Overcoming the problem of uncertain tracking errors in an AR navigation application

Bachelor thesis to obtain a bachelor's degree in the
Bachelor of Science Program "Computer Science"
Taught at the Faculty of Computer Science and Engineering Science of
TH Köln – University of Applied Sciences

Submitted by Oliver Scheibert
Matriculation No. 11112394
Contact scheibert.oliver@gmail.com

First Examiner Prof. Dr. Martin Eisemann
Second Examiner Prof. Dr. Gerhard Hartmann

Gummersbach, 15.09.2020

Table of Contents

Table of Figures

Table of Tables

1 Abstract

This thesis proposes four different Augmented Reality (AR) display techniques that can overcome the problem of uncertain tracking errors. Where most other papers focus on limiting the amount of error that a user tracking system generates, this paper proposes visualizations that limit their impact instead. For this, four different display techniques have been developed and evaluated using criteria found in related literate. Two of the four proposed solutions consistently ranked better within these criteria. The first, Bending Words, uses a mix between spatially registered information and turn instructions to achieve that result. The second, WIM, mostly relies on context information of the building instead.

2 Introduction

User tracking systems will always carry some degree of error that accumulates over time [1]. Until this error is not within very small margins, user experience can suffer from having to manually compensate for them. The goal of this work is to find an AR interface (display technique), that can provide useful navigation information when used with such an imperfect tracking solution.

Augmented reality can provide location-aware user experience by overlaying spatially registered, digital information on a screen to present real-time interaction with the physical and virtual environments simultaneously [2], [3]. This can be especially useful for pedestrian navigation. With recent advancements in user tracking technologies and modern smartphones having sufficient processing power, indoor navigation has become more and more available [4].

However, the technology still has not reached a point where it can be depended on for an extended amount of time [5]–[7]. According to McIntyre et al., the problem of accumulating errors within a tracking system will not be solved in the near future [8]. Relying on such a system can then result in digital objects being too far off their supposed position, making them less useful the further off they are. A work by Pankratz et al. approached this challenge by increasing the user's awareness of the tracking imperfections using different display techniques [9]. They proposed using 3D arrows that can change in color and shape, a digital avatar that guides the user with a varying display of confidence, and a colorful cloud that also changes color and spreads around the navigation path. The result of their own study suggests that "while error visualization has the potential to improve AR navigation systems, it is difficult to find suitable visualizations".

This work builds upon that idea and proposes Display Techniques that instead of increasing the awareness of tracking errors, reduce their impact instead. It does not focus on the technical aspect of localization systems but proposes a solution that utilizes the ability of humans to compensate for them. The criteria and display techniques listed are a first step towards research in this direction.

This thesis first defines the problem of uncertain tracking errors in the context of user tracking systems. Then, Display Techniques are designed using visualization categories defined in chapter 5.1 and implemented within a previously developed navigation application for the Campus Gummersbach. Study 2 evaluates these techniques based on several criteria found in related literature. Due to the current limitation of physical contact in place to prevent the spread of the current Covid-19 pandemic, one of the two studies proposed here only provides an outline for performing such a study in the future.

3 Related Work

Pedestrian navigation has advanced from outdoor solutions relying solely on GPS to taking advantage of more complex localization systems that can be used indoors [10]. However, localizing a user within a building requires higher accuracy than previously needed. Especially augmented reality systems rely on this accuracy to correctly display information. A wrongly placed arrow can point towards the wrong intersection, or a "you have reached your goal" message that is displayed at the wrong time can be confusing.

There is still no generally accepted solution for localization systems [4]. But to help understand the current state of the technology, a study by Adler et al. analyzed and categorizes 183 indoor localization papers published between 2010 and 2014 [11]. These categories are; 1) The underlying user tracking system (chapter 4) or technology, 2) how the localization system is evaluated within the paper itself, 3) the reference material that is used as the ground truth, 4) which metrics are used in the evaluation (such as accuracy or sample size), and 5) if the author established a baseline for their work and argue an improvement over it. These categories help to accurately reproduce the findings of each paper. And reproducibility is a vital part of scientific research where independent researchers validate works using data provided by such categories. For example, it is important to have a previous method that was improved on, to evaluate a work, thus providing a baseline.

The following paragraphs provide this baseline for this thesis. They describe current approaches in tracking technologies that are discussed further in section 4, provide insight into AR interface design possibilities, and explain user's behavior when using them.
The vision-based localization system used by Kim and Jun [12] consists of two parts: a head-mounted display (HMD) and a remote PC. The environment has to be scanned using a camera, that sends the video to the remote PC. It detects special identifiers (markers and color information) within the feed and stores that information in a location dictionary. During use, the HMD will send its own video the PC which then detects markers and matches image sequences to recognize the user's location to provide them with navigation information. An alternative to using image sequencing techniques is the approach by Liu et al. [13]. By previously scanning the geomagnetic fingerprint of a building, it is possible to use the magnetometer to locate a user within the building. This geomagnetic map can achieve an accuracy of 2m-6m.

The user's perception of this range of accuracy is important. If the perceived tracking accuracy is higher than the actual accuracy of the localization system, confusing interactions might occur. To decrease such an impact of inaccurate tracking solutions, Pankratz et al. introduced a simple concept of using different visualizations that automatically adjust for the rate of tracking errors [9]. To develop them, they summarized the visualizations into three categories; 1) Discrete information which describes navigation hints as discontinuous steps, 2) Continuous information that shows uninterrupted hints, and 3) Guiding information represented by a virtual entity that guides the user towards their destination and only shows the next few meters of the path. An evaluation of the different techniques showed that the location of the visualization was more important than the appearance. Therefore, this thesis will also focus more on how and when to utilize different visualizations instead of their exact appearance.

While this thesis focuses on smartphones to implement and evaluate different display techniques, other types of devices can be used. For example, Head-Mounted-Displays (HMD) can be used free-handed and are always within view. This allows for a different design approach for each device. A comparative analysis by Rehman et al. looks at the technical and human factors that influence design decisions on both of these platforms [10]. Factors such as: “perceived accuracy, navigation time, subjective comfort, subjective workload, and route memory retention.” The advantage of HMDs is that the inbuilt sensors track the head orientation and position without user input. Whereas handheld devices need to be held at a particular orientation and position for proper navigation view, requiring a higher cognitive and physical demand on the user. Their system uses a giant green-colored arrow in combination with auditory and text-based instructions, such as “turn left” or “go straight”, to guide the users.

Another extensive study of interacting with AR interfaces by Möller et al. combines an AR interface and a VR panorama view to accommodate for differences in tracking errors and comfort when using the device [4]. One key aspect of choosing between different display methods is the typical position that the smartphone has to be held when interacting with the application. Most vision-based approaches require a high quality of visual references that can oftentimes only be found when pointing the camera forward. However, a more typical orientation when using the phone is at a 45° angle downwards. A higher angle can be inconvenient over time or seem unnatural to use. A good UI system might be inclined to motivate the user to change the orientation of the device regularly, or when the error rate increases, to improve localization accuracy when needed. To do that, Möller et al. chose to highlight points of interest (POI) within the AR video feed using object recognition techniques that don’t require additional markers. To not distract the user by showing too many animations and to hide the exact degree of tracking error, they chose to color the area of the POI with blurry edges. These points of interest can include posters, showcases, doors, elevator controls, etc. In addition to the highlight feature, the AR view also superimposes a directional arrow on top of the video and the distance to the next turn. When the tracking accuracy drops too much or when the device is held at a downward angle, the VR panorama view is displayed. It shows the surroundings which are rendered from a prerecorded image-set that is also used for localizing the user. Since it shows a prerendered scene, errors in the position of navigation information are less noticeable. The user can locate themselves by comparing landmarks within the panorama with their surrounding view.

4 Limitations of User Tracking Systems

4.1 Localization

There are several different user tracking systems that use a variety of sensors to track a users' position, or more precisely, the position of the device that they use. These systems rely on physical properties of either signals, such as their speed traveling through the air, or otherwise measurable forces, such as the earth's magnetic field. Even light can be considered as a very fast signal which is measured by calculating the angle of light hitting the camera sensor to map a 3D scene into a 2D image [14]. There have been many new approaches that advanced the industry in recent years: New statistical methods that combine several different sensor information, such as Wi-Fi and magnetic signals [15]. Or using a priori knowledge, such as a map of the environment, to further increase the accuracy of these models [7], [16].

Despite the constant improvements in localization techniques over time, creating a perfect solution is almost impossible [17]. "The research community has not converged to a single, widely accepted solution that can achieve the desired accuracy at the required cost", so Lymberopoulos et al. They showed that top-performing solutions often take more than 2 years to develop and take 5 years for a solution that can consistently provide the same localization error over the tracking space.

A high degree of localization accuracy is important when guiding a user. The less information is known about the position, the more cognitive processing power of the user is needed to correct the system's misalignment. This effect can be especially prevalent in AR applications where the positioning of visual media within the camera feed can be distracting at best or misleading at worst [8].

Applications for localization systems not only include pedestrian indoor navigation, but also robotics, dynamic personalized pricing, product placement, advertisement, fleet management or intelligent spaces [5], [17], [18].

4.2 Error

Each localization system is subject to some degree of error. The physical nature of the sensors involved in this process introduces a range of variables that add up to imprecise measurements. Even if a perfect sensor would exist that does produce perfect linear performance over its tracking space, it might still be affected by unpredictable, strong outside forces [1]. For example, a magnetometer, no matter how perfect it is at localizing the earth's magnetic north pole if exposed to a strong magnet it too will start to point towards that. Other sources of nonlinear errors include vibrations and temperature differences.[19], [20]. Each step produces small vibrations throughout the body that can affect the gyroscope performance of handheld devices. And while there are error estimation techniques that take advantage of the walking pattern of humans, they typically do not consider the vibration of the sensor directly and instead generalize different sources of errors. For example, Zero Velocity Potential Update (ZUPT) techniques are used to detect the moment when a foot is making contact with the earth [21]. This moment is then used as a baseline for a noise filtering algorithm.

Not only do inaccurate measurements introduce errors, the time it takes to sample a sensor and interpret its data is affected by the system latency, too. It is an unavoidable delay that can also never be completely eliminated but only compensated by using robust and efficient statistical methods [8]. But no matter which model is used to predict these errors, the exact degree of orientational and translational error at a time is never known. If it

were, it could be corrected to eliminate it [1]. Which is why MacIntyre et al. refer to the term "error" as a range of possible errors.
This range of possible errors, or uncertain tracking errors, is a key topic to this work. The Display Techniques proposed in Chapter 5 are evaluated based on their ability to accommodate for this varying degree of error.

4.3 Example Systems

Different tracking systems have different degrees of errors. For example, signal-based localization techniques (Wi-Fi, GPS, Bluetooth, etc.) are more prone to errors deriving from changes in the environment. Cluttered environments introduce multipath, non-line-of-sight, and shadowing artifacts that affect either the arrival time, angle, or signal strength of a ray reaching the sensor [5], [10], [22]. Even the human body can have such an effect on tracking accuracy [23].

This chapter provides an overview of several tracking techniques commonly found in pedestrian navigation systems and lists their expected accuracy (Table 1).

Tracking System	**Accuracy**
UWB	1 cm – 0.3 m
Wi-Fi Based Positioning System	1 – 10 m
Magnetic Positioning System	1 – 8 m
GPS	1 – 10 m
Bluetooth	0.5 – 10 m
Vision-Based	1 mm – 1m
PDR	n/a

Table 1: Accuracy of user tracking systems

4.3.1 Ultra-Wideband (UWB) Positioning System

Ultra-wideband positioning systems rely on commodity hardware to usually locate a special tag using Time of Arrival or similar techniques. By transmitting ultrashort pulses over multiple frequencies, this system's performance can achieve higher accuracy than most other techniques. Under ideal circumstances, accuracies within centimeters can are possible, making it one of the best-performing systems [5], [24]. However, due to the high cost of installing and maintaining the infrastructure, most of the use-cases are special industry applications instead of pedestrian navigation [25].
The accuracy of the Ultra-Wideband positioning system range from 0.01 to 0.3 meters [5], [24].

4.3.2 Wi-Fi Based Positioning System

Every smartphone is able to detect a Wi-Fi signal and almost every building has more than one access point that sends out these signals. This is one of the biggest advantages of using Wi-Fi-based positioning systems and why they are one of the most known solutions for indoor localization [5]. To calculate the user's position, several techniques can be used: Time of Arrival, Angle of Arrival (AoA), Received Signal Strength (RSS), or Fingerprinting [5], [22], [26], [27]. RSS is typically used because it is easier to acquire than the other information [28]. Wi-Fi systems are also commonly part of hybrid tracking solutions that combine different positioning systems for better accuracy.
The disadvantage of using the Wi-Fi infrastructure is that it is often managed by the

networking department which has a different purpose for the hardware [16]. In addition to that, surveying the location of all access points in order to use these for the calculations is a time-consuming process. Especially since their location is not always documented or somewhere visible.
The accuracy of Wi-Fi-based positioning systems can range from 1 to 4 meters [17], [22], [26].

4.3.3 Magnetic Positioning System

The magnetometer is not only influenced by the earth's magnetic field, but also by changes in the geomagnetic field within a building [15]. This means that using one as a compass might not always accurately detect the magnetic north pole indoors. Shu et al. reached an accuracy of up to 0.9 meters by recording the geomagnetic fingerprint of a building and using that pattern to localize the user over time. This approach is only possible because the location-specific magnetic fields are usually temporal stable and can, therefore, be mapped. However, this approach does have its limitations. Moving objects or even people within the building or changes in the temperature can affect magnetometer and reduce their accuracy non-uniformly [1], [20]. At the very least, it can be used as an approximation of the orientation of a user within a building [16], [18].
The accuracy of magnetic-based positioning systems can range from 1 to 8 meters [15], [16], [18].

4.3.4 Global Positioning System (GPS)

GPS, GLONASS, BeiDou, and Galileo are satellite systems that enable real-time positioning by Time of Arrival methods [29]. These satellites are constantly orbiting the earth and sending out data downwards so that receiver devices can calculate their position relative to them. Its accuracy is relatively similar to the other positioning systems (1 – 10 m), but the signals used by the satellites are affected by buildings in such a way that the indoor performance suffers to sometimes unusable quality [5].
The accuracy of the GPS can range from 1 to 10 meters [1], [5], [29].

4.3.5 Bluetooth Positioning System

By distributing Bluetooth beacons across a building, devices within are able to receive several of these signals and use trilateration algorithms to determine their position relative to these beacons. The approach to calculating the user position is very similar to Wi-Fi solutions, in that they use RSS, AoA, or similar techniques [30], [31]. The main disadvantage of Bluetooth positioning systems is they rely on special hardware compared to the already installed Wi-Fi access points.
The accuracy of Bluetooth positioning systems can range from 0.5 to 10 meters [30], [31].

4.3.6 Vision-Based Positioning System

As mentioned above, it is possible to map a 3D scene into a 2D image by calculating the angle of light hitting a camera sensor [14]. Information about the surrounding world can then be used to calculate the position of the camera within that world that is constantly recording it. This localization can be done by gathering unique feature points within the image and match them to a dataset that contains the position of reference images and the correlating position of the camera within the building [12]. Another approach that does not rely on a priori knowledge is using these feature points to track the position of the camera and generating a map of the building dynamically [32]. It is also very common

for vision-based positioning systems to incorporate other positioning systems that rely on different sensors. For example, smartphones have a host of sensors available to them, such as accelerometers, gyroscopes, and magnetometer [33]. These can all be used together to compensate for their individual shortcomings.
The advantages of this system include its ability to transition between outdoor and indoor navigation seamlessly and that it doesn't rely on extra infrastructure, which makes it scalable and cheaper than other approaches [4]. A disadvantage is that the image quality is affected by lighting conditions and especially occlusion more than any other technique. Even reflecting surfaces can drastically reduce tracking accuracy [34].
The accuracy of vision-based positioning systems can range from 1 mm to 1.5 m [35]–[37]. This high accuracy is only possible in a controlled and small environment. Using such a system in a bigger space reduces that accuracy.

4.3.7 Pedestrian Dead Reckoning (PDR) Positioning System

The pedestrian dead reckoning system assumes that the position of a pedestrian is changed by step movements. The step pattern is typically recognized using an accelerometer and refined through various techniques, such as zero velocity updates. They can be implemented using a foot-mounted sensor that communicates with another processing device or by a smartphone with its inbuild sensors [27], [38], [39]. This system can only locate the user relative to a baseline position. Additionally, it suffers from accumulated errors the most out of all systems. This means that the accuracy of the PDR system constantly declines over time. Because of this, it is normally used together with other positioning systems that can provide this baseline.
The accuracy of these systems is usually reported as a percentage of the distance traveled. As such, this positioning system has no numerical value in Table 1 [18], [38], [39].

4.3.8 Other Systems

There are many other positioning systems that can be used to localize a user within a building. However, these were not included in the summary due to their strong reliance on proprietary hardware or because they are very rarely used. For example, a system that utilizes sound, light beacons, FM radio, or RADAR [5].

5 Augmented Reality Navigation Display Techniques

Pedestrian navigation services have seen an increase in interest for several years now [40]. They evolved from using 2D paper maps to utilizing digital maps on a mobile device to give location-based turn-by-turn instructions [10], [41], [42]. With the help of modern real-time positioning systems, more precise instructions were possible. Adapting how these instructions are delivered based on the user's needs and the performance of the positioning system can greatly increase the performance of users when navigating through unknown environments [10]. But according to Rehman et al., interface design has not been a topic on most indoor navigation applications which instead focuses on the localization technique.
This paper puts display techniques at the forefront instead and compares different implementations based on their ability on how they perform given non-perfect tracking solutions.
Tracking errors within an AR application can have a negative impact on the user

experience. Misaligned or constantly jittering visual elements can be distracting at best or misleading at worst [8]. A common issue for misaligned information within navigation applications is a wrongly assumed position of the user. The impact of this error varies between different display techniques. For example, by providing more context information through a map, users are able to evaluate their position independently and account for wrong instructions [43]. The display techniques proposed in this work are designed to still provide useful information, even if the assumed location of the user is wrong to some extent.

The three types of information provided by these techniques include:
1) **The path**, which is a list of 3D coordinates within the building starting from the first waypoint and ending at the goal. This can be displayed as virtual points within the building or as turn-by-turn instructions that describe the path.
2) **The building geometry**, which includes walls, doors, stairs, or elevator(s). The building geometry can be shown as an abstract visualization on a map or by displaying the complex 3D geometry as a model. It is also possible to only display relevant parts of the building within the application, for example, to "see" behind corners.
3) **The user's pose** that represents their position and rotation within the building. Acquiring knowledge over landmarks, routes and their relationship between each other is an important step to form a mental model of the environment [44]. How they are presented often depends on the representation of the user's pose. AR inherently support an egocentric viewpoint, but providing an exocentric view can also support building the user's mental model [45].

Another factor for developing a display technique is to manage the user's expectation of the system. If a technique is perceived as very accurate, based on past experiences of the user with similar techniques, an inaccurate tracking state could be perceived as overly distracting [46]. Whereas a less common technique which makes an imperfect tracking state less obvious could more easily prepare the user for the actual experience when using the system [1].
Display techniques that incorporate one or more of the three types of information (see above) can be categorized into three categories; Discrete Information, Guiding Information, and Context Information. This work will explore several examples for each category and compare them. Furthermore, for a more pronounced comparison, a number of these categories will be selected carefully

5.1 Navigation Visualization Categories

The categories presented here are a generalized extension of the visualization techniques introduced by Pankratz et al. [9]. The extended categories are defined as the following:

- **Discrete Information**: Navigation hints are shown as the exact location of the next or a series of next steps
- **Guiding Information**: Navigation hints only show the direction towards the next step
- **Context Information**: The area around the user is shown in an exocentric view

5.1.1 Discrete Information

Discrete information is the middle ground between guiding and context information by providing some understanding of the overall path in addition to providing immediate steps on where to go next. Typical navigation hints include information about the path, starting from the user's position. A common example of this category is lines on the ground (DI 1). It is a simple illustration of the path directly projected onto the ground either in the form of a colorized line or a string of arrows pointing towards the next corner. Instead of drawing a line towards it, only displaying the next corner using a Waypoint Marker (DI 2) is also possible. A drawback of this technique is finding a balance of displaying enough information but not too much to avoid cluttering the screen with information. Such techniques are also prone to tracking errors since they do not provide much context information to manually adjust for them. It is obvious when waypoints don't align with the environment due to a rotational error, for example. The limited amount of information provided by a simple line makes it sometimes hard to guess where the system wants the user to go.

To account for this, the first proposed display technique "**Bending Words**" (DI 3, Figure 1), incorporates explicit instructions into the display. It is based on the notion that most waypoints on a path are corners where the user has to either go left or right. Instead of showing a line that goes around the corner, a three-dimensional sentence is used that bends around it with the respective instruction as its words. This should allow the user to quickly see where they are headed to, based on the position of the words while the words themselves provide clear instructions in case of unclear situations. Instead of wondering where to go based on the form of the information, a user can resort to the written information and try to recover the instructions manually.

5.1.2 Guiding Information

The principle behind guiding information display techniques is to help users reach their goal step-by-step. Navigation hints only point towards the very next corner of the path and don't show any specific locations. The user's cognitive resources required when engaging with a display technique is one criterion when evaluating it [47]. Using a technique, such as guiding information-based techniques, that use an egocentric point of view can improve the performance of users [48]. An example of such a technique is a Guiding Arrow (GI 1) that start at the user and point towards the next waypoint. Arrows are a well-understood concept that will make it easy for the user to understand the intention of the display technique. Another example is a Shining Light (GI 2) that points toward the next waypoint and fades away in the distance. This has the advantage of being perceived as less accurate than an arrow, which can dictate the users' expectations beforehand. The starting point of the light could also be in the middle of the device instead of the user and act as a directional flashlight.

The low amount of information provided by guiding information techniques can be used to design a display technique that can help visually impaired people. One challenge of such a system is the limited information that can be conveyed each second [49]. The other is that only haptic and auditory cues can be used under these circumstances, which requires a slightly different approach to the interface design. The general idea is still to guide the user toward the next waypoint and instead of using an arrow or a light, different rates of vibrations have to be used. The proposed display technique "**Haptic Feedback**" (GI 3), guides the user towards the right direction by sending out frequent vibration impulses.

The last technique introduced here that falls under egocentric navigation information is the use of a “**Digital Avatar**” (GI 4, Figure 1) to guide the users. It could be human avatars, as well as animals, fictional characters, or even objects that hover in the air. To provide the navigation information, they would stand between the user and the next waypoint and try to make the user follow them. Reacting to tracking errors using them can be done in several different ways; For example, they could move around a little and not stand in the exact position between the user and their next waypoint to increase the variety of information provided to the user. Or it could be implemented in the form of idle animations that directly portray the confidence of the tracking system as the confidence of the avatar in where to go next.

5.1.3 Context Information

In this work, information that exceeds the knowledge about the users’ position and their very next waypoint is categorized as context information. This typically includes the traditional paper map (CI 1) or other forms of representations of the environment. Users are shown the context in which they are positioned in instead of the direct solution for navigation through it. The advantage of showing context information is that it often does not require orientation information but only the location of the user [4]. Their position can be displayed through various methods - a dot, a small avatar or a big arrow are all generally recognizable depictions of the user’s position. By providing landmarks or other identifiable information about their surroundings, users are able to orient themselves using these techniques even without displaying their approximate location [50]. However, the best landmark is still the user’s position itself. The goal of providing contextual information is to help them understand the large-scale environment around them and incorporate that knowledge into the mental model of the user for the cost of an increased cognitive load at that moment.

The spatial knowledge of humans can be described as a set of disconnected components with routes that can be followed in one direction and not the other [51]. The additional observation point provided by context-based display techniques can then be integrated with the users’ egocentric point of view to help them learn to understand the connections between the spaces and eventually help them form their own route memory [10].

An example would be the World In Miniature (**WIM**) model (CI 2, Figure 1) [7], [47]. It is essentially a small model of the building that the user can have a bird-eye view of. It is possible to easily localize yourself by locking the orientation of the miniature to the orientation of the building and showing a marker for the position of the user within that model.

Discrete Information	DI 1: Lines on the ground DI 2: Waypoint Marker **DI 3: Bending Words**
Guiding Information	GI 1: Guiding Arrow GI 2: Shining Light **GI 3: Haptic Feedback** **GI 4: Digital Avatar**
Context Information	CI 1: Paper Map **CI 2: WIM**

Table 2: Display Technique examples

5.2 General Interface Design

The following chapters will describe four implementations of these techniques that are then evaluated using several factors, as described in chapters 6 and 7.

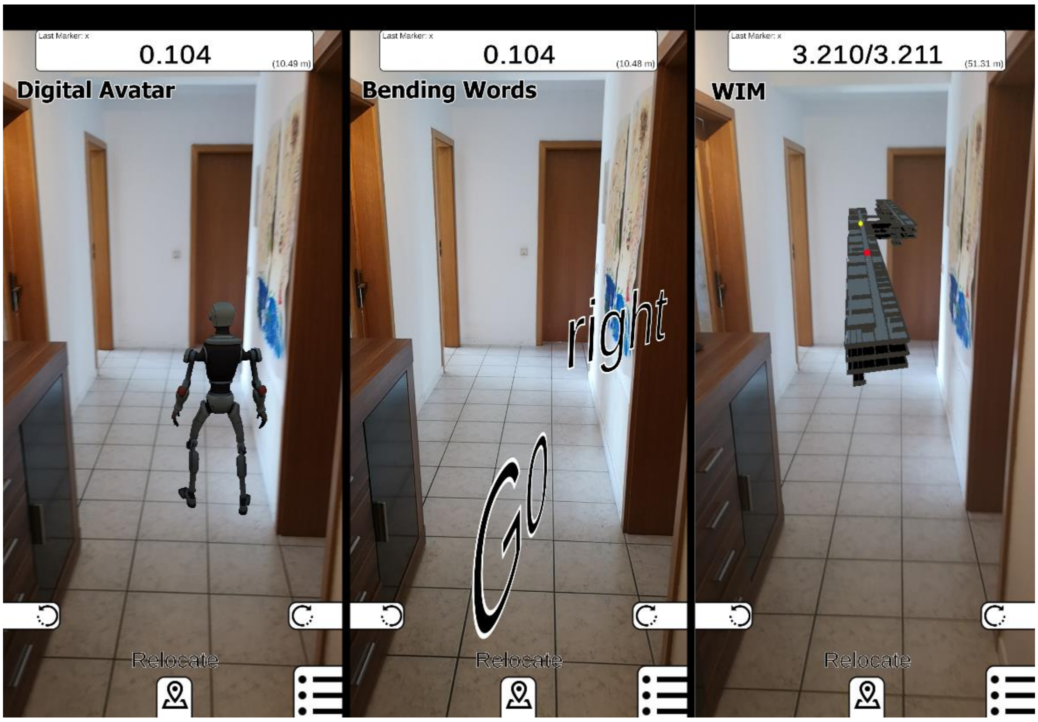

Figure 1: Display Techniques developed for this thesis

5.3 Technique 1: Bending Words

The Bending Words technique describes a new approach to the traditional line display. Instead of showing the navigation path as a line on the ground, a three-dimensional text is displayed at key locations. If the current task is to keep going straight, the text will reflect that. This is to improve the confidence of the user that the system is still working as indented and that they are on the right track. If they approach a corner, the display changes so that the words bend around the corner which is also reflected by the text of the words.
The advantage of this technique is that it not only shows the user where to go, it tells them in text form, too. This is to increase the user's ability to understand the system's intention even when it suffers from low accuracy. The drawback is that it's a relatively unknown approach. This could mean that users might not initially understand how to interact with the system.

5.4 Technique 2: WIM

The World In Minatare, or WIM, model shows the surrounding area as a small model within the view. The rotation of the model is kept the same as the building rotation perceived by the system. This consistency promises to improve the performance of users when using this technique [48]. The location of the user is represented as a big, red dot somewhere within the building. Rotating the phone around to look at different angles of the building reveals that the center of the rotation is the position of the user. The walls are

only half as tall to reveal some information about the structure behind them and visualize the number of floors below the user. If the user is located in one of the middle floors, the ones above are also hidden so that the current floor is on top of the model. Additionally, the very next step of the navigation path is displayed with a different marker so that the user knows where to go. It is possible to display this marker as discrete information within the camera feed as well, but to focus this work on each individual technique, only the model and its information is shown.
Tracking errors are most notable in their rotation. Due to the small scale of the model, a misaligned position is not as visible as a model that is not parallel to the building. The model developed for this work is kept very simple and only depicts the walls and shape of the building. The small scale mentioned before is also a drawback of this technique. The information displayed within is not easily readable and might lead to confusion in complex buildings.

5.5 Technique 3: Digital Avatar

The avatar as seen in Figure 1, is a humanoid robot that guides the user towards the next corner of the path. It walks in front of them and waits for the user if they are not moving anymore. Having the avatar in front of the user also has the potential of positively affecting the tracking capabilities of the system by motivating the user to keep their phone in an upright position. This can drastically improve the tracking accuracy by allowing the location system to acquire more feature points within the video feed [4].
However, this technique also suffers from accumulated tracking errors. Since the avatar is relatively close to the user, any movements through walls or obstacles are immediately noticeable. Managing the expectation of the user can also provide a challenge. Humanoid characters can seem more intelligent than they are. This disparity between expectation and reality can be disappointing for users.

5.6 Technique 4: Haptic Feedback

Traveling in an unknown environment is already a challenge for many people because of their lack of knowledge about the new area and the inability for humans to quickly acquire that knowledge. Blind and visually impaired people struggle even more with this information acquisition challenge [52]. The purpose of this display technique is to aid these people to not only navigate through an unknown environment but also provide them with information that they would need much more time acquiring otherwise.
When removing the visual components, the best interaction interfaces left are through haptic feedback and auditory cues. If the phone is pointed towards the right direction, it vibrates at a higher rate instead of when it is pointed to either the left or right side of the next corner. Audio cues could also be used to improve the users' performance by announcing if the device has to be pointed more to the left or right side. The downside of such a technique is that obstacle avoidance is left to the user and no feedback on the current system state is provided. However, since obstacle avoidance is an everyday task for this user group, it should not have a significant impact on their performance.

(Note that this work only focuses on the techniques themselves and not the application interaction from onboarding to inputting the destination.)

6 Study 1: Display Technique User Evaluation Study

The limited availability of the campus building and social distancing rules in response to the current Covid-19 pandemic made it impossible to properly conduct this survey with human participants. Therefore, this chapter will only describe such a survey and make a hypothesis of its outcome. To account for the lack of results, chapter 7 will look at the display techniques from a more theoretical standpoint and compare them using criteria found in other literature. The purpose of this chapter is to provide an outline for performing a survey in the future.

6.1 Methodology

The aim of this study is to evaluate the display techniques introduced in chapter 5 for their human factors, such as their performance and workload [10]. The rooms within the campus building are numbered in such a way that you can roughly determine the direction of your destination without the need of any navigational aids. However, this system is relatively unintuitive and relies on the user to confirm the number of sometimes two or more rooms to evaluate the direction in which their numbers ascend. This traditional method of navigating should prove as a baseline to evaluate the other techniques against. An alternative baseline could be provided by a paper printout of the floor plans. The advantage of this paper map is that it does not rely on a priori knowledge about the building.

6.2 Participants

The target audience for a campus navigation application can include new students who have not yet acquired the special knowledge to navigate the building independently, but also visitors who will only be at the university very rarely. As such, it is important to include a wide variety of participants that have a varying degree of experience with technology, especially augmented reality.

6.3 Settings and Materials

The setting for this study is the main building of the TH Köln – University of Applied Sciences Campus Gummersbach. Two different routes (Figure 2) are presented here that each involves two turns. The routes are contained within one floor to keep the accumulated tracking error as controlled as possible and because traveling between floors includes UI interactions that are not part of the evaluation. This reduces the possible number of turns available for each route. Route 1 requires the participant to navigate from room 1248 to room 1123 and Route 2 from room 3219 to room 3114.

The application would be deployed using a handheld device (Huawei P20 Pro) with screen recording software installed that can be used to record the exact view of the user when interacting with the application. If they agreed, the microphone audio could also be recorded for later evaluation. In addition to the device, a clipboard for the questionnaires and paper maps would also be useful.

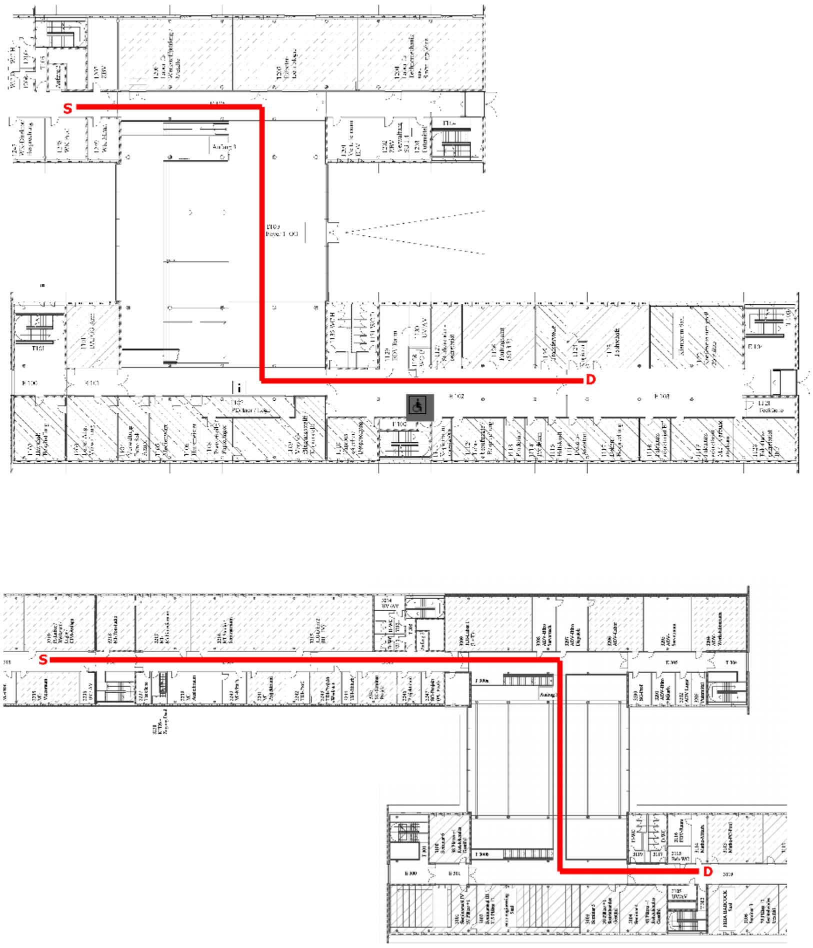

Figure 2: Route 1 (top) & 2 (bottom)

6.4 Variables and Data Collection

This study is designed to use a between-subject study design due to a large number of independent variables but could be adapted to a within-subject design if not enough participants are available [53]. The independent variables are the types of navigation techniques; DI 2, GI 3, GI 5, CI 1 (as listed in
Table 2) and Route 1 & 2 (as showing in Figure 2). The dependent variables include perceived comprehensibility and manipulability measured with the questionnaires proposed in 6.6, as well as the time required to navigate the route.
The task given to the participants is “to arrive at the destination as quickly as possible with a reasonable and safe walking speed” [10]. Any incident where they are requesting assistance should also be recorded with information about how often, when, and which kind of directions have been given. To better document these interactions and the other data points, a screen and audio recording can be used [54].

6.5 Procedure

The experiment should take place on different floors of the main building to reduce the impact of memory retention for each participant [10]. The aim of the study should also be presented after the participants received instructions on how to use the application [55].
First, the participants should fill out a pre-evaluation questionnaire and the required

consent forms. Then they should be led to the starting position of a route, preferably in a way so that they do not cross any points of it. To keep the focus on the display technique interaction, the application should be handed over with the first localization steps set-up and the navigation technique activated. The initial localization error should then be noted and compared to the error at the end of the procedure. At most two minutes should be enough to let the user familiarize themselves with the device and application before they are given their task. The data collection described previously starts with the beginning of the task and ends as soon as the user identified their goal and is acquired by walking behind the participant with a one-meter distance. After that, the end error state of the system should be recorded while the participants are given the questionnaire.

6.6 Questionnaire Design

There would be two questionnaires given to the participants. The first one is to ask them about their experience with AR and the campus. This is to put the results of the second questionnaire into perspective and to possibly change the route if the participant would be able to navigate to the destination independently.
The pre-evaluation questionnaire (Table 3) includes open questions that can be answered freely. It is derived from questionnaires proposed by other indoor navigation studies [10], [43], [55].

Santos et al. developed a usability scale for handheld augmented reality (HARUS) that scores the comprehensibility and manipulability of an application on a seven-point Likert scale where 1 is “strongly disagree” and 7 is “strongly agree” [56]. The evaluation questionnaire (Table 4) is based on this scale and was extended to include the subjective evaluation of the user experience listed by Xu et al [54].

Personal skills	Q1: Have you experienced Augmented Reality applications before this test? Q2: How often do you use mobile devices? Q3: How familiar are you with the Campus Gummersbach? Q4: Do you know how to get to the destination on your own?

Table 3: Pre-Evaluation Questionnaire

Comprehensibility statements	S1: I think that interacting with this application requires a lot of mental effort. S2: I thought the amount of information displayed on the screen was appropriate. S3: I thought that the information displayed on the screen was difficult to read. S4: I felt that the information display was responding fast enough. S5: I thought that the information displayed on the screen was confusing. S6: I thought the words and symbols on the screen were easy to read. S7: I felt that the display was flickering too much. S8: I thought that the information displayed on the screen was consistent.

Manipulability statements	S9: I think that interacting with this application requires a lot of body muscle effort. S10: I felt that using the application was comfortable for my arms and hands. S11: I found the device difficult to hold while operating the application. S12: I found it easy to input information through the application. S13: I felt that my arm or hand became tired after using the application. S14: I think the application is easy to control. S15: I felt that I was losing grip and dropping the device at some point. S16: I think the operation of this application is simple and uncomplicated.
Subjective evaluation statements	S17: The guide/tool is useful for the navigation task. S18: Using the guide/tool during the task is easy. S19: It was enjoyable to use the guide/tool. S20: I was stressed during the course of the task. S21: The guide/tool could make decisions like a human. S22: I could trust the guide/tool to guide me.

Table 4: Evaluation Questionnaire

6.7 Hypothesis

(H1) It is expected that the context information interaction technique (CI 1) requires more cognitive processes than the others which will be reflected in S1 and S2. (H2) Due to the focus on the individual techniques, it is also possible that the information provided by each one alone is not enough to satisfy the user's needs. (H3) Using digital navigation aids should also be easier to use compared to the baseline. (H4) However, due to the potential of tracking errors and other technical difficulties, the baseline techniques are expected to have a greater trust value than the digital counterparts. (H5) Lastly, using the application should not introduce any discomfort due to the short usage time.

7 Study 2: Display Technique Criteria Analysis

For a more objective approach to evaluating the proposed display techniques, this chapter presents criteria found in other augmented reality-related works and ranks the display techniques based on their ability to satisfy these criteria. The aggregated list of criteria is categorized under; visibility, interaction, as well as relevant meta criteria that don't fall under the first two categories. Each criterion was chosen based on the relevance to not only AR interfaces in general, but also to solve the problem of uncertain tracking errors described above. In addition to looking at the purely technical aspect of the techniques and how they were implemented during this work, this study also considers human factors, such as the angle they hold their device at or the cognitive affordance related to a UI element [57].

The display techniques have been implemented within a previously developed indoor navigation application for the Campus Gummersbach. These implementations are used as a basis for analyzing and comparing the techniques to each other. The techniques are then ranked based on their performance.

7.1 Visibility

The first category describes elements that influence the ability of a user to see the navigation information and process it as accurately as possible. It includes the visibility of elements so that they are perceivable within the video feed, how often and how much of this information is located within the screen, and how effective it is in guiding the user.

7.1.1 Contrast

Contrast measures the ability to easily discern elements within the screen [10]. A challenge of overlaying digital information within the video feed is finding a balance between blending these elements perfectly with the scene or making them distinct enough so that the user can easily identify important information. If the contrast of the display technique is high enough, even the effect of visual disturbances such as the trembling of the device while walking can be minimized [58]. The measurement to compare the different display techniques will be the luminance contrast ratio as calculated by Gabbard et al. [59]:

$$\frac{(L_{max} - L_{min})}{(L_{max} + L_{min})}$$

Equation 1: Calculation for luminance contrast

"Where L_{max} and L_{min} are taken from the Y-value in CIE XYZ color space and represent the highest and lowest luminance" on the edges of the display technique within an image taken during an analysis (Figure 3). The images couldn't be taken in the university due to the current prohibitions in place due to the Covid-19 pandemic. As a result, the exact degree of contrast ratios across the building could not be covered. Since this study only compares the different techniques with each other, it should still be sufficient to analyze their contrast within the same, albeit limited, conditions.

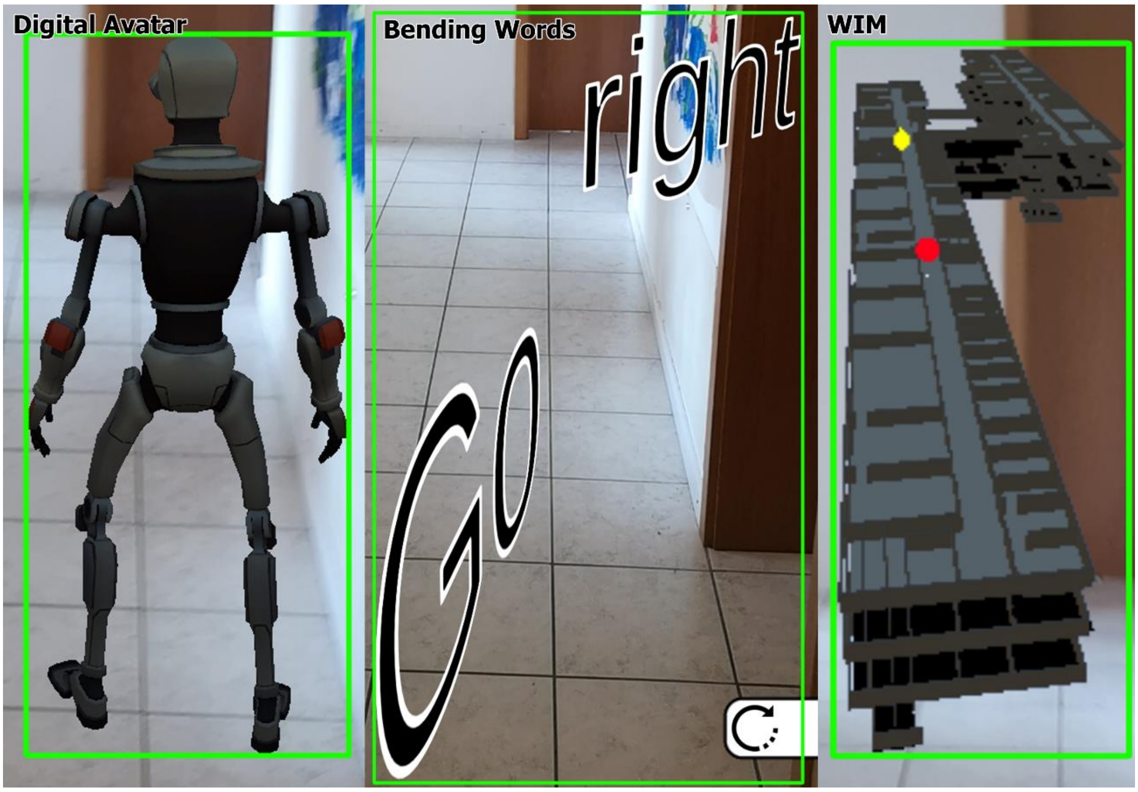

Figure 3: Contrast evaluation of Display Techniques with a green box indicating edges for the contrast calculation

7.1.2 Deviation Range

The deviation range describes the range of positions that a display technique can have within the video feed to still be perceived as displaying an acceptable accuracy [10]. For example, an arrow pointing towards a point in the distance can be rotated slightly to the left and right to still be considered pointing towards this point. A higher alignment between the desired position within the world and the position as shown in the video feed can lower the chance of user failures such as taking a wrong turn and following the wrong path [60]. Yeh et al. found that subjects performed best when the rotational error of virtual cues was less than 7.5° but still performed “poor” at an error of up to 22.5° [61]. The display techniques can be evaluated using these thresholds.

7.1.3 Visibility of Digital Information

An important factor in augmented reality applications is the method of integrating digital information with the environment. Since the screen is the main medium of conveying this digital information, it is naturally limited by its small size relative to the peripheral vision of humans. This means that in order to show a large amount of information, the user might have to move the device around a lot. To reduce the impact of this issue, navigation instructions have to be comprehensively distributed across the path and visible during the appropriate time [10], [62]. An overview of the upcoming tasks and especially the ability to easily see the next target location can also increase the performance of users [60]. To measure the impact of this issue within an indoor navigation context, the visible percentage of the path is measured.

7.2 Interaction

A big part of any display technique is how it promotes interaction with the user. This interaction can not only influence the user experience, but it can also have an impact on the tracking system itself. Vision-based tracking systems rely on the camera to extract the 3D geometry of the environment. If this camera is pointed at the ground, this information about the geometry is limited. The following chapters will explain the different criteria related to user interaction and their relevancy.

7.2.1 Device Orientation

The advantage of using a mobile phone for an AR application is that it is a very commonly used device that has a camera and screen already built-in. However, the everyday use of such a device does not include using the camera. If there is no reason to hold the phone upwards, it is usually held at a 45° angle towards the ground [63]. This limits the number of feature points visible to the camera which can negatively impact the performance of a vision-based tracking system (chapter 4.3.6). This criterion evaluates the Display Techniques based on how a user is expected to hold their device while interacting with the technique. If the user is prompted to hold their phone upwards, the respective technique will be ranked higher.

7.2.2 Instant Feedback

The time between a user-input and the moment of navigation-relevant information being presented to the user is an important factor for overall performance and user experience [10], [62], [64]. For example, if the tracking system corrects and error found, the display technique should be able to instantly adapt to the new situation. To measure this time, a video of the navigation application developed for this work was analyzed. A typical user interaction with the system is choosing the target destination and then identifying the navigation information that leads the user along the way. The exact time from when the user touches the screen to when the navigation information is displayed in the right position within the video feed is the value for this criterion.

7.2.3 Environmental Awareness

The awareness of the surrounding area is an important factor when using an AR application or digital navigation aids in general [10], [50], [58], [65]. Not only does it increase the knowledge of autonomously traversing the building over time, but it also improves the safety of users so that they can avoid dangerous situations and obstacles. A main contributing factor for this is inherently the navigation visualization category (chapter 5.1) of the display technique. A technique can increase this environmental awareness by using landmarks as part of the localization method or by making the environment stand out more through effects like occlusions.

7.2.4 Multimodality Count

"Multimodality Count" in this work is a simple measure counting the number of natural sensory receptors being utilized to convey information [65]. Or more simply, how many different senses are taken advantage of by the display technique. The main categories of modalities relevant to an AR application at this point in time are: visual, auditory, and haptic senses [66]. The visual task is the most known and used method to interact with an application. Most of the information is provided over the screen as either objects of different shape and size, or texts directly conveying information. Voice augmentations are the second most common instruction form to communicate with the user. Within a navigation context, they typically include turn instructions or to provide more general information about points of interest [10], [67]. To accommodate visually impaired people, haptic feedback has also been used in combination to voice augmentations. However, it is not very common in a mobile application since the amount of tactile feedback from a smartphone is fairly limited [49]. Using several different modalities together increases the number of information channels utilized to convey information, which in turn can increase the performance of the user [68].

7.3 Meta qualities

The last category includes one abstract criterium that evaluates and predicts the user's integration with the system. If one goal of a navigation application is to teach users knowledge on how to not rely on it, then this category is based on the notion that display techniques can be integrated into the toolkit of the user to achieve that goal.

7.3.1 Novelty

Novelty is an abstract measurement that tries to predict the likelihood and speed of a new design being accepted into the user's domain knowledge [64]. Already existing navigation applications have introduced a set of display techniques that are widely accepted and understood. For example, arrows on walls or lines on a paper map represent a common form of display technique that users will understand without the need for an explanation or without interacting with them. Using these known forms can help users understand the system's intention when being guided towards their goal. This criterion is the only one that is not directly measurable without an extensive study of the navigation state-of-the-art and users' response to it. Therefore, the ranking of the display techniques is solely based on argumentative comparisons. A better ranking for this criterium indicates a more common approach of displaying navigation information.

7.4 Evaluation

The results of this evaluation are displayed as a ranking of the different display techniques. A lower number means a better ranking in the respective category.

	Contrast	Deviation Range	Visibility of Digital Information	Device Orientation	Instant Feedback	Environmental Awareness	Multimodality Count	Novelty*	**Average Rank**
Bending Words	1	1	2	1	1	2	2	3	**1,63**
WIM	3	2	1	4	1	1	2	2	**2,00**
Digital Avatar	2	4	3	1	4	4	2	1	**2,63**
Haptic Feedback	n/a	3	4	3	1	3	1	4	**2,71**

Table 5: Rankings of the different display techniques (left) within a criterium (top)
*(A higher rank in novelty denotes a more common approach)

To compare the **contrast** of each technique, still-images of the display techniques have been evaluated (Figure 3). These images were taken in a stand-in location for the Campus Gummersbach and their illuminances contrast values calculated using Equation 1. The average of four points within the image (top, left, right, bottom) determined their ranking in this category. Bending Words performed the best with an average of 0.92, followed by Digital Avatar (0.37) and Bending Words (0.24). Haptic Feedback is excluded in this category because it does not take advantage of any visual elements.
A main contributing factor for the result of Bending Words was that this specific implementation has a white outline to better differentiate the black text from the background. These results are expected to be very similar in the campus building since the interior consists of mostly grey, or muted colors in general. This means that a bright display technique has an advantage in this category. The darker color of the other techniques could be favorable on an environment that is expected to be brightly lit.

The **deviation range** is obtained by measuring the range of rotational error that a display technique can be exposed to until it no longer overlaps the position of the next turn. This overlap was simulated in steps of 2.5° to acquire the respective error values. The Digital Avatar and Haptic Feedback could still display accurate navigation information within a 5° rotational error. Bending Words would go up to 10°. The value for Haptic Feedback has been obtained by measuring the range in which the vibration rates change. In this implementation, the vibrations adapt in 3 different values: Slow when pointing the phone away from the screen, fast when the next corner is within the outer third of the screen, and fastest when the middle of the screen is close to the next corner (Figure 4). These zones can be changed based on user feedback. Because of this ability to easily increase the deviation range, it has been ranked 3rd.
The WIM model would also stop overlapping after a 5° error. However, it can still be used to identify the next steps by comparing the landmarks within the model with landmarks of the surroundings. This advantage influenced the ranking of the technique even though its exact extent cannot be measured. However, since this study is only a comparative analysis, it should still hold value in this evaluation.

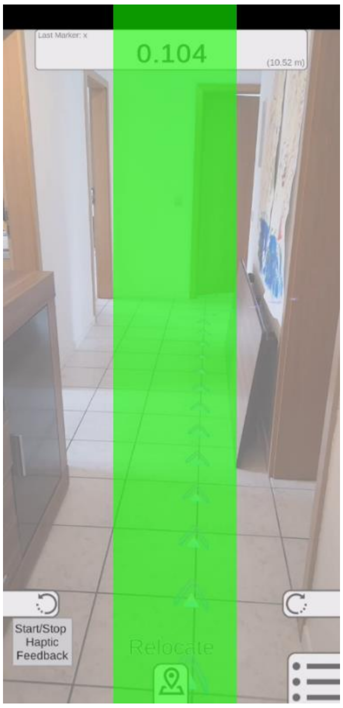

Figure 4: Haptic Feedback zones.

The visible percentage of the path makes up the **visibility of digital information**. Each navigation visualization category inherently limits this information. For example, Guiding Information techniques, such as Haptic Feedback and Digital Avatar, only show the direction towards the very next corner and not more information about the rest of the path. The difference between these two techniques is that Digital Avatar can also show the exact position of the corner when the user is very close to it. Because of this, it is ranked above Haptic Feedback as shown in Table 5. Contrasting to that, Bending Words always displays the exact location of the next corner and part of the path before and after it. This additional information makes it rank at the 2nd place behind WIM. It not only displays navigation information but also rooms behind corners and other possible paths that might be used to go around unforeseen obstacles. This advantage of displaying context information ranks it in the first place. The current implementation of WIM only shows one corner of the path but can be easily extended to display the whole path without cluttering the video feed with information.

The **device orientation** during normal use is mostly influenced by where and how digital information is displayed on the screen. Most techniques that rely on spatially registered information would support a device orientation of 90°, such as Bending Words and Digital Avatar. Therefore, both techniques are ranked first. Since Haptic Feedback is designed to aid visually impaired people, it does not use visual cues. Its rate of vibration is also solely based on the yaw of the device and not the pitch. As such, it is assumed that the device will be held at a natural angle of 45°, ranking it below the former two. WIM on the other hand changes its visibility of digital information through the rotation of the device. The most information is visible when looking down at the model, making a rotation of 0° more likely to occur. It also never prompts the user to point the camera in the direction of travel, making it rank the lowest.

The data for the **instant feedback** evaluation was obtained by analyzing a video recording of the application. The time between pressing a button to choose a display technique and the moment where it is loaded in the right position is used to evaluate this criterium. The test was done in two stages with one repeat at each stage. The first stage simulated the first time choosing a destination, while the second stage simulates walking through the building (see Appendix A.1). The Digital Avatar is the only technique that has time-consuming animations, such as walking. This means that it is inherently the slowest technique to display the navigation information. Additionally, loading the avatar model for the first also introduces an additional delay. Due to the low visual fidelity of the other techniques, they did not suffer from this issue. The other techniques did not show a significant time difference during this test, ranking all three in the first place.

As described above, **environment awareness** heavily relies on how much information is provided by the display technique and how much information has to be acquired in addition to that. The World In Miniature model supports this notion in that it provides context instead of integrating with information about the surroundings, ranking it in the first place. Bending Words also provides some context information in the form of text. For example, “Go right” can initiate a search within the environment for opportunities to turn on the path ahead. This motivation to scan the environment is rated to be second to the WIM. Haptic Feedback was rated third because, while it also only leads the user towards the goal – similar to the Virtual Avatar – it provides the least amount of information. This limitation on navigation information is expected to motivate the user to scan their environment more to compensate for its lack. Since the avatar leads the user towards the goal by staying close to them, it is considered the least awareness-friendly method. It does not inherently support the notion of looking around but rather focuses the user’s attention on the model and animation.

The **multimodal count** provides a very easy numeric analysis of the display techniques. All of the visual-based techniques only rely on visual information, while Haptic Feedback not only vibrates but also generates acoustic feedback to accommodate for the limited information available through only haptic feedback.

The **novelty** of each technique is based on how often elements of them are found within other navigation media or every-day situations. The most common medium for a long time were paper maps. However, this context information-based medium has since been replaced by digital navigation aids, such as Google Maps (Figure 5). These modern applications often rely on a mix between context information in the form of an abstract street map with lines displaying the path and turn-by-turn instructions. Bending Words lends itself mostly to these instructions, whereas the WIM resembles a new approach to paper maps. On the other hand, the Digital Avatar comes from a more human approach. It resembles a situation where the user has to follow someone through an unknown environment. However, dynamic feedback during these situations is an important feature that can also help set the right expectation of this technique. Due to the lack thereof, it is currently ranked below the former two techniques mentioned but could surpass them with future improvements. Haptic Feedback is the most uncommon approach as it relies on the assumption that users can correlate an increase in vibration speed to positive feedback.

Figure 5: Example of Google Maps navigation

8 Discussion

Mobile augmented reality applications rely on a robust tracking system to accurately place digital information within the environment. These systems combine information from a camera image with the information from a host of other sensors found within a mobile phone to track users [33], [69]. Due to their physical limitations, these sensors can accumulate errors over time, making it more difficult to place digital information in a meaningful way.

This thesis has identified four possible display technique that could still provide comprehensive navigation information when suffering from these errors. The names of these techniques are Bending Words, WIM, Digital Avatar, and Haptic Feedback (chapter 5.3 ff.). The following paragraphs will discuss the advantages and drawbacks of each technique.

Bending Words and WIM ranked the highest within the Display Technique criteria analysis. They especially achieved the most differentiating results within the “Visibility” category. Bending Words produced a high contrast using white outlines and combines that with a very wide digital visualization that spans approximately 20° in the deviation range category. This makes it easy to identify the display technique in different navigation scenarios.

While the **WIM** model does not have a very high contrast ratio, it is always visible within the screen and provides additional information about the environment. This means that identifying navigation-relevant information is only a matter of cognitive resources. This is essentially a tradeoff between the cognitive load reduction provided by AR interfaces for a more error-robust display technique [70]. To account for this loss of spatially registered information within the video feed, the view of the model itself is centered on the user. This turns the device into a bird-eye-view of the building that is constantly following them, making it easier for a user to recognize their own position. The increase in abstraction also has a positive effect on environmental awareness. By delegating a part of the localization task to the user, they have to scan their surroundings and match it to the model. This active process can support the learning process of spatial knowledge required to independently navigate the building [51].

The biggest advantage of this miniature model not utilized in this work is its ability to show digital information from a second point-of-view; Other techniques, like the Digital Avatar, could be displayed simultaneously within the model as well as the video feed. This can combine the advantages of both techniques in that it completes the egocentric experience provided by the camera alone with the exocentric context from the model.

Bending Words, on the other hand, retains the advantages of spatially registered information. It also reduces the abstract display of navigation information even further by providing more details in the form of words. In its core, this technique takes advantage of a person’s ability to naturally follow turn-by-turn instructions [41], [71]. However, these instructions have one drawback; they cannot accurately describe the exact location of the next turn [72]. To reduce the impact of this, they are displayed within the environment close to where this turning point is. Using a white outline for this technique also increases the contrast between the visualization and the background. The white outline was chosen to better differentiate the black text from the background. This mainly increases visibility in darker environments, such as the campus building. Other techniques could also benefit from this effect. However, outlining the avatar or the model could reduce the perceived spatial registration with the environment, making it harder to understand location-specific information. The extend of that effect should be evaluated in a future study.

One factor for evaluating the techniques is the time required to get feedback on the current navigation state. A user tracking system constantly updates the user's position. When an error within that position estimate is detected, correcting it can sometimes cause bigger jumps within that position. If a display technique is slow to react to this change, navigation information could be delayed at best or lost at worst. Most techniques don't introduce such a delay, except for **Digital Avatar**. It is the only technique that uses animations that take time to complete. The avatar walks towards the desired spot in response to the user's constant change in position during their navigation task. These walking animations were initially introduced to support the humanoid appearance of this technique and adapted to move faster than the average walking speed of up to 5.5 km/h [73]. If the user's position changes due to error correction, the avatar has to still walk towards the user's new position. This delay also occurs when the technique is initiated for the first time; The user has to wait for the avatar to walk in front of them before it is clear which direction is the right way to go.

Haptic Feedback is the most unique display technique out of the four. It is aimed at visually impaired people and as such does not rely on visual AR elements. This means it cannot rely on the vast amount of information that visual media can provide and instead has to use vibrations to convey information. However, the amount of information that can be convey using only vibrations is very limited. As such, only the direction of the next turning point can be described and nothing else. An unexpected side-effect of using the haptic motor of a smartphone was the mechanical sound produced by it. This vibration sound was not part of the initial design but still adds to the technique by increasing the modality count. Hearing the vibrations in addition to feeling them makes it easier to differentiate the changing vibration speed indicating the direction of travel. Audio cues were not included to only focus on the usability of haptic feedback alone. Using these cues to provide turn-by-turn instructions, for example, does have drawbacks. It takes a significant amount of time to deliver auditory instructions, which in turn can promote rotational errors caused by misunderstandings [49], [74].

A limitation of this study was the focus on the display techniques individually. Using more than one at a time could improve the shortcomings of each of them. For example, the Digital Avatar could not only be improved by adopting a color scheme that promotes higher contrast, but it could also be combined with an indicator that tells the user where the avatar is currently waiting for them. Some of the criteria for Study 2 could also not be evaluated with participants. For example, Environmental Awareness could only be evaluated by comparing the reliance on landmarks and other outside wayfinding cues. A study by Fenech et al., that evaluated acoustic turn-by-turn instructions, suggests that there are more human factors involved in environmental awareness [75]. Conducting the proposed user study could show such correlations.

9 Conclusion

To accommodate the research on user tracking systems, this thesis provided a set of Display Techniques that can reduce the impact of tracking errors within a navigation application. Each different technique provides a new approach to display the path, the building geometry, or the user's position. These approaches have then been evaluated based on criteria found in related literature. The techniques Bending Words and WIM performed the best within these criteria. Bending Words expands the spatially registered information provided by an AR display with precise instructions that can be easily interpreted. The approach of the WIM model relies on a person's ability to read a map and supports this task using a location-aware model of the environment.
In the future, new Display Techniques can be constructed for each visualization category, using the criteria presented in this work. The proposed human study can provide new insight into how these criteria and other human factors intersect with each other.

10 Personal Development

This work has shown me not only how much I have learned about writing an academic paper, but also how I can apply my skills as a software developer to quickly implement a desired feature within an application. Having experienced the research method program of a masters' course during my exchange semester has greatly helped me work through the unbelievable amount of papers related to the topics presented here. I learned how to quickly understand, not only the context but also the quality of an academic paper and search it for related information quickly. Writing all my papers in English starting with my exchange semester has allowed me to gain enough experience so that I can freely write out my thoughts in a way that (hopefully) can be understood.

Throughout this work, I have learned a lot about the current state of AR indoor navigation systems and have closely followed AR technology in general. This field is an exciting opportunity to bring together many different other technologies. And I hope to experience that first-hand.

11 References

[1] B. MacIntyre and E. Machado Coelho, "Adapting to dynamic registration errors using level of error (LOE) filtering," in *Proceedings IEEE and ACM International Symposium on Augmented Reality (ISAR 2000)*, 2000, pp. 85–88, doi: 10.1109/ISAR.2000.880927.

[2] R. T. Azuma, "A Survey of Augmented Reality," *Presence Teleoperators Virtual Environ.*, vol. 6, no. 4, pp. 355–385, Aug. 1997, doi: 10.1162/pres.1997.6.4.355.

[3] J. Grubert, T. Langlotz, S. Zollmann, and H. Regenbrecht, "Towards Pervasive Augmented Reality: Context-Awareness in Augmented Reality," *IEEE Trans. Vis. Comput. Graph.*, vol. 23, no. 6, pp. 1706–1724, Jun. 2017, doi: 10.1109/TVCG.2016.2543720.

[4] A. Möller, M. Kranz, R. Huitl, S. Diewald, and L. Roalter, "A mobile indoor navigation system interface adapted to vision-based localization," in *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia*, Ulm, Germany, 2012, pp. 1–10, doi: 10.1145/2406367.2406372 [Online]. Available: https://doi.org/10.1145/2406367.2406372. [Accessed: 24-Apr-2020]

[5] A. Yassin *et al.*, "Recent Advances in Indoor Localization: A Survey on Theoretical Approaches and Applications," *IEEE Commun. Surv. Tutor.*, vol. 19, no. 2, pp. 1327–1346, Secondquarter 2017, doi: 10.1109/COMST.2016.2632427.

[6] T. Feigl, A. Porada, S. Steiner, C. Löffler, C. Mutschler, and M. Philippsen, "Localization Limitations of ARCore, ARKit, and Hololens in Dynamic Large-scale Industry Environments," in *VISIGRAPP*, 2020, doi: 10.5220/0008989903070318.

[7] D. HALLAWAY, S. FEINER, and T. HÖLLERER, "Bridging the Gaps: Hybrid Tracking for Adaptive Mobile Augmented Reality," *Appl. Artif. Intell.*, vol. 18, no. 6, pp. 477–500, Jul. 2004, doi: 10.1080/08839510490462768.

[8] B. MacIntyre, E. M. Coelho, and S. J. Julier, "Estimating and adapting to registration errors in augmented reality systems," in *Proceedings IEEE Virtual Reality 2002*, 2002, pp. 73–80, doi: 10.1109/VR.2002.996507.

[9] F. Pankratz, A. Dippon, T. Coskun, and G. Klinker, "User awareness of tracking uncertainties in AR navigation scenarios," in *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2013, pp. 285–286, doi: 10.1109/ISMAR.2013.6671807.

[10] U. Rehman and S. Cao, "Augmented-Reality-Based Indoor Navigation: A Comparative Analysis of Handheld Devices Versus Google Glass," *IEEE Trans. Hum.-Mach. Syst.*, vol. 47, no. 1, pp. 140–151, Feb. 2017, doi: 10.1109/THMS.2016.2620106.

[11] S. Adler, S. Schmitt, K. Wolter, and M. Kyas, "A survey of experimental evaluation in indoor localization research," in *2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2015, pp. 1–10, doi: 10.1109/IPIN.2015.7346749.

[12] J. Kim and H. Jun, "Vision-based location positioning using augmented reality for indoor navigation," *IEEE Trans. Consum. Electron.*, vol. 54, no. 3, pp. 954–962, Aug. 2008, doi: 10.1109/TCE.2008.4637573.

[13] K. Liu, G. Motta, and T. Ma, "XYZ Indoor Navigation through Augmented Reality: A Research in Progress," in *2016 IEEE International Conference on Services Computing (SCC)*, 2016, pp. 299–306, doi: 10.1109/SCC.2016.46.

[14] W. Förstner and B. P. Wrobel, "Geometry and Orientation of the Single Image," in *Photogrammetric Computer Vision: Statistics, Geometry, Orientation and Reconstruction*, W. Förstner and B. P. Wrobel, Eds. Cham: Springer International Publishing, 2016, pp. 455–545 [Online]. Available: https://doi.org/10.1007/978-3-319-11550-4_12. [Accessed: 17-Jun-2020]

[15] Y. Shu, C. Bo, G. Shen, C. Zhao, L. Li, and F. Zhao, "Magicol: Indoor Localization Using Pervasive Magnetic Field and Opportunistic WiFi Sensing," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 7, pp. 1443–1457, Jul. 2015, doi: 10.1109/JSAC.2015.2430274.

[16] J. C. A. Herrera, P. G. Plöger, A. Hinkenjann, J. Maiero, M. Flores, and A. Ramos, "Pedestrian indoor positioning using smartphone multi-sensing, radio beacons, user positions probability map and IndoorOSM floor plan representation," in *2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2014, pp. 636–645, doi: 10.1109/IPIN.2014.7275538.

[17] D. Lymberopoulos, J. Liu, X. Yang, R. R. Choudhury, V. Handziski, and S. Sen, "A realistic evaluation and comparison of indoor location technologies: experiences and lessons learned," in *Proceedings of the 14th International Conference on Information Processing in Sensor Networks*, Seattle, Washington, 2015, pp. 178–189, doi: 10.1145/2737095.2737726 [Online]. Available: https://doi.org/10.1145/2737095.2737726. [Accessed: 05-May-2020]

[18] A. R. Jiménez, F. Seco, J. C. Prieto, and J. Guevara, "Indoor pedestrian navigation using an INS/EKF framework for yaw drift reduction and a foot-mounted IMU," in *Navigation and Communication 2010 7th Workshop on Positioning*, 2010, pp. 135–143, doi: 10.1109/WPNC.2010.5649300.

[19] H. Weinberg, "Gyro Mechanical Performance: The Most Important Parameter," p. 5, 2011.

[20] V.-T. Pham, D.-C. Nguyen, Q.-H. Tran, T. Chu, and D.-T. Tran, "Thermal Stability of Magnetic Compass Sensor for High Accuracy Positioning Applications," vol. 0, no. 0, p. 8, 2015.
[21] R. P. Suresh, V. Sridhar, J. Pramod, and V. Talasila, "Zero Velocity Potential Update (ZUPT) as a Correction Technique," in *2018 3rd International Conference On Internet of Things: Smart Innovation and Usages (IoT-SIU)*, 2018, pp. 1–8, doi: 10.1109/IoT-SIU.2018.8519902.
[22] C. Yang and H. Shao, "WiFi-based indoor positioning," *IEEE Commun. Mag.*, vol. 53, no. 3, pp. 150–157, Mar. 2015, doi: 10.1109/MCOM.2015.7060497.
[23] F. Askarzadeh, K. Pahlavan, S. Makarov, Y. Ye, and U. Khan, "Analyzing the effect of human body and metallic objects for indoor geolocation," in *2016 10th International Symposium on Medical Information and Communication Technology (ISMICT)*, 2016, pp. 1–5, doi: 10.1109/ISMICT.2016.7498903.
[24] A. Alarifi *et al.*, "Ultra Wideband Indoor Positioning Technologies: Analysis and Recent Advances," *Sensors*, vol. 16, no. 5, p. 707, May 2016, doi: 10.3390/s16050707.
[25] "Ultra-wideband Localization Examples." [Online]. Available: https://www.ultrawideband.io/en/examples.php. [Accessed: 17-Jun-2020]
[26] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, "SpotFi: Decimeter Level Localization Using WiFi," in *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*, London, United Kingdom, 2015, pp. 269–282, doi: 10.1145/2785956.2787487 [Online]. Available: https://doi.org/10.1145/2785956.2787487. [Accessed: 05-May-2020]
[27] S. He and S.-H. G. Chan, "Wi-Fi Fingerprint-Based Indoor Positioning: Recent Advances and Comparisons," *IEEE Commun. Surv. Tutor.*, vol. 18, no. 1, pp. 466–490, Firstquarter 2016, doi: 10.1109/COMST.2015.2464084.
[28] H. Wang, H. Lenz, A. Szabo, J. Bamberger, and U. D. Hanebeck, "Enhancing the Map Usage for Indoor Location-Aware Systems," in *Human-Computer Interaction. Interaction Platforms and Techniques*, Berlin, Heidelberg, 2007, pp. 151–160, doi: 10.1007/978-3-540-73107-8_17.
[29] X. Li *et al.*, "Accuracy and reliability of multi-GNSS real-time precise positioning: GPS, GLONASS, BeiDou, and Galileo," *J. Geod.*, vol. 89, no. 6, pp. 607–635, Jun. 2015, doi: 10.1007/s00190-015-0802-8.
[30] "Enhancing Bluetooth Location Services with Direction Finding." [Online]. Available: https://www.bluetooth.com/bluetooth-resources/enhancing-bluetooth-location-services-with-direction-finding/. [Accessed: 12-May-2020]
[31] M. E. Rida, F. Liu, Y. Jadi, A. A. A. Algawhari, and A. Askourih, "Indoor Location Position Based on Bluetooth Signal Strength," in *2015 2nd International Conference on Information Science and Control Engineering*, 2015, pp. 769–773, doi: 10.1109/ICISCE.2015.177.
[32] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: part I," *IEEE Robot. Autom. Mag.*, vol. 13, no. 2, pp. 99–110, Jun. 2006, doi: 10.1109/MRA.2006.1638022.
[33] V. Bettadapura, I. Essa, and C. Pantofaru, "Egocentric Field-of-View Localization Using First-Person Point-of-View Devices," in *2015 IEEE Winter Conference on Applications of Computer Vision*, 2015, pp. 626–633, doi: 10.1109/WACV.2015.89.
[34] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse, "MonoSLAM: Real-Time Single Camera SLAM," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 6, pp. 1052–1067, Jun. 2007, doi: 10.1109/TPAMI.2007.1049.
[35] K. Guan, L. Ma, X. Tan, and S. Guo, "Vision-based indoor localization approach based on SURF and landmark," in *2016 International Wireless Communications and Mobile Computing Conference (IWCMC)*, 2016, pp. 655–659, doi: 10.1109/IWCMC.2016.7577134.
[36] R. T. Rodrigues, M. Basiri, A. P. Aguiar, and P. Miraldo, "Low-Level Active Visual Navigation: Increasing Robustness of Vision-Based Localization Using Potential

Fields," *IEEE Robot. Autom. Lett.*, vol. 3, no. 3, pp. 2079–2086, Jul. 2018, doi: 10.1109/LRA.2018.2809628.
[37] A. Motroni *et al.*, "SAR-Based Indoor Localization of UHF-RFID Tags via Mobile Robot," in *2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2018, pp. 1–8, doi: 10.1109/IPIN.2018.8533847.
[38] W. Kang and Y. Han, "SmartPDR: Smartphone-Based Pedestrian Dead Reckoning for Indoor Localization," *IEEE Sens. J.*, vol. 15, no. 5, pp. 2906–2916, May 2015, doi: 10.1109/JSEN.2014.2382568.
[39] H. J. Ju, M. S. Lee, C. G. Park, S. Lee, and S. Park, "Advanced Heuristic Drift Elimination for indoor pedestrian navigation," in *2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2014, pp. 729–732, doi: 10.1109/IPIN.2014.7275553.
[40] A. J. May, T. Ross, S. H. Bayer, and M. J. Tarkiainen, "Pedestrian navigation aids: information requirements and design implications," *Pers. Ubiquitous Comput.*, vol. 7, no. 6, pp. 331–338, Dec. 2003, doi: 10.1007/s00779-003-0248-5.
[41] H. J. Kim, "Turn-by-turn navigation system and next direction guidance method using the same," US7844394B230-Nov-2010 [Online]. Available: https://patents.google.com/patent/US7844394B2/en. [Accessed: 22-Feb-2020]
[42] A. Nix and A. W. Gellatly, "Turn-by-turn navigation system with enhanced turn icon," US7546207B209-Jun-2009 [Online]. Available: https://patents.google.com/patent/US7546207B2/en. [Accessed: 22-Feb-2020]
[43] D. D. McMahon, C. C. Smith, D. F. Cihak, R. Wright, and M. M. Gibbons, "Effects of Digital Navigation Aids on Adults With Intellectual Disabilities: Comparison of Paper Map, Google Maps, and Augmented Reality," *J. Spec. Educ. Technol.*, vol. 30, no. 3, pp. 157–165, Sep. 2015, doi: 10.1177/0162643415618927.
[44] A. W. Siegel and S. H. White, "The Development of Spatial Representations of Large-Scale Environments," in *Advances in Child Development and Behavior*, 1975, vol. 10, pp. 9–55, doi: https://doi.org/10.1016/S0065-2407(08)60007-5 [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0065240708600075. [Accessed: 24-May-2020]
[45] A. Dünser, M. Billinghurst, J. Wen, V. Lehtinen, and A. Nurminen, "Exploring the use of handheld AR for outdoor navigation," *Comput. Graph.*, vol. 36, no. 8, pp. 1084–1095, Dec. 2012, doi: 10.1016/j.cag.2012.10.001.
[46] B. MacIntyre, J. D. Bolter, E. Moreno, and B. Hannigan, "Augmented reality as a new media experience," in *Proceedings IEEE and ACM International Symposium on Augmented Reality*, 2001, pp. 197–206, doi: 10.1109/ISAR.2001.970538.
[47] R. Stoakley, M. J. Conway, and R. Pausch, "Virtual reality on a WIM: interactive worlds in miniature," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Denver, Colorado, USA, 1995, pp. 265–272, doi: 10.1145/223904.223938 [Online]. Available: https://doi.org/10.1145/223904.223938. [Accessed: 23-Feb-2020]
[48] C. D. Wickens, C.-C. Liang, T. Prevett, and O. Olmos, "Egocentric and Exocentric Displays for Terminal Area Navigation:," *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, Nov. 2016, doi: 10.1177/154193129403800105. [Online]. Available: https://journals.sagepub.com/doi/10.1177/154193129403800105. [Accessed: 24-Apr-2020]
[49] X. Zhang, X. Yao, Y. Zhu, and F. Hu, "An ARCore Based User Centric Assistive Navigation System for Visually Impaired People," *Appl. Sci.*, vol. 9, no. 5, p. 989, Jan. 2019, doi: 10.3390/app9050989.
[50] M. Brown and J. Pinchin, "Exploring Human Factors in Indoor Navigation," p. 7, 2013.
[51] B. Kuipers, "The 'Map in the Head' Metaphor:," *Environ. Behav.*, Jul. 2016, doi: 10.1177/0013916584142005. [Online]. Available: https://journals.sagepub.com/doi/10.1177/0013916584142005. [Accessed: 24-Apr-2020]

[52] Y. Tian, X. Yang, C. Yi, and A. Arditi, “Toward a computer vision-based wayfinding aid for blind persons to access unfamiliar indoor environments,” *Mach. Vis. Appl.*, vol. 24, no. 3, pp. 521–535, Apr. 2013, doi: 10.1007/s00138-012-0431-7.
[53] G. Charness, U. Gneezy, and M. A. Kuhn, “Experimental methods: Between-subject and within-subject design,” *J. Econ. Behav. Organ.*, vol. 81, no. 1, pp. 1–8, Jan. 2012, doi: 10.1016/j.jebo.2011.08.009.
[54] Q. Xu, L. Li, J. H. Lim, C. Y. C. Tan, M. Mukawa, and G. Wang, “A wearable virtual guide for context-aware cognitive indoor navigation,” in *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services*, Toronto, ON, Canada, 2014, pp. 111–120, doi: 10.1145/2628363.2628390 [Online]. Available: https://doi.org/10.1145/2628363.2628390. [Accessed: 02-May-2020]
[55] G.-D. Voinea, F. Girbacia, C. C. Postelnicu, and A. Marto, “Exploring Cultural Heritage Using Augmented Reality Through Google’s Project Tango and ARCore,” in *VR Technologies in Cultural Heritage*, Cham, 2019, pp. 93–106, doi: 10.1007/978-3-030-05819-7_8.
[56] M. E. C. Santos, J. Polvi, T. Taketomi, G. Yamamoto, C. Sandor, and H. Kato, “Toward Standard Usability Questionnaires for Handheld Augmented Reality,” *IEEE Comput. Graph. Appl.*, vol. 35, no. 5, pp. 66–75, Sep. 2015, doi: 10.1109/MCG.2015.94.
[57] R. Hartson, “Cognitive, physical, sensory, and functional affordances in interaction design,” *Behav. Inf. Technol.*, vol. 22, no. 5, pp. 315–338, Sep. 2003, doi: 10.1080/01449290310001592587.
[58] A. Morrison *et al.*, “Like bees around the hive: a comparative study of a mobile augmented reality map,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Boston, MA, USA, 2009, pp. 1889–1898, doi: 10.1145/1518701.1518991 [Online]. Available: https://doi.org/10.1145/1518701.1518991. [Accessed: 24-Apr-2020]
[59] J. L. Gabbard and J. E. Swan II, “Usability Engineering for Augmented Reality: Employing User-Based Studies to Inform Design,” *IEEE Trans. Vis. Comput. Graph.*, vol. 14, no. 3, pp. 513–525, May 2008, doi: 10.1109/TVCG.2008.24.
[60] A. Mulloni, H. Seichter, and D. Schmalstieg, “Handheld augmented reality indoor navigation with activity-based instructions,” in *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, Stockholm, Sweden, 2011, pp. 211–220, doi: 10.1145/2037373.2037406 [Online]. Available: https://doi.org/10.1145/2037373.2037406. [Accessed: 24-Apr-2020]
[61] M. Yeh, C. D. Wickens, M. J. L. Merlo, and D. L. Brandenburg, “Head-Up vs. Head-Down: Effects of Precision on Cue Effectiveness and Display Signaling,” *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, vol. 45, no. 27, pp. 1886–1890, Oct. 2001, doi: 10.1177/154193120104502707.
[62] A. G. Sutcliffe and K. D. Kaur, “Evaluating the usability of virtual reality user interfaces,” *Behav. Inf. Technol.*, vol. 19, no. 6, pp. 415–426, Jan. 2000, doi: 10.1080/014492900750052679.
[63] A. Möller *et al.*, “Experimental evaluation of user interfaces for visual indoor navigation,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Toronto, Ontario, Canada, 2014, pp. 3607–3616, doi: 10.1145/2556288.2557003 [Online]. Available: https://doi.org/10.1145/2556288.2557003. [Accessed: 24-Apr-2020]
[64] A. Sutcliffe and B. Gault, “Heuristic evaluation of virtual reality applications,” *Interact. Comput.*, vol. 16, no. 4, pp. 831–849, Aug. 2004, doi: 10.1016/j.intcom.2004.05.001.
[65] R. Grasset, J. Looser, and M. Billinghurst, “A step towards a multimodal AR interface: a new handheld device for 3D interaction,” in *Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR’05)*, Vienna, Austria,

2005, pp. 206–207, doi: 10.1109/ISMAR.2005.7 [Online]. Available: http://ieeexplore.ieee.org/document/1544696/. [Accessed: 24-Apr-2020]
[66] M. A. Livingston, “Evaluating human factors in augmented reality systems,” *IEEE Comput. Graph. Appl.*, vol. 25, no. 6, pp. 6–9, Nov. 2005, doi: 10.1109/MCG.2005.130.
[67] K. Yu, J. Chiu, M. Lee, and S. Chi, “A mobile application for an ecological campus navigation system using augmented reality,” in *2015 8th International Conference on Ubi-Media Computing (UMEDIA)*, 2015, pp. 17–22, doi: 10.1109/UMEDIA.2015.7297421.
[68] N. O. Bernsen, “Multimodality Theory,” in *Multimodal User Interfaces: From Signals to Interaction*, D. Tzovaras, Ed. Berlin, Heidelberg: Springer, 2008, pp. 5–29 [Online]. Available: https://doi.org/10.1007/978-3-540-78345-9_2. [Accessed: 14-Jun-2020]
[69] R. Ban, K. Kaji, K. Hiroi, and N. Kawaguchi, “Indoor positioning method integrating pedestrian Dead Reckoning with magnetic field and WiFi fingerprints,” in *2015 Eighth International Conference on Mobile Computing and Ubiquitous Networking (ICMU)*, 2015, pp. 167–172, doi: 10.1109/ICMU.2015.7061061.
[70] A. Tang, C. Owen, F. Biocca, and W. Mou, “Comparative Effectiveness of Augmented Reality in Object Assembly,” *NEW Horiz.*, no. 5, p. 8, 2003.
[71] A. Mulloni, H. Seichter, and D. Schmalstieg, “Indoor navigation with mixed reality world-in-miniature views and sparse localization on mobile devices,” in *Proceedings of the International Working Conference on Advanced Visual Interfaces*, Capri Island, Italy, 2012, pp. 212–215, doi: 10.1145/2254556.2254595 [Online]. Available: https://doi.org/10.1145/2254556.2254595. [Accessed: 23-Jun-2020]
[72] M. Pielot and S. Boll, “‘In Fifty Metres Turn Left’: Why Turn-by-turn Instructions Fail Pedestrians,” p. 3.
[73] K. Fitzpatrick, M. A. Brewer, and S. Turner, “Another Look at Pedestrian Walking Speed:,” *Transp. Res. Rec.*, Jan. 2006, doi: 10.1177/0361198106198200104. [Online]. Available: https://journals.sagepub.com/doi/10.1177/0361198106198200104. [Accessed: 24-Apr-2020]
[74] D. Ahmetovic, U. Oh, S. Mascetti, and C. Asakawa, “Turn Right: Analysis of Rotation Errors in Turn-by-Turn Navigation for Individuals with Visual Impairments,” in *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*, Galway Ireland, 2018, pp. 333–339, doi: 10.1145/3234695.3236363 [Online]. Available: https://dl.acm.org/doi/10.1145/3234695.3236363. [Accessed: 23-Jun-2020]
[75] E. P. Fenech, F. A. Drews, and J. Z. Bakdash, “The Effects of Acoustic Turn-by-turn Navigation on Wayfinding,” *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, vol. 54, no. 23, pp. 1926–1930, Sep. 2010, doi: 10.1177/154193121005402305.

12 Appendix

A.1 Study 2 Result Analysis

To conduct Study 2, a video analysis has been made for each Display Technique. Due to the Covid-19 pandemic, the videos could not be made within the Campus building. This mainly affects the contrast results.

Contrast

RBG values extracted from an image during normal use has been converted into XYZ values using an online tool (https://www.nixsensor.com/free-color-converter/).

Bending Words

	Y-value inside	Y-value outside	Result
Top	1.0000	0.0153	0.98
Left	0.9911	0.1446	0.85
Right	1.0000	0.0158	0.98
Bottom	0.9989	0.1211	0.88

Average: **0.92**

WIM

	Y-value inside	Y-value outside	Result
Top	0.0394	0.4431 / 0.1163	0.40 / 0.08
Left	0.0333	0.3291	0.30
Right	0.0382	0.0685	0.03
Bottom	0.0053	0.4131	0.41

Average: **0.24**

Digital Avatar

	Y-value inside	Y-value outside	Result
Top	0.0927	0.0619	0.03
Left	0.0153	0.3645	0.35
Right	0.0077	0.7428	0.74
Bottom	0.0006	0.3676	0.37

Average: **0.37**

Haptic Feedback

Not available due to no visuals.

Deviation Range

The deviation range was tested by manipulating the rotational tracking error in 2.5° steps at a time. These are the results on how long the display techniques overlapped compared to their original orientation.

Bending Words

	Left	Right
Deviation range	10°	10°

WIM

	Left	Right
Deviation range	5°	5°

Digital Avatar

	Left	Right
Deviation range	5°	5°

Haptic Feedback

	Left	Right
Deviation range	5°	5°

Instant Feedback

The timing between the screen press and a stable navigation information state has been measured using a screen recording of the software. The button press is visualized by enabling the corresponding developer option within the smartphone's software settings.
Each technique was tested twice with two repeats each test. The first test made after starting the application and then choosing a destination. The second test was made after walking with a destination already selected.

Bending Words

First Run

	Button press frame time	Frame time of stable navigation information	Resulting delta time
Repeat 1	2.026s	2.284s	0.258s
Repeat 2	6.451s	6.636s	0.185

Second Run

	Button press frame time	Frame time of stable navigation information	Resulting delta time
Repeat 1	1.174s	1.338s	0.164s
Repeat 2	3.692s	3.877s	0.185s

Result average: 0,198s

WIM

First Run

	Button press frame time	Frame time of stable navigation information	Resulting delta time
Repeat 1	3.218s	3.378s	0.160s
Repeat 2	6.234s	6.402s	0.168s

Second Run

	Button press frame time	Frame time of stable navigation information	Resulting delta time
Repeat 1	1.800s	1.992s	0.192s
Repeat 2	5.835s	6.044s	0.209s

Result average: 0,183s

Digital Avatar

First Run

	Button press frame time	Frame time of stable navigation information	Resulting delta time
Repeat 1	2.950s	5.412s	2.462s
Repeat 2	7.723s	7.870s	0.147s

Second Run

	Button press frame time	Frame time of stable navigation information	Resulting delta time
Repeat 1	1.101s	2.298s	1.197s
Repeat 2	4.027s	4.192s	0.165s

Result average: 0,993s

Haptic Feedback

First Run

	Button press frame time	Frame time of stable navigation information	Resulting delta time
Repeat 1	1.601s	1.759s	0.158s
Repeat 2	6.086s	6.290s	0.204s

Second Run

	Button press frame time	Frame time of stable navigation information	Resulting delta time
Repeat 1	2.451s	2.585s	0.134s
Repeat 2	5.915s	6.119s	0.204s

Result average: 0,175s

A.2 Digital Appendix

Study 2 Evaluation

The accompanying files include the video files used for the evaluation of Study 2. The files ending with "1" and "2" correlate to the stages described in chapter 7. These are the File paths:
Study_2_Evaluation/Avatar_1.mp4
Study_2_Evaluation/Avatar_2.mp4
Study_2_Evaluation/BendingWords_1.mp4
Study_2_Evaluation/BendingWords_2.mp4
Study_2_Evaluation/Deviation Range tests.mp4
Study_2_Evaluation/HapticFeedback_1.mp4
Study_2_Evaluation/HapticFeedback_2.mp4
Study_2_Evaluation/WIM_1.mp4
Study_2_Evaluation/WIM_2.mp4

Navigation Application

The project that was used to implement the four display techniques can be found in "ARIndoorNav_Project.zip". This project is also online available (https://github.com/Oscheibe/ARIndoorNav). The four implementations developed for this work can be found within the compressed archive under:
\ARIndoorNav Project\Assets\Scripts\View\ARVisuals_Avatar.cs
\ARIndoorNav Project\Assets\Scripts\View\ARVisuals_BendingWords.cs
\ARIndoorNav Project\Assets\Scripts\View\ARVisuals_HapticFeedback.cs
\ARIndoorNav Project\Assets\Scripts\View\ARVisuals_WIM.cs

Sources

As not every source has a Digital Object Identifier (DOI) link, offline sources have also been provided for these. Their titles are as follows:

- Gyro Mechanical Performance: The Most Important Parameter [19]
- Thermal Stability of Magnetic Compass Sensor for High Accuracy Positioning Applications [20]
- Ultra-wideband Localization Examples [25]
- https://www.bluetooth.com/bluetooth-resources/enhancing-bluetooth-location-services-with-direction-finding [30]
- Turn-by-turn navigation system and next direction guidance method using the same [41]
- Turn-by-turn navigation system with enhanced turn icon [42]
- Exploring Human Factors in Indoor Navigation [55]
- Comparative Effectiveness of Augmented Reality in Object Assembly [70]
- In Fifty Metres Turn Left': Why Turn-by-turn Instructions Fail Pedestrians [72]

These can be found in the "\Sources" folder.

13 Notice of Change

This version of the thesis was slightly changed from the work that was graded on 18th of August 2020. These changes include:

- The contact email address and phone number
- The date shown on the first page
- This chapter
- Small spelling mistakes and formatting issues throughout the work