Visualization Research Center (VISUS)

University of Stuttgart
Allmandring 19
D–70569 Stuttgart

Masterarbeit

Exploring the Extension of AR-Visualization and Interaction for Large Physical Spaces

Britta Schulz

Course of Study:	Informatik
Examiner:	Prof. Dr. Michael Sedlmair
Supervisors:	Jonas Haischt, M.Sc. Nina Dörr, M.Sc.
Commenced:	April 11, 2023
Completed:	October 11, 2023

Acknowledgements

I would like to express my sincere gratitude to my supervisors, for their invaluable guidance and support throughout this project. Additionally, I want to thank all my co-workers who consistently provided me with thoughtful feedback and ideas and were readily available to motivate me when I was struggling. Moreover, my heartfelt appreciation goes out to the participants including pilots who took part in the user study and showed much interest in my work. Finally, I want to thank the research coordination and operation managers in ARENA2036 to take their time to answer all my questions during the problem identification, were willing to participate in my study and promote it as well as sharing my results.

Abstract

While present-day Augmented Reality (AR) often focuses on the interaction and superimposition of clearly reachable objects, navigating vast, information-rich environments presents unique challenges, especially when virtual objects are distant, overlapping, or in complex surroundings. Crafting a solution that minimizes visual clutter while conveying essential information has yet to be fully realized. This thesis explores the application of AR for visualizing information-dense large areas like production halls. It addresses the limitations of existing AR solutions and proposes an approach using a predefined viewpoint as a vantage point. The implementation takes place in a factory shop floor, utilizing AR Head-Mounted Displays (HMD) for interactive data superimposition. Challenges include suitable tracking techniques, data visualization in large physical dimensions and user interaction with partly occluded information of various distant locations. A user study evaluates suiting interaction techniques and the overall utility. On the whole, this research aims to enhance visualization and interaction in industrial and architectural contexts. Results show that the system is perceived as helpful for its target application and that an overview 2D menu, due to its reliability is desired as an omnipresent option for interaction. For a better user experience, head gaze or eye tracking is preferred. Head Gaze enables more precise pointing while eye gaze tends to lead to a better experience, immersion and less physical effort. The widespread far hand ray interaction performed worst, because the exact movement was difficult to learn.

Kurzfassung

Während sich Augmented Reality (AR) heute oft auf die Interaktion und Überlagerung von klar erreichbaren Objekten konzentriert, stellt die Navigation in riesigen, informationsreichen Umgebungen eine besondere Herausforderung dar, vor allem wenn die virtuellen Objekte weit entfernt sind, sich überlappen oder sich in einer komplexen Umgebung befinden. Die Entwicklung einer Lösung, die die visuelle Unübersichtlichkeit minimiert und gleichzeitig die wichtigsten Informationen vermittelt, ist noch nicht vollständig realisiert. Diese Arbeit untersucht die Anwendung von AR für die Visualisierung von großen Bereichen mit hoher Informationsdichte, wie zum Beispiel Produktionshallen. Sie befasst sich mit den Einschränkungen bestehender AR-Lösungen und schlägt einen Ansatz vor, der einen vordefinierten Blickpunkt als Aussichtspunkt verwendet. Die Implementierung findet in einer Fabrikhalle statt, wobei AR Head-Mounted Displays (HMD) zur interaktiven Datenüberlagerung eingesetzt werden. Zu den Herausforderungen gehören geeignete Tracking-Techniken, die Visualisierung von Daten in großen physikalischen Dimensionen und die Interaktion des Benutzers mit teilweise verdeckten Informationen von verschiedenen entfernten Orten. Eine Nutzerstudie evaluiert die geeigneten Interaktionstechniken und den Gesamtnutzen. Insgesamt zielt diese Forschung darauf ab, die Visualisierung und Interaktion in industriellen und architektonischen Kontexten zu verbessern. Die Ergebnisse zeigen, dass das System für seine Zielanwendung als hilfreich wahrgenommen wird und dass ein 2D-Übersichtsmenü aufgrund seiner Zuverlässigkeit als allgegenwärtige Interaktionsmöglichkeit gewünscht wird. Für ein besseres Nutzererlebnis wird Head Gaze oder Eye-Tracking bevorzugt. Head Gaze ermöglicht eine präzisere Zielführung, während der Eye Gaze zu einem besseren Erlebnis, mehr Immersion und weniger körperlicher Anstrengung führt. Die weit verbreitete Hand Ray Interaktion schnitt am schlechtesten ab, weil die exakte Bewegung schwierig zu lernen war.

Contents

1 Introduction 15

2 Background and Related Work 17
  2.1 Large-space Localization 17
  2.2 Visualization Techniques for High Information Density 20
  2.3 User Interaction in AR 21

3 Use Case Identification 25
  3.1 Interviews 25
  3.2 Chosen Use Cases 27

4 Implementation 33
  4.1 Software and Hardware Components 33
  4.2 Tracking 33
  4.3 Visualization 35
  4.4 User Interaction 36

5 User Study 43
  5.1 Research Questions and Expected Answers 43
  5.2 Design 44
  5.3 Procedure 44
  5.4 Pilot Study 44
  5.5 Participants 45

6 Results and Discussion 49
  6.1 NASA TLX 49
  6.2 SUS 52
  6.3 User-Reported Data and Feedback 58
  6.4 Discussion 59

7 Summary and Outlook 65
  7.1 Conclusion 65
  7.2 Limitations 66
  7.3 Outlook 66

Bibliography 71

A Additional Material 77

List of Figures

2.1 Comparison Space Pins and Image Targets 18
2.2 Head vs Eye Error Rate 22

3.1 ARENA2036 factory floor 26
3.2 5G Network Animation 28
3.3 Intelligent Floor Interface 29
3.4 FlexCAR Driving Animation 30
3.5 Safety Icons Shopfloor 31

4.1 The distribution of the four markers we placed on the balcony. Their recognition from the device has been highlighted by green squares. . . . 35
4.2 Shop floor 37
4.3 2D Menus 39
4.4 Visual Feedback for the Hand Ray Interaction 41

5.1 Age distribution of the study participants. 46
5.2 VR/AR experience of the participants. 46
5.3 Visual impairment of the participants. 47

6.1 QQ-Plots 50
6.2 Confidence intervals 54
6.3 NASA-TLX Boxplots for Head Gaze and Eye Gaze 61
6.4 Shows the mean and confidence interval of the overall utility rating for the application on a 5-point Likert scale, grouped by everybody and experts. 63

A.1 Histograms for data normality 77
A.2 Rolling Chassis and FlexCAR 3D models 78
A.3 Tooltip for 5G Traverse 78
A.4 Company Info Panels, shows occlusion problems 79
A.5 Boxplots for SUS 80
A.6 Boxplots for TLX 81

List of Tables

3.1 List of Interview Questions 26

6.1 Means and Standard Deviations for NASA TLX 51
6.2 Shapiro-Wilk and Levene’s tests for NASA-TLX 51
6.3 Kruskal-Wallis Test for Differences across User Interaction 53
6.4 Dunn’s test for *physical demand* 53
6.5 Dunn’s test for *performance* 54
6.6 User Interaction Technique SUS Scores 55
6.7 Means and Standard Deviations for SUS 56
6.8 Shapiro-Wilk and Levene’s tests for SUS 57
6.9 Dunn’s test for *easy* 57
6.10 Dunn’s test for *inconsistency* 58

A.1 List of Audio Interview Questions 79

1 Introduction

In factory planning and for operational duties, supervisors often need a quick and clear overview of the shop floor, the machinery, and components inside. Further large-scale use cases might include sight-seeing stops with annotations in the landscape, museums, planning and organizing setups for large events like fairs, or support for architects and building planners. Visualizing the necessary information and getting a quick overview can be achieved by using Augmented Reality (AR), which allows for superimposing real-world objects with virtual data. However, large spaces are not natively supported by AR solutions [Bil21], [FPS+20]. In general, the assistance of AR applications for the industrial context is constantly getting more attention, as it also brings enhanced opportunities to interact with robots [JBPH18], [CHS+20], [FCZ+23].

Previous work mainly tracked large rooms [HWW18] or the evaluation for larger spaces have only been researched for by walking around [HMG18], [FCZ+23], [TF22]. Their tracking evaluation still is an important foundation for this work. Regarding user interaction, many already mentioned the advantage of eye tracking [ŠIR+19], [BRP18], [PRSS16]. A related study with targets of different distances has been done by Pfeuffer et al. [PMMG17], but only in Virtual Reality (VR). A highly interesting work has been done by [QT17], which contrasting directs attention that pointing with the head might be much more precise.

This work implemented an AR approach for large-scale areas like factory floors. The solution used a predefined viewpoint to look down on the shop floor and has been implemented in a sample research campus to ensure a close-to-reality environment. Moreover, we researched and developed the possibility of using AR HMDs for interactive data superimposition of different parts or areas of the factory in real-time. We decided to use an HMD over a handheld AR device (for example, a tablet approach) to allow hands-free interaction and we assume a more comfortable, immersive user experience. The frequently implemented indoor-navigation approach [HMG18] is not sufficient due to security restrictions or physical limitations.

We first addressed the challenge of finding a suitable tracking technique for a factory area superimposition. Build upon that, we contribute research and implementation for the visualization of data within large physical dimensions or locations, following the concept of brushing techniques [BC87]. Our core contribution is the exploration of how large-scale dimensions can become "tangible" for users, drawing insight from earlier works like Yusof et al. [YHNI20], who showcased real hand gesture interaction in room-scale setups and Yoo et al. [YHC+10], who underscored the concern for fatigue in gaze and hand interactions in interactive wall displays. Adhering to Shneiderman's information-seeking mantra principle [Shn96], we aim to enable effective interaction with far-away or occluded data by different scalings of information.

Finally, we conducted a user study to evaluate the quality of the resulting application, gauging how effectively users can interact with data superimposed over large-scale physical environments, and how the incorporated user interaction techniques contribute to the usability, task load and overall user experience in a real-world factory setting. Through our work, we aspire to bridge the identified gaps and augment the scope of AR in facilitating intuitive, hands-free interactions in large-scale environments, catering to the distinct needs of industrial supervisors, planners and visitors. According to our results, such a system is regarded as helpful for the presented use case and beyond. Facing a suitable user interaction technique, a 2D menu is wanted to be an omnipresent method due to its reliability. To compensate for the lack of user experience (UX), a gaze method should be included as well. While pointing with the head works more precisely, eye tracking has the advantage of more comfort and UX. A distant interaction with free-hand gestures turned out to be difficult, because especially AR-unfamiliar users had problems to execute the gesture accurately enough for the hardware's requirement.

The subsequent thesis is structured as follows:

Chapter 2 – Background and Related Work discusses previous work in this area of research and places the thesis in that context.

Chapter 3 – Use Case Identification gives an overview of the collected factory use cases that can be visualized in AR and reasons the selection.

Chapter 4 – Implementation describes the challenges and chosen solutions for tracking, visualization and user interaction in the large factory space.

Chapter 5 – User Study states how the user study is designed, implemented and executed. Moreover, it provides the hypotheses, variables and pilot study results.

Chapter 6 – Results and Discussion reveals and discusses the results of the study's analysed and evaluated data.

Chapter 7 – Summary and Outlook summarizes the results of this thesis and its study and introduces possible future work based on it.

2 Background and Related Work

In this work, we focus on the three major components Large-space Localization, Visualization Techniques as well as User Interaction for large, information-dense physical spaces in AR. Therefore, this chapter introduces and reviews related work previously conducted in these research fields. This review aims to facilitate the implementation of an appropriate tracking technique for self-registration within the space and reasons for a suiting visualization technique for this application. Moreover, we use it as a base to define research questions and hypotheses on how we believe that existing user interaction findings can be generalized for our application.

2.1 Large-space Localization

Implementing AR on smartphones is a simple, obvious decision as they are widely available, affordable, easy to develop for, continuously evolving and ubiquitous [SLB11]. We want to extend it to be more immersive and allow for hands-free interaction. Regarding the comparably small display, we would have decided to use a tablet for our factory floor overview, however this brings the drawback of arm fatigue and missing immersion. Either way, there are quite similar challenges we face. Arth et al. [AS11] defined the key problem of registration and self-localization regarding the six degrees of freedom (6DOF) in AR. They state that neither the SLAM algorithm [DB06] nor the GPS methods were sufficient for an accurate large-scale environment localization. Furthermore, Gherghina et al. [GOT13] introduced a QR-Code marker-based tracking approach which still is a promising feasible solution for this work using a head-mounted display.

Another alternative large-scale AR realization is via spatial projections. Marner et al. introduced such a system to assist design and prototyping [MSP+11]. Even though such a projector only needs to be calibrated, it is not as portable and flexible as a HMD or smartphone. Their following paper introduces a user interface, which needs an additional tablet [MSWT14]. Consequently, in our field of application, we decided to face the self-localization difficulty of an HMD, but have much wider possibilities for immersive user interactions.

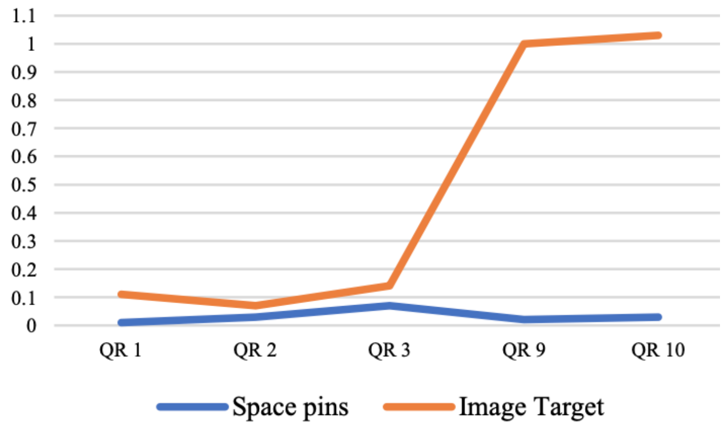


Figure 2.1: Compares the different deviation of Space Pins and Image Targets in meters, when walking along the Cathedral's south nave, which is 74m long, and passing the QR-Codes.

Hübner et al. [HWW18] provided a method to show that the Microsoft HoloLens[1] is capable of visualizing furniture for a large room with a good spatial accuracy when combined with tracking of multiple markers. They also found out that slow movement helps for orientation and registration. In the work of Teruggi et al. [TF22], using markers with HoloLens World Locking Tools (WLT)[2] have been compared to using Vuforia Marker Tracking[3]. Here, cathedrals and large buildings are used for benchmarking because of repetitive architectural patterns and difficult light incidence. Figure 2.1 shows that small orientation errors are accumulated to a larger shift for the Image Targets. They justify it by determining that "the stationary frame of reference does not allow tracking device errors" and that "rotations of the digital image target are manually set by the developer during deployment".

Our use case differs, as we only have a straight balcony to walk along with no additional environmental clues. Still, their reasoning for the shift holds and we also came to the conclusion that WLT space pins work best.

Either way, an industrial context will lead to new challenges compared to that room-scale showcase. Subsequently, Feigl et al. [FPS+20] evaluated to HoloLens, Google ARCore and Apple ARKit localization limitations for walk trajectories of $60m$ in a $1,600m^2$ context. They found out that "out of the box, these AR systems are far from useful even for normal motion behavior". They state that it would be better without the need

---

[1] Microsoft HoloLens: https://www.microsoft.com/de-de/hololens/hardware

[2] Microsoft WorldLockingTools Release: https://github.com/microsoft/MixedReality-WorldLockingTools-Unity/releases

[3] Vuforia Image Target Documentation: https://library.vuforia.com/objects/image-targets

of re-localization as in case of shorter, slower trajectories in one dimension. Still, the HoloLens using a combination of visual inertial simultaneous localization and mapping (VISLAM) and RGB-D outperforms the other HMDs. This finding inspired this work to use a HoloLens2 for implementation. Further, Hübner et al. [HCL+20] evaluated the capability of the HoloLens to use a triangle mesh for self-localization. They found out that among the multiple sensors, the most decisive one is the depth sensor. Moreover, having furniture or objects present in a room improves the tracking as they serve as texture markers. Finally, they reminded that the HoloLens is "not primarily designed as an indoor mapping device". Unfortunately, we cannot make use of the inbuilt triangle mesh localization method, as the shop floor is too far away to be recognized as a mesh. At present, it is recognized that previous AR benchmarks are not sufficient for more realistic and challenging data [SDS+22]. Others try to improve collaborative localization and 3D mapping via moving these tasks to a cloud [DCHR22] or include a deep learning approach for 6DOF tracking [GL17].

Moreover, there is much related research that focuses on augmented reality for human-robot interaction in manufactoring tasks. Our work also holds a potential for this extension, as robot tasks can be depicted from above and they can be sent to different places while having an overview. Juraschek et al. [JBPH18] already identified the possibilities that virtual, mixed and augmented reality can bring for the use case of a learning factory. They concluded that it can "improve imparting knowledge and skills". Piardi et al. proposed virtual laser range finders to operate with multi-robot systems, however only in a small-scale warehouse environment. Eventually, the work of Chan et al. [CHS+20] is closely related to this thesis as they also conducted a UI study in a manufacturing environment. For registration, they used a virtual model of the robot together with AR markers, that can also be used for re-calibration. In the study, where they compared their AR interface with a joystick interface and a manual method, the AR interface yields promising results to "improve task efficiency and reduce physical load". Similarly to our user interaction study, they used a ray cast from the head orientation to set trajectory points for the robot, confirming the input with voice commands.

Most recently, Fang et al. [FCZ+23] systematically reviewed HMD AR in manufacturing. According to them, current hybrid tracking in AR HMDs lacked the reliability and precision necessary for extended manual tasks in manufacturing settings. As a result, marker-based tracking techniques, which they describe as having a "low computational complexity and [being able to] provide real-time and reliable pose estimation", were frequently employed as a supplementary approach in production environments. Additionally, different interaction techniques have been reviewed. Regarding hand interaction, the device-assisted method is compared with the bare-hand one. Problems with extra devices are that they might be less affordable as well as invasive and uncomfortable. We also favored to use a bare-hand interaction. Still, discomfort in the shoulder region is likely to stay as a drawback, direct touch would be better and faster.

Unfortunately, direct touch is no possibility for our viewpoint application. In terms of head gaze, they mentioned the ergonomic problems of long-term use. As our overview application is expected to be used for a shorter amount of time as it would be for factory workers, this might still be a feasible option. Eye-tracking is mentioned as naturally and fast, but error-prone and suffering from "excessive user-specific calibration". The drawback of voice interaction, according to them, is the problem of noisy factories and disturbance on the shop floor. Most recommended is a multi-modal interaction that has the potential to compensate for various single interaction drawbacks, thus being more flexible and reliable which could lead to a better UX. Concretely, the combination of gesture, gaze and speech is recommended, like we implemented in our user study.

A different approach to get a better understanding of such a shop floor's structure is to walk around with AR functioning as indoor navigation. Hube et al. [HMG18] presented a typical example prototype for an exhibition navigation. Earlier, Oskiper et al. [OSK12] already presented a method by combining a monocular camera, an inertial measurement unit and GPS signals to be able to self-localize on a large space for navigation. However, due to possible security restrictions or physical limitations to reach every spot in a factory context, we recognized the need to have an overview of the hall from a vantage point. Compared to our case, it was easier because nearby areas can be modelled. Here, we have a large distance to the floor. Moreover, we can not explore much by just walking around, as we can just walk along the balcony at the short side of the building.

Rompapas et al. [RSP+19] introduced a large scale high fidelity collaborative AR experience. It combines a huge outdoor space with collaboration and interaction. For interaction, they used an indirect approach where users are represented by avatars. By contrast, our application is implemented in a more dense space where visualizations have to accurately superimpose physical objects. We have to deal with information overflow and research for a suiting interaction technique from a viewpoint about nine meters above the target floor.

2.2 Visualization Techniques for High Information Density

A highly important research paper for the visualization of this work is Shneiderman's famous work where he introduced the information seeking mantra [Shn96]. It serves as a guiding principle for designing effective visualizations and user interfaces and succinctly states: "Overview first, zoom and filter, then details-on-demand." *Overview* provides users with a broad view of the content to help them understand the available data, *Zoom and Filter* allows users to zoom into areas of interest and filter out unimportant information and *Details-on-demand* provides more detailed information for users when they request it. In a nutshell, this mantra emphasizes a progressive disclosure of information, guiding users from a general understanding to a more detailed exploration as per their needs. In this work, we implemented a minimized version for all use case visualization as well as colored boxes with company name tooltips at their respective locations, which addresses the *Overview* level. For *Zoom and Filter*, a 2D menu is implemented which enables to filter for different shop floor zones as well as specific company and research areas. Here, also the well-established brushing and linking method [BC87], widely spread in information visualization, can be applied. The *brushing* concept refers to the 2D menu view, where several areas or categorizations can be selected, while *linking* can be interpreted as the corresponding cuboids on the real factory floor that will light up when selected. Regarding *Details-on-demand*, several detailed company information panels and showcases can be enabled individually. Since readable panels will occlude much space of the shopfloor as well as other interactable symbols, it is not possible to clearly show them all at once.

2.3 User Interaction in AR

The endeavor to develop intuitive interaction mechanisms within large, information-dense spaces in augmented or virtual reality is a complex challenge. Previous efforts have laid a foundation by exploring various interaction modalities, particularly focusing on gaze and hand gestures, in differing contexts and environments. Shi et al. [SWQ+23] investigated region selection in AR environments. They compared Gaze-Finger, Gaze-Pinch, Pinch only, and Eye only techniques and found that Pinch only (PO) and Eye only (EO) techniques were suitable depending on whether the users' hands were available or not, thus providing an insightful perspective on designing interaction techniques based on user availability and preference.

The early efforts of Hales [HRM13] and Park et al. [PLC08] provided a foundation for gaze and hand gesture interactions within AR, albeit within a more confined spatial domain. Despite the limited scope, their work hinted at the potential for these interaction techniques within augmented reality settings. Pfeuffer et al. [PACG14; PMMG17] extended the gaze and hand gesture interaction paradigm to VR environments. Their efforts demonstrated the feasibility of manipulating objects either near or far using gaze selection and freehand gestures, thus broadening the scope of gaze and hand-based interactions beyond AR. Marques [MAN+20] explored various interaction methods for assembly procedures and highlighted the acceptance and usability of different techniques like touch gestures and mobile device movements, which could be insightful for designing intuitive interactions within larger spatial domains. The comparative analysis of eye gaze and head gaze in collaborative games by Spakov et al. [ŠIR+19] and in VR/AR systems by Blattgerste et al. [BRP18] and Palinko et al. [PRSS16] revealed that eye gaze

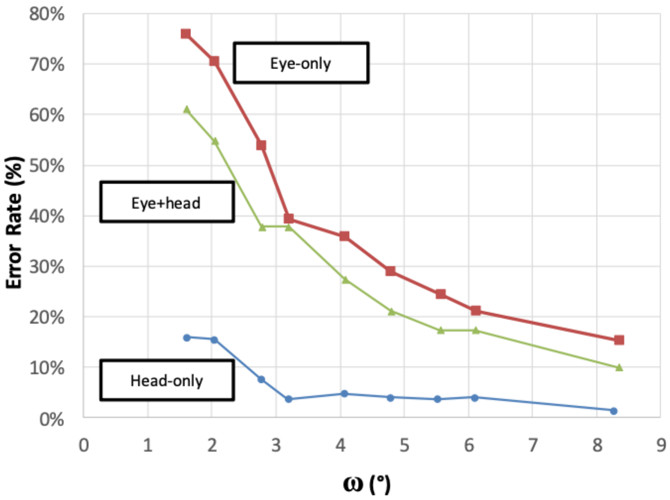


Figure 2.2: Shows the extremely lower error rate in percent of head-only compared to eye+head and eye-only, depending on the angular size of the target in degrees [QT17].

could be faster and less exhausting than head gaze, providing a potential direction for optimizing user interaction within large spaces. Several other works [HDP11; LRP+22; QT17; SJ00; WLM+23; YHC+10; YHNI20; ZSSB15] also explored gaze and hand gesture interactions across different contexts, adding to the body of knowledge on how these modalities can be employed effectively.

Several other studies have delved into the dynamics of gaze and hand gesture interactions across varying contexts, each contributing to the comprehensive understanding of how these modalities can be effectively utilized. Notably, the work by Qian et al. [QT17] serves as a pivotal reference, as it diverges from the common favor toward eye tracking observed in earlier studies. Through an empirical comparison of head-based and eye-based selection in virtual reality, they discovered that head-only selection significantly outperformed eye-only selection in terms of error rate, selection times, and throughput. A comparison between eye-only, eye+head and head-only in terms of the error rate in percent depending on the angular size of the target in degrees is shown in Figure 2.2. Moreover, head-only selection was "strongly preferred by participants," thus challenging the prevailing notion and illuminating a potential avenue for head-gaze-based interaction in large, information-dense spaces. Their work also attempted a combination of head and eye gaze, which unfortunately did not yield improved results, providing a crucial insight into the nuanced interplay between these interaction modalities.

Yusof et al. [YHNI20] showcased a room-scale setup where handheld AR facilitated real hand gesture interaction, providing a glimpse into the limitations associated with handheld AR in large-scale setups. Similarly, Zhang et al. [ZSSB15] provided an exploration into the costs and benefits of combining gaze and hand gestures for remote interaction, while Wagner et al. [WLM+23] demonstrated the decreased performance of gaze and finger alignment in increased depth scenarios. The investigation by Yoo et al. [YHC+10] on an interactive wall display situated three meters away underscored the concern for fatigue in gaze and hand interactions, despite the effectiveness of the combination. Lystbaek et al. [LRP+22] further explored this combination for menu selection nearby, highlighting the challenge posed by motion parallax.

Collectively, these diverse explorations provide a rich basis for understanding the limitations and potentials of gaze and hand-based interactions. The contrasting findings, especially from Qian et al. [QT17], present a nuanced landscape where head-gaze may offer unique advantages, particularly in scenarios involving large information-dense spaces with far away or occluding objects.

In light of the foregoing studies, our work extends the interaction paradigms to large information-dense spaces, aiming to address the specific challenges such as distant object interaction, occlusion, and visual clutter which haven't been extensively tackled before. Through our research, we aspire to facilitate more intuitive and visually uncluttered interactions within such complex environments, building upon the valuable insights provided by the preceding works.

To the best of our knowledge, this has not been done before in large, information-dense environments. The challenge we face is that simply walking to a cluster of virtual objects and directly touching them is not feasible due to the distance. Furthermore, we encounter difficulties with faraway objects that partially occlude each other, as well as with complex boundary objects, moving objects, and large, nearby objects. It is challenging to devise a method that eliminates visual distractions while conveying all essential details.

3 Use Case Identification

This work contributes in two important ways. First, we aim to implement a problem-driven utility approach to improve the understanding of a research campus's structure and ongoing projects through the use of an AR HMD. Second, we conducted a usability analysis of various interaction techniques tailored for information-dense, large physical spaces, which is presented in the following chapters.

To be able to visualize suitable use cases, we started with conducting stakeholder interviews to identify, which important companies and projects should be included. We also factored in our decision that we obtain different representations instead of only textual information and that they are located at different distances to cover the design space. The factory floor serving as the sample use case is ARENA2036 e.V.[1] at the University of Stuttgart, while ARENA is short for *Active Research Environment for the Next generation of Automobiles*. In this collaborative research campus, large companies, start-ups, and research institutes work together on a shared shop floor, as illustrated in Figure 3.1. This setting gives rise to several standard use cases, such as categorization highlighting and displaying various metrics like the number of employees, membership dates in ARENA2036, and current projects along with their associated partner companies. In addition to static textual information, the incorporation of interactive elements and animations is essential for enhancing both user experience and learning outcomes. Further details about our following visualization decisions are explained in Section 4.3.

3.1 Interviews

To perform a complete problem identification for the final utility of the application, we conducted semi-structured interviews with eight different stakeholders, including the managing director, shop floor- and project managers, company contact persons, research coordinator and technical manager of the research campus.

---

[1] ARENA2036 e.V. Homepage: https://www.arena2036.de/de/

Figure 3.1: The factory floor of the ARENA2036.

The following questions in Appendix A served as a guideline, while an open discussion has been preferred rather than asking predefined questions one by one.

Interview Questions
- Which companies can be interviewed and seen well from the balcony?
- How much is changing at the companies' locations?
- What data do they produce that could be visualized?
- Where could the information be displayed?
- What security aspects are sensible to show?
- Are there existing interfaces to exchange data?
- What is worthwhile to show from the balcony, what would be information overflow?

Table 3.1: List of Interview Questions

We received input that it is sensible to group the shop floor areas to locations of start-ups, big companies and to split between economy and university spaces. Moreover, the shop floor consists of three zones with different degrees of fixation, ranging from quite fix to often restructured. We decided to include a zone visualization as transparent, colored planes in our menu to be switched on and off. Regarding single companies, the information should contain the accession date, headquarter, contact person, scale and important projects, according to the project managers. We have chosen to include a selection of the most influential companies, where we have taken care of covering close, large areas, small ones that are far away and some that will occlude each other, when additional information is unfolded. This should properly cover the design space.

Regarding projects, we decided for a well covering selection of influential projects with many partners that could also be linked when opening the project visualization. In order to cover diverse use cases and visualization possibilities, we chose a project that can be directly controlled by data exchange, one with a moving object that will also be challenging to point at for the user interaction study, and one where we can display a network animation. Furthermore, we talked to the technical manager to identify suitable safety instructions. We decided to include escape doors, toilets, fire extinguishers, central light control panels and first aid boxes. Unfortunately, for the scope of this work, we had to discard an available laser scan as it was too large be be rendered and used in real-time.

3.2 Chosen Use Cases

In this section, the four different use cases, that we have decided to include in the resulting application are presented one by one. They have been chosen as they are among the most relevant projects with many partner companies. Additionally, they allow for different kinds of animations like parabolas with animated color gradient, a moving car, remote control of real floor tiles and the expansion of 3D safety signs. The figures presented in this section are all screenshots captured by the HoloLens. As there is only one single camera that is located above the eyes, the perspective differs from the user's perception [HWW18] and stereo vision is not possible.

3.2.1 5G Synergy Region

In 2020, ARENA2036 became the first Baden-Württemberg university research site to have a 5G structure accessible to all research partners. The network operates in the 3800 MHz frequency range and covers the hall area as well as an adjoining outdoor area [2]. Hence the transmitter, which is Nokia[3], and the receivers all are located in the same hall, this results in a promising use case for visualization. We decided to visualize it as parabolas with animated color gradient coming from the Nokia Technology Traverse and ending at the partners which use the network. The animation can be activated by clicking on the WiFi signal Gif at the bottom of the traverse. This might be hard as it is located directly on top of the Nokia company information cube and if activated, the car animation described in Section 3.2.3 will drive through the traverse as well, leading to the requirement of aiming precisely on the correct object. Figure 3.2 depicts the activated 5G animation starting from the animated image under the physical traverse system.

---

[2] 5G Synergy Region project page: https://arena2036.de/en/SynergieRegion

[3] Nokia Corporation, Homepage: https://www.nokia.com/

Figure 3.2: The 5G network animation shown as color gradient parabolas. Each parabola starts at the transmitter and ends at the receiver's place. The red arrows are not included in the application but added afterwards for a better visibility on the screenshot.

3.2.2 Intelligent Floor

The Intelligent Floor[4] by Robert Bosch GmbH[5] is another use case that applies to multiple companies on the ARENA2036 shop floor. The interactive floor plates can sense what stands on them and are even able to guide self-propelled transport boxes. It can also tell where a human stands or walks to avoid collisions. For demonstration purposes, the edges of these plates can be illuminated in different modes. For instance, they can track a person's movements by lighting up the plates as they are stepped on.

Its coloring showcase can be controlled by a MQTT interface that can be reached via a website within ARENA WiFi. It is simple in structure hence it consists of oblong blue buttons for each visualization mode. When pressing the "Intelligent Floor" button in the main menu or aiming with gaze or hand at the play button superimposed on their huge video wall, respectively, a browser is opened and the user is able to control the floor lighting via the interface. Moreover, there is a picture overlaid on the video wall.

---

[4] Intelligent Floor project page: https://arena2036.de/en/interactive-bosch-floor

[5] Robert Bosch GmbH, Homepage: https://www.bosch.com/de/

Figure 3.3: The Intelligent Floor being controlled via a 2D-interface. In this picture, the "rainbow" color effect is shown.

Figure 3.3 shows the opened control interface along with the resulting animation of the floor panel borders.

3.2.3 FlexCAR

The FlexCAR[6] project's vision is to use an open vehicle platform that is able to run driverless and provides the possibility to alter the interieur depending of the specific use case. The app is able to overlay the physical rolling chassis with an interieur vision designed for business meetings. When the car is intersected with gaze or hand ray or in the menu case, its respective button pressed, it follows a predefined invisible bezier curve to simulate driving outside the hall gate. To stop it, users will need to "catch" its bounding box while it drives in a loop. Figure 3.4 shows the FlexCAR model on its way outside the hall. One can see how it crosses the WiFi Gif and the cuboid locating the traverse system, hence this use case is expected one of the hardest regarding user interaction evaluation. An easy extension is to add possibilities that allow to switch

---

[6] FlexCAR project page: https://arena2036.de/en/flexcar

Figure 3.4: The FlexCAR, captured in a screenshot during its movement, as it crossed both the 5G animation target and the Nokia information target. The red arrow is not included in the application but added afterwards for a better visibility on the screenshot.

between the rolling chassis and the complete FlexCAR model or being able to grab it or let it fly towards the user so it can be regarded at in detail on the balcony.

3.2.4 Security

Another corporate use case is to highlight safety equipment. In this implementation, we chose to show the location of fire extinguishers, central light control panels, first aid boxes, emergency exits as well as toilets. To highlight the location, 3D object clones have been chosen for fire extinguishers, first aid boxes and toilets, 2D signs have been used and virtually placed on the wall to show the emergency exits and for the light control panels, 3D light bulbs will appear. This is generally applicable to any building and might save lives. The exemplary right side of the factory shop floor, overlaid with those items, is shown in Figure 3.5. Furthermore, as the balcony occludes the entrance area, a floor plan has been added there so the user is able to locate the safety items there as well.

To toggle the presence of all the icons, which of course will lead to visual clutter when always on, a "parent" fire extinguisher object is always present and one can expand all

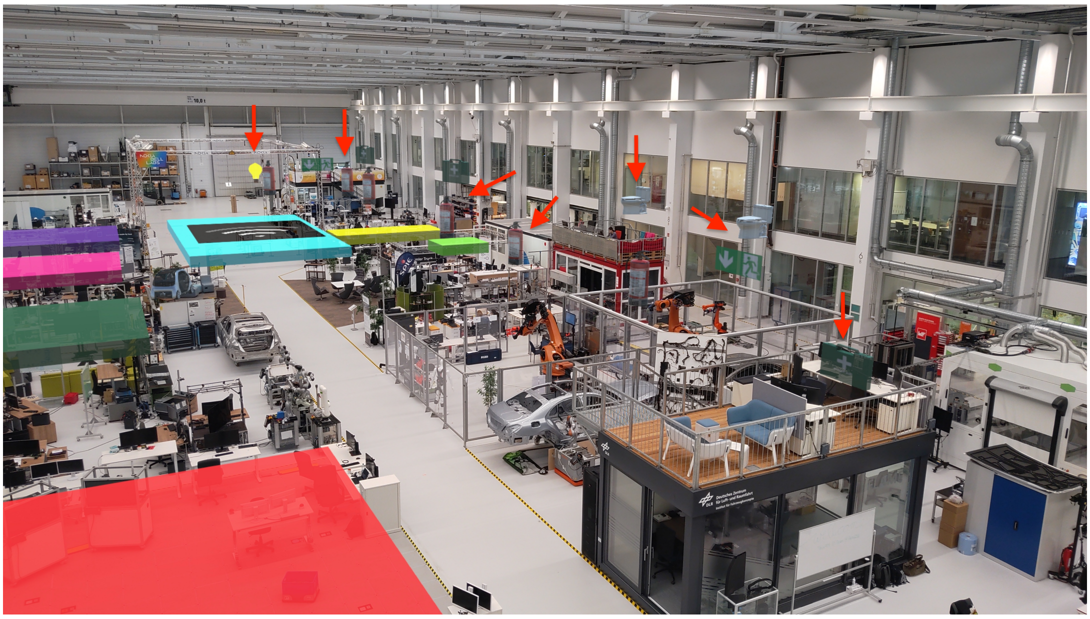


Figure 3.5: The safety icons, including the light bulb to show the central light control panel, the toilet symbols showing their location in the second floor office clip as well as many escape doors, first aid boxes and fire extinguishers. The red arrows are not included in the application but added afterwards for a better visibility on the screenshot.

the items by clicking on it. Similarly as the rolling chassis use case, it is comparably hard because its boundaries are cylindrical with a handle instead of the simple cube bounding boxes.

4 Implementation

This chapter describes how the resulting AR application is implemented. A core piece we first needed to explore is a suitable tracking technology to obtain stable holograms on the large factory floor. Subsequently, the use cases described in Chapter 3 have been visualized in an appropriate way. To prepare for the study in Chapter 5 that evaluates different interaction techniques, we eventually describe our four different interaction methods and reason for the design decisions.

4.1 Software and Hardware Components

As an AR HMD, we chose the HoloLens2, released in 2019 by Microsoft [Mic]. It orients itself in the world using VISLAM, an extension of the classical simultaneous localization and mapping (SLAM) algorithm [DB06]. To implement the AR application that will be deployed on the AR headset, Unity 3D[1] has been chosen together with several plugins from Microsoft's Mixed Reality Toolkit (MRTK)[2].

4.2 Tracking

To enable the visualization of a large-sized factory floor, several considerations must be made. In this specific use case, the ground plane has a size of 130 * 46 meters, which exceeds the range, where a digital twin of the factory floor can be created on the fly from approximately nine meters above the floor. Unfortunately, this distance also exceeds the limit for the HoloLens to be able to recognize physical meshes, which could have been used to compute intersection rays as applied from Hübner et al. [HCL+20]. Thus, the nearby approach is to use markers that allow the hardware to use them as reference points to align the virtual objects in the physical world.

---

[1] Unity Homepage: https://unity.com/

[2] MRTK Documentation: https://learn.microsoft.com/de-de/windows/mixed-reality/mrtk-unity/mrtk2/?view=mrtkunity-2022-05

4.2.1 Image Targets

In order to align the virtual content with the physical hall, we first started by attaching highly recognizable markers with many features to the balcony using them as Vuforia Image Targets. With this method, the exact placement of markers is predefined during the developing process and all virtual objects are oriented in the physical world as children from that reference image target. Once the event is triggered that all markers have been tracked, the interpolation point between them is calculated and used to align the virtual overlay for the factory floor.

However, this initial approach involving Vuforia Image Tracking was discarded due to several challenges. It required manual setup of multiple targets, exhibited instability, and posed difficulties in managing the placement and scaling of child objects relative to the targets. Details have already been given in Section 2.1, when referring to the work of Teruggi et al. [TF22].

4.2.2 World Locking Tools

The MRTK offers an array of world locking tools that enhance the spatial consistency and interaction experience in mixed reality environments. These tools facilitate the alignment of virtual content with the physical world, ensuring that digital elements remain accurately positioned relative to the user's surroundings. Utilizing methods such as spatial anchors and QR-code tracking, MRTK provides developers with tools to develop applications that integrate virtual and real-world elements in a coherent manner. This world locking functionality not only enhances the stability of augmented reality experiences but also enables the persistence of spatial data across different sessions, providing users with a consistent and reliable mixed reality application.

4.2.3 Spatial Anchors

In light of the issues with Image Targets as explained in Section 4.2.1, a decision was made to opt for QR-Code Spatial Anchors. This choice was driven by its compatibility with the Mixed Reality Toolkit (MRTK), its reliable performance, the capability to store anchors consistently across different sessions, and the advantage of being able to simultaneously track multiple QR codes or anchors. Now, after the first deployment of the application, it is required to walk around the balcony and scan the four QR-Codes with the headset's tracking camera. We distributed the printed QR-Codes with several meters between each other on the balcony wall and floor. The recognized markers, visualized as shining green squares, are shown in Figure 4.1. We decided for four codes,

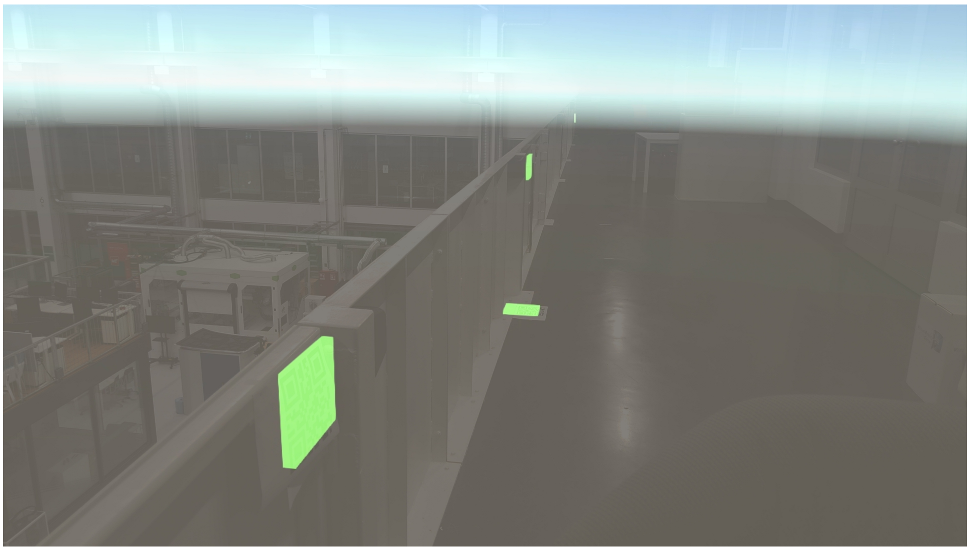


Figure 4.1: The distribution of the four markers we placed on the balcony. Their recognition from the device has been highlighted by green squares.

because three codes, when placed on two orthogonal planes like the balcony wall and floor, will lead to a unique determination of position and rotation in 3D space and the fourth one will work as a backup. Anyway, this only has to be done the very first time after opening the newly deployed app. After that, no more manual scanning is required.

4.3 Visualization

As a starting point, we overlaid the arena floor with a planar mesh of the factory floor map. The map serves as an overview of where the companies and projects are located. It is shown in Figure 4.2. However, in the resulting application release, this map has been removed, so that no real objects will be occluded unnecessarily. Still, the points of interest are highlighted in different ways.

In our visualization system, we have incorporated various dynamic and interactive features tailored to improve user engagement and comprehension. Firstly, we use distinctly colored cuboids, also included in Figure 3.2, to represent different companies, making it immediately evident to the viewer where each company is located. This straightforward visual differentiation is further enhanced by an intuitive tooltip-label system. When a user hovers over a cuboid, either using gaze or hand movements, a label appears displaying the company's name, allowing for quick and efficient identification. Furthermore, our visualization captures the 5G network research at the ARENA2036. We represent this network through animated color gradient parabolas that span from producer to consumers, offering a visually arresting depiction of data flow and connectivity.

An integral part of our visualization is the presentation of FlexCAR's digital twin. Rendered in detail, this digital representation showcases the vision to expand the physical rolling chassis with the suitable FlexCAR attachment, providing users with a comprehensive visual understanding of its design and functionality.

To accentuate safety and operational aspects, we've also integrated digital twins of essential safety objects, such as fire extinguishers. These digital replicas ensure that users have a clear visual guide for emergency equipment locations, reinforcing both safety and operational procedures. Lastly, our system incorporates an image projected on the physical video wall near the intelligent floor. It leads to the illusion that the floor covers a whole factory with drones flying around. Besides, it features a prominently centered play button. Once activated, this play button initiates the intelligent floor interface, allowing to control the physical LED lighting interactively to provide a richer, more integrated user experience. Together, these visualization tools not only provide an aesthetic and functional interface but also enable users to navigate, understand, and interact with the represented data in an efficient and user-friendly manner.

4.4 User Interaction

Regarding the user interaction with the displayed content, there are two major challenges. First, the tracking of hand rays to interact with objects at that large distance does not work out of the box and the mesh of the hall cannot be recognized by the AR device. To overcome this, one possibility is to manually set the ray to a much larger distance or to infinity until it intersects with a hologram. Second, the user interaction needs to be intuitive for those who are not familiar with AR head-mounted display interaction, which will be the case for most of the target group. Ray interaction and the so-called "air taps" are often challenging and take users a while to adapt to. A more intuitive approach might be the interaction with a 2D menu, as subjects are used to doing when interacting with touch screen devices such as their mobile phones. Microsoft's second-generation HoloLens lets you tap directly on menu items, without the distraction of distant rays and air taps. Unfortunately, users might lose immersion with a 2D menu and are not able to interact directly with the environment. To overcome this problem, we also implemented two different gaze versions, one of which uses eye tracking, while the other uses the camera's direction of view above the eyes, called head gaze. Consequently, the following subsections will describe the implementation of the resulting four interaction techniques one by one.

Figure 4.2: The map showcases the layout of ARENA2036's shop floor, which is segmented into distinct areas, that house start-ups, large enterprises, research institutes, and collaborative projects. This layout implicitly illustrates the extent of visual clutter and occlusions we need to address. The dimensions of the building ground are 130 m x 46 m which is another grand challenge for implementation. The balcony is located at the large red line on the right, which is where users will walk around and look down on the shop floor. It is at 9m height.

4.4.1 2D Menu

As mentioned previously, the basic intuitive way to get an interactable overview is to show a 2D menu. This could be done by connecting a tablet or, to keep interaction hands-free, showing it via the HMD. For the latter, it has to be taken into account that occlusion of anchored visualizations is unwanted.

Hence, one might consider using the balcony, where the markers are located, to fix a menu. This ensures that the menu will never occlude parts of the factory, but it can be tiring to look down and having to track another marker can reduce the performance. Another option is to show and hide it when a certain gesture or voice input is recognized. The latter has the advantage of appearing in the users' peripheral vision, eliminating the need to look up and down. Moreover, due to hologram transparency, the factory can still be seen in the background, giving a preview of what the highlighting will look like.

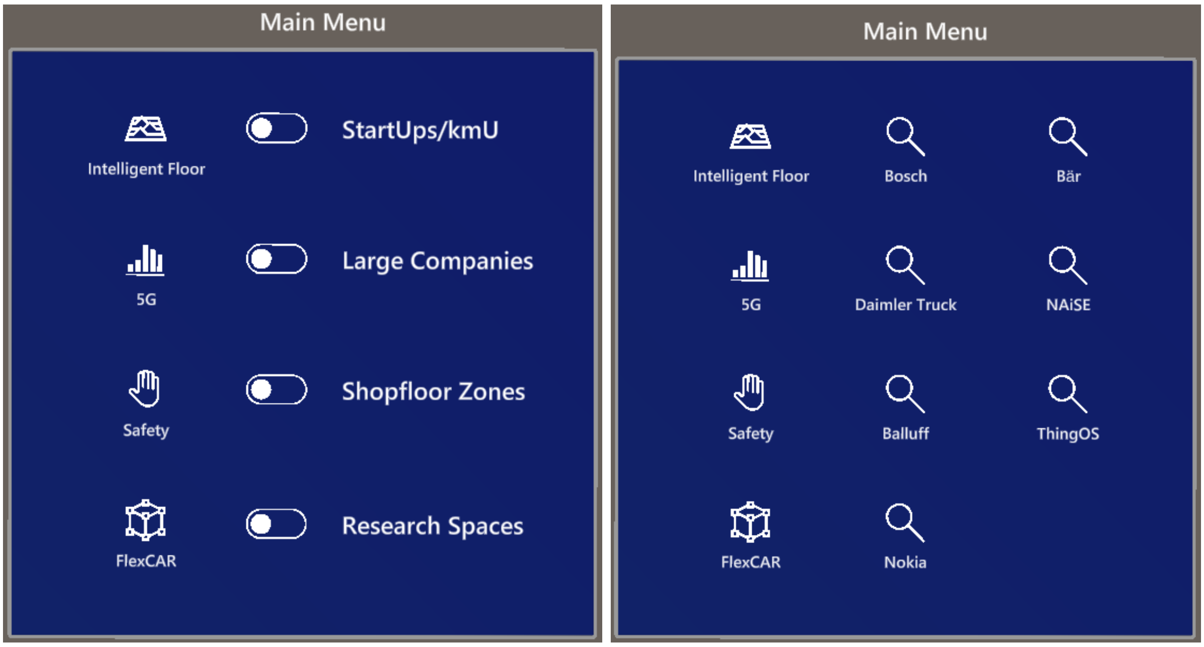
Finally, for menu interaction, we chose a nearby hand menu which will appear within an arm's length distance. The two possibilities to open it are to look at one's palm or to say "Show hand menu" as a speech command. Subsequently, the menu will be world locked until the user drags and drops or resizes it via direct touch interaction. Another possibility is to look at the palm at the desired position for the menu to appear. It can be closed by saying "Close hand menu" or by pressing the "x"-Button on the upper right corner. All buttons serve as toggles, which is displayed by a lighter blue button, if activated and a dark blue, if not currently active. Hence, the users can always keep an overview about what can be chosen and whether every use case has been explored yet.

The 2D menu has been implemented in two different versions, depicted in Figure 4.3. The first design version came from a linking and brushing approach, where switches on the right hand side allow for categorisation of company or research spaces, respectively. To head for specific companies, either further menu pages or a combination with another user interaction technique will be required. When designing the user study, we came up with the second design version. The main point has been to ensure equality of functionality between the indirect manipulation via the 2D menu and the following strategies namely distant hand ray or gaze interaction directly performed on the target objects. To this end, the left column still represents the special use cases while the center and right column show the prototypically chosen companies where detailed information can appear.

4.4.2 Hand Ray

Hand interaction with distant content presents unique challenges [WLM+23]. The combination of hand ray and air tap offers an intuitive solution to bridge this spatial gap [FCZ+23]. The hand ray interaction involves extending an imaginary line or "ray" from the user's hand towards the AR content, effectively acting as a remote pointer. This allows users to target and select items that are far away without the need to physically approach them. Once the desired content is targeted using the hand ray, an air tap, a gesture made by mimicking a tap in the air, can be employed to trigger an action, such as selecting or activating the targeted content. This method of interaction enables efficient and ergonomic interaction with far-reaching AR content, reducing the need for users to constantly move closer to objects of interest, thus enhancing the user experience in expansive AR scenarios.

In terms of visual feedback, a dashed line—either straight or parabolic—extends from the user's palm, serving as the starting point. At the conclusion of this line, a ring is present that aligns with the bounding boxes of objects when they intersect. Upon detection of an air tap movement, the ring transitions into a smaller, solid circle. This provides clear visual feedback, ensuring users are aware of their tap recognition. Any potential errors can then be attributed to imprecise selection of bounding boxes.



(a) (b)

Figure 4.3: (a) shows the variant as a supplementary menu, where the highlighting of company and research locations as well as shop floor zones can be switched on and off additively. This feature enhances the overall view and minimizes visual clutter. (b) presents the standalone version, which is designed to be functionally equivalent to the other user interaction options for the purpose of the user study.

4.4.3 Eye Tracking

To utilize the Eye Tracking technique, one must first adhere to the calibration guide. We opted against using a cursor because our pilot study indicated that it tends to lag remarkably within the target space, causing confusion. This delay is likely due to the extensive distances over which it needs to be moved and rendered. When an object is gazed upon, a tooltip displaying the company name emerges, providing subtle feedback. For object selection, we offer two options: either verbalizing the word "select" or executing an air tap. This air tap is reminiscent of the hand ray interaction, but it can be performed anywhere within the HMD’s field of view. We chose not to implement a prolonged dwell time for selections because it could disrupt natural eye movements. Requiring a nodding behavior to toggle information on and off seemed counterintuitive.

Upon self-empirical testing, we discovered that reducing the dwell time often resulted in unintentional activation of information and specific use case animations. This unintentional triggering led to visual clutter and necessitated manually deactivating everything, which means capturing everything once more with one's gaze. Various findings from previous research comparing head gaze, which will be described in the next subsection, and eye tracking suggest that eye-based interaction is generally faster, more favorably evaluated by users, less exhausting, and allows for a more natural interactive behavior [BRP18; PRSS16; ŠIR+19].

4.4.4 Head Gaze

For the subsequent interaction method, head gaze, the targeting and selection/manipulation techniques largely mirror those of the eye-tracking interaction. The primary distinction lies in how objects are targeted: Through a virtual ray based on the position and orientation of a camera situated just above the eyes. We hypothesized this method to offer greater accuracy and reliability than eye-tracking targeting for this use case, as it should be similar than the VR-Study case of Qian et al. [QT17]. However, a clear disadvantage is that it necessitates more pronounced head movements from the user, potentially leading to physical discomfort, such as neck strain. The physical effort compared to eye movements has been already stated by Blattgerste et al. [BRP18].

In the context of the head gaze technique, the inclusion of a cursor proves beneficial for immediate user feedback. Given that head movements are generally not as fast as eye saccades, the lag issue associated with cursors is mitigated. The presence of a cursor, therefore, boosts user confidence in accurately selecting the intended item, which has also been recommended based on Microsoft's research[3]. The visual feedback for the cursor closely resembles that of the hand ray, with the notable exception of the dashed line. Users will only observe the ring on objects. Upon executing an air tap, this ring transitions into a solid circle. In instances where speech input is utilized, the cursor remains unchanged. Instead, a tooltip displaying the recognized word, in this case "select", emerges and gradually fades away.

---

[3] Gaze and Commit Recommendations: https://learn.microsoft.com/en-us/windows/mixed-reality/design/gaze-and-commit

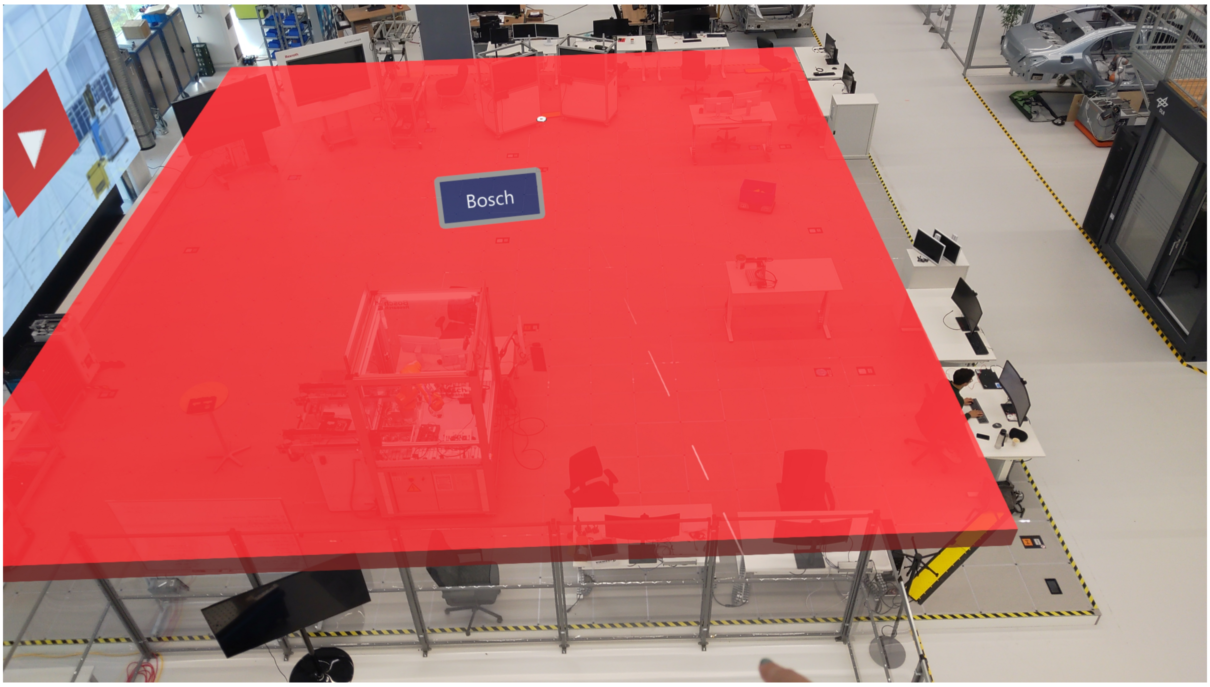


Figure 4.4: Shows the hand ray cursor, where the dashed line from the user's finger to the target object is visible

5 User Study

The primary objective of this experiment is to evaluate four distinct interaction techniques, focusing on a combination of usability and utility assessment. Usability is assessed through the metrics NASA Task Load Index (TLX)[1] and System Usability Scale (SUS) [Bro96], while utility is evaluated through qualitative feedback obtained via post-interaction questions, among others with use case experts.

5.1 Research Questions and Expected Answers

In our user study, we aim to obtain a detailed insight about what user interaction can be recommended for large, information-dense physical spaces and how the overall utility of the system is evaluated. We guide our analysis by the following research questions and allocated answers which we expect based on our literature review.

RQ.1 Which interaction is preferred regarding UX and effectiveness?

RQ.2 Is a combination of techniques desired?

RQ.3 How is the utility of the system evaluated?

EA1.1 The 2D menu will be easiest, fastest and most successful to use, but yield the lowest immersion and UX.

EA1.2 The hand ray interaction is easy and intuitive.

EA1.3 Head gaze is more precise than Eye gaze, while eye gaze is physically more comfortable.

EA.2 A combination with an other technique and the 2D menu is desired.

EA.3 The system is regarded as helpful for its target application.

---

[1] NASA TLX Homepage: https://humansystems.arc.nasa.gov/groups/tlx/

5.2 Design

The experiment employed a Latin-Square design to ensure an equitable distribution of conditions among participants. It also followed a within-object approach, with each participant experiencing all four interaction conditions: hand ray, eye gaze, head gaze, and 2D menu. Those conditions form the characteristics of the independent variable.

5.3 Procedure

Participants were provided with a verbal description of the task, including all special use cases and considerations as described in Chapter 3. Before commencing the experiment, participants completed a pre-questionnaire, offering anonymized demographic information.

The experiment consisted of four iterations, with each iteration involving one of the four interaction techniques. After each iteration, participants completed NASA-TLX and SUS questionnaires. Additionally, they responded to three qualitative questions:

1. What did you like about this user interaction?
2. What did you dislike about this user interaction?
3. What would you change about this user interaction?

To gain further insights, audio-recorded interviews were conducted with participants, guided by predefined questions, exploring their experiences and perspectives. In total, we collected 2:39:51h of audio material. This structured experimental approach allowed for a comprehensive evaluation of the four interaction techniques, encompassing both quantitative and qualitative feedback from the diverse pool of participants.

5.4 Pilot Study

Before our regular study, we first conducted a pilot study to identify bugs, test the procedure and be able to identify a potential weakness of the questionnaire. We tested our user study with two participants, one male and one female, where both have extensive experience with conducting VR/AR user studies themselves.

Luckily, nothing but a small bug has been found where two objects did not respond to gaze selection. This could be fixed for the final study. Moreover, there has been an improvement of the virtual object alignment for the final study. In total, for both of them, the study took less than 60 minutes.

Both pilots already knew and liked the hand ray interaction, "as it is intuitive and working nicely" and "usually easy to use as gesture is widely known". They did not like that the eyes have to keep looking at the target objects while saying the speech command. Head Gaze was preferred. Finally, they liked the 2D menu, because they had an "overview of which items were active at that time, it felt like a control panel that is easy to access all of the time".

Regarding the issue with voice interaction, we introduced the alternative to select gaze-hovered objects with simply touching thumb and index finger inside the field of view. One pilot suggested to use a hand-held clicker as a selection technique for the eye and head-gaze scenarios. However, we opposed this suggestion as this will require again holding some extra device in the hand while the objective of using the HoloLens was to be able to do a natural, immersive user interaction.

5.5 Participants

In this study, 20 participants were engaged, along with two pilot participants. Regular participants fell within the age range of 20 to 38, with an average age of 25.2 and a standard deviation of 4.38. A histogram of the age distribution is shown in Figure 5.1. The gender distribution among regular participants consisted of 14 males and 6 females. Three of them were left-handed. Regarding visual impairment, three participants wore contact lenses during the experiment, nine of them wore glasses (Figure 5.3). One subject needed to take off their glasses (-4.5 dioptries) for the eye tracking calibration, because it did not calibrate otherwise. Consequently, the whole eye tracking condition has to be done without glasses for this participant.

Moreover, one participant had a color weakness. Most participants have used VR/AR for few hours at some demos, while three have it included in their daily work and one more used it for about 30 hours. A histogram of the VR/AR experience is shown in Figure 5.2. Participants were organized into groups using the Latin square permutation method, with each group containing five participants, representing different orders of experimental conditions. The participant pool included students, doctoral researchers, and research campus employees, the latter selected for their expertise in providing qualitative feedback and utility evaluations.

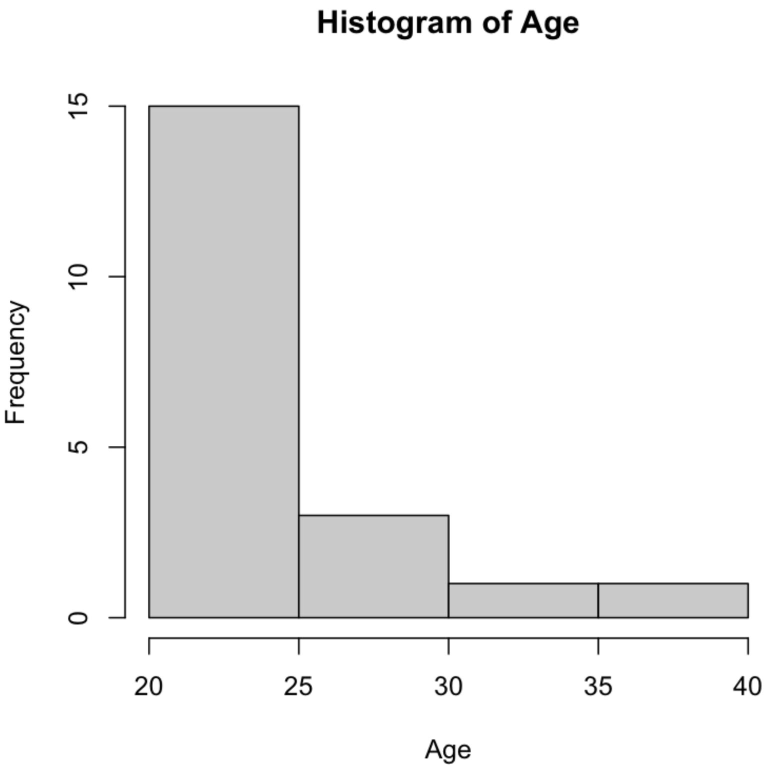


Figure 5.1: Age distribution of the study participants.

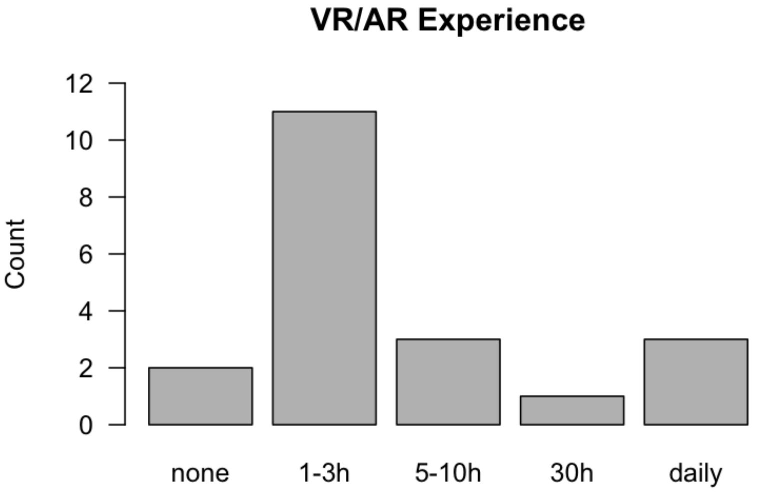


Figure 5.2: VR/AR experience of the participants.

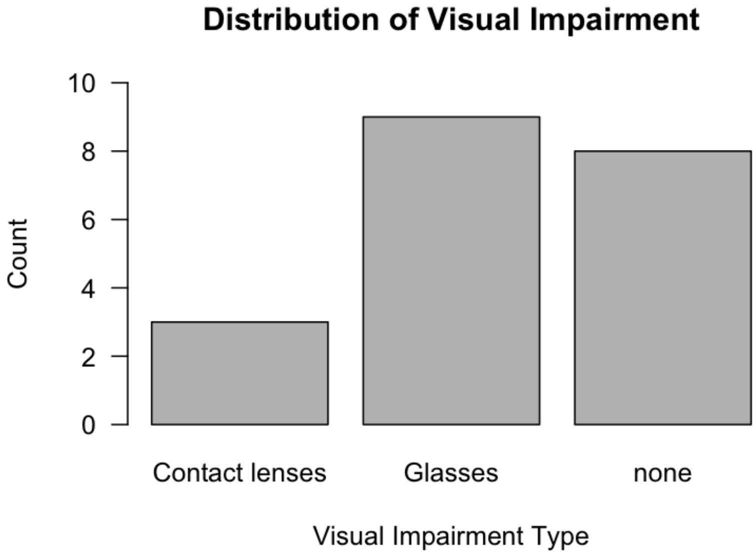


Figure 5.3: Visual impairment of the participants.

6 Results and Discussion

This section starts with a detailed analysis of the data, consisting of NASA-TLX, SUS, free-text answers after each condition and an audio interview after completing all conditions. We then continue with a comprehensive discussion of our findings, where we refer our results to the expected answers for our research questions. Following, we derive design recommendations and provide generalization of our findings and put it in the context of related work again.

6.1 NASA TLX

The NASA-TLX is a widely recognized subjective workload assessment tool developed by the Human Performance Group at NASA's Ames Research Center. Designed to provide a comprehensive measure of perceived workload, NASA-TLX evaluates both the demands of a task and the interactions among those demands from a human operator's perspective. Through a multidimensional rating procedure, this instrument assesses six key aspects: mental demand, physical demand, temporal demand, performance, effort, and frustration level. The abbreviations we use in the following tables are derived from the questionnaire:

- How *mental*ly demanding was the task?
- How *physical*ly demanding was the task?
- How hurried or rushed was the *pace* of the task?
- How *successful* were you in accomplishing what you were asked to do?
- How hard did you have to work to accomplish your level of *performance*?
- How insecure, discouraged, irritated, stressed, and *annoyed* were you?

The scale values reach from 1 to 20 for each aspect. For all aspects except *successful*, a small value is desired. Its flexibility in application across various domains and tasks, combined with its proven reliability, has cemented NASA-TLX's position as a benchmark in human factors research, ergonomics, and usability evaluations.

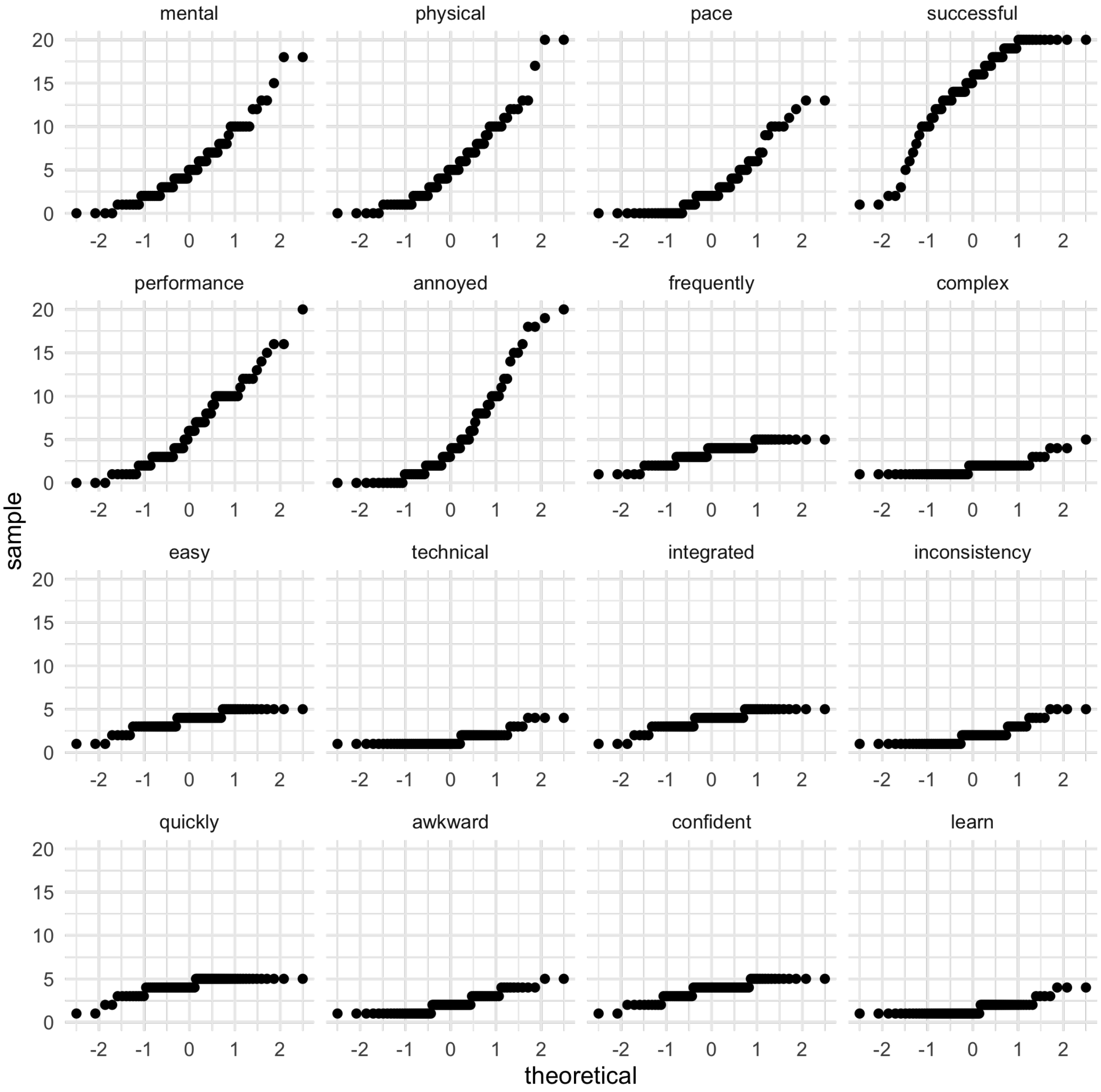


Figure 6.1: QQ-Plots for data normality. The first six graphs show the plots for NASA-TLX, the next then for SUS. The sigmoidal-shaped curves for the NASA-TLX data indicate, that there are more extreme values in the data than would be expected from a normal distribution. This might be a common inherent property of the 0-20 scale and the careful design of an appropriate user interaction.

Variable	**Hand Ray**	**Eye Gaze**	**Head Gaze**	**2D Menu**
Mental	5.60 (4.08)	7.15 (4.30)	5.25 (3.63)	4.05 (3.75)
Physical	9.20 (5.64)	3.60 (3.25)	5.20 (3.52)	4.30 (2.96)
Pace	3.70 (3.76)	2.65 (2.74)	3.55 (3.59)	3.10 (3.82)
Successful	13.40 (6.02)	12.60 (5.56)	15.55 (3.32)	16.70 (3.39)
Performance	7.75 (4.89)	6.75 (4.05)	5.60 (3.94)	5.05 (4.38)
Annoyed	6.95 (6.46)	6.50 (5.49)	3.95 (4.15)	3.10 (3.13)

Table 6.1: The means and standard deviations (in parentheses) for NASA TLX

We started with descriptive statistics, where we computed the means and standard deviation for all variables. The results are given in Table 6.1, while it is important to note that for all values except *successful*, a small value is better while for *successful*, a larger value is better.

Variable	Shapiro-Wilk		Levene's	
	W	p-value	F	p-value
mental	0.922	1.19×10^{-4}	0.1253	0.9449
physical	0.907	2.43×10^{-5}	3.1221	0.03079*
pace	0.838	6.54×10^{-8}	0.6483	0.5864
successful	0.885	3.02×10^{-6}	2.9785	0.03669*
performance	0.939	8.05×10^{-4}	0.4448	0.7217
annoyed	0.854	2.18×10^{-7}	2.3031	0.0837.

Table 6.2: Results of the Shapiro-Wilk and Levene's tests for each variable in NASA-TLX. The p-values of the Shapiro-Wilk test clearly indicate, that with a significance level of $\alpha = 0.05$, we need to reject the Null Hypothesis H_0 that the data is normally distributed. Levene's Test for Homogeneity of Variance uses the convension that the period (.) indicates marginal significance at the 10% level as given by the variable *annoyed* and the asterisk (*) indicates significance at the 5% level as obtained at the variables *physical* and *successful*. As the results suggest violated homogeneity of variance at least for some variables additionally to the rejection of normally distributed data, we need to use a non-parametric variant to test our hypotheses.

After the descriptive statistics, we continued with tests for normality and homogeneity of variance across groups. The corresponding p-values are given in Table 6.2. All of the p-values for Shapiro-Wilk lied below the significance level of $\alpha = 0.05$, so we reject the

Null Hypothesis H_0 that the data is normally distributed. To visualize the deviation from the normal distribution even better and understand its shape, we also fed the data to Quantile-Quantile(QQ)-Plots. Those are shown in Figure 6.1. Regarding the variance homogeneity across groups, we reported the F-values and p-values of Levene's test in Table 6.2 as well. For the variables *physical* and *successful*, a marginal significance at the 5% level was found and for *annoyed*, a marginal significance at the 10% level was found, so evidence of heterogeneity of variances for those variables has been revealed.

Given the results above, there is not sufficient evidence to support the assumptions of normality and homogeneity of variance. Consequently, the non-parametric equivalent of the one-way ANOVA, the Kruskal-Wallis test was employed to find significant differences across the groups. The results, presented in Table 6.3 revealed that for the NASA-TLX variables *physical demand* ("How physically demanding was the task?) and *performance* ("How successful were you in accomplishing what you were asked to do?"), there is a significant difference among the different interaction groups.

Starting with *physical demand*, the Kruskal-Wallis results revealed a statistically meaningful disparity, as evidenced by $\chi^2(3) = 14.678, p < 0.001$. To further investigate the pairwise differences, Dunn's test with Bonferroni correction was conducted. The analysis indicated a significant difference between the `EyeGaze` and `HandRay` groups ($p = 0.0009$), as well as between the `HandRay` and `TwoDMenu` groups ($p = 0.0118$). Other pairwise comparisons did not yield significant results, with adjusted p-values exceeding the 0.05 threshold. The exact Z-values that measure the difference between each two groups in terms of standard deviations and the adjusted p-values after the Bonferroni correction are shown in Section 6.1.

For *performance*, we obtained $\chi^2(3) = 8.0878, p = 0.04$. Post-hoc analysis using Dunn's test with Bonferroni correction was performed to identify the pairwise differences. A significant difference was observed between the `EyeGaze` and `TwoDMenu` groups ($p = 0.0223$). No significant differences were found in other pairwise comparisons as the adjusted p-values were greater than the significance level of 0.05. Here, the exact Z-values that measure the difference between each two groups in terms of standard deviations and the adjusted p-values after the Bonferroni correction are shown in Section 6.1. The corresponding confidence intervals are shown in Figure 6.2.

6.2 SUS

The SUS stands as one of the most established tools for assessing the usability of a wide range of products and systems. Introduced by John Brooke in 1986 [Bro96], this quick, ten-item questionnaire produces a comprehensive gauge of perceived usability. Unlike

Variable	**χ^2-value**	**p-value**
mental	7.0032	0.0718.
physical	14.678	0.002114*
pace	0.82853	0.8426
successful	8.0878	0.04423*
performance	4.6989	0.1952
annoyed	6.358	0.09543.
frequently	6.7745	0.07944.
complex	3.3827	0.3363
easy	7.9231	0.04763*
technical	3.7156	0.2939
integrated	3.1967	0.3623
inconsistency	8.3426	0.03944*
quickly	4.3883	0.2225
awkward	3.71	0.2945
confident	7.5071	0.05738.
learn	6.4353	0.09225.

Table 6.3: Kruskal-Wallis Test for Differences across User Interaction using the convention that the period (.) indicates marginal significance at the 10% level and the asterisk (*) indicates significance at the 5% level. We have four groups tested for each variable, hence the degrees of freedom are always three. Significant differences at the 5% level are indicated at the variables physical, successful, easy and inconsistency.

	EyeGaze	**HandRay**	**HeadGaze**
HandRay	-3.626903 0.0009*		
HeadGaze	-1.461689 0.4315	2.165213 0.0911	
TwoDMenu	-0.744505 1.0000	2.882397 0.0118*	0.717184 1.0000

Table 6.4: Pairwise comparison of *physical demand* by group computed with Dunn's test. Each entry consists of the Z-value and below, the adjusted p-value with Bonferroni correction. The asterisk(*) denotes significance at 5% significance level.

	EyeGaze	**HandRay**	**HeadGaze**
HandRay	-0.721131 1.0000		
HeadGaze	-1.660995 0.2901	-0.939863 1.0000	
TwoDMenu	-2.676048 0.0223*	-1.954916 0.1518	-1.015053 0.9302

Table 6.5: Pairwise comparison of *performance* by group computed with Dunn's test. Each entry consists of the Z-value and below, the adjusted p-value with Bonferroni correction. The asterisk(*) denotes significance at 5% significance level.

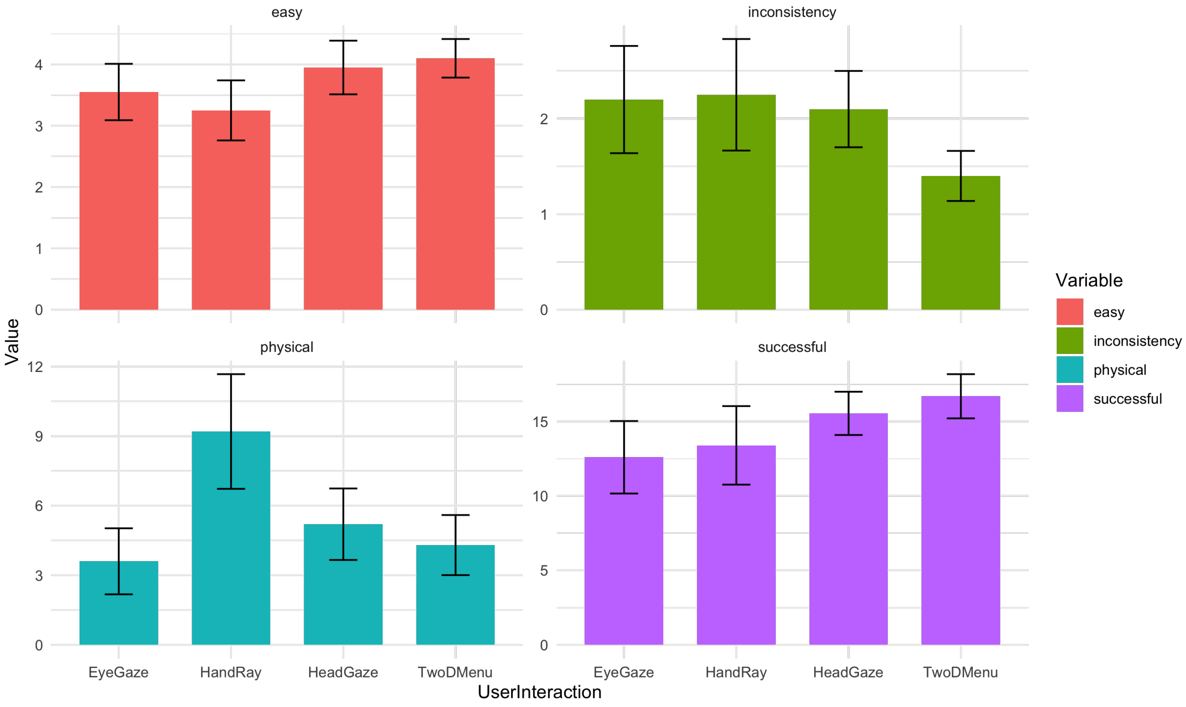


Figure 6.2: Confidence intervals for the variables where Kruskal-Wallis found significant differences among the interaction groups. The plot shows the mean and the error bars depict the 95% confidence intervals.

the scale of NASA-TLX, here the scales for each item reach from 1 to 5. While its simplicity and brevity have contributed to its popularity, the SUS's effectiveness in offering a reliable, high-level snapshot of user satisfaction is its key strength. Deployed in a variety of contexts, from software interfaces to physical products, the SUS provides designers, researchers, and practitioners a standardized method to capture users' experiences and sentiments, making it a cornerstone in usability research and evaluation.

First, we computed the SUS score from the means for each user interaction which is given by

$$\text{SUS Score} = 2.5 \times \left(\sum_{i=1}^{10} x_i \right) \tag{6.1}$$

where

$$x_i = \begin{cases} \text{user response} - 1 & \text{for odd } i \\ 5 - \text{user response} & \text{for even } i \end{cases}$$

.

The following table Table 6.6 depicts the mean SUS scores over all participants, subdivided by the corresponding user interaction method and ordered ascending.

Technique	**SUS Score**
Hand Ray	66.25
Eye Gaze	73.375
Head Gaze	76.625
2D Menu	82.375

Table 6.6: User Interaction Technique SUS Scores, in ascending order

To continue with more descriptive statistics, the means and standard deviation for each variable in the SUS questionnaire is given in Table 6.7.

Variable	Hand Ray	Eye Gaze	Head Gaze	2D Menu
Frequently	2.95 (1.28)	3.75 (1.16)	3.30 (1.22)	3.75 (0.64)
Complex	2.10 (1.21)	1.55 (0.69)	1.50 (0.61)	1.65 (0.59)
Easy	3.25 (1.12)	3.55 (1.05)	3.95 (1.00)	4.10 (0.72)
Technical	1.80 (1.06)	1.65 (0.75)	1.35 (0.75)	1.45 (0.60)
Integrated	3.55 (1.15)	3.55 (1.05)	3.90 (0.97)	4.05 (0.76)
Inconsistency	2.25 (1.33)	2.20 (1.28)	2.10 (0.91)	1.40 (0.60)
Quickly	3.85 (1.09)	4.20 (1.01)	4.40 (0.60)	4.40 (0.88)
Awkward	2.45 (1.36)	2.10 (1.02)	2.30 (1.08)	1.75 (0.79)
Confident	3.40 (1.27)	3.45 (1.00)	3.75 (0.91)	4.20 (0.52)
Learn	1.90 (0.97)	1.65 (0.81)	1.40 (0.50)	1.30 (0.57)

Table 6.7: The means and standard deviations (in parentheses) for SUS

Again, we tested for normality and for homogeneity of variance across groups. The W-value and p-value of Shapiro-Wilk's test as well as F-value and p-value of Levene's test are reported in Table 6.8. For the given Shapiro-Wilk p-values and the significance level of $\alpha = 0.05$, we need to reject the Null Hypothesis H_0 that the data is normally distributed. The deviation from the normal distribution is also visualized in a QQ-Plot (Figure 6.1). Levene's test results in significance at the 5% level for the variables *frequently* and *confident* and significance at the 10% level for the variables *awkward* and *learn*.

Hence, analogously to Section 6.1, Kruskal-Wallis tests with subsequent Dunn's tests with Bonferrioni correction was used to evaluate the data. The Kruskal-Wallis results are reported in Table 6.3. It indicated a significant difference between the groups for the variables *easy* ("I thought the system was easy to use") and *inconsistency* ("I thought there was too much inconsistency in this system"), so Dunn's test was used to further evaluate which groups hold the differences.

Firstly, for the variable *easy*, the result yielded $\chi^2(3) = 7.9231$ with a p-value of 0.05. Subsequent pairwise comparisons were performed using Dunn's test with Bonferroni correction. Analogously to the NASA-TLX tests, Z-value as well as the adjusted p-value for each pairwise comparison is given in Section 6.2. Here, a significant difference was observed between the `TwoDMenu` and `HandRay` groups ($p = 0.0009^{*}$).

Next, we computed Kruskal-Wallis for *inconsistency*. Results are $\chi^2(3) = 8.3426$ with a p-value of 0.04. Using again Dunn's test with Bonferroni correction, we revealed the difference between the groups `TwoDMenu` and `HeadGaze`. Detailed results are given in Section 6.2. Again, the 95% confidence intervals are depicted in Figure 6.2. For a

Variable	Shapiro-Wilk		Levene's	
	W	p-value	F	p-value
frequently	0.898	9.97×10^{-6}	3.2156	0.02748*
complex	0.735	1.02×10^{-10}	1.0175	0.3897
easy	0.876	1.37×10^{-6}	1.3897	0.2524
technical	0.691	1.12×10^{-11}	2.0413	0.1152
integrated	0.864	5.07×10^{-7}	0.6342	0.5953
inconsistency	0.794	3.15×10^{-9}	2.1155	0.1052
quickly	0.768	6.56×10^{-10}	0.4475	0.7198
awkward	0.856	2.69×10^{-7}	2.303	0.08372.
confident	0.867	6.52×10^{-7}	3.9471	0.01132*
learn	0.707	2.42×10^{-11}	2.3867	0.07559.

Table 6.8: Results of the Shapiro-Wilk and Levene's tests for each variable in SUS. Again, it is obvious that for the given Shapiro-Wilk p-values and the significance level of $\alpha = 0.05$, we need to reject the Null Hypothesis H_0 that the data is normally distributed. Levene's Test for Homogeneity of Variance uses the convension that the period (.) indicates marginal significance at the 10% level, which occurred at the variables *awkward* and *learn* and the asterisk (*) indicates significance at the 5% level, which is reflected at the variables *frequently* and *confident*. As the results suggest violated homogeneity of variance at least for these variables additionally to the rejection of normally distributed data, we need to use a non-parametric variant to test our hypotheses.

	EyeGaze	**HandRay**	**HeadGaze**
HandRay	0.766783 1.0000		
HeadGaze	-1.308881 0.5717	-2.075664 0.1138	
TwoDMenu	-1.711888 0.2608	-2.478671 0.0396	-0.403007 1.0000

Table 6.9: Pairwise comparison of *easy* by group computed with Dunn's test. Each entry consists of the Z-value and below, the adjusted p-value with Bonferroni correction. The asterisk(*) denotes significance at 5% significance level.

	EyeGaze	**HandRay**	**HeadGaze**
HandRay	-0.050726		
	1.0000		
HeadGaze	-0.192035	-0.141308	
	1.0000	1.0000	
TwoDMenu	2.271811	2.322537	2.463846
	0.0693	0.0606	0.0412

Table 6.10: Pairwise comparison of *inconsistency* by group computed with Dunn's test. Each entry consists of the Z-value and below, the adjusted p-value with Bonferroni correction. The asterisk(*) denotes significance at 5% significance level.

significance level of 10%, *frequently*, *confident*and *learn*, Kruskal-Wallis also showed significant differences among the groups.

6.3 User-Reported Data and Feedback

Apart from the NASA TLX and SUS data, which work on numerical scales, participants were asked three open-ended qualitative questions, previously defined in Section 5.3:

1. What did you like about this user interaction?
2. What did you dislike about this user interaction?
3. What would you change about this user interaction?

Furthermore, during the audio-recorded interviews, users had the opportunity to elaborate on their ratings. The outcomes of these responses are discussed in this section. When answering the three free-text questions after each interaction, there were multiple statements that repeated several times across the participants. A general remark that occurred for all techniques except the 2D menu one is that people would have liked to have larger targets or a possibility to zoom in. We decided against the zooming possibility, as this would have lead to a drastic loss of quality.

Regarding the hand ray interaction, they mentioned the physical demand, which especially occurred due to the vast movements to aim at the large area. Moreover, many of them had problems with accepting the fact that the ray is not consistent with where they feel like they are really pointing as they would when showing something to someone without AR. Another disadvantage is that the finger tapping technique has to be learned, some participants struggled noteably in learning the finger tap gesture. Nevertheless, it was still referred to as an overall intuitive technique to select objects out of reach.

For eye gaze, the most comment criticism was that people did not like the speech interaction. Some mentioned that they would have preferred a better visual feedback. This is restricted by the fact that in the large space, cursors are due to fast saccades of the eyes much slower and would introduce a delay, as mentioned in Section 4.4. An interesting limitation we found is that for one participant with glasses of -4.5 dioptries, calibration was impossible until she took her glasses off. For such cases, the technique is still not advanced enough to be used without a fallback possibility. Nevertheless, the eye gaze interaction received much positive subjective feedback. Participants referred to it as an "enjoyable experience", "easy to use, no hand gestures need to be learned", "no movements necessary for selecting" and there were several comments like "I felt like I'm in some cool futuristic movie" and "it feels extremely futuristic".

Regarding the head gaze technique, its "high precision", "eas[e] to aim", no need of using hands and the "quick and safe selection of desired targets" have been positively highlighted. From one participant, it was also dubbed as "very futuristic". Drawbacks noted were the slight discomfort experienced when pointing with the head and similar to the eye gaze interaction, the voice input was not well-received.

Moving on to the 2D menu, the subjective user experience and immersion ratings differed from the task load and usability ratings from before. Positive aspects were that it was "easy to target the individual objects", "very easy to hit the buttons and since everything was in one field, it was easy to keep track of all the possibilities and not have to search for them". It was also compared to a tablet and people liked that you can move it around so it will not occlude the shop floor. Negative voices were that "the modern research character is the least", that "you feel like VR/AR could offer more than this established UI" and even "since there was little interaction so far I can then also save myself this, then I would rather sit on the laptop".

6.4 Discussion

After analyzing our collected data using graphs and statistical tests, it is important to relate our findings to the expected answers to our research questions and to related work. We start by the analysis of Expected Answers, continue with deriving Design Recommendations and conclude with a joint discussion.

6.4.1 Analysis of Expected Answers

In this section, we relate the statistical tests shown in Section 6.1 and Section 6.2, as well as the user-reported data and feedback from Section 6.3 to the expected answers to the research questions.

Before analyzing the Expected Answers, it is noticeable that the 2D Menu has been present in nearly every Dunn's test. Only for *Physical Demand*, there was also a significant difference between Eye Gaze and Hand Ray, but still between 2D Menu and Hand Ray as well. It also achieved the highest overall SUS score. Furthermore, it is remarkable that Hand Ray shows the highest variance for almost every variable in SUS and NASA-TLX. We assume that this is caused by the factor that some participants, maybe by chance or by high VR/AR experience, immediately understood the movement, which results in an easy usage. The 2D menu, by contrast got the lowest variance for most of the time. Here, we assume that this is caused by its similarity to an interaction with a tablet and because all buttons are placed within reach for a direct touch. This causes the 2D menu to work very reliably. Another interesting observation is that the Hand Ray interaction had a significant difference in the physical task load compared to two other techniques and got a mean value of 9.2 assigned there. This effect might be caused by the large physical space, where sweeping movements must be executed. An interesting comparison would be the difference to a room-scaled setup with the same distant Hand Ray interaction. In order to refine our first impressions, we continue with a detailed analysis of the Expected Answers to our Research Questions.

EA1.1: The 2D menu will be easiest, fastest and most successful to use, but yield the lowest immersion and UX. Starting with the variables *easy* and *successful*, which we measured quantitatively, Table 6.7 shows that *2D Menu* has the overall highest mean for easy (4.10) compared to head gaze (3.95), eye gaze (3.55) and hand ray (3.25) as well as the overall highest mean for successful (16.7) compared to head gaze (15.55), eye gaze (12.6) and hand ray (13.4) as shown in Table 6.1. For both measures, Kruskal-Wallis also found significant differences across groups, depicted in Table 6.3. The 2D Menu is significantly easier rated than Hand Ray and significantly more successful than eye gaze, as the following pairwise comparison using Dunn's test showed. Detailed pairwise comparison results are also presented in Section 6.2 and Section 6.1. The formulation *fast* is meant as the response time, which is evaluated subjectively in the audio interview. During the interviews, it was very often explicitly mentioned, that the response time in the 2D Menu was the best and that it was the fastest to get used to. The subjective ratings as presented in Section 6.3, however, indicated that immersion and user experience of the 2D Menu is rather disappointing, especially when compared to the excitement about the "futuristic" interaction techniques.

Figure 6.3: NASA-TLX Boxplots for Head Gaze and Eye Gaze

EA1.2: The Hand Ray interaction is easy and intuitive. After the pilot study and due to the common pointing motion, we expected that this interaction will be easy and intuitive. Unfortunately, many participants had problems in performing the gesture as exactly as it is required to be recognized reliably. Their feedback included that it took some time to get used to the movement and a delay of the visual Hand Ray feedback was noticed. The analysis of the collected SUS data showed that for hand ray, the smallest mean for *easy*, which is 3.25 occurred (Table 6.7). Dunn's test in Section 6.2 revealed a significance compared to the 2D menu for *easy* as well. Even though the concept itself was mainly regarded as intuitive during the audio interview, the inconsistency of correctly reacting to the user input has been criticised.

As the pilots had daily experience in VR/AR, we assume that this have caused the difference and have to conclude, that for the application use cases of this research, Handy Ray interaction would be least recommended, because it is hard for AR-unfamiliar users.

EA1.3: Head gaze is more precise than Eye gaze, while eye gaze is physically more comfortable. We define the precision by the *performance* measure of NASA-TLX, asked via "How successful were you in accomplishing what you were asked to do?" Physical effort is also measured via the NASA-TLX questionnaire. For both measures, Kruskal-Wallis found significant differences across the groups (Table 6.3). Focusing on Eye Gaze, Dunn's test found a significant pairwise difference in *physical demand* between Eye Gaze and Hand Ray (Section 6.1) and a significant difference between Eye Gaze and 2D Menu in *successful* (Section 6.1). Watching the means and standard deviations shown in Table 6.1 and the free-text feedback in Section 6.3, Eye Gaze has to be significantly less physically demanding than Hand Ray, but is significantly less successful for task completion than 2D Menu. By contrast, Dunn's test only computed a corrected p-value of 0.43 between Head Gaze and Eye Gaze for physical demand (Section 6.1) and of 0.29 between Head Gaze and Eye Gaze for Section 6.1. Looking into means and standard deviation again (Table 6.1) and taking the TLX boxplots for Eye and Head Gaze into account (Figure 6.3), both parts of the hypothesis tend to be the case, but the ratings are too similar to find a statistical significance. Regarding Eye Gaze's variances in the Boxplot, it appears to be part of a personal preference. Moreover, both interaction techniques are implemented similarly with identical selection techniques.

EA.2: A combination with an other technique and the 2D menu is desired. Regarding the qualitative audio interview after users experienced all techniques, they have been asked whether they would have liked a combination of techniques. Here, about 76.5% of the responses explicitly mentioned that they would have liked to have the 2D Menu always included as an optional fallback technique. This proportion is found to be significantly different from 50% (p-value = 0.04904). We are 95% confident that the true proportion of these responses in the population is between 50.1% and 93.2%.

EA.3: The system is regarded as helpful for its target application. To address this Expected Answer, we posed the question "How helpful could this system be in the use case on a Likert scale from 1-5 with 1 meaning strongly disagree and 5 meaning strongly agree?" during the audio-recorded interview. For the result analysis, we distinguish between three application experts that already took part in the stakeholder interviews described in Chapter 3.

The mean answer of experts was 4.33, the one of all participants was 3.95. A graphical representation that also contains the confidence intervals is given by Figure 6.4. As we are 95% confident that the true value is around four for all participants and also includes five for the experts, but always starts above 3.5, we can say that it is agreed that the system is regarded as helpful for its target application.

6.4.2 Design Recommendations

Our findings allow us to derive three design recommendations on how to interact with large, information-dense physical spaces from a distant viewpoint.

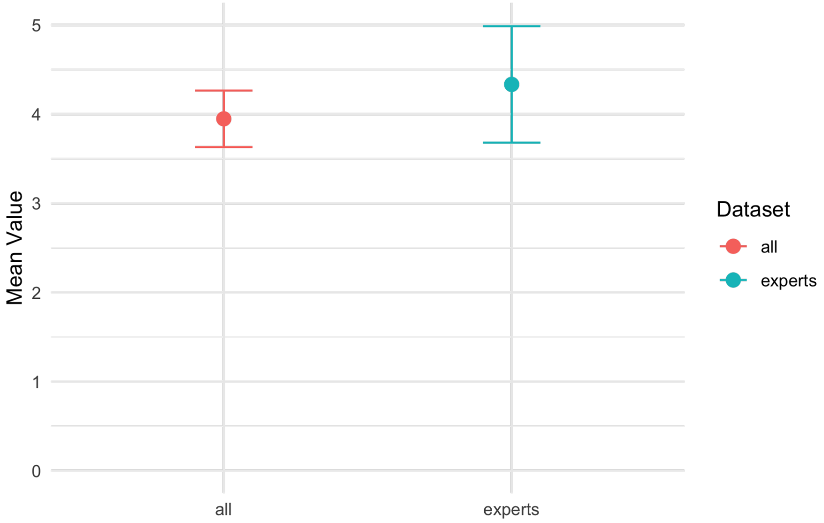


Figure 6.4: Shows the mean and confidence interval of the overall utility rating for the application on a 5-point Likert scale, grouped by everybody and experts.

DR.1 It is recommended to let users open an additional 2D menu whenever they need an overview or a precise, reliable interaction even with objects that are difficult to reach. The possibility to move the menu around and to open it at different locations either via speech or gesture input was evaluated as sensible.

DR.2 To allow for a direct interaction with the target object, which also improves UX and immersion, a gaze input variant is recommended. When precision is crucial and the wearing time is short, we recommend using the Head Gaze method. For a longer usage, the Eye Gaze method is a good alternative as it barely requires any muscle movements. Then, it will also be worth to take the time for a precise calibration, which is not the case for a very quick demo for maybe many people.

DR.3 A hand gesture method for distant interaction can only be recommended if a more robust variant is explored. The current gesture recognition is not robust and responsible enough, especially for inexperienced users.

6.4.3 Implications

To sum it up, we found that the 2D menu is a reliable technique, where response time and acclimatization were very fast. It has the drawback that immersion and UX are not as "futuristic" as the direct interactions, however it is desired to be present as an on-demand fallback method. The Hand Ray interaction, which we expected as easy and intuitive, lead to problems during the user study. We encountered comparably high variances in learning the movement. Moreover, a delay in the visual representation of the hand ray has been mentioned. We can agree to Fang et al. [FCZ+23] that far hand interaction performs worse than direct touch interaction, such as pressing buttons in the menu, and leads to physical fatigue. Our expansion in a very large space amplified the delay of the visual feedback.

To continue with gaze interaction, we got better results in both quantitative measures and open feedback compared to Hand Ray. We did not find a statistically significant difference between Head Gaze and Eye Gaze, but we did find several trends in the feedback. Eye Gaze was often described as "futuristic" and was rated with the lowest physical demand. Previous work often find advantages of the Eye Gaze technique compared to Head Gaze [ŠIR+19], [BRP18], [PRSS16]. When it comes to precision, however, our findings rather agree with Qian et al. [QT17], where Eye Gaze and Head Gaze have been evaluated in a Fitt's law study with targets of different depths. Our collected data also suggests that Head Gaze is more precise and reliable. For any use case that requires a quick demonstration of functionalities in a large, information-dense space, we would rather recommend the usage of Head Gaze, also for the reason that a calibration is always necessary to use Eye Tracking. For a longer usage, Eye Tracking might be favored due to its low physical demand and because calibration then will also take a negligible amount of time.

We conclude that for a user interaction in a large space, gaze selection is favored over distant air tapping. The specific variant of the gaze technique is dependent on the use case. For information-density, it is recommended to give the opportunity of an additional 2D menu, as it gives a complete overview, allows for filtering and via direct touch, its interaction is reliable.
On the whole, our visualization technique and interaction suggestions are regarded as helpful for visitors to understand the structure of such a factory hall. Large stadiums, also for referee decisions, museums, military applications or helicopter flights might be further use cases where our findings can be applied.

7 Summary and Outlook

This chapter summarizes the findings of our research. In the first section, we recap our design considerations, implementation and user study, along with our results and design recommendations. Then we continue with discussing limitations of our setup and conclude with an outlook of what can be done in the future based on our research presented in this thesis.

7.1 Conclusion

In this thesis, we investigated how a large, information-dense physical space can be tracked, how different use cases can be visualized and mainly, what kind of user interaction we can recommend for hands-free interaction with a HMD. We came to the conclusion that marker tracking with multiple QR-Codes along with Space Pins work best. Regarding the visualizations, we decided for information panels to display specific data for certain areas and four further use cases, demonstrating grand important projects as well as safety instructions. The visualizations are characterized by their variance in size, distant, shape complexity and whether they are moving or static. This also leads to varying difficulties for the user interaction techniques.

For those interaction techniques, we implemented four different methods. Those consist of pointing and selecting with arm movements and a hand gesture, using either eye tracking or pointing with the head, combined with speech selection or a hand gesture and, finally, a 2D menu that can be opened via voice or gesture and interacted with by direct touch input. To evaluate task load, usability, user experience and more open feedback, we conducted a user study with a quantified subjective questionnaire followed by a detailed audio-recorded interview.

Results show that the resulted application is regarded as helpful for receiving information and structure of a factory floor. Moreover, the study revealed that gaze methods provide a functional yet immersive interaction. Depending on the duration of the interaction and the required precision, either eye tracking, which is physically less exhausting, provides better immersion but needs careful calibration or pointing with the head, which allows for more precise and reliable selection, is favored. The gestures for interaction with a ray coming from the hand palm, however, turned out to be less reliable, more difficult to learn and physically more exhausting. As it turned out as the easiest, fastest and most successful interaction technique, it always should be possible to access a 2D menu, as it provides an overview and its buttons can be pressed by direct touch interaction. Since this interaction yield the lowest immersion and UX, we recommend to combine it with a gaze technique that suites the specific use case.

7.2 Limitations

While our study provides valuable insights, it's important to note the limitations we encountered during our research. First, out study focuses only on subjective impressions, even though we quantified them with TLX and SUS scores. We could have expanded the study by also including objective quantitative measures like actual response time of the interaction input and how long it takes to try out all use cases with each interaction. A normal distribution of additional data would extend the statistical analysis possibilities. Moreover, it would have been interesting to collect concrete data depending on size, distances, complexity of bounding boxes and static versus moving objects instead of only telling them to go into it at the free-text question and asking during audio interviews.

It is worth noting that training the participants for the hand ray presented challenges. During the course of the experiment, we got better in understanding the problems of inexperienced users and thus, we already addressed previous users' problems when explaining the motion. This might have coursed some variance that later participants rated this interaction better than the first ones. In any case, it can be rated as a general drawback of the interaction technique if the instructor needs to be trained for the perfect explanation for the single optimal motion beforehand. Finally, we note that all participants had an academic background, either as students or as managers of the innovation-driven research campus. This leads to a general openness to new, innovative topics such as AR.

7.3 Outlook

Looking ahead, several extensions can be researched based on the findings of this work on the hardly explored area of information-dense large spaces, viewed from above. Future work can extend this field of research by integrating newly explored interaction techniques that have only been evaluated for other use cases so far. Moreover, a following study can also compare handheld AR with the HMD version or include an ability to zoom in. Another interesting extension is to implement a collaboration functionality where two HMDs are connected, allowing an instructor to directly show areas in a sensible structure and order.

In addition, this work raises many possibilities for extended use cases. The coordinate system can be matched to the one used by transport robots to display their tasks direclty above them and sending them to specific locations. In general, many functional interfaces with the company data can be supplemented. Moreover, a second proof-of-concept study can be conducted that measures task completion time and tests participant's understanding of the shop floor's structure in a between-subject manner. The HMD condition can be compared to an explanation only or to a tablet AR approach as well. It also can be studied how different data visualization affect the utility and usability.

A great improvement can be made when the hardware allows to use a digital twin of the whole factory in real-time. Collisions can be detected precisely and more detailed fitting digital augmentations probably also improve the user experience. If also eye-tracking might get better in the future in providing faster and more reliable calibration, a similar study can be conducted again to see how this shifts the results.

This concludes a nearby suggestion of several extensions that can be made in future work to broaden the practical and theoretical research contribution in this rather unexplored field.

Remarks

In the process of composing this thesis, I enlisted the assistance of OpenAI's language model ChatGPT [Ope21] for some text reviews and refinements. Specifically, I utilized the "can you improve spelling and grammar?: [text]" prompt to solicit suggestions on enhancing the clarity and coherence of my prose. While the feedback provided informed some of the phraseology employed in this document, I have taken meticulous care to ensure that the originality of the ideas and the integrity of the content remained unaltered. The essence of the research, the analysis, and the conclusions drawn herein are entirely my own, with the AI serving merely as a tool to polish the articulation of these concepts.

Bibliography

[AS11] C. Arth, D. Schmalstieg. “Challenges of large-scale augmented reality on smartphones.” In: *Graz University of Technology, Graz* 1.4 (2011) (cit. on p. 17).

[BC87] R. A. Becker, W. S. Cleveland. “Brushing Scatterplots.” In: *Technometrics* 29.2 (1987), pp. 127–142. doi: 10.1080/00401706.1987.10488204. eprint: https://www.tandfonline.com/doi/pdf/10.1080/00401706.1987.10488204. url: https://www.tandfonline.com/doi/abs/10.1080/00401706.1987.10488204 (cit. on pp. 15, 21).

[Bil21] M. Billinghurst. “Grand Challenges for Augmented Reality.” In: *Frontiers in Virtual Reality* 2 (2021). issn: 2673-4192. doi: 10.3389/frvir.2021.578080. url: https://www.frontiersin.org/articles/10.3389/frvir.2021.578080 (cit. on p. 15).

[Bro96] J. Brooke. “Sus: a “quick and dirty’usability.” In: *Usability evaluation in industry* 189.3 (1996), pp. 189–194 (cit. on pp. 43, 52).

[BRP18] J. Blattgerste, P. Renner, T. Pfeiffer. “Advantages of Eye-Gaze over Head-Gaze-Based Selection in Virtual and Augmented Reality under Varying Field of Views.” In: *Proceedings of the Workshop on Communication by Gaze Interaction*. COGAIN ’18. Warsaw, Poland: Association for Computing Machinery, 2018. isbn: 9781450357906. doi: 10.1145/3206343.3206349. url: https://doi.org/10.1145/3206343.3206349 (cit. on pp. 15, 21, 40, 64).

[CHS+20] W. P. Chan, G. Hanks, M. Sakr, T. Zuo, H. Machiel Van der Loos, E. Croft. “An Augmented Reality Human-Robot Physical Collaboration Interface Design for Shared, Large-Scale, Labour-Intensive Manufacturing Tasks.” In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2020, pp. 11308–11313. doi: 10.1109/IROS45743.2020.9341119 (cit. on pp. 15, 19).

[DB06] H. Durrant-Whyte, T. Bailey. “Simultaneous localization and mapping: part I.” In: *IEEE Robotics Automation Magazine* 13.2 (2006), pp. 99–110. doi: 10.1109/MRA.2006.1638022 (cit. on pp. 17, 33).

[DCHR22] N.-D. Duong, C. Cutullic, J.-M. Henaff, J. Royan. "AR Cloud: Towards Collaborative Augmented Reality at a Large-Scale." In: *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 2022, pp. 733–738. DOI: 10.1109/ISMAR-Adjunct57072.2022.00155 (cit. on p. 19).

[FCZ+23] W. Fang, L. Chen, T. Zhang, C. Chen, Z. Teng, L. Wang. "Head-mounted display augmented reality in manufacturing: A systematic review." In: *Robotics and Computer-Integrated Manufacturing* 83 (2023), p. 102567. ISSN: 0736-5845. DOI: https://doi.org/10.1016/j.rcim.2023.102567. URL: https://www.sciencedirect.com/science/article/pii/S0736584523000431 (cit. on pp. 15, 19, 38, 64).

[FPS+20] T. Feigl, A. Porada, S. Steiner, C. Löffler, C. Mutschler, M. Philippsen. "Localization Limitations of ARCore, ARKit, and Hololens in Dynamic Large-scale Industry Environments." In: *VISIGRAPP (1: GRAPP)*. 2020, pp. 307–318 (cit. on pp. 15, 18).

[GL17] M. Garon, J.-F. Lalonde. "Deep 6-DOF Tracking." In: *IEEE Transactions on Visualization and Computer Graphics* 23.11 (2017), pp. 2410–2418. DOI: 10.1109/TVCG.2017.2734599 (cit. on p. 19).

[GOT13] A. Gherghina, A.-C. Olteanu, N. Tapus. "A marker-based augmented reality system for mobile devices." In: *2013 11th RoEduNet International Conference*. 2013, pp. 1–6. DOI: 10.1109/RoEduNet.2013.6511731 (cit. on p. 17).

[HCL+20] P. Hübner, K. Clintworth, Q. Liu, M. Weinmann, S. Wursthorn. "Evaluation of HoloLens Tracking and Depth Sensing for Indoor Mapping Applications." In: *Sensors* 20.4 (2020). ISSN: 1424-8220. DOI: 10.3390/s20041021. URL: https://www.mdpi.com/1424-8220/20/4/1021 (cit. on pp. 19, 33).

[HDP11] F. Hülsmann, T. Dankert, T. Pfeiffer. "Comparing gaze-based and manual interaction in a fast-paced gaming task in virtual reality." In: *Proceedings of the Workshop Virtuelle & Erweiterte Realität 2011*. 2011 (cit. on p. 22).

[HMG18] N. Hube, M. Müller, R. Groh. "Facilitating Exploration on Exhibitions with Augmented Reality." In: *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*. AVI '18. Castiglione della Pescaia, Grosseto, Italy: Association for Computing Machinery, 2018. ISBN: 9781450356169. DOI: 10.1145/3206505.3206585. URL: https://doi.org/10.1145/3206505.3206585 (cit. on pp. 15, 20).

[HRM13] J. Hales, D. Rozado, D. Mardanbegi. “Interacting with objects in the environment by gaze and hand gestures.” In: *Proceedings of the 3rd international workshop on pervasive eye tracking and mobile eye-based interaction*. 2013, pp. 1–9 (cit. on p. 21).

[HWW18] P. Hübner, M. Weinmann, S. Wursthorn. “Marker-based localization of the microsoft hololens in building models.” In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLII-1 (2018), pp. 195–202. DOI: 10.5194/isprs-archives-XLII-1-195-2018. URL: https://isprs-archives.copernicus.org/articles/XLII-1/195/2018/ (cit. on pp. 15, 18, 27).

[JBPH18] M. Juraschek, L. Büth, G. Posselt, C. Herrmann. “Mixed Reality in Learning Factories.” In: *Procedia Manufacturing* 23 (2018). “Advanced Engineering Education Training for Manufacturing Innovation”8th CIRP Sponsored Conference on Learning Factories (CLF 2018), pp. 153–158. ISSN: 2351-9789. DOI: https://doi.org/10.1016/j.promfg.2018.04.009. URL: https://www.sciencedirect.com/science/article/pii/S2351978918304815 (cit. on pp. 15, 19).

[LRP+22] M. N. Lystbæk, P. Rosenberg, K. Pfeuffer, J. E. Grønbæk, H. Gellersen. “Gaze-Hand Alignment: Combining Eye Gaze and Mid-Air Pointing for Interacting with Menus in Augmented Reality.” In: *Proc. ACM Hum.-Comput. Interact.* 6.ETRA (May 2022). DOI: 10.1145/3530886. URL: https://doi.org/10.1145/3530886 (cit. on pp. 22, 23).

[MAN+20] B. Marques, J. Alves, M. Neves, I. Justo, A. Santos, R. Rainho, R. Maio, D. Costa, C. Ferreira, P. Dias, B. S. Santos. “Interaction with Virtual Content Using Augmented Reality: A User Study in Assembly Procedures.” In: *Proc. ACM Hum.-Comput. Interact.* 4.ISS (Nov. 2020). DOI: 10.1145/3427324. URL: https://doi.org/10.1145/3427324 (cit. on p. 21).

[Mic] Microsoft Corporation. *HoloLens2*. URL: https://www.microsoft.com/de-de/hololens/hardware (cit. on p. 33).

[MSP+11] M. R. Marner, R. T. Smith, S. R. Porter, M. M. Broecker, B. Close, B. H. Thomas. “Large Scale Spatial Augmented Reality for Design and Prototyping.” In: *Handbook of Augmented Reality*. Ed. by B. Furht. New York, NY: Springer New York, 2011, pp. 231–254. ISBN: 978-1-4614-0064-6. DOI: 10.1007/978-1-4614-0064-6_10. URL: https://doi.org/10.1007/978-1-4614-0064-6_10 (cit. on p. 17).

[MSWT14] M. R. Marner, R. T. Smith, J. A. Walsh, B. H. Thomas. “Spatial User Interfaces for Large-Scale Projector-Based Augmented Reality.” In: *IEEE Computer Graphics and Applications* 34.6 (2014), pp. 74–82. DOI: 10.1109/MCG.2014.117 (cit. on p. 17).

[Ope21] OpenAI. *Chatgpt*. Language model developed by OpenAI. 2021. URL: https://openai.com/ (cit. on p. 69).

[OSK12] T. Oskiper, S. Samarasekera, R. Kumar. "Multi-sensor navigation algorithm using monocular camera, IMU and GPS for large scale augmented reality." In: *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2012, pp. 71–80. DOI: 10.1109/ISMAR.2012.6402541 (cit. on p. 20).

[PACG14] K. Pfeuffer, J. Alexander, M. K. Chong, H. Gellersen. "Gaze-Touch: Combining Gaze with Multi-Touch for Interaction on the Same Surface." In: *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*. UIST '14. Honolulu, Hawaii, USA: Association for Computing Machinery, 2014, pp. 509–518. ISBN: 9781450330695. DOI: 10.1145/2642918.2647397. URL: https://doi.org/10.1145/2642918.2647397 (cit. on p. 21).

[PLC08] H. M. Park, S. H. Lee, J. S. Choi. "Wearable augmented reality system using gaze interaction." In: *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*. 2008, pp. 175–176. DOI: 10.1109/ISMAR.2008.4637353 (cit. on p. 21).

[PMMG17] K. Pfeuffer, B. Mayer, D. Mardanbegi, H. Gellersen. "Gaze + Pinch Interaction in Virtual Reality." In: *Proceedings of the 5th Symposium on Spatial User Interaction*. SUI '17. Brighton, United Kingdom: Association for Computing Machinery, 2017, pp. 99–108. ISBN: 9781450354868. DOI: 10.1145/3131277.3132180. URL: https://doi.org/10.1145/3131277.3132180 (cit. on pp. 15, 21).

[PRSS16] O. Palinko, F. Rea, G. Sandini, A. Sciutti. "Robot reading human gaze: Why eye tracking is better than head tracking for human-robot collaboration." In: *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2016, pp. 5048–5054. DOI: 10.1109/IROS.2016.7759741 (cit. on pp. 15, 21, 40, 64).

[QT17] Y. Y. Qian, R. J. Teather. "The Eyes Don't Have It: An Empirical Comparison of Head-Based and Eye-Based Selection in Virtual Reality." In: *Proceedings of the 5th Symposium on Spatial User Interaction*. SUI '17. Brighton, United Kingdom: Association for Computing Machinery, 2017, pp. 91–98. ISBN: 9781450354868. DOI: 10.1145/3131277.3132182. URL: https://doi.org/10.1145/3131277.3132182 (cit. on pp. 15, 22, 23, 40, 64).

[RSP+19] D. C. Rompapas, C. Sandor, A. Plopski, D. Saakes, J. Shin, T. Taketomi, H. Kato. "Towards large scale high fidelity collaborative augmented reality." In: *Computers Graphics* 84 (2019), pp. 24–41. ISSN: 0097-8493.

DOI: https://doi.org/10.1016/j.cag.2019.08.007. URL: https://www.sciencedirect.com/science/article/pii/S0097849319301359 (cit. on p. 20).

[SDS+22] P.-E. Sarlin, M. Dusmanu, J. L. Schönberger, P. Speciale, L. Gruber, V. Larsson, O. Miksik, M. Pollefeys. "LaMAR: Benchmarking Localization and Mapping for Augmented Reality." In: *Computer Vision – ECCV 2022*. Ed. by S. Avidan, G. Brostow, M. Cissé, G. M. Farinella, T. Hassner. Cham: Springer Nature Switzerland, 2022, pp. 686–704. ISBN: 978-3-031-20071-7 (cit. on p. 19).

[Shn96] B. Shneiderman. "The eyes have it: a task by data type taxonomy for information visualizations." In: *Proceedings 1996 IEEE Symposium on Visual Languages*. 1996, pp. 336–343. DOI: 10.1109/VL.1996.545307 (cit. on pp. 16, 20).

[ŠIR+19] O. Špakov, H. Istance, K.-J. Räihä, T. Viitanen, H. Siirtola. "Eye Gaze and Head Gaze in Collaborative Games." In: *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*. ETRA '19. Denver, Colorado: Association for Computing Machinery, 2019. ISBN: 9781450367097. DOI: 10.1145/3317959.3321489. URL: https://doi.org/10.1145/3317959.3321489 (cit. on pp. 15, 21, 40, 64).

[SJ00] L. E. Sibert, R. J. K. Jacob. "Evaluation of Eye Gaze Interaction." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '00. The Hague, The Netherlands: Association for Computing Machinery, 2000, pp. 281–288. ISBN: 1581132166. DOI: 10.1145/332040.332445. URL: https://doi.org/10.1145/332040.332445 (cit. on p. 22).

[SLB11] D. Schmalstieg, T. Langlotz, M. Billinghurst. "Augmented Reality 2.0." In: *Virtual Realities: Dagstuhl Seminar 2008*. Ed. by G. Brunnett, S. Coquillart, G. Welch. Vienna: Springer Vienna, 2011, pp. 13–37. ISBN: 978-3-211-99178-7. DOI: 10.1007/978-3-211-99178-7_2. URL: https://doi.org/10.1007/978-3-211-99178-7_2 (cit. on p. 17).

[SWQ+23] R. Shi, Y. Wei, X. Qin, P. Hui, H.-N. Liang. "Exploring Gaze-Assisted and Hand-Based Region Selection in Augmented Reality." In: *Proc. ACM Hum.-Comput. Interact.* 7.ETRA (May 2023). DOI: 10.1145/3591129. URL: https://doi.org/10.1145/3591129 (cit. on p. 21).

[TF22] S. Teruggi, F. Fassi. "Mixed reality content alignment in monumental environments." In: *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 43 (2022), pp. 901–908 (cit. on pp. 15, 18, 34).

[WLM+23] U. Wagner, M. N. Lystbæk, P. Manakhov, J. E. S. Grønbæk, K. Pfeuffer, H. Gellersen. "A Fitts' Law Study of Gaze-Hand Alignment for Selection in 3D User Interfaces." In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI '23. Hamburg, Germany: Association for Computing Machinery, 2023. ISBN: 9781450394215. DOI: 10.1145/3544548.3581423. URL: https://doi.org/10.1145/3544548.3581423 (cit. on pp. 22, 23, 38).

[YHC+10] B. Yoo, J.-J. Han, C. Choi, K. Yi, S. Suh, D. Park, C. Kim. "3D User Interface Combining Gaze and Hand Gestures for Large-Scale Display." In: *CHI '10 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '10. Atlanta, Georgia, USA: Association for Computing Machinery, 2010, pp. 3709–3714. ISBN: 9781605589305. DOI: 10.1145/1753846.1754043. URL: https://doi.org/10.1145/1753846.1754043 (cit. on pp. 16, 22, 23).

[YHNI20] C. S. Yusof, N. A. A. Halim, M. N. A. Nor'a, A. W. Ismail. "Finger-Ray Interaction using Real Hand in Handheld Augmented Reality Interface." In: *IOP Conference Series: Materials Science and Engineering* 979.1 (Nov. 2020), p. 012009. DOI: 10.1088/1757-899X/979/1/012009. URL: https://dx.doi.org/10.1088/1757-899X/979/1/012009 (cit. on pp. 16, 22).

[ZSSB15] Y. Zhang, S. Stellmach, A. Sellen, A. Blake. "The Costs and Benefits of Combining Gaze and Hand Gestures for Remote Interaction." In: *Human-Computer Interaction – INTERACT 2015*. Ed. by J. Abascal, S. Barbosa, M. Fetter, T. Gross, P. Palanque, M. Winckler. Cham: Springer International Publishing, 2015, pp. 570–577. ISBN: 978-3-319-22698-9 (cit. on pp. 22, 23).

All links were last followed on October 11, 2023.

A Additional Material

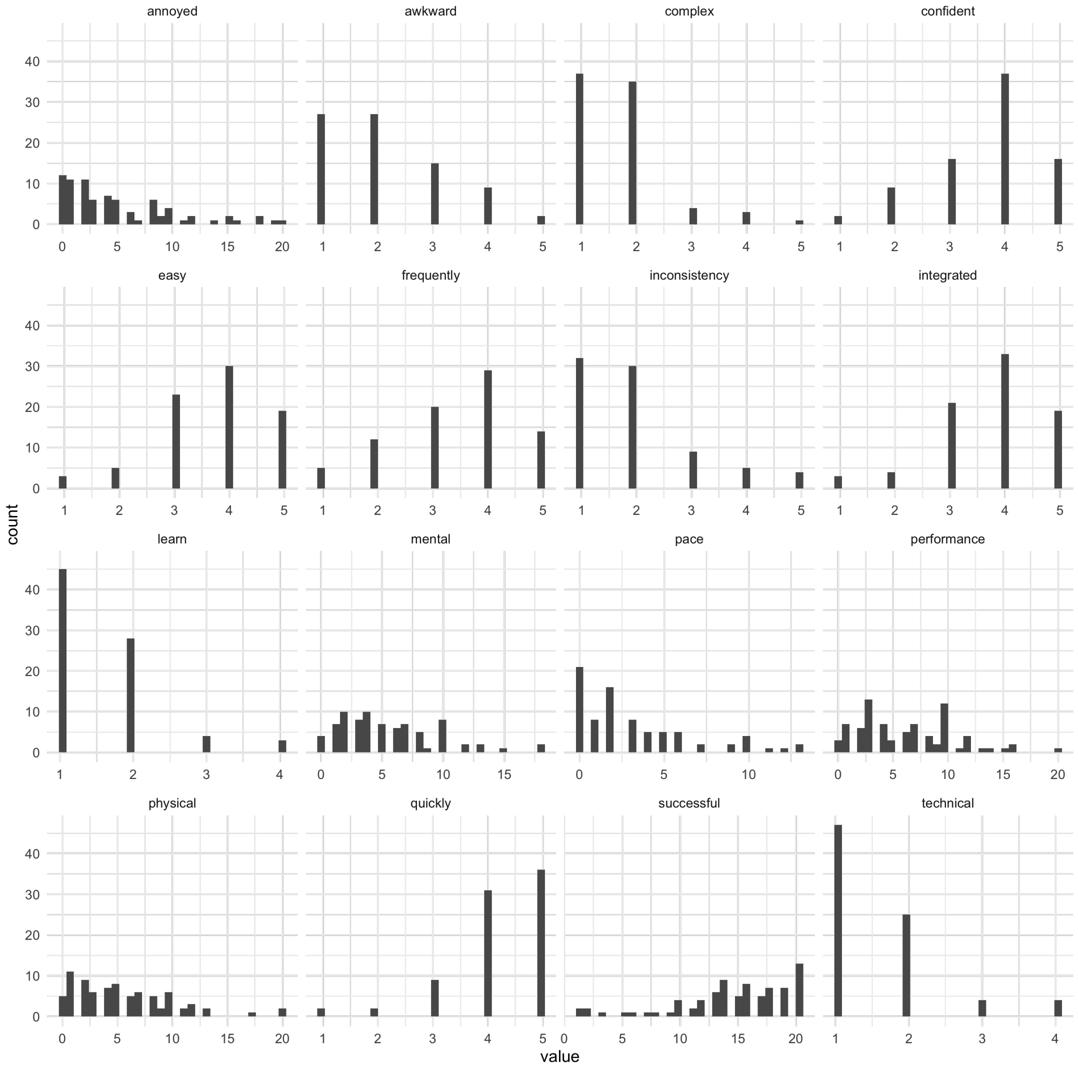
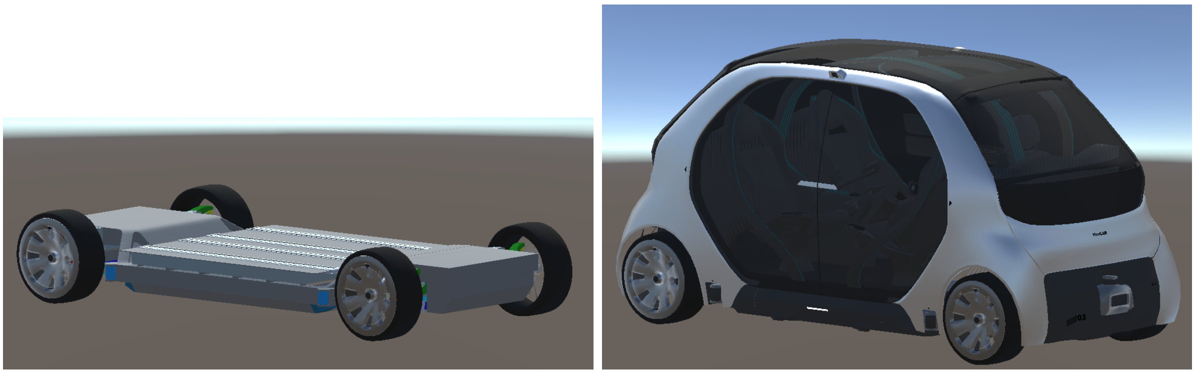


Figure A.1: Histograms for data normality



(a)

(b)

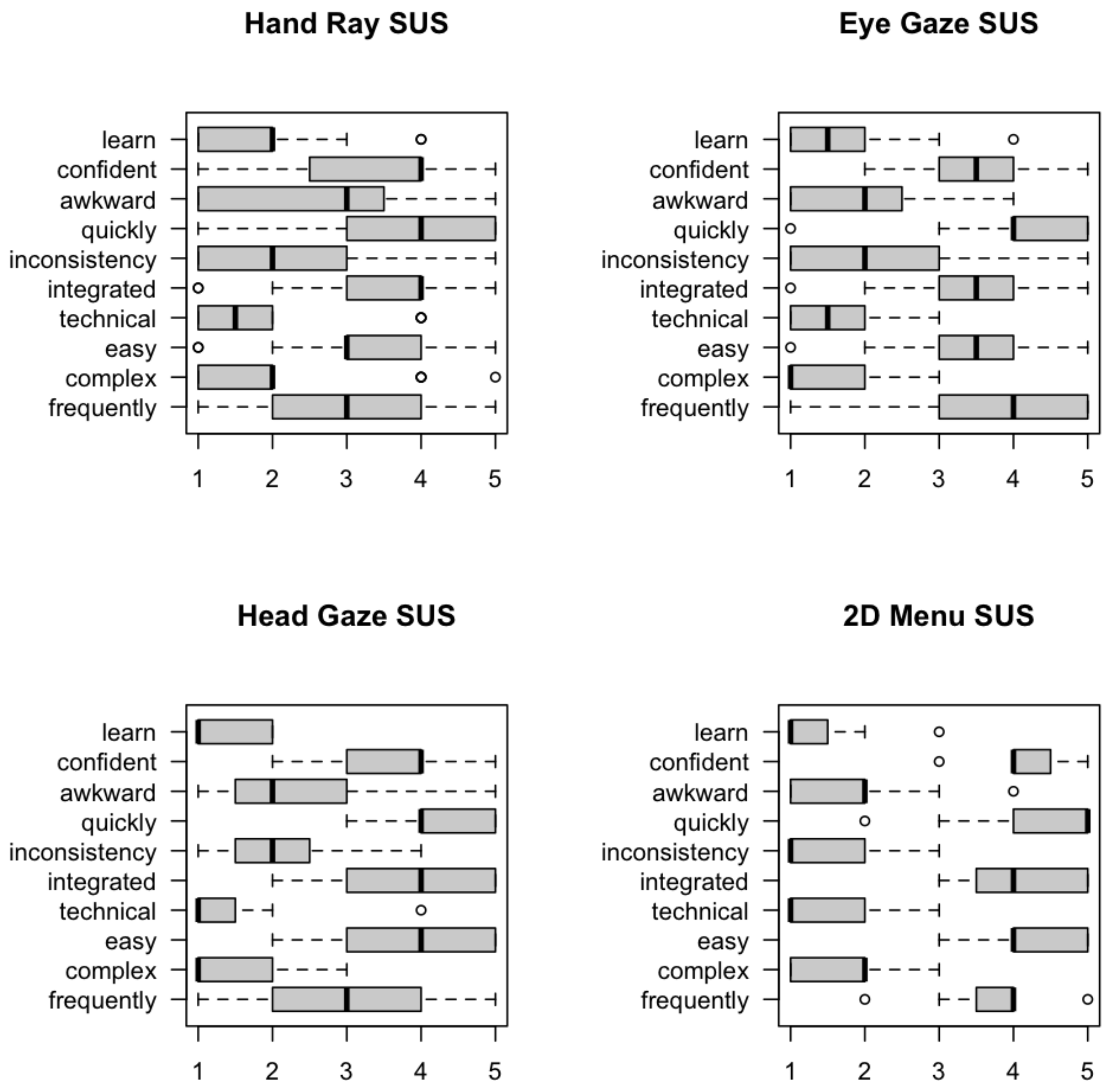
Figure A.2: Rolling Chassis and FlexCAR 3D models

Figure A.3: Tooltip for 5G Traverse

Figure A.4: Company Info Panels, shows occlusion problems

Audio Interview Questions
- How good was the response time of the user interaction? - Was it intuitive? - Did it respond the way you imagined? - How long did it take you to get used to it? - In which order would you rank the interaction techniques? Why? - Rank likeability of each on Likert Scale from 1-5 - Would you like a combination of different actions? - What do you think of the idea of such a visualisation in general? - What use cases could you imagine for it? - Do you feel like you can keep the information better with the AR tool? - Do you have any other suggestions/input for large information-dense space visualization/interaction? - How helpful could this system be in the use case on a Likert scale from 1-5? - Do you see problems in the application in normal working life?

Table A.1: List of Audio Interview Questions


Hand Ray SUS

Eye Gaze SUS

Head Gaze SUS

2D Menu SUS

Figure A.5: Boxplots for SUS

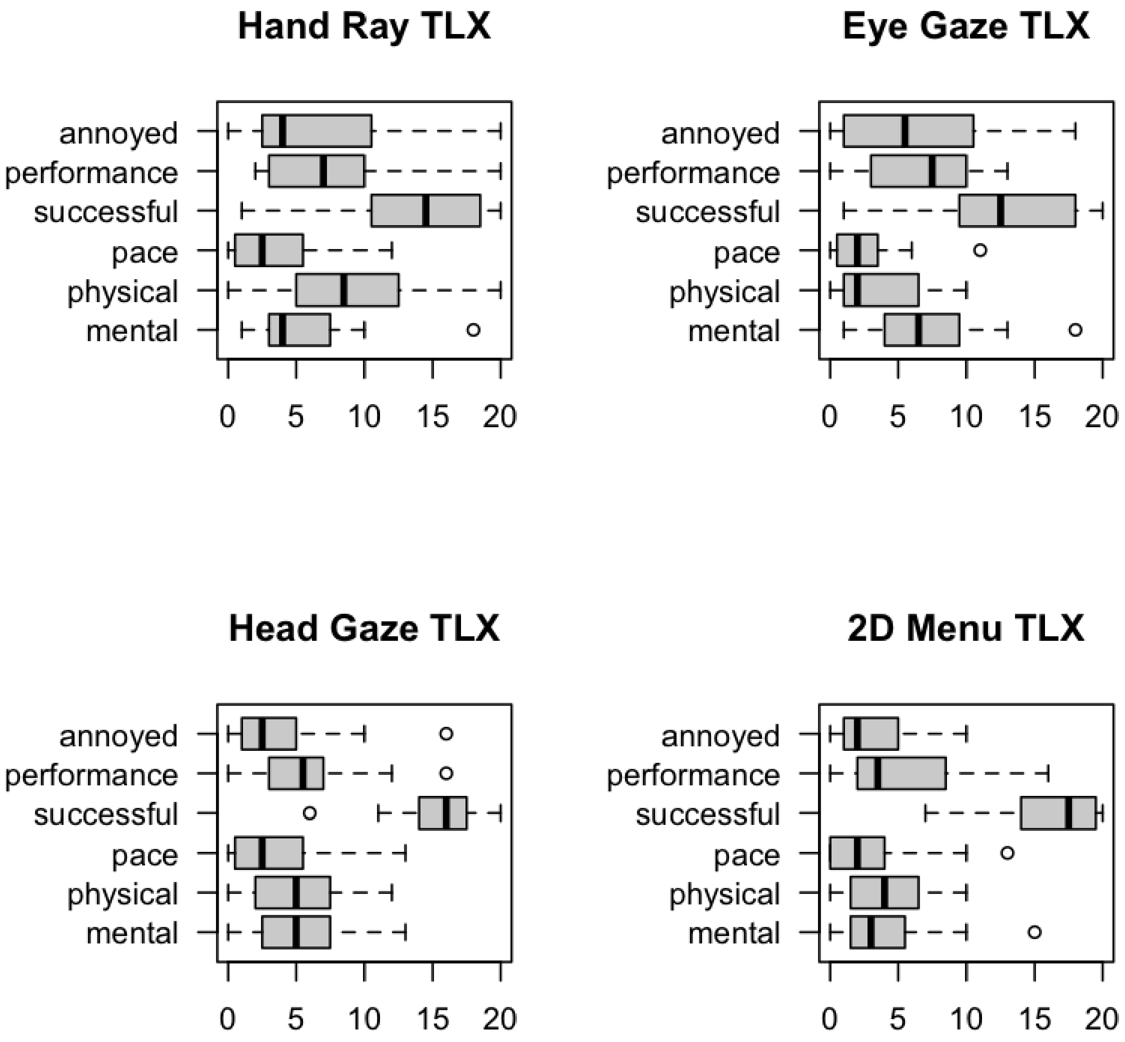


Figure A.6: Boxplots for TLX