Institute for Visualization and Interactive Systems

University of Stuttgart
Pfaffenwaldring 5a
D–70569 Stuttgart

Bachelorarbeit Nr. 154

Augmented Reality for Order Picking Using Wearable Computers with Head-Mounted Displays

Sebastian Pickl

Course of Study:	Softwaretechnik
Examiner:	Prof. Dr. Albrecht Schmidt
Supervisor:	Dipl.-Inf. Markus Funk Dipl.-Inf. Sven Mayer Dipl.-Inf. Lars Lischke
Commenced:	April 14, 2014
Completed:	October 14, 2014
CR-Classification:	H.5.2

Abstract

Order picking - the task of assembling orders in a warehouse is still the biggest expenditure in modern warehouses, therefore it is worth it focusing new research on the improvement of technologies and strategies which support this task. The recent development and the enhanced availability of head-mounted displays (HMDs) prompts further research regarding HMDs as a supporting technology for industrial applications. Based on research in related topics and the precise analysis of the use case we composed a concept for a system that supports the order picking process with the use of HMDs. In order to evaluate the concept a prototype that implemented the features of the concept, was built. The range of functionalities contains real time visual user guiding, visual task support, context and much more important step awareness that exempts the user from additional work steps and instant feedback in case of wrong actions of the user which enables direct reaction to errors. Tough a comparative user study, we compared the prototype with two traditional order picking methods and another order picking support system that uses projections for user guiding. The results show, that our prototype shows the lowest occurrence of long therm errors with the downside of a higher task completion time. However the study proved that the concept is worth being taken into consideration and be improved in further research.

Kurzfassung

Komissionierung, die Aufgabe, Bestellungen in einem Lager zusammenzustellen, ist immer noch der größte Kostenpunkt in modernen Warenlagern. Deshalb lohnt es sich, Forschung auf die Verbesserung von Technologien und Strategien, die diese Aufgabe unterstützen zu fokussieren. Der jüngste Aufstieg und die verbesserte Verfügbarkeit von Head-Mounted Displays (HMDs) fordert Nachforschung im Bezug auf HMDs als unterstützende Technologie für industrielle Anwendungen. Basierend auf Forschung in verwandten Themen und der genauen Analyse des Anwendungsfalles haben wir ein Konzept zusammengestellt, welches den Komissionierprozess mit Hilfe von HMDs unterstützt. Um dieses Konzept auszuwerten wurde ein Prototyp gebaut, der die Eigenschaften des Konzepts umsetzt. Der Umfang der Funktionalitäten umfasst visuelle Echtzeitbenutzerführung, visuelle Aufgabenunterstützung, Bewusstsein für den aktuellen Kontext und, viel wichtiger den momentanen Arbeitsschritt, was den Benutzer von zusätzlichen Arbeitsschritten befreit und sofortige Rückmeldung im Falle falscher Handlungen des Benutzers, was direkte Reaktion auf Fehler ermöglicht. Dieser Prototyp wurde in einer Benutzerstudie mit zwei traditionellen Systemen zu Komissionierungsunterstützung und einem weiteren System, das Projektionen zur Benutzerführung benutzt verglichen. Die Ergebnisse zeigen, dass mit unserem Prototyp am wenigsten Langzeitfehler auftreten, was eine längere Zeit zur Erledigung der Aufgabe zur Folge hat. Die Studie hat bewiesen, dass dieses Konzept in zukünftiger Forschung mit beachtet werden sollte und sich eine Weiterentwicklung lohnt.

Contents

List of Figures

List of Tables

List of Algorithms

List of Abbreviations

HMD head-mounted display

M mean

PbL Pick-by-Light

PbPa Paper-Picking

PbPr Pick-by-Projection

PbVi Pick-by-Vision

PbVo Pick-by-Voice

SD standard deviation

SE standard error

1 Introduction

In times where single companies like Amazon.com offer over 270.000.000 items to buy[1], where a big share comes from their own warehouses and single cars are built from thousands of single parts in thousands of different variations, fast and reliable warehousing gets increasingly important. Parts must be delivered just in time for production or delivered as fast as possible, in order to beat out competitors. Being able to provide a three hours delivery service, as cyberport is currently advertising[2] is not only the result of fast transportation, but also of intelligent warehouse management and a efficient order picking process.

The characteristics of efficiency in view of order picking is not only speed, but also accuracy and flexibility. It is not enough to deliver items fast, it is also necessary to deliver the correct amount of the right item. This work is in the most cases done by human workers, who are supported by some kind of guiding technique. These techniques are worth to be explored and improved in order to enhance picking speed and accuracy, while keeping up the flexibility of a human picker.

MotionEAP is a project at the University of Stuttgart in cooperation with the Hochschule Esslingen and partners from the industry that focuses on an increase of efficiency and assistance in production processes. Therefore, motion detection and in-situ projection are used to support production processes[3][FKS14]. As part of this project, an assembly table that uses motion detection and projection provides context sensitive augmented reality support for the current work step. The goal of this thesis is, to evaluate, if this attempt can be transferred to the warehouse environment by building a prototype, that realizes step aware order picking support using head-mounted displays and evaluating it in a study, similar to Guo et al.[GRX+14].

The recent rise of HMDs leads to a lot of companies, that try to be the first to build a consumer product, that serves the users needs best. A good example is the Google Glass[4], but other companies, like Vuzix[5] and Sony[6] are also presenting their HMD solutions. HMDs enable developers to build augmented reality applications that have not only the power to deliver a

[1]http://www.amazon.com/
[2]https://www.cyberport.de/eilt!
[3]http://www.motioneap.de/
[4]https://www.google.com/glass/start/
[5]http://www.vuzix.com/consumer/
[6]https://pro.sony.com/bbsc/ssr/mkt-digitalcinema/resource.latest.bbsccms-assets-mkt-digicinema-latest-EntertainmentAccessGlasses.shtml

(a) A Picker is wearing a HMD for order picking support.

(b) The view of the picker through the glasses can be seen here.

Figure 1.1: The prototype, that is designed, built and evaluated during this thesis.

completely new kind of entertainment and information, but also to actively support working processes as for example manufacturing or repairing processes [CM92].

HMDs seem to be the perfect solution for mobile order picking support. In order to evaluate this theory, a concept of such a system is created in the context of this thesis. This is followed by the development of a prototype, which is shown in Figure 1.1 that includes the main technical properties of the concept and offers an accurate user experience of the possible later product. A comparative study shows how the prototype can be improved and if it can compete with traditional methods and other recent developments.

The concept presented in this work provides:

- Context Sensitivity

  The provided information is dependent on the users location, the environment and the current work step in order to avoid information overloading. The picker is supplied with all relevant information at the right time.

- Visual Task Support

  Visual hints are indicating the next work step in an intuitive way. This is done to reduce the mental workload and direct the focus to the actual task.

- Task Evaluation and Feedback

  The execution of the current works step is monitored automatically. As far as possible, the actions of the user are evaluated and proper feedback is provided based on this evaluation. The instant feedback will help to reduce long term errors, because it points the worker to correct them right when they happen. Additionally, the learning phase with the system can be reduced by providing the possibility to discover the impact of certain actions.

Apparently, the system can not provide all this functionalities while only running on a pair of glasses. Therefore, the prototype contains a whole infrastructure of sensors and computation devices, which provides the relevant information for the glasses. The focus there is to integrate them into the existing environment and provide an effortless mean of interaction, that is completely based on natural interaction with the working environment. Work steps are implicitly triggering actions of the system in a context aware manner [SBG99]. Therefore the user is not required to perform additional actions to control the supporting system.

This thesis aims to evaluate, if this kind of support is suitable for the order picking use case in form of a comparative study. Objective measurements are therefore included as well as subjective feedback from participants of the user study. Objective data contains error rates and task completion time, which helps to compare the performance directly with other supporting and guiding methods. These informations are be completed by subjective user feedback, that makes it possible to predict the acceptance for the product by later users, long term performance and user satisfaction.

Structure of the Thesis

This thesis is structured in the following way:

Chapter 1 – Introduction: The introduction explains the motivation for this thesis and provides a short overview over the following chapters.

Chapter 2 – Related Work: An Overview over the related work, done in all related topics, which are order picking, HMDs, augmented reality and indoor navigation.

Chapter 3 – Concept and Design: The chapter concept and design describes the properties of a product that serves the order picking purpose best.

Chapter 4 – Technologies and Tools: A prototype that implements the previously and all implemented features are described from a technical point of view.

Chapter 5 – Evaluation: The evaluation method and the results of the evaluation are presented in this chapter.

Chapter 6 – Conclusion and Future Work: All findings and experiences are summed up in the conclusion and extended by advices for future work.

2 Related Work

This chapter covers the topics that are relevant in order to understanding the reach of the subject and to get an overview over the state of the art of existing technologies. This includes order picking, since it is the work task, that we aim to improve, it is important to be aware of all aspects of this topic. Second, there are head-mounted displays, because it represents a very good tool for supporting work tasks. It also enables Augmented Reality applications, which are discussed in Section 2.3. Indoor navigation is also an important topic to explore in order to develop the ideal order picking support. The discussion will finally summarize the findings of the previous sections and sets them into the context of this thesis.

2.1 Order Picking

Since order picking is the most expensive task in warehouses, with 55% of all warehouse costs, optimizing the picking task is a worthwhile goal for warehouse owners. Li et al. [LCTM12] mention two major challenges: minimizing the time, to find the proper part bin and minimizing the error rate for the picking task. They substantiate their statement according to the task completion time with findings from Tompkins [Tom10], who points out, that pickers spend 50% of their work time with travelling and additional 20% with searching. Another expensive entry mentioned above are errors. They can cost time, for correcting the error, money because the customer wants the error in his order corrected or may even lead to loss of customers.

The time, a picker needs per pick depends on two main factors. The first is the picking strategy, which is one of the main differences. The second is the picking method, which includes the guiding technology. In order to optimize the task completion time, it is important to take a close look on both subjects.

First of all, the picking strategies are mainly dependent on the warehouse type. There are two different systems, which are in use: man-to-goods and goods-to-man picking. The two types differ in the way, the picker gets to the desired part bin. At the man-to-goods strategy, the picker travels along the shelves, in order to collect the required parts out of static part bins. Compared with that, the goods-to-man system works the other way round: The worker has a fixed work place and the part bins travel towards the picking station, where the collecting and sorting work is done [VHCW13].

Batch Picking describes a process where one worker is handling several orders in a parallel manner. This can either be realized by first picking all the parts for all assigned orders and sorting them in an additional step or by sorting the items in the same step as the picking process itself takes place. This can for instance be done with a picking cart, that holds a box for every order. The advantage of this method is that the pickers travelling time can be reduced by organizing all picks from all registered orders in an optimal way [Bau13].

For *zone picking*, the warehouse is divided in several smaller zones. In that way, multiple pickers can process different parts of the same order at the same time and bring it together in the end. This also allows to skip zones, where nothing needs to be picked [Bau13].

The second big topic do discuss is as stated above, the picking method. There is a wide range of picking methods, which are either used in actual work environments, or existing as concepts, which need to be tested and improved to get ready for commercial use. The most important ones will be described in the following.

Paper-Picking (PbPa) is a method where the worker gets the work instructions handed over in paper form. The paper can either show a list of items to pick or a graphical representation, like simplified pictures of the shelves with marked fields on it. In order to identify the correct parts, the shelves and part bins are marked with letters and numbers (e.g. shelf A bin 5). A variation of the PbPa system is the use of picking labels which contains for each picking task only one instruction. The advantages of PbPa systems are the high flexibility and the low costs. On the other hand are the high error rate and lower picking speed which are the result of missing pick validation processes and guiding methods [Bau13, WBS+10].

For *Pick-by-Vision (PbVi)* special equipment for the picker is needed. One thing is some kind of HMD, to visualize the picking instructions and the second one is a computation device, in order to generate and process these information. The HMD offers different possibilities to support the work process. With minimal hardware expenditure, the worker can be provided with the same representations as with the PbPa method, which are text form and graphics that visualize the tasks. With additional hardware, like a camera or a motion tracking system, this picking system can be extended to an Augmented Reality system, which provides all information in an visual real world overlay, which expands the physical environment. The benefit of HMD Picking is the fast availability of the required information and the possibility of hands free guiding and working as well as the low costs, that are the result of the high scalability and high flexibility [WBS+10, SRG+09].

Pick-by-Voice (PbVo) also requires some additional hardware: A headset for input and output and a wearable computer are needed. The wearable computer generates the picking information and voice output provides the picker with instructions. The system can be controlled by voice commands. Therefore a speech recognition software identifies simple commands to confirm, repeat or skip tasks. This method enables hands-free working but is strongly dependent on environmental factors like noise [WBS+10, RGSK09].

Pick-by-Light (PbL) is also a computer-assisted picking system, which guides the picker to the correct part bin. Every part bin is equipped with a light signal, that indicates which and how many parts have to be picked. Taljanovic et al. [TSP12] mention a higher picking speed, the hands free working and a short familiarization period as the main advantages of this system. The downside is the poor scalability, which results of the additional hardware, that is needed for every new shelf and part bin [RGSK09].

This and other, less common methods where compared in other studies before. Guo et al. [GRX+14] discovered, that a cart-mounted display offers a better performance, than the other compared methods, which are PbPa, PbL, pick-by-HMD, tightly followed by pick-by-HMD. While Schwerdtfeger et al. [SRG+09] evaluated an Augmented Reality HMD system, that was however not step-aware against a 2D HMD system and the traditional PbPa list. They discovered a better performance of the Augmented Reality prototype, compared to the other methods.

As mentioned above, a second field with big optimizing potential beside improving the picking speed is error prevention. The reason is that every error causes financial disadvantages, which reach, dependent on the severity of the error and when it is detected, from a delay of some seconds to loss of customers and bad public perception. In order to identify, classify and weight errors properly, it is helpful to define categories.

Lolling identifies four types of errors for picking tasks [Lol03] :

- Amount-error: The picker picked the wrong amount of the correct part
- Skipping-error: A picking task was skipped by the picker
- Type-error: The picker picked the wrong item
- State-error: A picked item has not the required nature

In order to prevent those errors some approaches have been developed in commercial use and research. Baumann et al. [BSI+11] used a laser rangefinder, to detect picks and compare them with the picking instructions of a HMD picking system. This allows the system to detect, if the picker picked from the correct bin. If the picker reaches into a wrong bin, the system detects the error and notifies the worker. It is impossible to proceed, until the system has detected that the picker has corrected the error and reached again into the bin. The systems limitations are, that only skipping-errors and type errors can be prevented. Amount-errors and state-errors are still undetected. Baumann et al. [BSI+11] also found, that a bad error detection could confuse the picker and they detected a slower picking speed because of this confusing situation.

Another method, introduced by Li et al. [LCTM12] is based on monitoring the bin content. Because of the high hardware costs, this method is more likely to be used in goods-to-man warehouses. In this specific scenario, a Kinect sensor is used to collect RGB and depth information of the present bin. The pictures before and after the pick are compared and the missing items are identified. Together with the data, stored in a database, the system can

(a) Epson Moverio BT-200 (Source: Epson[1]). **(b)** Google Glass (Source: wikimedia[2]). **(c)** Oculus Rift (Source: Oculus Vr[3])

Figure 2.1: This image sequence shows a row of currently available HMDs. Each product represents an HMD category.

evaluate, if the correct parts where picked. Some limitations are, that the items that stay in the box should not be moved, in order to recognize the picks correctly. Currently it is only tested with box shaped items. Li et al. state, that the method would also work on other common shapes, like cylinders. Irregular shapes, such as unpacked screws cannot be detected by the system. If the system can be developed to an industry ready state, it could detect amount-errors, skipping-errors and type-errors. State-errors can only be found in a limited scope.

2.2 Head-Mounted Displays

A HMD is a device, that is placed in a low distance to the users eyes and either contains two small displays or projectors that project the images directly on the users retina. For order picking support, optical see through HMDs are more suitable. In contrast to non digital see-through, or Virtual Reality devices, they do not block or delay the view on the real, physical world, which is important for tasks where proper eye-hand coordination is important. Particularly, when non-blocked real world view is safety-critical [HG07]. HMDs have been a subject of research for some years now. Ivan E. Sutherland has built the soft- and hardware for a three dimensional HMD for Augmented Reality applications in 1968 and even he referenced some earlier work [Sut68].

A state of the art HMDs is for example the Google Glass, shown in Fig. 2.1b, which is still in a pre-public release program. The official specification refers to the used display only as the "High resolution display is the equivalent of a 25 inch high definition screen from eight feet

[1]http://www.epson.com/cgi-bin/Store/jsp/Product.do?sku=V11H560020

[2]By Tim.Reckmann (Own work) [CC-BY-SA-3.0 (http://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commonsg

[3]https://www.oculusvr.com/order/

away."[4]. However Google Glass uses only one eye for visualisation. The area of application for the Google Glass is bringing smartphone functionality directly to the eyes. This means that the glass is mostly dependent on an external computation device which provides mobile network connection and an interface for simplified device management.

In Fig. 2.1a, one can see another recently launched device, the Epson Moverio BT-200 which consist of two see-through displays mounted in glasses. The Epson device comes with an additional computation device and controller, which has a wired connection to the glasses. In contrast to Google Glass, the Epson Moverio runs a completely working customized Android operating system. It however does not support phone functionalities such as sending text messages, making a call or mobile networks. Its network connection is dependent on Wi-Fi. The coverage of the Moverio BT-200 goes from entertainment to medical science[5].

There are also non-see-through HMDs on the market. A popular example is the Oculus Rift by Oculus VR(Fig. 2.1c). This product is also in an open beta phase, where it is available to developers and early adopters, but hardware and software are still under development. The Oculus Rift is a Virtual Reality headset, designed for gaming applications. It offers a resolution of 960 x 1080 pixels per eye, positional tracking, a 100 degrees field of view and various Sensors. The headset is wired, which means, that the computation is done by an external device. This could be a computer or gaming console. However, the use case, the device was designed for is not a portable one.[6] Samsung tried to overcome this disadvantage in cooperation with Oculus VR and developed the Samsung Gear VR. This device is completely wireless and can be connected to mobile devices.[7]

2.3 Augmented Reality

Augmented Reality is a term that is widely used and abused as a buzzword for modern applications on mobile devices, that are in any case context sensitive. This blurring of the term makes it nearly impossible delimit it to certain properties, or clearly assign devices or applications. Therefore a proper definition that provides a clear demarcation is needed.

Augmented Reality applications aim to extend or supplement the real environment of the device, they are running on, in a visual way. In contrast to Virtual Reality, they combine real world objects and virtual objects in real time, in a way that results in a logical meaningful picture. Which means, that objects that are further away appear smaller and objects which are partly covered, are not displayed completely [A^+97].

[4] https://support.google.com/glass/answer/3064128?hl=en
[5] http://www.epson.com/cgi-bin/Store/jsp/Product.do?sku=V11H560020
[6] http://www.oculusvr.com/
[7] http://www.oculusvr.com/blog/introducing-the-samsung-gear-vr-innovator-edition/

Ronald T. Azuma [A+97] defines Augmented Reality as a system, that meets the following requirements:

1. Combines real and virtual
2. Is interactive in real time
3. Is registered in three dimensions

The application areas for Augmented Reality are very varied. It can be used for entertainment. A completely different mean of interaction offers many possibilities to develop games that combine real world information with virtual environments or objects. This also enables educational use by simulating experiments or situations.

Another big sector is supporting work processes in many fields, like manufacturing or medical tasks by displaying instructions or visualizing additional data. This can improve the productivity and security for tasks that are rarely done or vary often, for example repairing a car or perform an operation. The advantages in using Augmented Reality for supportive tasks can also be used for real world search and guiding scenarios [FBP+14, SK08]. Since Augmented Reality allows to highlight objects that are invisible or difficult to see, it is useful for military applications, for example in helmet visors of pilots [A+97].

2.4 Indoor Navigation

An important part of a order picking supporting system is a intuitive and accurate navigation technology and visualization. The problem of guiding the user must be solved in large scale as on room or corridor level as well as in small scale, when the users view has to be guided to look on the correct compartment.

The subject of indoor navigation breaks down into two major topics: the visualisation and the technology to locate the user. In the following existing solutions for the visualisation will be described in detail.

Löchtefeld et al. [LGSK10] use an approach to accomplish this task which is similar to the visualization GPS based large scale navigation devices are using. A map shows the current position. This makes the representation independent from the environment and the generic design allows to generate maps for places, where only basic information like a floor plan are given, without precise knowledge or dependency on the environment.

Another possibility is to provide the user with a list of graphical or textual instructions to follow, in order to reach the target. The problem with this kind of representation is that it requires the user to divide attention between walking and reading the instructions. In addition, the user has to map the instructions to the surrounding environment, which is increasing the mental demand [BKS+10]. Mulloni et al. [MSS11] tried to overcome this obstacle by adding an additional view, where the commands where graphically inserted in-situ into an Augmented Reality view which shows a real time video of the front camera of the used smartphone.

There are also solutions that directly address Augmented Reality devices. One of them is the so called "Attention Funnel" [BTOX06]. The concept behind it is to provide an omnidirectional mean of guiding that minimizes the mental demand, by displaying a tunnel directly into the users field of view. The tunnel guides the user to the desired target, while it is updated for every displayed frame in order to provide up-to-date information. A study, conducted by Biocca et al. [BTOX06] discovered, that the attention funnel was faster in helping people finding objects than verbal instructions and visual highlighting as proposed by Bonanni et al. [BLS05]. They determined also a reduced perceived mental workload for their participants. The big advantage of the attention funnel is, that it works in small and big scale. An other interesting fact is, that the attention funnel has been used for a order picking prototype before by Reif and Günthner and Schwerdtfeger et al.[RGSK09, SRG$^+$09]. Reif and Günthner discovered a long learning phase for the PbVi system. They also found, that PbVi is significantly faster, than PbPa and a lower error rate for PbVi, which is statistically not reliable because of the small sample size. However both of the existing prototypes are providing context awareness according to the location of the user, but lack of automation regarding the users work steps and task validation.

In order to notify the user of the location of off-screen objects or events, Baudisch and Rosenholtz [BR03] introduced Halo. A system, that marks off screen objects with circles, that barely reach over the edge of the field of view. The advantages of this technique are, that the user can determine the direction, where the target is located and its distance to the screen from the location and the size and curvature of the circular segment. However, a downside of this approach is, that a large amount of circles quickly result in cluttering when circles start to overlap. A possible solution to this problem could be, to replace the circles with wedges that reach into the display, similar to the circles in Halo [GBGI08].

In order to provide a proper and precise visualisation, it is also important to reliable locate the user. There are various methods available to locate the user in space in indoor environments. Some are more, others are less accurate or dependent on additional infrastructure or hard- and software.

A system called LANDMARC utilizes active RFID tags and estimates the position of the target dependent on the detected signal strength measured on a cluster of RFID readers. Reference tags are used to improve the results. The benefits of LANDMARC compared to similar systems are the reduced amount of used RFID readers, the flexibility in terms of environment and changes in the environment and an improved accuracy and reliability. On the other hand, the system also has its downsides. One is the high computational demand, another one is, that the accuracy is dependent on the amount and density of deployed reference tags, while a higher density of tags is accompanied by a higher interference effect among the tags [JLP06]. With consideration for the special use case, LANDMARC is not suitable, when three dimensional location information is needed.

Other methods focus on Wi-Fi signal strength. Lim et al. are suggesting a system, similar to the LANDMARC system, where they install an array of smart antennas. In this case, triangulation is used, because of the better accuracy compared to the nearest neighbour method. This technique still has a high deviation rate, which makes it only limited suitable for reliable location tracking. It also requires a lot of additional hardware and software [LWNS07].

Since there are existing methods to use QR codes as AR markers, it has to be possible to determine the position of the camera that recognized the code [KTC09]. This could deliver position information as long as at least one code is visible in the current frame. This should work for small scale location detection. For a larger scale, the individual codes can be equipped with information about their location. In addition, a dead reckoning approach as described by Steinhoff and Schiele can be used to bridge over the time, where no QR code is visible to the camera [SS10]. The advantage of such a system would be the low costs, because only a camera and some QR codes are needed, the scalability, because a bigger area only requires more QR tags and not more cameras and that the view direction is considered. However, the downside is, that one has to rely on guessing in the dead reckoning phases.

A far more accurate method is provided by commercial systems like the OptiTrack or the motion tracking system by qualisys[8]. In this case, infrared cameras track the position of passive markers in space. The downsides of this approach are the high costs of the system and the low scalability. In order to extend the tracked space, more stationary cameras have to be added to the system.

2.5 Discussion

We investigated the previously discussed topics in order to acquire a knowledge base for building a product by combining state of the art technologies. The section Order Picking shows, that a efficient order picking supporting system should optimize two factors. First of all, the task completion time needs to be minimized, by guiding the user as fast as possible to the proper destination and preventing manual overhead during the task execution by automating the system controls. The second important point to minimize is the error rate. Errors can lead to massive costs and delays. This can be achieved by monitoring the users actions and providing adequate feedback. HMDs are serving best for this task, because they enable hands-free interaction and provide a mean of constantly displaying important information to the user. However, it has a big impact, what kind of HMD is used. Optical see-through HMDs are providing the necessary free view on the working environment, that is not blocked by any computation process, as it is the case with non-see-through displays. This kind of hardware enables us to develop a system that is based on Augmented Reality. The challenges are, to include the visualisation seamless into the users real world view and provide non intrusive

[8] https://www.naturalpoint.com/optitrack/

feedback. The current state of the art in the subject of indoor navigation helps us to build a product, that serves the users needs best. Previously built prototypes for PbVi systems are using a concept, that is called attention funnel, which appears to be promising. However, the context awareness of current prototypes is restricted to the users location. Work progress has to be reported manually, which represents a distraction for the user. It is important to prevent such distractions in order to generate a streamlined workflow. Therefore we will provide a step aware system, that monitors the users actions and reacts directly to them.

3 Concept and Design

In order to provide a convenient and usable method to speed the order picking process up and reduce the error rate for minimizing overall warehouse costs using augmented reality with wearable computers and head mounted displays, several approaches are possible. In this chapter, the design decisions that where made in the development process are described and explained in order to give insight into the background and reasons for the decisions.

In order to understand those decisions, it is important to be aware of the underlying scenario and the use case. The software and hardware components of the system should all be designed to optimize the order picking process. Therefore a concrete scenario has to be defined. The following section describes the scenario, that is the basis for all design decisions.

The location is a large warehouse, consisting of several rows of shelves. Every shelf has a number of compartments, where goods are stored, while each compartment holds exactly one item type. The goods are collected and sorted into orders by workers called pickers. Every picker is processing several orders at the same time and the goods are sorted into the picking cart. The picking cart belongs to the working equipment of every picker. They are holding different compartments, which are each assigned to an order. Every picker gets an amount of orders assigned. The task is now, to fulfil every order correctly, which means picking the right amount of an item from the correct compartment in the correct shelf of the warehouse and put it into the correct box on the cart. When a worker has finished all orders, he returns to the start position and hands in the picked items. As mentioned in Section 2.1, the focus is on a fast picking speed and a low error rate.

3.1 Functionalities and Requirements

Since the supporting system is used to improve the order picking process, there are some requirements to meet. First of all, the interaction with the system should provide a maximum of implicitness, which means that there should be no additional steps necessary to control the system. The system is rather meant to support the user and reduce task complexity, than to be another expenditure that adds complexity to the workers every day work. The desired effect is a natural and easy work process without distractions. Therefore, the system should automatically recognize and evaluate the users actions and decide independently what would be the appropriate response. It should further enable hands-free interaction, in order to set the focus on the work task and not on handling hardware and software of a system that should actually act supporting. The requirement is, that the system merges seamlessly into the surrounding environment. This will reduce the time, the worker spends on one task, because he does not have to check for a paper list or manually request the next step, like it is the case in current systems. Another requirement is, to reduce the error rate, in order to provide efficiency and accuracy by direct feedback and clear instructions.

3.2 Solution and Concept

The first step to achieve implicit and distraction-free interaction can be done by using a wearable display. The requirements however are that the display does not block the users normal view but rather extends it, therefore a see-through display serves best. It offers the possibility to enrich the users perception with additional information without setting aside the benefits of a clear view on the work area. Since glasses can be worn without the use of hands, the hands stay free to accomplish the task. The worker does not have to switch between lifting big or heavy objects and viewing information. It can all be done at the same time, which should also lower the task completion time, because the mental overhead is cancelled out as well as the physical expenditure. To prevent information overload, the system should provide context sensitivity and only display the information, that is of use in the current situation.

Another important design decision is, how to communicate the task to the user in an intuitive way. If we take a look on order picking tasks, they are mainly consisting of reaching into boxes or shelf compartments and either picking or dropping a defined amount of parts. The first step to create an appropriate task representation, would be to find a metaphor to represent the correct boxes as like numbers and letters represent the boxes in currently used systems as PbPa and PbVo. The big advantage of head-mounted displays is, that they offer the possibility to extend the reality with virtual objects, that are projected into the users view in real time. This can be used, to highlight the proper box. Due to inaccuracies in determining the current position of the glasses, it could get confusing, to represent the boxes by a highlighted model of the box on the screen, if it does not fit correctly. Therefore a sphere that hovers in the centre of the compartment or the box to highlight will be less distracting. The sphere can be placed with more tolerance inside the box.

To completely make use of the potential of augmented reality systems, the spheres can also be used to highlight the compartments, where the worker has made an error and the boxes of the last pick, in order to provide the possibility to correct amount errors. If the spheres are utilized for more than one purpose, the meaning of the sphere has to be stated in an other way. A simple way to do this is to use colours. The user can be pointed to errors by red spheres, to the next target by green spheres and so on. While this might appear as an easy way to display information to the user, one should always keep in mind, to keep the visualization as simple as possible, because the user still has to take care on his environment and the task itself.

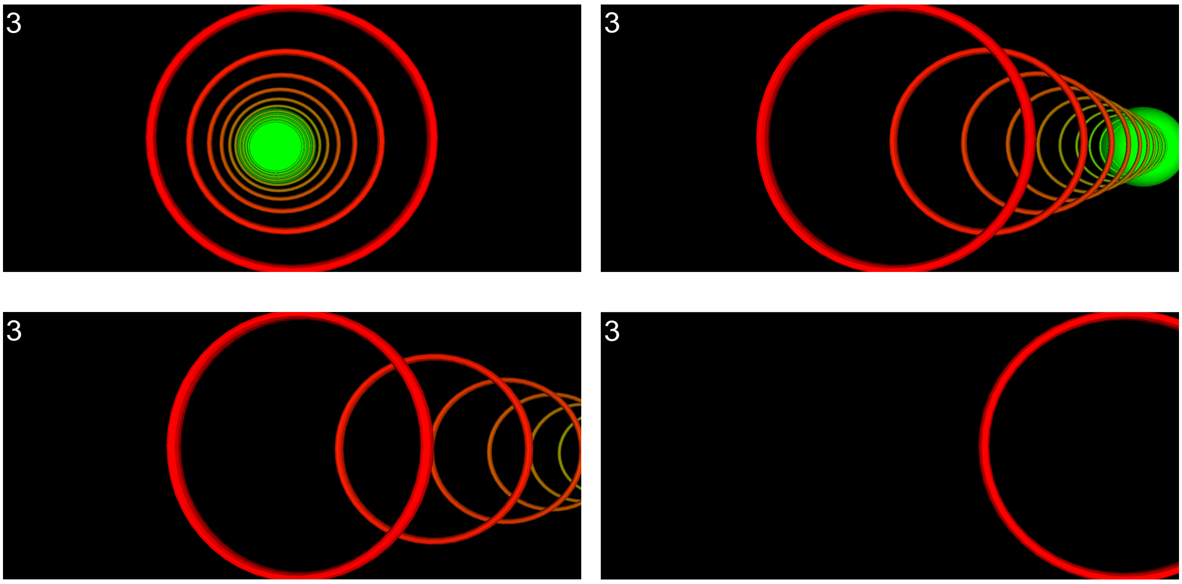

Figure 3.1: This series of screenshots shows the behaviour of the attention funnel, when the users turns their head.

The final problem to solve is how to display the amount of parts to pick. This can easily be achieved with a number in a corner of the screen. There it will always be visible and nearly without any effort readable, but it does not block or distract the users view.

3.3 Guiding the User

While designing the task representation, it was not important if the user could see the highlighted spheres because the position of the spheres depends on the actual location of the compartments in the real world. But now, another question comes up: how can we point the user to the current task, if it is off-screen or even in a different part of the warehouse? In Section 2.4 means of guiding and possible visualizations are discussed. The requirements for the guiding technique are, that it should work in large and small scale in order to guide the user to the right shelf and then identify the correct compartment. Since the user with the glasses can move freely in every direction and angle, the system should provide clear instructions for every position in space.

It turns out that the attention funnel meets the requirements best. It points the user to the target, no matter how far away it is. Nevertheless, some modifications took place. In order to be independent of orientation of the users head and the target, the tunnel consists of toruses. The round shape guarantees, that the tunnel fits always to its start and target, independent of their orientation. A step to simplifying the model and optimizing the computation time was to use a straight tunnel. Therefore the tunnel starts in front of the users field of view and ends in front of the target. If the user looks directly into the direction of the target, the target will be displayed at the end of the tunnel. When this is not the case, it is clear as to what direction the tunnel is pointing. In order to strengthen the impression of depth and adding another realistic degree, the tunnel should be a 3d object that follows the rules of perspective. The perception of depth can be intensified by slowly fading the colour from tunnel segment to tunnel segment from red to green. Since red is well known to express something wrong and green on the other hand represents something correct, this can support the guiding process. An exception is if the viewing angle in relation to the target is greater than 90 degrees left or right, the tunnel disappears and gets replaced by a circle that protrudes into the field of view either on the left or right side of the screen as it can be seen in Fig. 3.1. On which side the circle appears depends on, in which direction the user has to turn his head in order to get the tunnel back into his field of view. The system should always try to suggest the shortest possible way. In this case, the attention funnel is used for a small and large scale scenario. Therefore, it would be desirable, to provide a medium, to estimate the distance to the target. This can be realized by assembling the tunnel from a fixed amount of circles and vary the distance between the circles. The user can see if the circles have a big distance between each other, which means a high distance to the target or if the opposite is the case. The circles should however not get too intense if the user stands close, therefore a good value proved to be ten.

3.4 Picking Cart

Part of a pickers equipment in a large warehouse is a picking cart. This cart can be utilized as an augmented cart, that enables context aware order picking support without any additional user interaction, so that the picker can concentrate completely on his task and is not bothered by the overhead and providing input for a program that should actually support the work. There are two areas that can be monitored. The first one are the boxes that are mounted to the cart and where the picker places the picked parts, the second one is the shelf in front of the cart, where the picker takes the parts from.

In order to recognize, in which box on the cart the picker has reached in, some sort of light barrier should recognize, when and where a pick took place. The main challenge is to get the proper coordinates, where the hand penetrates the monitored plane. This will bring a major advantage in evaluating the users action and giving direct feedback. The design of the cart should allow to mount some kind of sensing device straight above the box entrances for producing the best results. A minimum height is required to provide an adequate coverage of the boxes.

The shelves represent a way bigger challenge: they are not in a fixed position to a cart mounted sensing device, since the cart moves along the shelve and can be placed in any position or rotation to the shelf in front. Another problem in this case is the user, who could possibly block the way between a sensing device and the monitored area. To improve the results and minimize the risk of unrecognised picks because of blocked ways, the sensing device should be mounted as near as possible to the shelf. This would result in a similar situation as the pick detection on the cart. In this case, the monitored plane cannot be aligned with the direction of view of the sensing device, therefore the plane has to be adjusted whenever the cart is moved.

This kind of augmented cart enables context aware order picking support, without any additional user interaction.

3.5 Instant Feedback

It is important to inform the user about the correctness of his actions in order to avoid errors and prevent long term effects of mistakes, such as complaints about wrong orders by customers up to stopped assembly lines because of missing parts which could result in high costs or customer loss. The speed of the feedback is critical for the task completion time. The optimal case would be, if the user gets notified of errors while or even before making them. This would help to interrupt wrong actions.

Another desired effect is to provide an implicit learning phase. When the user gets introduced to the device, a try and error approach teaches the user what are correct actions and what actions are wrong. Since exhortations only appear in case of errors, the system will automatically adapt to the users skill level.

Feedback is necessary in two situations: Firstly, if the user performs a correct and secondly if they perform a wrong one. In case of a correct action, for example a correct pick, the feedback includes, that the system displays the next work step. This can be combined with a colour change, to make the correct action visible. However, incorrect actions need a special kind of feedback. It is not enough, to highlight the wrong compartment in some way, because the user should be able to recognize an error even if it does not happen in his field of view. A solution to this problem could be a flashing red lettering that reads error in the users view, independent of the viewing direction.

4 Technologies and Tools

The prototype, built in the context of this thesis consists of a set of completely new developed software and hardware, combined with existing products and has been assembled to a working system. In the following all components and their interaction will be described and explained in order to give a detailed view of the functionality and implementation details. Starting with the Hardware Setup, where all the used hardware is described in detail, followed by the Software Setup, where the corresponding software and the interactions between different parts of the system are described.

4.1 Hardware Setup

This section describes the hardware setup of the system, which consists of a specially designed cart, the motion tracking system, two Kinects, a related computer and a pair of smart-glasses. All this components are integrated into one system that is controlled by a set of software, that will be described in Section 4.2.

4.1.1 Cart

The biggest component of the prototype is a specially designed cart for order picking tasks. It provides holders for 49 small load carriers on each side, where the picked goods can be placed in. This modular construction makes it possible to rearrange and exchange the boxes easily for fast unloading. The top of the cart is designed in a way, that enables a maximum of flexibility. Dependent on the use case and the environment, the hight of the cart can be adjusted up to 3382 millimetres. However, a lower hight results in worse performance and coverage for the pick detection for both, the shelf and the boxes on the cart. A construction on the top of the cart carries two Kinects, that monitor the boxes on the cart a well as the shelf in front of it. Therefore, a bar reaches 1100 millimetres from the middle of the cart over the working area. The technical drawing and a photo of the end product are shown in Fig. 4.1. Computational power for the pick detection is provided by a computer, that rests on the linkage on the bottom of the cart. The mobility of the vehicle is guaranteed by four wheels, of which two have fixed directions and only the rear two allow the user to steer. A handle on each side makes it possible to push or pull the cart in two directions.

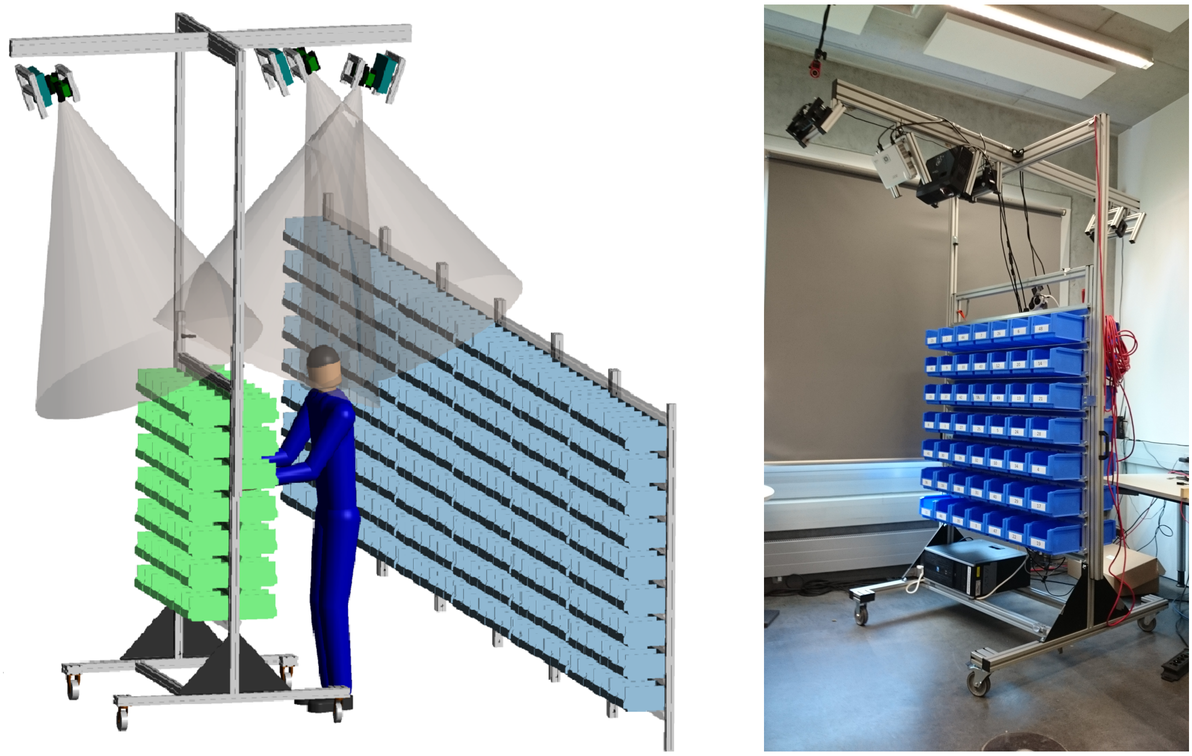

Figure 4.1: The picking cart is with equipped with monitoring sensors as shown here.

4.1.2 Depth Sensor

As a depth sensor the Kinect is used which is a sensor device, that is developed and sold by Microsoft. It was designed to provide extended control possibilities for the Xbox game consoles(360 and One) and Windows applications[1]. For this purpose, it contains a standard RGB camera with a resolution of 640x480 pixels which allows to capture image data and an infrared depth field sensor. The depth field sensor consists of a infrared projector, which projects a special pattern of shapes on the environment. These shapes are unique in every point of the projection. The pattern is recognized by an infrared camera and the device generates a 3D point cloud with a resolution of 640x480 dependent on the size and distance of the recorded shapes. The Kinect also provides an array of four microphones. This arrangement enables the Kinect not only to record, but also to locate the source of the recorded sound. Especially the colour and depth sensors make the Kinect the optimal device for the previously described requirements.

[1]http://www.xbox.com/en-US/kinect

Figure 4.2: The Epson Moverio BT-200, equipped with passive infrared tracker for the motion tracking system.

The first Kinect sensor is used to detect pick events on the warehouse shelves. It is mounted on the outermost point of the cart in order to achieve an optimal coverage an minimize the cases, where the worker covers the shelf or several boxes with his body as seen in Fig. 4.1.

Another Kinect is used to detect interactions with the parts carriers on the cart which are: placing an object into a part bin and correcting errors by picking an object out of a part bin. This sensor is placed orthogonally to the floor above the part bins and monitors the space in front of the boxes. That way the confidence of the pick detection reaches a maximum, which is critical for a properly functioning system.

4.1.3 Head-Mounted Display

The used HMD are the Epson Moverio BT-200 smart-glasses. In Fig. 4.2 they can be seen ready for the study with infrared position markers. They provide two see-through colour displays each with a resolution of 960x540 pixels, which potentially enables 3D visualisation and provides the possibility to build Augmented Reality applications and real time data overlays into the user's usual view. According to Epson, the glasses generate a perceived screen size of 80 inches at 5 meters distance and 320 inches at 20 metres distance. Since the glasses only contain the displays, a camera and additional sensors which includes a compass, gyroscope and accelerometer, an additional device which is connected to the glasses by wire, provides the calculation power, a touchpad buttons for interaction. The glasses run a customized Android operating system in version 4.0.4. Network communication via Wi-Fi and Bluetooth 3.0 is

[2] https://www.naturalpoint.com/optitrack/products/flex-3/

possible. Furthermore the device provides a 1.2 GHz Dual Core CPU and 1 GB RAM and a battery life time of approximately 6 hours in video mode.[3] These glasses fulfil the requirements, which are stated in Section 3.2.

4.1.4 Location Tracking System

The OptiTrack motion tracking system consists of a set of infrared cameras, that are connected to a computer and the associated software motive, that provides extensive functionality, from system calibration to recording, monitoring, organizing and streaming of tracking data. In this special case, 17 cameras track the positions of passive infrared markers in space. Markers can be grouped to rigid bodies, which enables the system not only to deliver the position, but also orientation and look direction of the tracked object. An object is represented by a 3D coordinate as a point in space and the orientation in Euler angles. The objects are tracked with a frame rate of 100 fps and a latency of 0.2 to 10 milliseconds.

For the test scenario of the prototype, the cart and the glasses had to be equipped with markers. The cart had seven markers distributed on a bar above the small load carriers. This wide distribution made it possible, to achieve a high stability of the cart position and angles on two axis, which is enough for this special use case, because all dimensions of freedom of movement of the cart are covered.

A second set of markers that where grouped to a rigid body was attached to the glasses(Fig. 4.2). In this case it was important to use the maximum of seven markers, allowed by the motive tracking software in order to guarantee stable and reliable knowledge about the position and orientation of the glasses and therefore the user's head, even if the trackers are only visible to half of the cameras or completely covered, for example by the user's head and surrounding objects, like the cart or a shelf. Head tracking is far more difficult than tracking the position of the cart, because of a faster moving speed that requires real time tracking and does not tolerate any latencies and an extended degree of freedom in position and rotation around three axis.

A downside of OptiTrack is that its accuracy can be harmed by sunlight or other interference sources, that emit or reflect infrared light.

4.1.5 Projectors

The two projectors, that can be seen in Section 4.1.1, are used for the comparative study. In the compared use case, they provide the data overlay directly projected in the environment and on the part bins. For this purpose, the projector for the environment overlay is a special short

[3] http://www.epson.com/cgi-bin/Store/jsp/Product.do?sku=V11H560020s

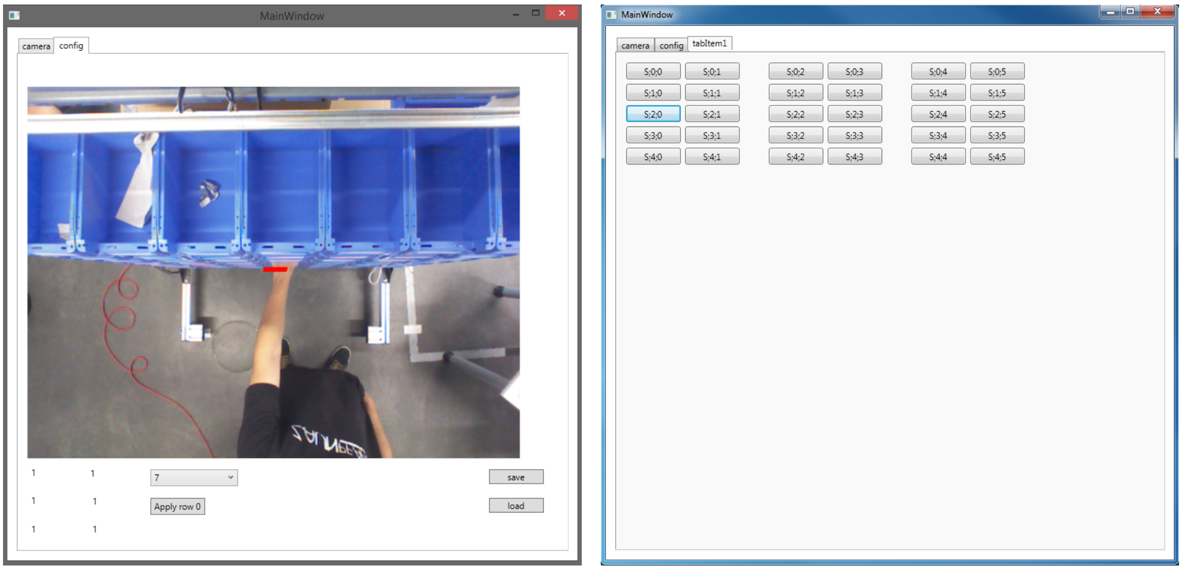

Figure 4.3: A specially built pick detection software provides the necessary information to implement context awareness.

throw projector. The projectors are mounted to the extension on the top of the cart, similar to the Kinects.

4.2 Software Setup

The software being utilized is a mixture of hardware related software(SDKs etc.) that existed before and newly written or adapted software for the special use case. This section describes its functionality and the interactions between the different parts.

4.2.1 Pick Detection for Cart

The Kinect that monitors the space in front of the part bins as mentioned above, is controlled by a self-written pick detection software, that streams pick events via TCP into network. The TCP streaming is necessary to make sure that every pick is recognized by the system. For this purpose the depth picture of the Kinect is used. Every part bin has a trigger box in the space in front of it, similar to a light barrier. Every time something enters one of this areas the software reports a pick into the related bin. New picks are only detected when the object leaves the space and enters it again. In Fig. 4.3 the picture of the Kinect is displayed, while a pick is detected. In order to eliminate inaccuracies of the Kinect depth sensor, the software offers a calibration function. For calibration, the height and width of every part bin row is measured with the depth sensor and a calibration bar which is placed on every row. This procedure allows a flexible design of the cart because it doesn't matter in which height the Kinect is placed, as long as it is in the range of the depth sensor. Another advantage is the flexible number and size of boxes per row and different row sizes, which enables a maximum of freedom.

A big limitation with this method is the strong dependency on a fast WI-FI connection to the HMD in order to provide fast feedback, which could not always be guaranteed. This effect was noticeable on all streaming procedures, but had worse effects on the TCP streams.

4.2.2 Pick Detection for Shelf

In the case of the pick detection for the shelf, a Wizard of Oz [DJA93] approach was introduced in order to achieve the required accuracy and reliability in order to get clear results at the user study. Nevertheless the interface for the Wizard of Oz is integrated into the software that also runs the detection on the cart. This is realistic because both Kinects are mounted to the cart and connected to the same PC. They also use the same TCP stream to transmit the pick events to the augmented reality headset. The prototype should be as close to a final product as possible. Therefore is the Wizard of Oz component embedded into the system, as the actual pick detection software would be. The interface consists of a button for each compartment on the warehouse shelve that where used for the user study as it can be seen in Fig. 4.3.

4.2.3 Motion Tracking Software Motive and NatNet SDK

Motive is an "Optical motion capture software"[4]. It manages the installed infrared cameras and provides several features, that allow the user to calibrate the system, manage and record motion data and stream it to other devices or pieces of software. Therefore, the NatNet SDK allows developers, to use the tracking data in client applications. The Information is streamed via Multicast into network[5]. Motive enables the user to create rigid bodies from up to seven markers, which have a fixed position to each other. The advantage of rigid bodies is, that they can be tracked, even if not every marker of the body is visible to the system. It also allows the system, to calculate and display the orientation of the tracked object, as well as streaming it to other applications.

[4] https://www.naturalpoint.com/optitrack/products/motive/
[5] https://www.naturalpoint.com/optitrack/products/natnet-sdk/

4.2.4 UDP Streaming for Motion Detection

The motion tracking system delivers the position and orientation of rigid bodies in a special format. This set of data is transformed to a more use case related format and streamed to the glasses with an UDP stream in order to prevent data congestions and enable a real time system. This sort of middleware is necessary because it was not possible to integrate the NatNet SDK into the software that runs on the glasses. It additionally provided some useful debug information about the transferred data.

4.2.5 Logic and Visualisation

The biggest piece of software is the component that is responsible for the logic and visualization on the smart-glasses. After an extensive task analysis and some experiments with the Android canvas, the choice fell on the game development ecosystem Unity. Unity provides the opportunity to create 3D models and dynamically move objects in space and modify them. Therefore, one can attach a script to an object, which controls its behaviour, location and appearance.

In the first step a model of the experimental warehouse was generated with Unity whilst the compartments of the shelves where simulated by spheres. In the same way the small part carriers on the cart where modelled. Every sphere that represents a compartment of a shelve was placed on a fixed position, with a fixed size, dependent on its actual position in the real world. The position and orientation of the cart model is controlled by an attached script, that uses the real-time streaming data from the motion tracking system to place the objects in a correct way. In the same way, the position of the camera is controlled. The camera position and orientation follows the parameters of the glasses in the real world. In that way, the user of the glasses has a direct view into the 3D model, overlayed to the view of the real world. In order to provide a correct mapping of the object positions, the position data OptiTrack delivers has to be mapped to the Unity coordinates and the real world measures had to be converted to Unity units. The model is shown in Fig. 4.4.

Additionally to the model, the logic plays a big role for the system. The main part of the software that is running on the glasses is the order manger. This component is responsible for organizing the orders, evaluating the actions of the user and giving appropriate feedback. The order manager holds a list of tasks, which contain information about the shelf compartment to pick from, the box on the cart, where the picked parts should be placed in and the amount of parts that need to be picked. That is basically everything a warehouse manager would have to put into the system in order to keep the work running. The focus here was to keep the manual overhead as simple and small as possible. However the system itself has to store a few more parameters in order to provide an optimal range of support for the worker. These parameters are a list of errors, the previous task and the current target. The guiding software communicates with the pick detection and motion tracking software via streams as described

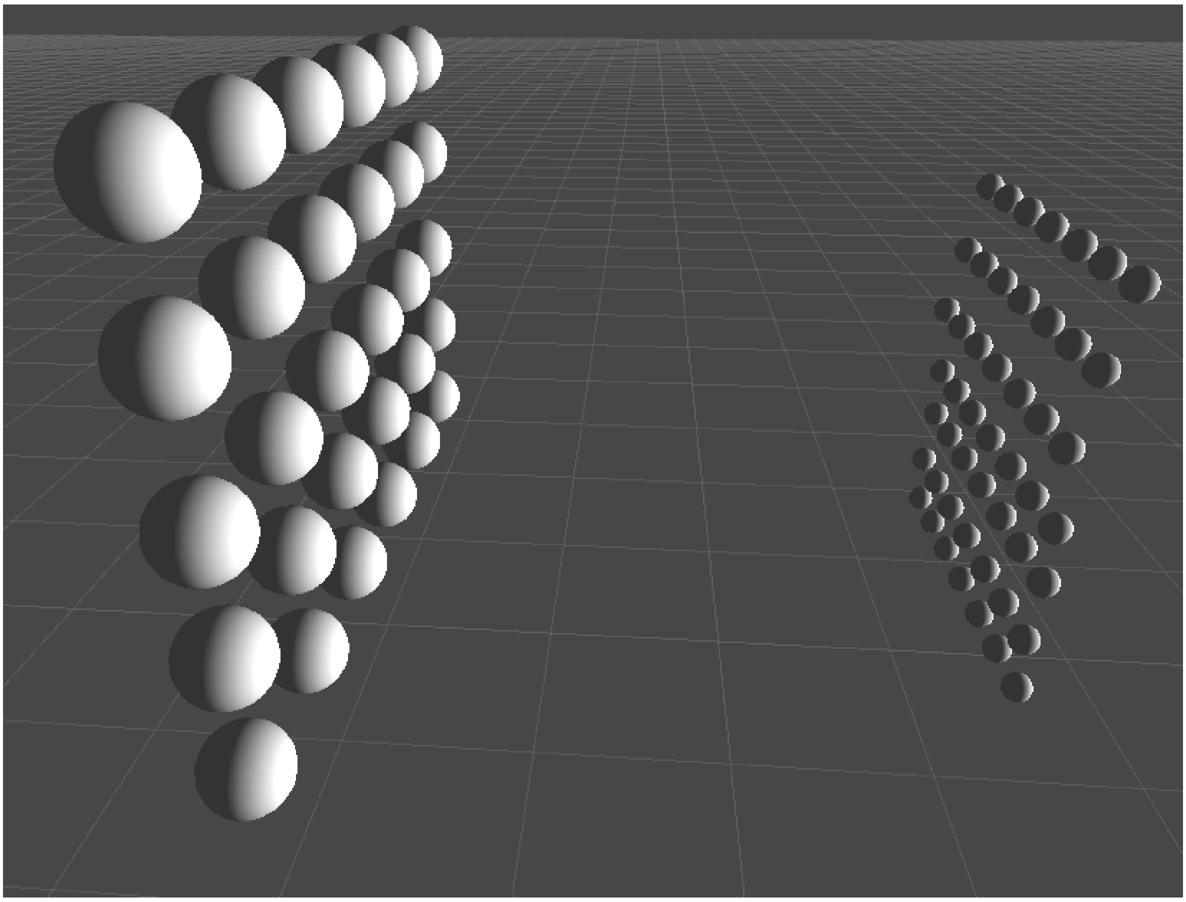

Figure 4.4: The 3D model of the experimental setting was created with Unity and provides options to visualize picking tasks.

above. Whenever a pick event is revived it is pushed on a stack and stays there until it is processed. In order to store the location of the glasses and the cart, static variables were introduced, that store the latest position and get replaced when a package with a newer timestamp arrives. This is necessary because the receivers are running in different threads than the evaluation logic. This also provides robustness against dropouts on the data stream. The received information gets evaluated and processed once per calculated frame.

The workflow, that is described in algorithm 4.1 will give a deeper insight into the functionality of the system. However,it is not an exact representation of the code, but it describes the steps that are processed during the execution of the program. These steps are performed in advance of the rendering of every frame in a just in time preparation.

In order to provide the user a consistent and intuitive handling of the system, the visualization focuses on the important information and avoids excessive display use.

Algorithm 4.1 The workflow of the order manager.

```
if there are tasks left to complete then
    for all triggered objects on shelf or cart do
        if triggered object == current target then
            set the next target as current target
        else if triggered object == last task then
            ignore this pick in order to enable corrections on the last task
        else
            add the triggered object to the list of errors
        end if
    end for
    if nothing was triggered && the current target == null then
        set the next target
    end if
    guide the user to the current target
    if the error list is not empty then
        display error message
        for all errors in the error list do
            highlight the faulty compartment red
            fade from red to black
            if fading is finnished then
                remove the error from the error list
            end if
        end for
    end if
else
    stop the execution
end if
```

The main concept of the visualization is the representation of compartments and boxes as spheres as described in Section 3.2. The colour codes for the spheres are realized just as described in the concept. All other spheres on the cart and the shelf that are not used at the moment are not rendered in order to prevent overloading the user's view with unnecessary information, which could result in dangerous distractions.

The tunnel is also modelled and attached to a point in front of the camera. That way it is always visible in the user's view wherever the glasses are pointed. The other end is fixed at the current target and therefore sticks to the right sphere. Depending on the distance between these two points, the distance between the individual circles is calculated.

A number at the top left corner of the screen tells the picker how many parts to pick. It also gets updated every time the task is exchanged.

5 Evaluation

In order to evaluate the previously described prototype a study was conducted. The following chapter describes the study structure, details and results beginning with the study design. There the modalities and measurements of the study are described, followed by a description of the participants. The compared conditions are presented in a separate section. The section Apparatus describes the study environment and tools, followed by the procedure, where the concrete implementation of the study is described. The measured data is presented and discussed. Finally, the limitations of the prototype that were found during the study are listed.

5.1 Design

The study was a comparative study with a counterbalanced repeated measures design [VC96]. This helps to cancel out advantages and disadvantages that result of the order of the performed tasks. All participants had to perform a given task four times with a different supporting method each time. In order to cancel out possible learning effects or other effects that are triggered by the order of the tasks, a cyclic Balanced Latin Square [Bai08] of the order four was used to arrange the tasks for each participant. Every task was rotated every time and repeated in the pattern that is shown in Table 5.1. This pattern formed four groups that were each repeated four times, so every order was used four times.

The goal of the study was, to get both, subjective and objective feedback and data about the compared methods. All described measurements were performed manually. Time was

Participant	**Task order**			
A	PbVi	PbPa	PbPr	PbVo
B	PbPa	PbPr	PbVo	PbVi
C	PbPr	PbVo	PbVi	PbPa
D	PbVo	PbVi	PbPa	PbPr

Table 5.1: The Study Task Pattern.

measured with a stopwatch and a supervisor counted the errors. The questionnaires that were used are described below.

The measurements that were used during the study consists of objective and subjective measurements. To get an objective sight on the performance of the four methods, the overall time for 16 picking operations with each method was recorded. In order to prevent the participants from accomplishing a low time while making a lot of errors, errors were also logged. The errors were divided into five categories:

- Pick Error

  A pick error occurs, when the participant reaches into a wrong compartment on the shelves, where the parts are picked from and touches a wrong part.

- Filling Error

  Filling errors are when the participant reaches into the wrong box on the cart. Therefore it doesn't matter, whether the parts are placed in that box or not.

- Amount Error

  A wrong amount of parts places into the cart box resulted in an amount error.

- Part Error

  A part error appears when the wrong parts remain in the a box. This happens when a filling error is not corrected.

- Cart Error

  If the cart was not placed on the taped line or on the correct position, e.g. in front of the wrong shelf, a cart error was recorded.

The Errors can be grouped into three categories, dependent on their effect on the result of the study and on their interpretation and implications for later use:

- Short Term Errors

  Short term errors are errors, that are made during the picking process. Dependent on if they get corrected or not, they can also result in a long term error. As long as short term errors are corrected during the pick process, for example due to feedback given by the order picking supporting system, their effects on the entire work process is a slightly add to the task completion time. Short term errors are pick errors and filling errors.

- Long Term Errors

  Long term errors are the result of uncorrected short term errors. In contrast to short term errors, they do not lengthen the task completion time of the picker. Their effect however can result in serious aftereffects, dependent of the receiver of the picked order. For example, if the picker is assembling orders for customers, for example in a mail order warehouse, the consequences can reach from the costs of the correction up to loss of customers. If the picker is part of the supply chain for an assembly line, falsely picked parts can result in downtime of the complete line, which causes immense costs, for worker and machines, that have to be paid for doing nothing. Long term errors are amount errors and part errors.

- Study Restriction Violation

  The cart error is the only violation of study restrictions. It is counted as an error, because it results in a shorter task completion time, than if the conditions of the study were followed.

Additional to the objective measurements, subjective data was collected. To include this point of view into the evaluation of the prototype, impressions and opinions were collected with two questionnaires, which were handed over in German and can both be viewed in the appendix:

- NASA TLX (Appendix A)

  The NASA Task Load Index [HS88] is a short questionnaire that is used to determine how several factors of the task were perceived. Therefore the participant has to asses the mental, physical and temporal demand of the task, the performance, the effort and the frustration level while performing the task. In order to get detailed information, every question can be answered on a Likert scale with 20 answering opportunities on a scale from *sehr niedrig/gescheitert*(very low/failed) to *sehr hoch/perfekt*(very high/perfect).[1]

- concluding questionnaire (Appendix A)

  At the end of the study, the participants got a final questionnaire to fill out, that demanded an overall rating of all four methods, on a five points Likert scale from gefällt mir nicht(I do not like) to gefällt mir sehr(I really like). Free text fields allowed the participants to leave written comments. Three questions with free text fields were asked:

[1] http://humansystems.arc.nasa.gov/groups/TLX/

(a) PbPa **(b)** PbVo **(c)** PbPr **(d)** PbVi

Figure 5.1: Participants are performing the picking tasks with each of the compared methods.

5.2 Participants

16 people participated in the comparative study. Four of them were female, the remaining 12 were male. The youngest participant was 20 years old, the oldest was 43, their average age was 24.8 and the median age was 24 *(mean (M) = 24.81, standard error (SE) = 1.39, standard deviation (SD) = 5.58)*. In an introductory questionnaire the participants were asked for their profession. All of them were studying or employed at the university. Twelve of them were students, two postgraduates, one research assistant and a secretary.

5.3 Conditions

Four methods were compared in this study, which are PbPa, PbVo, PbVi and Pick-by-Projection (PbPr). The differences and properties of these methods are described in the following.

- Paper-Picking (PbPa) (Fig. 5.1a)

  For this Method, the participants got handed a paper list (Appendix A) where the compartment designation, the box designation and the amount to pick were registered together with some additional information, that were not relevant for the task. The participants were not allowed to use any additional aids. Additionally they were instructed to work of the list in the given order to keep the travelling distances, that were defined above.

- Pick-by-Voice (PbVo) (Fig. 5.1b)

  The PbVo method was realized by an Wizard of Oz approach. Therefore, the wizard rad the tasks to the participants. They could control the wizard with three predefined voice commands. It was possible to switch the task back and forth with the commands "previous" and "next". In order to hear the instructions again, it was possible to let them repeat with "repeat". The commands were presented in the form "pick X from Y and put them in Z". The task order was again fixed, as above.

- Pick-by-Projection (PbPr) (Fig. 5.1c)

  PbPr is a method, where the instructions are visually communicated via projections from two cart mounted projectors. One of the projectors was pointed to the boxes on the cart and the other illuminates the compartments on the shelf. The projection was also controlled by the motion tracking system and therefore was able to highlight the correct compartment from every position. The target compartment was highlighted in green colour and outfitted with a number, that constitutes the amount of parts to pick. After a correct pick, the box on the cart, where the pieces should be placed was highlighted. Feedback in case of errors was given. This was realized by highlighting the falsely picked compartments and boxes red. In order to recognize picks, this system used the same appliance as the prototype that is evaluated in the study and described in Chapter 4. It was only possible to proceed the task, if the participant reached into the correct box. The PbPr system was developed in collaboration with the prototype presented in this thesis.

- Pick-by-Vision (PbVi) (Fig. 5.1d)

  The PbVi approach was realized by the prototype, that was described above and was built in the context of this thesis. I order to give the participants time to get used to the technology and the hardware, a short introduction was given.

5.4 Apparatus

For all tasks and methods, the same picking cart was used. A detailed description and pictures can be found in Section 4.1.1. The methods were as described above PbPa, PbVo, PbPr and PbVi.

The environment where the study took place was a students working room at the university of Stuttgart. There, three shelves, with 5x2 compartments each, which makes a total of 30 compartments, were aligned to represent a shelf row in a warehouse. A line of tape parallel to the shelves showed the participants where the cart should be placed. Additionally a position in front of each shelf was marked. These precautions were made in order to get the best and most realistic result for the projection and to keep the travelling distances for the cart equal. The participants were instructed to keep the cart on the line and place it on the prepared positions. It was strictly required to use the given positions, in order to obtain comparable time measurements.

In order to correctly represent the realistic searching time for every task, the designations on the compartments and the cart boxes were not ordered and randomly addressed. The parts that the participants had to pick were Lego pieces in different colours and sizes. The text on the compartment label was not linked to its content.

amount of parts	amount of occurence
1	3
2	4
3	5
4	2
5	2

Table 5.2: The distribution of parts to pick.

All instructions were given in a numerical form and therefore language independent except for the PbVo commands. In order to offer equal conditions for every participant, they were asked if they want to use the PbVo system in German or in English.

5.5 Procedure

At the beginning of the study, the participants got a short introduction, where the topic order picking was described. All error types were explained and it was made clear that these instructions are valid for all of the following tasks. After the introduction a small introductory questionnaire(Appendix A) was filled in by the participants where age, gender and profession were recorded and the participant was asked to sign a consent form(Appendix A). After that procedure the tasks started. Subsequent to every task, the participant had to fill in a NASA TLX and after the final task, the previously described concluding questionnaire had to be filled in. The participants were compensated with 5 Euros.

The task was based on real warehouse activities in order to get a setup that is as close to reality as possible. Every run was composed from 16 picks. While every pick instruction consisted of a shelf compartment where the participant should take the parts from a box on the cart where the parts should be placed and an amount of parts to pick. Because the single runs should be comparable, parameters that stayed the same for every task were defined to keep the tasks as similar as possible and still cancelling out learning effects. These Parameters were the amount of parts to pick and the cart movements in distance and occurrence. The distribution of amounts of parts are illustrated in Table 5.2. It was important to define this value because a wrong amount of parts is counted as an error and the level of difficulty has to be the same for every task.

The cart movements also needed to be unified in order to achieve comparability of the tasks. Therefore three positions for the cart were defined and the participants were instructed to place the cart in front of the shelf where they want to pick a part. The distances are defined as the number of stations that had to be passed: zero means that the cart stays where it is,

distance	amount of occurrence
0	10
1	4
2	2

Table 5.3: The Moving Distances and their Occurrences.

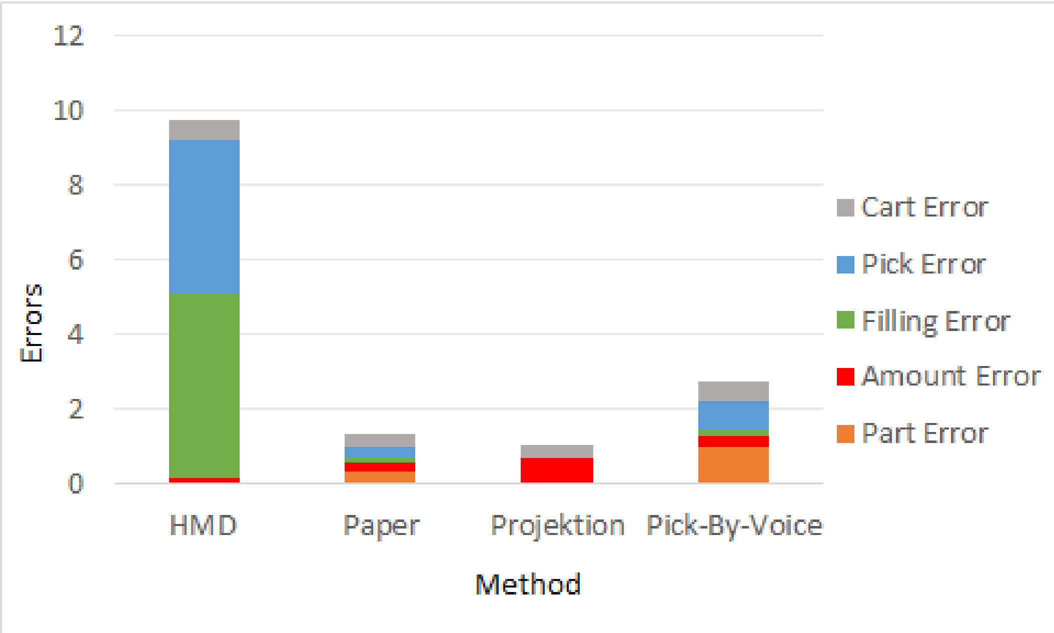

Figure 5.2: The errors for each method, grouped by error type, are shown in this chart.

one means it moves one position to the left or right. It applies in the same way to the two. Table 5.3 shows the distribution of the cart movements.

5.6 Results

The data that was collected during the previously described study is presented in the following. As mentioned above, objective data about the performance of the participants while using the different methods was collected as well as their subjective impressions while using the systems. At first the objective results are described. All the presented numbers and values refer to the whole study task, which consists of 16 picking actions for each method.

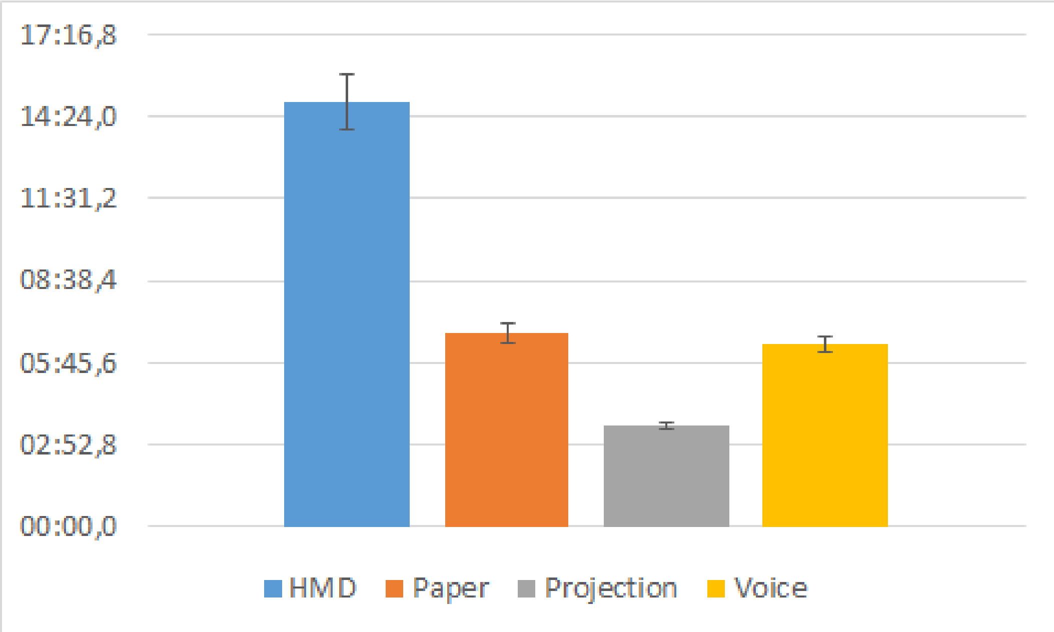

Figure 5.3: This chart visualizes the median task completion times and the related standard error.

The error rate was measured and recorded in order to make statements about the performance of the methods. Fig. 5.2 shows the mean errors that were made during performing the tasks. It appears that the participants made far more errors with the HMD method *(M = 9.75, SE = 1.52, SD = 6.07)* than with the PbVo method *(M = 2.75, SE = 0.64, SD = 2.57)*, PbPa *(M = 1.13, SE = 0.30, SD = 1.20)* and the Projection *(M = 1.00, SE = 0.81, SD = 3.25)*, which is the clear winner in the overall error rating. An analysis of the present data with Mauchly's Test of Sphericity shows, that the sphericity assumption is violated, for which reason the Greenhouse-Geisser correction was applied. A pairwise comparison shows the significance of the differences from the PbVi approach compared to all other methods, whilst the differences between the other methods does not show significant results at all. But if we take a closer look on the composition of these numbers, it stands out, that the errors for the HMD are mainly consisting of short term errors, that are pick errors and filling errors *(M = 9.06, SE = 1.44, SD = 5.76)*, while the amount of the long term errors, which are amount errors and part errors is very low *(M = 0.13, SE = 0.09, SD = 0.34)*, whilst PbVo has a mean of 1.25 long term errors *(M = 1.25, SE = 0.32, SD = 1.29)*, followed by projection with a mean of 0.69 errors *(M = 0.69, SE = 0.69, SD = 2.75)* and PbPa *(M = 0.56, SE = 0.22, SD = 0.89)*. The errors are intentionally not weighted, because weights would depend on the special use case.

Another big point while rating order picking methods is the task completion time, which was measured in minutes. The task completion time shows, what the error chart predicted: the high amount of short time errors at the HMD method resulted in a higher task completion

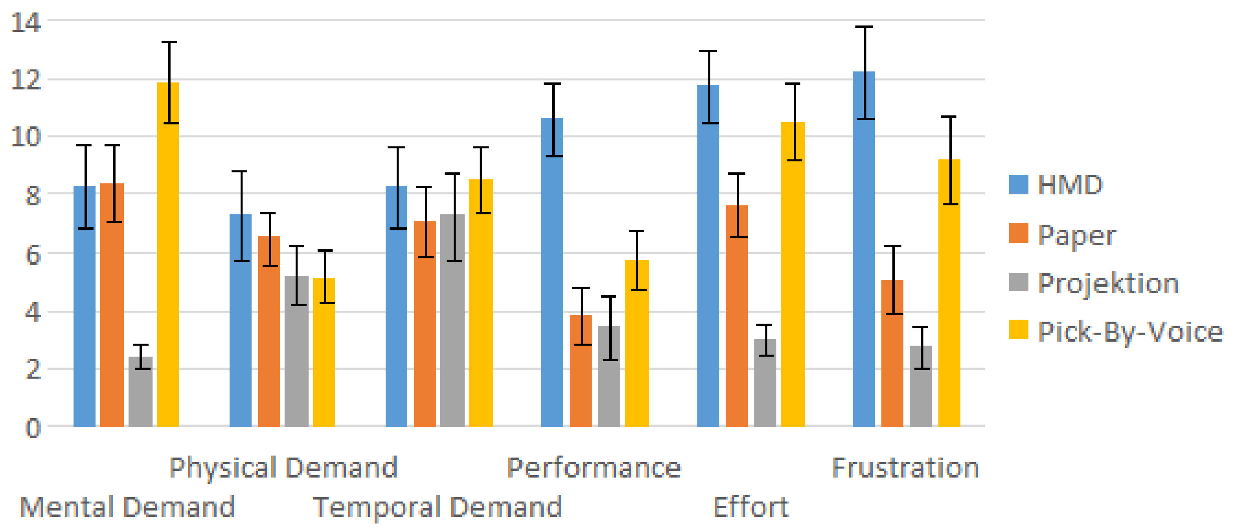

Figure 5.4: An overview over the NASA Task Load Index results provides insight to the users perception of the methods.

time *(M = 15:18.5, SE = 00:58.4, SD = 03:53.8)*, while the PbPa method *(M = 07:06.3, SE = 00:21.2, SD = 01:26.6)* and PbVo *(M = 06:48.3, SE = 00:16.8, SD = 01:07.2)* are close together. PbPr *(M = 03:33.1, SE = 00:05.7, SD = 00:22.7)* showed much faster results , with a mean time of 03:33.1, as it can be seen in Fig. 5.3. Again, Mauchly's Test of Sphericity shows, that the sphericity assumption is violated and the Greenhouse-Geisser correction is used. This time a pairwise comparison shows the significance of the differences between the task completion times of PbVo and all the other methods.

Additional to the objective measurements, questionnaires were introduced, in order to gather subjective impressions of the user. This helps to improve the usability and supports a user centred development. The results from the NASA TLX are displayed in Fig. 5.4. The overall TLX score which sums up the individual scores of all categories shows, that the least demanding method was the projection method *(M = 24.13, SE = 3.78, SD = 15.10)*, followed by the PbPa method *(M = 38.50, SE = 5.29, SD = 21.15)*, PbVo *(M = 51.00, SE = 5.21, SD = 20.83)* and HMD picking *(M = 58.44, SE = 6.32, SD = 25.30)*.

In the questionnaires that the participants were asked to fill out at the end of the study, they were asked to rate all methods on a scale from 1, which means I do not like it to 5, which means I like it. The most popular method according to this questionnaire was the projection method with a median rating of 5.0 *(M = 4.75, SE = 0.14, SD = 0.58)*. Second was the PbPa method with a median rating of 3.0 *(M = 3.00, SE = 0.22, SD = 0.89)*. HMD pickingwith a median of 2.o *(M = 2.19, SE = 0.34, SD = 1.38)* was voted slightly higher than PbVo with a median of 1.5 *(M = 1.13, SE = 0.29, SD = 1.15)*.

The question, what method the participant would like to use for a whole work day was answered 14 times with the projection method , one time with PbPa method and one participant does not want to use any of the presented method for a whole work day. Two participants added a comment, that they would want to use the HMD method, if it is going to be improved.

Regarding the question, what the participants want to see improved, many participants mentioned a smoother representation on the HMD, faster and more accurate motion and position tracking and faster feedback. Additionally some participants mentioned, that the Epson Moverio BT-200 is heavy and uncomfortable.

5.7 Discussion

The result of this experiment show that it is strongly dependent on the use-case, which method is the best.

It appears that at the current state of development, the projection method is the best pick, in regard to overall errors, task completion time and overall usability. But with consideration for the different error types the HMD method shows the best performance. Dependent on the area of application, the impact of the different error types can be estimated and an individual weighting can be created. For example environments, where accurate delivery is critical, like production lines, the longer task completion time can be a worthy trade-off for the lower long-term error rate.

According to the subjective measurements the projection method is the most favoured method among the participants, which suggests the use of this method from a user centred point of view. However it is important to keep in mind, that the state that was evaluated in the study is still a prototype state. This allows to adapt to the results of this experiment and adjust the future steps of development. Even if the usability rating of the HMD method was not convincing, a lower mental demand than with traditional methods was achieved. Additionally comments about the poor wearing comfort of the used device are encouraging to improve the system in order to score better results.

There is also room for improvement according to the systems performance. Some participants noted a delay of the representation when head movements are getting too fast, as well as delayed feedback, which had confusing effects on the participant. Cancelling this out would mean providing a much better usability and functionality and eliminate the insecurities of the user. However performance problems are caused by properties of the hardware and software of the system and the surrounding infrastructure, that can easily be improved. These improvements are worth a try because participants mentioned that they are curious about using the system, if its performance would increase.

Another finding that can however not be supported by the present data is that the performance of the participants was strongly dependent on their body size, which caused a wide spread of the numbers of errors. This is presumably caused by a poorly calibrated model. Further investigation on this topic can help to eliminate this disadvantage.

The learning effect with the system is also an unproven hypothesis, based on observations, because of a lack of detailed data. Nevertheless, the participants showed insecurity during the first few picks that occurred in spite of the introduction, the participants were given. The time the participants needed per pick decreased towards the end of the task and they seemed to get more familiar with the representation and gained confidence with every successful pick.

5.8 Limitations

The prototype has some limitations that had an impact on the usability and therefore also influenced the results of the study. However this points provide a starting point for improving the prototype.

The first and most obvious limitation of the current system is the Wizard of Oz approach that is used during the study. A reliable and fast way to detect picks for the shelf in front of the cart has to be found and implemented in order to increase the automation.

Another limitation that was discovered during the study was the calibration of the 3D model. In order to improve the user experience, a way to calibrate the used model for the visual representation has to be developed. As a first step it is important to discover if there is a perfect calibration for every room, that fits all users, no matter what size they are and how big the distance between theHMD and the users eyes is. If there is one, a simple way to generate models with little previous knowledge for example out of floor plans. But if different users need different calibrations, a fast and reliable way to calibrate the model has to be found.

The poor network speed and reliability resulted in bad performance of the whole system. It was in some cases not possible to transfer the necessary data fast enough from the motion tracking and pick detection systems to the glasses to provide a smooth real time experience to the user.

6 Conclusion and Future Work

This chapter gives a brief overview over this thesis and a summary of the most important points. It finishes with an outlook and suggestions for future work.

6.1 Summary

In this thesis the concept for an HMD based order picking supporting system was designed, a prototype was built and evaluated.

The concept describes an ideal product that serves best for the order picking use case. Every design decision is supported by the reason why it was designed that way. In order to provide a use case orientated solution, the warehouse use case was described and analysed in a first step. Based on this description, requirements for an optimal system where defined. The chapter Concept and Design further explains, why HMDs are suitable for this area of application. In the following, the Task representation with augmented reality spheres that represent compartments and colour codes which displays the state of the compartment, are described in detail. In order to guide the user, an adapted version of the attention funnel was presented. The environment and infrastructure, including the picking cart was explained subsequent to this section. In a last step, the necessity of instant feedback about the work process and its possible realization is discussed.

In order to evaluate the previously described concept, a prototype was built. It's technical implementation was described in chapter 4. In a first step the hardware setup, including the design of the cart, the Kinect as motion sensing device, the choice of HMD and the infrared positioning system are described, followed by the software setup. The first described component is the pick detection, that is responsible for delivering information about task progress and helps to validate the users actions. A combination of existing software and additional self-written code is responsible for streaming location data to the glasses. The last section focuses on the implementation of the piece of software that is running on the glasses. It is responsible for managing the tasks and over all progress and the visualisation for the user.

The prototype was evaluated in a comparative study where 16 participants where asked to perform 16 picks with four different supporting methods, that where PbPa, PbVo, PbPr and PbVi. The study task, the methods in detail, all restrictions and the environment, the objective and subjective measurements and the execution of the study where explained in detail in chapter 5. The experiment revealed that it is dependent on the use case, which method serves best however it helped to discover the starting points for further development.

6.2 Future Work

The goal of this thesis was to develop and evaluate a prototype for an order picking supporting system using HMDs and a special picking cart. Since the prototype is the first of its kind, the whole concept was based on related work, that covers similar topics or parts of the topic of this thesis. It was necessary to combine knowledge and experiences from other projects in order to form a starting point for a new kind of system. This prototype can be the basis for a whole row of similar systems, that take advantage of the findings in this thesis. In order to enlarge the benefits of the result of this experiment, they can be used to enhance the prototype and trigger research on the unclear points.

An evaluation that focuses on the current representation could help to improve the prototype. Therefore the use of the tunnel should be investigated. Following studies can discover, whether the straight line, the tunnel follows in the current system is the best solution or if this visual hint can be improved in any other way. Another thing that needs to be evaluated according to the tunnel is its behaviour, when the target lies behind the user.

As mentioned in Section 5.7, during the experiment, a learning effect was noticed, but there is no verification of this hypothesis in form of measured data. Therefore it would be interesting, to evaluate the learning curve of first-time users and find out, how long the performance of the participants is getting better and what level it can reach, in order to compare the HMD picking method against other methods with a shorter habituation time. A long-term study could bring the desired results.

Another topic, where some investigation needs to be done is the feedback for amount errors. Currently there is no such feedback if the user takes a wrong amount of parts from the shelf and places it into the cart. One could think of an approach including scales on the cart, which would limit the costs of such an upgrade to the amount of boxes on the cart.

A Appendix

The appendix contains all documents and questionnaires, that where used in the study. They are presented in the same order as they where presented to the participants. This chapter also contains the used paper picking list.

University of Stuttgart
Germany

Human Computer Interaction Group (MCI), VIS

Prof. Dr. Albrecht Schmidt

Einverständniserklärung

BESCHREIBUNG: Sie sind hiermit dazu eingeladen an der **Studie** über die **Benutzbarkeit von Assistenzsystemen bei der Kommissionierung** teilzunehmen.

ZEITAUFWAND: Ihre Teilnahme dauert ungefähr **60 Minuten**.

DATENERFASSUNG: Für die Evaluation des Systems werden Zeiten und die Fehlerrate gemessen. Zusätzlich werden während der Studie Fragebögen ausgefüllt. In dieser Studie wird das zu testende System geprüft - nicht die Teilnehmer!
Bilder:

- ▢ Ich bin damit einverstanden, dass Bilder von mir während der Studie gemacht werden.
- ▢ Ich bin **nicht** einverstanden, dass Bilder von mir während der Studie gemacht werden.

Videos:

- ▢ Ich bin damit einverstanden, dass Videoaufnahmen von dem Arbeitsprozess während der Studie gemacht werden.
- ▢ Ich bin **nicht** einverstanden, dass Videoaufnahmen von dem Arbeitsprozess während der Studie gemacht werden.

RISIKEN UND NUTZEN: Mit dieser Studie sind keine Risiken verbunden. Die gesammelten Daten werden sicher und anonym gespeichert. Die gesammelten Daten werden aggregiert und anonymisiert in einem wissenschaftlichen Bericht veröffentlicht. Ihre Privatsphäre bleibt erhalten. Die Teilnahme an der Studie hat keinen Einfluss auf Ihr Arbeitsverhältnis. Die Daten werden nur in aggregierter Form und anonymisiert an Ihren Arbeitgeber weiter gegeben.

RECHTE DER TEILNEHMER: Wenn Sie dieses Formular gelesen und sich dazu entschieden haben an dieser Studie teilzunehmen, ist diese Teilnahme weiterhin **freiwillig** und Sie haben das Recht, jederzeit Ihre Zustimmung zurückzuziehen und Ihre Teilnahme jederzeit abzubrechen. Sie haben das Recht spezifische Fragen nicht zu beantworten. Die Ergebnisse dieser Forschungsstudie werden möglicherweise bei wissenschaftlichen Konferenzen oder Expertentreffen präsentiert oder in wissenschaftlichen Zeitschriften veröffentlicht.

KONTAKT INFORMATIONEN: Bei Fragen, Bedenken oder Beschwerden über diese Forschung, die Abläufe, Risiken und Nutzen, kontaktieren Sie bitte folgende Personen:
Markus Funk (markus.funk@vis.uni-stuttgart.de)
Albrecht Schmidt (albrecht.schmidt@vis.uni-stuttgart.de)

Mit der Unterzeichnung dieses Dokuments stimme ich den oben genannten Bedingungen zu.

Name: ________________________ *Unterschrift, Datum:* ______________________

Figure A.1: The consent form that was used in the study.

Eingangsfragebogen:

Teilnehmer Nummer: ___________________________

Alter: ____

Geschlecht: ______________________________________

Beruf: ____

Figure A.2: The introducing questionnaire queries demographic information.

Proband ID: ______________ Bedingung: _______________ Datum: ___________________

Mentaler Aufwand Wie geistig anspruchsvoll war die Aufgabe?

Sehr niedrig Sehr hoch

Körperlicher Aufwand Wie anstrengend war die Aufgabe?

Sehr niedrig Sehr hoch

Zeitlicher Aufwand Wie hastig oder gehetzt war das gesetzte Tempo der Aufgabe?

Sehr niedrig Sehr hoch

Performance Wie erfolgreich waren Sie im Lösen der Aufgabe?

Perfekt Gescheitert

Aufwand Wie sehr mussten sie sich anstrengen, um Ihre Leistung zu erreichen?

Sehr niedrig Sehr hoch

Frustration Wie unsicher, entmutigt, irritiert, gestresst, und verärgert waren Sie?

Sehr niedrig Sehr hoch

Figure A.3: The NASA TLX was used in german language.

Bewerten Sie die angewandten Methoden auf der folgenden Skala:

Paperpicking	*gefällt mir nicht*	o	o	o	o	o	*gefällt mir sehr*
Pick-By-Voice	*gefällt mir nicht*	o	o	o	o	o	*gefällt mir sehr*
Head Mounted Display	*gefällt mir nicht*	o	o	o	o	o	*gefällt mir sehr*
Projektionen	*gefällt mir nicht*	o	o	o	o	o	*gefällt mir sehr*

Welche Methode würden Sie 8 Stunden am Tag (also ein Arbeitstag) benutzen wollen?

Was könnte man verbessern bzw. was würden Sie sich wünschen?

Was hat Ihnen weniger gefallen?

Figure A.4: In a final questionnaire the participants where asked to rate all methods and leave comments.

Pos	Article Number	Count	From	To	Price	Checked
1	757-567-346	3	16-22	43		Gelbes 2x2
2	999-234-313	1	91-22	30		Gelbes 2x1
3	640-324-346	1	33-42	46		Rotes 4x2
4	443-563-564	2	82-21	38		Rotes 2x2
5	172-391-134	4	44-22	42		Grünes 2x1
6	743-849-173	5	13-22	36		Weißes 3x2
7	920-876-776	2	39-32	9		Weißes 2x2
8	774-143-332	3	23-32	3		Hellblaues 4x2
9	867-174-651	3	59-42	10		Rotes 3x2
10	593-739-096	2	31-42	28		Orange 2x2
11	632-649-756	3	16-42	18		Grünes 3x2
12	113-413-073	5	24-22	32		Schwarzes 2x1
13	467-432-097	4	25-21	34		HellGrün 4x2
14	967-312-532	3	00-42	37		Grünes 4x2
15	777-654-333	1	02-21	11		Blaues 4x2
16	423-563-109	2	13-42	20		Weißes 4x2

Figure A.5: The paper picking list, that was used during the study.

Bibliography

[A^+97] R. T. Azuma, et al. A survey of augmented reality. *Presence*, 6(4):355–385, 1997. (Cited on pages 19 and 20)

[Bai08] R. Bailey. *Design of comparative experiments*, volume 25. Cambridge University Press Cambridge, 2008. (Cited on page 41)

[Bau13] H. Baumann. *Order picking supported by mobile computing*. Ph.D. thesis, University of Bremen, 2013. (Cited on page 16)

[BKS^+10] A. Brush, A. K. Karlson, J. Scott, R. Sarin, A. Jacobs, B. Bond, O. Murillo, G. Hunt, M. Sinclair, K. Hammil, et al. User experiences with activity-based navigation on mobile devices. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*, pp. 73–82. ACM, 2010. (Cited on page 20)

[BLS05] L. Bonanni, C.-H. Lee, T. Selker. Attention-based design of augmented reality interfaces. In *CHI'05 extended abstracts on Human factors in computing systems*, pp. 1228–1231. ACM, 2005. (Cited on page 21)

[BR03] P. Baudisch, R. Rosenholtz. Halo: a technique for visualizing off-screen objects. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 481–488. ACM, 2003. (Cited on page 21)

[BSI^+11] H. Baumann, T. Starner, H. Iben, A. Lewandowski, P. Zschaler. Evaluation of graphical user-interfaces for order picking using head-mounted displays. In *Proceedings of the 13th international conference on multimodal interfaces*, pp. 377–384. ACM, 2011. (Cited on page 17)

[BTOX06] F. Biocca, A. Tang, C. Owen, F. Xiao. Attention funnel: omnidirectional 3D cursor for mobile augmented reality platforms. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pp. 1115–1122. ACM, 2006. (Cited on page 21)

[CM92] T. P. Caudell, D. W. Mizell. Augmented reality: An application of heads-up display technology to manual manufacturing processes. In *System Sciences, 1992. Proceedings of the Twenty-Fifth Hawaii International Conference on*, volume 2, pp. 659–669. IEEE, 1992. (Cited on page 12)

[DJA93] N. Dahlbäck, A. Jönsson, L. Ahrenberg. Wizard of Oz studies: why and how. In *Proceedings of the 1st international conference on Intelligent user interfaces*, pp. 193–200. ACM, 1993. (Cited on page 36)

[FBP+14] M. Funk, R. Boldt, B. Pfleging, M. Pfeiffer, N. Henze, A. Schmidt. Representing indoor location of objects on wearable computers with head-mounted displays. In *Proceedings of the 5th Augmented Human International Conference*, p. 18. ACM, 2014. (Cited on page 20)

[FKS14] M. Funk, O. Korn, A. Schmidt. Assisitive augmentation at the manual assembly workplace using in-situ projection. In *Proceeding of the chi workshop on assistive augmentation*. 2014. (Cited on page 11)

[GBGI08] S. Gustafson, P. Baudisch, C. Gutwin, P. Irani. Wedge: clutter-free visualization of off-screen locations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 787–796. ACM, 2008. (Cited on page 21)

[GRX+14] A. Guo, S. Raghu, X. Xie, S. Ismail, X. Luo, J. Simoneau, S. Gilliland, H. Baumann, C. Southern, T. Starner. A comparison of order picking assisted by head-up display (HUD), cart-mounted display (CMD), light, and paper pick list. In *Proceedings of the 2014 ACM International Symposium on Wearable Computers*, pp. 71–78. ACM, 2014. (Cited on pages 11 and 17)

[HG07] H. Hua, C. Gao. Design of a bright polarized head-mounted projection display. *Applied optics*, 46(14):2600–2610, 2007. (Cited on page 18)

[HS88] S. G. Hart, L. E. Staveland. Development of NASA- (Task Load Index): Results of empirical and theoretical research. *Advances in psychology*, 52:139–183, 1988. (Cited on page 43)

[JLP06] G.-y. Jin, X.-y. Lu, M.-S. Park. An indoor localization mechanism using active RFID tag. In *Sensor Networks, Ubiquitous, and Trustworthy Computing, 2006. IEEE International Conference on*, volume 1, pp. 4–pp. IEEE, 2006. (Cited on page 21)

[KTC09] T.-W. Kan, C.-H. Teng, W.-S. Chou. Applying QR code in augmented reality applications. In *Proceedings of the 8th International Conference on Virtual Reality Continuum and its Applications in Industry*, pp. 253–257. ACM, 2009. (Cited on page 22)

[LCTM12] X. Li, I. Y.-H. Chen, S. Thomas, B. A. MacDonald. Using Kinect for monitoring warehouse order picking operations. In *Proceedings of Australasian Conference on Robotics and Automation*. 2012. (Cited on pages 15 and 17)

[LGSK10] M. Löchtefeld, S. Gehring, J. Schöning, A. Krüger. PINwI: pedestrian indoor navigation without infrastructure. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, pp. 731–734. ACM, 2010. (Cited on page 20)

[Lol03] A. Lolling. *Analyse der menschlichen Zuverlässigkeit bei Kommissioniertätigkeiten*. Shaker, 2003. (Cited on page 17)

[LWNS07] C.-H. Lim, Y. Wan, B.-P. Ng, C. See. A real-time indoor WiFi localization system utilizing smart antennas. *Consumer Electronics, IEEE Transactions on*, 53(2):618–622, 2007. (Cited on page 22)

[MSS11] A. Mulloni, H. Seichter, D. Schmalstieg. Handheld augmented reality indoor navigation with activity-based instructions. In *Proceedings of the 13th international conference on human computer interaction with mobile devices and services*, pp. 211–220. ACM, 2011. (Cited on page 20)

[RGSK09] R. Reif, W. A. Günthner, B. Schwerdtfeger, G. Klinker. Pick-by-Vision comes on age: evaluation of an augmented reality supported picking system in a real storage environment. In *Proceedings of the 6th International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa*, pp. 23–31. ACM, 2009. (Cited on pages 16, 17 and 21)

[SBG99] A. Schmidt, M. Beigl, H.-W. Gellersen. There is more to context than location. *Computers & Graphics*, 23(6):893–901, 1999. (Cited on page 13)

[SK08] B. Schwerdtfeger, G. Klinker. Supporting order picking with augmented reality. In *Proceedings of the 7th IEEE/ACM international Symposium on Mixed and Augmented Reality*, pp. 91–94. IEEE Computer Society, 2008. (Cited on page 20)

[SRG+09] B. Schwerdtfeger, R. Reif, W. A. Gunthner, G. Klinker, D. Hamacher, L. Schega, I. Bockelmann, F. Doil, J. Tumler. Pick-by-Vision: A first stress test. In *Mixed and Augmented Reality, 2009. ISMAR 2009. 8th IEEE International Symposium on*, pp. 115–124. IEEE, 2009. (Cited on pages 16, 17 and 21)

[SS10] U. Steinhoff, B. Schiele. Dead reckoning from the pocket-an experimental study. In *Pervasive Computing and Communications (PerCom), 2010 IEEE International Conference on*, pp. 162–170. IEEE, 2010. (Cited on page 22)

[Sut68] I. E. Sutherland. A head-mounted three dimensional display. In *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*, pp. 757–764. ACM, 1968. (Cited on page 18)

[Tom10] J. A. Tompkins. *Facilities planning*. John Wiley & Sons, 2010. (Cited on page 15)

[TSP12] K. Taljanovic, A. Salihbegovic, A. Pandzo. A Fast Manual Picking of Small Parts High Volume Orders. *Journal of Communication and Computer*, 9:1097–1103, 2012. (Cited on page 17)

[VC96] E. Vonesh, V. M. Chinchilli. *Linear and nonlinear models for the analysis of repeated measurements*. CRC press, 1996. (Cited on page 41)

[VHCW13] J. Verriet, R. Hamberg, J. Caarls, B. van Wijngaarden. Warehouse Simulation Through Model Configuration. In *ECMS*, pp. 629–635. 2013. (Cited on page 15)

[WBS+10] K. A. Weaver, H. Baumann, T. Starner, H. Iben, M. Lawo. An empirical task analysis of warehouse order picking using head-mounted displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1695–1704. ACM, 2010. (Cited on page 16)

All links were last followed on October 13, 2014.

Declaration

I hereby declare that the work presented in this thesis is entirely my own and that I did not use any other sources and references than the listed ones. I have marked all direct or indirect statements from other sources contained therein as quotations. Neither this work nor significant parts of it were part of another examination procedure. I have not published this work in whole or in part before. The electronic copy is consistent with all submitted copies.

place, date, signature