Evaluation of an Augmented Reality Assisted Manufacturing System for Assembly Guidance

Max Bode

University of Twente / KTH Royal Institute of Technology Enschede, The Netherlands / Stockholm, Sweden mspbode@gmail.com

ABSTRACT

Augmented Reality (AR) is a technology that enhances a real-world environment with computer-generated information and objects. The technology is under constant development and still rising in popularity. One of the popular applications is assembly guidance. AR enables to show the assembly instructions in the field of view, in real-time so it is always relevant to the situation. The operators can do the tasks at hand, without having to change their view, position and attention, and without having to focus on the instructions to figure out the next step. This paper covers the evaluation in the form of a user study of a newly developed AR assembly guidance system with a spatial display that projects instructions in the view of work of Scania truck assembly workers. The guidance system is compared with Scania's current assembly instructions (instructions printed on paper) in terms of effectiveness and usability. The effectiveness is measured by time-to-completion, number of errors and learning curve. The usability is measured by a SUS score. The study results in a significant difference in effectiveness, where AR guidance improves upon the current instructions. The SUS results of the current instructions indicate that it is the worst imaginable, a detractor and it would not be accepted by new users. The SUS results of the AR guidance indicate that it is excellent, a promoter and it would be accepted by new users.

SAMMANFATTNING

Augmented Reality (AR) är en teknik som förbättrar en verklig miljö med datorgenererad information och objekt. Tekniken är under ständig utveckling och ökar fortfarande i popularitet. En av de populära applikationerna är monteringsguider. AR gör det möjligt att visa monteringsinstruktionerna i synfältet i realtid så att det alltid är relevant för situationen. Operatörerna kan utföra uppgifterna till hands utan att behöva ändra syn, position och uppmärksamhet och utan att fokusera på instruktionerna för att ta reda på nästa steg.Denna rapport behandlar utvärderingen i form av en användarundersökning av ett nyutvecklat AR-monteringssystem med en rumslig display som projicerar arbetsinstruktioner för Scania lastbilsmontörer. AR-vägledningen jämförs med Scanias nuvarande monteringsinstruktioner när det gäller effektivitet och användbarhet. Effektiviteten mäts genom tid-till-slutförande, antal fel och inlärningskurva. Användbarheten mäts med en SUS-poäng. Studien resulterar i en signifikant skillnad i effektivitet, där ARvägledning visar på förbättring mot de nuvarande instruktionerna. SUS-resultaten från de nuvarande instruktionerna indikerar att det är det sämsta tänkbara, en "detractor" och att de inte skulle accepteras av nya användare. SUSresultaten från AR-vägledningen indikerar att den är utmärkt, en "promotor" och att den skulle accepteras av nya användare.

CCS CONCEPTS

• Human-centered computing → Mixed / augmented reality; User studies; Usability testing; Heuristic evaluations; User centered design;

KEYWORDS

Augmented Reality Assisted Manufacturing (ARAM), Augmented Reality Guidance, Light Guidance, Assembly Manufacturing, Assembly Instructions, Human-Computer Interaction

1 INTRODUCTION

Augmented Reality (AR) is a technology that enhances a real-world environment with computer-generated information and objects. The technology is under constant development and still rising in popularity. Ever since it emerged, one of the popular applications is guidance. Especially in the medical field, AR guidance has many applications already, e.g. for biopsy and other invasive procedures.

Another promising application field is assembly manufacturing. One of the most important aspects of a successful production is correct and efficient assembly. To achieve this, there is a need for proper assembly information, such as instruction manuals and schematics. However, these instructions are usually separated from the assembly process, and the operators therefore would need to alternate their attention between their task at hand and the instructions. This causes a loss of time, loss of focus and increased mental workload. It results in reduced productivity, errors and possibly injuries. AR enables to show the assembly instructions in the field of view, in real-time so it is always relevant to the situation. The operators can do the tasks at hand, without having to change their view, position and attention, and without having to focus on the instructions to figure out the next step[27][40].

One of the players in the industry that has shown its interest in AR guidance is Scania AB. In the Scania truck end assembly plant in Södertälje, Sweden, the operators (assembly workers) work with assembly instructions printed on paper. These instructions are coming from Scania's own developed Enterprise Resource Planning (ERP) tool, and are further referred to as the current assembly instructions or current instructions (see figure 5). Developments have been done to improve the assembly guidance, including digital instructions on screens and sequenced power tools. However, the problem that the assembly workers must alternate their attention between the task at hand and the instructions has not been solved.

Aimed at solving this problem of alternating attention, the current project developed assembly instructions that can be shown in the view of work. This was done with the use of AR technology. Since a universal term for using AR guidance in manufacturing did not exist yet, the current project attempted to create one, giving birth to the concept of Augmented Reality Assisted Manufacturing (ARAM). The project focused on a specific station in the assembly line. In this station the assembly workers need to place press screws in the frame (chassis) of the truck, which are used to attach all kinds of brackets later on down the assembly line. Because of Scania's service to its customers of highly customizable trucks, and constant developments with often new part introductions, the different kind of screw configurations is endless. The learning curve is long, the mental workload is high and human errors are often made. Next to the above mentioned issues with printed instructions, the current instructions bring another issue in this case. The information provided by Scania's ERP is not sufficient to understand the full assembly process. Another tool is necessary to be able to know where to place the right screws. This tool is a visualization tool, created for the design and production of the truck chassis. It visualizes the holes and brackets that need to be mounted on the chassis, and is referred to as chassis visualization tool. It has many usability issues on its own and, most importantly, it is not directly available for the assembly workers placing the screws. Only one person on every station, the team leader, has access to the tool. The goal was to design a solution to improve this situation with the use of AR guidance, make a prototype and test it with a user study.

Although the Master graduation project entails the full project mentioned above, the current paper (mainly) covers the last part of the project, the user study. Furthermore, it contains the literature study. The goal of the user study was to test with scientific relevance if the design of the AR guidance works, how it compares to/if it improves upon the current instructions in terms of assembly performance (effectiveness) and usability, and to discover usability issues. The current paper further briefly summarizes the work approach and methods used during the full case study, such as observations, user interviews and several design iterations. It is recommended to read the full report, which can be requested by contacting the author of the current paper.

1.1 Research Question

Following the goal of the user study and literature review, the following research question was formulated:

What are the benefits and costs of supporting Scania final assembly line workers with an interactive augmented reality guidance system for the press screw placement in the truck chassis as measured by timeto-task completion, number of errors and user feedback (perceived usability) captured through system usability scale (SUS) questionnaires and semi-structured interviews?

1.2 Report Organization

The graduation project of this Master thesis was a case study for Scania. Structured as a scientific paper, the current thesis is focused on the evaluation of the prototype that resulted from the case study. The current paper only reports on the part of the project that is scientifically relevant.

Section 1 introduces the topic and the current situation and goal of the full case study at Scania, and the goal of the current paper. Section 1.1 states the research question that the current paper aims to answer. Section 2 reports on the current scientific work out there about AR guidance in assembly manufacturing. Section 3 shortly summarizes the full case study to give some context, especially the work structure and methods used. It is advised to read the full report. Section 4 covers another theoretical background on currently used methods to evaluate an AR guidance system. Section 5 explains the method used in the current study to answer the research question. Section 6 reports on the results of the user study and analyses them to understand the results and be able to answer the question. Section 7 gives a discussion on the outcome of the study and future work. Lastly, section 8 answers the research question by giving the conclusions of the analysis and summarizing the study.

2 THEORETICAL BACKGROUND: AR GUIDANCE IN ASSEMBLY MANUFACTURING

AR guidance has been an interest of manufacturers for decades. Often the production process is too complex for automation because of variations in the procedures and the environment. Humans are far more capable than robots if it comes to dealing with these variations. However, humans can make mistakes and have a learning curve. The use of AR guidance can enhance the learning process and reduce the completion times and number of redundant process actions[28][27][21][25][21][23][4]. It can reduce the number of errors made[33][27][2][12][2]. Furthermore, AR guidance can be more effective than instructions on paper or a screen and decrease the mental work load[38][35][41][2][34][22]. Hence, a lot of research has been done on different kinds of applications of AR technologies in manufacturing to guide assembly workers.

One of the oldest AR guidance technology in manufacturing is a Heads Up Display (HUD). Caudell and Mizzel[7] created a Heads Up Display headset (HUDset) for assembly workers in the Boeing 747 factory. Similar to the Scania assembly line, the assembly workers must deal with countless parts and different kinds of tasks that are unfitted for automation. They created a proof of concept with a head-mounted display and a fixed platform, with four simple example applications. The set was calibrated with the platform and the orientation was tracked by measuring the motion in six degrees of freedom. They encountered many difficulties related to weaknesses of the HUDset (comfort, resolution) and a delay between the movement of the head and the image transition. Nevertheless, they have high expectations for future HUDset applications in manufacturing. They claim that the position sensing technology (tracking) is the ultimate limitation, as that is controlling the range and accuracy of possible applications.

2.1 Tracking

The most popular way of position sensing is marker-based augmentation [40]. By processing a live video stream captured by a camera a certain marker (figure or pattern) is detected and recognized. The virtual object is then projected on this marker. Sääski et al.[35] used this marker-based augmentation technique in a case study on the assembly of a Valtra Plc tractor accessory's power unit to test the effectiveness of AR guidance through a Head Mounted Display (HMD). Their results showed a lower task completion time with the AR instructions compared to a situation with instructions on paper. However, the HMD was uncomfortable and a part of the instructions needed to be shown on a monitor still because they were unreadable through the tiny display. They suggested a normal computer monitor to be a good alternative for the display. They found that the presentation and animation of how the part should be mounted in the right location and orientation were the most valuable features in AR instructions.

Many AR applications and tools such as the ARToolKit¹ can make use of a marker-based augmentation with several markers to detect the 3D orientation of objects out of a 2D image or video stream. These markers have drawbacks however, such as that they cannot always be placed on small and/or round objects like screws[43], sensitivity to different light conditions and inter-marker confusion (because the computer vision works with a probability threshold)[19][40]. An alternative is markerless-based augmentation. Several techniques are developed for this which include camera position orientation tracking by comparing the scene with predefined (prerecorded) features[13], rigid body tracking with a 3D model and texture matching[14], 2D to 3D feature translation combined with an inertial measurement unit [24] and a hybrid version that makes use of image feature detection, sample comparison and inertial measurement[26]. However, these techniques have not been tested in real-life industrial applications.

A promising markerless-based augmentation technique for manufacturing applications is developed by Platonov et al.[31]. They project 2D images on a 3D CAD model. This way they can compare the 2D features with the 3D features and connect them to get the 2D-3D correspondence. With an additional tracking system that does use a marker they create several key frames which include an image combined with the location relative to the marker. When this one-time learning process is done the markerless initialization and calibration is done by comparing these key frames with the 2D-3D correspondences from the live video stream. A downside of this method is that the problem with lighting comes back since initial key frames (recordings) are used. Another downside is that the position of the live stream camera must be near the position of the initial camera for the system to recognize that particular key frame. They introduced a training procedure for the initial key frame learning to improve the reliability of the key frames. They were able to track the right orientation of a real car engine with the use of a CAD model. The biggest usability flaw was a slow re-initialization when the user looked away for a moment.

A robust technique to build a 3D map within an unprepared environment is Simultaneous Localization and Mapping (SLAM). However, normal SLAM techniques cannot cope with a change in the environment (eg. moving the camera). Dong et al.[10] developed a SLAM technique that can also deal moving scenes, making relocalization possible. They do that by removing invalid 3D points, points that do not match the scene anymore. Still, their technique has some limitations. The major limitation is that it failed to track fast moving objects inside the frame. Nevertheless, this technique has huge potential for future AR applications.

Next to vision-based tracking approaches, there is sensor-based tracking, of which the above mentioned inertial measurement unit is an example. It is the oldest way of tracking and with the arrival of computer vision the least used. However it has been used the most and is considered to be a reliable tracking approach[40]. Sensorbased tracking can be done with (a combination of) different kinds of sensors including acoustic, inertial, optical magnetic and mechanical sensors. The key is that they need to have high accuracy and low latency [27]. Henderson and Feiner[17] used a set of 11 infrared sensors, mounted around the work area, to track three reflectors fixed on a HMD. With tracking software they were able to track the HMD as a 3D rigid body. They used AR to guide the assembly process of a Rolls-Royce Dart combustion engine. Vignais et al.[39] used IMU's combined with mechanical goniometers to track the posture of the user's body. Through a HMD the user got to see his or her own posture, a real-time ergonomic assessment, while performing manual tasks in a manufacturing environment.

2.2 Context-Aware Applications and Authoring

AR guidance has been developed for several purposes in assembly manufacturing, including real-time interactive instructions, assembly visualization/simulation, sequence planning, training, collaborative assembly and sequence- and product evaluation [27][40]. Zauner et al. [43] developed a Mixed Reality Assembly Instructor, a step-by-step guidance for furniture assembly. Context related actions are given (reacting on the current situation) to the user through a tablet with virtual objects and audio feedback that uses a camera and markers for tracking. The instruction are intuitive and proactive. For example, a misplaced part is virtually rotated and/or moved until it fits. A major drawback of this system is that the markers cannot always be placed on all parts. Parts like small screws and round surfaces cause problems. They also created Authoring Wizard, which is a simple graphical tool that allows people to create instructions for an assembly sequence without any programming skills. However, they found that the tool was still too complicated for an average person without any knowledge of programming and therefore suggest to use clearer buttons in future work. Nevertheless, their concept of having a software tool that allows authors to create AR instructions without programming skills was an important development for AR assembly guidance. As in manufacturing industries processes for assembly need to be flexible, it is important that content for AR assembly guidance can be created and re-authored fast and easily[3].

Another authoring method first generated instructions in Microsoft PowerPoint and then translated those into 3D AR content with PowerSpace Editor [16]. A more recent project at AIRBUS used an industrial Digital Mock-Up (iDMU) to create content for AR assembly guidance. The iDMU combined 3D models, metadata and other digital structured documents created by different design departments with different kinds of software including CATIA, Optegra, 3DVIA Composer, PRIMES and DELMIA[36]. There are even methods developed recently that automatically create instructions with the use of computer vision. Researchers created real-time AR guidance instructions by analyzing a video demonstration (video examples from the workflow) and training an image recognition model to recognize process steps, keep track of the progress and add instructions automatically at the right place and moment[30], even using the moving image of a head mounted camera[29]. Bhattacharya and Winer[3] used a similar computer vision approach

¹"ARToolKit". Philip Lamb. Accessed on May 2, 2019. http://www.hitl.washington.edu/artoolkit/

(but with a 3D camera generating RGB point cloud information) and combined it with a 3D model library with the different assembly components to recognize the process and generate real-time instructions.

2.3 Interactivity

By keeping track of the process, the above mentioned projects were able to not only give 3D context-aware instructions, also make them real-time and interactive. In assembly operations, the user is in need of correctly timed guidance. Poorly timed step-based instructions can stall the user or interfere with the production process. By making the 3D instructions interactive and hands-free, an intuitive and unobtrusive way of guidance can be created[40]. Andersen et al. keep track of the process of the assembly of a water pump with image processing. With feature extraction they compare the video stream with a CAD model of the pump[1]. However, the camera needed to be fixed, the pump could not move and the system was sensitive for occlusion. The latter happens often as the user often moves its hand over the pump, blocking the camera's view. Although these issues have been partly solved in later projects[8][37], there are other options as well. Keeping track of the process does not have to be fully automatic, it can also be done by user input. Charoenseang and Panjan[9] make use of an 5-finger exoskeleton glove with force feedback to give the user the ability to move virtual objects during assembly training. Yuan, Ong and Nee[42] developed a Virtual Interaction Panel that can be used in the field of sight so the user could control the AR system. Other options are voice control, motion tracking or sensors placed on/around the item that needs to be assembled[40].

2.4 Information Transfer (Displays)

Next to the tablet, there are other alternatives for the HUD or HMD. Another handheld display (HHD) application is using a smartphone. Hakkarainen et al.[15] created a mobile phone based AR assembly guidance system. They connected a Nokia N95 as client over Bluetooth to a server on a PC running a ARToolKit based image processing software. The mobile device made snapshots and send it to the server where the processing took place. The server send back an image with AR instructions. They found that users enjoyed using the system. The bottleneck was the connection speed, which took a few seconds for each image to send. Modern smartphones are far more powerful than the old N95 and the wireless connections have much greater speed and bandwidth, creating a lot more opportunities for a seamless live server-client based AR assembly guidance.

Another option is a spatial display, where the user is completely detached from the display, like a projector[40]. Aligned Vision² developed a system with a laser that projects cutting lines, placement borders and other assembly guidance. This is especially suitable for production with raw materials.

The Optimum Smart Klaus³ is a working station for small assembly. A camera with computer vision monitors and reports errors. The instructions are given on a screen.

Light Guide Systems⁴ developed a system that works with computer vision and a projector for inspection and guidance. Their Light Guide Pro system works with the HP Sprout Pro portable platform.

3 A CASE STUDY ON ASSEMBLY GUIDANCE FOR THE PRESS SCREW STATION

The prototype used in the user study of this project resulted from a case study for Scania AB, about the above described press screw station. The format of the case study was the Design Thinking format. The goal of the project was to improve the situation by implementing some sort of operator/assembly guidance, make a prototype and test it. The current paper covers the latter. Although not relevant for the current paper, below, the case study is briefly summarized to give the reader some background information on how the AR guidance system was developed. A full report can be requested by contacting the author of the current paper.

3.1 First Iteration: The Problem and a Concept

The case study started off with observations, data analysis (gathered data and existing data), semi-structured user interviews and expert interviews to discover the (underlying) problems. The process of mounting a screw and the flow of information between all people directly involved were mapped. The business need was captured and a concept for a solution was designed and presented. The concept included an interactive Augmented Reality Assisted Manufacturing (ARAM) assembly guidance system, to help the assembly workers improve their assembly performance. A Computer Vision based Quality Control would ensure that the screws are mounted correctly. Furthermore, the concept included a data processing application that would feed the system with the correct information for assembly. The remainder of the case study was focused on the design of the ARAM assembly guidance.

3.2 Second Iteration: Design of the ARAM Assembly Guidance

To design the proposed ARAM guidance, a Human Centered Design approach was used, which focuses on the needs of the end users. Following this approach, all stakeholders were identified of which the stakeholders directly involved in the press screw station were interviewed (structured and semi-structured) in a second iteration interview to get feedback on the proposed solution and an even better understanding of the assembly process, underlying issues, thoughts and feels of the people and requirements of the new system. A first prototype was used in the feedback sessions to illustrate the idea and interaction. More brainstorms were held with employees from different departments, a literature study was done (covered by the current paper), requirements were stated and a final design was made. The solution is an interactive AR Light

 $^{^2}Aligned$ Vision: Portfolio Accessed on 24 September 2019 https://aligned-vision.com/portfolio/

³Real time assembly guidance with vision validation and documentation. Accessed on 24 September 2019 https://www.liveaoi.com/

⁴Leading the transformation to Industrial Augmented Reality. Light Guide Pro. Accessed on 24 September 2019 https://lightguidesys.com/



Figure 1: Concept design of the ARAM assembly guidance. Three projectors linked together are hanging from the ceiling. They project instructions on the grey chassis of the truck. The white construction is mimicked from the real press screw station. The blue boxes contain the screws. A touchscreen with dashboard is placed next to the chassis for extra information and to control the system.



Figure 2: Concept design of the visualizations of the system. The grey frame is the chassis of the truck viewed from the side. The white construction is holding the blue boxes with screws. Symbols show where to place which screws on the chassis, by placing the symbols on the boxes and the corresponding holes. The dashboard view on the right shows extra information including the instructions region (indicated on the frame with two vertical yellow lines). The instructions region is the region on the frame that is highlighted by the system. Only in this region instructions are being projected. This is to ensure there is no information overload and to be able to make step-by-step instructions.



Figure 3: Final prototype (screen with dashboard not in picture). This is the right side of the chassis. The projector is hanging from a tripod, which is the black pole in the middle of the picture. The system projects on and above the boxes how many screws need to be picked and assigns a symbol to each screw. Those symbols are projected on the corresponding holes in the chassis.

Guidance system that projects the relevant assembly instructions on the chassis and boxes, while reacting on the progress of the assembly worker. The assembly progress is tracked by automatic user tracking and user input in the form of voice control or the user pressing buttons. There is a touchscreen dashboard that shows extra information about the chassis and can be used to control the system (see figure 1 and 2). A high level architecture was created and the new process was mapped in an activity diagram.

3.3 Third Iteration: The Final Prototype

The final prototype used in the user tests was created. It consisted of one projector that was hung up on a tripod, standing next to a test chassis. It was connected to a computer running the software (written in Javascript). A large touch screen was connected that showed the dashboard. A construction with a large white screen was placed on the chassis, at the same height as in the press screw station, that held ten boxes with screws (see figure 3 and 4). Third iteration interviews were held in the form of an unstructured feedback session with the end users (assembly workers of the press screw station). Next to above literature study, additional literature was reviewed for the design of the visualizations [32][11]. The prototype was evaluated with a user study. The current paper covers this user study.

4 THEORETICAL BACKGROUND: EVALUATING AR SYSTEMS FOR ASSEMBLY GUIDANCE

There is not a clear agreed upon framework to assess AR systems in assembly processes. There are no benchmarks for current or new assembly procedures and all the proven benefits achieved with AR are case specific. The evaluation studies performed so far can be divided into two categories; effectiveness- and usability evaluation [40]. The current study also evaluated on those two topics.

Conventional effectiveness evaluation studies in AR assembly guidance mainly focus on task completion time, assembly errors and work load. Quantitative and qualitative tests are used and the tests are performed in a controlled environment. The tests are comparative with a control group using the (case specific) conventional assembly instructing method and another group using the AR method. If there are several new methods tested, these are compared with a test group for each method [23][21][12][25][2][38][41][4][34]. Hou et al. [21] also compared the performance learning curve by measuring the number of trials until the assembly was completed without an error. Qualitative measures can be for example the type of assembly error that is being made, self-reported performance and self-reported work load. Some experiments also include a qualitative assessment of the level of technical assembly skills of the participants prior to the experiment. The participants perform the experiment usually once (with sometimes several trials in one experiment) and Gavish et al.[12] found that there can be a ceiling effect when their users had two sessions with their AR system.

Usability evaluation in AR assembly guidance is performed with user tests, user interviews and expert evaluations [40]. From the user tests, results are gathered from observations and subjective questionnaires, which commonly are filled in by the user after the experiment and make use of Likert scales. The focus lies on functionality and cognitive load [11][20][22]. The current study was designed based on above mentioned works.

The number of participants ranges from 6 - 75 (usually around 20) and can be dependent on the number of case specific (real) assembly workers that are involved in the assembly process if they are needed in the experiments. Other participants were (medium aged) office workers, researchers and students. The number of participants in one test/control group can be as small as 5. [21][12][23][25][2] [38][41][4][43][18][22]. The number of participants of the current study was based on above mentioned works and the resources and availability of suitable participants inside Scania.

5 METHOD

5.1 User Study Design

To be able to answer the research question, and put the measured benefits and costs of using an interactive AR guidance system in perspective, it was imperative to compare it with the current instructions system (referred to as current instructions). Therefore, the final prototype developed in the case study is used in a comparative between-subject user study to compare the new AR guidance instructions system (referred to as ARAM) with the current assembly instructions. The comparative user study was quantitative and focused on assembly performance (effectiveness) and usability. Furthermore, a qualitative usability test was done with only the new system to find usability issues.

5.2 Participants

The participants for the comparative study were Scania employees who did not have any experience with the press screw station. This way the experience of the participants could not be of influence on the assembly performance. Some of them knew about the station. The test was considered to be easy to execute, and Scania normally accepts all people with any skill set to become assembly worker, so



Figure 4: Setup of the user tests with the final prototype in the Smart Factory Lab.

the level of technical skills was not relevant. The same holds for gender, age and other demographics.

The participants only tested one system, not both. The participants testing the ARAM guidance did not get any explanation about the system and did never see the new system or concept image/drawing. The participants testing the current instructions did not have any experience with the chassis visualization tool and the current instructions for the press screw station. Some of them did know or heard of one or both tools.

The participants for the qualitative usability test were assembly workers that are currently working at the press screw station (station 8), and thus have experience with the current assembly process. All tests were anonymous, and the participants were asked to sign a consent form for their data.

5.3 Setup

Both tests were held in the Smart Factory Lab at Scania MS with a test chassis and the above-mentioned prototype. A table was placed next to the chassis that contained all the objects needed for the tests. The researcher was standing behind the participant during the test. The screen with the dashboard was operational (during tests with the new system), however, not part of the tests. There was a total of ten different kinds of screws, each in a separate box, labelled with the corresponding article number (see figure 4 and 3).

5.4 Comparative Study

5.4.1 Assembly Performance. A quantitative test was done to compare the ARAM instructions with the current instructions on

assembly performance (effectiveness). Concurrent with above mentioned literature, this study focused on assembly speed, accuracy and learning curve as dependent variables. This was measured by, respectively, the completion time, number of errors and number of trials that were necessary to make the test without errors. The independent variable was the kind of instructions system (Scania's ERP or ARAM).

Since the current instructions only list which screws need to be placed for which brackets, the participant testing the current instructions also received a printout of a visualization of the chassis visualization tool for each bracket for which they needed to place the screws. The combination of the current instructions and the chassis visualization tool gives enough information to know where to place which screw. The current instructions and chassis visualizations were specially made for the user test and looked exactly like the real tools that are daily used. The time used for placing the screws in the real station usually should not exceed four minutes (of the total tact time of seven minutes). Therefore, the maximum time period of the test was also four minutes. Although the test did not cover the whole chassis (only three meters), these four minutes were still considered to be the right amount as it would make the test a bit easier for the participants (since they did not have any experience).

The participants received safety gloves and glasses and were told how to mount a screw. In the real station, the screws are tightened with a power tool. However, since the screws and chassis needed to be re-used with every new test trial, the screws were considered to be mounted when they were placed in the frame without any tightening (with the tip pointing outwards). The participants were asked to mount all the screws according to the instructions, without any explanation about the instructions. They were told that they could skip screws if they got stuck. They could ask questions and the researcher only answered them when it would not influence the test.

The participants testing the current instructions received printouts of the chassis visualization tool (seven in total) on a stack without any particular order and a printout of the current instructions (see figure 5), after which the time would start. They were told to say 'stop' when they thought to be finished with placing all the screws and the time would stop.

Since tracking of the user's assembly progress was not yet implemented in the prototype, a Wizard of Oz technique was used for the interaction with the participants testing the ARAM system. This way interaction through voice control was mimicked. Voice control was the preferred type of interaction, as pressing a button is less obvious and could be missed by the researcher controlling the system. The user was told to say 'next' when he or she was finished with the current instructions region to go to the next region. The instructions region is the region on the frame that is highlighted by the system. Only in this region instructions are being projected (see figure 2). The researcher would press a button on a keyboard to initiate the 'next' command. There was a total of three regions. The user did not know beforehand how many regions there were. The time started when the first region appeared. When the user said 'next' after being finished with the third region the time would stop. When the time reached the four minutes the participant was not allowed to continue, and the test was finished.





Figure 5: Participants testing the current instructions received them as a printout on paper (top) and printouts of the chassis visualization tool (bottom). The prints were on A4 paper. These are fabricated for the user tests and not real examples.

After the test the time and the amount of errors were noted. The screws were numbered so the researcher could check whether the right screws were placed (see figure 6). An error was counted when the user placed a screw at the wrong place, placed the wrong screw at the right place or did not place a screw at a place where there should be a screw. There was a total of 17 screws that needed to be placed, so if for example the user would not place a screw at all, the number of errors would have been 17 errors. Every kind of screw had extra screws in the box. The test included six different kinds of screws.

If the participants made errors, they were asked to do the test again. This was to test the learning curve. Because of time constrains, it was decided to do a maximum of three trials for each participant. When there were no errors made in the first or second trial, the participants were asked to do the test again as fast as they could. This was to see whether the assembly performance changed when there was more pressure. These results are only considered as extra observations and are not mentioned in the results of the user study.



Figure 6: The screws used in the user tests were numbered so the researcher could check whether the right screws were placed.

5.4.2 Usability. To compare both systems on usability all participants were asked to fill out a System Usability Scale[6] (SUS) questionnaire after the user test. The SUS questionnaire is a widely used method to get a quick measurement of the perceived usability of a system[5]. The questionnaire consisted of five positively and five negatively asked questions to balance out biased opinions or people that do not pay attention. The questions were the standard SUS questions, adjusted so they were relevant for this system. The questions were presented on a laptop, using Survey Monkey, and were stored anonymously. See Appendix for the full questionnaire.

5.5 Qualitative Usability

To test the ARAM system for usability issues, a qualitative user test was conducted with the ARAM system and assembly workers of the press screw station. The same AR instructions as described in the comparative study were used. This time, a Think Aloud technique was used. Since the participants have had a demonstration and feedback session before, they already were introduced with the system and its functionality. The participants were asked to mount the screws by following the AR instructions, while saying out loud exactly what was going on in their minds. The researcher wrote down all their comments. After the test, they too were asked to fill out the SUS questionnaire.

6 RESULTS AND ANALYSIS

In this section the results of the comparative study and the qualitative usability test are reported and directly analysed to be able to answer the research question.

6.1 Comparative Study

The comparative study was spread over a total of four days. A total of 24 people participated in the study, 16 of them tested the ARAM system, eight of them tested the current instructions. The 16 participants testing the ARAM system had a total of 46 trials. The eight participants testing the current instructions had a total of 23 trials. It was not intentional to get less participants for the current instructions. However, since the participants could not have



Figure 7: Percentage of test trials that were finished within 4:00 minutes. 1/23 Current Instructions trials were finished in time. 46/46 ARAM Guidance trials were finished in time.

experience with the system, a greater group of participants could not be found in the available time. Each test took 20 – 30 minutes.

6.1.1 Assembly Performance: Completion Time, Duration. One out of 23 of the current instructions trials, 4,3%, finished in time (within four minutes). All trials of the ARAM guidance, 100%, finished in time (see figure 7).

Since the maximum time was almost always reached with the trials for the current instructions (except for one trial), a real average completion time cannot be calculated.

The average completion time for the ARAM guidance trials was two minutes and three seconds. The average completion time for the first trial was two minutes and 45 seconds. For the second trial two minutes and three seconds. For the third trial one minute and 17 seconds (see table 1).

	Trial 1	Trial 2	Trial 3	Overall
Current Instructions	-	-	-	-
ARAM Guidance	2:45	2:03	1:17	2:04

Table 1: Average completion times (duration) of the tests.

Since there is no average completion time for the current instructions, the difference cannot be compared and tested for significance. To still be able to quickly compare the two systems in a graph, the duration of the trials with the current instructions are put in as the maximum amount of time of four minutes. Note that this is not a real representation of the learning curve of the current instructions. The average completion times are shown in seconds (see figure 8).

6.1.2 Assembly Performance: Number of Errors. All trials with the current instructions consisted errors, with an average of 11,4 errors per trial. When 17 errors are considered as the maximum



Figure 8: Average completion duration per trial. The average duration of the current instructions could not be calculated, and is put to 4:00 for illustration. The error bars show the standard error of the mean.

amount of errors that can be made (17 screws needed to be placed), this results in an overall accuracy of 32,9%. Accuracy is considered in this report as being the assembly performance of the user measured by the portion of correctly placed press screws.

The first trial with the current instructions had an average of 15,6 errors, an accuracy of 8,1%. The second trial had an average of 12,3 errors, an accuracy of 27,6%. The third trial had an average of 5,7 errors, an accuracy of 66,5

The ARAM trials had an average of 0,28 errors per trial, which gives an overall accuracy of 98,4%. The first trial had an average of 0,69 errors, an accuracy of 95,9%. The second trial had an average of 0,13 errors, an accuracy of 99,2%. In the third trial nobody made an error, an accuracy of 100% (see figure 9).

The kind of errors that occurred during the tests with the current instructions were:

- Wrong screw in right hole, misread instructions
- Wrong screw in right hole, mixed up screw numbers on box
- Right screw in wrong hole, misread instructions
- Hole left empty, out of time
- Hole left empty, did not understand instructions
- · Hole left empty, did not understand co-mounting

During the test with the ARAM guidance only one kind of error occurred:

• Hole left empty, did not understand co-mounting

To see if the accuracy of ARAM guidance is significantly higher, a one-tailed t-test was done with the following hypothesis:

H0: People using the ARAM guidance do not make less errors than



Figure 9: Average accuracy per trial. The error bars show the standard error of the mean.

people using the current instructions.

H1: People using the ARAM guidance make less errors than people using the current instructions.

As we just want to see to what extend the difference is significant, the threshold is not set.

The standard deviations of both groups are found with the following formula:

$$s = \sqrt{\sum((xi^{\check{}}\mu)2/(N^{\check{}}1))}$$

First, for each trial, the amount of errors made was subtracted from the average. These numbers were squared and summed. This was divided by the number of trials minus one. The square root of this gave the standard deviations:

$$s(ARAM)=0,46$$

$$s(current) = 5,25$$

The variance between the two (standard error) is:

$$sd = \sqrt{(((s1)2/N1) + ((s2)2/N2))}$$

$$= \sqrt{(((5.25^2)/23) + ((0.46^2)/46))} = 1,097$$

This gives a t-score of:

$$t = (\mu 1 \tilde{\mu} 2)/sd = (11.4 0.28)/1.097 = 10, 14$$

The degrees of freedom:

$$(N1 + N2) - 2 = (23 + 46) - 2 = 67$$

In the t-table⁵ we see that for 60 degrees of freedom (or higher) the t-score needs to be 4,460 or higher for a probability of 0,0005 and lower (the probability of H0). 10,14 > 4,46, so with a confidence of

⁵ T-Distribution Table (One Tail) Accessed on 23 September 2019 https://www. statisticshowto.datasciencecentral.com/tables/t-distribution-table/

more than 99,99% we can reject H0 and say the assembly performance is higher with the ARAM guidance.

Having seen there is a significant difference between all trials, it could be relevant to see if that is also the case between only all the first trials. To see if the accuracy of ARAM guidance is significantly higher during the first trials, a one-tailed t-test was done with the following hypothesis:

H0: People using the ARAM guidance do not make less errors than people using the current instructions.

H1: People using the ARAM guidance make less errors than people using the current instructions.

The standard deviations are:

$$(ARAM) = 0,48$$

S

$$s(current) = 2,07$$

The variance between the two is:

$$sd = \sqrt{(((2.07^2)/8) + ((0.48^2)/16))} = 0,742$$

This gives a t-score:

$$t = (15.6^{\circ}0.69)/0.742 = 20, 1$$

The degrees of freedom:

$$(8+16) - 2 = 22$$

In the t-table we see that for 22 degrees of freedom the t-score needs to be 3,792 or higher for a probability of 0,0005. 20,1 > 3,792, so with a confidence of more than 99,99% we can reject H0 and say the assembly performance is higher with the ARAM guidance.

6.1.3 Assembly Performance: Learning Curve. On average, for the current instructions the improvement in accuracy in total was 58,3% (ranging from 8,1% to 66,4%). For ARAM the improvement in accuracy in total was 4,1% (ranging from 95,9% to 100%).

The learning curve of the current instructions (see figure 10) shows that there was a greater learning effect after trial 3. The participants kept on improving and learned faster with each trial.

The learning curve of the ARAM guidance cannot be analysed the same way, as there were no errors made during the third trial already. Because of the high accuracy of the ARAM trials, the learning curve of ARAM is almost none existent compared to the learning curve of the current instructions (see figure 10).

6.1.4 SUS Questionnaire. The SUS questionnaire gives a score between 0 – 100. A score above 68 is considered to be above average (average of all systems tested with SUS), below 68 is below average. However, this is not a percentage. By normalizing the score, a usability percentile rank can be calculated, and it can be compared to other systems⁶.

The average score of the current instructions is 15,94 (see figure 11). This score indicates that it is the worst imaginable. It scores better than 0 - 1.9% of all systems. It would not be accepted by new users and users would actively discourage others to use it (see figure 12).



Figure 10: Learning curves in accuracy. Blue: current instructions, orange: ARAM guidance.



Figure 11: SUS scores. Left: current instructions, right: ARAM guidance

The average score of the ARAM guidance is 83,75 (see figure 11). This score indicates that it is excellent. It scores better than 90 - 95% of all systems. It would be accepted by new users and users would actively promote it to others (see figure 12).

The participants could optionally leave an extra comment. These comments were compared on sentiment. Three random comments of both systems are stated below. In the full report all comments are reported.

Three comments about the current instructions were: "The maps were awful."

"There can be solid(better) reference to find the right holes in the 'chassis visualization tool' which will improve the usability of this system drastically."

⁶5 Ways to Interpret a SUS Score. Jeff Sauro. Accessed on 23 September 2019 https: //measuringu.com/interpret-sus-score/



Figure 12: SUS scores percentile ranks, usability compared to other systems.

"The names of the components should be in the same language for the sake of consistency."

Three comments about the ARAM guidance were:

"I think that this system is applicable to many of our processes. The lighting conditions would be interesting to discuss further. I like it!" "It's an easy way to learn how to assemble, by using this system, it's good if you new on this position."

"Instructions and information is a bit spread out so it can be missed easily. Would be really great to have an option to deduct what parts I have already mounted so I can see what I have left."

All comments were analysed with a Sentiment Analysis tool⁷. Sentiment Analysis is based on a deep learning algorithm that identifies and extracts subjective information from text while taking the context into account.

The comments of AR guidance have the following score; positive: 53,5%, neutral: 22%, negative: 24,5%. Since most part is rated as positive, the sentiment can be considered as positive.

The comments of the current instructions have the following score; positive: 7,1%, neutral: 14,2%, negative: 78,7%. The sentiment can be considered as negative (see figure 13).

6.2 Qualitative Usability

The qualitative usability was spread over a total of two sessions on two different days (they could come when the production line stopped). A total of four assembly workers participated.

The foremost result is about co-mounting. One participant did not understand the color coding of the co-mounting. He thought that one screw belonged to the different color with which the comounting was highlighted. This way, he could not find the right screw for this hole. His remark was that co-mounting could just be a normal color like all others. Two other participants were confused about the coding of the co-mounting and therefore could not find the right screw. One of them remarked that co-mounting just needs to be a symbol.





Another issue that came up was the similarity of the symbols. Two participants were distracted because the symbols looked like each other. Both suggested to make use of less similar symbols, and to use other colors to make a clearer distinction.

One participant suggested to have training or introduction sessions where users can see different options and try out different kind of visualizations.

6.2.1 SUS Questionnaire. Although the SUS questionnaire is a quantitative measure, these results below should be considered as qualitative as it is the opinion of only a few members of the team of station 8. The average score of the AR guidance, rated by the team members of the press screw station, is 90,83. This score indicates that it is best imaginable. It scores better than 96 – 100% of all systems. It would be accepted by new users and users would actively promote it to others.

As the focus here was on qualitative results, the individual scores for the questions are given in table 2, to see how the system scores on which points. The best possible score is 5. Reliability scores an average score of 3,33, indicating that the users do not always fully trust that the instructions are correct. Or that they do a good job by solely following the instructions.

The part about the usability threshold scored the lowest, with 2,67. This could mean that the system needs a more layered approach for usage. The first step that users undertake to use the system needs to be easier to understand.

An extra comment from a team member was: "*Easy to learn and it will change and help our minds a lot*". Indicating that the system could lower the mental workload of the users.

6.3 Technical Test

The prototype was tested on a 'real' chassis at the press screw station. When the line was stopped (after 16:00), the prototype

⁷Sentiment Analysis. ParallelDots. Accessed on 26 September 2019 https://www.paralleldots.com/sentiment-analysis

Q	Торіс	Average score
1	Likability/ acceptance	4
2	Complexity	4
3	Usability	3,67
4	Intuitive design	3,67
5	Function integration/ completeness	3,67
6	Consistency	3,67
7	Learning curve	4
8	Functionality	3,67
9	Reliability	3,33
10	Usability threshold (layered approach)	2,67

Table 2: Individual scores of the SUS questions.

was installed and tested. The necessary information sources were gathered in the same way as was done for the test chassis in the Smart Factory Lab.

All necessary information could be loaded, and the system projected the holes accurately on the chassis. Also, it became clear that the system would fit nicely in the station. However, the lighting conditions were too bright for this prototype. The lights needed to be dimmed for the system to work properly.

7 DISCUSSION AND FUTURE WORK

Although a HUD or HMD often gives a more richer immersive AR experience, a projector was used as display. This was done because of the requirements that the users must be able to move freely in a large environment and also must be able to bend over the frame. However, if the technology matures and the HMDs become smaller, lighter and more portable, this could also be a good option. The assembly workers already have to wear safety glasses, and also indicated that they are willing to wear a HMD. Following the work and vision of Reinhart et al.[33], a HMD could be a good alternative for the projector. Since the projectors necessary for the installation need to be very powerful, their costs would be very high. Thus, worth it to look into comparing the two kinds of displays in a later stage.

Next to AR there is also Mixed Reality (MR). Where AR places virtual objects on the real world (in terms of visuals), MR really blends in the virtual objects with the real objects. The virtual and physical objects co-exist and interact with each other in real-time. These virtual objects are always fixed according to the physical world and do not move when the user turns its head. Therefore, one could argue that the AR guidance described in the current paper is actually MR, as the virtual objects (visualizations) co-exist and interact in real-time with the physical objects (frame). However, MR is also often considered to be fully immersive, and the virtual objects blend in seamlessly with the physical objects in such a way one could not tell the difference anymore. Since the AR guidance is not an immersive experience and the visualizations are relatively simple (figures and text) and clearly distinguishable (they must be in order for the system to work), the system is considered to be AR.

The SUS scores gave a clear difference between the current instructions and ARAM guidance. However, a standard deviation could not be calculated because of the way the data was gathered from Survey Monkey (only the total points per question was known, so only the average score could be calculated). Therefore, the significance could not be calculated. Still, it is such a big gap in the percentile rank graph (figure 12) one could guess there is a significant difference. Also, the overall sentiment in the extra comments of the current instructions are negative, where the overall sentiment in the comments of the ARAM guidance are positive. Although this gives some insight in the overall sentiment, it would be better to know exactly which comments are positive and negative. The content of these comments could expose more usability issues and lead to further improvements. In this study, the comments are analysed all together. Future work could analyse every comment separately. It should be noted however, that most sentiment analysis algorithms have trouble analysing short sentences.

Almost none of the participants of the current instructions tests completed the test within the maximum amount of time of four minutes. Therefore, it was not clear how much time the participants would still need to complete and an average could not be calculated. Although there seems to be a clear difference in time-to-completion, its significance could not be calculated. New experiments where the users have all the time they need can have better results. This was considered in this project, however it was decided to have a maximum amount of time because of time and resources constraints.

Almost all participants had three trials. This was to be able to measure the learning curve and gather more data. These trials were always the same. Therefore, the learning curve could be influenced by the ability of the participant to remember where the screws needed to be placed in the second and third trial. Having this said, it should also be considered that remembering the screw configurations is also part of the learning curve of the assembly workers in real life and was beneficial for both systems. These results are therefore still considered to be representative. To rule out this variable, future tests could be held with a different screw configuration for each trial.

The learning curve of the current instructions is steeper compared to the learning curve of ARAM and seems to keep on improving with every trial. However, both learning curves are difficult to compare since during the ARAM trials the most amount of errors made by one person in the first trial was one error. Often, the second trial was already without errors. With such a high accuracy at the start, the learning curve can never be steep. Obviously, it becomes less steep after trial 2 since trial 3 had an accuracy of 100%. Therefore, it seems futile to compare both curves. Perhaps tests with more participants, more different trials and over a longer period of time can result in more comparable learning curves.

Although it is an important part of AR guidance, also covered in the literature review, time and resources did not allow to implement tracking in the prototype. This is the next step in the further development of the system. From literature and experience we know that the best option would be the use of computer vision and possibly additional markers that are placed on the chassis.

The user tests done in this project only covered the effectiveness and usability of one prototype. More research and user studies must be done to be able to design the best user interface and visualizations. The qualitative usability tests were fruitful, but not extensive. Although the group of end users is small, the number of participants of the qualitative usability tests could not be representative. More tests with more end users are encouraged.

The current project can contribute to a better production process which lowers costs and usage of materials, which could eventually improve the position of Scania in the industry. It can improve the working conditions of assembly workers by lowering the complexity and mental workload. It is important, however, to also consider possible negative consequences. Employees could get too reliant on the system or become redundant because of automation. Investments and actions taken for this project could be harmful for the company and its employees. It is the responsibility of Scania and its innovators (including the author of this paper) to make decisions that are beneficial for the company and its employees.

8 CONCLUSIONS

This study investigated the research question: What are the benefits and costs of supporting Scania final assembly line workers with an interactive augmented reality guidance system for the press screw placement in the truck chassis as measured by time-to-task completion, number of errors and user feedback (perceived usability) captured through system usability scale (SUS) questionnaires and semi-structured interviews?

To be able to put the measured benefits and costs in perspective, they were compared with the current system (current assembly instructions coming from Scania's ERP tool). The final prototype developed in the case study was used in a comparative user study to compare the newly developed Augmented Reality Assisted Manufacturing (ARAM) guidance system with the current instructions. Furthermore, a qualitative usability test was done with only the new system to find usability issues.

From the comparative study it can be concluded that the ARAM guidance scores better on assembly performance than the current instructions. There is a clear difference in completion time, as all ARAM trials were finished in time against only one trial of the current instructions. With a very high confidence there can be said that the accuracy of ARAM guidance is significantly higher than that of the current instructions. Because of the high accuracy of the ARAM trials, the learning curve of ARAM is almost none existent compared to the learning curve of the current instructions. However, the curve of the ARAM guidance starts at a very high accuracy and flattens out to an accuracy of 100%. Therefore, it is difficult to draw conclusions from this comparison.

The results of the SUS questionnaire indicate that there is a huge difference in usability. The current instructions score very low in the percentile rank. They would not be accepted by new users and people would actively discourage others from using them. The ARAM guidance scores very high in the percentile rank. It would very much be accepted by new users and people would actively promote it to others. The overall sentiment of the extra comments about the current instructions is negative, indicating that people gave negative comments about current. The overall sentiment of the comments about the ARAM guidance is positive, indicating people gave positive comments about the ARAM guidance.

Furthermore, the average completion time during the first ARAM trial of 2:44 indicates that a novice assembly worker could be able to perform all tasks within the official tact time without any prior training.

An important usability issue found in the qualitative usability test is the implementation of co-mounting with its corresponding visualization. Users did not see the extra information on the big board and did not understand the meaning of the visualization. All errors made during the ARAM guidance tests were caused by co-mounting. It is best to visualize a hole meant for co-mounting in the same way as a normal hole.

Other usability issues were found, such as a high similarity between symbols and a high usability threshold. These issues are documented and can be improved in the next iteration.

The current design would fit and function properly in the press screw station in assembly. One issue that needs to be solved is the brightness of the lighting. With the use of a more powerful projector this issue can possibly be solved.

ACKNOWLEDGEMENTS

I would like to show my gratitude to Mario Romero for guiding me through this thesis project. I would like to thank Klara Jönsson, Emil Ljungberg, John Ekström and Jim Tolman from Scania for giving me the opportunity and helping me during the project. I would like to thank Dennis Reidsma, my supervisor from the University of Twente. Lastly, I would like to thank my examinor Tino Weinkauf and my opponent Daiman Khan.

REFERENCES

- Martin Andersen, R Andersen, C Larsen, Thomas B Moeslund, and Ole Madsen. Interactive assembly guide using augmented reality. In *International Symposium* on Visual Computing, pages 999–1008. Springer, 2009.
- [2] Kevin Michael Baird. Evaluating the effectiveness of augmented reality and wearable computing for a manufacturing assembly task. PhD thesis, Virginia Tech, 1999.
- [3] Bhaskar Bhattacharya and Eliot Winer. A method for real-time generation of augmented reality work instructions via expert movements. In *The Engineering Reality of Virtual Reality 2015*, volume 9392, page 93920G. International Society for Optics and Photonics, 2015.
- [4] Andrew C Boud, David J Haniff, Chris Baber, and SJ Steiner. Virtual reality and augmented reality as a training tool for assembly tasks. In 1999 IEEE International Conference on Information Visualization (Cat. No. PR00210), pages 32–36. IEEE, 1999.
- [5] John Brooke. Sus: A retrospective. J. Usability Studies, 8(2):29-40, February 2013.
- [6] John Brooke et al. Sus-a quick and dirty usability scale. Usability evaluation in industry, 189(194):4–7, 1996.
- [7] Thomas P Caudell and David W Mizell. Augmented reality: An application of heads-up display technology to manual manufacturing processes. In *Proceedings* of the twenty-fifth Hawaii international conference on system sciences, volume 2, pages 659–669. IEEE, 1992.
- [8] Himanshu Chandel and Sonia Vatta. Occlusion detection and handling: a review. International Journal of Computer Applications, 120(10), 2015.
- [9] Siam Charoenseang and Sarut Panjan. 5-finger exoskeleton for assembly training in augmented reality. In *International Conference on Virtual and Mixed Reality*, pages 30–39. Springer, 2011.
- [10] Z. Dong, G. Zhang, and H. Bao. Robust monocular slam in dynamic environments. In 2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pages 209–218, Oct 2013.
- [11] Michele Gattullo, Antonio E Uva, Michele Fiorentino, and Joseph L Gabbard. Legibility in industrial ar: text style, color coding, and illuminance. *IEEE computer graphics and applications*, 35(2):52–61, 2015.
- [12] Nirit Gavish, Teresa Gutiérrez, Sabine Webel, Jorge Rodríguez, Matteo Peveri, Uli Bockholt, and Franco Tecchia. Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks. *Interactive Learning Environments*, 23(6):778–798, 2015.
- [13] Y. Genc, S. Riedel, F. Souvannavong, C. Akinlar, and N. Navab. Marker-less tracking for ar: a learning-based approach. In *Proceedings. International Symposium* on Mixed and Augmented Reality, pages 295–304, Oct 2002.
- [14] Philippe Gérard and André Gagalowicz. Three dimensional model-based tracking using texture learning and matching. *Pattern Recognition Letters*, 21(13-14):1095– 1103, 2000.
- [15] Mika Hakkarainen, Charles Woodward, and Mark Billinghurst. Augmented assembly using a mobile phone. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 167–168. IEEE Computer Society, 2008.
- [16] Matthias Haringer and Holger T Regenbrecht. A pragmatic approach to augmented reality authoring. In Proceedings. International Symposium on Mixed and Augmented Reality, pages 237–245. IEEE, 2002.
- [17] S. J. Henderson and S. K. Feiner. Augmented reality in the psychomotor phase of a procedural task. In 2011 10th IEEE International Symposium on Mixed and Augmented Reality, pages 191–200, Oct 2011.
- [18] Steven J Henderson and Steven Feiner. Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. In 2009 8th IEEE International Symposium on Mixed and Augmented Reality, pages 135–144. IEEE, 2009.
- [19] Martin Hirzer. Marker detection for augmented reality applications. In Seminar/Project Image Analysis Graz, pages 1–2, 2008.
- [20] Lei Hou and Xiangyu Wang. Experimental framework for evaluating cognitive workload of using ar system for general assembly task. In Proceedings of the 28th International Symposium on Automation and Robotics in Construction, 2011.
- [21] Lei Hou and Xiangyu Wang. A study on the benefits of augmented reality in retaining working memory in assembly tasks: A focus on differences in gender. *Automation in Construction*, 32:38–45, 2013.
- [22] Lei Hou, Xiangyu Wang, Leonhard Bernold, and Peter ED Love. Using animated augmented reality to cognitively guide assembly. *Journal of Computing in Civil Engineering*, 27(5):439–451, 2013.
- [23] Lei Hou, Xiangyu Wang, and Martijn Truijens. Using augmented reality to facilitate piping assembly: an experiment-based evaluation. *Journal of Computing* in Civil Engineering, 29(1):05014007, 2013.
- [24] Reinhard Koch, Kevin Koeser, Birger Streckel, and Jan-Friso Evers-Senne. Markerless image-based 3d tracking for real-time augmented reality applications. In The 7th Int. Workshop on Image analysis for multimedia interactive services, Montreux, Switzerland, 2005.
- [25] Barbara Odenthal, Marcel Ph Mayer, Wolfgang Kabuß, and Christopher M Schlick. A comparative study of head-mounted and table-mounted augmented vision

systems for assembly error detection. Human factors and ergonomics in manufacturing & service industries, 24(1):105-123, 2014.

- [26] T. Okuma, T. Kurata, and K. Sakaue. A natural feature-based 3d object tracking method for wearable augmented reality. In *The 8th IEEE International Workshop* on Advanced Motion Control, 2004. AMC '04., pages 451–456, March 2004.
- [27] SK Ong, ML Yuan, and AYC Nee. Augmented reality applications in manufacturing: a survey. International journal of production research, 46(10):2707–2742, 2008.
- [28] N. Pathomaree and S. Charoenseang. Augmented reality for skill transfer in assembly task. In ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005., pages 500–504, Aug 2005.
- [29] Nils Petersen, Alain Pagani, and Didier Stricker. Real-time modeling and tracking manual workflows from first-person vision. In 2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pages 117–124. IEEE, 2013.
- [30] Nils Petersen and Didier Stricker. Learning task structure from video examples for workflow tracking and authoring. In 2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pages 237–246. IEEE, 2012.
- [31] Juri Platonov, Hauke Heibel, Peter Meier, and Bert Grollmann. A mobile markerless ar system for maintenance and repair. In 2006 IEEE/ACM International Symposium on Mixed and Augmented Reality, pages 105–108. IEEE, 2006.
- [32] Rafael Radkowski, Jordan Herrema, and James Oliver. Augmented reality-based manual assembly support with visual features for different degrees of difficulty. *International Journal of Human-Computer Interaction*, 31(5):337–349, 2015.
- [33] G Reinhart and C Patron. Integrating augmented reality in the assembly domainfundamentals, benefits and applications. CIRP Annals, 52(1):5-8, 2003.
- [34] Horacio Rios, Mauricio Hincapié, Andrea Caponio, Emilio Mercado, and Eduardo González Mendívil. Augmented reality: an advantageous option for complex training and maintenance operations in aeronautic related processes. In *International Conference on Virtual and Mixed Reality*, pages 87–96. Springer, 2011.
- [35] Juha Sääski, Tapio Salonen, Marja Liinasuo, Jarkko Pakkanen, Mikko Vanhatalo, Asko Riitahuhta, et al. Augmented reality efficiency in manufacturing industry: a case study. In DS 50: Proceedings of NordDesign 2008 Conference, Tallinn, Estonia, 21.-23.08. 2008, pages 99–109, 2008.
- [36] Javier Serván, F Mas, JL Menéndez, and J Ríos. Using augmented reality in airbus a400m shop floor assembly work instructions. In Aip conference proceedings, volume 1431, pages 633-640. AIP, 2012.
- [37] Mengu Sukan, Carmine Elvezio, Ohan Oda, Steven Feiner, and Barbara Tversky. Parafrustum: Visualization techniques for guiding a user to a constrained set of viewing positions and orientations. In Proceedings of the 27th annual ACM symposium on User interface software and technology, pages 331-340. ACM, 2014.
- [38] Arthur Tang, Charles Owen, Frank Biocca, and Weimin Mou. Comparative effectiveness of augmented reality in object assembly. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '03, pages 73–80, New York, NY, USA, 2003. ACM.
- [39] Nicolas Vignais, Markus Miezal, Gabriele Bleser, Katharina Mura, Dominic Gorecky, and Frédéric Marin. Innovative system for real-time ergonomic feedback in industrial manufacturing. *Applied ergonomics*, 44(4):566–574, 2013.
- [40] Xiangyu Wang, Soh K Ong, and Andrew YC Nee. A comprehensive survey of augmented reality assembly research. Advances in Manufacturing, 4(1):1–22, 2016.
- [41] Stefan Wiedenmaier, Olaf Oehme, Ludger Schmidt, and Holger Luczak. Augmented reality (ar) for assembly processes design and experimental evaluation. *International journal of Human-Computer interaction*, 16(3):497–514, 2003.
- [42] ML Yuan, Soh-Khim Ong, and Andrew YC Nee. The virtual interaction panel: an easy control tool in augmented reality systems. *Computer Animation and Virtual Worlds*, 15(3-4):425–432, 2004.
- [43] J. Zauner, M. Haller, A. Brandl, and W. Hartman. Authoring of a mixed reality assembly instructor for hierarchical structures. In *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality*, 2003, pages 237–246.

APPENDIX **SUS Questionnaire**

Press Screw Us	er Test			
* 1. Are you part of the Jobbar du på Press S	team of the press screw Skruv Stationen,	station ?	?	
⊖ Yes				
O No				
* 2. Which system did Vilket system provac	you test? le du?			
Conventional	instructions			
🔿 AR assembly guida	nce			
* 3. To what extend c	o you agree on the follow	ving:		
I think that I would	like to use this system fre	equently for the	press screw assembly	
Hur väl instämmer Jag kan tänka mig a	du i nedanstående påståe att använda detta system	ende: för press skruv	montering frekvent.	
Strongly Disagree				Strongly Agree - hålla med
0		\bigcirc	\bigcirc	
* 4. I found the system	unnecessarily complex.			
Jag tycker att syster	net är för komplex.			
Strongly Disagree	0	\bigcirc	0	Strongly Agree

.

* 5. I thought the system was easy to use. Jag tyckte att systemet var lätt att använda.

Strongly Disagree		Strongly Agree

 * 6. I think that I would need the support of the team leader or Andon to be able to use this system on the real assembly line.

I tror jag kommer att behöva stöd från support teamet eller andon för kunna använda det i den verkliga monteringsline.

Strongly Disagree			Strongly Agree
0	\bigcirc	\bigcirc	\bigcirc

* 7. Everything I needed to place the right screws in the right holes was well integrated in the system. Allt jag behövde för att montera rätt skruvar i de rätta hålen var väl anpassat i systemet.

Strongly Disagree		Strongly Agree
\bigcirc		

* 8. I thought there was too much inconsistency in this system. Jag trodde att det var för mycket inkonsekvens i detta system.

Strongly Disagree		Strongly Agree
\bigcirc		

* 9. I would imagine that most people would learn to use this system very quickly. Jag tror att dem flesta människor skulle kunna lära sig att använda systemet mycket snabbt.

Strongly Disagree		Strongly Agree

* 10. I found the system very cumbersome (slow and/or complicated) to use. Jag tycker att systemet är väldigt besvärligt (långsamt och/eller komplicerat) att använda.

Strongly Disagree		Strongly Agree

* 11. I felt very confident that I did everything correct using this system.

Jag kände mig väldigt säker på att jag gjorde allt rätt när jag använde systemet.

Strongly Disagree			Strongly Agree
0	\bigcirc	\bigcirc	0

* 12. I needed to learn a lot of things before I could get going with this system. Jag behövde lära mig massa saker innan jag kunde komma igång med med systemet.

Strongly Agree

Nånting annat ?

Färdig