Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Engineering
Department of Mechanical and Industrial Engineering

Lorenzo Andreoli

Smart Workstation with a Kinect-Projector Assistance System for Manual Assembly

Master's thesis in Global Manufacturing Management
Supervisor: Fabio Sgarbossa
June 2019

NTNU
Norwegian University of Science and Technology

Lorenzo Andreoli

Smart Workstation with a Kinect-Projector Assistance System for Manual Assembly

Master's thesis in Global Manufacturing Management
Supervisor: Fabio Sgarbossa
June 2019

Norwegian University of Science and Technology
Faculty of Engineering
Department of Mechanical and Industrial Engineering

NTNU
Norwegian University of
Science and Technology

Executive Summary

Although automation has increased more and more in manufacturing companies over the last decades, manual labor is still used in a variety of complex tasks and is currently irreplaceable, especially in assembly operations. The problem of assisting and supporting the human worker during potentially complex assembly tasks is, therefore, very relevant. Clear and easy-to-read assembly instructions, error-proofing methods, and an intuitive user interface for the worker have the potential to reduce the cognitive workload of the operator, increase the productivity, improve the quality, reduce defects, and consequently reduce costs.

Digital technologies such as augmented reality and motion recognition sensors can help assembly workers in their tasks and have been subject of research with increased interest over the years. In this thesis, we develop a prototype of a smart workstation equipped with a Kinect-projector assistance system for manual assembly. By following a V-model for systems development approach, we identify key requirements for assistance systems for both continuously supporting workers and teaching assembly steps to workers. Thereby, based on the identified requirements, we design and build a functional prototype with the following features: in-situ projection visualization, Pick-by-Light, picking error-proofing, and gesture user interaction.

A case study in a laboratory is conducted to assess the prototype in terms of performance and accuracy, user acceptance and mental workload. The participants tested the assistance system together with two established methods in manual assembly (i.e. paper-based manuals and a mounted monitor) by assembling LEGO models of different product complexity. The results reveal that our system performs significantly better in terms of number of errors made by the workers, user acceptance, and mental workload required by the user to execute the assembly tasks. Further, they show that product complexity is a major factor in deciding whether adopting the assistance system. However, the measured task completion times were higher than using methods such as mounted monitors, therefore highlighting the limitations of the prototype and suggesting that further research is necessary in this regard.

Preface

This report is the result of a Master's Thesis conducted at the Department of Mechanical and Industrial Engineering at the Norwegian University of Science and Technology (NTNU), in Trondheim. The Master's Thesis is a final assignment in the Master of Science program in Global Manufacturing Management. This project is coded as TPK4930 and is part of the Production Management research group.

The research was conducted in connection with the Logistics 4.0 Laboratory at NTNU and it is a multidisciplinary project that combines digitalization within the logistics and manufacturing context.

It is assumed that the reader has a level of knowledge from industrial engineering in line with the Master of Science program.

Trondheim, 2019-06-11

Acknowledgements

This project is carried out in cooperation with my supervisor, Professor Fabio Sgarbossa. First of all, I want to thank him for providing me this incredible project, and further for the time he has devoted to me throughout the semester. His support and guidance has been crucial. In addition, I want to thank PhD candidate Giuseppe Fregapane for being my co-supervisor and assisting me superbly in many steps of the project. I also want to thank PhD candidate Mirco Peron for his precious help and advice, and master's student Tom Lowder for his valuable inputs and collaboration during our work in the lab.

Several professors and researchers at the Department of Mechanical and Industrial Engineering at NTNU have provided feedback, and for this I want to thank them.

I want to thank my parents for providing me with continuous support throughout my years of study and especially in the past two years away from home.

Finally, I want to express my deepest appreciation to Norway and its people. They have welcomed me with open arms and have made my student experience by far the best years of my life.

L.A.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AR	Augmented reality
H-CPPS	Human cyber-physical production system
HMD	Head-mounted display
MTP	Department of Mechanical and Industrial Engineering
NASA-TLX	NASA Tax Load Index
NTNU	The Norwegian University of Science and Technology
PbL	Pick-by-Light
SAR	Spatial augmented reality
TCT	Task completion time

1 Introduction

This chapter introduces the thesis by presenting the background and motivation, problem description, project scope, along with the objectives, research questions and project limitations, and finally the structure of the thesis.

1.1 Background and Motivation

Production effectiveness and efficiency are crucial in the industrial context. In order to seek improvements along those lines, automation and robotics have profoundly changed the way products are manufactured over the last 50 years (Funk et al., 2018). Furthermore, it is expected that over the next decades a further revolutionary shift to more integrated and flexible production systems will take place, as outlined in the Smart Factory (Lucke, 2008) and Industry 4.0 (Hermann et al., 2015) initiatives.

In fact, manufacturers are confronted with an increasingly customer-driven market. Such a market demands customized products that need to be released and manufactured in even shorter innovation cycles. At the same time, customers expect a high degree of technological innovation in their products. This leads to shorter time-to-market requirements for new products (Krammer et al., 2011). These changes also lead to an increasing number of product variants and more frequent releases of new products (Loch et al., 2016). In order to address these requirements, manufacturing companies aim at more flexible and adaptable production lines.

In most domains, production is not fully automated and human workers still play an essential role. For instance, in the automotive industry, cars and their components are produced by a cooperation of human workers and robots through a series of complex assembly processes (Funk et al., 2018). Considering the increasing degree of customization, product variants, and complexity, the capabilities of human workers are and will be needed in manufacturing (Loch et al., 2016). Humans are creative and have great skills when manipulating objects, therefore they bring an unparalleled degree of flexibility and improvisation (Funk et al., 2018, Gewohn et al., 2018).

However, dealing with a large number of variants and high degrees of complexity is cognitively demanding for the human worker. High level instructions are required. Workers have to understand which product variant they are creating and what steps are needed. With frequent changes and small lot sizes, traditional training and teaching approaches are not suitable, as one cannot learn all possible product variants upfront (Funk et al., 2018). A better option is to provide the information required for production when the worker needs them. For this reason, assistance systems that support the workers in complex assembly processes are an important field of research (Loch et al., 2016).

Manufacturers not only strive to guarantee product quality for their customers; they aim at ensuring especially production quality (Gewohn et al., 2018). This means producing as few defects as possible, while keeping the production rate high and achieving the quality requirements.

In that direction, assistance systems have a great potential in manufacturing and assembly activities. They can support the human worker through correct information visualization, error-proofing methods, and intuitive user-interfaces (Gewohn et al., 2018).

Augmented Reality (AR) (Bannat et al., 2008a), Pick-by-Light (PbL) (Funk et al., 2015) and motion detection sensors are among the technologies more often used in assistance systems. The manual workstation (Zamfirescu et al., 2014) is an example of a solution that provides step-by-step assistance for manual assembly processes.

Nevertheless, new assistance systems need to prove their benefit over established methods, for instance written documentation or instructional visualization on a monitor, in order to be considered for practical application (Loch et al., 2016). This report presents the study and the development of a smart workstation using a projector and a Microsoft Kinect motion sensing device for assisting manual assembly tasks. It shows an adequate way to visualize assembly-related information for the worker, as well as to prevent defects in a user-oriented manner.

This study was conducted in the Logistics 4.0 Laboratory (NTNU, 2018) at the Department of Mechanical and Industrial Engineering (MTP) at the Norwegian University of Science and Technology (NTNU). The intent of the lab is to “create value from digitalization in logistics” (NTNU, 2018). It merges digital technologies with traditional production and logistics systems, and enables the replication of real-life operations and material handling activities.

1.2 Problem Description

Although automation has increased more and more in manufacturing companies over the last decades, manual labor is still used in a variety of complex tasks and is currently irreplaceable. For instance, many assembly operations need to be executed by human workers.

Therefore, the problem of assisting and supporting the human worker during potentially complex assembly and manufacturing operations is relevant. Clear and easy-to-read assembly instructions, error-proofing systems, and an intuitive user interface for the worker have the potential to reduce the cognitive workload of the operator, increase the productivity, improve the quality of the assembled part, prevent defects, and therefore reduce costs.

Digital technologies can help the human workers in assembly and manufacturing activities, and their potential has been studied over the past few decades with increased interest. Augmented Reality (AR) is one of the most promising technology for information visualization in the context of Industry 4.0; AR can be achieved through different devices such as mobile displays, head-mounted displays (HMD), and projectors. Cameras, sensors, and motion detection devices are another branch of digital technologies that can support a human assistance system by measuring the manufacturing processes.

The above-mentioned technologies have been the subject of many researches, which led to the development of several and different prototype solutions. In this study, we are going to develop a solution combining an AR device (a video projector) and a motion sensing device (Microsoft Kinect) applied into a common assembly workstation in order to create a system with the following functionalities: in-situ projection visualization, Pick-by-Light, error-proofing, and gesture user interaction.

This solution consists of two relatively low-priced technologies (i.e. video projector and Microsoft Kinect) and has the potential to reduce the costs of manufacturing and re-manufacturing. Through in-situ projection, assembly instructions are projected directly on the workstation desk where and when the operators need them, leading potentially to a mitigation of their cognitive workload. Through a Pick-by-Light approach, the projector sheds light on the bins which the operator need to pick a component from, with potential benefits to the worker's cognitive workload. Through an error-proofing approach, the Microsoft Kinect detects the movement of the operator's arms and verify whether the right component is picked, while the projector can warn the operator with an error message whenever a component is picked from the wrong bin, potentially resulting in defects prevention at an early stage of the assembly. Finally, an intuitive gesture user interface has the potential benefits of lessening the worker’s cognitive workload and speeding up the process.

The developed solution is then compared with established and traditional methods in assembly manufacturing, such as paper-based instructional manuals and workplace-mounted monitors, through a series of assembly tests conducted in the laboratory.

1.3 Project Scope

In this section the objectives, research questions, and limitations of this thesis are stated.

1.3.1 Objectives

Assistance systems using in-situ projection for providing instructions at the workplace by means of affordable technologies such as a Kinect and a projector, have the potential for becoming available in many assembly environments. Therefore, research is required to identify potentials and limitations of such system.

The objectives of this thesis are identified as follows:

- Build a solution for manual assembly that combines a Kinect device and a projector in order to assist the worker during assembly operations;

- Develop an assistance system for manual assembly with the following functionalities: in-situ projection visualization, Pick-by-Light, error-proofing, and gesture user interaction;

- Assess whether the newly developed solution bring benefits to the assembly process in terms of assembly performance and accuracy, user acceptance and mental workload.

1.3.2 Research Questions

RQ1	How can a Kinect-projector assistance system for manual assembly include functionalities such as in-situ projection visualization, Pick-by-Light, error-proofing, and gesture user interaction?
RQ2	Is a Kinect-projector assistance system better in terms of assembly performance and accuracy, user's acceptance and mental workload, compared with traditional methods?
RQ3	How does product complexity impact the assembly performance and accuracy when using a Kinect-projector assistance system?

Table 1.1: An overview of the research questions

The research questions addressed in this thesis are presented in Table 1.1. For answering RQ1, the V-model approach is followed. First, the requirements of the new system are identified, as well as the general architecture of the system. Then, the system physical design and logical architecture are implemented in the form of a smart workstation for manual assembly. Later, the newly developer Kinect-projector assistance system is tested by users in order to evaluate the system and, in this way, answer RQ2 and RQ3. For answering RQ2, assemblies of LEGO models were performed and measured by means of the new system and two traditional methods, i.e. a paper-based manual and a workplace-mounted monitor. Further, the results were compared and analyzed. Whilst for answering RQ3, two different LEGO models with different levels of complexity were involved in the tests.

1.3.3 Project Limitations

The main limitation of this thesis is to be carried out in an artificial environment. Both the development and the evaluation of the new solution was conducted in a laboratory, which is inherently different from a real-world factory. Thereby, the results of the evaluation are affected by factors such as the environment, the users, and the products. In fact, the tests took place in a laboratory, the test population came from the academia, and the products consisted of simple LEGO models. Therefore, this study would surely have benefitted from a real manufacturing context, real assembly operators, and real-world assembly products.

This thesis is looking into applying digital technologies into the manufacturing context. Thereby, competences within the area of computer science, programming and coding above all, are surely needed in order to develop an assistance system. The author's background in industrial engineering, production management and logistics lacks of those above-mentioned skills. Therefore, the project was affected by this aspect in a way. For instance, an already existing motion recognition software was implemented to the assistance system, instead of programming a motion recognition algorithm, specific for this use case.

Finally, one other limitation of the project comes from the five-month time frame for this master's thesis. For instance, a longer time frame would have allowed for the implementation of corrective actions to the assistance system based on the results of the case study, or the design of more thorough assembly tasks for real-world products with higher level of complexity for the tests in the laboratory.

1.4 Thesis Structure

This thesis is written as a research report and consists of 9 chapters. The next chapter is Chapter 2, where the methodology that was followed in this thesis is presented, along with the research process and specific methods that were used. Chapter 3 introduces the theoretical background upon which the thesis is built, such as terms and definitions used throughout this report, and related work from the literature in the main areas of augmented reality and worker assistance systems. Accordingly with the V-model approach for systems development, Chapter 4 includes the requirements, a general architecture and the main concepts of our assistance system. Successively, Chapter 5 presents the physical and logical design of the Kinect-projector assistance system based on the defined requirements, along with a description of the implementation of the software for motion recognition. Chapter 6 introduces the experimental evaluation phase of the thesis by presenting the case study, the methods for assessing the system, and the metrics that were measured during the experiment. Chapter 7 gives an overview of the results of the case study. Chapter 8 includes a discussion about the findings from the study in relation to the research questions, and also provides a discussion about the limitations of the findings and of the assistance system itself. Finally, conclusions and recommendations for further work are given in Chapter 9.

2 Methodology

This chapter introduces the methodology used in the thesis and explains how the research process was carried out. A V-model for system development lifecycle process was followed for developing a smart workstation with a Kinect-projector assistance system. Further, both qualitative and quantitative methods were used in the process, and they are described later in the chapter.

2.1 Research Process

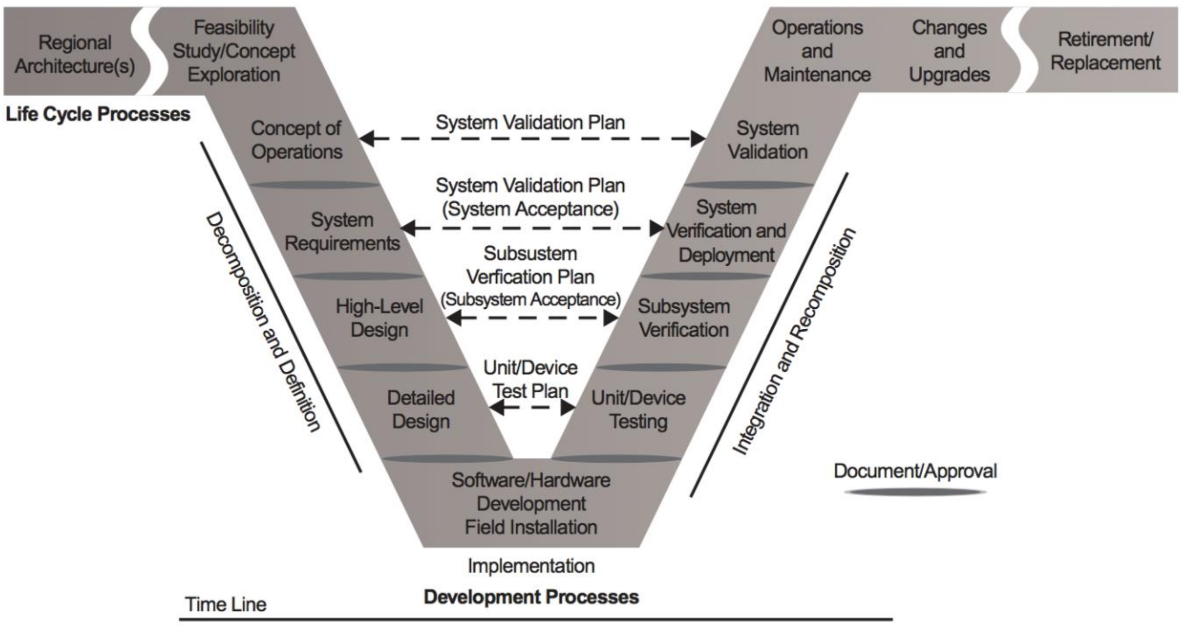


Figure 2.1: The V-model systems engineering approach

The research process of the thesis followed a V-model approach (Kossiakoff et al., 2011), which is shown in Figure 2.1. The V-model comes from systems engineering and is used for system development. It summarizes the main steps to be taken in a system development lifecycle. It consists of the following main steps: project definition through concepts, requirements, and architecture; system design; system implementation; test and validation. This process was followed for the development of a smart workstation equipped with a Kinect-projector assistance system for manual assembly. First, a literature study on augmented reality and worker assistance systems was carried out in order to gain relevant knowledge on the topic and the technologies involved in the project, and more specifically to define the use cases, the requirements, the general architecture and the concepts for our assistance system. Second, the system was designed in details and implemented by physically building the smart workstation in a laboratory and implementing the software for motion recognition and the underlying logical architecture. At a later stage, the newly developed assistance system was tested by users in a case study for validating and assessing the effects of the new system in comparison with more traditional methods that are generally used in assembly manufacturing environments. Finally, the results of the tests were discussed and the limitations of the system were introduced, along with suggestions for future developments of the system.

2.2 Methods

Along the research process, a combination of both qualitative and quantitative research methods was adopted.

A literature study was conducted to gain knowledge on augmented reality, worker assistance systems and the state-of-the-art solutions used in manual assembly activities. This literature study allowed to define the main requirements and concepts of the assistance system to be developed. Books and articles from the research databases Oria, Scopus, and Google Scholar were considered. The keywords that were searched in the research databases in several combinations and building blocks are summarized in Table 2.1.

assistance system, worker assistance system, assistive technology
human machine interaction, smart operator
digital manufacturing, industry 4.0
assembly, manual assembly
visualization methods, information visualization
augmented reality, spatial augmented reality
augmented workplace
in-situ projection, in-situ instructions
projector, video projector, high lumen projector
error detection, error-proofing
motion recognition, detection, capture
Kinect, motion sensor, depth sensor, camera

Table 2.1: An overview of the keywords

A case study was conducted by testing the newly developed assistance system. The tests were performed in a laboratory by a group of 15 students and researchers. The test consisted of performing the assembly of a set of products using the Kinect-projector assistance system and other two traditional methods for instruction visualization in manual assembly (i.e. paper-based manuals and a workplace-mounted monitor). The experimental phase, as well as the whole development process, took place in the Logistics 4.0 Laboratory in the MTP department at NTNU. The tests allowed for quantitative data collection. Specifically, the time for completing the assembly tasks and the number of errors performed during the assembly were measured. These measures allowed for an assessment of the performance and accuracy of the assistance system prototype. Further, after completing the assembly tests at the smart workstation, the testers were asked to fill in a questionnaire. The questionnaire asked to give feedback on the three different methods that were used during the assemblies in terms of perceived ease of use, perceived enjoyment, mental workload, effort, and frustration. The results from the questionnaire allowed to assess the Kinect-projector assistance system in terms of user's acceptance and mental workload.

3 Theoretical Background and Related Work

The theoretical background and related work that is relevant for assisting workers during manual assembly work tasks is based on different areas of research in human-computer interaction, mechanical engineering, and psychology. In this chapter, terms and definitions that are used throughout the thesis are introduced. Further, relevant theory and related approaches are presented, and are assigned to the following sections: human workers in Industry 4.0, augmented reality, sensing interaction, worker assistance systems, assembly instructions, parts picking, and assembly errors.

3.1 Terms and Definitions

Industry 4.0
The term Industry 4.0 means using information technology in general for improving industrial manufacturing processes. Further, the improvements can take place in any stage of the process. For instance, using information technology to improve the human-computer interface for a machine with digital components, or using sensors and actuators to monitor and be aware of the manufacturing status at all times. The term itself refers to the 4th industrial revolution, which describes the current trend of automation and data exchange in manufacturing (Hermann et al., 2015).

Augmented Reality
Even though there are a lot of different definitions for augmented reality (AR), in this thesis it is defined as the combination of digital information with real world scenarios or physical objects according to real world circumstances. Thereby, the information can have any modality and is not limited to visual information (Nee et al., 2012).

Spatial Augmented Reality
In this thesis, spatial augmented reality (SAR) is used to describe visual augmented reality information that is registered at a fixed point in the physical space (Raskar et al., 1999). In this way, the technology used for spatially augmenting the reality is not specifically defined, as SAR can be achieved, for instance, by using head-mounted displays (HMDs), projectors, or hand-held screens (Zhou et al., 2011). By using SAR, the position of the

visual information in the space can either be a fixed coordinate, or spatially attached to a movable object.

Assistance System

An assistance system is defined as an interactive system that uses a technology to give instructions or feedback to a worker while performing work tasks. Therefore, the modality that is used for presenting the instructions and the technology that is used to present them is not specified. An assistance system can use one or many modalities and technologies to present instructions or feedback during work tasks. Further, an assistance system can be context-aware, i.e. reacting to a user's actions or tasks (Korn et al., 2012).

3.2 Human Workers and Industry 4.0

The 4th industrial revolution is transforming and will continue to transform the industrial workforce and their work environment. This will have significant implications on the nature of work in industry, as Industry 4.0 will transform design, manufacture, operation, and service of products and production systems (Hermann et al., 2015). At the same time, the demography is changing, especially with the integration of new migrant workers with different skills and educational levels. Because of these challenges, manufacturing enterprises must take a socio-technical system approach and make improvements in order to assist apprentice workers by using advance digital technologies. On the other hand, considering the developments from a technical perspective, new connectivity and interaction technologies among parts, machines and humans will make production systems more lean, agile, traceable, and adaptable (Romero et al., 2016b).

To successfully embrace the Industry 4.0 paradigm in a socially sustainable way, manufacturing enterprises will need to accompany its technological transformations with training and development programs for their workforce. Furthermore, new working environments such as the "cyber-physical factory" will directly affect the operator and the nature of work, creating new interactions not only between humans and machines, but also between digital and physical worlds. Accordingly, Romero et al. (2016b) introduced the concept of Operator 4.0, which can be described as a smart and skilled worker who performs not only cooperative work with robots, but also work aided by machines, by means of human cyber-physical system. Furthermore, a human cyber-physical production system (H-CPPS) is defined as a work system that improves the workers' abilities thanks to a dynamic interaction between humans and machines by means of intelligent human-machine interfaces. H-CPPS will use human-computer interaction techniques designed to fit the operators' cognitive and physical needs, and improve human's physical, sensing and cognitive capabilities, through various enhanced technologies (Romero et al., 2016a).

One of the technologies that enrich and improve the original human capabilities is augmented reality. Augmented reality (AR) can enrich the real-world factory environment of the workers with digital information and media (e.g. graphics, video, sound, GPS data, etc.) that is overlaid in real-time in their field of view, for instance through head-gear, smartphones, tablets, or projectors. Hence, AR can be considered a key enabling technology for improving the transfer of information from the digital to the physical world of the workers in a non-intrusive way (Romero et al., 2016b). AR technology may offer significant advantages, e.g. faster cycle times, reliability, and reduced failure rate. It can support the workers in real-time during manual operations by becoming a digital assistance system. In this way, it allows for reducing human errors and at the same time reducing the dependence on paper-based work instructions, computer screens, and thus the operator's mental workload (Gorecky et al., 2013). For example, AR can provide intuitive information and enable digital error-proofing systems for work-intensive tasks in order to reduce defects, rework and redundant inspection, whilst improving the quality of work. In the following section, AR applications from the literature are introduced.

3.3 Augmented Reality

Augmenting the reality with information goes back to Sutherland (1969). In his head-mounted prototype, he overlaid the view of participants with objects that are close and objects that appear to be far away. According to Milgram and Kishino (1994), AR starts from real environments and successively adds additional information into a real world scene. This overlaying of real world scenes can be applied to any scene in any context. For instance, in a desktop scenario, Wellner's DigitalDesk (Wellner, 1991) was the first system that combined a camera and a projector for creating an augmented table that could merge digital information with physical objects that are placed on the desk. The DigitalDesk uses an RGB camera to detect the position of a paper on the desk and to detect where a user is pointing. Further, a projector is used to highlight information directly on papers that are placed on the desk.

The idea of augmenting work processes with visual information has been around for more than two decades. In 1992, Claudel and Mizell (1992) suggested using HMDs for displaying drilling spots and instructions for a manufacturing task. Over the years, research has defined sub-categories of augmented reality according to the different use cases and the ways of presenting information. For example, SAR (Raskar et al., 1999) is referred to an object that is being displayed directly on the physical space around the user. An example of SAR is the Everywhere Displays Projector (Pinhanez, 2001), where information is projected directly into the physical world by means of a projector and a rotatable mirror. Further, studies by Nee et al. (2012) show that AR can be used to support almost every aspect of a manufactured product's life-cycle: from design and training to manufacturing and inspection. For instance, Zhou et al. (2011) use in-situ projection for highlighting welding spots in manual welding tasks for quality control. Moreover, Raskar et al. (2003) created a geometrically aware camera-projector system where the projected images are transformed and corrected to be viewed without distortion even on non-planar surfaces. They further used their system to project feedback onto picking bins. Schwerdtfeger et al. (2008) use head-mounted laser projectors to display information in a welding context. Their findings comprise that head-mounted projectors are too heavy to be used in long-term tasks e.g. at workplaces.

On the other hand, in-situ projection has already been used in other domains to teach and instruct learners. For example, in the domain of learning how to play instruments, Weing et al. (2013) used in-situ projection on a piano to support learners in playing the piano.

3.4 Sensing Interaction

Apart from presenting information, sensing interaction is the most important aspect of building an interactive assistance system. Traditionally, interactive systems are operated by using graphical user interfaces following the WIMP (windows, icons, menus, pointer) paradigm. However, more recently user interfaces for interactive systems have been proposed to be made tangible (Ishii and Ullmer, 1997), or additionally allow for being operated using natural interaction (Jacob et al., 2008).

A major part of creating a natural user interface is the possibility to recognize gestures. For detecting two-dimensional gestures, touch events need to be detected by a surface. An example is the Touchlight system (Wilson, 2004), which uses two RGB-cameras to detect touch input on a projected surface. Thereby, a user is able to interact with projected content. With the proliferation of Microsoft Kinect depth cameras in 2010, sensing touch on projected interfaces became easily possible on arbitrary surfaces. Therefore, Wilson (2010) suggested an algorithm that observes the depth data in close proximity to surfaces. Whenever a user touches a surface with a finger, a touch event can be detected.

For detecting gestures in three-dimensional space, a user usually has to be equipped with a sensor or carry a sensor (Schlömer et al., 2008). However, in the domain of assembly, detecting 3D gestures with their full trajectory is not needed for interacting with an assistance system. For example, Bannat et al. (2008a) present a framework using a top-mounted RGB camera to detect bins. Once the position of the bins is known, their system uses the RGB camera to detect the position of the worker's hand. Thereby, the 3D movement of the hand is simplified to just using the current position of the hand and defining interactive zones in the camera image. In their system, assembly instructions are shown on a monitor close to the work area, while the bins to pick parts from are highlighted by using a top-mounted projector. Korn et al. (2013a) extended this approach by using a top-mounted depth camera instead of an RGB camera and a top-mounted projector in production environments. The position of the bins and the position of an assembled part have to be defined manually using a graphical editor. Their system then highlights the bin to pick parts from. As their system cannot automatically detect the correct assembly in each step, it uses projected buttons so that the user can manually advance the projection to next steps. Instead of augmenting the assembly parts, other research proposed mobile systems by augmenting the users with sensors. For example, Ward et al. (2006) equipped the user with body-worn microphones and accelerometers to infer the user's current activity in an assembly environment. However, a body-worn system unfortunately cannot detect if a part is assembled correctly.

Overall, previous work on sensing interaction uses either 2D surfaces for detecting gestures that are performed on the surface, uses body-worn sensors to detect gestures that are performed in 3D space, or creates simplified abstractions from 3D trajectories. The sensed interaction is not only used to directly interact with systems but also to indirectly monitor the user's actions, such as picking from bins or assembling a part in a certain place.

3.5 Worker Assistance Systems

Assistance systems for workplaces have been proposed to facilitate collaborative work, give a continuous support to the workers, and for providing cognitive assistance during complex tasks. These systems are implemented using many different technologies. One of these technologies is presenting assembly instructions on a mobile display. These mobile displays are either carried or worn by workers during assembly tasks. Echtler et al. (2004) use a display that is mounted directly at a welding gun to provide information about the exact position of welding spots. Other research suggested presenting assembly instructions using chest-worn displays, nearby screens, mobile phones, tablet computers, or smartwatches.

Considering stationary displays, Korn et al. (2013b) conducted a study with impaired workers, where they compared in-situ pictorial instructions to instructions that are presented on a nearby screen. They found that pictorial in-situ instructions lead to a faster assembly, but workers were making more errors. Also Marner et al. (2013) compared in-situ projected instructions to instructions that are shown on a screen. They conclude that in-situ instructions are faster and lead to less errors.

Other assistance systems using AR for manufacturing are the ones presenting instructions on HMDs. For example, Tang et al. (2003) showed that spatially overlaying the assembly workplace with AR instructions using an HMD reduces the error rate in assembly tasks by 82% compared to paper-based instructions.

Other assistance systems focus on using in-situ projection to display information directly onto the workplace. An assistance system using a top-mounted projector and a top-mounted camera was introduced by Bannat et al. (2008b). They use an RGB-camera to detect which bin the worker is picking the parts from. Further, they equipped the worker with a grasping sensor. This sensor ensures that the worker actually picked up an item from the bin and that the system did not just register the position of the worker’s hand above the bin. Korn et al. (2012) suggested using motion and voice input for sensing and triggering events at an augmented workplace using in-situ projection. Further, they suggested to use gamification elements in order to motivate workers during their tasks. More recently, Büttner et al. (2015) presented an assistance system using a top-mounted projector which is displaying picking information directly onto the picking bins. In their use case, the workers perform all assembly steps in their hands without using a workpiece carrier, while the assembly instructions are projected onto an instruction area at the workplace. Further, their system provides a foot pedal, which the worker can press to advance to the next work step.

3.6 Assembly Instructions

When designing assistance systems for providing instructions, the design of the presented instructions is very important. Several projects focused on how to generally visualize instructions. One of the ways to provide instructions for manual assembly tasks is using textual descriptions. However, compared to text, pictorial instructions are more widespread as they are language independent. Lancioni et al. (2000) experimented with pictorial instructions for performing tasks. In a study, they compared instructions on a computer-aided palm device with instructions on cards. Participants using the computer-aided palm device to view the pictorial instructions performed better and they were also preferred by the participants. Korn et al. (2013a) and Bannat et al. (2008b) used a camera-projector system. In their systems, they both use pictorial instructions in a manufacturing environment for assembling LEGO models. The images used in their projected instructions look exactly as the ones in printed manuals.

Another branch of research focuses on how to build easily understandable pictorial instructions (Agrawala et al., 2003). Studies suggested building hierarchical pictorial instructions where the reader can see the action that is being performed. Step-by-step instructions enable the reader to better identify the step that is being performed. Furthermore, the parts should be oriented in a way that all important features are visible to the reader.

Considering video-based instructions, Rüther et al. (2013) use video-based interactive in-situ instructions using a projected user interface by means of a camera-projector system. Moreover, Suzuki et al. (2016) use in-situ projection for displaying the hand movements of expert workers. In this way, novice workers can learn assembly tasks by mimicking the hand movements of expert workers.

3.7 Parts Picking

Systems for supporting workers during parts picking tasks and systems supporting users in finding objects have been the topic of various studies. Li et al. (2012) used a stationary Kinect together with computer vision algorithms to identify picked objects based on their shape and visual appearance. In their approach, the worker has to explicitly place the object in front of the camera, which results in an extra work step and might increase the TCT. More recently, Bächler et al. (2015) investigated how beneficial a parts picking system using in-situ projection might be for workers with cognitive impairments. Their results reveal that 85.9% of the interviewed persons benefitted from an interactive system. Furthermore, in a comparative study, Bächler et al. (2016) evaluated four different picking visualizations: Pick-by-Projection, Pick-by-Light, Pick-by-Display and a Pick-by-Paper baseline. Their results reveal that the Pick-by-Light method was significantly faster than the other methods used in their study.

The parts picking task is so far the only scenario where long-term evaluations of HMDs have been conducted. For example, Schwerdtfeger et al. (2009) tested their approach in a two-hours study to get insights about long-term usage of AR in production environments. After using the HMD for two hours, the participants reported headaches, problems to focus on the instructions shown on the HMD, and they needed a 15-minute break from using the HMD.

3.8 Assembly Errors

Errors in the assembly process are caused by many factors, which can be categorized in assembly system factors (e.g. high repetitiveness of tasks, poor ergonomics), product factors (e.g. products with many or similar components, high variety), and operator factors (e.g. the worker’s memory, mental ability, training level, experience) (Michalos et al., 2013). Assembly errors can cause increases in production time and cost, production waste and a deterioration in the quality level of the product, resulting in serious damage to the entire production system. To minimize the number of manufacturing defects in the assembly process, these factors must be analyzed in order to identify tools that reduce the probability of human errors (Dalle Mura et al., 2016).

Several methods have been implemented in the industry to face these problems. In recent years, the scientific literature has been mainly focused on the development of methods involving sensors and AR. Sensors can be positioned on the arm (or hand) of the operator or, alternatively, on the tool. The analysis of the movement is realized by devices that are capable of transforming kinematic and dynamic quantities into electrical nature quantities, which can be captured, digitalized, and then processed by a computer (Hartmann, 2011). A studies by Zaeh et al. (2009) has shown that providing spatial information to the operator is a good starting point for the development of a valid support to the assembly activities. Dalle Mura et al. (2016) proposed a system based on the combination of a force sensor and AR equipment. This system gives to the worker the necessary information about the correct assembly sequence and alerts him/her in case of errors, leading to significant improvements compared with traditional methods for preventing and correcting human errors in assembly processes.

It is evident that several technological solutions of different nature are currently available to reduce the probability of human errors during the assembly process. AR in industry is still in an experimental phase, thus the field is still open to the investigation of new applications. In particular, AR can guide the operator to perform the correct action, also providing a support to recover any committed errors. An issue to further development relates to the integration of AR with other systems such as sensors, in order to create a synergistic system in which the limits of one may be filled by the other (Dalle Mura et al., 2016). The aim of this thesis is therefore to propose a configuration of manual assembly workstation based on the use of a sensing device and augmented reality equipment (i.e. Kinect-projector system), able to guide the actions carried out by the worker.

4 Requirements and Concepts

In order to design and develop an assistance system, the requirements must be first identified. The human worker is the subject who is meant to be assisted by the system and who is going to interact with the system. Therefore, the interaction concepts must be also identified. The requirements and concepts that are introduced in this chapter are technology independent and do not rely on the underlying hardware of the assistance system.

To support workers appropriately at the workplace with an assistance system, the system needs to be able to distinguish between correctly performed work steps and being able to detect errors (e.g. picking the wrong part or assembling a part at the wrong place). The fact that an assistance system can distinguish between correctly and incorrectly performed steps might be one of the biggest benefits of an assistance system (Funk et al., 2014). It takes the decision whether a task was correctly performed or not away from the worker, and automatically decides about the completion of the task. This will lead to a reduction of the cognitive effort that is needed by the worker at the workplace.

Regarding the visualization of the instructions, the system should provide instructions that are easy to understand. In fact, understanding the instructions should result in the least possible cognitive effort.

4.1 Use Cases

Assistance systems might have a great potential for different areas of applications in the whole process of manual assembly. Therefore, use cases (U) are defined. These use cases would benefit from using an assistance system and, thus, they will be addressed in the thesis.

U1: Continuous worker support
Assistance systems can be used to continuously support the workers during a task. This could be beneficial if the task is very cognitively demanding (e.g. when producing different product variants in lot size one) or when a worker needs continuous support (e.g. inexperienced workers).

U2: Training workers for a new task
Another use case for assistance systems at the workplace is the training of new workers or already experienced workers, in learning how to execute

new tasks. Instead of learning from another worker, the assistance system teaches the worker directly on the task by providing context-aware instructions and reacting upon errors that are made.

4.2 Requirements

The requirements (R) for an assistance system at the workplace are defined. These requirements are general requirements that an ideal assistance system should have, regardless of the underlying technology of the system (Korn et al., 2012). The requirements are either functional requirements, which describe a function that is necessary for the assistance system to work, or non-functional requirements, which cannot be directly quantified but are important for the maintenance and long-term usage of the assistance system.

4.2.1 Functional Requirements

R1: Provide understandable instructions

The goal of an assistance system is to provide instructions at the workplace to help workers performing the task. The presented instructions should be easily understandable, focusing on the currently performed task, should be visualized at the right position and at the right time, resulting in no unnecessary motions by the workers, should be context-sensitive (i.e. reacting upon the worker's actions and errors), and should result in no additional cognitive effort.

R2: Provide intuitive user interface

The worker should interact with the assistance system in a simple and intuitive way. The user interface should result in very little additional cognitive effort and should not generate frustration to the worker.

R3: Detect picked parts

Picking parts is an activity that needs to be performed at many workplaces where manual assembly tasks are done. Parts are usually stored in picking bins, shelves, or storage boxes. An assistance system would need to detect when a worker is picking a part.

R4: Detect incorrectly picked parts

An assistance system should also detect if a worker is picking a wrong part from the storage box. The system should warn the worker when it detects that a wrong part is picked. In this way, an error can be prevented before a wrong part is assembled. This is especially useful if a task consists of many parts that are similar and can be easily mixed up.

R5: Detect correct assembly
The assistance system needs to detect when a task is correctly performed. One of the tasks at an assembly workplace is the assembly of parts. Thus, a requirement for the assistance system is to be able to detect the correct assembly of parts. The system should be able to check if the assembly is in its defined final position after each work step. Only when correct assembly can be detected, the system can provide feedback after a step has been performed correctly and advance to the next task.

R6: Detect assembly errors
A work step should not be advanced until the system can detect that it was performed correctly. As the same is true for presenting picking errors feedback, the system has to be able to detect assembly errors. In this way, the system can present error feedback when an assembly mistake is made.

4.2.2 Non-Functional Requirements

In addition to the previously defined functional requirements, the assistance system should also include some non-functional requirements. First, the system should run stable and reliably as it will be deployed in an assembly environment. Depending on the shift plan of the factory where the assistance system is used, the minimum run time will be eight hours per workday, i.e. one shift per day. As some factories produce in up to three shifts per day, the assistance system needs to be available up to 24 hours per day. Another requirement is that the system should not need being re-calibrated during the use. Once a calibration is set, it should be valid until a parameter is changed. Considering maintenance, the assistance system should be also easy to maintain.

4.3 General Architecture

According to the previously outlined requirements for assistance systems, a general architecture for implementing an assistance system is introduced. Figure 4.1 shows an overview of the general architecture. Three different types of components are considered: motion recognition components, feedback components, and logical components.

The motion recognition components include a pick detection (R3, R4), an assembly detection (R5, R6), and a user interface (R2). The feedback component consists of providing understandable instructions and feedback (R1). Lastly, the logical components include defining a workflow for the assembly process. This architecture combines all these components creating the building blocks for an assistance system that fulfills the requirements. As of now, the building blocks are generic, independent from the use case, and independent from the underlying technology.

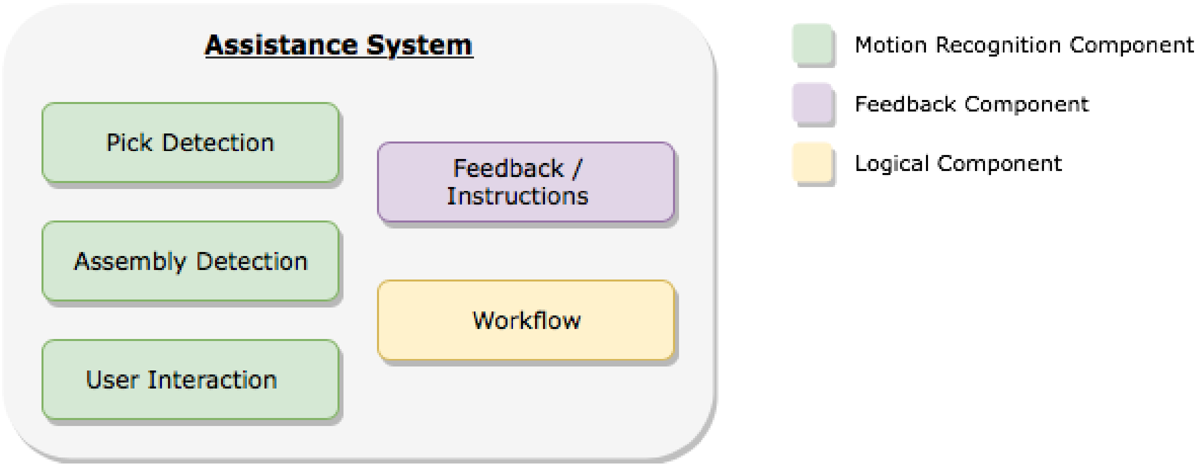


Figure 4.1: A general architecture for an assistance system

4.4 Interaction Concepts

The previously introduced requirements and the general architecture allow for using the motion recognition components as an input for interacting with the assistance system. Overall, the design goal of the system is to use as many natural and intuitive interactions as possible (Funk et al., 2014). The following paragraphs describe two interaction concepts that can be implemented in the assistance system: implicit interaction with instructions, and explicit interaction with projected buttons.

4.4.1 Implicit Interaction with Instructions

During the daily use of the assistance system, workers must interact with the system implicitly, just by performing the work steps that they would normally do without using an assistance system. This interaction concept does not require the worker to operate a graphical user interface anymore (e.g. on a touch screen display) as the interaction with the system is simply based on performing physical actions at the workplace. The implicit interaction concept also uses the motion recognition components that were outlined in the system's general architecture (Figure 4.1). In this way, the system can advance instructions when a worker picks from a correct bin, or assembles a part correctly. In contrast, if the worker picks a part from an incorrect bin or an error is made in the assembly, the system is able to implicitly display an error warning. The error state is exited again if the worker either picks a part from the correct bin or resolved the assembly error. Using this concept, the workers are not required to acquire knowledge about operating the assistance system, as they interact only by performing actions on the workplace.

4.4.2 Explicit Interaction with Projected Buttons

For some complex assembly tasks, the assistance system might not be able to detect whether the part was correctly assembled or not. If the system cannot assess the correctness of the worker's assembly, then the system cannot automatically advance to the next work step. For this reason, an explicit interaction concept should be implemented. This interaction concept would allow the worker to deliberately advance to the next assembly step, and even go back to the previous assembly step, in case a corrective action is to be taken. A projected button is the object dedicated to this type of interaction. It is a visual representation of a button which is projected on the workstation's desk, and each of them are linked to a specific function (e.g. "go to next work step", or "go to previous work step"). As the assistance system is able to detect motion in particular areas of the workplace, then the system can detect whenever the area of a projected button is touched by the worker's hand, in the same way the worker would press an actual button. This explicit interaction results in an additional cognitive effort for the worker, but it is necessary when the implicit interaction is ineffective and the assistance system lacks quality control of the assembly.

5 System Design and Implementation

In this chapter, the previously outlined requirements are addressed and an implementation for the introduced concepts is provided. As for implementing the requirements, a top-mounted camera-projector system for providing activity recognition and giving in-situ projected feedback was used, similarly to related approaches (Bannat et al., 2008b, Korn et al., 2013b). In fact, the required technology was totally mounted at the workplace instead of requiring workers to wear sensors, allowing for better working conditions for workers. Figure 5.2 shows the physical design of the assembly workstation equipped with the Kinect-projector assistance system.

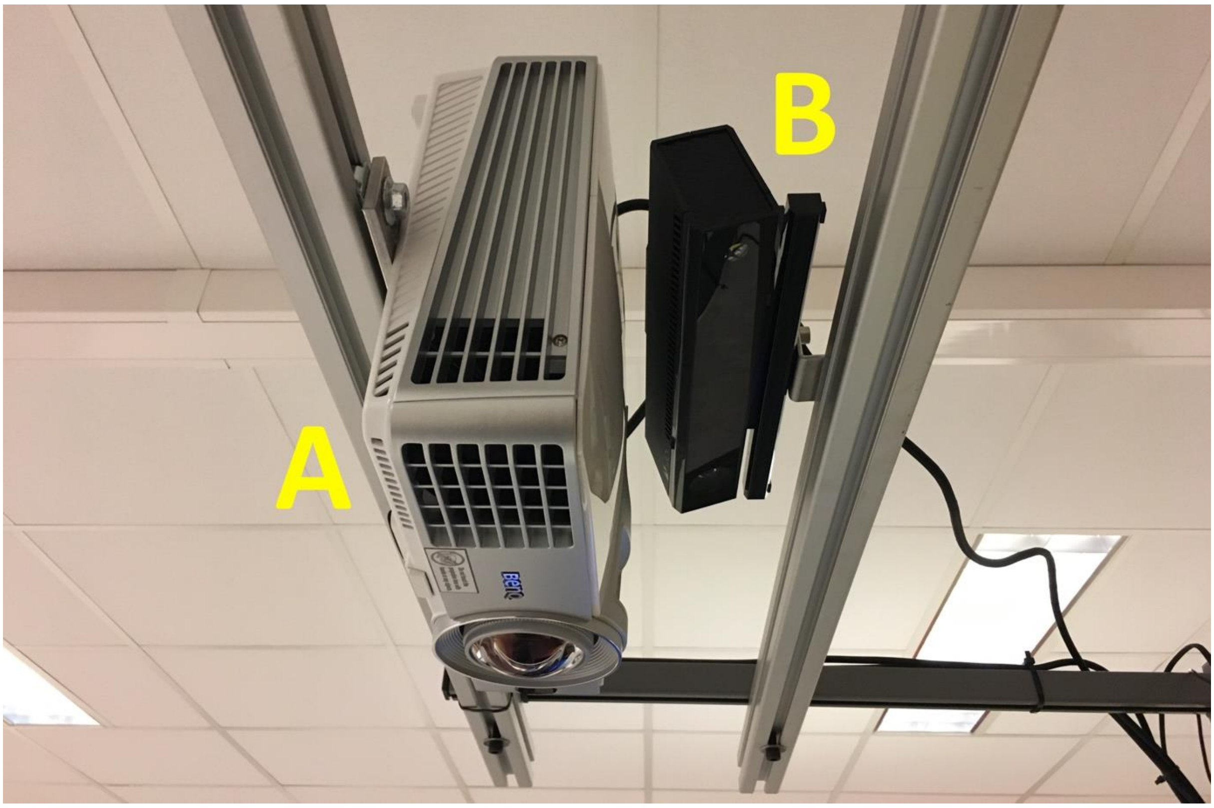


Figure 5.1: The top-mounted projector (A) and Kinect (B)

A top-mounted depth camera (Microsoft Kinect for Xbox One) (Figure 5.1 (B)) is observing the work area and the bins where the parts are stored. The Kinect's depth image does not rely on the environment's light. It has its own infra-red laser projector, making the sensor immune to light changes in the environment (e.g. shadows, bad lighting, lights on/off in a room).

Further, a top-mounted projector (BenQ MW632ST) (Figure 5.1 (A)) is able to project in-situ feedback onto the storage bins and the work area. The projector has a brightness of 3200 lumens, making it suitable for delivering bright and clear images even in rooms with ambient light.

Both, depth camera and projector, need to be mounted at a distance so that the boxes and the work area are both covered by the projection area and the field of view of the depth camera.

In the following, the assembly workstation's physical design is presented. Then the feedback projection system is introduced, as well as the implementation of the software components for the motion recognition. Finally, the workflow logical component is presented. These elements are designed technology-dependently, i.e. requiring a camera-projector setup.

5.1 Physical Design

The manual assembly workstation is the place where a worker is working autonomously and performs a defined number of work steps. This manual assembly workstation is considered to not have direct dependencies on other workplaces, as work pieces required for the assembly are started to be assembled at this workstation.

The physical design of the manual assembly workstation (Figure 5.2) is based on a regular assembly workbench as currently used in production environments. The workbench used for this setup has a surface of dimensions 120x60 cm. The setup of the assistance system consists of a Microsoft Kinect for Xbox One and a BenQ MW632ST projector that are mounted on top of an aluminum construction in a way that they are both facing downwards (see Figure 5.2 (A)). They are mounted at 102 cm and 92 cm above the work area, respectively. Further, a variable amount of picking bins are placed on a mobile bins rack at the back of the assembly area, which represent the picking area (see Figure 5.2 (B)). The system uses the depth data that is recorded in the picking area to check for correct picking using the motion recognition software. This setup allows for a maximum of 8 bins per raw, and a maximum of 2 rows of bins. In fact, the picking bins can be stacked on top of each other, as the pick detection also works when there is a difference in height between two picking bins. Also considering the in-situ projection, stacking the picking bins in a 2-row configuration is feasible as the projector is mounted with an angle in a way that the projection will not be occluded by the top-most picking bins. For this setup, small picking bins with dimensions 18x10x7 cm are used. On the bottom of the assistance system, there is an area for assembling the parts. This is called the assembly area (see Figure 5.2 (C)). Further, the areas on the right and left sides of the workbench represent the interaction areas (see Figure 5.2 (D)). Here, augmented buttons with specific functions are projected, and the motion recognition system can detect when they are pressed by the worker. Finally, a touch screen monitor (Microsoft Surface Pro) is mounted with an arm support on the left of the workbench, and it is used to configure the system (see Figure 5.2 (E)).

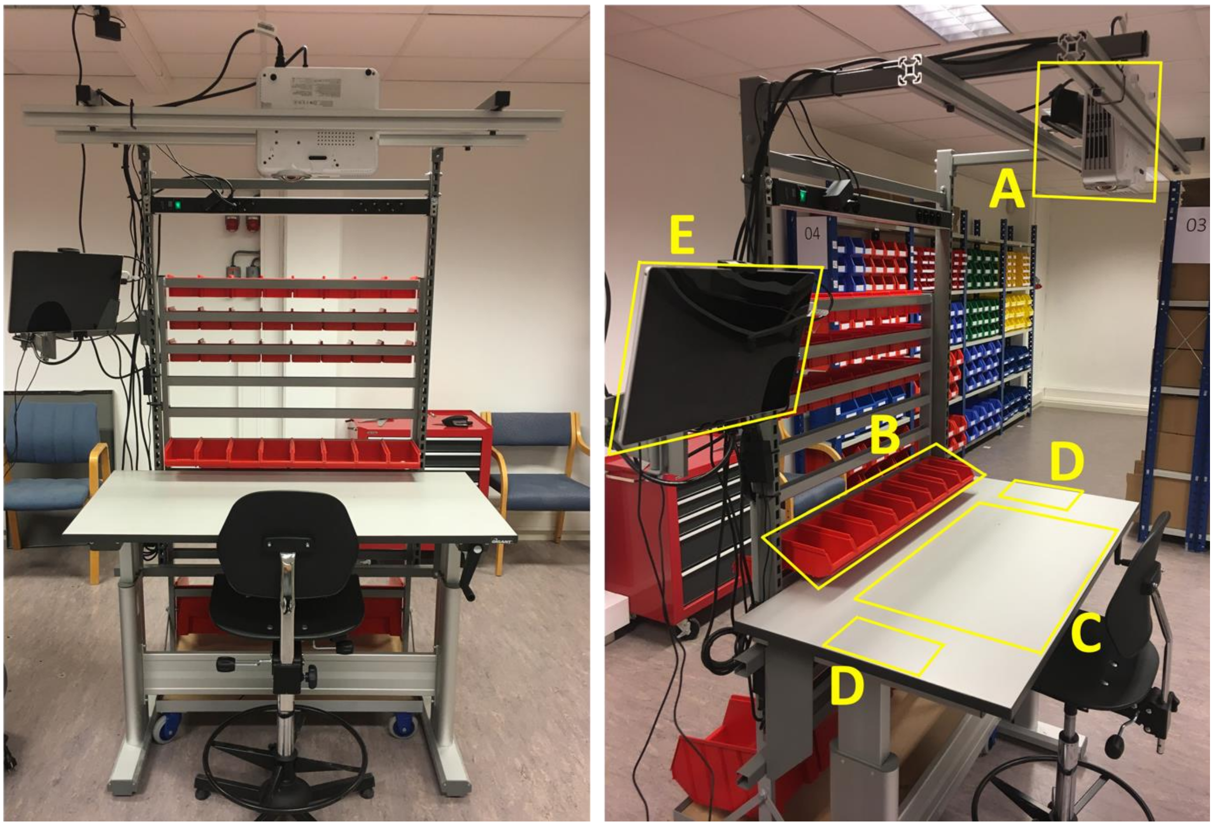


Figure 5.2: The physical design of the smart workstation

In this setup, both the depth image of the Kinect and the projection of the projector are covering the picking area, the assembly area, and the interaction areas. However, there is a limit to the number of rows of bins that is possible to set in the picking area, as the projection in this configuration can cover up to 2 rows.

5.2 Feedback through In-Situ Projection

In-situ projection is used as the method for giving instructions. The top-mounted projector is used to highlight the correct picking bins, and to project the assembly visual instructions on the workbench surface and the projected buttons at the right and left sides of the workbench surface.

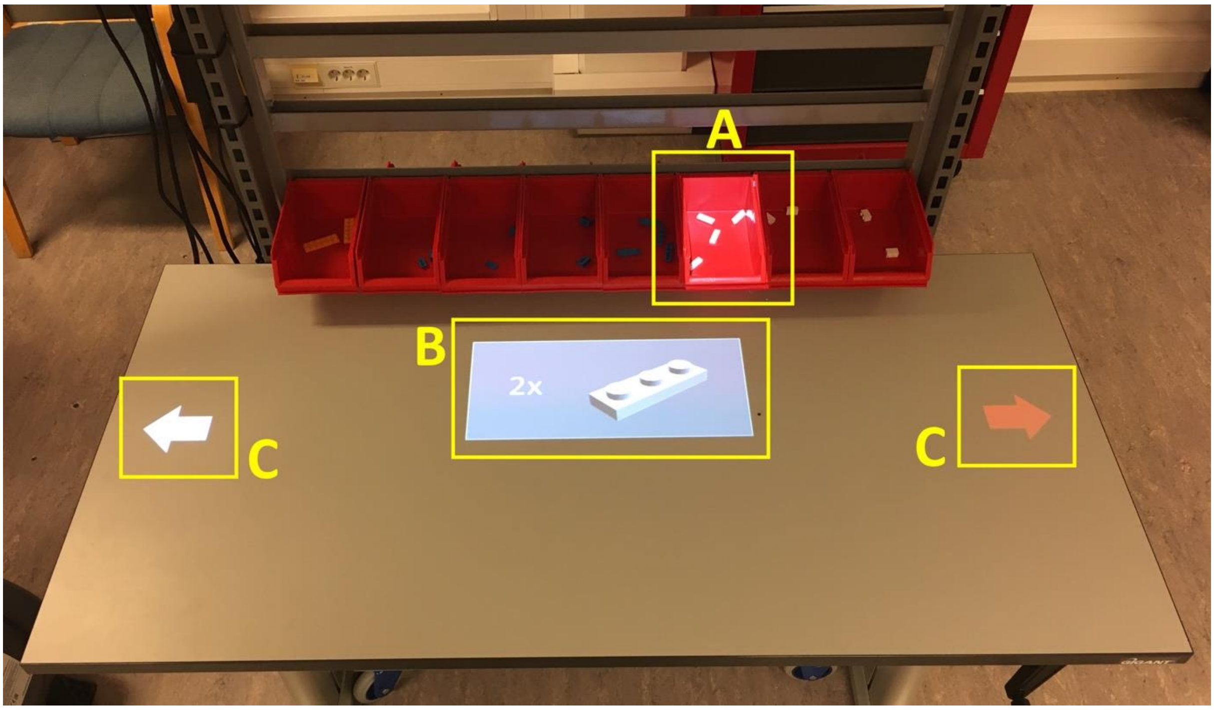


Figure 5.3: Example of in-situ projection instructions

To give picking instructions, a Pick-by-Light approach is used (Figure 5.3 (A)). The picking bin is highlighted by a white light beam, which indicates which bin the worker needs to pick the parts from. If a picking error is made, a red-stripe warning is projected on the whole picking area, while having the correct picking bin highlighted by the white light beam.

The assembly visual instructions are projected at the center of the workbench surface, in a way that they are placed right in front of the worker and are not covered by the worker's motions while performing the assembly (Figure 5.3 (B)). In this way, the worker does not have to turn his/her head to look at a side monitor. Instead, the worker can easily see the instructions in front of him/her, right where the instructions are needed. The visual design of the assembly instructions is simple and easy to understand. No text is used. Instead, arrows and visualizations of the parts are combined to create easy to follow assembly instructions.

The projected buttons are placed at the right and left sides of the workbench surface and they are needed by the worker to confirm a picking or assembly step and advance to the next work step (Figure 5.3 (C)). The position of the augmented buttons is quite intuitive: a "Right Arrow" button is placed at the right side of the workbench and, if pressed, it confirms a work step and advances to the next one; a "Left Arrow" button is placed at the left side of the workbench, which activates the function of going back to the previous work step. They have been placed in a position in a way that they are at an arm's length distance, thus easily reachable by the worker; and at the same time, they are not interfering with product assembly activities. The augmented buttons are red to indicate that they cannot be pressed because the correct part has not been picked yet. They are green to indicate that the correct part was picked, so now the button can be pressed to advance to the next work step.

These instructions and feedback projections are created in Microsoft PowerPoint. They are created in a way that each PowerPoint slide represents a step in the workflow, i.e. picking the part, picking error, confirm the picking, and assemble the part. The slides consist of a black background and a series of objects, shapes and images that represent the instructions or feedback projections. For instance, the picking instruction is represented by a white quadrilateral which is shaped appropriately in a way that its projection fits the shape of the storage bin. A template was developed in order to create all the different slides for a particular product in a fast and intuitive way. Finally, the slides are exported from PowerPoint to a set of JPEG images, which then are showed through the projector.

5.3 Software for Motion Recognition

As it was chosen to equip the workplace with sensors rather than requiring the worker to be equipped with body-worn sensors, the depth image of the Microsoft Kinect is used to detect activities that are performed at the workstation.

The software that was chosen for the motion recognition is called Webcam Zone Trigger Pro (OmegaUnfold, 2016b), developed by Omega Unfold, a Canadian R&D software company. It is a rather intuitive and easy to use software that is able to detect motions via a depth camera source, and to match a motion in a particular *zone* of the image to a desired *action*. The zones where the software looks for motion are called Hot Spots. A description of a Hot Spot and an Action is presented in the following.

Hot Spot

A Hot Spot is a zone on the image that is specified by the user. This is where the software will look for motion. It is possible to define many Hot Spots, and each spot can have a different action to execute when it detects motion. When the video source is a depth image, the Hot Spots are in “3D Presence” mode, meaning that both their distance from the camera and their depth can be set as desired (see Figure 5.4 as an example).

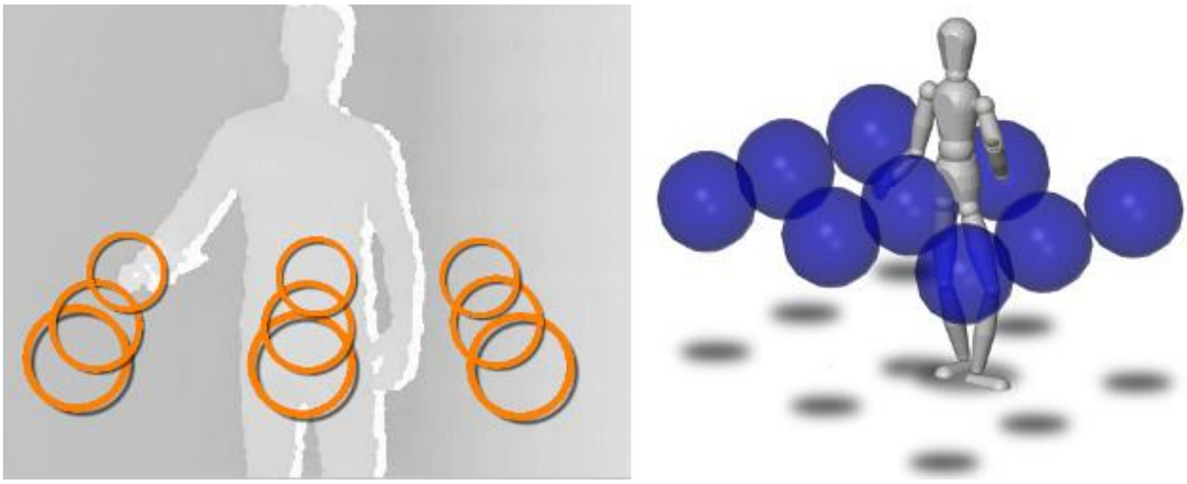


Figure 5.4: Example of 3D Hot Spots in the depth image (OmegaUnfold, 2016a)

Action
Whenever something moves in or out of the space occupied by the Hot Spot, the software will execute an Action. Possible actions that can be implemented with the software are: execute keyboard hot key, play sound, run program, count, record video, etc.

With this software, the assistance system is able to perform two types of activity recognition using the depth image: pick detection (R3, R4) and augmented buttons detection (R2). However, this implementation is not a perfect reflection of the requirements presented in Chapter 4. It suffers from two flaws: there is no recognition of correct assembly (R5) and no recognition of assembly errors (R6). This is due to technical limitations of the software: while the software allows for motion recognition in specific areas of the depth image, it cannot detect whether a part was assembled correctly or not. For this reason, this prototype of assistance system cannot provide feedback on the quality of the assembly. This constraint is mirrored in the realization of the workflow logical component, presented later in this chapter.

5.3.1 Pick Detection

For detecting the picks from the storage bins, one Hot Spot for each bin was created and added to the depth image in the software (Figure 5.5). Thus, a total of eight Hot Spots for the eight bins was created. The Hot Spots were positioned in the area in front of the bins, where the hand of the worker is supposed to pass through when picking the parts from the bins. The diameter of the Hot Spots was set to cover most of the area in front of the bin, while the depth of the Hot Spots was set to 5 cm as suggested by Zone Trigger project lead, in order to prevent noise issues in motion detection.

Because of the possibility to add multiple video sources in the software, the RGB image of the Kinect was used for defining the position of the interactive areas in front of the bins, while the depth image of the Kinect was used for detecting the picks. It is possible to add a new Hot Spot from the Menu, or to duplicate an already existing one. The area of the Hot Spot can be adjusted by dragging and dropping the borders of the circles in a way that they fit the dimensions of the bins.

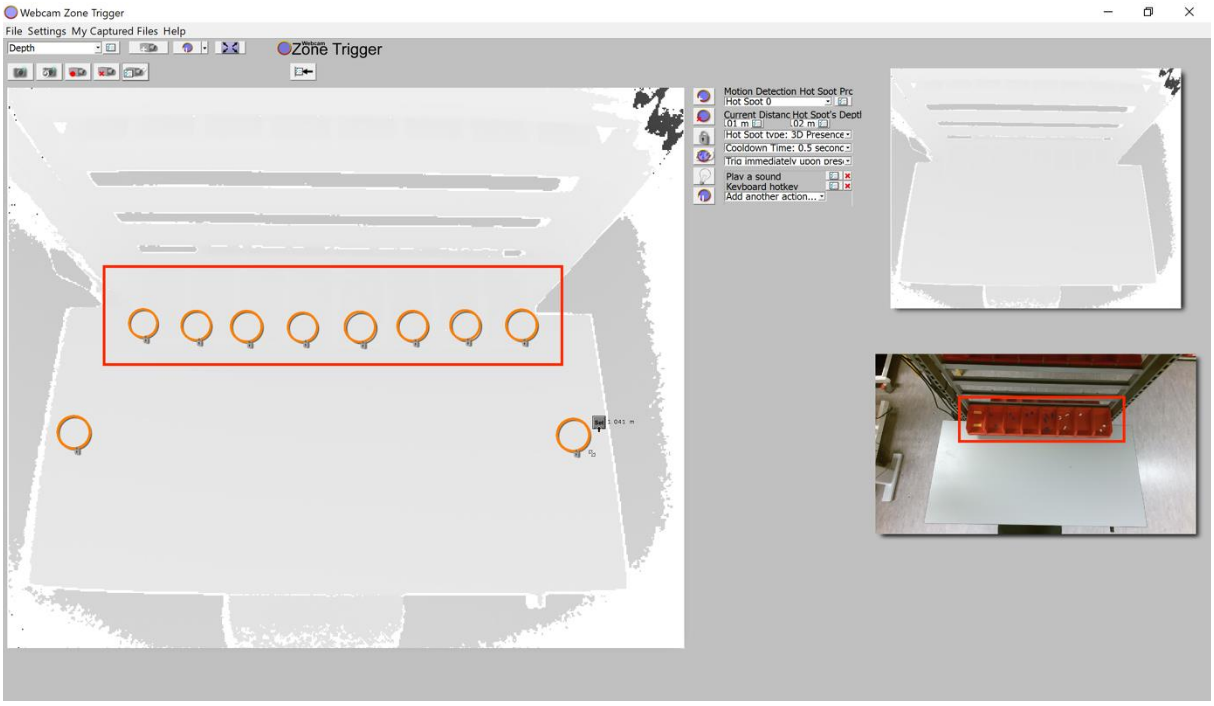


Figure 5.5: The picking Hot Spots in the depth image

As it was mentioned above, each Hot Spot can be linked to a specific Action, which is triggered whenever motion is detected in the Hot Spot three-dimensional space. For the picking Hot Spots, the action "execute keyboard hot key" was chosen, specifically the hot key "Alt+1" for Hot Spot 1, "Alt+2" for Hot Spot 2, and so on until "Alt+8" for Hot Spot 8. This system of different hot keys linked to different picking bins was used in order to manage the correct sequence of work steps in the logical component of the assistance system, which is described later in the chapter.

Considering the trigger of an Action, picking the parts from a box results only in sending the pick trigger once, i.e. when the hand of the operator enters the Hot Spot space. It was designed this way to prevent triggering two consecutive work steps that might have the same trigger.

5.3.2 Projected Buttons Detection

For detecting whether the projected buttons are pressed, one Hot Spot for each button was created and added to the same depth image of the Kinect

(Figure 5.6). Thus, two Hot Spots were added to the right and left extremities of the workbench, in a way that they would be overlaid to the projection of the projected buttons.

The idea is to trigger an Action upon the pressing of the projected button's area on the surface of the workbench. Therefore, the Hot Spots were placed on the workbench surface, with a diameter that was set to cover the dimensions of the button, and a depth that was set to 2 cm. In this way, the Hot Spot detects motion only when the hand of the worker touches the workbench surface, as the worker would do in pressing a real button. Placing the button Hot Spots on the workbench surface means that waving the hand above the projecting button without touching the surface would not interfere with the Hot Spot space, thus no motion would be detected, and no action would be triggered.

Figure 5.6: The projected buttons Hot Spots in the depth image

As it was the case for the picking Hot Spots, the buttons Hot Spots will execute an Action whenever motion is detected in their space. The Action "execute keyboard hot key" was chosen for these two interactive areas as well: in particular, the hot key "Alt+9" is linked to Hot Spot Left, and "Alt+0" to Hot Spot Right.

5.4 Workflow

To show the instructions and feedback in the right order and in a structured manner, a structure called logical loop is defined to represent a logical sequence of work steps in the assistance system. In fact, the logical loop is designed to program the sequence of work steps in a structured manner. Four different types of instruction slides were defined. Finally, in order to show the instructions through the projector and implement the logical conditions, a set of HTML files was created.

Figure 5.7 shows the logical loop, which is the logical component of the assistance system and is responsible for showing the instruction projections in the right order. One loop of instructions represents the sequence of the instructions covering one work step. A work step is considered to include the following physical activities at the workstation: pick the part, confirm that the part is picked, assemble the part, confirm that the part is assembled.

In the logical loop, four different types of instruction slides are defined: Pick Slide, Error Slide, Picked Slide, and Assembly Slide. These four types are also repeated in each work step. They are described in the following.

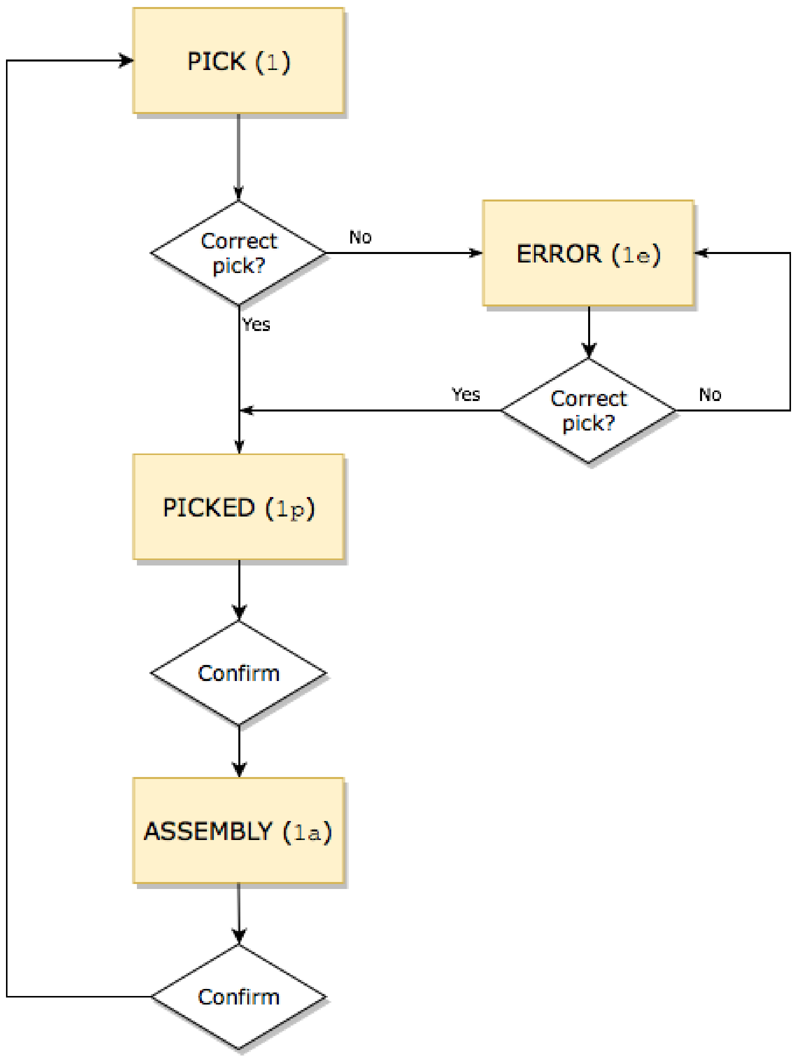


Figure 5.7: The logical loop

Pick Slide

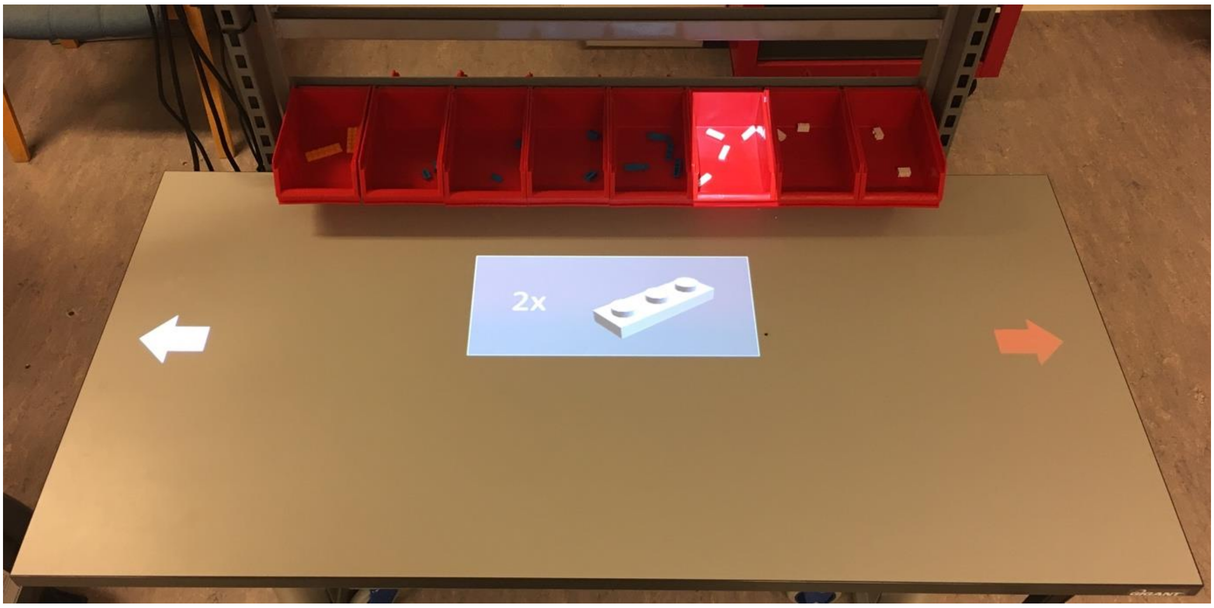


Figure 5.8: Example of a Pick Slide projected onto the workplace

This slide of instructions is the first step in the loop and it shows information on what kind of part needs to be pick, the quantity (e.g. 1x, 2x, and so on), and the position of the bin where to pick the part from. The following code is an example of the HTML file of a Pick Slide. With the HTML Attributes <a> href and accesskey, the logical conditions that allow to advance to the next slides are defined. In the example below, the keyboard hot key “Alt+6”, which is automatically executed by picking the part in storage bin number 6 (i.e. the correct bin), will make the program advance to the Picked Slide. Instead, all the other keyboard hot keys, which are executed by picking the part in storage bins number 1, 2, 3, 4, 5, 7, and 8 (i.e. the wrong bins), will make the program advance to the Error Slide.

```
<!DOCTYPE html>
  <html>
    <body bgcolor="#000000" topmargin="0" leftmargin="0" marginheight="0"
marginwidth="0">
      <div style="height:100px; width=100px">
        <img src="C:/Users/ProdLedLab1/Desktop/sidepod/images/2.jpg" width="1368"
height="912" alt="">
      </div>
      <a href="C:/Users/ProdLedLab1/Desktop/sidepod/2e.html" accesskey="1"</a>
      <a href="C:/Users/ProdLedLab1/Desktop/sidepod/2e.html" accesskey="2"</a>
      <a href="C:/Users/ProdLedLab1/Desktop/sidepod/2e.html" accesskey="3"</a>
      <a href="C:/Users/ProdLedLab1/Desktop/sidepod/2e.html" accesskey="4"</a>
      <a href="C:/Users/ProdLedLab1/Desktop/sidepod/2e.html" accesskey="5"</a>
      <a href="C:/Users/ProdLedLab1/Desktop/sidepod/2p.html" accesskey="6"</a>
      <a href="C:/Users/ProdLedLab1/Desktop/sidepod/2e.html" accesskey="7"</a>
      <a href="C:/Users/ProdLedLab1/Desktop/sidepod/2e.html" accesskey="8"</a>
    </body>
  </html>
```

Error Slide

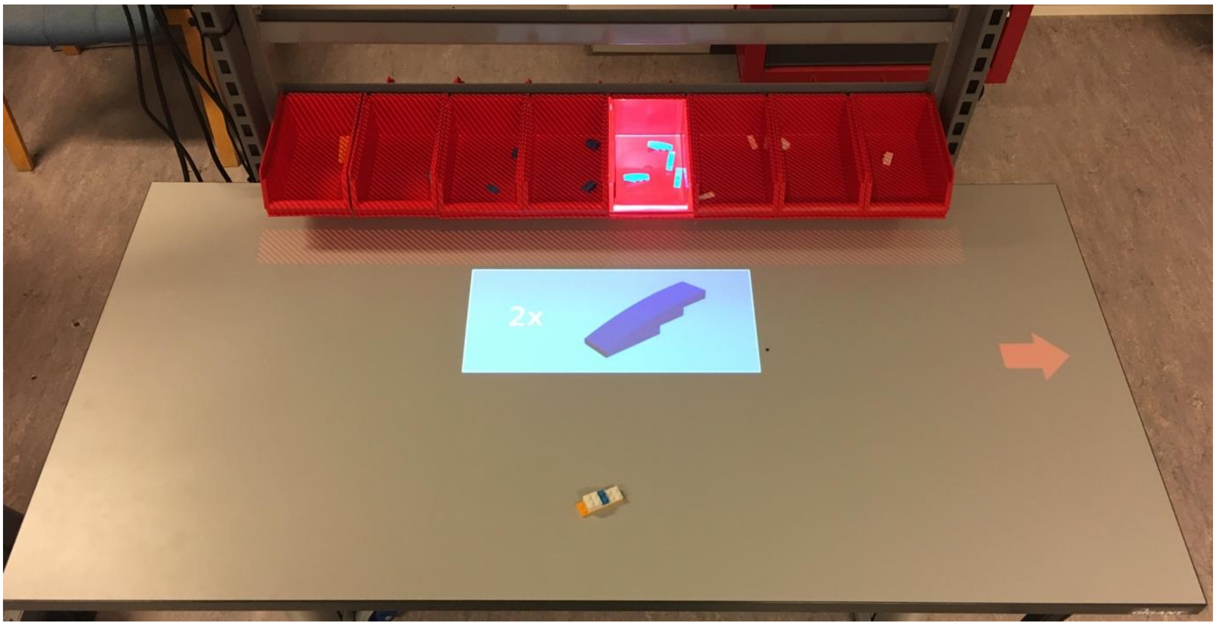


Figure 5.9: Example of an Error Slide projected onto the workplace

In case of a picking error, the assistance system advances from a Pick Slide to an Error Slide. This slide shows the same information of the previous Pick Slide, but with the addition of a red-stripe error warning that covers the whole picking area. An example of the HTML file code of an Error Slide is shown below. The only logical condition states that the assistance system will not advance to the next slide until the keyboard hot key “Alt+6” is executed, which represents the picking from the storage bin number 6 (i.e. the correct bin). When the correct pick is made, the system will advance to the Picked Slide.

```
<!DOCTYPE html>
  <html>
    <body bgcolor="#000000" topmargin="0" leftmargin="0" marginheight="0"
marginwidth="0">
      <div style="height:100px; width=100px">
        <img src="C:/Users/ProdLedLab1/Desktop/sidepod/images/2e.jpg" width="1368"
height="912" alt="">
      </div>
      <a href="C:/Users/ProdLedLab1/Desktop/sidepod/2p.html" accesskey="6"</a>
    </body>
  </html>
```

Picked Slide

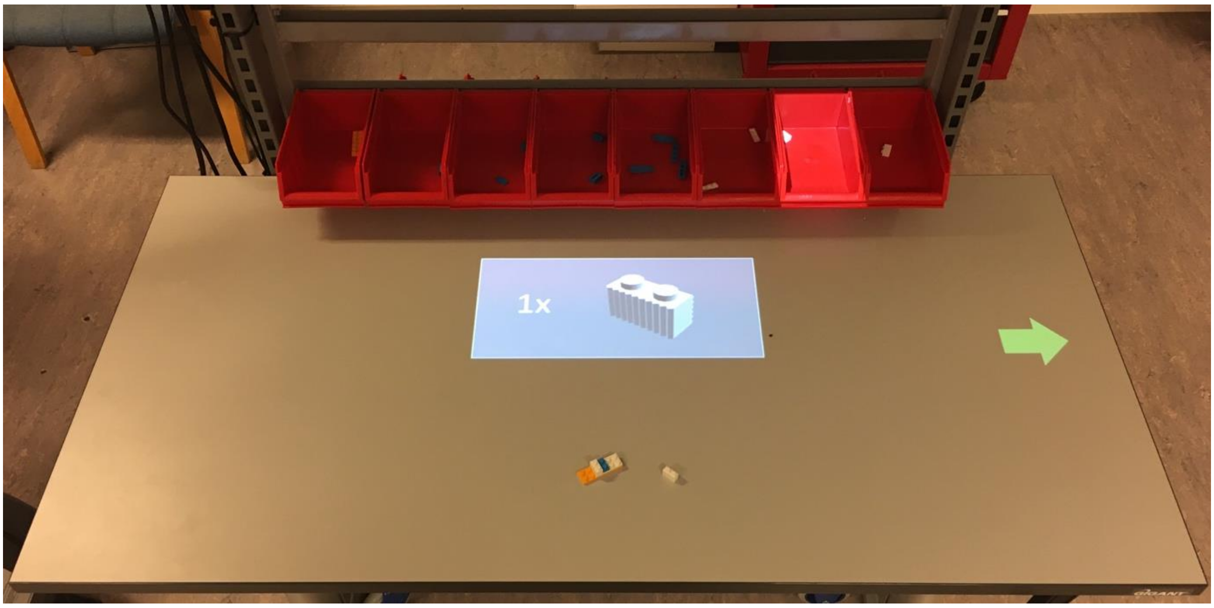


Figure 5.10: Example of a Picked Slide projected onto the workplace

When the correct picking is performed, the assistance system advances to the Picked Slide. In this slide of instructions, the Right Arrow projected button turns its color from red to green, meaning that it can be now pressed by the worker in order to confirm the performed picking of the parts and advance to the Assembly Slide. An example of the HTML file code of a Picked Slide is shown below. The only logical condition states that the assistance system will not advance to the next slide until the keyboard hot key “Alt+0” is executed, which represents the pressing of the green “Right Arrow” augmented button. When the worker presses the projected button, the system will advance to the Assembly Slide.

```
<!DOCTYPE html>
  <html>
    <body bgcolor="#000000" topmargin="0" leftmargin="0" marginheight="0"
marginwidth="0">
      <div style="height:100px; width=100px">
        <img src="C:/Users/ProdLedLab1/Desktop/sidepod/images/2p.jpg" width="1368"
height="912" alt="">
      </div>
      <a href="C:/Users/ProdLedLab1/Desktop/sidepod/2a.html" accesskey="0"</a>
    </body>
  </html>
```

Assembly Slide

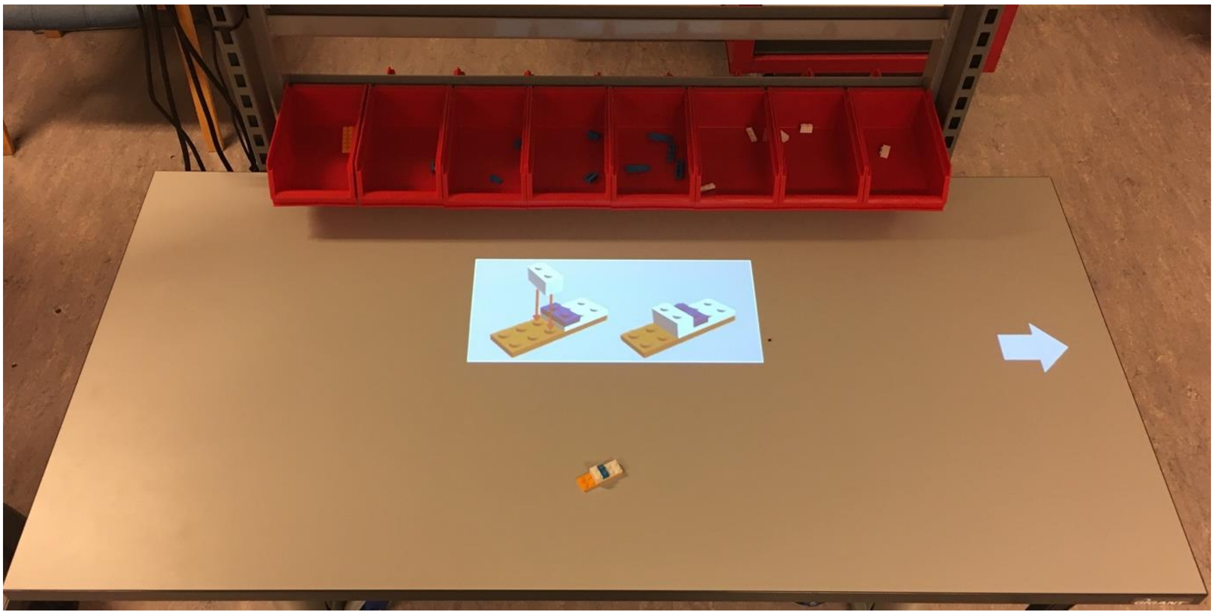


Figure 5.11: Example of an Assembly Slide projected onto the workplace

This slide of instructions is the last step in the loop and it shows information on how to assemble the part. An example of the HTML file code of an Assembly Slide is shown below. The logical condition states that the assistance system will not advance to the next slide until the keyboard hot key “Alt+0” is executed, which represents the pressing of the “Right Arrow” projected button. When the worker presses the augmented button, the system will advance to the next Pick Slide, and the loop will start again from the beginning.

```
<!DOCTYPE html>
  <html>
    <body bgcolor="#000000" topmargin="0" leftmargin="0" marginheight="0"
marginwidth="0">
      <div style="height:100px; width=100px">
        <img src="C:/Users/ProdLedLab1/Desktop/sidepod/images/2a.jpg" width="1368"
height="912" alt="">
      </div>
      <a href="C:/Users/ProdLedLab1/Desktop/sidepod/3.html" accesskey="0"</a>
    </body>
  </html>
```

6 Experimental Evaluation

After developing and implementing the assistance system for a smart manual assembly workstation, a user study was conducted. The aim of the study is to assess the effects of in-situ instructions and motion recognition provided by the assistance system if compared to traditional methods for providing instructions, such as paper-based manuals and workplace-mounted monitors.

The study consisted of assembly tests in a laboratory artificial environment. The data was collected from measurements during the tests and from questionnaires. In the following, the case study is presented. Later in the chapter, the metrics are introduced, and the specific procedure of the tests is described.

6.1 Case Study

To support the assessment of the proposed assistance system, a user case study in a laboratory environment was conducted. The assistance system that was tested is the one that was presented in the previous chapter with the same functionalities: in-situ projection visualization, Pick-by-Light, picking error-proofing, and user interaction with projected buttons.

The purpose of the case study was to assess the potential of the assistance system in terms of productivity and mental workload of the workers if compared with two traditional information visualization methods that are still used in most production environments: paper-based manuals and workplace-mounted monitors. The specific metrics that were measured during the tests are introduced later in the chapter.

The test population consisted of 15 participants, who carried out a series of assembly tests at the workstation by performing two different LEGO assemblies. LEGO construction models were chosen because they do not require specific tools or a certain level of expertise to perform the assembly. Besides, most of the people is already familiar with LEGO bricks and LEGO assemblies.

The workstation configuration used for the test is the same Kinect-projector system described in Chapter 5, with a total of eight storage bins which were organized in a single row. The single row configuration was chosen because of the better robustness of the motion recognition software, compared to the 2-row configuration. This 8-bin and 1-row configuration presented limitations in terms of the different parts that could be used in the assembly tests. In fact, the assemblies were made by using a maximum of eight different LEGO parts.

The smart assembly workstation was developed in the Logistics 4.0 Laboratory at the Department of Mechanical and Industrial Engineering at the Norwegian University of Science and Technology (NTNU), and here is where the tests were performed.

6.2 Metrics

Assistance systems in the industrial domain have to fulfill requirements regarding production performance and human factors. Production performance can be measured by the effects of a newly introduced assistance system on the output. Those effects can be measured by the number of produced goods or their quality. Human factors are defined by ease of use or the mental workload. Both aspects correlate but also conflict to some extent (Loch et al., 2016).

Any assistance system needs to satisfy both aspects to be suitable for practical adoption. However, it has to promise economic benefits at first to justify the initial investment. To take both aspects into account, a methodology that measures performance and accuracy, perceived ease of use and mental workload was developed. The metrics are discussed in the following.

Accuracy and Performance

A measurement for the quantity and the quality of the produced goods with the evaluated systems is necessary. This was addressed by measuring the time to complete a task and by counting the errors made by the tester. Improvements of quantity and quality also yield benefits not only for companies, but also for assembly workers since inadequate qualification, for instance caused by improper instructions, can lead to errors and stress (Loch et al., 2016).

Perceived Ease of Use

A second aim was to determine whether factory workers would accept an AR-based assistance system. This metric is important since workers are more likely to adopt and support a system that they accept. The construct "perceived ease of use" was used for this respect. According to the TAM 3 model (Venkatesh and Bala, 2008), it is one of two determinants of the use intention of a technology. Perceived ease of use is a function from six determinants of which perceived enjoyment was taken into account.

“Perceived enjoyment” is defined as the extent to which “the activity of using a specific system is perceived to be enjoyable in its own right, aside from any performance consequences resulting from system use” (Venkatesh, 2000). Perceived enjoyment was chosen since the working conditions of assembly operators can be characterized by monotonous and repetitive tasks. Hence, the influence of an assistance system that uses modern, probably more attractive, technologies on the use intention was addressed.

Mental Workload

The benefits of an AR-based assistance system, for instance the expected increase of efficiency and quality, are also expected to lower the mental workload of the users. Mental workload summarizes all aspects of the working context that affect the physiological and the psychological condition of an employee (Loch et al., 2016). Research indicates that excessive mental workload leads to an increase of psychological illnesses. The NASA task load index (NASA-TLX) (Hart and Staveland, 1988) is a tool to assess mental workload and rates categories such as mental demand, effort, and frustration level. This tool was also used in comparable experiments by Tang et al. (2003).

6.3 Method

This section presents the participants and the methodology of the experiment. The results of the experiments are presented in Chapter 7.

15 people were recruited for the experiment. The test population consisted of master’s students, PhD candidates and Professors from the Department of Mechanical and Industrial Engineering at NTNU. The participants aged between 23 and 38 years old.

The first part of the experiment consisted of making the tester assemble two different LEGO designs by using three different information visualization methods:

- A paper-based manual,
- A workplace-mounted touch screen monitor (Microsoft Surface Pro), and
- The Kinect-projector assistance system presented in this thesis.

Two LEGO designs (

Figure 6.1) were chosen for the tests and consisted of two components of a LEGO Formula One car: a front wing, which is assembled from 3 different parts and 4 work steps; and a side pod, which is assembled from 8 different parts and 8 work steps. The decision of testing two different designs comes from RQ3, that is to assess the assistance systems with two different product complexities. In this way, it can be assessed whether there is a correlation between product complexity and the metrics introduced in the previous section.

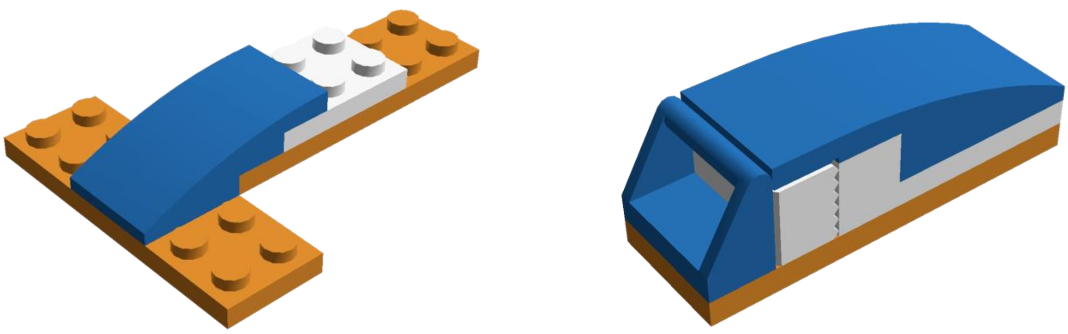


Figure 6.1: The LEGO Front wing (on the left) and the LEGO Side pod (on the right)

The visual instructions' graphics were the same for all three methods and were designed by using LEGO Digital Designer, a computer program developed by the Lego Group that allows users to build model using virtual LEGO bricks. The visual instructions' graphics that were used for testing can be found in Appendix A and Appendix B.

Before starting the test, each participant was instructed to the objectives of the experiment and got to familiarize with the assistance system. Then he/she was asked to complete the following task: "Assemble the LEGO component in the correct way, as fast as possible". The participant was asked to perform the task with all three systems (i.e. paper-based manual, workplace-mounted monitor, and Kinect-projector system).

In the second part of the experiment, at the end of the tasks, the participant was asked to answer a questionnaire based on the metrics that were described earlier in the chapter, in order to assess the perceived ease of use, the perceived enjoyment, and the mental workload. A copy of the questionnaire that was used for the case study is found in Appendix C. The answers to the questionnaire were given by using a scale from 1 to 10. The participant was seated at the assembly workstation during the whole experiment. Test completion times were measured by using a stopwatch. Moreover, the numbers of errors (i.e. picking errors and assembly errors) made to complete the task were collected.

7 Results

In this chapter, the results of the experimental tests conducted in the laboratory are presented. Repeated-measures analyses were performed by considering the type of technology used to provide the instructions as the independent variable (i.e. paper-based manuals, workplace-mounted monitors, and Kinect-projector system). The dependent variables were the Task Completion Time (TCT), the number of errors, and the simplified NASA-TLX questionnaire. Each participant tested the three types of technology on two different assembled products: a LEGO Front wing and a LEGO Side pod. These two assemblies are characterized by different degrees of complexity: the LEGO Front wing has a lower degree of complexity (3 different parts, 4 work steps), while the LEGO Side pod has a higher degree of complexity (8 different parts, 8 work steps). Therefore, the complexity of the assembly is also considered in the results.

7.1 Task Completion Time

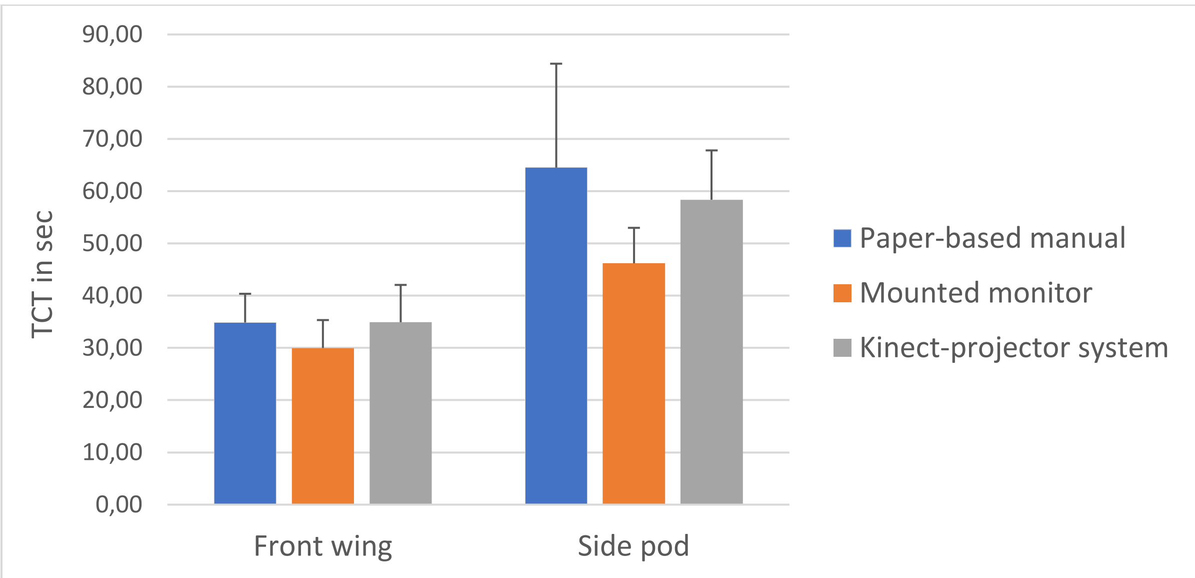


Figure 7.1: Task Completion Time

For both Front wing and Side pod, the instructions on the mounted monitor gave the fastest results (*M* = 29.98s, *SD* = 5.35s and *M* = 46.19s, *SD* = 6.79s, respectively), followed by the Kinect-projector system (*M* = 34.92s, *SD* = 7.14s and *M* = 58.36s, *SD* = 9.45s, respectively). Using the paper-based manual, it took the longest time to assemble (*M* = 34.96s, *SD* = 5.50s and *M* = 64.52s, *SD* = 19.89s, respectively). Further, for both the Front wing and the Side pod assemblies, the t-test shows that there was a significant difference between the TCTs by using Kinect-projector system and mounted monitor (p = 0.025 for Front wing, p = 0.001 for Side pod). Instead, the difference between Kinect-projector system and paper-based manual was not significant (p = 0.963 for Front wing, p = 0.125 for Side pod). An overview of the TCTs according to the different systems is depicted in Figure 7.1.

7.2 Number of Errors

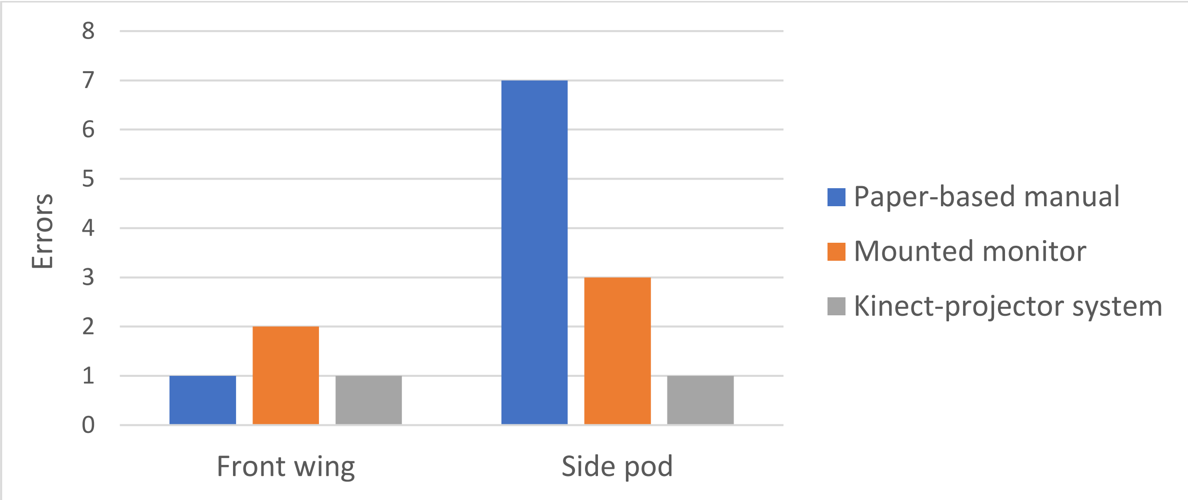


Figure 7.2: Number of errors

During the assembly tests, the number of errors was counted. Picking errors and assembly errors were considered. A few number of errors were detected on average for the assemblies. This is probably due to the relative simplicity of the considered tasks. However, the observed errors were quite distributed in the test populations. In fact, 10 out of 15 participants made at least 1 error. Therefore, the total number of errors is considered in these results. Regarding the lower complexity assembly (i.e. the Front wing), 1 error was detected during the tests with paper-based manuals, 2 errors with the mounted monitor, and 1 error with the Kinect-projector system. Instead, regarding the higher complexity assembly (i.e. the Side pod), 7 errors were observed during the tests with paper-based manuals, 3 errors with the mounted monitor, and 1 error with the Kinect-projector system. While the number of errors for the Front wing was very similar between the three methods, the difference in the number of errors performed for the Side pod between the three methods is evident. For instance, only 1 error was performed with the Kinect-projector system, against 7 and 3 errors performed with paper-based manuals and the mounted monitor, respectively. Thus, these results show a benefit in using a Kinect-projector system when the product complexity is higher. An overview of the number of errors according to the different systems is depicted in Figure 7.2.

7.3 Simplified NASA-TLX

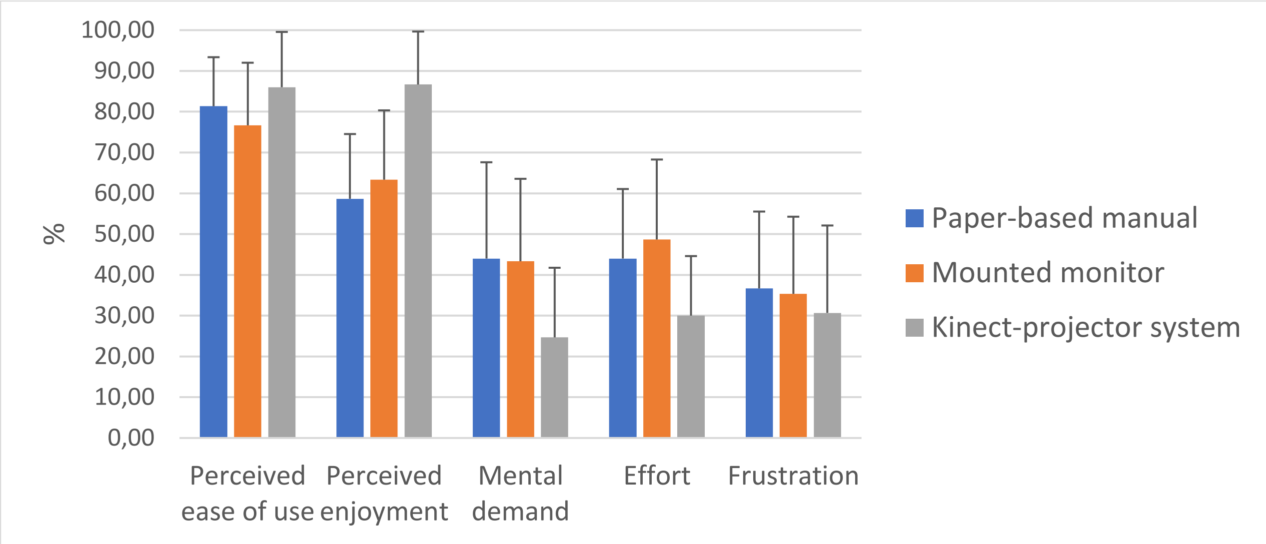


Figure 7.3: Simplified NASA-TLX

After performing all the assembly tests, each participant filled out a simplified NASA-TLX questionnaire, which can be found in Appendix C. They were asked to give feedback for the three different systems (i.e. paper-based manuals, mounted monitors, and Kinect-projector system) on the following five dimensions: perceived ease of use, perceived enjoyment, mental demand, effort, and frustration. A detailed presentation of the results is in the following sections. An overview of the results of the simplified NASA-TLX according to the different systems is depicted in Figure 7.3.

Perceived ease of use

A better perceived ease of use of the Kinect-projector system was measured (86%), compared to 81.3% of the paper-based manuals and 76.7% of the mounted monitors. Although the Kinect-projector system was perceived easier in use on average compared to the other two systems, the t-test shows no significant differences between Kinect-projector and paper-based manuals (p = 0.301), and between Kinect-projector and mounted monitor (p = 0.110).

Perceived enjoyment

Feedback regarding the perceived enjoyment gave the best score to the Kinect-projector system (86.7%). Instead, the other two systems scored lower results: 58.7% for paper-based manuals and 63.3% for the mounted monitor. An increase in perceived enjoyment registered with the Kinect-projector system is evident. In fact, the t-test shows significant differences between Kinect-projector and the other two traditional systems.

Specifically, between Kinect-projector and paper-based manuals ($p < 0.001$), and between Kinect-projector and mounted monitor ($p < 0.001$).

Mental workload
Similar values of mental workload were registered for the two traditional systems (44% for paper-based manuals and 43.3% for mounted monitor). Instead, the mental workload was lower by using the Kinect-projector system with a score of 24.7%. The benefit regarding mental workload of using the Kinect-projector system is evident. In fact, the t-test shows significant differences between Kinect-projector and the other two systems. Specifically, between Kinect-projector and paper-based manuals ($p = 0.006$), and between Kinect-projector and mounted monitor ($p = 0.002$).

Effort
According to the scores, the most effort (48.7%) was put for the assemblies using the mounted monitor, followed by paper-based manuals (44%). The least amount of effort, thus the best result, was registered for the tests when the Kinect-projector system was used (30%). Moreover, the t-test shows significant differences between Kinect-projector and both the other two systems. Specifically, between Kinect-projector and paper-based manuals ($p = 0.004$), and between Kinect-projector and mounted monitor ($p = 0.001$).

Frustration
Although the frustration feedback scored better for the Kinect-projector system, the values for the three different systems are quite similar. Specifically, 30.7% for the Kinect-projector system, 35.3% for the mounted monitor, and 36.7% for the paper-based manual. Further, the t-test showed no significant difference between Kinect-projector and mounted monitor ($p = 0.461$) and between Kinect-projector and paper-based manual ($p = 0.315$).

8 Discussion

In this chapter, a discussion about the results of the case study involving the newly developed assistance system for manual assembly is presented. Later in the chapter, the limitations of the assistance system and of the results from the study are introduced.

8.1 Discussion of the Results

The results of the case study suggest that a Kinect-projector assistance system has several advantages over the traditional methods for information visualization used in manual assembly. However, the results of the study also revealed some inconveniences, which showed that this prototype of assistance system is still not ready for being used in a production environment in the industry.

Performance
Considering the Task Completion Times, in fact, the workplace-mounted touch screen monitor performed significantly better than the Kinect-projector system. The tests executed with the monitor were faster by a significant margin. Thus, it remains the better solution in terms of productivity performance in production environment. However, this method requires that the workers drag their finger on the touch screen in order to advance to the next slide of instructions. But this is not always possible, since in several manufacturing environments the workers wear work gloves or safety gloves and deal with dirt while manufacturing the products. In a situation like this, thus, the Kinect-projector system could be used. In fact, it registered slightly better levels of productivity than by using paper-based manuals. A reason for the relatively high TCTs when using the Kinect-projector system comes from the way the system's steps was conceived. In fact, in each work step, the worker must perform the following actions: pick the part, confirm the picking, assemble the part, confirm the assembly. Thus, there are two confirmation steps that are time-consuming and inevitably have a negative impact on the TCT. However, the system can be improved by at least removing the picking confirmation. The picking detection could make the system automatically advance to the assembly step, without having the worker confirming the picking with the interactive button. On the other hand, the assembly confirmation step is currently needed since there is no quality assessment feature in the assistance

system developed in this thesis that can verify the correctness of the assembly.

Use Cases

The results regarding productivity bring us to another crucial matter, that is the use case of the assistance system. Although the results of the study revealed that the Kinect-projector system was not the fastest, it could be applied for different purposes. In fact, the assistance system can find use in the training of the workers (U2). Further, different types of information visualization methods could be integrated for training purposes, such as projecting demonstration videos, or by combining pictures and videos, and so on. On the other hand, the Kinect-projector system can be also a good solution for continuous support during productions (U1) when many product variants are manufactured, similar parts are used, and a lot size one approach is adopted. In these situations, the worker cannot learn and remember the work steps to assemble each product configuration, but needs a system that can assist him/her real-time through each work step.

Accuracy

No evident benefits were revealed by the study in terms of error-proofing for low levels of complexity when using the Kinect-projector system. However, the results also showed that increasing the level of complexity leads to more errors when using paper-based manuals and mounted monitors. As a matter of fact, with high levels of complexity, the Kinect-projector system becomes beneficial, reducing the number of errors significantly (RQ3). This was the result from a still relatively simple assembly (i.e. the LEGO Side pod: 8 different parts, 8 work steps), so the benefits could be even greater for more complex products, e.g. products with many work steps, parts that are similar to each other, work steps difficult to remember, and so on. Particularly, the Pick-by-Light feature of the assistance system prevented many picking errors. It can be especially beneficial when the parts are very similar to each other, and the worker may be confused when picking and may need to pay extra attention.

Economy

The evaluation of a setup that relies on affordable hardware appears valuable to support the adoption of an augmented reality solution in practice. Measuring performance and accuracy only partly indicates the effect of an assistance system on production costs. Indirect monetary savings could arise from lowering training costs. On the other hand, the efforts for creating and maintaining AR-based instructions compared to traditional methods such as paper-based manuals or videos have to be considered as well.

Acceptance

The Kinect-projector system got good results in perceived ease of use. They were quite comparable to the other more traditional methods, despite being a digital and more advance piece of technology. In fact, it is important that such a new system will find acceptance among the assembly workers. In this regard, the results were positive, particularly in terms of perceived enjoyment, where the assistance system scored significantly higher than the traditional methods. Manual assembly tasks can become monotonous, therefore an interactive Kinect-projector assistance system resulted more pleasant and enjoyable in the eyes of the participants, compared to traditional methods. This is relevant both for production and for training. It can be used for training purposes, for instance for product ramp-ups, making the training sessions more enjoyable and effective; or as an assistance system within production. Whether the measured effect of perceived enjoyment can be observed in practice depends on the use case. When used within the daily routine, the positive effect may wear off. If used in a separate training context, the effect of novelty and enjoyment may last and conserve the benefits of an AR-system.

Cognitive Effort

The case study revealed positive results in terms of both mental workload and effort. Much lower mental workload and effort were used for performing the tasks with the Kinect-projector system. Pick-by-Light and picking error-proofing substantially contribute to reduce mental workload. Having the instructions projected right in front of the worker in a comfortable way leads to less effort used by the worker. Instead, more effort is used, for instance, to turn the pages of a paper-based manual, or to swipe on the touch screen in order to advance to the next slides, and more attention is paid to pick parts that are very similar to each other. Besides, there are still potential improvements that can be achieved by adding an assembly recognition feature to the assistance system. In this way, the system will automatically assess the correctness of the assembly, leading to an even greater reduction of the cognitive effort by the worker. Finally, comparable levels of frustration were found for the three different methods. In this regard, the Kinect-projector system did not give significantly better results than the other traditional methods. This is probably due to technical reasons. The motion recognition software was at times not responsive and failed to detect the worker’s motion, leading to a certain level of frustration. Improvements in this aspect are evidently needed.

8.2 Limitations

In this section, in view of the results of the tests in the laboratory, the limitations of the newly developed Kinect-projector assistance system and the limitations of the case study itself are discussed.

8.2.1 Limitations of the Assistance System

One of the biggest limitations of the Kinect-projector assistance system that is presented in this thesis is the inability to detect the correctness of the assembly after each assembly step. The motion recognition software that was implemented for this system is, in fact, incapable of recognizing whether a part was assembled in the right place. In order to do that, we need a different software that can also use the Kinect camera as the video source and that can compare the current assembly with images of correct assemblies that are stored in a database. A quality assessment after each assembly step is crucial to reduce the number of defects to a minimum, but also it can reduce the mental workload of the workers even more. On the contrary, it can be time-consuming to check the assembly correctness after each assembly step, thus a trade-off between productivity and error rate should be achieved.

The assistance system is able to prevent picking errors, more specifically it signals which part needs to be picked through Pick-by-Light and it warns whether a wrong part was picked. In fact, the system detects the presence of the worker's hand whenever it picks parts from the storage bins. A limit of this functionality is that the system cannot detect if the correct amount of parts was picked. In fact, in some picking step, multiple pieces of the same part need to be picked. To cope with this issue, a picking confirmation step was added in the working flow, so that the worker can confirm through the projected button that the right parts were picked in the right amount.

That brings us to another issue of this assistance system, which is the many confirmation steps that need to be performed to advance in the work process. In fact, a picking confirmation and an assembly confirmation from the worker are needed since the system cannot detect when a work step is over. These confirmation steps lead to higher task completion times, so this can be an issue in production environment. Instead, they do not seem to represent an issue if the assistance system is used for training purposes.

One issue that was already mentioned earlier in the chapter is the effectiveness of the motion recognition software. At times, it accidentally fails to recognize the action of the operator, causing a certain level of frustration to the worker. Therefore, especially for production purposes, a more robust version of motion detection should be achieved. On the other hand, this prototype represents an affordable solution and it adequately works for demonstrations, laboratory tests, and research purposes.

Another limitation of this assistance system prototype is about the product type to be assemble. The product cannot be too big and its parts must fit in the relatively small storage bins. Further, there is a limit on the number of storage bins that can fit in the workstation, therefore products that are composed by many different parts would not be suitable. In this prototype, a configuration with 2 rows of bins was identified as the maximum number of rows that could both fit in the depth image of the Kinect and be covered by the projector's light.

8.2.2 Limitations of the Results

The results from the user study present some limitations that need to be addressed. First of all, the tests were conducted in a laboratory, which is an artificial environment, inherently different from a real factory. Further, the assembly tests were performed on LEGO constructions, which is a simple form of assembly where, for instance, working tools are not needed. Therefore, the conclusions of the presented study need validation in real-world prototypes. Furthermore, using a more realistic task would allow to consider further requirements of the industrial domain, for instance more diverse parts.

A second limitation of the study concerns the way the assembly tasks were conducted. The participants that came to conduct the test in the laboratory performed the assemblies with the different methods (i.e. paper-based manuals, mounted monitor, Kinect-projector system) at short time intervals between them. Therefore, a certain degree of learning curve effect was present, especially considering the relatively simple assembly designs. One way to avoid this would be to make the testers come to the laboratory more than once, with some days between each visits, so that they can perform the assembly using a different method each visit and, therefore, would not be affected by learning curve effects. Another way to avoid this problem is by increasing the complexity of the assembly tasks, so that the learning curve effect has a lower impact on the results. However, this way conflicts in some extent with the workstation physical design, which does only allow a limited number of storage bin, thus the use of a limited number of different parts.

Finally, a limitation of the results comes from the test population. The participants, in fact, were students, researchers, or professors, therefore they all came from the academia and had no experience as factory workers. The experiment showed, for instance, a strong effect of perceived enjoyment on the use intention. A judgment of whether real assembly workers perceive this effect at all cannot be provided, since the experiment focused on initial users. Accordingly, a test made by real operators and assembly workers would be beneficial, especially for better assessing the user acceptance of such assistance system in real-world operations.

9 Conclusions and Recommendations for Further Work

This chapter concludes the report by summarizing the research contributions and answering the research questions that were addressed in this thesis. Further, recommendations on the next steps for future work are given.

9.1 Conclusions

The main contributions of the thesis can be summarized as follows.

By reviewing the literature, the requirements, architecture and concepts for a general assistance system for manual assembly were defined. These findings are also applicable to any assistance system no matter the underlying technology.

Then, a prototype was built by using affordable technologies such as a Kinect, a projector, and a simple but intuitive motion recognition software. The prototype allows for testing and further research opportunities.

By combining those technologies, an assistance system was developed with all of the following functionalities: in-situ projection visualization, Pick-by-Light, picking error-proofing, and gesture user interaction.

The case study involving the new assistance system was carried out to assess the system in comparison with the established methods for manual assembly assistance such as paper-based manuals and mounted monitors. The study measured the performance and the accuracy of the system, along with the user acceptance and cognitive effort.

The results showed that this Kinect-projector assistance system leads to higher task completion times than providing instructions with mounted monitors, but lower that with paper-based manuals. However, it leads to fewer mistakes, preventing both picking errors thanks to the Pick-by-Light and error-proofing features, and assembly errors thanks to assembly instructions presented in a comfortable way to the worker. Further, the system also seems to be well accepted, scoring very well in perceived ease of use and perceived enjoyment. Finally, the results showed that the system contributes to significantly reduce the mental workload and cognitive effort of the user.

One other aspect that was monitored in the case study is the impact of product complexity on our assistance system performance. This was measured by testing the system on two different LEGO models with different levels of complexity. The results showed that the workers benefit the most from the assistance system when the product complexity is higher, while for low complexity no significant benefits were registered.

Therefore, the results of the study indicate which situations can benefit the most by the assistance system. That is for the manufacturing of complex products, characterized by many different work steps and assembly parts, and for manufacturing of many product variants where teaching each and every configuration beforehand is not possible.

Regarding the possible use cases of such Kinect-projector assistance system, it can be deployed in the trainings of the workers without any drawbacks. However, the results of the study revealed that this system is still not ready for continuously supporting workers during production operations. Further research needs to be done in this regard.

9.2 Recommendations for Further Work

In this thesis, a smart workstation equipped with an affordable Kinect-projector assistance system was built, and later the use of the system for manual assembly tasks in a workstation scenario was investigated. As the conducted research and the presented prototype build a solid foundation for the workstation scenario, further interesting areas of research that were beyond the scope of this thesis are identified. These areas for future research are presented in this section.

First of all, the study could be extended to the investigation of assembly performance on real-world manufacturing products. In this way, the system would be assessed and validated also in this regard.

The assistance system developed in this project is lacking of an assembly quality assessment feature which is able to recognize whether the assembly step was correctly performed or not. Therefore, further research in this regard is needed. For instance, the implementation of a different software for assembly recognition could be investigate. Such a software could use the RGB image of the Kinect to assess the assembly quality by comparing it with pre-stored images of correct assembly steps in a database.

Another important direction of research is authoring tools that facilitates the creation and maintenance of AR-based instructions. Advanced authoring tools could make these tasks manageable for users without a background in information technology, for example instructors.

Moreover, further research needs to be addressed to the study of a business case about the benefits and costs of implementing this Kinect-projector assistance system. Many aspects such as cost of equipment, cost of system maintenance, cost of avoided defects, savings from workers training, need to be considered in order to evaluate the system from the economical point of view of manufacturing companies that may be interested in this kind of technology solution for their operations.

Bibliography

AGRAWALA, M., PHAN, D., HEISER, J., HAYMAKER, J., KLINGNER, J., HANRAHAN, P. & TVERSKY, B. 2003. Designing effective step-by-step assembly instructions. *ACM Transactions on Graphics,* 22**,** 828-837.

BAECHLER, A., BAECHLER, L., AUTENRIETH, S., KURTZ, P., HOERZ, T., HEIDENREICH, T. & KRUELL, G. A comparative study of an assistance system for manual order picking - called pick-by-projection - with the guiding systems pick-by-paper, pick-by-light and pick-by-display. Proceedings of the Annual Hawaii International Conference on System Sciences, 2016. 523-531.

BAECHLER, A., KURTZ, P., HOERZ, T., KRUELL, G., BAECHLER, L. & AUTENRIETH, S. About the development of an interactive assistance system for impaired employees in manual order picking. 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments, PETRA 2015 - Proceedings, 2015.

BANNAT, A., GAST, J., RIGOLL, G. & WALLHOFF, F. 2008a. Event analysis and interpretation of human activity for augmented reality-based assistant systems.

BANNAT, A., WALLHOFF, F., RIGOLL, G., FRIESDORF, F., BUBB, H., STORK, S., J MÜLLER, H., SCHUBÖ, A., WIESBECK, M. & F ZÄH, M. 2008b. *Towards Optimal Worker Assistance: A Framework for Adaptive Selection and Presentation of Assembly Instructions*.

BÜTTNER, S., SAND, O. & RÖCKER, C. Extending the design space in industrial manufacturing through mobile projection. MobileHCI 2015 - Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct, 2015. 1130-1133.

CAUDELL, T. P. & MIZELL, D. W. 1992. Augmented reality: an application of heads-up display technology to manual manufacturing processes.

DALLE MURA, M., DINI, G. & FAILLI, F. 2016. An Integrated Environment Based on Augmented Reality and Sensing Device for Manual Assembly Workstations.

E. SUTHERLAND, I. 1969. *A head-mounted three dimensional display*.

ECHTLER, F., STURM, F., KINDERMANN, K., KLINKER, G., STILLA, J., TRILK, J. & NAJAFI, H. 2004. The Intelligent Welding Gun: Augmented Reality for Experimental Vehicle Construction. *In:* ONG, S. K. & NEE, A. Y. C. (eds.) *Virtual and Augmented Reality Applications in Manufacturing.* London: Springer London.

FUNK, M., KORN, O. & SCHMIDT, A. An augmented workplace for enabling user-defined tangibles. Conference on Human Factors in Computing Systems - Proceedings, 2014. 1285-1290.

FUNK, M., LISCHKE, L., MAYER, S., SHIRAZI, A. S. & SCHMIDT, A. 2018. *Teach Me How! Interactive Assembly Instructions Using Demonstration and In-Situ Projection*.

FUNK, M., SHIRAZI, A. S., MAYER, S., LISCHKE, L. & SCHMIDT, A. Pick from Here- An interactive mobile cart using in-situ projection for order picking. UbiComp 2015 - Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2015. 601-609.

GEWOHN, M., BEYERER, J., USLÄNDER, T. & SUTSCHET, G. 2018. Smart Information Visualization for First-Time Quality within the Automobile Production Assembly Line. *IFAC PapersOnLine,* 51**,** 423-428.

GORECKY, D., CAMPOS, R., CHAKRAVARTHY, H., DABELOW, R., SCHLICK, J. & ZÜHLKE, D. 2013. *Mastering Mass Customization - A Concept for Advanced, Human- Centered Assembly*.

HART, S. G. & STAVELAND, L. E. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *In:* HANCOCK, P. A. & MESHKATI, N. (eds.) *Advances in Psychology.* North-Holland.

HARTMANN, B. 2011. Human worker activity recognition in industrial environments. KIT Scientific Publishing.

HERMANN, M., PENTEK, T. & OTTO, B. 2015. Design Principles for Industrie 4.0 Scenarios: A Literature Review. Unpublished.

ISHII, H. & ULLMER, B. 1997. Tangible bits: towards seamless interfaces between people, bits and atoms. *In:* PEMBERTON, S. (ed.) *CHI '97.* ACM.

JACOB, R., GIROUARD, A., HIRSHFIELD, L., HORN, M., SHAER, O., SOLOVEY, E. & ZIGELBAUM, J. 2008. Reality-based interaction: a framework for post-WIMP interfaces.

KORN, O., SCHMIDT, A. & HÖRZ, T. 2012. Assistive systems in production environments: exploring motion recognition and gamification. *Proceedings of the 5th International Conference on PErvasive Technologies Related to Assistive Environments.* Heraklion, Crete, Greece: ACM.

KORN, O., SCHMIDT, A. & HÖRZ, T. Augmented manufacturing: A study with impaired persons on assistive systems using in-situ projection. ACM International Conference Proceeding Series, 2013a.

KORN, O., SCHMIDT, A. & HÖRZ, T. The Potentials of In-Situ-Projection for Augmented Workplaces in Production. A Study with Impaired Persons. Conference on Human Factors in Computing Systems - Proceedings, 2013b. 979-984.

KOSSIAKOFF, A., SWEET, W. N., SEYMOUR, S. J. & BIEMER, S. M. 2011. *Systems Engineering Principles and Practice: Second Edition*.

KRAMMER, P., NEEF, D. & PLAPPER, P. 2011. Advanced Manufacturing Technologies for General Assembly.

LANCIONI, G. E., O'REILLY, M. F., SEEDHOUSE, P., FURNISS, F. & CUNHA, B. 2000. Promoting Independent Task Performance by Persons with Severe Developmental Disabilities through a New Computer-Aided System. *Behavior Modification,* 24**,** 700-718.

LI, X., CHEN, I. Y. H., THOMAS, S. & MACDONALD, B. A. Using kinect for monitoring warehouse order picking operations. Australasian Conference on Robotics and Automation, ACRA, 2012.

LOCH, F., QUINT, F. & BRISHTEL, I. Comparing video and augmented reality assistance in manual assembly. Proceedings - 12th International Conference on Intelligent Environments, IE 2016, 2016. 147-150.

LUCKE, D. 2008. *Smart Factory - A Step towards the Next Generation of Manufacturing,* London, London: Springer London.

MARNER, M. R., IRLITTI, A. & THOMAS, B. H. Improving procedural task performance with Augmented Reality annotations. 2013 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2013, 2013. 39-48.

MICHALOS, G., MAKRIS, S. & CHRYSSOLOURIS, G. 2013. The effect of job rotation during assembly on the quality of final product. *CIRP Journal of Manufacturing Science and Technology,* 6**,** 187-197.

MILGRAM, P. & KISHINO, F. 1994. Taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems,* E77-D**,** 1321-1329.

NEE, A. Y. C., ONG, S. K., CHRYSSOLOURIS, G. & MOURTZIS, D. 2012. Augmented reality applications in design and manufacturing. *CIRP Annals - Manufacturing Technology,* 61**,** 657-679.

NTNU. 2018. *Logistics 4.0 Laboratory – Logistics of the Future* [Online]. Available: https://www.ntnu.no/mtp/laboratorier/logistics-4.0-laboratory [Accessed 10/06 2019].

OMEGAUNFOLD. 2016a. *Using Microsoft Kinect with software on PC* [Online]. Available: https://www.zonetrigger.com/articles/Kinect-software/ [Accessed 10/06 2019].

OMEGAUNFOLD. 2016b. *Webcam Zone Trigger* [Online]. Available: https://www.zonetrigger.com/motion-detection/ [Accessed 10/06 2019].

PINHANEZ, C. 2001. The everywhere displays projector: A device to create ubiquitous graphical interfaces.

RASKAR, R., VAN BAAR, J., BEARDSLEY, P., WILLWACHER, T., RAO, S. & FORLINES, C. 2003. iLamps: Geometrically aware and self-configuring projectors. *ACM Trans. Graph.,* 22**,** 809-818.

RASKAR, R., WELCH, G. & FUCHS, H. 1999. Spatially augmented reality. *Proceedings of the international workshop on Augmented reality : placing artificial objects in real scenes: placing artificial objects in real scenes.* Bellevue, Washington, USA: A. K. Peters, Ltd.

ROMERO, D., BERNUS, P., NORAN, O., STAHRE, J. & FASTH BERGLUND, Å. 2016a. The Operator 4.0: Human Cyber-Physical Systems & Adaptive Automation towards Human-Automation Symbiosis Work Systems.

ROMERO, D., STAHRE, J., WUEST, T., NORAN, O., BERNUS, P., FAST-BERGLUND, Å. & GORECKY, D. Towards an operator 4.0 typology: A human-centric perspective on the fourth industrial revolution technologies. CIE 2016: 46th International Conferences on Computers and Industrial Engineering, 2016b.

RÜTHER, S., HERMANN, T., MRACEK, M., KOPP, S. & STEIL, J. An assistance system for guiding workers in central sterilization supply departments. ACM International Conference Proceeding Series, 2013.

SCHLÖMER, T., POPPINGA, B., HENZE, N. & BOLL, S. 2008. Gesture recognition with a Wii controller.

SCHWERDTFEGER, B., PUSTKA, D., HOFHAUSER, A. & KLINKER, G. 2008. Using laser projectors for augmented reality.

SCHWERDTFEGER, B., REIF, R., GUNTHNER, W. A., KLINKER, G., HAMACHER, D., SCHEGA, L., BOCKELMANN, I., DOIL, F. & TUMLER, J. 2009. Pick-by-Vision: A first stress test.

SUZUKI, G., MURASE, T. & FUJII, Y. Projecting recorded expert hands at real size, at real speed, and onto real objects for manual work. International Conference on Intelligent User Interfaces, Proceedings IUI, 2016. 13-17.

TANG, A., OWEN, C., BIOCCA, F. & MOU, W. 2003. Comparative effectiveness of augmented reality in object assembly.

VENKATESH, V. 2000. Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model. *Information Systems Research,* 11**,** 342-365.

VENKATESH, V. & BALA, H. 2008. Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decision Sciences,* 39**,** 273-315.

WARD, J. A., LUKOWICZ, P., TROSTER, G. & STARNER, T. E. 2006. Activity Recognition of Assembly Tasks Using Body-Worn Microphones and Accelerometers. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 28**,** 1553-1567.

WEING, M., SCHAUB, F., RÖHLIG, A., KÖNINGS, B., ROGERS, K., RUKZIO, E., GUGENHEIMER, J. & WEBER, M. P.I.A.N.O.: Enhancing instrument learning via interactive projected augmentation. UbiComp 2013 Adjunct - Adjunct Publication of the 2013 ACM Conference on Ubiquitous Computing, 2013. 75-78.

WELLNER, P. 1991. The DigitalDesk calculator: tangible manipulation on a desk top display. *Proceedings of the 4th annual ACM symposium on User interface software and technology.* Hilton Head, South Carolina, USA: ACM.

WILSON, A. 2004. TouchLight: an imaging touch screen and display for gesture-based interaction.

WILSON, A. 2010. Using a depth camera as a touch sensor.

ZAEH, M., WIESBECK, M., STORK, S. & SCHUBÖ, A. 2009. A multi-dimensional measure for determining the complexity of manual assembly operations. *Research and Development,* 3**,** 489-496.

ZAMFIRESCU, C.-B., PIRVU, B.-C., GORECKY, D. & CHAKRAVARTHY, H. 2014. Human-centred Assembly: A Case Study for an Anthropocentric Cyber-physical System. *Procedia Technology,* 15**,** 90-98.

ZHOU, J., LEE, I., THOMAS, B., MENASSA, R., FARRANT, A. & SANSOME, A. 2011. Applying spatial augmented reality to facilitate in-situ support for automotive spot welding inspection.

Appendices

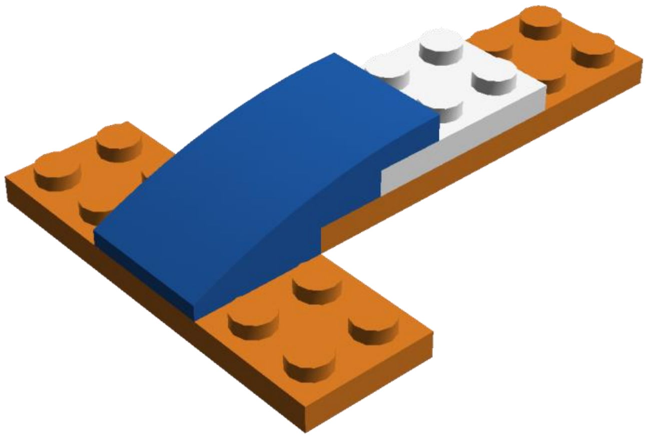
Appendix A: LEGO Front wing assembly instructions

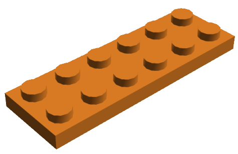
Appendix B: LEGO Side pod assembly instructions

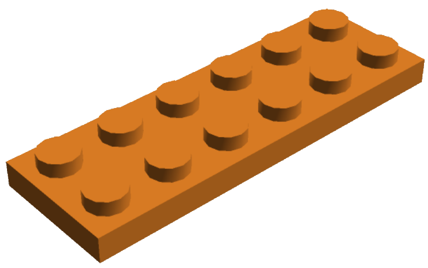
Appendix C: Questionnaire

Appendix A

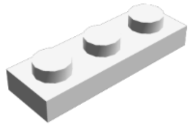
LEGO Front Wing

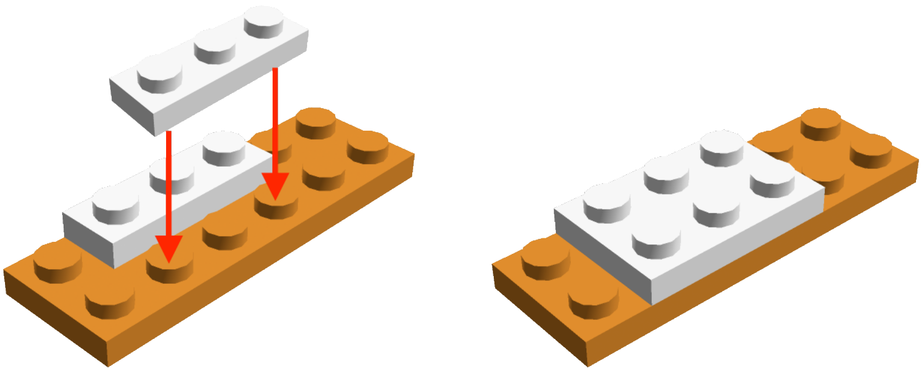


1x	

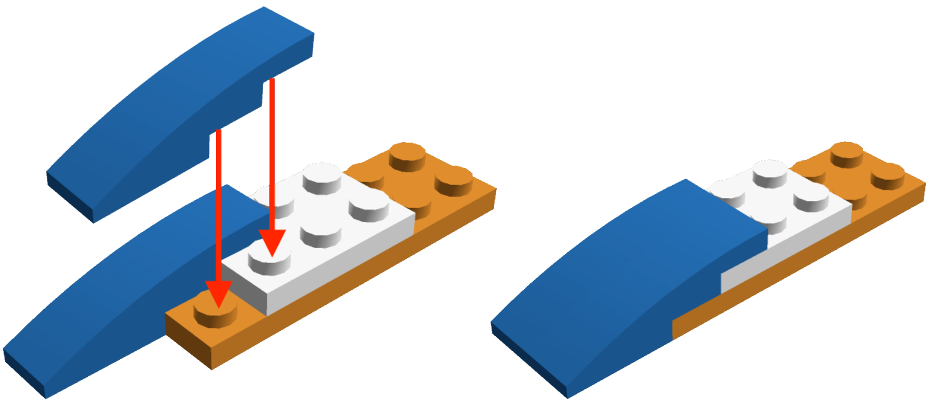


2x

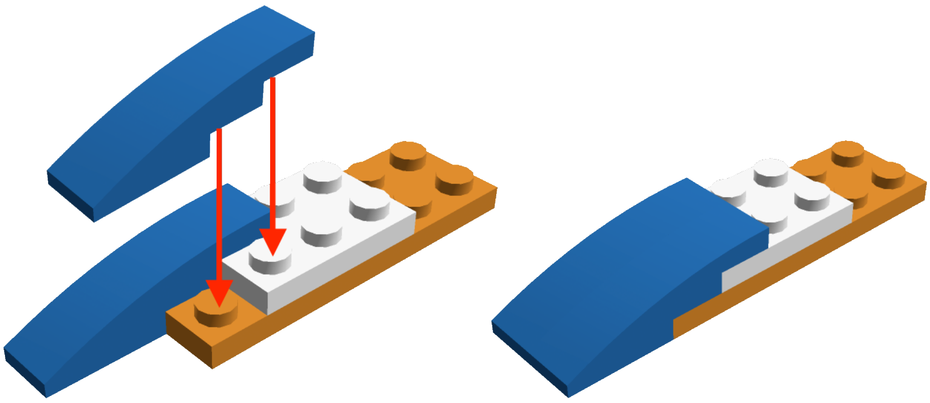


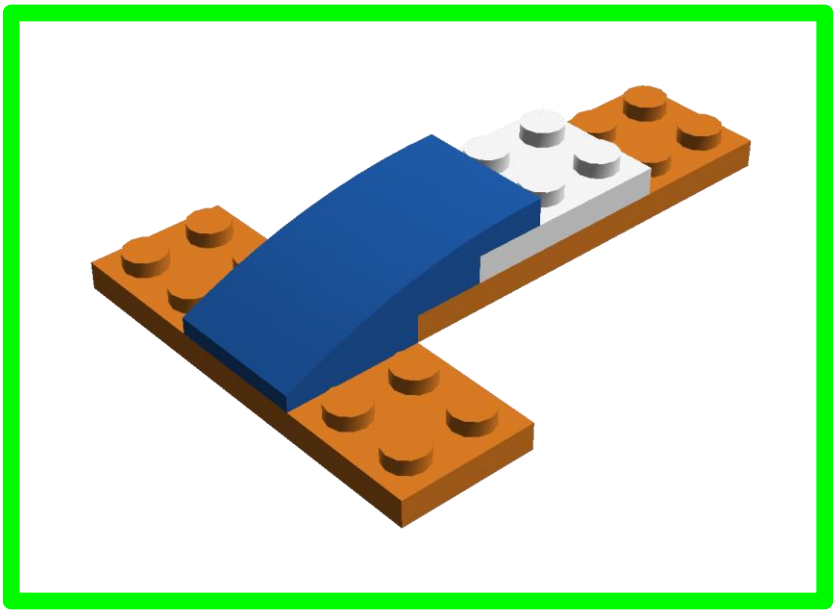


2x



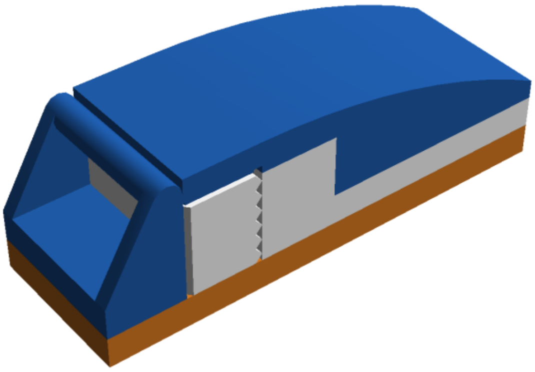
2x

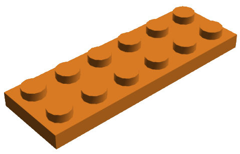


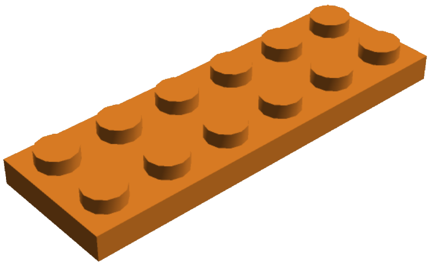


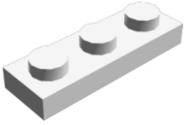
Appendix B

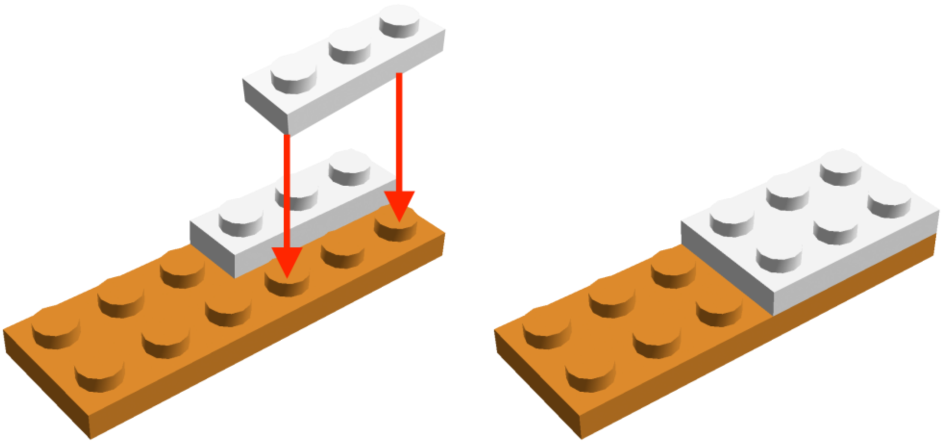
LEGO Side Pod

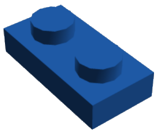


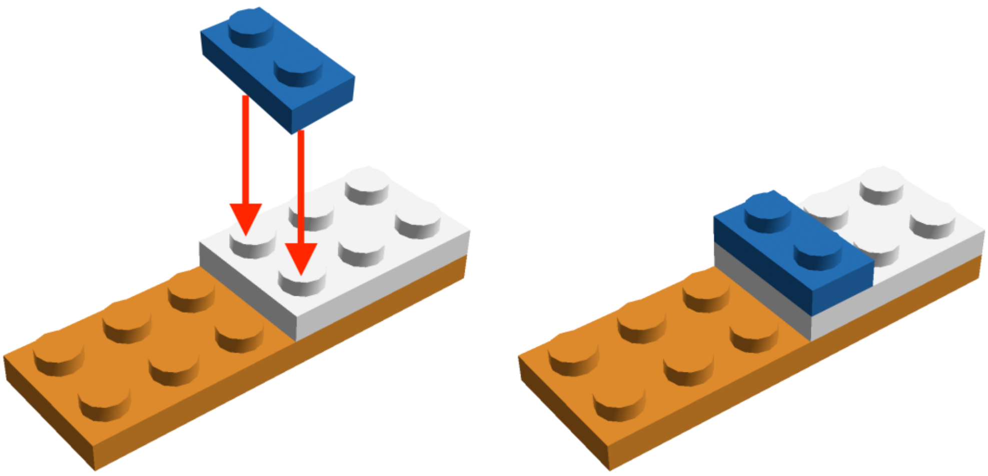
1x	

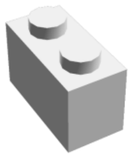


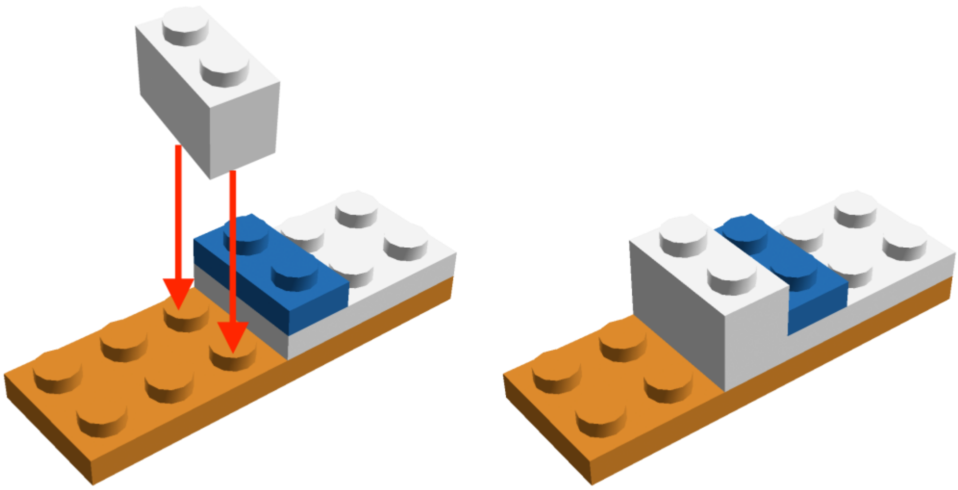
2x	

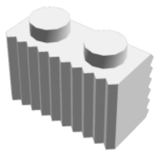


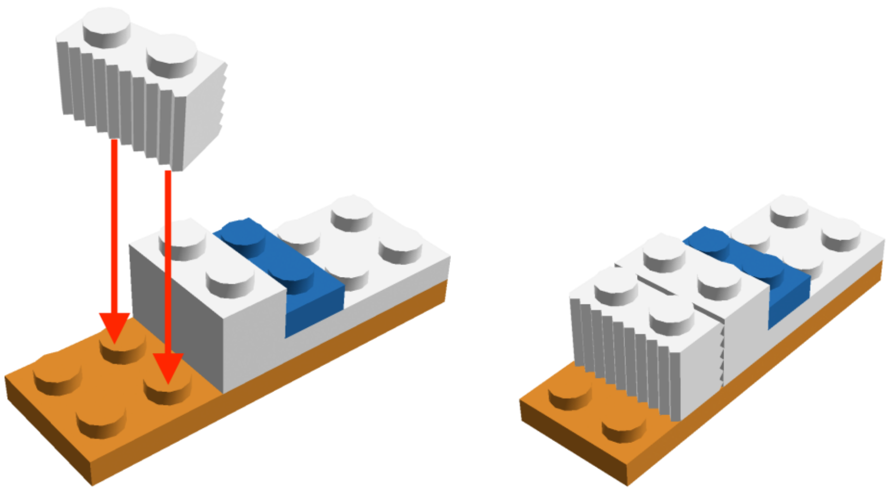
1x	



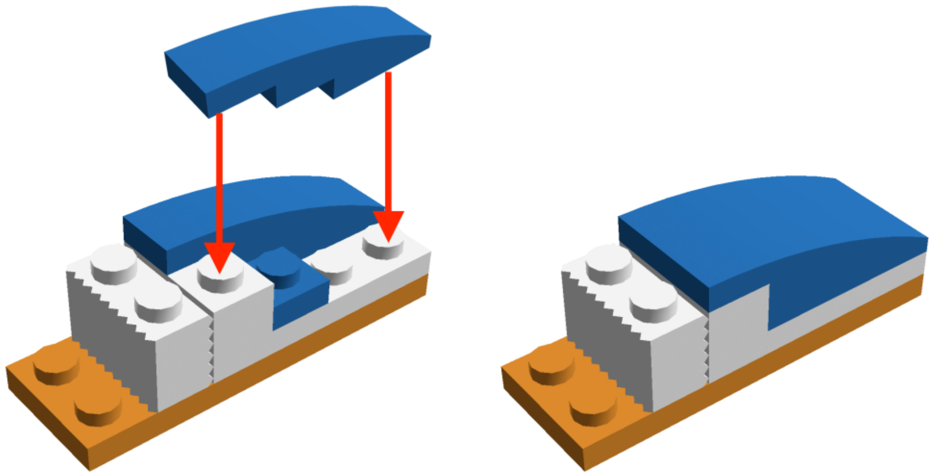
1x	

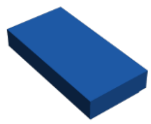


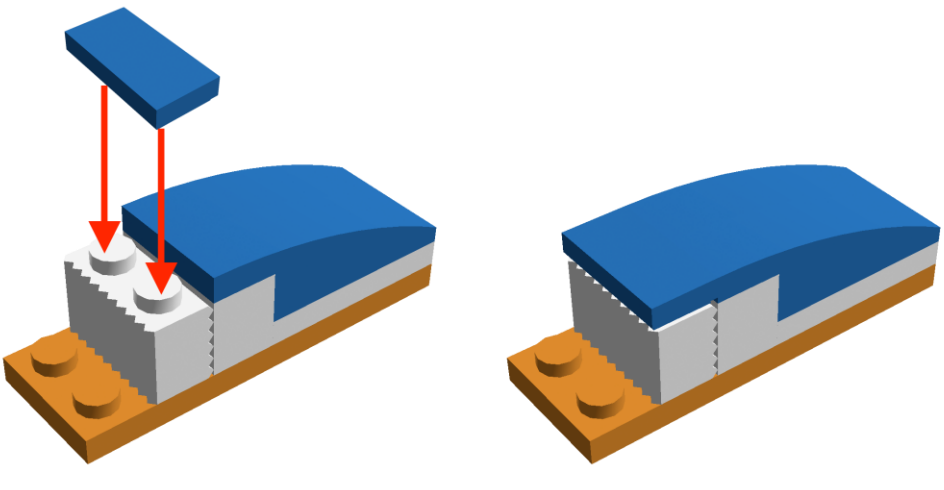
1x	



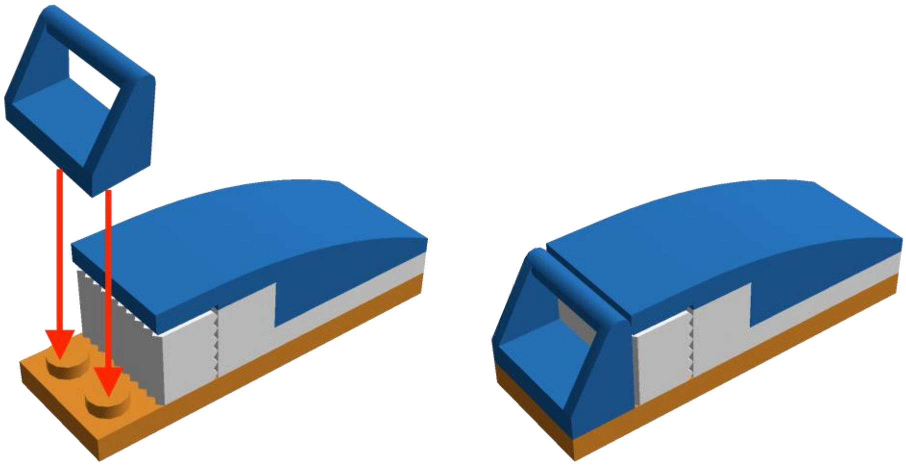
2x	

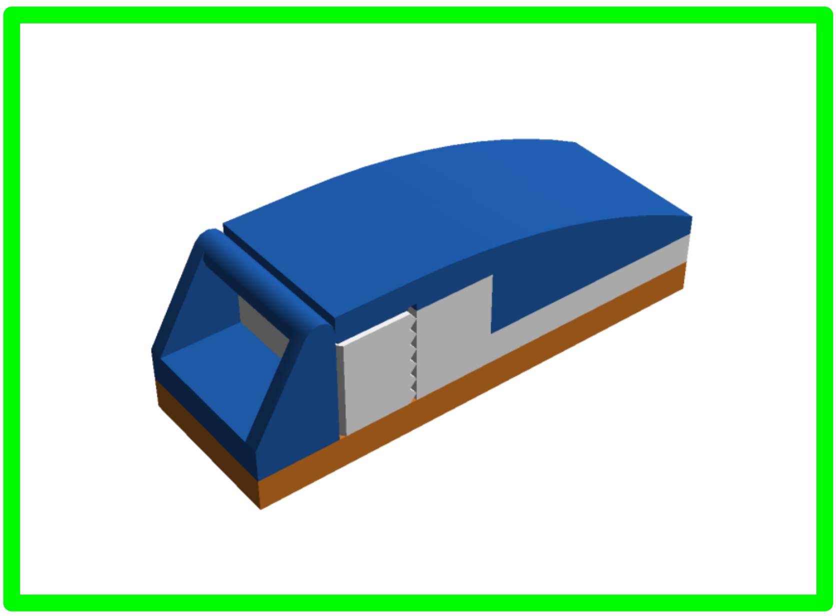


1x	



1x	





Appendix C

PAPER-BASED MANUAL	*Low*									*High*
Perceived ease of use:	1	2	3	4	5	6	7	8	9	10
Perceived enjoyment:	1	2	3	4	5	6	7	8	9	10
Mental Demand: How mentally challenging (e.g. thinking, searching, deciding) was the task?	1	2	3	4	5	6	7	8	9	10
Effort: How much energy was put forth to achieve your level of performance in the task?	1	2	3	4	5	6	7	8	9	10
Frustration: How discouraged, bothered, and annoyed were you because of the task?	1	2	3	4	5	6	7	8	9	10

TOUCH SCREEN MONITOR	*Low*									*High*
Perceived ease of use:	1	2	3	4	5	6	7	8	9	10
Perceived enjoyment:	1	2	3	4	5	6	7	8	9	10
Mental Demand: How mentally challenging (e.g. thinking, searching, deciding) was the task?	1	2	3	4	5	6	7	8	9	10
Effort: How much energy was put forth to achieve your level of performance in the task?	1	2	3	4	5	6	7	8	9	10
Frustration: How discouraged, bothered, and annoyed were you because of the task?	1	2	3	4	5	6	7	8	9	10

KINECT-PROJECTOR SYSTEM	*Low*									*High*
Perceived ease of use:	1	2	3	4	5	6	7	8	9	10
Perceived enjoyment:	1	2	3	4	5	6	7	8	9	10
Mental Demand: How mentally challenging (e.g. thinking, searching, deciding) was the task?	1	2	3	4	5	6	7	8	9	10
Effort: How much energy was put forth to achieve your level of performance in the task?	1	2	3	4	5	6	7	8	9	10
Frustration: How discouraged, bothered, and annoyed were you because of the task?	1	2	3	4	5	6	7	8	9	10

NTNU

Norwegian University of
Science and Technology