MscThesis

University of Turku
Department of Computing
Master of Science Thesis
Laboratory Name
February 2025
Botond Ortutay

UNIVERSITY OF TURKU
Department of Computing

Botond Ortutay: MscThesis

Master of Science Thesis, 16 p.
Laboratory Name
February 2025

Keywords: tähän, lista, avainsanoista

TURUN YLIOPISTO
Tietotekniikan laitos

Botond Ortutay: MscThesis

Pro gradu -tutkielma, 16 s.
Labran nimi
Helmikuu 2025

Asiasanat: here, a, list, of, keywords

Contents

1 Introduction **1**

1.1 The goals of this thesis 1

1.2 Research Questions 1

1.3 Methodology Overview 2

2 Background (IF NOT COVERED BY 1.1) **3**

3 Literature review **4**

3.1 Client-Server Architectures } } 4

3.2 Computer Vision (CV) } } (COMBINE IF NEEDED) 4

3.3 Augmented Reality (AR) } 4

3.4 Prototypes Similar to Ours 4

4 Architecture Description **11**

4.1 Perceived Challenges 11

4.2 Proposed Architecture 11

5 (IMPLEMENTING AN ARCHITECTURE FOR A SOFTWARE SYSTEM WITH AR AND CV) **13**

6 (USABILITY) **14**

7 (FEASIBILITY) **15**

8 Conclusion and summary **16**

8.1 Overview of Results 16

8.2 Answering Research Questions 16

8.3 Summary 16

References **17**

List of Figures

4.1 Visual Representation of the Proposed Architecture 12

1 Introduction

1.1 The goals of this thesis

This thesis aims to design and implement a system that integrates a backend powered computer vision with an Augmented Reality (AR) interface. The concept involves a device capturing an image feed through its camera, which is then transmitted to a processing unit. Here, computer vision algorithms analyze the data to extract meaningful information, which is subsequently sent to the AR interface to give feedback to user. In this thesis, I document the journey of developing this system and assess its performance.

1.2 Research Questions

The thesis aims to answer the following Research Questions:

RQ1: What are the technological challenges in combining advanced computer vision algorithms with an AR user interface?

RQ2: Can a system with a backend computer vision system and an AR user interface be used in a cooking environment?

RQ3: Can such a system provide satisfactory user experience?

1.3 Methodology Overview

The work covered in this thesis mostly consist of the following three phases: firstly I conduct a literature review in chapter 3. The purpose of this is to learn about the technologies involved in this project, to find out what perceived challenges were found by other people working with these technologies and to search for projects similar to ours, conducted by the scientific community. Starting with a literature review should also provide a firm scientific basis to the later phases.

The second phase of the work concerns architecture design and prototype development. All relevant technological challenges found during the literature review are collected to chapter 4.1 for further analysis. Based on all these findings we then propose an architecture for a prototype in chapter 4.2. The prototype is then implemented and the whole development process, the technologies used, as well as anything notable that happens during development is described in chapter 5.

The third phase of the work is an empirical usability study conducted on the developed prototype. Here the finished prototype is given out to test subjects to measure how well the system performs. Usability is measured both through asking the opinions of the test subjects through a questionnaire and through measurements made by the prototype software. All the collected data and the questionnaire used can be found in the attachments. This phase of work is more thoroughly described in the chapters 6 and 7.

2 Background (IF NOT COVERED BY 1.1)

Notes:

- This may be in 1.1 or as a separate chapter (2)
- Right now it is written as a separate chapter, decide precise datastructure later.
- The purpose of the background is to justify the research questions!
  - Phases described in #8 :
    * RQ1 answered by phases 1 & 2 (chapters 3,4,5)
    * RQ2 has to be considered in phase 2 (chapters 4 & 5) but is only fully answered in phase 3 (chapters 6 & 7)
    * RQ3 is answered by phase 3 (chapters 6 & 7)

Tasks:

- Create a more in depth "skeleton" of the bg chapter
- identify, look at, take notes & read through sources you need for this chapter
- Write the chapter

3 Literature review

3.1 Client-Server Architectures } }

3.2 Computer Vision (CV) } } (COMBINE IF NEEDED)

3.3 Augmented Reality (AR) }

3.4 Prototypes Similar to Ours

Before designing and describing our own software architecture it is worthwhile to look at applications other people have developed that solve a problem similar to ours. A great example of such an application is the work done by Pylvänäinen, Solis, Toivola, *et al.* [1].

The work conducted by Pylvänäinen, Solis, Toivola, *et al.* [1] started from a simple observation: They coulnd't find any mobile apps for microscopy education that incorporated AR and Virtual Reality (VR) features and step-by-step guidance. They then sent out a needs assesment survey to students to map out demand for such a software and found that 70% of the respondents showed interest in using such an app in their microscopy studies.[1] They then outlined goals for such an app and started to develop it. Pylvänäinen, Solis, Toivola, *et al.* [1] state that their application should:

1. Be a useful tool in teaching microscopy
2. Help its users to operate a microscope
3. Be a helpful tool with troubleshooting microscopy related issues
4. Be a tool that could be used to revive microscopy knowledge after a long pause in practicing microscopy skills

The app developed by Pylvänäinen, Solis, Toivola, *et al.* [1] consists of three sections: "Teach me microscopy", "Help me at the microscope" and "Help me to troubleshoot". What's relevant to this thesis is under "Help me at the microscope". In that section there is the option to view a 3D-model of a specific microscope commonly used in laboratories (Leica DM RXA microscope). Interactive step-by-step tutorials are also available for this microscope on various things, such as microscopy parts, setting optimal Köhler alignment and focusing the microscope on the sample. These tutorials are also usable outside the virtual microscope in the real world as the "Help me at the microscope"-section also acts as a marker-based AR-environment.[1] AR-markers put on the microscopes can also help students find different parts of the microscope and learn of their functions. This system can also be used to integrate microscope-specific information into this AR-environment.[1]

Pylvänäinen, Solis, Toivola, *et al.* [1] also conducted a questionnaire-based usability study on the app they developed, and found that using the app during microscopy education increased the students' confidence at later using the same microscope independently, without assistance. Furthermore 64% of the students reported that the app definitely helped them at learning microscopy and 90% of them reported that the app helped them recall microscopy skills later.[1]

Another quite similar project is the work carried out by Estrada, Paheding, Yang, *et al.* [2]. Their work was built on a simple goal: To enable students to have a better experience when learning how to use electrical engineering laboratory equipment. They aimed to do this by offering the students AR-based tutorials for various electrical engineering lab equipment, and by using Deep Learning (DL) methods to detect such equipment in the laboratory.[2] The long term goal of this project was to create a framework that could be later used to easily develop interactive smartphone apps for different laboratory devices integrating this concept.[2]

Essentially Estrada, Paheding, Yang, *et al.* [2] developed a superimposition-based AR-app with an integrated DL-model to be used for object detection. This application can be used for a template to create any AR-based tutorial for any device. Creating a tutorial for a device using this framework consists of three steps of work:

1. Training the DL-model to recognize the device
2. Creating an Augmented Reality User Interface (AR-UI) of the device, consisting of a 3D-model and User Interface (UI) panels
3. Creating step-by-step instructions that can be displayed using the AR-UI defined in the first step.

On top of designing and defining such a framework, Estrada, Paheding, Yang, *et al.* [2] actually developed an application using it. In this application the DL-model was trained to detect various types of multimeters, oscilloscopes, wave generators and power supplies. They also created AR-UIs and step-by-step tutorials for using real multimeters.[2]

The application logic of the app by Estrada, Paheding, Yang, *et al.* [2] is as follows: First the DL-model detects (so classifies and localizes) the equipment. Next if a tutorial is available for that equipment, the UI notifies the user of this. If the user decides to view the tutorial, the AR-based tutorial gets loaded and the AR-UI gets superimposed on top of the real object. Then UI-panels are used to display the tutorial content.

Stepping outside of the laboratory setting, Van Gestel, Van Aerschot, Frantz, *et al.* [3] developed an AR-application for helping orthopedic procedures. Namely they wanted to create a tool that would act as a new real-time AR-based safety solution and guidance technique in the use of power tools in surgery. Prior to this there did exist other camera and AR-based surgery systems but Van Gestel, Van Aerschot, Frantz, *et al.* [3] noted them to be physically too large, expensive and time-consuming, which was said to limit their usefulness in assisting with performing surgeries. Van Gestel, Van Aerschot, Frantz, *et al.* [3] wanted to develop a solution that could run on a head-mounted display (HMD) so that a surgeon could use it while working with his/her hands. However one problem was that HMDs typically couldn't do accurate enough tracking for surgical use.[3]

Van Gestel, Van Aerschot, Frantz, *et al.* [3] don't describe the structure and logic of the software built for their task very deeply. They do mention building it for Microsoft's HoloLens headset, and circumventing the poor performance of its camera's tracking ability by using the built in infrared sensor instead.[3] They measured the infrared sensor's tracking accuracy to be below 1mm, accurate enough for surgical work.[3] Their software technically uses marker-based AR, however it must be noted that the markers they use are not physical markers, rather markers registered by an infrared-tracked stylus at the key positions to the surgical operation.[3] These virtual markers were then used to show the users an AR-based guidance vector representing the desired drilling direction when performing the surgery. The vector would also change color to represent whether the current drilling direction was correct or not. If the surgeon was drilling in the right direction the vector would be green, otherwise the vector would be orange or red.[3]

This software was then tested by letting 18 people perform mock surgeries on wooden models.[3] Three surgery guidance techniques were compared: freehand surgery without guidance, proprioception-guided surgery and surgery using the new

AR-tool.[3] The mock surgeries were quite simple: the wooden bone-models had defined entry and exit points between which the surgeons had to drill, with parts of the models being covered by a cloth to better simulate real conditions.[3] The success of mock surgeries was examined by measuring the distances and angles between the desired exit point and the actual exit points drilled by the surgeons.[3] They also performed statistical analysis on the drilling session results, considering the angle between the actual and planned exit points a variable dependent on the experience level of the surgeon, the guidance technique used and the desired drilling direction.[3] They found that the AR guidance tool they developed improved the surgeons' output regardless of experience level.[3] The AR tool was an especially useful tool when performing oblique, complex and angled drilling paths. With these drilling paths the difference between AR-guidance and traditional guidance was even more pronounced.[3]

Another prototype was developed by Monroy Reyes, Vergara Villegas, Miranda Bojórquez, *et al.* [4]. Their aim was to develop an AR system to aid novice users in using milling and lathe machines in a school manufacturing laboratory, and use this as a basis to measure the acceptance rate and performance of such a system in the field of education.[4] Since they designed a tool to aid with the physical operation of industrial machinery, it was important to them that their tool was hands-free as the users would need to operate the industrial machinery simultaneously.[4]

Monroy Reyes, Vergara Villegas, Miranda Bojórquez, *et al.* [4] emphasized that their system was a Mobile Augmented Reality (MAR) system, so essentially it was technically an application built for the Android Operating System. Contentwise their application included tutorials for a milling and a lathe machine which would guide the users in tool setup, working material setup, machinery setup and starting the machines in question.[4] The AR-elements in this app included 3D-models of the machines themselves as well as additional tools, such as spanners and Allen wrenches, as well as text instructions with descriptions on how to perform the basic tasks, labels for helping the user in locating machinery components and tools, 3D arrows to indicate flow direction, and real time videos of task explanations performed by experts.[4]

Since the two-hand requirement made it undesirable to run the application on a smartphone, two AR-devices were chosen to run the application in the experimental phase. Firstly ORA-1 AR glasses, which are optical see-through glasses where the real world is observed through transparent mirrors placed in front of the eyes of the user, and the VR-PRO AR HMD, which is just a HMD with video see-through.[4] Their AR solution was simply Marker-based MAR with two kinds of markers: Frame markers (FM), which are traditional AR-markers, so frame patterns with encoded data in the frame, and Item Targets (IT) which are real world objects that the system tries to match to a 2D image using traditional CV-algorithms such as edge detection and corner detection.[4] The purpose of the markers in this system is to help the system detect the machinery, and provide the system with placement information for the augmentations and explanation videos.[4]

Monroy Reyes, Vergara Villegas, Miranda Bojórquez, *et al.* [4] describe the flow of the application as follows: first the user needs to start the system by scanning the main menu marker. Then they will enter the main menu. From here the user can either start the application or get system help information. The system help information gives the users info on how to use the app. Once the application is started, the user can scan the device markers, which then lets the user scan the individual lesson markers and multimedia markers for that device. This lets the user access the interactive tutorials and educational videos which actually help him/her to learn the operation of that machine.[4]

This application was then tested by 16 students and teachers in the university manufacturing laboratory for an experiment.[4] In the experiment the subjects were first given a general introduction to AR, after which they were introduced to the AR-system developed for this study. Then they were asked to complete the lessons included in the AR-systems on one of the two hardware options that was available. Finally the participants were asked to complete a survey to gather results.[4] The survey was a 10-question survey with a Likert-scale scoring-system. The first five questions related to acceptance metrics (satisfaction, precision, understandability, explanativeness, attractiveness). The next three questions related to performance metrics (interface, speed, marker system). This was followed by a generic "Have you used mobile AR systems before?"-type of question. The last question was an oper answer field where they asked for feedback and comments.[4]

For each question they calculated a score by doing a Likert-scale to point conversion where "very bad" became 1, "regular" became 3, and "very good" became 5, and then calculating the average from all the answers. The scores for the different metrics were: satisfaction: 4; precision: 4; understandability: 4,5; explanativeness: 4; attractiveness: 4,5; interface: 3,5; speed: 4 and marker system: 4,5. Out of these interface was the weakest element, with only 50% of the respondents rating it as good or very good.[4] According to the feedback from the open answer field the teachers and the lab staff participating in this study were interested in using the AR-system developed for this experiment as an educational tool later. The students also saw it as an appealing way to learn machine operation basics.[4] With this it is pretty clear to say that the system developed for this project has potential to be a real-life teaching-learning tool.

4 Architecture Description

4.1 Perceived Challenges

- Do this based on 3
- Mention challenges encountered by others possible solutions if needed
- Add as many subsections as needed

4.2 Proposed Architecture

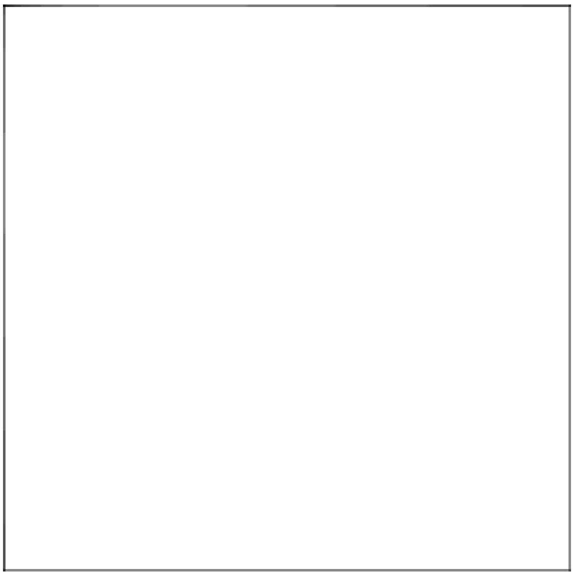

Figure 4.1: Visual Representation of the Proposed Architecture

5 (IMPLEMENTING AN ARCHITECTURE FOR A SOFTWARE SYSTEM WITH AR AND CV)

6 (USABILITY)

7 (FEASIBILITY)

8 Conclusion and summary

8.1 Overview of Results

8.2 Answering Research Questions

8.3 Summary

References

[1] J. Pylvänäinen, J. Solis, D. Toivola, and P. Kankaanpää, "Supporting microscopy learning with ocul-ar, a virtual and augmented reality-powered mobile application", *CEUR Workshop Proceedings*, vol. 3393, pp. 28–37, May 2023, Technology-Enhanced Learning in Laboratories workshop (TELL) ; Conference date: 27-04-2023, ISSN: 1613-0073.

[2] J. Estrada, S. Paheding, X. Yang, and Q. Niyaz, "Deep-learning-incorporated augmented reality application for engineering lab training", *Applied Sciences*, vol. 12, no. 10, 2022, ISSN: 2076-3417. DOI: 10.3390/app12105159. [Online]. Available: https://www.mdpi.com/2076-3417/12/10/5159.

[3] F. Van Gestel, F. Van Aerschot, T. Frantz, *et al.*, "Augmented reality guidance improves accuracy of orthopedic drilling procedures", *Scientific Reports*, vol. 14, no. 1, p. 25 269, Oct. 2024, ISSN: 2045-2322. DOI: 10.1038/s41598-024-76132-3. [Online]. Available: https://www.nature.com/articles/s41598-024-76132-3#citeas.

[4] A. Monroy Reyes, O. O. Vergara Villegas, E. Miranda Bojórquez, V. G. Cruz Sánchez, and M. Nandayapa, "A mobile augmented reality system to support machinery operations in scholar environments", *Computer Applications in Engineering Education*, vol. 24, no. 6, pp. 967–981, 2016. DOI: https://doi.org/10.1002/cae.21772. eprint: https://onlinelibrary.wiley.com/doi/pdf/

10.1002/cae.21772. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/cae.21772.