Andrea Slyngstadli
Shahitha Sothinathan

Possibilities Using AR for Indoor Navigation

Master's thesis in Engineering and ICT
Supervisor: Terje Midtbø
Co-supervisor: Torbjørn Morland
June 2021

Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Engineering
Department of Civil and Environmental Engineering

NTNU
Norwegian University of Science and Technology

Andrea Slyngstadli
Shahitha Sothinathan

Possibilities Using AR for Indoor Navigation

Master’s thesis in Engineering and ICT
Supervisor: Terje Midtbø
Co-supervisor: Torbjørn Morland
June 2021

Norwegian University of Science and Technology
Faculty of Engineering
Department of Civil and Environmental Engineering

NTNU
Kunnskap for en bedre verden

NTNU
Faculty of Engineering Science and Technology
Department of Civil and Environmental Engineering

Master thesis

(TBA4925 - Geomatics, Master thesis)

Spring 2021
for

Andrea Slyngstadli and Shahitha Sothinathan

Possibilities Using AR for Indoor Navigation

BACKGROUND

AR is a new technique that potentially could be used to help users get a better user experience in wayfinding. Yet the technology behind AR is in early stages and it has several challenges.

TASK DESCRIPTION

The goal of the assignment is to implement and test 1-2 different technologies for AR and evaluate the usability and scalability of the technology for large campuses like universities.

Specific tasks:

- Evaluate existing AR solutions
- Select 1-2 AR solutions to implement together with MazeMap's indoor mapping platform
- Evaluate the technology and implementation with user testing

ADMINISTRATIVE/GUIDANCE

The work on the Master Thesis starts on January 15th, 2021.

The thesis report as described above shall be submitted digitally in INSPERA at the latest at June 11th, 2021.

External supervisor: Torbjørn Morland

Supervisors at NTNU and professor in charge:
Terje Midtbø

Trondheim, April, 2021

This page is intentionally left blank

Abstract

Most outdoor positioning and navigation systems adopt Global Navigation Satellite System (GNSS) signals to position a user's device, but within indoor spaces, these signals are not available, and alternative solutions have to be implemented. 2D floor plans or maps are the most commonly used interface in navigation systems, but these systems often require a high cognitive load, as the user has to make the connection between the real world and the interface. Augmented reality (AR) based indoor navigation systems have the potential of assisting their users without requiring great cognitive load, due to the navigation instructions being presented directly on top of the real world. This thesis tries to identify the current status and the possibilities of using AR as a way to aid users navigating within indoor spaces. To do this, theory regarding indoor positioning, extended realities, and augmented reality is presented. An AR-based indoor navigation system was implemented using tools like ARCore, IndoorAtlas SDK, and data provided by MazeMap, and then tested and compared to MazeMap's 2D digital map solution through a user test. The implemented AR solution is also compared to other AR solutions. The results from the user test showed no significant difference between the implemented AR solution and the 2D digital map solution in terms of average speed from start to destination, but significant differences related to the level of familiarity and genders were found. However, a majority of the participants preferred the AR solution. Based on data acquired from the user test, advantages, disadvantages, and other feedback regarding the AR-based indoor navigation system are presented and discussed.

Keywords: Indoor Navigation, Augmented Reality, Experimental Study

This page is intentionally left blank

Sammendrag

De fleste utendørs posisjonerings- og navigasjonssytemer bruker GNSS signaler til å posisjonere en enhet. Derfor må andre alternative metoder taes i bruk. 2D plantegninger eller kart er de mest brukte grensesnittene i navigasjonssystemer, men disse systemene krever ofte en stor kognitiv belastning for brukeren, siden vedkommende selv må knytte det som blir presentert i brukergrensesnittet opp mot den virkelige verden. Utvidet virkelighet (AR) er et annet alternativ til plantegninger og kart, hvor det ikke kreves en like stor grad av kognitiv belastning. Ved bruk av AR blir instruksjonene kommunisert til brukeren vha. informasjon plassert direkte over oppfatningen av den ekte verden. Denne oppgaven prøver å identifisere foreløpig status og muligheter rundt bruken av AR i innendørs navigering. For å gjøre dette blir teori angående inndendørs posisjonering, "extended realities" og AR presentert. Et AR-basert innendørs navigasjonssytem ble utviklet ved bruk av verktøy som ARCore, IndoorAtlas SDK og data hentet fra MazeMap, og deretter testet og sammenlignet med MazeMap sin 2D kart-applikasjon gjennom en brukertest. Den implementerte AR-løsningen blir også sammenlignet med andre AR-løsninger. Resultatene fra brukertesten viste ingen signifikant forskjell mellom gjennomsnittsfarten ved bruk av den implementerte AR-løsningen og 2D kart-løsningen, men signifikante forskjeller ble funnet når det kom til ulike nivå av bekjentskap til lokasjonen og kjønn. Allikevel kom det frem at et flertall av deltakerne i brukertesten foretrakk AR-løsningen. Fordeler, ulemper og andre tilbakemeldinger fra brukertesten, som omhandler AR-løsningen, blir også presentert og diskutert.

Nøkkelord: innendørs navigering, utvidet virkelighet, eksperimentell studie

This page is intentionally left blank

Preface

This paper is a master's thesis written for the Department of Civil and Environmental Engineering at the Norwegian University of Science and Technology (NTNU) in Trondheim, Norway. It is part of the study program Engineering and ICT, with a specialization in Geomatics. The thesis, and its coherent work, was conducted in the spring of 2021.

We would like to thank our supervisor, Terje Midtbø, for his feedback and guidance during the writing of this master's thesis. We would also like to thank MazeMap for all valuable assistance and help, especially during the implementation of the application. All participants in the user test were also highly appreciated. We also would like to thank our families for always supporting us. Lastly, we would like to thank Trondheim for providing us with the greatest of friends and being a great city to study in, leaving us with lots of fond memories.

Trondheim, June 2021

Andrea Slyngstadli
Shahitha Sothinathan

This page is intentionally left blank

Table of Contents

Abstract iii

Sammendrag v

Preface vii

List of Figures xiii

List of Tables xv

1 Introduction 1

1.1 Introduction 2

1.2 Motivation 4

2 Theory & Background 7

2.1 Indoor positioning 8

2.1.1 Signal properties 8

2.1.2 Positioning algorithms 10

2.1.3 Positioning technologies 11

2.1.4 Simultaneous localization and mapping 15

2.2 Extended Reality 16

2.2.1 Brief history of AR 17

2.2.2 Commercial AR 18

2.3 Augmented Reality 21

2.3.1 Sensors 21

2.3.2 Processor 23
2.3.3 Display 23
2.3.4 Augmented reality experience 24
2.4 AR Technologies 26
2.4.1 ARKit 26
2.4.2 ARCore 29
2.4.3 Vuforia 31
2.4.4 AR Foundation 33
2.4.5 AR indoor navigation technology 35
2.5 Existing solutions 37
2.5.1 Templates 37
2.5.2 Related work 38
2.5.3 Google Maps 39

3 Method 41

3.1 Implementation 42
3.1.1 IndoorAtlas 42
3.1.2 Fingerprinting 43
3.1.3 Applications 45
3.1.4 MazeMap 46
3.2 Implemented AR solution 47
3.3 Testing 51
3.3.1 Usability testing 51
3.3.2 Location 52
3.3.3 Equipment 53
3.3.4 Objective 53
3.3.5 Pilot test 54
3.3.6 Participants 54
3.3.7 Data acquisition 55
3.3.8 Infection control measures 56

4 Results **59**

4.1 User test 60

4.1.1 Pre-experimental questions 60

4.1.2 Practical part of the experiment 61

4.1.3 Post-experiment interview 66

5 Discussion **69**

5.1 Evaluation of AR solutions 70

5.2 Participants in the user test 72

5.3 Comparison of results 74

5.3.1 2D map and AR solutions 74

5.3.2 Level of familiarity 75

5.3.3 Gender 75

5.3.4 Combination 76

5.4 Implemented AR solution 77

5.4.1 Initial calibration 77

5.4.2 All-consuming 78

5.4.3 Camera angle 78

5.4.4 Points-of-Interest 79

5.4.5 How-to guide 79

5.4.6 Symbols and other information 80

5.4.7 Other improvements 81

6 Conclusion & Further Work **83**

6.1 Conclusion 84

6.2 Further Work 86

Bibliography **87**

Appendix **95**

A Pre-experimental questions 95

B Post-experimental interview 97

List of Figures

2.1 Signal properties: AOA and TOA 9
2.2 Signal properties: TDOA and RSSI 9
2.3 Extended realities 17
2.4 Camera coordinate system 27
2.5 Examples of image rating 32
2.6 Concept of AR Foundation 34
2.7 Modules of AR indoor navigation 36

3.1 Start page of application 47
3.2 User interface of the application 48
3.3 Instruction in applications 49
3.4 Point-of-Interest in application 50
3.5 2D, AR, and combination 56

4.1 Familiarity with building 61
4.2 Average speed by route 61
4.3 Average speed by familiarity 62
4.4 Average speed by gender 62
4.5 Preferences between the different solutions 67

List of Tables

2.1 Feature support in AR Foundation 35

3.1 Distribution of participants 55

4.1 Participants previous experience 60

4.2 Advantages and disadvantages of the 2D map solution 66

4.3 Advantages and disadvantages of the AR solution 66

4.4 Advantages and disadvantages of a combined solution 67

This page is intentionally left blank

Chapter 1

Introduction

In this chapter, the goal of the literature study and accompanying experiment is introduced, followed by the motivation behind the thesis.

1.1 Introduction

A vital and fundamental part of humans is being able to move from one place to another. The goal of the movement could be anything from acquiring resources, such as food, water, and shelter, or going to a meeting or birthday party. The variations are countless. This goal-directed movement from A to B through an environment is called navigation (Montello 2005). The origin of the word navigation stems from the late 1500s and was used regarding ships moving on water from one point to another (Merriam Webster 2021b). However, this has evolved in recent decades and is today used when describing any movement from one place to another, whether it is on land or on water (Merriam Webster 2021a).

In this day and age, two-thirds of the world's population have smartphones (Hollander 2017) where most of these have implemented GPS. As the use of smartphones and its popularity has increased over the recent years, maps and navigation apps have become a fundamental part of the smartphone experience. This has resulted in people worldwide becoming dependent on these kinds of services (He 2019), e.g. Google Maps or Apple Maps, which provides functionality for both drivers and pedestrians. However, pedestrian navigational systems have the added issue of navigating indoors.

Navigating indoors in unfamiliar buildings and environments can often result in over-use of time and a substantial amount of attention from the user (Fallah et al. 2013). In order to decrease the users' cognitive workload, external aids have been used in complex buildings, such as hospitals, universities, and shopping malls (Dong et al. 2021). Most current navigation systems on smartphones, specifically for pedestrians, provides either visual or audio of map-based turn-by-turn instructions (Amirian et al. 2016). However, this could require more of the user's attention and increase the cognitive load.

Augmented reality (AR) has in the recent years become more and more popular in various tasks in the field of geomatics (Dong et al. 2021). Particularly in navigational tools on mobile devices, which combines the video stream from the camera with the data from positioning and orientation sensors (Amirian et al. 2016). An electronic 2D-map on smartphones has symbols that represent geographical entities and requires the user to interpret the symbols and instructions that are given, whereas an AR system shows real scenes in the world with added, and preferably simple, symbols.

The aim of this thesis is to answer the research questions:

1. How can AR be used in indoor navigation?
2. What are the advantages and disadvantages with an AR-based indoor navigation system?

To answer these research questions, technology from IndoorAtlas and ARCore was used to map a building and to implement an AR-based indoor navigation system. A user test was conducted to test the usability of the AR solution and compare it to a 2D digital map. In addition to this, various AR technologies and solutions was evaluated.

The rest of the thesis is structured as follows: The remainder of chapter 1 presents the motivation of the thesis. Chapter 2 provides relevant background and theory. Chapter 3 presents the implementation process and the final design of the AR solution, as well as the setup of the user test. Chapter 4 presents the results of the user test. Chapter 5 provides discussion of the results, and finally, chapter 6 contains concluding remarks and further work.

1.2 Motivation

Companies are always on the lookout for technology that makes their product better for the users. This also applies to developers within the geomatics field. Throughout the last few years, the augmented reality technology has been a popular topic in the geomatics field (Dong et al. 2021). Using AR in order to create a more seamless and effective solution for the user is highly valuable.

Navigation systems have commonly been used in outdoor environments. However, navigation systems for indoor environments are still in the early development stages. In order to improve these solutions, the potential of combining AR technology and indoor navigation has been the subject of various research over the past few years. There are many reasons why using AR-based navigational aids benefits and improves already-existing solutions. The possible use cases for an indoor navigation system are numerous. Shopping malls, industrial facilities, educational institutions, hospitals, and museums are just some of the places where an application like this could be advantageous.

A 2D map requires the user to recognize abstract symbols and perform complex map-learning tasks, such as locating themselves on the map and route planning. When implementing AR with already-existing navigation technology some of the cognitive effort is eliminated (Dong et al. 2021). In such an application, the user is presented with simple virtual 3D models showing the way in front of a video stream captured by the device, and the navigation becomes more intuitive. The users would only have to understand the simple guidance shown on the screen and would not have to interpret any abstract symbols, recognize landmarks, or use mental rotation (Dong et al. 2021). This results in an easier and seamless navigation experience, which is very much needed and wanted, especially for people with no prior knowledge or reduced cognitive capacity.

This is something MazeMap saw potential in. MazeMap is a digital wayfinding platform that offers solutions for large campuses, hospitals, offices, hotels, and event venues. They already have a 2D digital map solution that is implemented on different campuses, including NTNU. When offering solutions for e.g. large campuses and offices, the desire is to increase the efficiency for the users. One of the main interests with this thesis was studying the usability and scalability of AR technology. This included testing a solution that could decrease the time a student uses to class or the time an employee uses to a meeting. Whether or not an implementation of AR is a feature users would prefer could be decided by exploring and researching this further.

Chapter 2

Theory & Background

In this chapter, the theoretical background and related work will be presented. This will give the reader the necessary and basic understanding needed in order to understand the rest of the paper.

2.1 Indoor positioning

Most outdoor positioning and navigation systems adopt Global Navigation Satellite System (GNSS) signals to position a user's device (H. Huang and Gartner 2010). Within indoor spaces, the signals from the satellites are not available as they, in most cases, cannot penetrate building structures (Kunhoth et al. 2020). As a result of this, indoor positioning systems have to use alternative solutions.

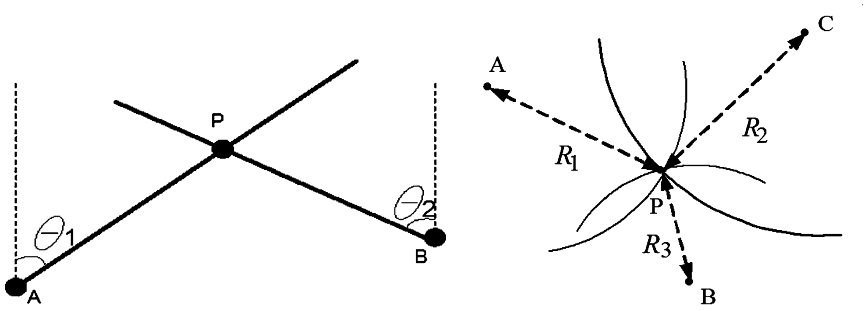
H. Huang and Gartner (2010) accentuated three key aspects when looking further into indoor positioning; *signals*, *signal metrics*, also referred to as signal properties, and *positioning algorithms*. Different kinds of signals can be Bluetooth, UWB, RFID or visible light, while examples of signal metrics are Received Signal Strength Indication (RSSI), Time of Arrival (TOA), Angle of Arrival (AOA), and Time Difference of Arrival (TDOA) (H. Huang and Gartner 2010). Positioning algorithms are used to derive the actual position from the recorded signal properties. Commonly used algorithms for this purpose are proximity, triangulation, or location fingerprinting. Sakpere et al. (2017) did a similar survey of indoor positioning systems and looked further into both indoor positioning techniques and technologies. Indoor positioning techniques are the properties and algorithms used to determine an estimate of the actual position, while indoor positioning technologies are the actual signal(s) used. Similar to the research by H. Huang and Gartner (2010), the indoor positioning techniques are divided into signal properties and positioning algorithms.

2.1.1 Signal properties

Signal properties are geometrical parameters, which consist of metrics such as the angle, distance, and signal, that can be used in various calculations to determine an object's position (Sakpere et al. 2017). The most commonly used are, as previously mentioned, AOA, TOA, TDoA, and RSSI.

Angle of Arrival: The AOA is the direction from which a signal is received (Wiig 2010). To obtain the AOA, two or more different reference points are used. The direction (angle and distance) between each of the reference points and the transmitter (object to be positioned) is used to find the intersecting point

(Sakpere et al. 2017). Figure 2.1a illustrates positioning based on AOA measurements.

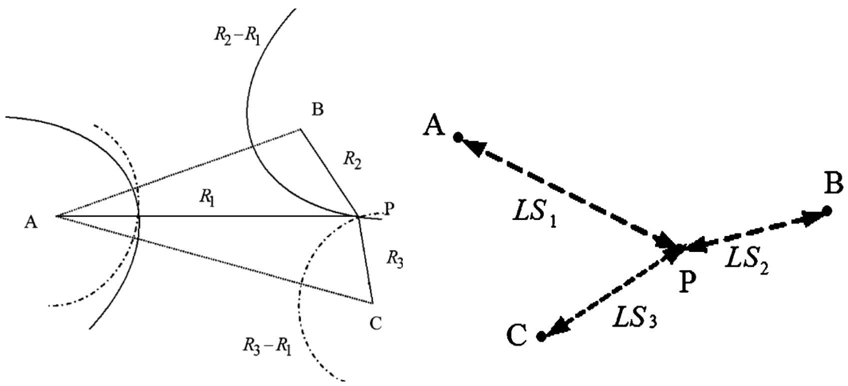


(a) AOA (Liu et al. 2007) (b) TOA (Liu et al. 2007)

Figure 2.1: Signal properties: AOA and TOA

Time of Arrival: TOA, which is a distance-based metric, is the time it takes a signal to arrive at a receiver from a transmitter (Sakpere et al. 2017). Multiple receivers then convert the measured TOA into the distance, and trilateration is used to determine the position (Wiig 2010). Figure 2.1b illustrates positioning based on TOA measurements.

Time Difference of Arrival: TDOA uses the time difference between the arrival of the signal at various receivers (Wiig 2010). The receivers record the time at which the signal arrives and then determine the relative position of the transmitter (Sakpere et al. 2017). In other words, TDOA is a measure of the difference in TOA at two different receivers. Figure 2.2a illustrates positioning based on TDOA measurements.



(a) TDOA (Liu et al. 2007) (b) RSSI (Liu et al. 2007)

Figure 2.2: Signal properties: TDOA and RSSI

Received Signal Strength Indication: RSSI is a measure of signal strength from a transmitter to several receivers (Wiig 2010). A higher RSSI indicates a better signal quality and a lower reduction or loss of signal strength. The recorded signal strength can be converted to distance, and used to determine the position (Sakpere et al. 2017). Figure 2.2b illustrates positioning based on RSSI measurements.

2.1.2 Positioning algorithms

Positioning algorithms are, as already mentioned, used to derive the actual position from the recorded signal properties (H. Huang and Gartner 2010), and the most commonly used approaches are proximity, triangulation, trilateration, and fingerprinting.

Proximity: Positioning by proximity does not provide an absolute or relative position (Sakpere et al. 2017), but rather an approximate position through the use of so-called anchor nodes. Anchor nodes emit a signal and have a known location. If a user, or other objects, receives a signal from an anchor node, it is assumed to be very close to the anchor node, thus the position of the user is set to be the same as the known position of the anchor node (Correa et al. 2017). If the user is within proximity of multiple anchor nodes, the position is set to be the same as the anchor node with the strongest received signal.

Triangulation: Positioning by triangulation (also known as angulation) uses the geometric properties of a triangle when trying to position an object (Sakpere et al. 2017). At least two reference points with known positions are required, from which an angle direction line can be derived from each of the points. The object is then assumed to be positioned at the intersection of the two (or more) angle direction lines (Correa et al. 2017).

Trilateration: Trilateration (also known as lateration) is quite similar to triangulation, but instead of focusing on angles, trilateration uses distance measurements to position an object (Sakpere et al. 2017). The position of an object is determined by finding the intersection between three (or more) circles. The circles are created by using the known position of reference points as the center, and the distance between each reference point and the object to be positioned as the radius (Correa et al. 2017).

Fingerprinting: Fingerprinting methods assume that at all positions, a unique combination of signals can be received (Correa et al. 2017). Positioning by fingerprinting involves creating a database containing the signals received at different positions, and then using this database as a tool when trying to find the position of an object (Sakpere et al. 2017). The received signals by e.g. a user's device are compared to the data stored in the database, and the position that best matches the received signals are assumed to be where the user is located (Correa et al. 2017).

2.1.3 Positioning technologies

Numerous technologies have been tested and applied for indoor positioning, with varying results. They all have their advantages and disadvantages, which result in some being more commonly used than others. An indoor positioning system (IPS) consists of either only one technology or a combination of multiple technologies (Wiig 2010). Brena et al. (2017) divided indoor positioning technologies into four different groups; *radio frequency, optical, sound*, and *passive* (i.e. without embedded information). Some examples of each type of technology are included below.

Optical technologies: Technologies that are restricted by line-of-sight constraints (Brena et al. 2017).

Infrared (IR) is a technology that uses electromagnetic radiation with longer wavelengths than visible light. An IR positioning system consists of an emitter, that emits an IR signal (non-visible light), and a receiver, which detects the emitted light pulses (Brena et al. 2017). IR systems can be divided into passive and active systems. An active IR system requires an IR tag, which emits a unique IR code regularly. The emitted code is then received by an IR sensor (Brena et al. 2017). A passive IR system uses e.g. thermal IR sensors, which measure the thermal radiation from humans within close proximity (Sakpere et al. 2017). In this way, the system can keep track if e.g. a person enters a room, as this leads to a change in the thermal radiation. A limitation with this approach is the existence of other sources of heat other than human beings, which may lead to disturbances (Sakpere et al. 2017). IR systems also have a limited coverage range and accuracy, due to e.g. the line-of-sight restriction.

Visible Light Communication (VLC) is a technology that transfers data via visible light, which is possible due to the light's ability to be switched on and off again within very short intervals (Brena et al. 2017). A positioning system using VLC consists of one (or more) light source(s), an image sensor and a line-of-sight communication channel (Sakpere et al. 2017). Each light source has a fixed position and a different flicker encoding, which is then received by the sensor. The sensor's position is determined to be the same as the most dominant light source (Brena et al. 2017).

Sound-based technologies: Technologies that use sound signals, which consist of pressure waves (Brena et al. 2017).

Ultrasound (US) does not penetrate solid walls (Svalastog 2007). The US waves are mechanical, have short-wavelength (Fallah et al. 2013), and do not interfere with electromagnetic waves. Positioning systems using US consists of ultrasonic tags, which are placed on objects (or users), and these tags can serve as either receivers or transmitters. Tags used as receivers have a fixed, known position, and transmitters are in motion (Sakpere et al. 2017). The position of an object is determined by using the time used for the signal to travel from transmitter to receiver (Brena et al. 2017). The transmitter emits a radio signal and an ultrasonic wave at the same time. The radio signals arrive at the receivers almost instantly, and then the receivers measure the time difference between the arrival of the radio signal and the ultrasonic wave. The time difference is converted to distance (Mainetti et al. 2014), and used to determine the position.

Audible sound is used to position an object by encoding information within the audible sound signals (Brena et al. 2017). This can be done by e.g. using watermarking of audio signals. By watermarking already available sound, like music playing at a mall, it appears undetectable by humans, but a receiver will be able to detect the different watermarks. Each watermark represents a different transmitter, and signal strength is used as an indicator of the distance between a transmitter (speaker) and a receiver (microphone) (Brena et al. 2017).

Radio frequency technologies: Radio-based technologies that employ radio signals (Brena et al. 2017).

Wi-Fi, which is also known as Wireless Local Area Network (WLAN), can be used to position an object within the network (Mainetti et al. 2014). A WLAN is a wireless high-speed network that transmits and receives data by the use of electromagnetic waves (Sakpere et al. 2017). Various approaches to determine the position of an object by the use of Wi-Fi signals have been implemented, and the most commonly applied ones are fingerprinting and propagation methods. Wi-Fi is a widely adopted positioning technology due to WLAN infrastructure being quite widespread in indoor environments (Mainetti et al. 2014). Fine Timing Measurement (FTM) is a protocol included in the IEEE 802.11 WLAN standard, which introduced Wi-Fi RTT (Round Trip Time). Wi-Fi RTT uses the time of flight instead of received signal strength (Huilla 2019). Time of flight measurements are supposedly more accurate than received signal strength, thus gives more accurate estimates of distance which can be used to determine position. Gjøvåg (2020) researched Wi-Fi RTT and conducted an experiment using a Google Pixel 3a and Google Wi-Fi base stations together with FTM to further investigate the accuracy. However, it was not found compelling evidence of the 1-meter accuracy claimed.

Bluetooth is a wireless technology that exchanges data through digitally embedded information on radio frequency signals (Brena et al. 2017). A Bluetooth positioning system consists of Bluetooth devices, Bluetooth tags or sensors, server, and WLAN (Sakpere et al. 2017), and the position of a Bluetooth device can be acquired with various techniques, like e.g. proximity, trilateration, or fingerprinting (Svalastog 2007).

RFID, which is an abbreviation for Radio Frequency Identification, is a technology that transmits the identity of an object through radio waves (Svalastog 2007). Objects wear RFID tags, which are scanned and identified by an RFID reader. An RFID tag consists of an antenna, that receives and transmits radio frequency signals, and a microchip, which stores and processes information (Wiig 2010). There are two types of RFID tags; active tags and passive tags. Active tags have a battery and they automatically transmit signals. Passive tags, on the other hand, do not have any battery installed and only transmit a signal when powered by an RFID reader (Fallah et al. 2013). An RFID positioning system can be used as a way to position either tags or readers, depending on which of the two already has a known position and which needs to be localized (Mainetti et al. 2014).

Ultra-Wideband (UWB) is a short-range radio technology that can be used to transmit information at very low levels of energy by using a large portion of the radio spectrum (Wiig 2010). Time of Arrival and Time Difference of Arrival, which was explained in section 2.1.1, can be used as measures to find the distance between the target and a reference point in a UWB positioning system (Brena et al. 2017). Also UWB-based positioning systems can be classified as either active or passive. An active system consists of mobile battery-powered UWB tags, fixed UWB sensors, a central software controller, and WLAN (Sakpere et al. 2017). The position of an object is acquired by the tags emitting ultra-short UWB pulses that are received by the sensors. A passive system uses signal reflection instead of active UWB tags, and fixed transmitters and receivers are installed in the area/room of interest. If a person enters the room, their body reflects the signals emitted (Sakpere et al. 2017), and then the receivers receive the reflected signals.

Passive technologies: Technologies that rely on naturally occurring signals, thus does not contain any embedded information (Brena et al. 2017).

Magnetic field can be measured by using a magnetometer and used to determine the position of an object (Brena et al. 2017). A magnetic positioning system consists of fixed transmitters and mobile receivers, which receive magnetic signals from the transmitters (Sakpere et al. 2017). Most modern positioning systems use Earth's natural magnetic field for the localization process, and they use the magnetometer to measure variations in the magnetic field (Brena et al. 2017). The position of an object is then acquired by using techniques like fingerprinting.

Computer vision uses the camera capture as a way of providing a position. The object to be localized carries a device with a camera that captures the surrounding environment (Fallah et al. 2013). The images captured by the camera are then compared against a database with images, with corresponding positions, portraying the environment, using image matching methods.

Hybrid systems use a combination of different technologies (Sakpere et al. 2017).

Commonly, one of the technologies in a hybrid system is more relevant for determining the position of an object, while the rest of the technologies are considered complementary (Brena et al. 2017).

2.1.4 Simultaneous localization and mapping

Simultaneous localization and mapping (SLAM) is the process of determining the location of an object, while concurrently creating a map (i.e. model) of the surrounding environment (Z. Wang et al. 2011). SLAM makes it possible to give an estimate of the position, without having any prior knowledge of the surroundings. The most commonly used setting for SLAM is a mobile robot moving through an environment consisting of various features (Correa et al. 2017). Equipped with numerous sensors, the robot is able to measure its own motions, in addition to the relative location between itself and the surrounding features (Z. Wang et al. 2011). As a result of this, it is able to estimate both its own position and the position of the features, within the map concurrently created of the surroundings. SLAM is regarded as one of the most important problems in the pursuit of developing and building autonomous mobile robots (Thrun 2008). Different methods of SLAM have been researched, and two of them are visual SLAM and LiDAR SLAM. Visual SLAM uses images captured by cameras or other image sensors, while LiDAR SLAM uses primarily laser sensors (MathWorks 2021).

SLAM have a variety of possible applications, and Z. Wang et al. (2011) mention some in their research. In e.g a hostile subsea environment there is usually a lack of infrastructure. Autonomous Underwater Vehicles (AUVs) need a system for navigating unexplored underwater environments, and for this purpose, SLAM could be very useful. Another similar application is aerial observations. Unmanned Aerial Vehicles (UAVs) can use SLAM to map unknown terrain, even when the position of the UAV also is unknown (Z. Wang et al. 2011).

2.2 Extended Reality

Extended Reality (XR) is an umbrella term that covers the different types of immersive technologies that can merge the real and a virtual world (Marr 2019). The range of this umbrella term goes from having a fully immersive virtual environment to only having a few virtual components brought into the perception of the real world. The most commonly known terms covered by XR are Virtual Reality (VR), Mixed Reality (MR), and Augmented Reality (AR).

VR experiences fully immerse the user into a simulated virtual environment (Marr 2019). To achieve this, a VR headset or another kind of head-mounted display (HMD) is required so that a 360-degree view of the virtual environment is obtainable. VR can be used for many different purposes, but some of the most common ones are gaming and entertainment (Bardi 2019). VR video games allow the user to fully immerse and interact with the virtual world. Other use cases for VR are educational purposes, like surgeons learning to perform brain surgery or a pilot learning to fly, and in construction, where e.g. contractors can discuss building models together regardless of their physical location. Porras et al. (2018) researched the use of VR in the rehabilitation of balance and gait. They concluded that VR has the potential of leading to improvement of balance and gait in patients with neurological disorders, especially when combined with conventional rehabilitation. The most popular and well-known VR systems are Sony's PlayStation VR, Facebook's Oculus Rift, and HTC Vive (Bardi 2019).

AR experiences let the user view the real world with virtual objects and information presented on top of it (Marr 2019), as a way of integrating the digital with the real. These kinds of objects and information can be text, images, 3D models, animation, etc., and their purpose is to enhance the user's perception of the real world, rather than replacing it like in VR (Ariso 2017). One of the most well-known AR experiences is *Pokémon Go*, which made AR grew rapidly in popularity in 2016. Pokémon Go is a mobile game that uses the camera of a mobile device to capture the real world and then lets the user collect different digital creatures, depending on, among other things, the user's location in the real world (Niantic 2021). Other use cases for AR are educational uses. Cabero-Almenara et al. (2019) implemented an application to further research the use of AR within educational uses and concluded that it benefited the learning process itself.

VR allows the user to interact with the virtual environment but lacks real-world interaction. AR allows the user to interact with the real world, but interaction with virtual objects is either non-existent or very restricted. *MR experiences* combine elements from VR and AR, lets the user interact with both the real and the virtual environment, and also let objects from the two environments interact with each other in real-time (Marr 2019). Pan et al. (2006) explored the use of virtual learning environments (VLEs), which include VR, AR, and MR, and discussed the possibilities of VLE by viewing different examples of existing applications. In Kyoto in Japan, a virtual sand garden designing system was presented as an aid for human mind therapy. The system uses MR to allow the user to interact with the virtual sand garden but removes the need for physical sand (Pan et al. 2006). MR is currently not as available as VR and AR to all kinds of users, and its use cases are not as widely developed. However, several companies are working on MR technology and already have developed MR products, like Microsoft's HoloLens, Lenovo Explorer, Samsung Odyssey, and Acer Windows Mixed Reality (Marr 2021).

Figure 2.3 shows a summary of the three types of extended realities; VR, AR, and MR, and indicates how they are different.

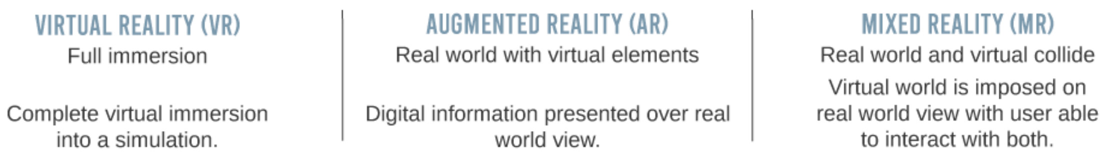


Figure 2.3: Overview of extended realities (GIGXR 2021)

2.2.1 Brief history of AR

Ivan Sutherland is by many considered "the Father of AR" (Picard 2020), and he gets this credit by creating the first prototype of an AR head-mounted display system, called "The Sword of Damocles", in the 1960s. The fundamental idea behind this HMD was, as explained by Sutherland himself, to "display the user with a perspective image which changes as he moves" (Sutherland 1968). The user would get an impression that he or she was viewing a 3D object, while in fact, he/she was only seeing a series of 2D images that changed accordingly to the change of position and rotation of the user's head, and thus the change of the position and rotation of the HMD. These illusions of 3D objects were laid on top of the real world, and therefore creating something very similar to what is now called AR.

However, the early stages of AR date back to even before Sutherland's work in the 1960s. The first aircraft heads-up display, which was implemented by the British military in their Mark VIII Airborne Interception Radar Gunsighting project in World War II (Vaughan-Nichols 2009). This system overlaid information regarding nearby aircraft, in addition to a radar screen, on the pilot's windshield.

The actual phrase "augmented reality" was coined by Tom Caudell back in the 1990s. He was working for Boeing at the time, with his colleague David Mizell, helping the workers wiring aircraft. They implemented a heads-up, see-through, head-mounted display, which was meant to "improve the efficiency and quality of human workers in their performance of manufacturing activities" (Caudell et al. 1992). The user was presented with the necessary information, related to their current task, as a way of augmenting their visual field.

2.2.2 Commercial AR

AR has continued to grow in popularity since the 1960s, and all the big tech names, like Google, Apple, Amazon, Facebook, and Microsoft, have explored AR in one way or another in recent years (Makarov 2021).

Google has developed several services using AR that are meant to make the user's everyday life more effortless (Google 2021c). *Google Lens* takes the search bar one step further, by searching using the camera lens. The camera lens captures the physical world around the user, and by pointing the camera at the subject of the search query, Google Lens will try to find relevant results (Google 2021c). The subjects of these searches can be clothing items to be found in online stores, equations to be solved, text to be translated, plants to be identified, or restaurant menus where reviews or photos of dishes are desired (Google 2021c). *Google Search* has added functionality that allows the user to use AR to place 3D digital objects in their own space. *Google Maps* have started exploring the possibilities of AR within navigation, where 3D models are overlaid the real world to help guide the user towards their destination (Google 2021c). In addition to this, Google has created a developer platform, *ARCore*, which provides developers with tools to create new AR applications.

Apple has produced hundreds of millions of devices that are AR-enabled (phones, tablets, etc.), and in Apple's App Store thousands of AR apps can be downloaded (Apple 2021d). This makes Apple one of the world's largest AR platforms. *ARKit* is Apple's own developer framework, which provides tools for developing AR applications for Apple devices (Apple 2021d), much like Google ARCore does for Android development.

Amazon has released an application that allows its users to design and decorate their homes with the use of AR (Amazon 2021). *View in your room* is an application that lets the user preview products, and see how it fits in a room, before buying it. The 3D models of the products are sized and rendered to scale, which gives a realistic fit in the real world (Amazon 2021). In addition to this, Amazon has launched an application called *Amazon Augmented Reality* in October 2020 (Perez 2020). Amazon Augmented Reality is an application that provides the users with an interactive and shareable AR experience when scanning QR-codes on the company's shipping boxes. Examples of these AR experiences are drawing a pumpkin on the cardboard box and then transforming it into an AR object, or turning the cardboard box into a blue car when the code is scanned (Perez 2020).

Facebook strives to keep creating content that will keep the users intrigued. One of the trends within social media is filters to use when either taking a picture or having a video call. These filters may use AR effects, like placing an accessory on top of the user's face, or making a digital robot appear in the room. Facebook has released a platform where the users can create these filters and effects, and this platform is called *Spark AR* (Facebook 2021). Spark AR allows the user to take advantage of expanded libraries of AR assets and makes it easy to share the effects on social media. According to Facebook, hundreds of millions of users use AR experiences monthly across all of the Facebook family of apps, which are Facebook, Messenger, Instagram, and Portal (Slater 2019).

Microsoft has been focusing on Mixed Reality (MR) lately. MR is similar to AR in some ways, but there are also some differences. The differences between AR and MR were already looked deeper into in section 2.2. Nonetheless, in MR digital elements are also brought into our perception of the real world. But instead of only enhancing this perception, MR allows interaction and manipulation of both the real and the virtual environment in real-time (intel 2021). *HoloLens* is a pair of MR smartglasses (i.e. wearable computer glasses) that are developed and manufactured by Microsoft. These smartglasses allow the user to e.g. touch, grasp, and move holograms similar to how real objects can be interacted with (Microsoft 2021a). HoloLens can also be used to interact with others through *Microsoft Mesh*. Mesh is an MR platform that allows people located in different physical locations to interact through the use of collaborative and shared holographic experiences (Microsoft 2021b). Mesh can be accessed by HoloLens, VR headsets, mobile phones, tablets, and PCs.

2.3 Augmented Reality

Augmented reality can be displayed in many different formats. As mentioned, all the big tech companies have used AR to their advantage and been innovative. However, all of these have the underlying technology in common where it all comes down to location and identification (Franklin Institute 2021).

According to Alan B. Craig (2013), the process of AR applications can be divided into two key elements that need to take place in each time step. The two steps are listed below:

1. An application has to determine the current state of the physical world and the current state of the virtual world.
2. An application has to then display the virtual world in alignment with the real world in a manner that will cause the user to sense the virtual world elements as part of their physical world.

At the end of step 2, the application would return to step 1 to move on to the next time step. Doing this over and over again will generate a display that merges the virtual and the physical world. To achieve a well-working AR system that supports the steps listed above, there is a need for three components (Craig 2013). The three components are sensor(s), processor, and display.

2.3.1 Sensors

Sensors are needed to obtain information about the real world in real-time. One of the sensor's primary and most common applications is to track the position, which includes both location and orientation of the user. This is crucial as AR depends on being spatially registered (Craig 2013). Common sensors that can be used to track the position are cameras, accelerometers, gysoscopes and GNSS, where it is worth noting that the latter is first of all not effective in indoor locations and consumes more power than the other sensors mentioned (Saha et al. 2015).

The technology behind the camera is computer vision, as this allows the application to determine the user's position. The camera registers the real world seen through the lens, and based on that, determines the location and orientation (Craig 2013). However, this requires that landmarks are placed in the environment in question, where they are used by the software to position the user.

According to Franklin Institute (2021), we can separate AR into being either marker-based or markerless. Marker-based AR uses image recognition to identify items that are predefined in the given AR application. Items that are simple whilst being distinct are preferable, such as a QR code. Marker-based AR firstly converts the camera feed to greyscale to speed up the processing time (Franklin Institute 2021), and then goes on to detecting markers and compares the recognised markers with markers stored in its database. After recognising a marker, it will mathematically be able to place the user at a known position and display AR models in the right location on the screen. An example is when having a stroll through a museum and being able to hover the camera over a QR code and get supplementary information on the specific artifact.

On the contrary, markerless is somewhat more complicated as there are no predefined items in the application (Franklin Institute 2021). This results in the application needing to recognise items, surfaces, patterns, colours, and other features on natural features in the environment as the user walks with the device. As Alan B. Craig (2013) explains, the technology of natural feature detection is developing rapidly. An example of this is the new feature within Google Maps called Live View Beta. In this feature, they have integrated AR in walking navigation to "..quickly orient yourself and know which way to go.." (Google 2021c). The user is initially asked to point the camera at buildings and signs across the street, which indicates the use of a recognition algorithm.

Other use cases of the sensors are to gather information of the present environment and to gather input from the user. Sensors used for gathering environmental information are less commonly used in AR systems (Craig 2013). They can obtain information about the physical world in present time, such as information on the temperature and humidity. However, these are not that frequently used in common AR applications. The previously mentioned sensors relies on passive users. Yet, there are several sensors that requires the user to actively take action. These kinds of sensors are touchscreen, keyboards, and other components that are typical in a user interface. Using these will increase the interaction with the user and allow the user to have an input in their experience of the application (Craig 2013).

2.3.2 Processor

Another component needed to achieve a satisfactory augmented system is *a processor* (Craig 2013). This component is vital as it coordinates and analyses the information obtained by the sensors, pulls and fetches, and performs the appropriate tasks. One thing that Alan B. Craig (2013) emphasizes is that the processor needs to have enough computational power to perform tasks in real-time. He continues by saying that the system must provide a reaction presented in a display whenever the user acts, and the response must be performed without any noticeable lags. This is called feedback and goes under one of Jakob Nielsen's 10 heuristics for user interface design. Nielsen (1994) states that the application always should keep users updated on what is going on by giving the feedback within a reasonable amount of time. By taking feedback into account, the application is secured to run smoothly in the eyes of the user, which is one of the main tasks of the processor.

2.3.3 Display

Besides sensors and processors, the last component that is listed by Alan B. Craig (2013) as a necessity in an AR system is *a display*. Various types of displays can be utilized in AR. When categorizing by the sense they affect, the two most common displays are audio, such as headphones and speakers and visual display, such as a computer monitor. Otherwise, another categorization is to separate on whether or not the display is attached to the user or not. An example on a display that is attached to the participant is a head-mounted device, whereas an example on a display that not is attached to the participant is a computer monitor (Craig 2013). A display that could go under both of these categories is a smartphone, since the visual display is held in the hand when needed.

2.3.4 Augmented reality experience

The three vital components presented earlier are only a smaller part of the full AR experience. Alan B. Craig (2013) includes six building blocks needed in an optimal AR experience, which are listed below.

1. Augmented reality application
2. Content
3. Interaction
4. Technology
5. Physical world
6. Participant(s)

The *AR application* is explained as the computer program that fundamentally controls and executes the different tasks in the application and is connected to the sensors, devices, and displays that are used in the provided experience (Craig 2013). An AR application should therefore be able to be used in multiple scenarios where the basic task is similar. An example of two scenarios where the same AR application can be utilized, is for postal services to examine the size of a parcel and for children to explore various animals. In both these examples, the scenarios have their basic task in common; placing and examining an object. *Content* is different from AR application defined above and includes everything that does not fall under any of the other six building blocks. It includes, inter alia, the story, objects, how the objects should behave, game rules and simulations (Craig 2013).

Interaction is defined as an occasion where people or objects communicate with or react to each other (Cambridge Dictionary 2021). When considering interaction in an AR experience, Alan B. Craig (2013) mentions possibilities. Some of these being between participant and AR application, virtual world and the real world, participant and virtual world, and participant and real-world. Every AR experience has to be interactive. That could be as simple as allowing the user to press buttons to interact with the experience. A building block that is obvious and almost inevitable is *technology*. As Alan B. Craig (2013) says "Every AR experience does involve technology". The technology needed in a basic application is made up of the three important components mentioned above, which are sensors, processors, and displays.

The two last building blocks in an AR experience are the *physical world* and *participant(s).* All AR experiences take place in the real world, contrary to VR which is a full immersion into the virtual world. However, the physical world could be both a specific place or generic space that is used to represent the physical world (Craig 2013). The potential in AR technology makes it possible to make the physical world have components that give the participant the perception of it being the virtual world. The art of an AR experience is to convince the participant that something is happening that, in reality, is not (Craig 2013).

2.4 AR Technologies

As newer handheld mobile devices have the capacity to host AR experiences, AR has been embraced by various markets in the past few years, including the consumer market (Chandra 2019). As the popularity of AR increases, so does the demand for development tools needed for AR applications. To meet these demands, big tech companies have released powerful platforms (ISL 2018).

Two major tech companies, Apple and Google, each made their own development platform. Respectively, ARKit in 2017 and ARCore in 2018 (Leon 2020). The competition between the two also creates a stride to keep up in the industry (Makarov 2020). In addition to ARCore and ARKit, SDKs such as Vuforia and AR Foundation have become available for developers, which has made developing AR applications for mobile devices easier than ever.

2.4.1 ARKit

ARKit is a framework created to aid developers when creating AR applications for iOS devices (Chandra 2019). Applications made by using ARKit can only be used on iOS devices, which implies that it can not be used on e.g. Android devices (Leon 2020). Apple has now released its 3rd iteration of ARKit and has each time brought new features to the table. The platform can be divided into three layers that all work simultaneously together. These layers are tracking, scene understanding, and rendering (To 2021).

A key feature of ARKit is the ability to create and track a correlation between the real world and the virtual world (Apple 2021c). ARKit follows a right-handed convention when it comes to both the world and camera coordinate system, with the y-axis pointed upwards, the z-axis pointing toward the viewer, and the x-axis pointing right, see figure 2.4. ARKit uses a technology called visual-inertial odometry (VIO), which combines information from the device's motion-sensing hardware with computer vision analysis on the scenes available through the lens of the device (Apple 2021c). The sensing hardware are cameras, gyroscope, accelerometer and motion sensors (Ridland 2019). These inputs obtained by VIO allow the device to correctly sense how it moves in the room, without needing additional calibrations

(AppleInsider 2021). Along with tracking the movement, it also understands the scene. It analyzes and understands the content of the room and detects surfaces and planes (Apple 2021c). The latter is a feature called *planeDetection* that ARKit has in its arsenal which can be enabled in the session configuration. When enabled, the developer can easily add virtual objects on the different planes in the virtual world.

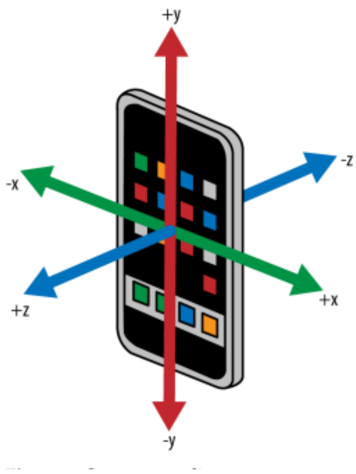


Figure 2.4: Camera coordinate system (Valvo et al. 2021a)

Rendering is possible by using various technologies. Apple provides developers with three APIs where each of them has various area of use. *SceneKit* is used to render 3D overlay content, whilst *SpriteKit* is used to render 2D overlay content. The last option is *Metal*, which is used when wanting to create a rendering engine for the AR experience (Apple 2021b). With the latter, the developer gets more control over the rendering of the app's virtual content displayed over the camera stream. The technologies mentioned are Apple's own technologies. However, ARKit is also compatible with multiple third-party tools, such as popular Unreal Engines or Unity IDE (Ridland 2019).

It is given that the real world environment lacks consistency, which may lead to the occurrence of unpredictable factors. To still produce quality AR experiences, Apple presents some advice and limitations one should be aware of before creating such experiences in their documentation (Apple 2021c). When it comes to lighting conditions, Apple requests developers to "design AR experiences for predictable lighting conditions". Being as ARKit uses image analysis when tracking, it needs a clear image to work properly. This implies a lacking tracking features if the scene is too dark or the camera is pointed at a blank wall. However, ARKit has a feature that estimates the light of the environment and makes it possible to apply lighting effects to virtual objects. Being able to change the lighting and shadowing on virtual objects makes it look more realistic (AppleInsider 2021).

Another piece of advice Apple presents is to let the user know how to resolve low-quality tracking situations. There are a few factors that badly impacts tracking. As mentioned, the lighting is a uncontrollable condition that makes the tracking more difficult for ARKit. Besides that, the tracking uses device motion in image analysis (Apple 2021c). Using the motion of the device gives a better understanding of the scene. An excessive motion may result in reduced tracking quality, as it can result in blurry images or a larger distance for tracking features between frames. Giving the user information on how to better the quality will decrease the amount of bad experiences. The ARKit framework's class *ARCamera* is helpful in these situations, as this provides possible causes for lacking quality by using the position tracking from frames captures by the camera (Apple 2021a). The last piece of advice Apple gives in their documentation is regarding plane detection. They encourage developers to allow enough time for accurate plane detection and clear results but disable plane detection when the wanted results are achieved. If allowed for longer than that, ARKit may continue changing the plane anchor's position, extent, and transform still after, e.g., an object is placed on the plane (Apple 2021c).

The various features mentioned above are some reasons as to why ARKit has been used to create multiple augmented reality applications throughout the years. ARIANNA+ is a system created by using ARKit, and is specifically designed for visually impaired people for both indoor and outdoor localization and navigation (Valvo et al. 2021a). Valvo et al. (2021) previously developed a system which was based on the recognition of landmarks, such as QR codes and physical paths (e.g. coloured tapes). However, by utilizing the libraries of ARKit and SceneKit, the researchers were able to step away from landmarks and build an entire virtual path. The SceneKit library contains components such as SCNScene, SCNNode, and SCNGeometry, which are used when creating the path within the 3D world. The path is later saved as a map. It should be noted that visually impaired people cannot see the line on the device. However, the creators of this system state that they include tactile vibration for when the line is located at the center of the camera (Valvo et al. 2021a). Similar to ARIANNA+, Fusco et al. (2018) created a wayfinding system for people with visual impairments. However, they combined VIO with computer vision-based sign recognition. Sign recognition is used when signs are visible and recognized, while VIO is used when no signs are visible on the camera stream. They mention that the reason for combining these two technologies was because several other indoor localization technologies often require installing and maintenance of physical infrastructure, or an update of existing systems (Fusco et al. 2018).

Apple has made an effort to keep everything related to AR easily understandable and generally easy for developers to utilize. It should be noted that there might be less resistance when working with ARKit for iOS devices, as Apple software often is closely tied to Apple hardware (Leon 2020). They have done so by releasing several ARKit tutorial materials and updates with every new version of iOS (Program-Ace 2019). However, as mentioned, applications created with ARKit are only compatible with iOS devices. Additionally, ARKit has a price of $99 a year for AppStore distribution (Leon 2020).

2.4.2 ARCore

ARCore is similar to ARKit by being developed by a platform for a specific type of devices, where the platform is Google and the devices are Android devices. However, in contrast to ARKit, ARCore work on both Apple and Android devices. The ARCore platform is compatible with Unity3D and Unreal Engine (Leon 2020). Similar to ARKit, ARCore has three main features that make it possible to create a wholesome AR application that integrates the virtual content with the real world. The three key features are motion tracking, environmental understanding, and light estimation (Google 2020).

Motion tracking is necessary to understand where the device is placed relative to the world surrounding it (Google 2021e). ARCore uses the process previously mentioned called SLAM. The platform detects distinct features, called feature points, in the camera stream captured by the device. The feature points are then used to compute the device's change in location. To estimate the position and orientation of the camera relative to the world, the visual information is merged with inertial measurements from the device's IMU (inertial measurement unit). Aligning the position and orientation of both the device's camera and the virtual camera, allows developers to render virtual content in the real world that looks realistic from all angles (Google 2021e).

Understanding the environment involves detecting both horizontal and vertical surfaces and planes. ARCore tries to look for clusters of feature points that lie on common surfaces, such as tables or walls, and make them available as planes. In addition to that, ARCore has the possibility to determine each plane's boundary which allows the user to place objects resting on flat surfaces. As ARCore uses feature points to detect planes and flat surfaces, white walls without texture may not be detected (Google 2020). The last key feature that ARCore uses to visualize the virtual content more realistically, is *light estimation*. ARCore obtains information about the lighting of the surroundings and provides the developer with the average intensity and colour correction of the given camera image. This allows the application to make an illusion for the user where the lighting on the objects matches the environment. It also increases the sense of realism (Google 2020). Even though these features are only three of multiple features in ARCore, they are key in creating such a wholesome AR experience.

ARCore advertises with being compatible with Android (Java), Android NDK (C/C++), Unity, Unreal and iOS, and provides software development kits (SDK) for each of these. These SDKs provide native application programming interfaces (API) for features that are important when building AR experiences, like motion tracking, environmental understanding, and light estimation (Google 2021d).

ARCore has since 2018 been used in many applications. Zhang et al. (2019) proposed an assistive navigation system for visually impaired people that utilizes ARCore's feature to attain computer vision-based localization. The feature being discussed is SLAM, which the system used to track motion to create a scene understanding and acquire a better mapping and tracking (Zhang et al. 2019).

Google offers great in-depth documentation on how ARCore could be used within the various platforms (Leon 2020). In addition to the official documentation by Google, multiple unofficial guides thoroughly explain each step in the build of an ARCore application. For instance, Hendrickx (2019) wrote an article explaining how to create an ARCore powered indoor navigation application in Unity. A great example of the potential of ARCore is the new indoor navigation AR feature in Google Maps that navigates the user of the application by displaying arrows and markers on top of the camera stream (Haselton 2021). It does not cost anything to use ARCore (Leon 2020).

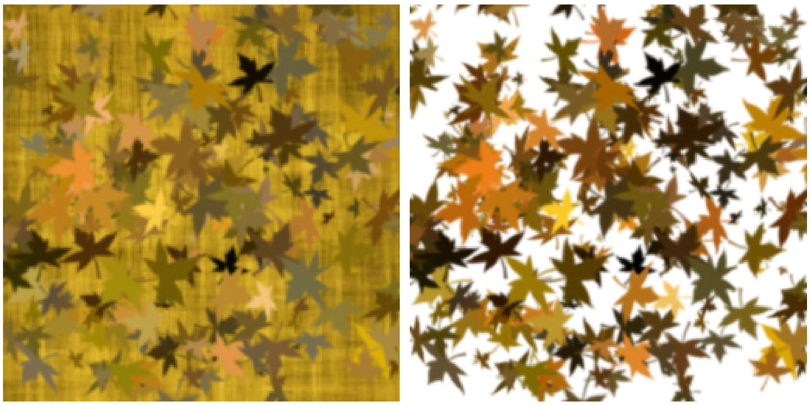
2.4.3 Vuforia

Vuforia is a popular SDK which supports Android, iOS, UWP (Universal Windows Platform) and Unity. Furthermore, Vuforia can run on iPhone's older models that even ARKit is not compatible with (Josh 2020). Vuforia has in recent years expanded its line of AR tools, which includes products such as Vuforia Engine and Vuforia Studio. Vuforia utilizes computer vision technology to recognize and track 3D objects in real-time, which is one of the reasons it is popular (Nikitin 2020)

The latest versions of Unity have Vuforia built in, which makes it easier for developers as they do not have to integrate it themselves. It is worth noting that Vuforia Engine, one of the previously mentioned products, is mainly used with the Unity3D game engine. There is extensive developer documentation on how to use Vuforia Engine with Unity on both Unity3D's forum and Vuforia's website (Leon 2020).

The SDK of Vuforia can be divided into three parts: image recognition, object recognition, and additional features (Nikitin 2020). As mentioned, the ability to *recognize images* in real-time is something that stands out with Vuforia. Vuforia offers an application called *Vuforia Image Target*, which rates an uploaded image from 0 to 5, see example in figure 2.5. This allows the developer to immediately get feedback on images and choose an image with a higher score (Nikitin 2020).

In addition to Vuforia Image Target, Vuforia offers VuMark. VuMark is an image similar to a QR code, created by Vuforia's own developers with a goal to naturally fit into an AR app and be less distracting (Nikitin 2020). The Vuforia markers have some drawbacks, as presented by Delfa et al. (2016) in their research. Drawbacks such as that the source code not being accessible, which makes it impossible to modify the predetermined algorithm, and that the number of markers is fixed, which gives the developer low flexibility. The last drawback they present is that it is not possible to reduce the size of the marker without having inaccuracy and lack of



(a) 0 stars (Nikitin 2020) (b) 5 stars (Nikitin 2020)

Figure 2.5: Example of an image rated 0 stars and of an image rated 5 stars.

performance (Delfa et al. 2016).

Vuforia offers two methods of data storage for these specific target images: local and cloud. With the local database the developer would have to download the whole database, add the database to the given application and connect it. This is something that would have to be repeated each time the content of the database is updated. Whereas, with the cloud method, the developer would only have to add/remove a picture from the database. Given that the device has an internet connection, the update would be detected in the cloud as well (Nikitin 2020).

Object recognition is offered by Vuforia. The technology they call *Model Target* can identify a 3D object in real-time (Nikitin 2020). For the object recognition to reach its full potential, the developer has to upload a 3D CAD model or a 3D scan of the object obtained to Vuforia Target Manager and follow some requirements set by the framework regarding attributes, such as the coordinate system and textures/colours. When it comes to *additional features*, Nikitin (2020) mentions extended tracking and ground plane. Extended tracking provides a target's position and orientation even though the target is no longer in the field of view. Meaning that the position and orientation of the target are maintained by the application even though it is out of sight, with respect to the real world (Vuforia 2021a). Ground plane enables content and objects to be placed on horizontal planes, such as floors and top of tables. It also allows the placement of content in mid-air, by using Anchor points (Vuforia 2021b).

By using computer vision, Vuforia offers both marker-based and markerless AR.

This SDK has features that make it one of the best for object recognition and 3D modelling (Program-Ace 2019). Some of these features are *ground plane*, which supports adding content to horizontal surfaces, and *VuMarks*. VuMarks designs are unique designs created for every unique object and offers developers a scalable way to identify objects to have a trackable AR target in the virtual world (Vuforia 2021d). As mentioned, Vuforia offers marker-based AR which Romli et al. (2020) used in their research, where they created a mobile marker-based AR indoor library navigation application. They registered a target image to Vuforia which included the placement of the virtual object that would appear after the target image was scanned. The database needed to be downloaded and imported to the Unity3D libraries for it to properly work (Romli et al. 2020).

Vuforia has a free developer plan that makes it possible to use before the app is deployed. However, after the app is deployed, the pricing for the basic plan starts at $42 per month, which is higher compared to other presented AR tools (Vuforia 2021c).

2.4.4 AR Foundation

AR Foundation is a package within *Unity* that helps build cross-platform AR applications (Lee 2021). Without this, the developer would have to use SDKs for iOS and Android, which results in twice as much development effort and a more complicated codebase. AR Foundation includes ARKit and ARCore XR packages, which allows the developer to develop an AR application in Unity and go on to build it for Android and/or iOS (Leon 2020). Figure 2.6 illustrates the concept of AR Foundation and how it unites ARKit and ARCore SDKs.

The framework supports features such as those presented in table 2.1 (Unity 2021a). Unity encourages to refer to this table when developing to understand what parts of AR Foundations are relevant on the various platforms. The original table includes Magic Leap and HoloLens. However, these are excluded as they are dedicated to HMD, which makes them less relevant when discussing indoor navigation on mobile devices (Leon 2020).

Some features are implemented by only one of the specified platforms, which could have been a problem. Nevertheless, AR Foundation has a solution for some of these

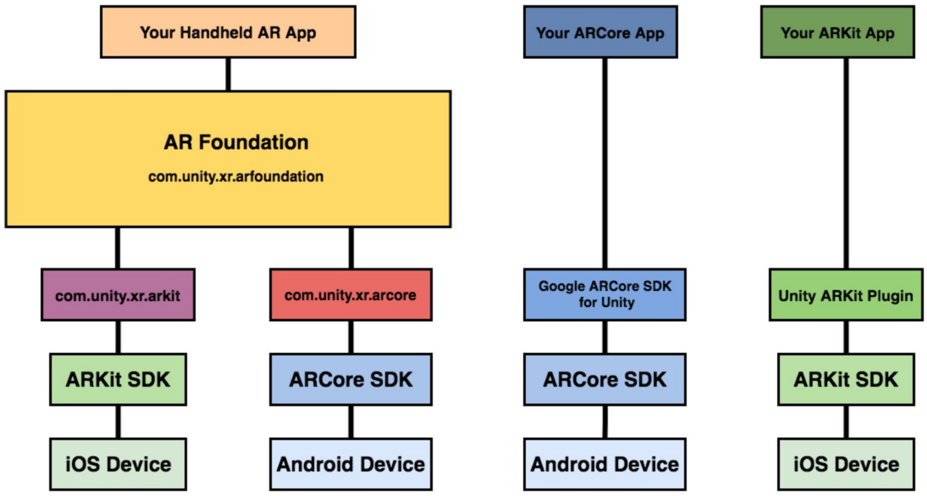


Figure 2.6: Concept of AR Foundation (Miller et al. 2018)

obstacles. An example is if a developer builds an iOS (ARKit) AR application that uses a 3D-object tracking feature that recognizes real world-objects of a certain shape. While ARKit supports such a feature, ARCore does not. For the feature to work with both of these platforms, AR foundation adds special hooks on the objects which result in an application that works on Android devices as well, despite the lack of feature. Whenever this feature is available with ARCore, the developer only needs to update the AR Foundation app package without needing to rebuild the app (Leon 2020).

AR Foundation is built on subsystems. Subsystems are defined as a *platform-agnostic interface* for various types of information (Unity 2021a). A platform-agnostic interface is described as an interface that works equally well across multiple platforms (Techopedia 2021). Each of these subsystems treats a separate and specified functionality. It should be noted that the subsystem defines method instances, and does not have any implementation details. The implementation details are written in different platform-specific XR plugins, such as ARKit and ARCore (Lee 2021). A concrete implementation of a subsystem is called a provider. E.g. the *ARKit XR Plugin* package contains the ARKit implementation for many of the AR subsystems. These subsystems have a descriptor that specifies which features the specific subsystem supports. This is helpful since different providers have varying support for specific features (Unity 2021a).

Table 2.1: Feature support per platform in AR Foundation

	ARCore	**ARKit**
Device tracking	✓	✓
Plane tracking	✓	✓
Point clouds	✓	✓
Anchor	✓	✓
Light estimation	✓	✓
Environment probes	✓	✓
Face tracking	✓	✓
2D image tracking	✓	✓
3D object tracking		✓
Meshing		✓
2D & 3D body tracking		✓
Collaborative participants		✓
Human segmentation		✓
Raycast	✓	✓
Pass-through video	✓	✓
Session management	✓	✓
Occlusion	✓	✓

To use AR Foundation, the developer has to install at least one platform specific package, such as ARKit XR or ARCore XR. Similar to ARCore, AR Foundation has extensive documentation available on the Unity document website and contains thorough information on how to use and download the framework (Leon 2020). AR Foundation is a part of Unity3D, which consists of a large community with videos, tutorials and forums that provides guidance to developers on their way to a well-working AR experience.

Given that AR Foundation is a framework within Unity, a standard Unity pricing system applies. It is free unless the project makes more than $100 000 in a year. If so, it can cost anything from $40 to $150 per month, depending on the type of functionalities (Unity 2021b).

2.4.5 AR indoor navigation technology

The technology surrounding AR indoor navigation consists of three vital modules to develop a complete AR indoor navigation application (Makarov 2019). The three modules presented by Makarov (2019), are positioning, mapping and rendering. This is illustrated in figure 2.7.

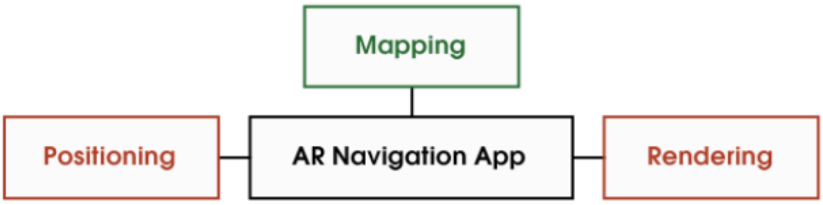


Figure 2.7: Three modules that AR indoor navigation technology consists of (Makarov 2019)

Mapping includes the process and information required before making an estimate of a route. Two vital items in mapping are the map and coordinates (Makarov 2019). Having these in place makes it easy to draw a route. In addition, this module is easy to upgrade and customize with respect to the use case in question. The module called *rendering* includes the layout of the AR content displayed on the screen (Makarov 2019). Matching the virtual objects with the real world, when e.g. drawing the route, is dependent on the accuracy of the positioning. The rendering is done by the various AR technologies. *Positioning* is more challenging to implement compared to the other modules as there are several ways of determining the location of a user indoors (Makarov 2019). However, there are multiple technologies available today where the accuracy is admissible.

There are various ways to creating maps and integrating mapping in the AR technologies such as Apple's ARKit and Google's ARCore. One method of creating these maps is to draw a detailed map, add visual markers that are supported by multiple AR SDKs, build the route using graph theory and drawing polylines and/or arrows in the real world using AR (Makarov 2019). Visual markers, known as AR markers or ARReferenceImage, are images recognized by AR SDKs, such as ARCore, ARKit, and Vuforia, that let the application know where to place the specific AR content. The scalability of this method is an issue as drawing such a detailed map of a large campus is time-consuming and the increase of markers can decrease performance (Makarov 2019).

2.5 Existing solutions

Many existing indoor navigation systems are developed specifically for a certain venue or a company, and are often only available through applications or websites that belong to this specific venue/company. Data belonging to outdoor spaces are usually more obtainable, thus also making it easier to create outdoor navigation systems which cover all outdoor space. The localization of devices outside is also generally more obtainable, as satellites provide coordinate data, while within indoor spaces other positioning methods have to be applied (Kunhoth et al. 2020). Data and information regarding indoor spaces are usually only acquirable by the proprietor of the specific building, which also means that the ones that possess this information have to specifically provide it or request the service. Examples of this are indoor navigation systems created for a company where only the employees have access or a system that provides navigation instructions to visitors at a mall and is only available through the mall's website or mobile application. As a 2D floor map is the most commonly used interface (B. Huang et al. 2020), not many indoor navigation systems with an AR interface exist as of today, especially not ones developed for the general public. The request for navigation services at various large buildings, that has to be set up and customized to each venue, has increased in the past few years (Kunhoth et al. 2020). Several templates and building blocks, that makes it less complicated to implement a fitting system for the specific venue, have also emerged.

2.5.1 Templates

Indoor navigation with AR technology has been explored as an opportunity with much potential for venues like museums, where the user can be guided through the building, in addition to being presented with information regarding the exhibition. ViewAR is one of the providers of templates for AR indoor navigation systems for e.g. museums. ViewAR's INDOAR template provides guidance and assistance through a virtual character (ViewAR 2021), which leads the user to the desired destination and can provide additional information of any point of interest. The template has two different variants of positioning and/or tracking; QR codes and visual recognition. The QR codes are marker-based, where the user scans QR codes placed on the walls, floors, or similar for localization, and the codes work as anchors with a known location (ViewAR 2021). For venues that do not visually change, visual recognition can be used after creating a 3D scan of the building. ViewAR is just one of many tools that provide technology and templates that make it easier to implement AR indoor navigation. Other similar products can also be obtained through the use of e.g. WaypointAR (Gnanamoorthy 2021), Wemap (Wemap 2021), and IndoorAtlas (IndoorAtlas 2021c).

2.5.2 Related work

AR-based indoor navigation systems have also been the subject of various research-based works. B. Huang et al. (2020) developed the navigation system ARBIN (Augmented Reality Based Indoor Navigation system) as an alternative to the commonly applied 2D navigation map. When using a 2D navigation map, with route instructions, the user has to make the connection between the instruction and the real environment. However, with the use of an AR-based navigation system, there is no need for the user to make this connection (B. Huang et al. 2020), which in turn could decrease the cognitive load. ARBIN conducts positioning by utilizing Bluetooth beacons and RSSI, which are located at intersections and points of interest, and uses Google's ARCore to create AR 3D models and obtain data to place the models in the real world. The route is planned by the use of Dijkstra's shortest path algorithm and then presented by a combination of 3D models and navigation instructions (B. Huang et al. 2020). INSAR is another AR-based indoor navigation system, presented by Alnabhan et al. (2014), which uses Wi-Fi-fingerprinting to provide low-cost positioning. Wi-Fi fingerprinting is regarded as a low-cost positioning method due to the existing infrastructure of Wi-Fi network and Wi-Fi chips in most modern smart devices. In their research, Alnabhan et al. (2014) found that both the software and hardware of the device used affected the accuracy of their results. Sato (2018) evaluated the accuracy of indoor navigation with AR markers, and also addressed the problem of battery drain, as it is necessary to keep the camera on when navigating with these markers. When continuously operating the camera of the mobile device, the battery is drained at a faster rate than by intermittent camera operation. A dynamic approach to camera operation was concluded to reduce battery consumption by approximately 30% (Sato 2018). Some research-based work investigates the need and possibilities regarding systems specialized towards groups with special needs. Valvo et al. (2021) implemented a localization and navigation system, called ARIANNA, which is the predecessor to the mentioned system ARIANNA+. ARIANNA is explicitly designed for visually impaired people and provides automatic guidance along a route, by the use of haptic, speech, and sound feedback (Valvo et al. 2021b). The system takes advantage of AR technology to eliminate the need for physical support, by using Apple's ARKit.

2.5.3 Google Maps

Google Maps was originally an outdoor navigation system but has in the last few years started to launch functionality for indoor use as well. Unlike applications made specifically for different venues, Google Maps is created for the general public, as it covers places open to all.

Live View is a feature in Google Maps, which was released in 2019, that allows the user to see which way to walk with the help of AR (Inman 2019). Live View is intended for pedestrian use, and it requires that the device used is compatible with ARKit or ARCore and that the area of interest has a good Street View coverage. In March 2021, Google announced that they were launching AR directions for indoor spaces (Glasgow 2021), which means that the Live View feature now includes indoor navigation, as well as outdoor navigation. Live View takes advantage of a technology called *global localization*, which is a way of improving navigation. Global localization combines Street View, Visual Positioning Service (VPS), and machine learning, and is then able to calculate and determine the user's position and orientation (Reinhardt 2019). Google Street View contains tens of billions of images, that visualize the world, and by scanning these images using AI, a more accurate position and orientation can be achieved. VPS creates maps by analyzing imagery and finding key visual features, like outlines of buildings. The key visual features are then stored in a large searchable index, which makes it possible to compare them to input provided by a device's camera. The VPS index is in this case created by analyzing the Street View images, which have a known location (Reinhardt 2019). As the Street View data set is rich and contains billions of images, covering over 93 countries across the globe, it is suitable for creating strong reference points. These reference points can be used to determine the position and orientation of a device by the use of triangulation. Machine learning is applied to detect features in the images that are most likely to be permanent and disregard temporary features, like trees that change depending on the season. Global localization increases accuracy (Reinhardt 2019), and the increased precision enables the possibility of building AR experiences with platforms like ARCore. Indoor Live View works very much the same way as Live View does outdoors, but is currently only available in certain malls, airports, and transit stations in some of the biggest cities in the world. The navigation instructions are provided to the user through the use of arrows and accompanying directions, placed on top of the real world in real-time, that points the user in the right direction (Glasgow 2021).

Chapter 3

Method

In this chapter, the structure of both the implemented application and the user test is looked deeper into. In section 3.1, the process of implementing the AR navigation system, and the technologies used, are explained more thoroughly, while in section 3.2 the final design and functionality of the solution is presented. Section 3.3 presents the content of the user test is in focus, which was conducted to evaluate the implemented application.

3.1 Implementation

To further investigate and get a more practical insight into the current status and possibilities of AR within the field of indoor navigation, a navigation system was implemented. This was done by using Google's ARCore, IndoorAtlas API, and data provided by MazeMap. ARCore is previously explained in section 2.4.2, while the other technologies will be looked further into in this section.

3.1.1 IndoorAtlas

IndoorAtlas is primarily a platform that provides indoor positioning for mobile applications, but also has additional features like wayfinding, geofencing, asset tracking, and location intelligence. According to their web page, their vision is "to create a seamless indoor world that is discoverable for people and businesses around the world" (IndoorAtlas 2021c). IndoorAtlas offers an SDK that integrates with an application and the sensors of a smart device to help provide e.g. indoor positioning. When trying to position a device indoors, the IndoorAtlas SDK observes each data source, e.g. Wi-Fi, magnetic, beacon, or any other sensor, conditionally. It also observes how much each data source's contribution varies based on the changing accuracy and probabilistic models (IndoorAtlas 2021d). The built-in inertial sensors of the smart device track the movement of the device, and by the use of dead reckoning a relative position can be obtained. Earth's magnetic field, Bluetooth beacons, and Wi-Fi signals are used together to obtain a more absolute position. The built-in barometric sensor can indicate whether the device has moved from one floor to another, given changes in the atmospheric pressure (IndoorAtlas 2021d).

With the release of version 3.4 of the IndoorAtlas SDK, the Augmented Reality API was introduced (IndoorAtlas 2021b). The IndoorAtlas AR fusion API merges the information gained from the use of IndoorAtlas positioning SDK and AR frameworks, like ARCore and ARKit. Both these sources of information are tracking the movement of a mobile device, but they focus on some different elements. The IndoorAtlas positioning SDK focuses on the *global* position and orientation of the device, which means that it will provide information about the floor level and the position within this level. The AR frameworks, on the other hand, focus on the *local* position of the device and tracks the motion in a local metric coordinate system. This means that the change of position can be recorded quickly and accurately, but there is no information on the global position. By combining these two sources of information, both the approximate global position and accurate change of this position can be obtained and used (IndoorAtlas 2021b). The AR fusion API also assists with the conversion between the global and local coordinate systems, which is essential when placing geographically referenced content in the AR world. The global coordinate system consists of latitude, longitude, floor number, and heading/elevation, while the local coordinate system is assumed to be a right-handed 3D metric coordinate system where the Y-axis points up (IndoorAtlas 2021b).

3.1.2 Fingerprinting

To be able to use IndoorAtlas SDK and API, a location, map, and a positioning technique had to be settled. IndoorAtlas provides a web application where creators may create new locations. A location is a container for related floor plans (IndoorAtlas 2018d), and typically corresponds to a building. Each location contains one or more floor plans for each floor of the specific building or area. The floor plans are geo-referenced on the world map, and by that bound to the global coordinate system WGS-84 (IndoorAtlas 2018a). Floor plans are also given a floor number to determine the correct vertical ordering within a location.

As previously mentioned, fingerprinting (also called mapping) is the process of gathering signal data from a target venue, and by doing so, enabling indoor positioning (IndoorAtlas 2019a). Positioning by fingerprinting assumes that there is an almost unique signal that can be recorded for every position. By measuring the signals across buildings, the received signals can be compared to existing records, and then the position can be estimated based on these signal measurements (AVSystem 2019). The process of fingerprinting is quite an important part of the positioning process, as the fingerprinting quality directly affects the accuracy of the positioning. In other words, high-quality mapping is the key to quality indoor positioning (IndoorAtlas 2019a).

Before doing the actual fingerprinting, different waypoints has to be placed out on the floor plans. These waypoints act as confirmation points for the data collection paths (IndoorAtlas 2018b), which are paths planned out based on where potential users are assumed to move within the floor plans. When placing the waypoints, it is expedient to place them every 5-10 meters along a path, and near distinct features, like pillars, doors, and stairs. By doing so, it will most likely ease the mapping effort (IndoorAtlas 2018b). The distance between the waypoints is an important factor when trying to achieve the best possible results (Alnabhan et al. 2014). If the distance is too short, the signal recorded at the waypoints may be very similar, which will affect the accuracy of the positioning. If the distance is too long some areas may not be covered adequately (Alnabhan et al. 2014).

After the waypoints are placed out, the actual fingerprinting can be performed. To record a path, IndoorAtlas' mobile mapping application, *MapCreator 2*, can be used. MapCreator 2 is connected to the IndoorAtlas web application and provides information regarding the floor plans and waypoints. The process of fingerprinting starts by first calibrating the device, which consists of two steps (IndoorAtlas 2018c). The first step is calibrating the gyro sensors, and this is performed by placing the device on a flat, stable non-metal surface. The next step is performed by rotating the device around its different axes, which aids the calibration of the magnetic sensors (IndoorAtlas 2018c). The gyro sensors help the device sense angular velocity, i.e. the change in rotational angle per unit of time (Epson 2021). A magnetic sensor, on the other hand, is a sensor that detects the magnitude of magnetism and geomagnetism that is generated by a magnet, in this case, Earth's magnetic field, and is commonly used by e.g. a compass (AKM 2021).

When the calibration is performed adequately, the device must be positioned at the exact location of the starting waypoint of the planned path. Then the path is recorded by walking from waypoint to waypoint, and checking in when arriving at every new waypoint along the path (IndoorAtlas 2018e). While doing this, it is important to keep the top of the mobile device pointing in the direction of movement, as well as keeping the orientation of the device stable. After the last check-in, the path is stored and appended to the floor plan's existing map. The recorded paths should cover (nearly) all walkable areas to achieve a good positioning performance (IndoorAtlas 2018e).

Several conditions may affect the results of fingerprinting. The coverage of the Wi-Fi data might not be adequate, or the quality of the recorded paths might be inaccurate due to mapping errors. IndoorAtlas provides an Analytics tool in both their web tool and in the MapCreator application. Analytics visualize the coverage and quality of the performed fingerprinting by presenting overlays on top of the floor plan (IndoorAtlas 2020b). Each overlay is given a color, representing the current status of the data. Green means a good quality or coverage, while red indicates a bad quality or coverage. The Mapping Quality analytics describes how accurately nearby paths have been collected, i.e. to which degree the signals from the paths match. A mapping error may lead to inaccurate results. Such mapping errors can be inaccurate check-ins, inaccurate placement of waypoints, too few mapping paths in an area, or some other condition that leads to a bad quality (IndoorAtlas 2020b). To increase Mapping Quality, a path should be recorded several times and all paths should follow instructions, like ensure the check-in is performed when accurately positioned. By collecting several paths, the Wi-Fi Mapping Coverage is often improved as well. A green overlay of the Wi-Fi Mapping Coverage indicates that there is a sufficient amount of samples of the Wi-Fi data (IndoorAtlas 2020b). To be able to use this data as a positioning method, strong quality observations of Wi-Fi access points are essential.

After following the mentioned steps, and ensuring that all vital overlays in the Analytics tool indicate a sufficient result, the fingerprinting results can be used in an application to position a device.

3.1.3 Applications

The IndoorAtlas web tool allows for the creation of applications. In this context, an application is a container for API keys, which can be used to identify and authenticate with the IndoorAtlas SDK (IndoorAtlas 2020a). Each API key created within an application has three possible scopes: *SDK*, *Positioning API*, and *Data API*, and these scopes are used to constrict the permissions of the key. The Data API enables use in data API, the Positioning API enables use in positioning API, and the SDK enables use in mobile clients using the IndoorAtlas SDK (IndoorAtlas 2020a). The SDK scope also brings options for data storage, which regards the data acquired from a session created with the API key. The three options are *Do not store location data* (location data does not leave the device), *Store location data* *temporarily* (location data is available for up to four days) and *Store location data permanently* (location data is permanently available) (IndoorAtlas 2020a). If the data is stored, it is available through the Session Viewer in the web tool and the Data API.

Such an API key is then used to run the IndoorAtlas SDK on an Android device, by using the credentials (API key and secret) in the back-end of the mobile application to be implemented. The connection that is created, allows the mobile application to use the floor plans, maps, and positioning enabled by the IndoorAtlas web tool.

3.1.4 MazeMap

The floor maps used in the implementation, which were added to the IndoorAtlas web application, were provided by MazeMap. These floor plans cover the first and second floor of the building called Byggtekniske laboratorier at campus Gløshaugen of NTNU in Trondheim, Norway.

MazeMap also provides Data APIs, which can be used to fetch Point-of-Interest (POI) data. POIs are named points or rooms, which can be fetched as GeoJSON objects (MazeMap 2021). The POI data contains information about e.g the campus, building, floor, coordinates, etc., as well as a title (a room name or similar) and a description. In the implemented application, the POIs corresponding to a defibrillator was fetched using the Data API. Then the geographical information was used to place the POI in the AR application.

3.2 Implemented AR solution

By using the technologies mentioned, an AR navigation system was developed. The application was built locally on an Android phone, Google Pixel 1, and consisted of the start page with different buttons and the actual AR navigation screen. The start page is the user's initial meeting with the application and can be seen in figure 3.1.

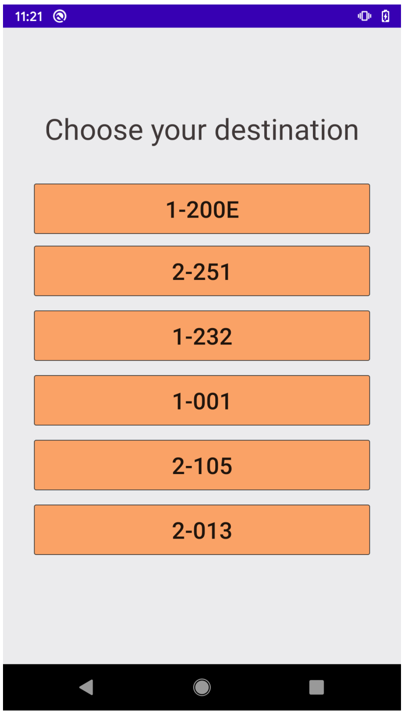
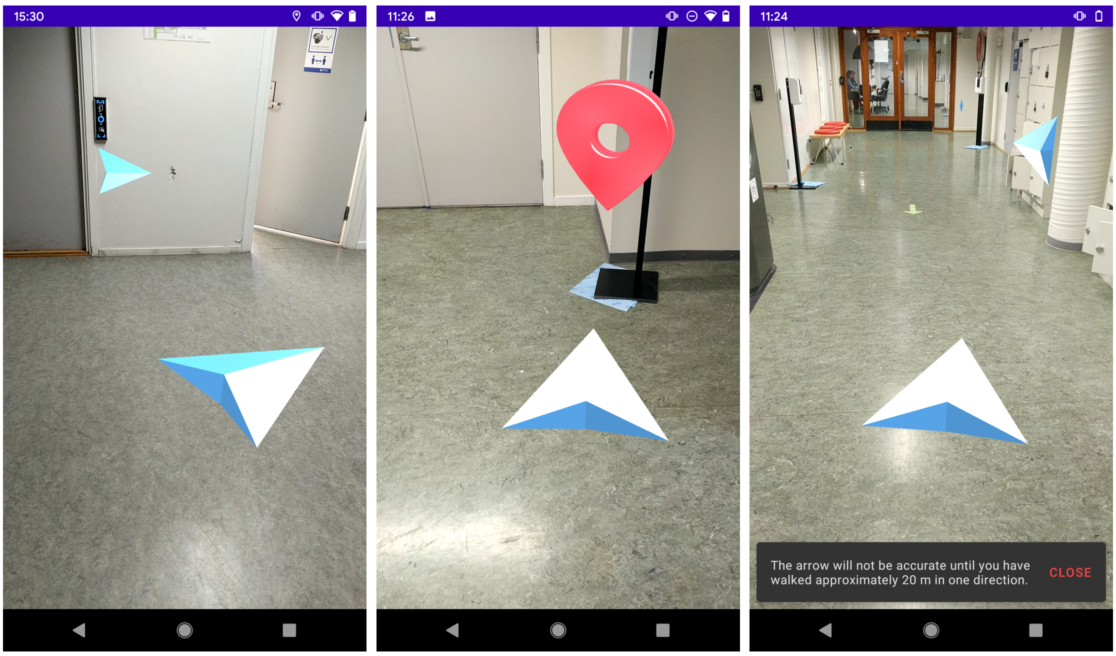


Figure 3.1: Start page of application

The start page includes text telling the user to choose a destination and buttons that represent different rooms at *Byggtekniske laboratorier* at NTNU in Trondheim, Norway. When clicking one of the buttons, a corresponding wayfinding request is sent to the AR session. An AR session is an IndoorAtlas interface that includes methods for handling AR objects, coordinate systems, and integration with external AR solutions, like ARCore (IndoorAtlas 2021b). The AR session then calculates and presents the user with a route leading to the desired destination. A wayfinding request consists of the floor number, longitude, and latitude of the destination.

At the building used in this implementation, the rooms are commonly known by room names rather than room numbers among students and employees at campus. This was taken advantage of, and resulted in the destination being presented as only the room numbers. The objective of this decision was to prevent that participants in the user test, that was familiar with the building, would navigate using personal experience rather than following the instruction given by the navigation systems.

The next screen the user is presented with is the AR navigation screen. This user interface uses the camera lens of the user's device and presents the camera capture. Different kinds of information are then presented on top of this live camera capture and assists the user in their navigation from start to finish. Figure 3.2 visualize some of the information the user receives. In figure 3.2a the two types of arrows in the application can be viewed. The arrow in the back, pointing to the right, is a "turn-by-turn" arrow. Turn-by-turn arrows are navigation arrows that appear and disappear depending on where along a calculated route the user is presently located. The arrows are rendered statically along the route and are typically placed in turns where the user has to make a change of direction, but also along long distances, like hallways. The second arrow visible in figure 3.2a is the compass arrow. The compass arrow also assists the user along the calculated route, but, in contrast to the turn-by-turn arrows, is constantly located right in front of the user (on the screen) and turns clearly with the camera. This arrow points approximately to the direction the user should walk, which is either towards the destination or the location of the next turn-by-turn arrow.



(a) Arrows (b) Destination-pin (c) Arrows with instructions

Figure 3.2: User interface of the application

Before starting the actual navigation process, the system has to accurately position and orient the device used, thus also the user itself. To position the device, the current signal measurements received by the device is constantly compared to the pre-collected data from the building. The orientation is then obtained by measuring approximate positions when moving forward in one direction. This also helps the system position the user more accurately. To ensure accurate estimations of position and orientation, the user was given instructions to walk approximately 20 meters in one direction before starting the actual navigation. The instruction was given using a Snackbar message, which is a temporary message located towards the bottom of the screen (Material 2021). Figure 3.3 illustrates the instruction, zoomed in, and it can also be viewed in figure 3.2c, which illustrates the instruction as it was actually presented to the user.

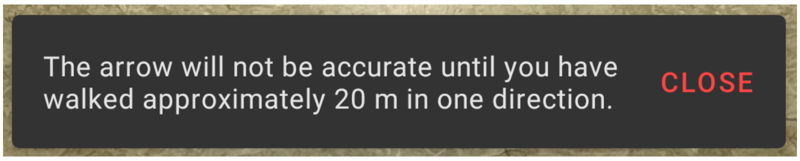
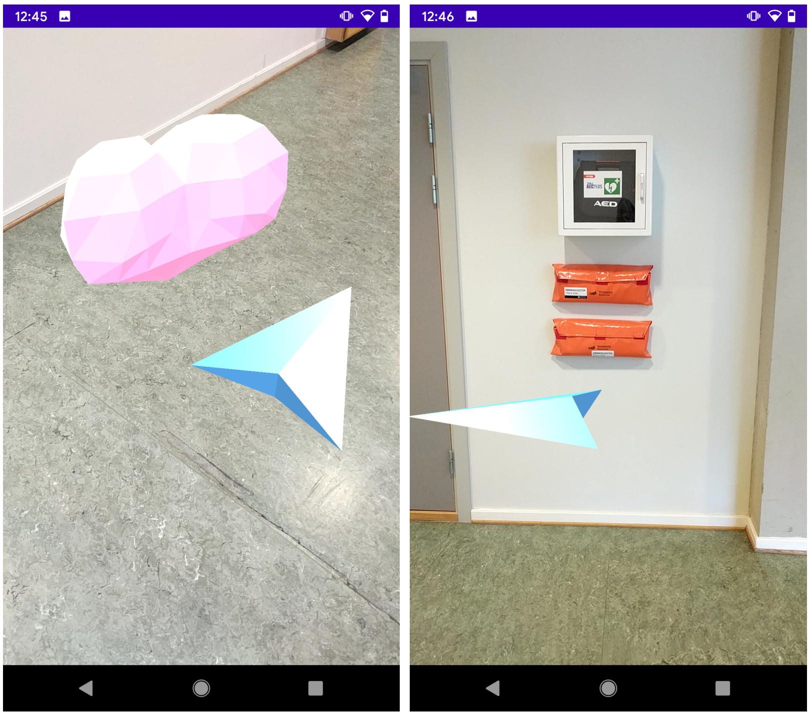


Figure 3.3: Instruction

As the user approaches the destination, a red destination pin appears, which is illustrated in figure 3.2b. ARCore provides a Depth API for Android, which created depth maps. Depth maps, also called depth images, allows occlusion within AR and make it easier to create immersive and realistic user experiences (Google 2021b). Virtual objects can then accurately appear behind or in front of real-world objects. However, the Depth API is only available through the use of specific devices, and the device used in this implementation was not included in the list of supported devices (Google 2021a). As the AR objects, like the destination pin, would show through walls and from great distances, other methods of creating a similar effect were implemented. The destination pin was, as a result of this, not rendered before the user was located at the same floor and within a certain radius of the destination.

The final AR object introduced in the AR application was a Point-of-Interest (POI). A POI is an object that consists of coordinates (latitude and longitude) and floor numbers. It has a stationary position, and was, like the destination, decided to be rendered only when the user was located within certain parameters. The application did end up being unresponsive for some milliseconds when the user was close enough for the POI to be rendered, but it only happened when the POI initially appeared.



(a) Point-of-Interest (POI) (b) Defibrillator

Figure 3.4: POI marking a defibrillator

In other words, the unresponsiveness was not present throughout the rest of the navigation process. In the implemented application, only one kind of POI was included. By using a MazeMap API to collect geographical data about defibrillators located at NTNU, a 3D model in the shape of a heart was placed out to mark the position of each defibrillator. The result can be seen in figure 3.4. Figure 3.4a visualize the pink heart placed along the route, and then the corresponding defibrillator can be seen in figure 3.4b.

3.3 Testing

An experiment was conducted to test the implemented AR solution and to check whether this approach to indoor navigation has any potential, according to a selection of potential users. To collect data from the experiment in a clear and orderly manner, and to ensure near-equal conditions for each participant, a test plan was developed. The test plan consisted of predefined pre-experimental questions, concrete information to be given to the participants, instructions to be given to the participants in the practical part, tasks of the instructor, and questions used in the post-experimental interview. It also contained information about where the location of the experiment was, who the participants were, what kind of equipment were to be used, and what the objective was.

3.3.1 Usability testing

Usability metrics reveal something about the interaction between the product to be tested and the user (Albert et al. 2013). According to the ISO standard *Ergonomics of human-system interaction* (*ISO 9241-11:2018* 2018), usability is "the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use". A usability metric should include effectiveness, efficiency, and satisfaction in some aspects. *Effectiveness* is a measure of the completeness, and accuracy, with which a user performs specified tasks. *Efficiency* is a measure of how much effort is required to perform specified tasks, and *satisfaction* is a measure of how pleased the user is with the experience while performing specified tasks (Albert et al. 2013).

When conducting a usability test, decisions regarding what *kind* of experiment is favorable have to be made. There are different types of usability tests, and some of them are moderated vs. unmoderated, remote vs. in-person, and laboratory vs field (hotjar 2020). A moderated test is supervised either in person or remotely by a researcher. The researcher instructs the participants, answer potential questions, and may ask follow-up questions. On the contrary, an unmoderated test is conducted without direct supervision. The purpose of the experiment was to observe the participants, but not interfere and affect the participant's perception and experience as little as possible. As the purpose also was to get in-depth results and the reasoning behind the participant's actions, a moderated test was favored.

Remote testing is normally performed over the internet or the phone, while in-person testing requires physical presence (hotjar 2020). In-person testing often comes with a larger cost, both in terms of resources and time. As the application to be tested requires presence at the test location, and the participants' reasoning is highly valued, the test was decided to be in-person.

A laboratory experiment is an experiment that is conducted within a highly controlled environment (McLeod 2012). Nearly all conditions are predetermined and planned, and it is easy to replicate. A field experiment, on the other hand, takes place in a more realistic environment, i.e. real life, where all conditions, like possible distractions, cannot be planned and controlled. This type of experiment is more likely to give results that will reflect real life, but is harder to replicate, as variables might change from participant to participant (McLeod 2012). When testing the implemented application, a field experiment was chosen, as the results were desired to reflect real-life usage.

3.3.2 Location

The location of the experiment was at the campus of Norwegian University of Science and Technology (NTNU) in Norway. NTNU is located in three Norwegian cities; Trondheim, Ålesund and Gjøvik. The headquarters is situated in Trondheim, which also is the location chosen for this experiment. The campus of NTNU was chosen as MazeMap already had indoor floor plans of the buildings. The specific building chosen was *Byggtekniske laboratorier*, which is located at the south-end of Gløhaugen Campus. Gløshaugen Campus is one of the seven campuses in NTNU Trondheim. Byggtekniske laboratorier has a total of four floors, but only two of them (first and second floor) were used in the implementation. As the building is a part of campus, it forms a realistic environment, with different kinds of distractions. It was chosen due to the fact that the division of Geomatics at NTNU is stationed there.

3.3.3 Equipment

The equipment used in this experiment was a Google Pixel 1, which is an Android smartphone that is designed, developed, and marketed by Google. The smartphone is one of the devices included in the list of supported devices for indoor mapping (IndoorAtlas 2021a), and was assumed to provide good quality results. It was used by all participants to ensure that all had the same prerequisites, e.g. accuracy, when conducting the tasks. The application was developed to fit this device's qualifications.

To record the data acquired from the experiment, two additional phones, and a laptop were used. The phones were used during the practical part to record the time used and write down all actions and comments provided from the participants, and the laptop was used to take notes during the post-experiment interview. In addition to this, the already existing Wi-Fi infrastructure at the location was used.

3.3.4 Objective

The objective of the experiment was to test the implemented AR solution and get potential users' opinions. To do this, the AR solution was put up against MazeMap's 2D digital map solution. The participants were to test both of the solutions, give their opinion on both of them individually, and compare them.

To have enough input to give their opinions, the participants were asked to follow a route with each of the two solutions. Before testing both solutions, the participants were asked to follow a test route with the AR solution. This route was common to all participants, and the purpose of it was to let the participants get more familiar with the AR solution, as none of the participants had tried it out previously. The test route also included a POI placed out along the path. The purpose of the POI was to show the participants what a possible POI could look like. Then the participants were divided into two groups, and each group consisted of 9 participants. One group was given navigational instructions to the first route presented in the form of AR, and the other group was presented these instructions by a route drawn on top of the 2D map. With the second route, the groups switched the form of presentation, i.e. the ones with AR on the first route switched to a 2D map on the second route and vice versa. All participants tested both the AR and the 2D digital map solution.

3.3.5 Pilot test

A pilot test is, as defined by Jhangiani et al. (2015), "a small-scale study conducted to make sure that a new procedure works as planned". To be confident that the procedure of the experiment would work the best way possible, such a pilot test was conducted. An informal recruitment of a classmate was made to find any potential weaknesses and improvements to the experiment. The pilot test was conducted in the same way that a real test would have been, and the pilot participant had the chance to give feedback regarding the tasks, the instructors' role, and the whole execution in general. It was made clear to the participant that all kinds of feedback and criticism were highly appreciated, as the purpose was to exclude errors and remove misunderstandings. The pilot test also gave a good indication of how much time it was necessary to devote to each participant and whether the data was recorded correctly. After this pilot test was completed, the procedure of the experiment was updated and improved to deal with the issues brought up in the feedback.

Examples of improvements that were made after the pilot test were the distribution of responsibilities between the instructors, the layout of the document used for taking notes during the test, and the content of the instructions given to the participant.

3.3.6 Participants

As this experiment was conducted during the COVID-19 pandemic, the access to a diverse selection of participants was limited. Due to the pandemic, no other than students or employees were allowed to stay on campus. Also, it was not favorable to gather students not normally belonging to the chosen building, in addition to the fact that the interaction between individuals was to be limited. This made it rather difficult to involve a variety of participants in terms of age and fields of study, which in normal circumstances would have been preferred.

A total of 18 participants participated in the experiment, where 11 of the participants identified themselves as male and the remaining as female. Three were in the age group 18-21, 13 in the age group 22-25, and two in the group 26 or older. The ages of the participants were, as mentioned, not that different as they all were students. The fields of studies of the participants were divided into geomatics, other technology-related, and non-technology-related. As the experiment was conducted at a building where the division of geomatics was located, and it was not advised to interact with many new people, a majority of the participants were geomatics students. More specifically 10 out of the 18 participants studied geomatics, but their progression varied from being 3rd-year students to 5th-year students. The rest of the group of participants consisted of six studying a technology-related subject and the remaining two studying a non-technology-related subject. A summary of the distribution of the participants can be found in table 3.1.

Table 3.1: Distribution of the participants

Participants	18
Male	11
Female	7
18-21	3
22-25	13
26+	2
Geomatics	10
Technology	6
Non-technology	2

All of the participants were already acquainted with the authors before conducting the experiment.

3.3.7 Data acquisition

Before conducting the more practical part of the experiment, the participants were to answer a number of pre-experimental questions. These questions were used to collect statistical information about the participants. The pre-experimental questions are listed, and explained more thoroughly, in appendix A.

During the practical part of the experiment, the participants were encouraged to think aloud as much as possible. They were asked to justify all their choices and speak about what information made them take which kinds of actions. All of these remarks were written down. The time the participants used to navigate from A to B was also recorded, and it was noted whether the participants took any wrong turns or if many stops were necessary. The time used was used as a metric measuring the efficiency of the application.

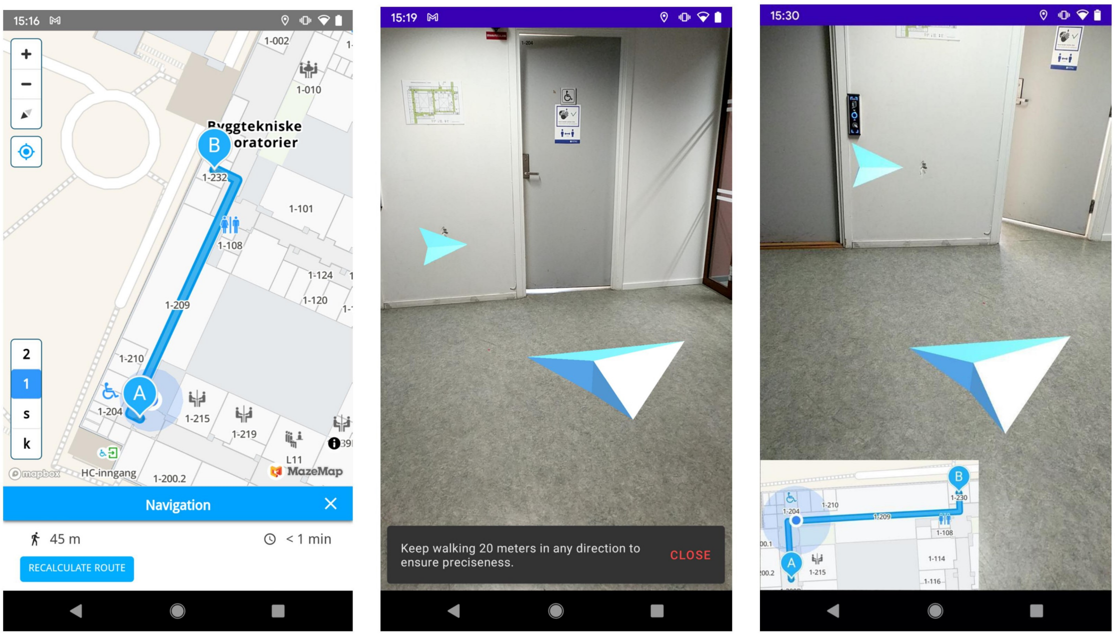


Figure 3.5: Visualizations of 2D map, AR and a possible combination of the two

After the practical part, an interview of the participants was carried out in a more controlled environment. The purpose of this interview was to get a more detailed insight into the participant's opinions and ideas. The participants were asked to give advantages and disadvantages with both the 2D map and the AR solutions, and a more detailed explanation about their impression and thoughts on the AR solution. Figure 3.5 visualizes the user interface of MazeMap's 2D map solution, the implemented AR solution and a possible combination of the two. This visualization was printed out and showed to the participants during the interview to help them get a more clear idea of what the different possibilities were or could be. An enlarged version of figure 3.5 and the specific questions used during the interview can be found in appendix B.

3.3.8 Infection control measures

Precautions to prevent a possible spread of infection were important to both establish and maintain throughout the user test. To conduct the user test in a safe matter, the participants were encouraged to cancel their appointment if they felt any kind of illness or if they had been in contact with someone possibly exposed to the disease.

As a part of the experiment, the participants filled out a digital form with the pre-experimental questions. To avoid as many common contact surfaces as possible, the participants used their own personal device to scan a QR-code and access the form. By filling out the form, personal information like the participants' full name was collected. This was important to be able to contact and inform everyone if an infection were to happen.

In the practical part of the experiment, the participants were using the same device. As this was a common contact surface for all of the participants, this was cleaned thoroughly with antibacterial wipes before and after each use. In addition to this, the participants were encouraged to use both antibacterial wipes and liquid hand sanitizer before, during, and after the experiment.

Chapter 4

Results

In this chapter the results of the user test is presented.

4.1 User test

To test and evaluate the implemented application, a user test was conducted. The following section includes the results from the user test.

4.1.1 Pre-experimental questions

The questions used to gather statistical data of the participants at the beginning of the user test, are explained more thoroughly in appendix A. The participants answered a total of 18 questions, and all 18 participants answered all questions.

The participants were asked to rate their previous experience with different technologies. Their answers varied from none (never used) to much (use regularly). Table 4.1 illustrates their answers. The questions started quite general and gradually got more specific.

Table 4.1: Participants previous experience with different technologies

Previous experience with ...	**Much**	**Some**	**None**
Mobile applications	18	-	-
Navigation systems	14	4	-
Pedestrian navigation systems	12	6	-
Indoor navigation systems	2	16	-
2D digital maps	16	2	-
2D map indoor navigation systems	14	4	-
AR applications	1	12	5
AR navigation systems	-	4	14

All of the participants stated that they were familiar with MazeMap's application. This was relevant as this application was used as the opponent to the implemented AR application in the user test.

A number of 15 of the 18 participants had been to the location of the test before, while three of them had never been. The participants who had been to the building before were asked to rate their familiarity on a scale from one to five, where one meant almost no knowledge of the building and five meant very familiar. The distribution of their answers is visualized in figure 4.1.

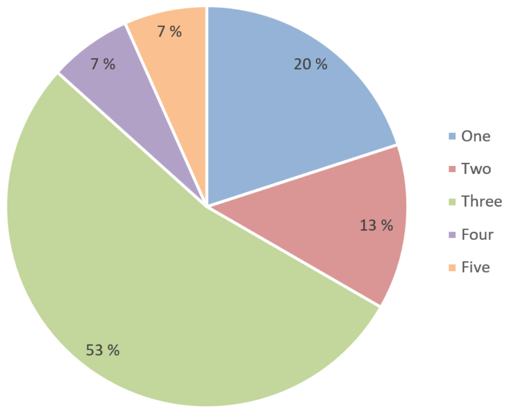


Figure 4.1: Participants familiarity with the building

4.1.2 Practical part of the experiment

While conducting the practical part of the experiment, data about the participants' performance and experience were acquired. One of the metrics used to compare the implemented AR solution with the 2D digital map navigation system was the time used to reach the destination. As the participants navigated through two different routes, with different lengths, the time was converted to average speed instead. This conversion was performed as a measure to make the collected data more suitable for comparison. Figure 4.2 visualize the average speed of the participants using AR and 2D for both routes, as well as the overall average speed regardless of route.

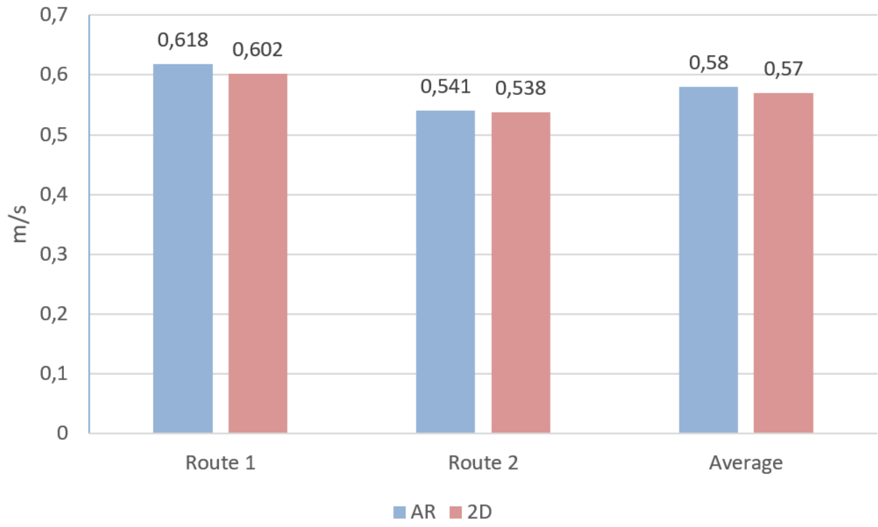


Figure 4.2: Average speed by route

The information gathered through the pre-experimental survey was then used to divide the participants based on different characteristics. The participants were divided into those who were familiar and those that stated that they were unfamiliar with the building used as the location of the user test. The group of unfamiliar participants consisted of those that had never been to the building before the experiment and those who rated their knowledge as a one, which was the case for a third of the participants (i.e. 6 out of 18). Figure 4.3 visualize the average speed of the participants using AR and 2D divided into groups based on their familiarity with the building. The participants were also divided based on their stated gender, and figure 4.4 visualizes these results. Seven of the participants classified themselves as female and the remaining 11 as male.

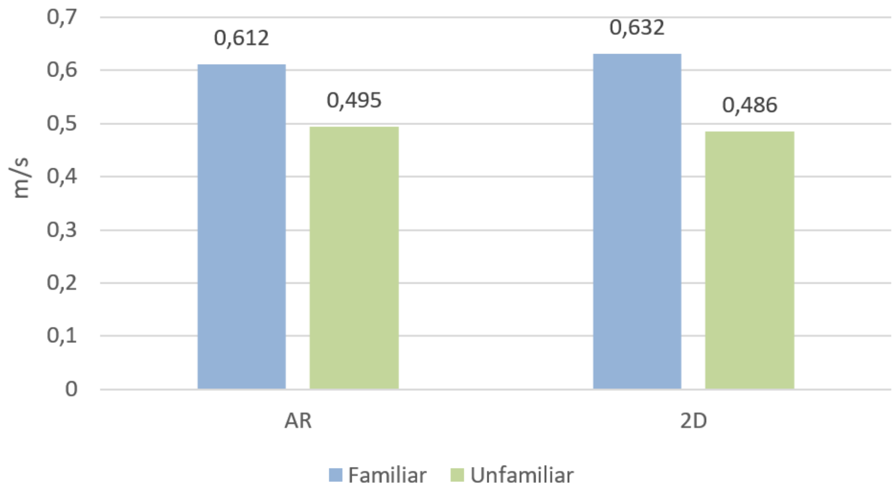


Figure 4.3: Average speed by familiarity

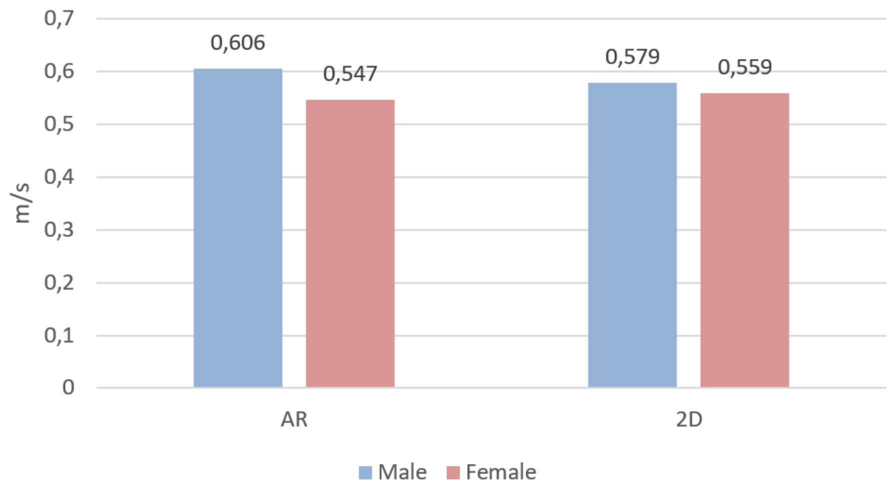


Figure 4.4: Average speed by gender

To further investigate the data acquired in the user test, a test of significance was conducted. A significance test is a way to estimate whether there is a real difference between data sets, or if the difference is purely coincidental (Kitchin et al.

2000). According to Kitchin et al. (2000), such a test usually consists of four generic components, and these components are null/research hypothesis, calculated/critical values of statistic, one/two-tailed tests, and significance level. The null hypothesis (H_0) usually states that there is no significant difference between the data sets, while the research hypothesis (H_1) states that there is a significant difference. As it is interesting to see whether there is either a negative or a positive difference, a two-tailed test is, in this research, preferred. A one-tailed test only investigates differences in one direction (i.e. only negative or only positive) (Kitchin et al. 2000). A t-test is a statistical hypothesis test, which requires a normal distribution. To meet the requirements of normal distribution, the data sets were transformed using the function $f(x) = \sqrt{x}$.

First the differences between the two solutions were investigated. As each participant navigated through a route in both AR and 2D digital map, and the data sets can be considered related, the related t-test was preferred to an unrelated one. A related t-test are defined by Kitchin et al. (2000) as a test that "compares the actual size of differences between the matched/related scores in each condition or time".

$$x_{\bar{AR}} = 0.582 \quad x_{\bar{2D}} = 0.570$$

$$\sum d = 0.235 \quad \sum d^2 = 1.0195 \quad (\sum d)^2 = 1.0393$$

x_{AR} is the average speed used when navigating with the AR solution, while x_{2D} is the average speed used when navigating with the 2D digital map solution. $\sum d$ is the sum of the differences, $\sum d^2$ is the sum of the differences squared, and $(\sum d)^2$ is the square of the sum of differences.

$$t = \frac{\sum_{i=1}^{n} d_i}{\sqrt{\frac{n\sum_{i=1}^{n} d_i^2 - (\sum_{i=1}^{n} d_i)^2}{n-1}}} \tag{4.1}$$

where d_i is the differences between the matched data sets, and n is the number of observations.

The value of t is calculated by using equation 4.1, which results in $t = 0.902$. The degrees of freedom DF is found by taking the number of observation and subtract 1 ($n - 1 = 16 - 1 = 15$). The critical value for a related two-tailed t-test at the 0.05

significance level and with a DF of 15, is 2.131, which is greater than the calculated value of 0.902. This means that the null hypothesis H_0 was not rejected, thus there is no significant difference between the average speed when using the AR and the 2D digital map solution.

A similar calculation can be performed on the data sets regarding differences between the different familiarity levels and genders. As AR is the focus of this research, only the average speed related to the AR solution was looked further into when looking at the possible differences within these groups. The data sets regarding familiar and unfamiliar participants are not related. For this reason, an unrelated t-test was applied, which according to Kitchin et al. (2000) "compares two unrelated data sets by expecting the amount of difference between their means and taking into account the variability of each data set". x_F is the average speed of a participant classified as familiar, and x_U is the average speed of a participant classified as unfamiliar.

$$\bar{x_F} = 0.790 \quad n_F = 10 \quad \bar{x_U} = 0.706 \quad n_U = 6$$

$$\sum x_F = 7.903 \quad \sum x_F^2 = 6.310 \quad (\sum x_F)^2 = 62.4512$$

$$\sum x_U = 4.235 \quad \sum x_U^2 = 3.045 \quad (\sum x_U)^2 = 17.9344$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}}{n-1} \tag{4.2}$$

where x_i is the value of an observation and n is the number of observations.

In an unrelated t-test, the variances of each data set is required. Equation 4.2 is used to acquire these values, which gives $s_F^2 = 0.0072$ and $s_U^2 = 0.0112$. The variances are then used to find the value of F, by dividing the greater variance with the lesser variance. As s_U^2 is greater than s_F^2, F is obtained by $F = s_U^2/s_F^2 = 1.5516$. The critical value of F with $DF_F = 9$ and $DF_U = 5$ is 4.77 at the 0.05 level. This means that $F < F_{critical}$, and a pooled version of the t-test can be used. The degrees of freedom is determined by $DF = n_F + n_U - 2 = 14$.

$$t = \frac{|\bar{x_a} - \bar{x_b}|}{\sqrt{(\frac{s_a^2 \times (n_a - 1) + s_b \times (n_b - 1)))}{n_a + n_b - 2}) \times (\frac{1}{n_a} + \frac{1}{n_b}))}} \tag{4.3}$$

The t-value is calculated with equation 4.3, and is determined to be $t = 7.0412$. The critical value of t at significance level 0.05 with 14 DF is 2.145, which is less than the calculated value of t. This means that the null hypothesis H_0 is rejected, and there is a significant difference. The critical value of t at significance level 0.01 with 14 DF is 2.977, which also is less than the calculated value of t. This means there is a *highly significant* difference between the average speed of the ones who are familiar with the location and the ones who are unfamiliar.

$$\bar{x_M} = 0.607 \quad n_M = 9 \quad \bar{x_F} = 0.558 \quad n_F = 7$$

$$\sum x_M = 5.462 \quad \sum x_M^2 = 3.401 \quad (\sum x_M)^2 = 29.8334$$

$$\sum x_F = 3.903 \quad \sum x_F^2 = 2.405 \quad (\sum x_F)^2 = 15.2334$$

$$s_M^2 = 0.0108 \quad s_F^2 = 0.0381$$

Similar to the data sets regarding familiarity, an unrelated t-test was used and calculated for the data sets regarding genders. x_M is the average speed of a male participant, and x_F is the average speed of an female participant. The calculated F is 3.526. The critical value of F with $DF_F = 8$ and $DF_U = 6$ is 4.15 at the 0.05 level. This means that $F < F_{critical}$, and a pooled version of the t-test can be used again. The degrees of freedom are determined to be $DF = 14$, and the t-value is determined to be $t = 2.611$. The critical value of t at significance level 0.05 with 14 DF is 2.145, which is less than the calculated value of t. This means that the null hypothesis H_0 is rejected, and there is a significant difference between the average speed of male and female participants. However, the critical value of t at significance level 0.01 with 14 DF is 2.977, which is greater than the calculated value of t. This means there is a significant difference, but not a highly significant difference.

While the participants navigated from start to finish with both AR and 2D, all their action and remarks were written down. Some of the most repetitive remarks made by the participants regarding both the solutions are summarized in section 4.1.3, and discussed further in chapter 5.

4.1.3 Post-experiment interview

During the post-experiment interview, the participants were asked questions to acquire more detailed opinions and reasoning behind their behavior. First, they were asked whether they preferred the AR or 2D map solution, and encouraged to give a justification for their answer. Four out of the 18 participants stated that they preferred the 2D map solution to the AR solution, while the remaining 14 stated that they preferred the AR solution. Figure 4.5a visualize the distribution. The participants were also encouraged to give advantages and disadvantages to both the 2D map solution and the AR solution. The most repetitive answers regarding the 2D map solution are included in table 4.2, while the ones regarding the AR solution are included in table 4.3.

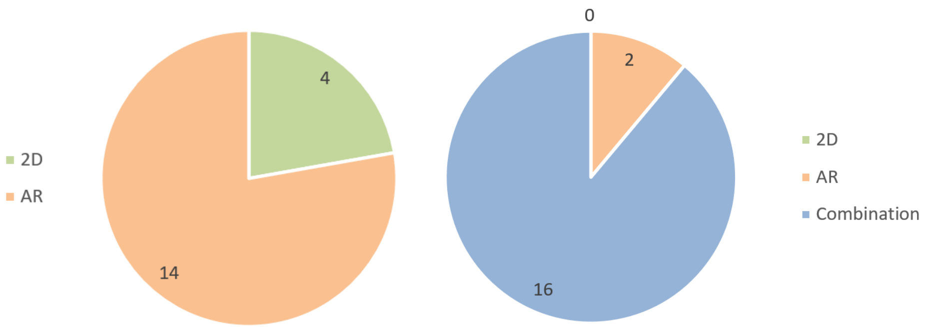
Table 4.2: Advantages and disadvantages of the 2D map solution

Advantages	**Disadvantages**
Gives context to the proposed route	No orientation with the blue dot
The user gets more control	No automatic change of floor plan
Provides view of the whole route	The map does not rotate with the user
Possible to plan ahead	Requires some understanding of maps

Table 4.3: Advantages and disadvantages of the AR solution

Advantages	**Disadvantages**
Easy to understand	Does not provide any context
Requires little cognitive load	Have to trust the arrows completely
Provides enough information to get to the destination	Not possible to plan ahead
Adapts if the user deviates from the route	Inaccuracy easily result in errors
	Uncomfortable to angle the phone upwards

The participants were also introduced to a possible combination of the 2D map and AR solution, which is visualized in figure 3.5. When asked whether they would prefer the 2D map solution, the AR solution, or a possible combination of the two, 16 of the 18 participants stated that they would prefer the combination. Two stated that they would still prefer the AR solution, while none preferred the 2D solution. Figure



(a) 2D map and AR solution (b) 2D map, AR and combination

Figure 4.5: Preferences between the different solutions

4.5b visualize the distribution. They were also encouraged to give some advantages and disadvantages to possible combinations, if they could think of any. The most repetitive answers are included in table 4.4.

Table 4.4: Advantages and disadvantages of a combined solution

Advantages	**Disadvantages**
Does not have to rely completely on one of the solutions	Harder to keep track of two systems at once
Possibility to switch between the systems, depending on needs	Little space on a mobile device screen

The opinions regarding the implemented AR solution were looked further into. More specifically the participants were asked to elaborate on whether there were any important defects, and how they would like to improve the application. Some of the most repetitive answers are included in the list below.

- More information regarding the *destination* on the screen.
  Room number/name or similar. Floor number.
- More information regarding *change of floor level* (or building).
  Current floor number. Number of floors up or down to destination.
  Warning when approaching a staircase.
- More information regarding *wrong turns* or direction.
  Notification in case of any deviations from the route.
  Vibration, sound, text message that alerts the user.

Chapter 5

Discussion

In this chapter, various AR solutions will be evaluated and compared to the implemented solution. In addition to this, the results from chapter 4 will be discussed. More specifically, the participants, data gathered from the practical part of the experiment, and feedback from the participants will be looked further into.

5.1 Evaluation of AR solutions

Even though AR is implemented in various applications throughout various markets, it is still not that common within the field of indoor navigation. Even so, there are multiple solutions available for indoor positioning for mobile applications. As mentioned in section 3.1, the tools and platforms used when developing the AR-based navigation system presented in this paper are Google's ARCore, IndoorAtlas and data provided by MazeMap.

IndoorAtlas is a platform that provides indoor positioning for mobile applications and determines the global position and orientation of a device. As beacons were required in the iOS implementation with IndoorAtlas and the location used in this experiment had no beacons installed, the obvious choice fell on an Android implementation. Therefore, the platform used was ARCore as it is compatible with Android devices. The IndoorAtlas AR API merges ARCore's ability to track the device in a local metric coordinate system, and IndoorAtlas' ability to determine the global position and orientation of the device (IndoorAtlas 2020c). The floor maps used in the implementation were provided by MazeMap. The technologies and tools used in this solution have some weaknesses.

The IndoorAtlas SDK provides accurate positioning. By fingerprinting the two floors used in the experiment, the system had an almost unique signal for each position throughout the floors and was able to position the user accurately by matching the signals acquired from the device with the signals in the database. However, there was a need for the user to walk in one direction for about 20 meters in order for the positioning service to accurately orient the user. This was, according to the participants, more visible in the AR solution as the compass arrow would point towards the wall as a result of inaccurate orientation. The 2D map used in the experiment did only display the positioning and not orientation of the user. When using the 2D digital map, the participants were able to orient themselves, but requires a greater cognitive load.

Scalability is an important factor when it comes to indoor navigation systems, as they often are implemented in larger areas such as universities and hospitals. Scalability describes how simple it is to expand the provided solution. The method presented in section 2.4.5 by Makarov (2020) creates floor maps by drawing these in detail, adding visual markers and building the possible route using graph theory. Drawing detailed maps and building the route as developers, may decrease scalability as it is time-consuming. However, by using already-existing floor plans, such as those provided by MazeMaps and using platforms that provide wayfinding, such as IndoorAtlas, the process of creating such an application gets more effective.

Nevertheless, the process of fingerprinting has some disadvantages that decreases the scalability. The algorithms supporting the Wi-Fi fingerprint positioning are heavy on the computation and complexity. In addition, the preliminary work has a high cost factor as someone has to physically walk around and map the building. Also, since it is dependent on Wi-Fi signals, which are easily affected by environmental changes, re-acquisition and re-modeling of the data is required when any changes takes place (Xia et al. 2017). Another method that could be used in place of Wi-Fi fingerprinting are Bluetooth beacons. Implementing beacons have low deployment cost, as the beacons use their inner batteries after being installed and will last about 6 months (Zuo et al. 2018). The range-based method in beacon-based positioning avoids fingerprinting and estimates the distance between the users and the beacons by using trilateration (Zuo et al. 2018). By avoiding fingerprinting, this method also avoids the high preliminary workload.

Google provides ARCore free of charge (Leon 2020). However, both IndoorAtlas and MazeMap provides plans with cost that varies with wanted features and terms of use. The positioning technology in use is Wi-Fi fingerprinting which is regarded as a low-cost positioning method due to the existing infrastructure of Wi-Fi network in buildings and Wi-Fi chip in most modern smart devices (Alnabhan et al. 2014).

All of the tools above, including ARCore and IndoorAtlas, had extensive documentation available which made the implementation of the AR-based indoor navigation application easier. Step-by-step tutorials on how to implement the tools and SDKs is highly valuable to a developer.

A possible alternative is to use AR Foundation when developing which allows iOS devices to also indulge in the AR application created. However, IndoorAtlas does, to date, only provide fingerprinting with iOS if the venue has beacons, which is less accessible than Wi-Fi network (IndoorAtlas 2019b).

5.2 Participants in the user test

Hinderer et al. (1998) said that ”effective participant recruiting is crucial to collecting reliable data during usability testing of high-tech products and services”. Participants in a user test should reflect the characteristics of the intended target group of the application (Hinderer et al. 1998). As previously mentioned, the participants were more similar than what would, under normal circumstances, have been preferred. The COVID-19 pandemic made it challenging to recruit a selection of participants that represented a variety of different people. The participants were all students at NTNU, and they were all between 18 to 28 years old, which means that they represent only a small selection of our society. As the test was conducted at the campus of NTNU, the most natural target users are in fact students. However, it would have been natural to include a more diverse group in terms of age and fields of study, and maybe some employees. Younger people usually have more experience with newer technology, especially those that study technology-related subjects, which was the case for 16 of the 18 participants. It is not unlikely to think that other results could have been acquired if a more diverse group of participants, in terms of age, were recruited. An example of this is the research of C. Wang et al. (2019), which concluded that there was a significant age-related difference in indoor wayfinding. A majority of the participants had been to the location before participating in the user test, and in other circumstances, it would have been sensible to include more unfamiliar participants, to get a more diverse selection.

As mentioned in section 3.3, national and local restrictions related to the pandemic did also lead to recommendations regarding limitations of the number of social contacts. This led to participant recruitment more or less within the social circles of the authors. NTNU also had its own restrictions regarding who had permission to stay at the campus, which included only students or employees.

The consequences of a homogeneous group of participants are not receiving important feedback from all kinds of potential users, and possibly getting homogeneous results. When reviewing the results of the user test, it was visible to see the differences between the majority of the group and the participants that were both unfamiliar with the location and not majoring in technology. As not enough unfamiliar participants majoring in non-technology-related subjects were partaking, it is difficult to draw some real conclusions related to these findings. It could have been a coincidence. However, it is clear to see that the user test would have even more reliable results with a more diverse group of participants.

5.3 Comparison of results

In chapter 4, the results regarding performance and preferences were provided by the participants through actions and statements. This section contains further discussion of the results regarding various comparisons.

5.3.1 2D map and AR solutions

During the practical part of the user test, the time each participant used to reach the destination of the route was measured and stored. The time and length of the route were then used to calculate the average speed. As seen in figure 4.2, the average speed when using the AR solution was slightly higher than when using the digital 2D map solution. However, by using a related t-test, it was calculated that this difference was not significant, which means that it could be purely coincidental that the average speed when using the AR solution was higher.

When asked whether they preferred the AR or the digital 2D map solution, a majority of the participants stated that they preferred the AR solution, as visualized in figure 4.5a. This means that even if the average speed when using the two solutions were quite similar, a majority still preferred one of the solutions over the other. The AR solution had the benefit of being a relatively new and undiscovered technology for a major amount of the participants. As seen in table 4.1, previous experience with AR applications was most commonly stated to have been minimal, while AR navigation systems were in most cases never even tried out. 2D maps, on the other hand, were more commonly used among the participants. Newer technology could be perceived as more exciting, and several of the participants drew associations to video games when using the AR solution. As the digital 2D map solution was, to various degree, familiar to the participants, they may already have made up some opinions before the actual experiment. The 2D map solution's familiarity could also have contributed beneficially, as the participants knew how to use it.

The participants were also aware that the authors had implemented the AR solution, and may have been affected by *subject bias*. Subject bias, also called participant bias, is a phenomenon that may happen in experiments where the participants either know or think they know what the expected outcome is, and then try to achieve this outcome (Duignan 2016). This can happen either as a conscious act or simply unconsciously. In this experiment, the participants may have been affected by this phenomenon and answered that they preferred the AR solution as they believed this was the desired answer.

5.3.2 Level of familiarity

The participants were divided into groups depending on their level of familiarity (familiar or unfamiliar) with the location of the user test. They were divided based on their own classifications from the pre-experimental questions, which means their level depended on subjective interpretation and evaluation. Two different users with the same amount of knowledge of the building may rate their familiarity completely different. Taking this into consideration, the results of this division were as presented in figure 4.3. As seen in the figure, the ones that classified themselves as rather familiar with the building had an average speed that was higher than the ones of those classified as unfamiliar. By performing an unrelated t-test, it was determined that the difference between the average speed of the two groups was highly significant. This means that it was not coincidental. H. Huang, Schmidt et al. (2012) researched the spatial knowledge acquisition of a group of participants by the use of different user interfaces. They found that there were significant differences between familiar and unfamiliar participants, and more specifically the group of familiar participants solved the tasks regarding spatial knowledge acquisition with significantly less errors than the unfamiliar ones (H. Huang, Schmidt et al. 2012).

5.3.3 Gender

The participants were also divided into groups based on their stated gender. As all participants stated that they were either male or female, this resulted in two groups. The result of this division can be seen in figure 4.4. By performing an unrelated t-test, it was determined that there was a significant, but not a highly significant, difference in the average speed of male and female participants. Males were moving with a higher average speed with the AR solution than the females.

Zhou et al. (2020) researched the relation between genders and time pressure when navigating within indoor spaces. They found that males took more risks than females under time pressure and that males also searched more quickly for information. The participants were not under time pressure in the user test presented in this thesis, but some still stated that they felt some kind of pressure due to the measurements of the time used to reach the destination and the presence of the instructors. They were ensured that there was no pressure of any kind and that only natural movements and responses were expected of them, but it is realistic to assume that some of the participants still felt the pressure to reach the destination with the least amount of error and the least amount of time possible. In a situation where the navigation system might have given slightly inaccurate instructions, the males might have taken more risks than the females, which could result in a faster average speed. However, these are only speculations.

C. Wang et al. (2019) researched the differences related to gender and age differences in indoor maps and wayfinding. They found that males and females had no significant differences in route reading and following, and stated that "there was no significant gender difference in indoor wayfinding performance" (C. Wang et al. 2019).

5.3.4 Combination

The participants were also asked whether they would prefer the AR solution, the 2D map solution, or a possible combination of the two. As seen in figure 4.5b, a majority of the participants answered that they would prefer the combination over the other two alternatives. Only two out of the 18 participants stated that they would prefer the AR solution, while none preferred the 2D map. During the post-experiment questions, the participants were encouraged to give advantages and disadvantages to both the AR solution and the 2D map solution. As these questions and answers were given before the questions regarding the possible combination, the combination may have been perceived as an obvious answer, as it would have the advantages from both solutions. However, some participants suggested a possible combination even before it was brought up by the authors, and others picked the combination but showed some ambivalence about their answer as they were uncertain if such a solution would work on a small screen like on a mobile phone.

5.4 Implemented AR solution

Various feedback regarding both the implemented AR solution and the 2D digital map solution was provided by the participants in the user test, and the ones regarding the implemented AR solutions are specifically looked closer into, as this research's main focus is the possibilities using AR for indoor navigation. This section contains the findings, regarding the AR solution, of most interest and value.

5.4.1 Initial calibration

As mentioned previously, the arrows in the AR solution did in some situations point slightly towards walls or doors that did not lead anywhere. This could lead to confusion and misunderstandings. In the AR solution, the arrows are the only source of information regarding the path provided to the users. This means that if the arrows are slightly inaccurate, the user may have problems navigating. This inaccuracy was especially a problem at the beginning of a new route. When starting a new navigational process, the AR application needs to orient and position the user correctly. To be able to do this, the user is required to move approximately 20 meters in one direction with its device. As an attempt to communicate the requirements to the participants, a Snackbar message was placed at the bottom of the user interface, which is visualized in figure 3.2c and figure 3.3. It was observed that a majority of the participants in the user test either forgot about this message rather quickly after seeing it or simply ignored it. In the post-experiment interview, several of the participants stated that they did not notice the message or that they thought it was an error message due to its placement and design. Others expressed that they did see it, but when the arrows appeared they thought the calibration was sufficient. As even small inaccuracies may lead the user in the wrong direction, such a message must be communicated in an explicit manner. The user's attentiveness towards the message could possibly benefit from a change in design and placement. If the message was blocking a greater portion of the screen and had to be actively removed before the user could start the actual navigation, it could force the user to actually read the instructions. This would also allow a more thorough explanation.

5.4.2 All-consuming

Another problem mentioned is the fact that the user can only rely on the arrows when it comes to navigating. This resulted in the user only looking at the screen, and paying little attention to the surroundings. Some of the participants in the user tests were so engrossed in their device, that they almost walked into obstacles or fell down the staircase. When reaching the destination, some participants were also initially confused as the destination pin had an inaccuracy in the position of a few meters. The majority of the participants did at this point look up from the screen and were able to locate the correct destination.

The participants also reported that they were way more consumed by the AR solution, than the 2D digital map solution. This due to the 2D digital maps solution providing the user with more context to the proposed route, which gives the user more insight and the possibility to take more control of the navigation process. Chung et al. (2016) studied the level of engagement in pedestrian navigation, and stated that "navigation systems are often followed mindlessly, as users may focus the attention on the device and not on the path". The user is more likely to keep attention to their surroundings if they have to choose the path and with the implemented AR solution the participants were only following the arrows. In the 2D digital map solution, the participants had to make the connection between the map and their surroundings, which forces the user to be more in control and engaged to be able to follow the path to the destination. However, Chung et al. (2016) also concluded that the projection-based AR interface resulted in the users paying more attention to their surroundings, as this interface required little cognitive load, which was not the case in the results presented in chapter 4.

5.4.3 Camera angle

Many participants stated that they felt uncomfortable holding the mobile device angled upwards, as they were afraid people located in their surroundings would think they were taking pictures or filming them. This resulted in some instinctively angling the device a bit too far down, which in turn made some of the arrows appear inaccurately or not at all.

5.4.4 Points-of-Interest

The POI that was placed out (see figure 3.4) led to some confusion during the practice route of the user tests. The participants were not given an explanation of what this 3D object, nor any of the other objects, was before the navigation, and many stopped to look more closely at it. As it was in the shape of a heart, a few of the participants drew associations to video games and thought it could be some way of collecting points or similar. All the participants eventually walked away from the POI after investigating. The purpose of the POI was to simply introduce the concept to the user. In the post-experiment interview, the participants were informed that the heart was meant to mark the position of a defibrillator, and then asked about their opinions and ideas regarding this kind of information. The response was somewhat divided. Many thought the idea of providing the user with additional information was intriguing. Specific situations that were mentioned as a good use case for this kind of information were the first time visiting a conference center or a new campus, and have POIs mark emergency exist, toilets or other interesting points. Others were more skeptical as this additional information could result in being more distracting than useful. Too many 3D models on the screen could take the user's attention away from the actual navigation arrows. Many of the participants stated that such a feature should have a toggle function, where each user has the choice of whether POIs should be visible and also which kinds of POIs are relevant.

5.4.5 How-to guide

A feature that was brought up quite repetitively, was that the participants requested some kind of guide explaining how to use the AR solution. In the user test, the participants were not given any explanation before usage, as it was desired to see how intuitive the application was and to find out what was not optimal. A "how-to guide" could be a useful feature in a navigational application. The guide would preferably contain explanations of what the different 3D objects (arrows, destination pin, POIs, etc.) on the screen both looked like and what they meant, but also demonstrate how the mobile device is supposed to be held, to ensure that e.g. the arrows are rendered correctly. A message communicating that the user has to walk approximately 20 meters before the orientation of the arrows is accurate could also be included in the guide. Such a how-to guide could pop up the first time the application is used, but also be available through a menu feature.

5.4.6 Symbols and other information

The participants were explicitly asked about their opinions and thoughts regarding the symbols and other information presented in the application. A majority of the participants stated that they felt the arrows were very intuitive. The general interpretation of any arrow was to follow it, which also was the intention behind it. However, as the arrows were 3D objects placed on top of their view of the real world, an unfavorable angle could mean that the arrow shape was not as recognizable. The compass arrow was placed right in front of the user at all times, and generally, there were no problems related to this object. The turn-by-turn arrows were rendered a little further ahead in the planned path, which meant that the participants in some situations saw the 3D object from behind. From behind, the turn-by-turn arrows had a diamond shape, and the user did not necessarily initially understand that this was an arrow as well, at least not until they saw the object from a different angle. A majority of the participants stated that they generally were pleased with the design and color of both types of arrows, and felt that the arrows provided enough information to navigate.

The red destination pin also received good feedback. The majority of the participants stated that they understood that they had reached their destination when the destination pin appeared. One of the participants said that they felt unsure due to the destination pin's red color, as red often indicates that something is wrong. This participant suggested that the pin could have a green color instead, as this color more often is associated with something being right. However, only one of the participants expressed any insecurities regarding this object.

Generally, most participants stated that they were satisfied with the number of symbols and the choice of colors presented in the application, and also said that more information could have been distracting.

5.4.7 Other improvements

The participants were also asked whether they could think of any improvements to make the user experience better. Several participants suggested a line visualizing the planned route to be followed. Such a line could provide the user with more insight into the next steps, and make the route even easier to follow. Numerous participants also suggested more information regarding the destination presented on the screen. After selecting the destination, the user had no information regarding the destination until the destination pin appeared. Many stated that the name of the destination (e.g room number) should be visual, and others also requested an estimation of the remaining length until arriving at the destination. This would provide the user with more insight into the navigation process, and possibly allow them to make decisions without being consumed with the application. Other suggestions regarding destination, was that the arrows could change color when the user was within a certain proximity to the destination or a message appearing when getting close.

Changing floors was a problem when navigating in both 2D and AR. A suggested improvement of the AR solution was to include information regarding which floor the user is supposed to end up, as well as how many floors the user is expected to move up or down. This could make the transition between floors easier to understand. If the user walks in the wrong direction according to the planned path, the only indication of this is that the compass arrow is pointing in the opposite direction. Some participants suggested that information should be provided to the user if this happens. This could be through vibrations, sound, or text telling the user to take a u-turn or make a change of direction.

Chapter 6

Conclusion & Further Work

In this chapter, a conclusion and suggestions for further work are presented.

6.1 Conclusion

The goal of this thesis is to research the possibilities using AR for indoor navigation. To do this, a literature study of indoor positioning, AR technologies and AR navigation systems was conducted. In addition to this, an AR-based indoor navigation system was implemented and then tested by a group of potential users. The results of this thesis provides some answers to the research questions defined in chapter 1.

1. How can AR be used in indoor navigation?
2. What are the advantages and disadvantages with an AR-based indoor navigation system?

AR can be used in indoor navigation to communicate the path to the user through 3D objects rendered on top of a live camera capture. In order to merge AR and indoor navigation, multiple technologies and tools are available. Technologies such as ARCore and ARKit combined with tools such as ViewAR and IndoorAtlas can be utilized when designing such applications. However, when choosing which tools to use, the specific needs and circumstances of the environment should be considered in order to create a seamless implementation.

The implemented AR-based indoor navigation system has a user interface consisting of arrows, pointing the user towards the destination, and a destination pin, symbolizing the arrival. By putting the solution up against a 2D digital map solution, several advantages and disadvantages were acquired. The AR solution was described as both intuitive and simple in the user test, but it also had several flaws. An indoor navigation system is an assistive technology, and its whole purpose is to ease the task of navigation through unfamiliar and complex buildings. AR-based indoor navigation has a lot of potential, which became clear by the feedback provided by the participants in the user test. However, improvements would have to be made to make the implementation presented in this thesis a good replacement for the commonly used 2D digital map interface. One of the most important disadvantages of the AR solution is the lack of context to the navigation instructions. As the user's only source of information is the arrows rendered over the live camera capture, the solution depends on good accuracy and precision. By adding functionality which gives the user more control, the all-consuming effect could be reduced and the user could be less passive. More information regarding the path should also be present on the screen. Examples of this is name and floor of destination, as well as distance left until arrival. A suggested way to add context to the AR solution is to implement a combination of an AR and 2D digital map interface, which either have combined both views on the screen at the same time or has the opportunity to switch between the two. This has already been implemented by Google Maps, and is likely to be seen adapted by others as well.

6.2 Further Work

This research tested an AR-based indoor navigation system on only a small selection of potential users. Further work should see if any other results can be acquired by testing an AR-based solution with a more diverse group, in terms of e.g. occupation and age. It would also be interesting to conduct user tests with a group of participant consisting of people with special needs, to identify what is required to meet their abilities and needs.

The possibilities using POIs in an AR-based navigation system was looked into during this research. However, a more thorough investigation of the subject could be interesting, as POIs may have potential within the area of e.g. advertisements or tourist information belonging to surrounding landmarks.

Other further work could explore the possibilities with the ARCore Depth API. As mentioned previously, the device used in the implemented solution did not support the Depth API. However, as devices get even more powerful and advanced, it is likely that more and more devices will support this feature. The Depth API could help create more realistic AR experiences, and has a variety of different use cases. Some of the functionality that the API powers are object occlusion, improved immersion and increased user interaction (Google 2021b).

New and more powerful technologies and devices are constantly emerging, which leads to more opportunities when creating navigation systems. Smartglasses are a wearable device that allows the user to experience AR hands-free, and this kind of device enables a variety of new possibilities. Further work could research the use of smartglasses within the field of indoor navigation. It could be interesting to see whether using a device like smartglasses could eliminate the issues regarding users feeling uncomfortable and instinctively angling the camera downwards.

Bibliography

AKM (2021). *What's a Magnetic Sensor?* URL: https://www.akm.com/eu/en/technology/technical-tutorial/basic-knowledge-magnetic-sensor/magnetic-sensor/ (visited on 22nd Apr. 2021).

Albert, W. and T. Tullis (2013). *Measuring the User Experience : Collecting, Analyzing, and Presenting Usability Metrics.* Morgan Kaufmann.

Alnabhan, A. and B. Tomaszewsk (2014). 'INSAR: Indoor Navigation System using Augmented Reality'. In: *GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems*, pp. 36–43.

Amazon (2021). *View in Your Room.* URL: https://www.amazon.com/adlp/arview (visited on 22nd Mar. 2021).

Amirian, P. and A. Basiri (2016). 'Landmark-Based Pedestrian Navigation Using Augmented Reality and Machine Learning'. In: *Progress in Cartography.* Ed. by G. Gartner, M. Jobst and H. Huang. Springer, Cham.

Apple (2021a). *ARCamera.* URL: https://developer.apple.com/documentation/arkit/arcamera (visited on 13th May 2021).

— (2021b). *Displaying an AR Experience with Metal.* URL: https://developer.apple.com/documentation/arkit/displaying_an_ar_experience_with_metal (visited on 13th May 2021).

— (2021c). *Understanding World Tracking.* URL: https://developer.apple.com/documentation/arkit/configuration_objects/understanding_world_tracking (visited on 12th May 2021).

— (2021d). *Dive into the world of augmented reality.* URL: https://developer.apple.com/augmented-reality/ (visited on 24th Feb. 2021).

AppleInsider (2021). *ARKit.* URL: https://appleinsider.com/inside/arkit (visited on 12th May 2021).

Ariso, José M. (2017). 'Is Critical Thinking Particularly Necessary when Using Augmented Reality in Knowledge Society? An Introductory Paradox'. In: *Augmented Reality.* De Gruyter.

AVSystem (2019). *What an Indoor Navigation System is and how to use it in your venue.* URL: https://www.avsystem.com/blog/indoor-navigation-and-indoor-positioning (visited on 22nd Apr. 2021).

Bardi, Joe (2019). *What is Virtual Reality?* URL: https://www.marxentlabs.com/what-is-virtual-reality/ (visited on 26th Mar. 2021).

Brena, R. F. et al. (2017). 'Evolution of Indoor Positioning Technologies: A Survey'. In: *Sensors* 17.359.

Cabero-Almenara, J. et al. (2019). 'Educational Uses of Augmented Reality (AR): Experiences in Educational Science'. In: *Sustainability* 11.4990.

Cambridge Dictionary (2021). *Interaction*. URL: https://dictionary.cambridge.org/dictionary/english/interaction (visited on 10th May 2021).

Caudell, T. and D. Mizell (1992). 'Augmented reality: An application of heads-up display technology to manual manufacturing processes'. In: *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences* 2.

Chandra, Harsh (2019). *Augmented Reality (AR) Development: Tools and Platforms*. URL: https://heartbeat.fritz.ai/augmented-reality-ar-development-tools-and-platforms-26a6de07d12e (visited on 11th May 2021).

Chung, J., F. Pagnini and E. Langer (2016). 'Mindful navigation for pedestrians: Improving engagement with augmented reality'. In: *Technology in Society* 45.

Correa, A. et al. (2017). 'A Review of Pedestrian Indoor Positioning Systems for Mass Market Applications'. In: *Sensors* 17.1927 (8).

Craig, Alan B. (2013). 'Chapter 2 - Augmented Reality Concepts'. In: *Understanding Augmented Reality*. Ed. by Alan B. Craig. Morgan Kaufmann, pp. 39–67.

Delfa, G.C. La et al. (2016). 'Performance analysis of visualmarkers for indoor navigation systems'. In: *Frontiers of Information Technology & Electronic Engineering* 17 (8).

Dong, W. et al. (2021). 'What is the difference between augmented reality and 2D navigation electronic maps in pedestrian wayfinding?' In: *Cartography and Geographic Information Science*.

Duignan, John (2016). *A Dictionary of Business Research Methods*. Oxford University Press.

Epson (2021). *Gyro sensors - How they work and what's ahead*. URL: https://www5.epsondevice.com/en/information/technical_info/gyro/ (visited on 22nd Apr. 2021).

Facebook (2021). *Spark AR*. URL: https://sparkar.facebook.com/ar-studio/ (visited on 24th Mar. 2021).

Fallah, N. et al. (2013). 'Indoor Human Navigation Systems: A Survey'. In: *Interacting with Computers 25*, pp. 21–33.

Franklin Institute (2021). *The Science of Augmented Reality*. URL: https://www.fi.edu/science-of-augmented-reality (visited on 27th Apr. 2021).

Fusco, G. and J.M. Coughlan (2018). 'Indoor Localization Using Computer Vision and Visual-Inertial Odometry'. In: *Computers Helping People with Special Needs*. Ed. by Klaus Miesenberger and Georgios Kouroupetroglou. Springer, pp. 86–93.

GIGXR (2021). *Extended Reality*. URL: https://www.gigxr.com/immersive-technology-in-healthcare-training (visited on 26th Mar. 2021).

Gjøvåg, Christopher Wennersteen (2020). 'WiFi RTT for Indoor Localization using Google WiFi and Google Pixel 3a'. MA thesis. Norwegian University of Science and Technology.

Glasgow, Dane (2021). *Redefining what a map can be with new information and AI*. URL: https://blog.google/products/maps/redefining-what-map-can-be-new-information-and-ai (visited on 11th May 2021).

Gnanamoorthy, Sabarish (2021). *WaypointAR*. URL: http://waypointar.com/ (visited on 14th May 2021).

Google (2020). *ARCore overview*. URL: https://developers.google.com/ar/discover (visited on 12th Apr. 2021).

— (2021a). *ARCore supported devices*. URL: https://developers.google.com/ar/devices (visited on 1st June 2021).

— (2021b). *Depth API overview for Android.* URL: https://developers.google.com/ar/develop/java/depth/overview (visited on 30th Apr. 2021).
— (2021c). *Augmented Reality.* URL: https://arvr.google.com/ar/ (visited on 24th Feb. 2021).
— (2021d). *Choose your development environment.* URL: https://developers.google.com/ar/develop (visited on 18th May 2021).
— (2021e). *Fundamental concepts.* URL: https://developers.google.com/ar/discover/concepts?hl=en (visited on 18th May 2021).
Haselton, Todd (2021). *Google Maps has a wild new feature that will guide you through indoor spaces like airports.* URL: https://www.cnbc.com/2021/03/30/google-maps-launches-augmented-reality-directions-for-indoor-spaces.html (visited on 18th May 2021).
He, Amy (2019). *People Continue to Rely on Maps and Navigational Apps.* URL: https://www.emarketer.com/content/people-continue-to-rely-on-maps-and-navigational-apps-emarketer-forecasts-show (visited on 9th Apr. 2021).
Hendrickx, Sam (2019). *Creating an ARCore powered indoor navigation application in Unity.* URL: https://blog.craftworkz.co/creating-an-arcore-powered-indoor-navigation-application-in-unity-cab5b53b8861 (visited on 21st May 2021).
Hinderer, D. and A. Arbor (1998). 'Challenges in Participant Recruiting for Usability Testing'. In: *IPCC 98. Contemporary Renaissance: Changing the Way we Communicate. Proceedings 1998 IEEE International Professional Communication Conference* 2.
Hollander, Rayna (2017). *Two-thirds of the world's population are now connected by mobile devices.* URL: https://www.businessinsider.com/world-population-mobile-devices-2017-9?r=US&IR=T (visited on 9th Apr. 2021).
hotjar (2020). *The different types of usability testing methods for your projects.* URL: https://www.hotjar.com/usability-testing/methods/ (visited on 9th Apr. 2021).
Huang, B. et al. (2020). 'ARBIN: Augmented Reality Based Indoor Navigation System'. In: *Sensors* 20.5890.
Huang, H. and G. Gartner (2010). 'A Survey of Mobile Indoor Navigation Systems'. In: *Cartography in Central and Eastern Europe*, pp. 305–319.
Huang, H., M. Schmidt and G. Gartner (2012). 'Spatial Knowledge Acquisition with Mobile Maps, Augmented Reality and Voice int the Context of GPS-based Pedestrian Navigation: Results from a Field Test'. In: *Cartography and Geographic Information Science* 39 (2).
Huilla, Sami (2019). 'Smartphone-based Indoor Positioning Using Wi-Fi Fine Timing Measurement Protocol'. MA thesis. University of Turku.
IndoorAtlas (2018a). *Adding a Floor Plan.* URL: https://indooratlas.freshdesk.com/support/solutions/articles/36000050481-adding-a-floor-plan (visited on 22nd Apr. 2021).
— (2018b). *Adding Waypoints for Fingerprinting.* URL: https://indooratlas.freshdesk.com/support/solutions/articles/36000050489-adding-waypoints-for-fingerprinting (visited on 22nd Apr. 2021).
— (2018c). *Calibration for Fingerprinting.* URL: https://indooratlas.freshdesk.com/support/solutions/articles/36000050488-calibration-for-fingerprinting (visited on 22nd Apr. 2021).

IndoorAtlas (2018d). *Creating a Venue (Location)*. URL: https://indooratlas.freshdesk.com/support/solutions/articles/36000050478-creating-a-venue-location- (visited on 22nd Apr. 2021).

— (2018e). *Recording Paths*. URL: https://indooratlas.freshdesk.com/support/solutions/articles/36000050494-recording-paths (visited on 26th Apr. 2021).

— (2019a). *Fingerprinting Overview*. URL: https://indooratlas.freshdesk.com/support/solutions/articles/36000103986-fingerprinting-overview (visited on 22nd Apr. 2021).

— (2019b). *Planning beacon deployments with IndoorAtlas*. URL: https://www.indooratlas.com/2019/02/04/planning-beacon-deployments-with-indooratlas/ (visited on 7th June 2021).

— (2020a). *Creating Applications and API Keys*. URL: https://indooratlas.freshdesk.com/support/solutions/articles/36000050559-creating-applications-and-api-keys (visited on 26th Apr. 2021).

— (2020b). *Improving Mapping Quality*. URL: https://indooratlas.freshdesk.com/support/solutions/articles/36000050540-improving-mapping-quality (visited on 26th Apr. 2021).

— (2020c). *IndoorAtlas goes AR*. URL: https://www.indooratlas.com/2020/11/13/indooratlas-goes-ar/ (visited on 29th May 2021).

— (2021a). *Devices Compatible with Fingerprinting*. URL: https://indooratlas.freshdesk.com/support/solutions/articles/36000054947-devices-compatible-with-fingerprinting (visited on 1st June 2021).

— (2021b). *SDK 3.4 Release Information*. URL: https://docs.indooratlas.com/technical/release-notes/sdk-34-release-information/ (visited on 12th Apr. 2021).

— (2021c). *About us*. URL: https://www.indooratlas.com/about-us/ (visited on 12th Apr. 2021).

— (2021d). *Last meter accuracy through technology fusion*. URL: https://www.indooratlas.com/positioning-technology/ (visited on 12th Apr. 2021).

Inman, Rachel (2019). *Take off to your next destination with Google Maps*. URL: https://blog.google/products/maps/take-your-next-destination-google-maps/ (visited on 11th May 2021).

intel (2021). *The differences between Virtual Reality, Augmented Reality, and Mixed Reality*. URL: https://www.intel.com/content/www/us/en/tech-tips-and-tricks/virtual-reality-vs-augmented-reality.html (visited on 24th Mar. 2021).

ISL (2018). *Every (Major) AR Development Platform and Why It's Important*. URL: https://medium.com/inborn-experience/every-major-ar-development-platform-and-why-its-important-fbd9bc300fd5 (visited on 11th May 2021).

ISO 9241-11:2018 (2018). *Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts*. Standard. International Organization for Standardization.

Jhangiani, R.S., I-C.A. Chiang and P.C. Price (2015). *Research Methods in Psychology*. BCcampus.

Josh (2020). *Comparing ARKit vs ARCore vs Vuforia: The Best Augmented Reality Toolkit*. URL: https://bluewhaleapps.com/blog/comparing-arkit-vs-arcore-vs-vuforia-the-best-augmented-reality-toolkit (visited on 19th May 2021).

Kitchin, R. and N. J. Tate (2000). *Conducting Research into Human Geography - theory, methodology & practice*. Pearson Education Limited.

Kunhoth, J. et al. (2020). 'Indoor positioning and wayfinding systems: a survey'. In: *Human-centric Computing and Information Sciences* 10.18.

Lee, Ken (2021). *AR Foundation in Unity: Getting Started.* URL: https://www.raywenderlich.com/14808876-ar-foundation-in-unity-getting-started (visited on 18th May 2021).

Leon, Melanie (2020). *Ultimate AR Comparison Guide: ARKit vs ARCore vs Vuforia vs AR Foundation.* URL: https://circuitstream.com/blog/augmented-reality-guide/ (visited on 12th May 2021).

Liu, H. et al. (2007). 'Survey of Wireless Indoor Positioning Techniques and Systems'. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 37 (6), pp. 1067–1080.

Mainetti, L., L. Patrono and I. Sergi (2014). 'A Survey on Indoor Positioning Systems'. In: *22nd International Conference on Software, Telecommunications and Computer Networks*, pp. 111–120.

Makarov, Andrew (2019). *How Augmented Reality-Based Indoor Navigation System Works.* URL: https://mobidev.biz/blog/augmented-reality-indoor-navigation-app-developement-arkit (visited on 20th May 2021).

— (2020). *10 Augmented Reality Trends in 2021.* URL: https://mobidev.biz/blog/augmented-reality-future-trends-2018-2020 (visited on 25th Mar. 2021).

— (2021). *Augemented Reality Development: Guide for Business Owners and Managers.* URL: https://mobidev.biz/blog/augmented-reality-development-guide (visited on 24th Feb. 2021).

Marr, Bernard (2019). *What Is Extended Reality Technology?* URL: https://www.forbes.com/sites/bernardmarr/2019/08/12/what-is-extended-reality-technology-a-simple-explanation-for-anyone (visited on 26th Mar. 2021).

— (2021). *The Important Difference Between Augmented Reality And Mixed Reality.* URL: https://bernardmarr.com/default.asp?contentID=1912 (visited on 26th Mar. 2021).

Material (2021). *Snackbars.* URL: https://material.io/components/snackbars#usage (visited on 30th Apr. 2021).

MathWorks (2021). *What is SLAM?* URL: https://se.mathworks.com/discovery/slam.html (visited on 19th May 2021).

MazeMap (2021). *Data: Pois.* URL: https://api.mazemap.com/js/v2.0.12/docs/#api-data:-pois (visited on 30th Apr. 2021).

McLeod, Saul A. (2012). *Experimental method.* URL: https://www.simplypsychology.org/experimental-method.html (visited on 9th Apr. 2021).

Merriam Webster (2021a). *Navigating.* URL: https://www.merriam-webster.com/dictionary/navigating#learn-more (visited on 9th Apr. 2021).

— (2021b). *Navigation.* URL: https://www.merriam-webster.com/dictionary/navigation#learn-more (visited on 9th Apr. 2021).

Microsoft (2021a). *HoloLens 2.* URL: https://www.microsoft.com/en-us/hololens/hardware (visited on 24th Mar. 2021).

— (2021b). *Microsoft Mesh.* URL: https://www.microsoft.com/en-us/mesh (visited on 24th Mar. 2021).

Miller, D., T. Mowrer and B. Weiers (2018). *Unity's Handheld AR Ecosystem: AR Foundation, ARCore and ARKit.* URL: https://blog.unity.com/technology/unitys-handheld-ar-ecosystem-ar-foundation-arcore-and-arkit (visited on 22nd May 2021).

Montello, Daniel R. (2005). *The Cambridge Handbook of Visuospatial Thinking*. Cambridge University Press.

Niantic (2021). *Play Where You Are?* URL: https://pokemongolive.com/en/play-where-you-are/ (visited on 19th May 2021).

Nielsen, Jakob (1994). *10 Usability Heuristics for User Interface Design*. URL: https://www.nngroup.com/articles/ten-usability-heuristics/ (visited on 5th May 2021).

Nikitin, Nikita (2020). *How To Choose A Technology For An AR Project. Part 1: Vuforia*. URL: https://medium.com/@ARchy_Team/how-to-choose-a-technology-for-an-ar-project-part-1-vuforia-65b92d7c7561 (visited on 19th May 2021).

Pan, Z. et al. (2006). 'Virtual reality and mixed reality for virtual learning environments'. In: *Computers & Graphics* 30.1.

Perez, Sarah (2020). *Amazon launches an AR app that works with new QR codes on its boxes*. URL: https://techcrunch.com/2020/10/12/amazon-launches-an-ar-app-that-works-with-new-qr-codes-on-its-boxes/ (visited on 22nd Mar. 2021).

Picard, Gabriel (2020). *1965 - Ivan Sutherland, Father of AR*. URL: https://atomicdigital.design/blog/1965-ivan-sutherland-father-of-ar (visited on 25th Mar. 2021).

Porras, C. et al. (2018). 'Advantages of virtual reality in the rehabilitation of balance and gait'. In: *Neurology* 90 (22).

Program-Ace (2019). *5 Top Augmented Reality Tools for App Development*. URL: https://program-ace.com/blog/augmented-reality-sdk/ (visited on 12th May 2021).

Reinhardt, Tilman (2019). *Using Global Localization to Improve Navigation*. URL: https://ai.googleblog.com/2019/02/using-global-localization-to-improve.html (visited on 11th May 2021).

Ridland, Michael (2019). *Introduction to Augmented Reality with ARKit*. URL: https://michaelridland.com/mobile/introduction-to-augmented-reality-with-arkit/ (visited on 13th May 2021).

Romli, R. et al. (2020). 'Mobile Augmented Reality (AR) Marker-based for Indoor Library Navigation'. In: *Conference Series: Materials Science and Engineering* 767.

Saha, S. et al. (2015). 'TrackMe - A Low Power Location Tracking System Using Smart Phone Sensors'. In: *2015 International Conference on Computing and Network Communications (CoCoNet)*. IEEE, pp. 457–464.

Sakpere, W., M. Adeyeye-Oshin and N.B. W. Mlitwa (2017). 'A state-of-the-art survey of indoor positioning and navigation systems and technologies'. In: *South African Computer Journal* 29, pp. 145–197.

Sato, Fumiaki (2018). 'Indoor Navigation System Based on Augmented Reality Markers'. In: *Innovative Mobile and Internet Services in Ubiquitous Computing*. Ed. by Leonard Barolli and Tomoya Enokido. Springer International Publishing, pp. 266–274.

Slater, Michael (2019). *Facebook Shares Major Spark AR Studio Update*. URL: https://developers.facebook.com/blog/post/2019/04/30/spark-ar-studio-update/ (visited on 12th May 2021).

Sutherland, Ivan E. (1968). 'A Head-Mounted Three Dimensional Display'. In: *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I*, pp. 757–764.

Svalastog, Mari Saua (2007). 'Indoor Positioning - Tehnologies, Servies and Architetures'. MA thesis. University of Oslo.

Techopedia (2021). *Platform Agnostic.* URL: https://www.techopedia.com/definition/23666/platform-agnostic (visited on 19th May 2021).

Thrun, Sebastian (2008). 'Simultaneous Localization and Mapping'. In: *Robotics and Cognitive Approaches to Spatial Mapping.* Ed. by Margaret E. Jefferies and Wai-Kiang Yeap. Springer, pp. 13–41.

To, Dara (2021). *Introduction to ARKit.* URL: https://designcode.io/arkit-intro (visited on 12th May 2021).

Unity (2021a). *About AR Foundation.* URL: https://docs.unity3d.com/Packages/com.unity.xr.arfoundation@4.1/manual/index.html (visited on 19th May 2021).

— (2021b). *Unity Personal.* URL: https://store.unity.com/products/unity-personal (visited on 19th May 2021).

Valvo, A. Lo et al. (2021a). 'A Navigation and Augmented Reality System for Visually Impaired People'. In: *Sensors* 21 (9).

— (2021b). 'Indoor Navigation Using Augmented Reality'. In: *Sensors* 21 (9).

Vaughan-Nichols, S. J. (2009). 'Augmented Reality: No Longer a Novelty?' In: *Computer* 42.12.

ViewAR (2021). *INDOAR - MUSEUM.* URL: https://museum.viewar.com/ (visited on 12th May 2021).

Vuforia (2021a). *Device Tracking Overview.* URL: https://library.vuforia.com/features/environments/device-tracker-overview.html (visited on 19th May 2021).

— (2021b). *Ground Plane.* URL: https://library.vuforia.com/features/environments/ground-plane-guide.html (visited on 19th May 2021).

— (2021c). *Vuforia Engine Pricing: The Right Plan to Build Your Vision.* URL: https://www.ptc.com/en/products/vuforia/vuforia-engine/pricing (visited on 19th May 2021).

— (2021d). *VuMark.* URL: https://library.vuforia.com/features/objects/vumark.html (visited on 19th May 2021).

Wang, C. et al. (2019). 'Gender and Age Differences in Using Indoor Maps for Wayfinding in Real Environments'. In: *ISPRS International Journal of Geo-Information* 8.11 (1).

Wang, Z., S. Huang and G. Dissanayake (2011). *Simultaneous Localization and Mapping: Exactly Sparse Information Filters.* World Scientific Publishing Company.

Wemap (2021). *Navigation with Maps and Augmented Reality.* URL: https://getwemap.com/navigation-wayfinding/ (visited on 14th May 2021).

Wiig, Thomas F. (2010). 'Assessment of Indoor Positioning System (IPS) technology'. MA thesis. University of Oslo.

Xia, S. et al. (2017). 'Indoor Fingerprint Positioning Based on Wi-Fi: An Overview'. In: *ISPRS International Journal of Geo-Information* 6.

Zhang, X. et al. (2019). 'An ARCore Based User Centric Assistive Navigation System for Visually Impaired People'. In: *Applied Sciences* 9 (5).

Zhou, Y. et al. (2020). 'How does gender affect indoor wayfinding under time pressure?' In: *Cartography and Geographic Information Science* 47 (4).

Zuo, Z. et al. (2018). 'Indoor Positioning Based on Bluetooth Low-Energy Beacons Adopting Graph Optimization'. In: *Sensors* 18.11.

Appendix

A Pre-experimental questions

The test subjects were asked to answer a number of questions before conducting the practical part of the experiment. The questions were used to collect statistical information about the test subjects. The questions are listed bellow.

1. What is your full name?
2. What is your age?
3. What is your sex?
4. What is your field of study?
5. Do you have previous experience with mobile applications?
6. Do you have previous experience with navigation systems?
7. Do you have previous experience with pedestrian navigation systems?
8. Do you have previous experience with indoor navigation systems?
9. Do you have previous experience with 2D digital maps?
10. Do you have previous experience with 2D map indoor navigation systems?
11. Do you have previous experience with AR applications?
12. Do you have previous experience with AR navigation systems?
13. If you are located in an unfamiliar building, do you tend to use indoor navigation systems?
14. Does using navigation systems to navigate through an unfamiliar building help you remember the path next time?
15. Do you have any prior experience with MazeMap's systems?
16. Have you ever been to *Byggtekniske laboratorier* before?
17. If so, how familiar would you say you are?

18. Do you give your permission so that data from this experiment can be stored?

Some of the questions had a number of predetermined answers that the subjects could choose between to answer. The purpose was to make the questions easier to understand, and to make the data acquisition more structured. Question 2 divided the test subjects into three age groups; *18-21*, *22-25* and *26+*. Question 3 had the options *male*, *female*, and *other*, where the participant had the option to write their own answer. Question 5 to 12 had the options *none*, *some* or *much*. As explained to the subjects, none meant *never used it*, some meant *have tried it* and much meant *use it on a regular basis*. Question 14, 15, 16 and 18 allowed only *yes* or *no* answers. If the subject had been to the building before (question 16), they had to answer question 17. Question 17 had the subjects rate their familiarity with the building on a scale from *1 to 5*, where 1 meant almost no knowledge of the building and 5 meant very familiar.

B Post-experimental interview

After the more practical part of the experiment, the test subjects were asked some investigating questions to find out more about their experience, impressions and thoughts. The questioning session took form of a interview, where the answers were more free than in the pre-experimental questions and the answers were written down. The questions asked are listed bellow.

1. Did you prefer the 2D map or the AR solution?
   - Why did you prefer it?
   - Give some advantages and disadvantages with both solutions.
2. If a combination of the 2D and AR was obtainable, would you prefer the combination or still just 2D or AR?
3. What is your opinion on the information presented to you in the AR application? *(Symbols, text, colors, etc.)*
4. What is your opinion on POIs in an indoor navigation application?
   - Do you find it distracting? Unnecessary? Beneficial? Why?
   - Can you think of any possible use cases were it would be helpful?
5. What kind of information did you miss and would like to see in this kind of application?
6. Do you have any additional comments or remarks about the application, experiment or anything else?

When the test subjects were asked question 2, a figure illustrating the three possible solutions were given to them for examination. This was to give them a visual aid to help them make up an opinion. The figure (figure 3.5) is repeated in an enlarged version in the following page.

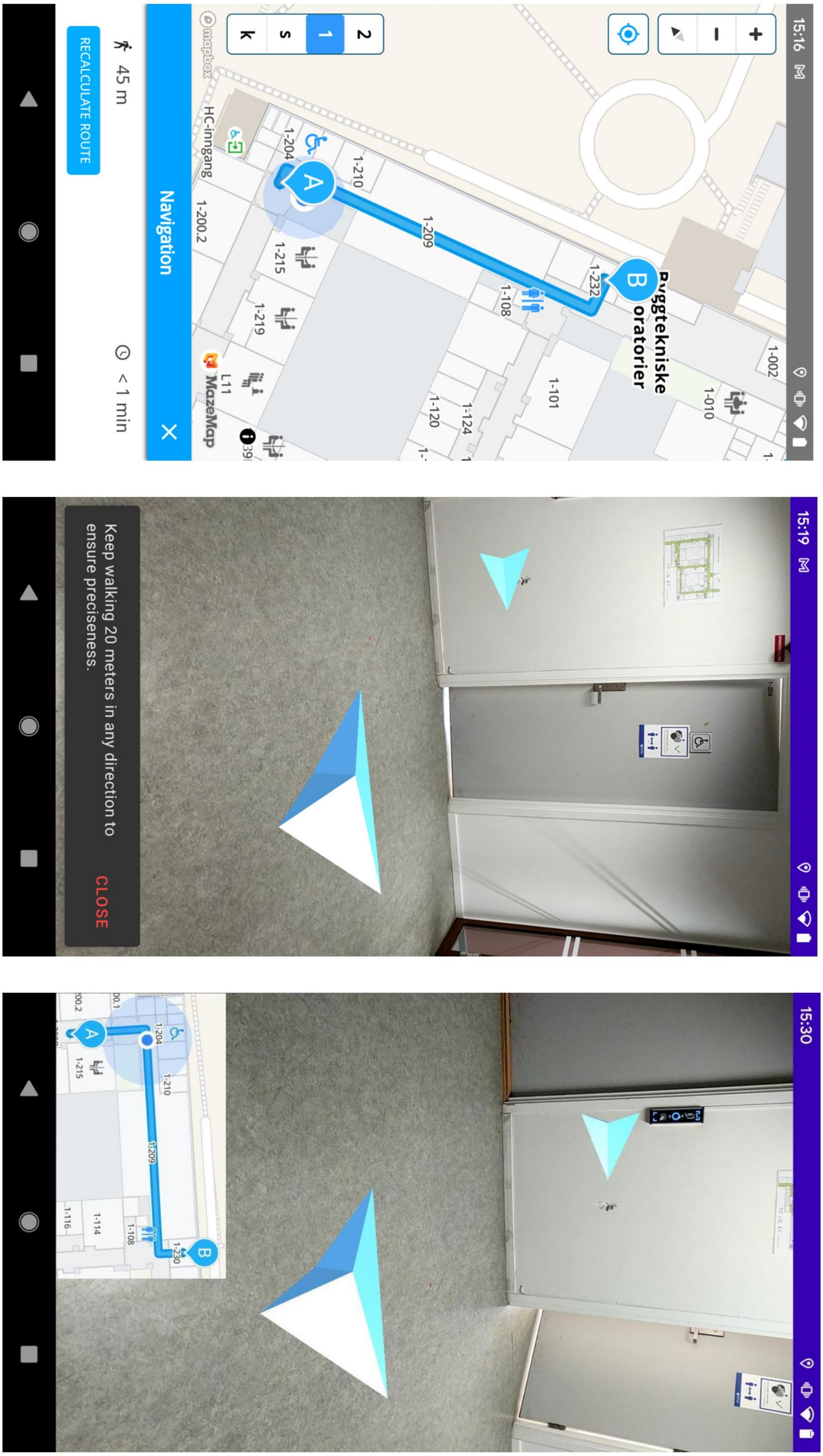


Figure 3.5: Visualizations of 2D map, AR and a possible combination of them (repeated from page 56)

Andrea Slyngstadli & Shahitha Sothinathan

Possibilities Using AR for Indoor Navigation

NTNU

Norwegian University of
Science and Technology