which could then have been merged with proximity data and input from the feed to give users rich and integrated feedback in a third-person perspective.

Another feature which was impacted by the robot's limited set of sensors and softwares was the automation segment. As the project began, automation was initially identified as a more promising approach for the primary navigation instead of manual driving, as it required less training and brought large benefits in terms of efficiency and safety. However, robot limitations meant that the automatic usage became more and more constrained and what would seem to the users as arbitrarily convoluted. Manual driving therefore took up a larger role in the intended usage, and appropriately gained a larger focus in the interface design.

While the findings relating to the research questions - such as the broader areas of teleoperated inspections of construction sites and usability of teleoperation interfaces - are applicable to any project seeking to implement a teleoperated robot onto construction sites, the design solution is less general in its application. Because the hardware dictates the interface design to such a significant degree, the final design of this project might not be applicable to other robots whose hardware differs from that of Droidio. Due to the considerations discussed in this section, the design decisions must therefore be considered in the context of designing for Droidio when applying the learnings to one's own project.

8.3 Ethical considerations

Throughout the project, a number of ethical considerations relating to the project as well as the idea at large have been identified. The most frequently identified one concerns the surveillance aspect of the robot, which users from the interviews regarded as a hurdle. Many users reflected on how they and their colleagues would react to being monitored by a robot on site. Though some of them relayed how stationary cameras were once seen as intrusive, but had since become a natural part of the site, others still emphasized that their employees would not want to be monitored while conducting their work. As the robot has autonomous capabilities, the question as to who is responsible for this monitoring also becomes relevant. Aside from adhering to laws and regulations, the interface should therefore likely provide the user with the means to control what is and what is not monitored, as well as give them the ability to erase or save certain data. It may also be a good idea to keep track of what user is operating the robot at any given time, and inform said user of their responsibilities as an inspector on site before they are given control of the robot. Furthermore, properly informing the on-site workers about the robot will be important to increase their acceptance towards it.

In addition to surveillance, another impact of the product is the reduction of in person site visits. Basically, having the option of inspecting the site remotely will reduce the need for going to sites personally and instead lead to operators staying in their office or at home. In addition to the health disadvantages of spending more time sitting still and being inside, users also expressed desire to be on site themselves and to socialize with their colleagues there. If a product like the one covered in this report were to become the standard for construction inspection, and users were incentivized to not come to sites in person, this could have negative implications for their physical and mental health. As it stands, the robot is not intended for socializing, and its role as a medium for functional communication is limited to certain use cases. Therefore, there will still be a need for on-site visits to socialize with co-workers and to cooperate in solving complicated on-site problems. Nevertheless, the robot fundamentally makes users spend less time on site. The question for the construction industry, therefore, is to what extent human personnel should be present on site. Given that an answer to that question exists, the features and usage of the robot can be tailored to incentivize users to go personally for certain tasks and to use robots for others.

In terms of accessibility, the idea at large should become a useful tool for those users who, for whatever reason, struggle with conducting site visits themselves. Being able to perform their inspections via the robot will therefore increase overall accessibility to the occupation of construction inspector. For the interface itself, the project has not been focused on maximizing accessibility, and therefore most likely has some flaws in this regard. The project has considered two primary user groups - young and old people. The assumption was that older people would have more difficulties with using a digital interface and that younger people who have more experience with video games and various web applications would perform better. This assumption was later supported by the results from both the interviews and the usability test.

For the interface, this meant that a lot of effort was placed on making it usable by the older and less experienced demographic, as these users make up a large portion of the target group of construction inspectors. Aspects that may improve usability for older people that also translate to better usability for other groups may therefore aid the accessibility despite the lack of focus on this portion of the design, although future evaluations of this would be appropriate. The readability of features on screen, being able to recover from errors as well as adherence to commonly used applications are all aspects of the design which benefit many different user groups.

Lastly, the topic of safety has also been considered during the project. The robot is not particularly large or fast and has features to prevent it from driving off edges or physically hitting things, and thus the risks involved were not deemed to be great enough to make safety a primary aspect of the interface design. Some thought has gone into it however, mostly related to ensuring that the users have good situation awareness and thus remain in better control of the situation, which lowers the risk of errors and accidents. Despite the low perceived risk, it would be appropriate to also brief the users about the robot and the safety aspects before they begin driving. Furthermore, there will be laws and regulations which will affect the design of the robot and its use, which in turn affects the design of the interface. As this project has been focused on the usability of the interface, these aspects have not been addressed, and thus would be necessary to do in the future.

8.4 Reflection on process

The project process, as specified in chapter five, served its purpose well for the most part. However, there are a few instances where the process could have been altered to better suit the needs of the project.

The ideation phase, which followed the user- and literature studies and was meant to produce a range of distinct concepts for the decision matrix evaluation, took up too much time. Each concept required a layout for the various elements and, while a lot of knowledge had been acquired regarding what elements to prioritize and what general principles to follow in terms of teleoperation interface design, data to motivate different layouts was still lacking. The literature study only provided general guidelines for the design and the user study was more concerned with the overall functionality and need identification. It was thus difficult to motivate which layout concepts were better and move on to the individual features. The fact that many features would drastically alter the layout if implemented in a certain way also made this phase more difficult. An approach that might have enabled well motivated decisions to be made faster is the inclusion of more prototypes and usability testing.

If the usability test had been performed earlier or an additional smaller test had been utilized in the middle of the ideation phase, better feedback could have been gathered in regards to the concepts' overall layout. In turn, concepts with better usability could have been identified quicker and more time would have been left to make additional iterations and refinements of the final concept. However, due to the extensive ideation phase, many problems were indeed discovered and new information was acquired where it was needed. Thus, there was a lot of iteration present and it is possible that some of these insights that were gathered would have been missed if this phase was sped up.

Another area which in hindsight could have been prioritized differently is the prototyping phase. Initially, it was unclear what kind of prototype would be possible to produce. As the authors had limited programming experience, a fully functional prototype with driving included was not deemed to be realistically implementable. However, after some research, the game development platform Unity was determined to be a possible solution and thus the driving and control functions were implemented using it, which significantly increased the realism and potential of the usability tests. As this type of simulation closely represents the final use of the real interface, it turned out to be a very useful tool for evaluating this type of teleoperation. It meant that semi-realistic control of the robot and driving could be tested and combined with the Figma prototype, covering all main aspects of the product's use and enabling them to be evaluated. This supported the validity of the tests which, despite the switching between the two softwares, closely resembled the desired interaction.

However, towards the end of the prototyping phase, it was clear that if more time had been put into the Unity prototype earlier, the entirety of the usability test could have been performed using it, instead of allocating some functions to the Figma prototype. This would have increased the validity even further and allowed for evaluation of even the complex interaction between the control of the robot and the use of the interface features. Thus, a better understanding of the concepts' usability could have been acquired, due to the increased understanding of how the navigational control of the robot and the menu interactions interplayed with each other.

8.5 Next steps

One important aspect which has not been tested, as this was a delimitation, is the actual teleoperation aspect. The prototype used in the usability test was a simulation running in real time on a computer in front of the user and thus there was no delay or lag and the robot model did not respond the same way a physical robot would. These aspects could affect how the control of the robot is perceived and if the current solutions for feedback are sufficient. Thus, testing a real robot would be necessary to truly evaluate these aspects and determine the real level of usability and what possible problems still exist in the interface design. The fact that users were aware of the fake nature of the robot also likely led to them not behaving in the same way as they would in a real scenario, mainly by driving recklessly. Real world testing would resolve this issue, though it would also make the consequences of reckless driving more severe.

During the literature study it was concluded that using multiple modalities, in particular audio, could be a good way of addressing the issue of visual overload in the interface. However, due to time limitations, this aspect of the interface has not been addressed in this project and is therefore something which could be investigated in the future.

Another activity which would be of high importance for future work is testing the final concept in a real context with real users. This is related to the second research question which could only be partially answered by the end of the project. To fully answer it, the final concept would need to be evaluated with more users from the intended user group, who would be able to provide additional valid feedback in terms of the product's usefulness for their specific work and other kinds of construction inspections.

To answer this research question, the concept would also need to be tested in a more realistic context. The site model used in the Unity prototype was based on knowledge acquired from the interviews, but was in many ways inaccurate due to an incomplete understanding and limited time for creating the prototype. For example, the model is very streamlined and does not contain any rough terrain or debris which was described as potential problems during the interviews. More importantly, the model was completely static during the tests, meaning it did not capture the dynamic nature of real sites with moving people and machines. Thus, the model failed to capture some of the key aspects of a construction site, aspects which would increase the difficulty of the teleoperation. To evaluate the interface's usability in relation to its intended context, the product therefore needs to be tested on a real site. A more realistic simulation could potentially also be used, however this would require extensive time, programming and 3-D modeling experience and knowledge of construction sites. Thus, testing a simple prototype on a real site may be a more economical and in many ways better solution.

Lastly, the concept needs to be further developed to fit the technical requirements and limitations that exist. As this project has not included the technical development of the interface solution, the final design might not be completely aligned with the final product. The technical solutions to the various features have been taken into account, but there is a chance that they could be altered in the future. Thus, some features might end up working in different ways than those used as the basis for the final concept and the interface might therefore need to be reworked to match the new features. Making the interface responsive to varying screen sizes, while retaining good usability, is also important to consider for future improvements, as there was not time to investigate this aspect in this project.

9

Conclusions

The area of teleoperated robotics is one of increasing significance, as advancements in the technologies and manufacturing processes of components and softwares brings new capabilities and lower costs each year. Construction inspection is an area that could see large benefits from the introduction of teleoperated robots, which would increase efficiency, mitigate the risks of accidents involving humans and further progress the industry's digitization. While this paints an optimistic picture of the future, many challenges still remain that need to be overcome before such a product is able to be implemented. Aspects such as maneuvering a robot through the ever changing conditions of construction sites, enabling even the most advanced inspection to be performed via a robot and making these robots usable by novices are all problems that need to be resolved, and ones that this project has sought to find solutions to.

Through the use of user- and literature studies, a number of needs and requirements were established to support the design of a teleoperated interface, the purpose of which was enabling remote inspections at construction sites with a high degree of usability. These findings related to the work tasks, the context of construction inspections as well as general usability. An initial concept of the interface was then designed and evaluated through usability testing, which further built upon and validated the findings.

While the most critical requirements and most fundamental aspects related to the usage of the robot and its interface are deemed to have been covered in this project, there are certainly additional ones which remain undiscovered. For future work, performing evaluations in a more realistic context would be the obvious way forward. Paired with a physical robot, this would provide a testing environment capable of evaluating the robot and interface in a much more robust way, capturing all the intricacies of real teleoperation.

The project concluded with a design proposal covering the main features of the interface and addressing the majority of the identified requirements and needs. Problems discovered during evaluations were also addressed to improve the usability of the interface. The authors' hope is that this report may be used as a source of inspiration for future development of interfaces for teleoperated inspection robots, particularly those aimed towards the construction industry.

References

Benyon, D. (2010). Envisionment. In *Designing Interactive Systems* (pp. 177-197). Addison-Wesley.

Braun, V., Clarke, V. (2012). Thematic analysis. *APA Handbook of Research Methods in Psychology: Vol. 2. Research Designs*, 57-71. https://doi.org/10.1037/13620-004

Chen, J. Y. C., Haas, E. H., Barnes, M. J. (2007). Human Performance Issues and User Interface Design for Teleoperated Robots. *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS, 37*(6). https://doi.org/10.1109/TSMCC.2007.905819

Cross, N. (2000). *Engineering Design Methods: Strategies for Product Design (3rd ed.).* John Wiley & Sons.

Delmerico, J. et al. (2019). The current state and future outlook of rescue robotics. *Journal of field robotics*, 36(7), 1171-1191. https://doi.org/10.1002/rob.21887

Endsley, M.R., Jones D. G. (2003). *Designing for Situation Awareness: An Approach to User-Centered Design***,** Taylor & Francis. Chapter 6: Principles of Designing for SA. https://doi.org/10.1201/b11371

Figma. (2022, May 30).
https://www.figma.com/

Hanington, B., Martin, B. (2019). *Universal Methods of Design.* Rockport Publishers.

International Organization for Standardization. ISO 9241-11:2018. Ergonomics of human-system interaction -- Part 11: Usability: Definitions and concepts. Retrieved from https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en

Jordan, P. (1998). Principles of Usable Design. In *An Introduction to Usability* (pp. 25-37). Taylor & Francis. https://doi-org.proxy.lib.chalmers.se/10.1201/9781003062769

Labonté, D., Boissy, P., Michaud, F. (2008). Comparative Analysis of 3D Robot Teleoperation Interfaces with Novice Users. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *40*(5)1331-1342. https://doi.org/10.1109/TSMCB.2009.2038357

Mast, M., Materna, Z., Španěl, M. et al. (2015). Semi-Autonomous Domestic Service Robots: Evaluation of a User Interface for Remote Manipulation and Navigation With Focus on Effects of Stereoscopic Display. *Int J of Soc Robotics,* 7, 183–202. https://doi.org/10.1007/s12369-014-0266-7

Mitropoulos, P., Abdelhamid, T. S., Howell, G. A. (2005). System Model of Construction Accident Causation. *Journal of Construction Engineering and Management, 131*(7). https://doi.org/10.1061/(ASCE)0733-9364(2005)131:7(816)

Nielsen, C. W., Goodrich, M. A., Ricks, R. W. (2007). Ecological Interfaces for Improving Mobile Robot Teleoperation. *IEEE Transactions on Robotics, 23*(5). 927-941. https://doi.org/10.1109/TRO.2007.907479

Project-Management.com. (2022 February 22). *Affinity Diagram - Kawakita Jiro or KJ Method.* Retrieved June 5, 2022, from https://project-management.com/affinity-diagram-kawakita-jiro-or-kj-method/

Rasmussen, J. Pejtersen, A. M., Goodstein, L. P. (1994). *Cognitive systems engineering.* Wiley. Wiley series in systems engineering.

Senft, E. et al. (2021). Task-Level Authoring for Remote Robot Teleoperation. *Front. Robot. AI 8:707149*. https://doi.org/10.3389/frobt.2021.707149

Sheridan, T. (2016). Human-Robot interaction: Status and challenges. *Human Factors The Journal of the Human Factors and Ergonomics Society 58*(4). https://doi.org/10.1177/0018720816644364

Schultz. (2020, 30 May). Reclaiming Innovation: Closing the Technology Gap. *The Bluebeam Blog.* https://blog.bluebeam.com/uk/technology-adoption-construction-industry/

Singh, A., Seo, S. H., Hashish, Y., Nakane, M., Young, J. E., Bunt, A. (2013). An Interface for Remote Robotic Manipulator Control That Reduces Task Load and Fatigue. *2013 IEEE RO-MAN*, 2013, 738-743. https://doi.org/10.1109/ROMAN.2013.6628401

Statista. (2022, 30th May). *Size of the market for industrial robots worldwide from 2018 to 2020, with a forecast through 2028.* https://www.statista.com/statistics/728530/industrial-robot-market-size-worldwide/

Unity Technologies. (2022, May 30). https://unity.com/

Wikberg Nilsson , W., Ericson, Å., Törlind, P. (2015). *Design: Process och metod.* Studentlitteratur.

Appendices

Appendix A: Usability test template

Introductory briefing

We're currently doing our thesis on the topic of designing a control interface for a remote controlled robot intended for use as an inspection tool at construction sites. The idea is that various actors can connect remotely to do their work tasks on the site without having to go there in person. The robot consists of a low body with track wheels and an arm extending upwards with a camera and various other tools.

For this test, we're using software that is at an early stage and does not contain all features and polish of the final product. Feel free to explain your thought process and what you would do in a real scenario and try to press the things on screen that you actually would, but when asked to do a task - don't make it more complicated than it needs to be (!) The size of the prototype scales exponentially with how many options we give the user.

This test is intended to evaluate and gather insight into the prototypes - we don't want you to just say what you think we want to hear. If there are problems you see with the interface, that's what we want to hear about! Also, please think out loud as you're completing the tasks, and try to place yourself in the scenario of someone using this interface for the first time.

You're an inspector at a construction company who is in charge of monitoring a site in New York. Today, you're not able to go to the site in person, and are instead aiming to complete your inspections using a new remote controlled robot that the company has purchased. You've been given a link to the control interface by a colleague and have just connected via your laptop.

Basic inspection

- You want to check up on how work is proceeding at zone A - you're not sure of the way to get there. You need to confirm whether any red wall segments have come up. You also want to see if the black boxes have been installed up on the three pillars. You want to go there manually.

- *Was it easy to complete this task? Why/why not?*

Automatic driving

- You now want to go to zone B, but would like to do so automatically. Please complete this task.

- *Was it easy to complete this task? Why/why not?*

Going to a different floor

- You have been asked to do a safety round on floor 5, a routine inspection round that is done every day. You are on floor 1 and need to locate a construction hoist somewhere on this floor.

- *Was it easy to complete this task? Why/why not?*

- You now need to tell the construction hoist operator which floor you want to go to.

- *Was it easy to complete this task? Why/why not?*

Complex inspection

- Drive to Zone C. You want to closely inspect the material meeting of a window panel and a wooden panel. You can see the segments ahead, but there is a lot of debris in the way.

- *Was it easy to complete this task? Why/why not?*

- You now want to take a picture of and compare it with one that was taken yesterday.

- *Was it easy to complete this task? Why/why not?*

- The concrete pillars in zone B have recently been constructed, you want to go there and find a pillar for testing. Select a pillar and get close to it.

- *Was it easy to complete this task? Why/why not?*

- You want to measure the heat of the pillar using the attached heat sensor. You know from experience that the sensor needs to be very close to the material to get an accurate reading.

- *Was it easy to complete this task? Why/why not?*

Finishing up

- You're now done with the work for today and need to head back to the charging station. You've been told to put the robot back on the charging pad with the arm folded down. Exit the interface when you feel done.

- *Was it easy to complete this task? Why/why not?*

Follow Up questions

- What are your overall thoughts and feelings after using the robot and its interface?

- What was **easy** and what was **difficult**?
- Was the structure of the program good? Would would have rearranged some of the features or information?

- Did you feel like you had **situation awareness** (were you aware of your surroundings?) Was your view obscured too much?
- Where did you look during the navigation?
- Which methods of control did you prefer? Why? Do you think this would change over time?

- Would you be able to perform some or all of your daily tasks remotely using this solution? **Why/why not?** (Features, layout, awareness, etc.)

- How would you describe the experience? Use the stress/positivity scale?
- Did you feel like the interface was helping or hindering you? Was anything frustrating?

Appendix B: Interview template

- Could you give us your job title and then describe in your own words what it is that you do?
    - What are some routine tasks you do every day?
        - What tools do you use / have on you when at the site?
    - What are some of the most challenging tasks?
    - How much time is spent on site vs off site?
        - What is the purpose of your visits to the site?
            - *Which inspections are you responsible for?*
            - *Which inspections are done by third party individuals?*
            - *What are the most important aspects of your visit? (E.g communication, visuals, measurements, team building etc.)*
        - Are you able to do any on site work remotely?

- What are some key characteristics / elements of the construction sites you work on?
    - Are all sites the same? - what varies?
    - What people are present on the site?
        - At different times of the day / project?
        - Interaction between you and them?
    - What robots and machinery are present on the site?
    - How long do sites stay active?
        - What period of time are you active in the project?
    - Are there any areas on site where you don't want to / can't go?
        - Hazardous
        - Time consuming
        - Cramped spaces
        - Night time
        - Bad weather
    - Do you do some setup of the site before commencing work? I.e setting up fences, marking areas etc.
        - Do you use BIM to get a digital representation of the area?
        - Do you have your own Wifi?
        - What other tasks do you do at the start of a project or when first setting up the site?

- Have you used robots or drones on site before?
    - For what purpose
    - Pros and cons

- Do you think you could complete some of your on-site tasks using a remote controlled robot?
    - Which ones are definitely doable? Why?
    - Which ones are definitely not doable? Why not?
    - If you were to use such a robot, what would the most important aspects of the robot be? What features would it need to have?
    - Do you see any legal (or otherwise problematic) issues with letting robots represent someone on site during an inspection or letting it take over some tasks from the people

on site?

- What are your thoughts about a robot being used for
    - Delivery
        - One idea is to have the robot deliver things around the site. What do you think about this?
        - What examples of material, tools etc. that could be delivered can you think about?
        - Is there a centralized hub for storing material and tools on site? Would a robot charging station be able to be implemented there?
        - Do you have specific roads on site for forklifts and trucks etc.?
            - Are there often obstacles / other traffic on the roads?
            - What could happen if a robot started driving on the roads?

    - Monitoring
        - Monitoring in this case simply means that the robot has cameras or sensors with which it or an operator can monitor something. What potential uses for this do you see?
            - Inspections
                - Visual
                - Thermal
                - Audial
                - etc.
            - Security

    - Communication
        - Communication at the construction site is another potential application, what do you think of using a robot to communicate with people on site?
        - How would you feel if you were approached by a robot on site? (show examples? Use existing robots? Spot + small RC + big RC + drone + Double 3 for example..)
            - One that is talking and that you are meant to talk to
            - One that broadcasts a video stream

    - Is there some use area we haven’t considered?

- What kind of interfaces do machines / robots on site currently use?
    - Thoughts?

- What kind of interfaces do **you** come into contact with?
    - Video conferencing
    - Camera feeds
    - etc.
    - What should an interface for controlling the robot be like?

- How should a distance controlled robot be controlled by an operator?
    - Show examples?

    - General thoughts, problems, preferences..

- Is there anything you thought we would talk about that we haven't covered?
    - Anything you'd like to add to what we've talked about?

- Thank you so much for your participation!

Appendix C: Functional analysis

Function			Prioritization 1=Critical 2=Necessary 3=Desireable 4=Potential
Verb	**Noun**	**Comment**	**1,2,3,4**
Facilitate	**Remote work**	**The solution should facilitate the operator's remote work at construction sites.**	**1**
Enable	**Monitoring**		**1**
Provide	Video feed		1
Provide	Multiple camera feeds		3
Enable	PTZ movement of camera		1
Provide	Data gathering	Data should be able to be saved, either automatically or manually, using the currently equipped tool / camera	2
Offload	Calculations/estimations	Interface should assist the user in performing calculations and estimations of for example number of items they are looking at or how much they should turn the camera to see something.	4
Enable	Automatic object count	Users shou ld be able to use a "count object" function on an area of or the entire camera feed	4
Enable	Picture/video/audio capture		2
Enable	Sorting of data	Data should be able to be sorted based on the user's needs	4
Enable	Reviewing of data	View videos/photos taken, see previous measurements.	4
Enable	Exporting of data	To other software or to other hardware	2
Provide	Preset program	Preset programs for how the monitoring should be handled during a specific route.	4
Provide	Self-inspection	The user should be able to point the cameras onto the robot itself to inspect it	4
Enable	**Movement/control**		**1**
Enable	Manual driving	Provide method for user to manually drive using the interface	1
Provide	Turning		1

Provide	Acceleration	Either a continuous slider or static speeds.	1
Provide	Breaking/stopping		1
Facilitate	Control of arm	Presets of arm positions or continuous control.	1
Provide	Manual control of arm	The user should be able to directly move the arm in all its degrees of freedom	1
Provide	Action level control	i.e "pull", "loosen", "stretch out"	3
Enable	Button pressing	For when going in elevators. The robot arm should be able to locate buttons and press them. Perhaps use RFID signs on elevator or next to button to make it easier for the robot to notice it.	4
Facilitate	Autonomous driving		2
Enable	Planning of route	Could be planning of both driving and use of of tools/cameras. Either through choosing a saved route or by drawing up your own route/using destination pathfinding	2
Enable	Selection of existing route	Utilizing the saved paths to save time	2
Enable	Destination Pathfinding	Place a destination and have the robot go there.	2
Enable	Saving of route	Routes should be able to be saved and then conducted automatically later on	3
Provide	Abort function	The user should be able to stop the robot during an autonomous route.	1
Support	Control	Help the user controlling the robot effectively and safely	2
Enable	Verification	Users should be able to verify that their inputs have been interpreted correctly before executing. While it is most important in movement, it is also important for all inputs the user gives to the robot.	2
Provide	Awareness of global position	Where on the site am I? Which floor am I on?	2
Provide	Awareness of local position	Where am I in relation to my immediate surroundings?	2
Provide	Awareness of robot status	Battery charge, connection, memory, name, current goals, joint positions etc. Should be provided directly, even if robot itself is delayed.	2
Provide	Awareness of site status	Areas being closed/opened, active work on site etc.	3
Support	Sensing of spatial features	Help user understand **depth**, speed, distances, size etc. (of both robot and environment)	3

Enable	Connection-loss recovery	There should be a way(s) for the user to resolve the problem of lost connection, possibly depending on what task they are undertaking.	4
Enable	Direct control	Allow users on site to access robot control without access to a computer	4
Provide	Automatic charging	Robot should be able to find its own way back to a charging station (and dock with it)**,** the user should be provided with an option for doing this.	3
Communicate	Non-trivial deviations	If an obstacle is encountered or the robot is unable to perform the user's instructions, this must be communicated to the user along with the reason for it	1
Minimize	Battery consumption	Possibly by using an alternate mode where the robot is stationary and only observing.	4
Provide	360 view around robot		3
Facilitate	**Communication**	**Talking to people, making others aware of the robot's presence and the operator's intent etc.**	**2**
Provide	Hearing		2
Broadcast	Sound	Operators should be able to talk into their microphone and broadcast the speech to on site personnel. Perhaps there should also be means of sending signals for alerting people on site to the robot's presence. Mute/unmute or push to talk	2
Target	Communication	Be able to direct your speech to a certain person	4
Reduce	Noise	Filter out background noise etc.	4
Control	Sound	Raising, lowering, muting etc.	2
Record	Sound		4
Enable	Contacting of operator	Personnel on site must know that there is an operator to contact, who it is and how to do it	3
Facilitate	Laser-pointing	Allow operators to point to objects on site to aid communication.	3
Facilitate	**Delivery**		**2**
Ensure	Awareness of payload	Within the interface the user should be able to tell if the robot is currently carrying a load	4
Provide	Ability to give description	To payload. I.e, "payloads consist of 3x boxes of 9mm screws, one walkie talkie"... Maybe time of loading and unloading, place, person etc.	4
Be	**Usable**	**Based on theory regarding usability, user friendliness, error and safety, ergonomics etc.**	1
Be	Intuitive / Guessable		2
Emulate	Control norms	I.e swiping, clicking, zooming..	3
Be	Simple to use		2

Be	Consistent	Make sure to use the same kind of signals and wording in all parts of the interface so that the user doesn't get confused	2
Be	Explicit		2
Be	Compatible	With currently used softwares and how they function/behave	2
Facilitate	Positive user experience		2
Convey	User consideration	Should feel like the product is adapted for you and that thought has been put into your experience.	3
Be	Aesthetically pleasing		3
Reduce	Anxiety of failure		2
Possess	Visual clarity	Location of information, color coding, sizes support a clear overview of information	2
Utilize	Spatial mapping	In an effective way so that the user understands the clear connections between their input and what happens in the real world.	3
Use	Symbolism	Symbolic rather than analog language. Reason is that many users will be novices who are not proficient in robotics or programming, and will therefore have a hard time operating the robot through analogic means.	3
Separating	Basic and advanced features	To assist users attention guidance and maintain a simple standard view.	2
Minimize	Cognitive load	Segmenting tasks, allowing focus on one thing at a time	3
Reduce	Attention workload	Users should not have to work hard to distribute their attention efficiently. The need for attention should be minimized throughout the interface. Possibly by integrating information.	3
Prioritise	Information and functions	The most critical functions and information are given more space and accessibility in the interface, for example through a toolbar or hierarchical menu structures	2
Offload	Low level tasks	Where possible, the solution should offload low level tasks from the user to the robot	3
Provide	Suggestions	The interface should be able to anticipate certain user needs and provide suggestions to alleviate their usage	4
Combat	Errors		2
Retain	Attention	Make sure the interface retains the user's attention even if distractions occur.	4
Mitigate	Error consequences		2
Counteract	User errors		2

Facilitate	(sense of) Control		2
Require	User confirmation	User should have to confirm having seen notifications of the robot interpreting danger. I.e, if the robot thinks it's about to run off a cliff, the human should not be able to ignore the prompt, but should have to confirm that they have seen the danger and deemed it to be safe. If it sees a slope it thinks it can't get over, the user should be made aware of the risks associated with trying to do it anyways.	2
Provide	Feedback		2
Provide	Relevant functionalities/information	For the task at hand. So, for instance, when you want to drive, all functionalities for doing so should be available to you	2
Facilitate	Efficient use	Interaction with robot and navigation of the interface should be quick and efficient in regards to the user's time and mental resources	2
Communicate	Robot's capabilities	Be clear about functionalities, dispel illusions surrounding robots	2
Manage	Operator confidence	Primarily related to most users likely being unsure of their ability to drive the robot and might thus need support to feel comfortable using it.	4
Minimize	Disruptions	Users should be able to maintain a good work flow.	3
Facilitate	**Administration**		
Facilitate	Planning		3
Display	On-site info	Weather, time, units present etc.	3
Provide	Scheduling service	The robot should be able to be booked. These bookings need to be viewable by different users with access to the robot.	3
Provide	Programmed route importing	Programmable routes with movement of robot, arm and cameras. Can be saved and presented in scheduling service.	4
Enable	Choice of unit	If there are many robots in the fleet, the user should be able to choose which one (s) to access and differentiate between them.	3
Enable	Personalization		4
Provide	Login service	When connecting to interface, the user should be able to identify themselves and access their account with its associated settings and customizations	3
Provide	Presets	Based on your profile, there should be presets to aid your usage of the product (i.e "I am an electrician - only show me info relating to electrical work"...)	4

Enable	Customization	For each user, there should be options for customizing the interface to their needs. Users should be aware of what default settings they are using, how to change the settings and how to save them for future use.	4
Support	Swapping (of attachments)	Need interface to match reality and provide tools for controlling and monitoring attachments	3
Provide	Help and tips	The user should be able to look up information on how to use the interface and robot. Possibly also about the site/project	3
Provide	Notifications	In certain instances, the interface should notify the user of different status updates. Charging complete, idling warning, low charge, bad connectivity, booked time almost over, dangerous area ahead etc. 5 minutes left until someone has booked the robot, take it back to charge.	2
Facilitate	Multiple users	Multiple users should be able to connect to a session to view the camera feed and other information.	4
Notify	Admin	If a robot gets activated unexpectedly (alarm goes off?), leaves the site, needs service etc.	4
Provide	Access to history	Operators should be able to access the history of where they've been throughout their usage of Droidio. Likewise, admins should be able to see where the robot has been over many uses.	3
Provide	Cleanliness check	If the robot gets muddy, snowy or otherwise dirty while driving, there is a risk that the operator will not notice and be able to inform someone so that they can clean it before charging. The interface should make the operator aware of this possibility.	4
Enable	Adding of new features	The interface should be flexible so that the user can add new camera feeds, tools, controller methods easily	3
Facilitate	**Safety**		
Establish	Responsibility	At no point should it be unclear who is responsible for the robot and its actions	2
Encourage	Safe use	Alert users of dangers, prompt users to drive safely, limit speed and force of certain features etc.	2
Inhibit	Unauthorized movement	A single individual on site should probably not be able to roll the robot around. Requiring two people to lift it limits the potential harm one person can do.	3
Communicate	Intent (to on-site personnel)	People on site should be able to understand what the current mission of the robot is and how it is likely to behave	3

DEPARTMENT OF INDUSTRIAL & MATERIALS SCIENCE
Division Design & Human factors
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2022
www.chalmers.se