Heads Up!

Guiding Attention to Moving Physical Hazards in 3D Space Using AR and UWB-Based RTLS

Linnéa Gadd and Oliver Levay

DEPARTMENT OF DESIGN SCIENCES
FACULTY OF ENGINEERING LTH | LUND UNIVERSITY
2026

MASTER THESIS

FORESIGHT
DATA MACHINES

Heads Up!

Guiding Attention to Moving Physical Hazards in 3D Space Using AR and UWB-Based RTLS

Linnéa Gadd and Oliver Levay

LUND UNIVERSITY

Heads Up!

Guiding Attention to Moving Physical Hazards in 3D Space Using AR and UWB-Based RTLS

Published by
Department of Design Sciences
Faculty of Engineering LTH, Lund University
Box 118
SE-221 00 LUND
Sweden

Subject: Virtual Reality and Augmented Reality (MAMM20)
Supervisor(s): Günter Alce, gunter.alce@design.lth.se and
Jiří Kučera, jiri@foresightdatamachines.com
Examiner: Joakim Eriksson, joakim.eriksson@design.lth.se

Abstract

In heavy industrial environments such as steel manufacturing plants, workers interact closely with moving machinery under conditions of limited visibility and high noise. A recent World Steel Association report attributes roughly half of industry fatalities to causes with a strong attentional component, motivating research into augmented-reality (AR) cues that direct workers' attention to hazards outside their field of view. This thesis presents ScrapAwear, a system combining an ultra-wideband (UWB) real-time locating system with a Meta Quest 3 headset to guide users toward out-of-view physical targets in 3D space. Two AR navigation cues were implemented: a 3D arrow that points directly at the target, and a top-down minimap that shows the target's position relative to the user. Their effectiveness was compared in a within-subject user study ($n = 16$) using a dual-task design where participants performed a 1-back letter task while detecting a UWB-tracked object moving behind them. The 3D arrow produced significantly faster detection times than the minimap (mean 1571 ms vs. 2464 ms, $p = 0.005$) and faster concurrent n-back reaction times ($p = 0.044$). NASA TLX workload was lower with the arrow on every subscale, though the overall score did not reach significance ($p = 0.052$). The arrow's advantage was largest for target placements outside the horizontal plane, reflecting the minimap's lack of vertical information. Subjective preferences matched objective performance for 13 of the 15 participants who expressed a preference. The findings show that the 3D arrow used in this study outperformed the minimap under dual-task conditions. They do not necessarily establish a general superiority of arrows over maps.

Keywords: augmented reality, ultra-wideband, real-time location system, attention guidance, navigation cues, user study

Sammanfattning

I tunga industrimiljöer som stålverk arbetar operatörer nära rörliga maskiner under dålig sikt och hög ljudnivå. En rapport från World Steel Association anger att ungefär hälften av dödsolyckorna i stålindustrin involverar en stark uppmärksamhetskomponent, vilket motiverar forskning kring augmented reality-signaler (AR) som riktar arbetarnas uppmärksamhet mot faror utanför synfältet. Detta examensarbete presenterar ScrapAwear, ett system som kombinerar ett ultra-wideband (UWB)-baserat realtidspositioneringssystem med en Meta Quest 3 för att vägleda användaren till fysiska objekt utanför synfältet i tredimensionellt rum. Två AR-navigeringsvisare implementerades: en 3D-pil som pekar direkt mot målet, och en minikarta sedd ovanifrån som visar målets position i förhållande till användaren. Deras effektivitet jämfördes i en användarstudie med within-subject-design ($n = 16$), där deltagarna utförde ett 1-back-bokstavstest samtidigt som de skulle upptäcka ett UWB-spårat objekt som rörde sig bakom dem. 3D-pilen gav signifikant kortare upptäcktstider än minikartan (medelvärde 1571 ms mot 2464 ms, $p = 0{,}005$) och snabbare reaktionstid i det parallella 1-back-testet ($p = 0{,}044$). Den uppskattade arbetsbelastningen på NASA TLX var lägre med pilen på samtliga delskalor, även om totalpoängen inte nådde signifikans ($p = 0{,}052$). Pilens fördel var störst för mål placerade utanför det horisontella planet, vilket speglar minikartans avsaknad av höjdinformation. Subjektiv preferens överensstämde med objektiv prestation för 13 av de 15 deltagare som uttryckte en preferens. Resultaten visar att 3D-pilen som användes i användarstudien överträffade minikartan vid kognitiv belastning. De visar inte nödvändigtvis att pilar generellt överträffar kartor.

Nyckelord: förstärkt verklighet, ultra-wideband, realtidspositionering, uppmärksamhetsstyrning, navigeringsvisare, användarstudie

Acknowledgements

We would like to thank our supervisor Günter Alce at LTH and our co-supervisor Jiří Kučera at Foresight Data Machines (FDM) for their guidance and feedback throughout this project. We are grateful to FDM for hosting this thesis work and for providing the industrial context that motivated it. This work built on DipFlip's open-source UWB positioning project, which provided the foundation for our positioning server, and Ai-Thinker's BU03-Kit, which enabled the UWB ranging hardware. Finally, we thank the 16 participants who took part in the user study.

Lund, May 2026
Linnéa Gadd and Oliver Levay

Table of contents

List of acronyms and abbreviations

ADB	Android Debug Bridge
AoA	angle of arrival
AR	augmented reality
BLE	bluetooth low energy
FDM	Foresight Data Machines
FOV	field of view
GNSS	global navigation satellite system
GPS	global positioning system
HMD	head-mounted display
HRTF	head-related transfer function
LiDAR	light detection and ranging
LOS	line-of-sight
MQDH	Meta Quest Developer Hub
MR	mixed reality
NASA TLX	NASA task load index
NLOS	non-line-of-sight
OHMD	optical head-mounted display
OTG	on-the-go (USB)
PDoA	phase difference of arrival
POI	point of interest
PWS	proximity warning system
RFID	radio frequency identification
RIR	room impulse response
RTLS	real-time locating system
RTT	round-trip time
SA	situational awareness
TDoA	time difference of arrival
ToA	time of arrival
ToF	time of flight
TWR	two-way ranging
UWB	ultra-wideband
VR	virtual reality
WARD	wearable augmented reality display
XR	extended reality

1 Introduction

The thesis was completed in collaboration with Foresight Data Machines (FDM) Ltd. This chapter introduces the context, research questions, and goals of the thesis.

1.1 Overview

Significant technological advances have been made in industrial environments over recent decades. There are, however, still processes relying on human operators working while sharing a workspace with robots and machinery [23]. These workplace conditions demand a high level of situational awareness (SA) to perceive and avoid hazards in their surroundings [44]. Continuing to find ways to support workers in maintaining both efficiency and safety therefore remains important for the industry [36].

Concepts such as *Industry 4.0* and *Industry 5.0* have been defined to describe the increasing integration of technology in industrial environments. *Industry 4.0*, introduced in 2011, describes an industrial environment optimised and digitalised through concepts such as cyber-physical production systems (CPPS), leading to smarter factories [33, 55]. *Industry 5.0* emerged a few years later, strongly connected to the technological growth and collaboration between humans and intelligent machines [55]. The increasing digitalisation introduces new opportunities, but also more complex workspaces. With a 'human-centric approach' as one of the core values, *Industry 5.0* emphasises the safety and well-being of the workers in the industry [55].

Foresight Data Machines (FDM) Ltd. is a technology company that develops data-driven and machine-learning-based systems for heavy industrial environments. The company focuses on industrial environments where materials of varying composition are moved intermittently using cranes or mobile equipment, making accurate measurement and tracking challenging. Due to the complexity of material handling and limited sensing coverage, it is difficult to track when material is moved or consumed. As a result, raw material behaviour frequently constitutes one of the least controlled variables in industrial operations.

To address these challenges, FDM develops systems for material tracking, material classification, and process monitoring using a combination of computer vision, sensor data, and machine learning models. These systems are designed to integrate material-level observations with downstream process data in order to support improved process understanding and decision-making. The company's solutions are deployed in operational industrial settings, where close collaboration between engineers, software developers, and domain specialists enables the development of systems tailored to site-specific constraints and requirements.

Many of the challenges addressed by FDM's systems arise in environments where operators interact closely with heavy machinery under conditions of limited visibility, high noise levels, and dynamic material movement. In steel manufacturing plants, safety management is a high priority. Numerous hazards exist that are often outside workers' immediate view, making it vital for workers to maintain a high level of SA to avoid accidents. Despite extensive preventive efforts in manufacturing plants, accidents continue to occur. According to the World Steel Association [43] report on safety and health in the steel industry, "Moving machinery" is the second most common cause of fatality over the past decade, see Table 1.1.

Several of the fatality causes mentioned in the report involve dynamic hazards such as falling objects, overhead cranes and moving machinery. These are examples of incidents where a worker's safety might partly depend on their ability to perceive, understand and predict the best action in time. There are various possible factors that could worsen a person's SA in these situations [44]. Reduced SA has been connected to an increased risk of human errors and accidents [44]. While preventative measures remain essential for improving workplace safety, having an additional system capable of warning workers about potential hazards could serve as a complement. Technological advancements have not only optimised the industry, but also enabled the development of tools that could improve safety. This thesis explores the use of augmented reality (AR) with a head-mounted display (HMD) for directing attention and warning workers about hazards outside of their field of view (FOV). Combining the AR-based visual warnings with a real-time locating system (RTLS) for tracking physical hazards aims to support the SA of workers in industrial environments.

Table 1.1 **Causes of fatalities in the steel industry (2015–2024).**

Rank	Cause	Fatalities
1	Fall from height	164
2	Moving machinery	139
3	On site road vehicle	77
4	Gassing and asphyxiation	76
5	Falling object	67
6	Overhead crane	45
7	Electrical	41
8	Fire	37
9	Hot metal	35
10	Explosion	30
11	Structural failure	29
12	Rail	28
13	Hot substances	28
14	Other mobile equipment	20
15	Forklift	14
16	Product handling storage	12
17	Manual tasks tools	10
18	Slips, trips and falls	6
19	Product loading	6
20	Off site road vehicle	6
21	Exposure to chemicals	3
22	Unknown	1

Note. Adapted from the 2025 safety and health report by the World Steel Association [43].

1.2 Goals and research questions

The goal of this thesis is to investigate whether, and under which conditions, AR-based visual attention guidance is safe and useful when applied to moving physical hazards in a noisy industrial environment.

RQ1: *How can the world-locking capabilities of the Meta Quest 3 simplify the process of calibrating out-of-view real-time locating systems?*

RQ2: *How can AR be used to evaluate the effectiveness of real-time locating systems?*

RQ3: *Which AR navigation cue, a 3D arrow or a minimap, better supports the detection of out-of-view physical objects?*

1.3 Sustainable development

This thesis contributes to United Nations Sustainable Development Goal 9 (see Figure 1.1), which is focused on infrastructure, innovation and sustainable industrialisation [49]. Researching the use of AR in an industrial setting aims to contribute to the technological capabilities in the industry, with a focus on workplace safety.

Figure 1.1 Sustainable Development Goal 9

1.4 Ethical considerations

The aim of this thesis is to contribute to the development of AR glasses that can be used to give warnings in an industrial environment. One issue that could arise when using the intended end product is automation bias, where the worker could start to favour the AR warnings and trust those more than their own perception of their surrounding environment. Research from 2024 that required participants to find real-world objects with the help of AR cueing showed that imperfect cueing led to overall lower performance and longer response times [40]. This could cause issues when the warnings are supposed to increase safety.

1.5 Usage of AI

AI tools such as Claude and Cursor have been used to assist in developing the software used for this thesis. AI tools have also been used for spell checking and proofreading this report.

2 Background

This chapter explains the technology and terminology used in the rest of the thesis.

2.1 Extended reality

Due to the rapid evolution of the technology in this field, different and often contradictory terminology definitions have been used [52, 51]. For the purposes of this thesis the following definitions are used for augmented reality (AR), mixed reality (MR), virtual reality (VR) and extended reality (XR).

- AR: according to the definition by Azuma it has three different properties [3].
  - Combines real and virtual
  - Interactive in real time
  - Registered in 3D
- MR: a view of the physical world with an overlay of digital elements where physical and digital elements can interact.
- VR: a fully-immersive digital environment.
- XR: an umbrella term that covers all these different technologies, including AR, MR, and VR.

This differentiation is visualised in Figure 2.1.

2.2 Head-mounted displays

A head-mounted display (HMD) is a display device worn on the head.

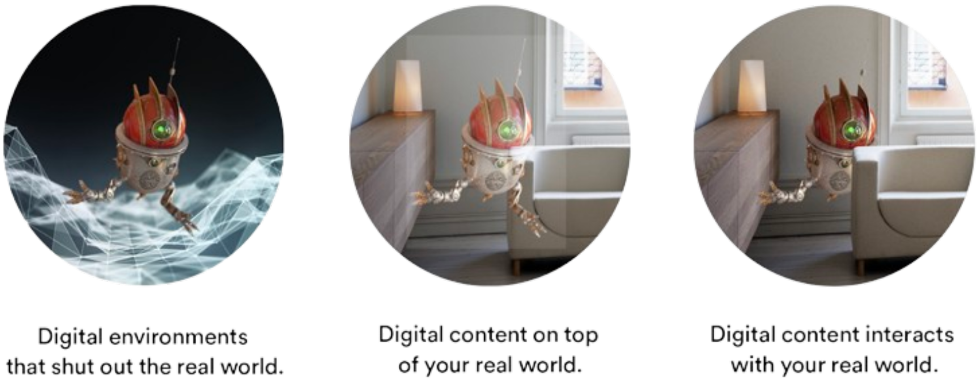


Figure 2.1 Magic Leap's definition for VR (left), AR (middle), and MR (right) [52].

2.2.1 Virtual reality headset

A VR headset is an HMD that uses positional tracking to provide a virtual environment to the user. Some VR headsets utilise VR pass-through technology. VR pass-through uses a combination of cameras and lenses to allow a user of a VR headset to interact with the physical world. If a VR headset has pass-through functionality it can be used for MR views [48]. Commercial examples of VR headsets include the Apple Vision Pro and the Meta Quest 3 depicted in Figure 2.2.

(a) Apple Vision Pro **(b) Meta Quest 3**

Figure 2.2 Two commercial VR headsets.
Source: Wikimedia Commons.

2.2.2 Optical head-mounted display

An optical head-mounted display (OHMD) is a wearable display that can reflect projected images and allows a user to see through it. Commercial examples include

MR headsets like Microsoft HoloLens and AR smart glasses like Google Glass and Meta Ray-Ban Display.

2.2.3 Smart glasses

Smart glasses are wearable computers that add information to the user in some capacity. Not all smart glasses include an OHMD. Those without a display interact with the wearer through other means, typically open-ear audio via built-in speakers and microphones, sometimes complemented by a camera and voice or touch controls. Commercial examples include AR smart glasses like Google Glass, Meta Ray-Ban Display but also tethered MR smart glasses like XREAL One Pro depicted in Figure 2.3. Most smart glasses with a display have AR functionality but not MR functionality.

(a) Google Glass **(b) Meta Ray-Ban Display** **(c) XREAL One Pro**

Figure 2.3 Commercial smart glasses.
Sources: Wikimedia Commons (a), Mrwhosetheboss on YouTube (b), and Tom's Guide (c).

2.3 Methods of attention direction

Senses such as vision, hearing and touch are all a part of an individual's sensory system [18]. The sensory system processes input from the surrounding environment. The sections below introduce the modalities that can be used to direct a person's attention. The cue designs chosen for this thesis, and prior studies that apply and compare these modalities for attention direction and hazard awareness, are discussed further in Chapter 3 (Related Works).

2.3.1 Visual

Visual cues are a common way of directing attention in AR, since the information is rendered directly in the user's field of view. Typical visual cues include arrows that point towards a target, highlights or outlines drawn on the object itself, and overview representations such as a top-down map. Because virtual visual cues are drawn on top of the real world, they compete with it for the user's attention. A cue can hide a safety-critical object behind it, or clutter the view and add to the user's mental workload [41, 39]. Balancing visibility against this risk is a central design consideration.

2.3.2 Spatial audio

Audio can also be used to give a sense of direction. Spatial audio simulates how sound is perceived in 3D space [57]. The spatial impression is created using models such as the head-related transfer function (HRTF), which captures how the sound reaching each ear depends on direction, and the room impulse response (RIR), which captures how sound is shaped by the environment [7, 57].

2.3.3 Tactile

Tactile feedback utilises a person's sense of touch to give cues, typically through vibration motors worn on the body. A key advantage is that it is independent of vision: it does not block the user's field of view and does not add to the visual workload, which is beneficial when the visual channel is already busy or limited [37]. Vibrotactile cues can convey not only the direction to a target but also how near it is, by varying the pulse rate or vibration intensity [37].

How well the skin can tell two nearby points of vibration apart varies across the body. The back, a common placement for wearable tactile displays such as vests, is relatively poor at this, so vibration motors must be spaced fairly far apart to be felt as separate rather than as a single buzz. This ability is also uneven on the back: it is about twice as sharp horizontally as vertically, so motors can be spaced more sparsely in the vertical direction than in the horizontal [38]. The direction of a vibration is also easier to judge for some orientations than others: on the back, the cardinal directions (up, down, left and right) are perceived more accurately than the diagonal ones [25]. These limits inform how many motors are needed and where they should be placed.

Tactile feedback can be used either on its own or as a complement to visual or auditory signals.

2.3.4 Multimodal

Cues from different modalities can also be combined, for example pairing a visual cue with a tactile or auditory alert. Combining modalities in this way has been shown to improve reaction times compared to a single modality [11, 46]. A redundant channel can also convey information without further loading an already busy sense.

2.4 Locating a point in space

A positioning principle is the underlying method used to determine the unknown location of a point from measurements taken relative to references at known positions. There are three different positioning principles: trilateration, triangulation and scene analysis [30]. The three principles can be used in isolation or in combination to create a system to track objects in real-time. A real-time locating system (RTLS) is defined in ISO/IEC 24730-2:2006 as a wireless system capable of determining the position of a transmitting device relative to fixed receivers in a defined space and reporting that position in or close to real time [22]. In this section the major position principles and their commonly used algorithms are described.

2.4.1 Trilateration

Trilateration uses distances to three or more known positions to calculate the unknown position coordinates of a point of interest (POI) [30]. The pre-defined known points are also known as anchor points (AP). As seen in Figure 2.4, trilateration in 2D space can be illustrated with the intersection of three circles from the APs, where the radius of each circle is the measured distance to the POI. Similarly spheres can be used to illustrate trilateration in 3D space. Looking at the illustration in Figure 2.5 there are two points of intersection. If the APs are sufficiently spaced apart, one of the points will always be underground or in space, allowing it to be excluded.

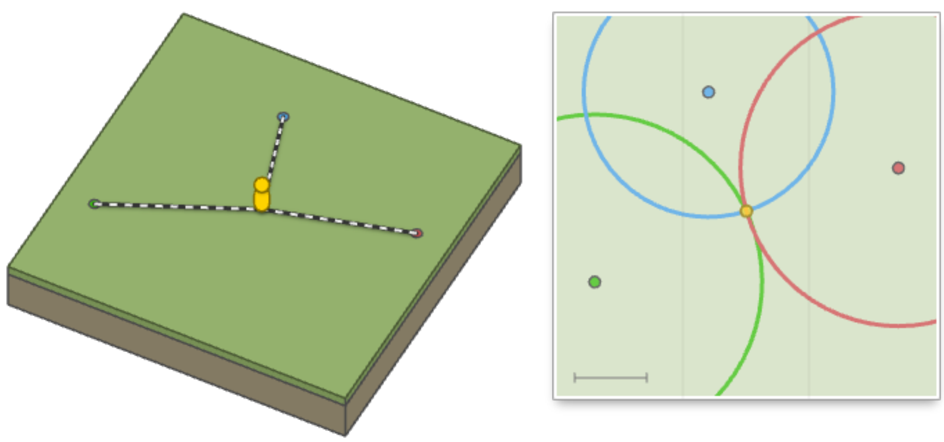


Figure 2.4 Illustration of trilateration using three distance measurements in 2D. Adapted from Bartosz Ciechanowski, *GPS*.

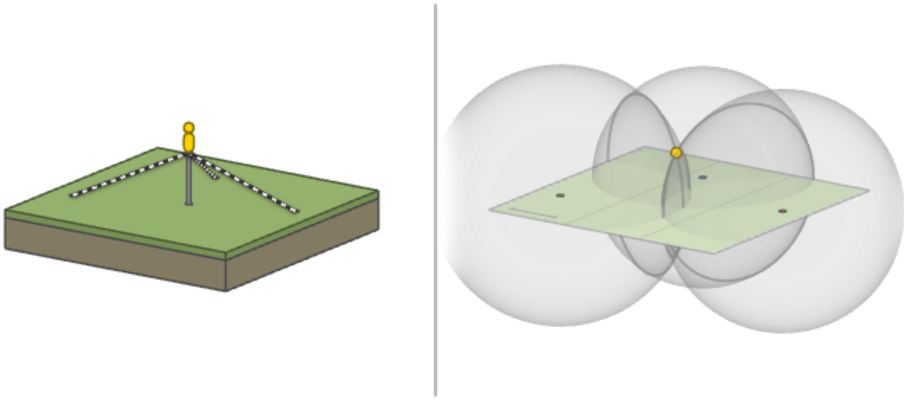


Figure 2.5 Illustration of trilateration using three distance measurements in 3D. Adapted from Bartosz Ciechanowski, *GPS*.

The global navigation satellite system (GNSS) is a popular implementation of trilateration. It is utilised by systems such as GPS and GLONASS and has achieved a position accuracy at the sub-metre level [31]. When used indoors the accuracy of GNSS is poor. It is sufficient for identifying which room an object is in, but not for reliably locating a fast-moving hazard [13].

2.4.1.1 Time of arrival

Time of arrival (ToA) is the simplest algorithm to achieve trilateration in theory. The distance between the POI and APs can be calculated by multiplying the time of flight (ToF) with the speed of the signals. While the calculation is simple in theory, it is impractical in practice because the clocks in the POI and APs have to be synced to the nanosecond. Additionally it is preferred to have a clear line of sight (LOS) to achieve high accuracy. When there is a non-LOS (NLOS) condition the signals may be delayed or blocked making the ToA calculation inaccurate [30].

2.4.1.2 Time difference of arrival

Time difference of arrival (TDoA) is a modified version of ToA [30]. Instead of measuring the absolute time a signal takes to travel from the POI to each AP, TDoA measures the difference in the time at which the same transmission arrives at different APs. Because taking this difference cancels out the unknown instant at which the POI sent the signal, the POI clock no longer needs to be synchronised with the rest of the system. This is the main advantage over ToA, which requires the POI and the APs to share a common time reference. The APs still require highly synchronised clocks with each other, but that is a more practical requirement since they are stationary and can use large and very precise clocks that are continuously monitored to ensure synchronisation. Each measured difference places the POI on a curve of possible positions defined relative to one pair of APs, and combining the differences from several AP pairs yields the position.

Both ToA and TDoA suffer from degraded accuracy in NLOS conditions. Several attempts at mitigating or identifying NLOS effects have been made using statistical corrections and machine learning, yet no universally reliable solution exists. There are methods that effectively mitigate NLOS conditions if the obstacle locations are known. To date mitigating effects of NLOS in dynamic environments remains an open problem [27].

2.4.1.3 Two-way ranging

Two-way ranging (TWR) utilises an exchange between the nodes in the system to calculate the round-trip time (RTT) as illustrated in Figure 2.6. With a TWR exchange the clocks do not need to be synchronised, considerably simplifying hardware requirements [35].

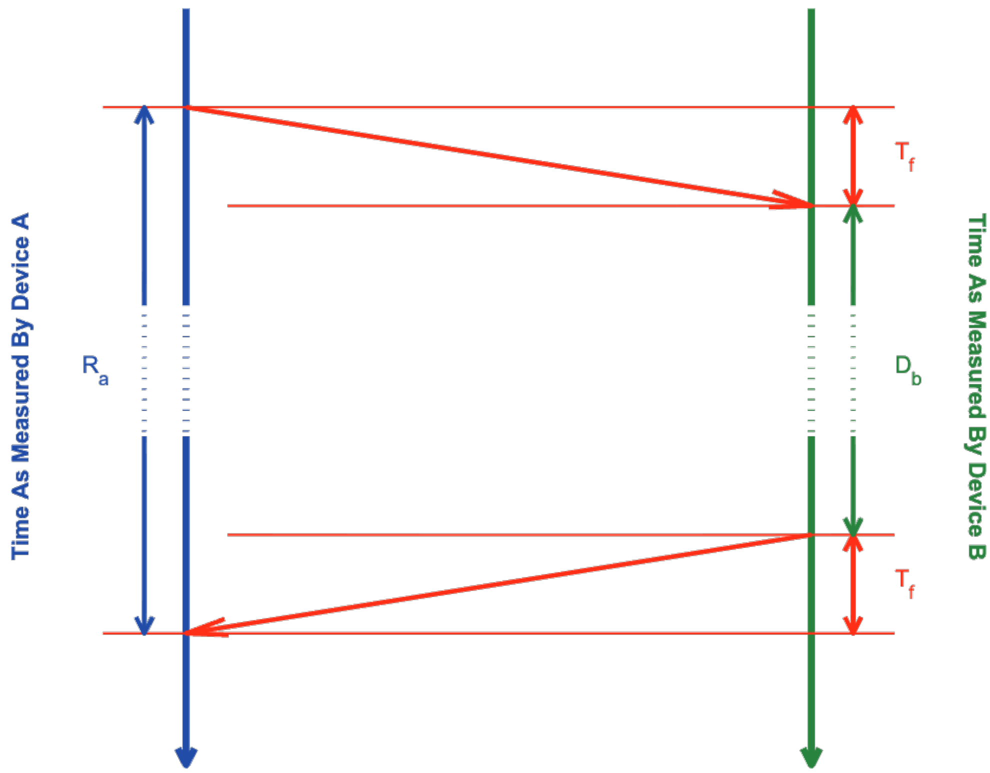


Figure 2.6 Two way ranging exchange.
Adapted from IEEE *An Alternative Double-Sided Two-Way Ranging Method*[35].

2.4.2 Triangulation

Triangulation estimates the position of the POI from angle measurements rather than distances [30]. Each anchor measures the angle of arrival (AoA) of the incoming signal, and the POI is located at the intersection of the resulting bearing lines. Because each bearing constrains the position to a line, two anchors are sufficient to determine a position in 2D. One way to obtain the AoA is to measure the phase difference of arrival (PDoA) of a signal across two or more closely spaced antennas on the same anchor, an approach that has been demonstrated on a commercial UWB module with an average angular accuracy of about 2.4 degrees over a 90-degree range [32].

2.4.3 Position estimation and filtering

The positioning principles above yield, in practice, an over-determined and noisy system of equations that must be solved numerically and then smoothed over time.

Given distance measurements to several anchors with known positions, the POI position is commonly estimated as the point that minimises the sum of squared differences between the measured and the predicted distances. This nonlinear least-squares problem can be solved iteratively by linearising it around a current estimate, an approach introduced by Foy as Taylor-series estimation, which gives a least-sum-squared-error solution to the linearised equations and converges to the correct solution from a reasonable initial estimate [17]. Because individual range measurements can be heavily corrupted (for example under NLOS conditions), a robust estimator such as Random Sample Consensus (RANSAC) can be used to fit the position model while rejecting outliers: it repeatedly estimates the model from random minimal subsets of the measurements and keeps the solution supported by the most inliers [16].

The resulting position estimates are still noisy from sample to sample. A Kalman filter is a recursive estimator that fuses a model of the system's motion with successive noisy measurements to produce a smoothed estimate of the state, and can also be used to predict the state between or ahead of measurements [24, 54]. With a constant-velocity motion model the state typically contains both position and velocity, allowing the filter to smooth the trajectory of a moving tag.

2.5 Positioning technologies

A range of wireless technologies can be used to build a real-time locating system (RTLS). They differ mainly in accuracy, robustness to obstacles and noise, cost, and power consumption, and the most suitable choice is application and environment specific [30]. This section surveys the technologies most commonly used for indoor positioning and motivates the choice of ultra-wideband for this thesis.

2.5.1 Global navigation satellite systems

GNSS such as GPS achieve sub-metre accuracy outdoors, but their signals are blocked or severely attenuated by building structures, so accuracy degrades sharply indoors [31, 13]. Specialised indoor GNSS variants exist, but they require additional infrastructure and still typically reach only sub-metre accuracy [30]. This makes GNSS unsuitable for the centimetre-level indoor tracking needed to flag a fast-moving hazard.

2.5.2 Bluetooth Low Energy

Bluetooth Low Energy (BLE) is attractive for its very low power consumption and low cost, as off-the-shelf beacons are inexpensive and BLE is available on most smartphones [30]. Its positioning accuracy is, however, only at the multi-metre level, typically estimated through signal-strength fingerprinting [30]. BLE-based real-time locating has nonetheless been proposed for proximity warnings that keep workers away from hazards on construction sites [21].

2.5.3 Radio-frequency identification

A radio-frequency identification (RFID) system consists of tags and antennas connected to a host computer [30]. Passive tags are cheap and battery-free but have a short range, while active tags reach considerably further at a higher cost. RFID can reach moderate to high accuracy, and specially designed systems attain centimetre-level accuracy under line-of-sight conditions, but it scales poorly because identifying a tag generally requires a search over the back-end database [30].

2.5.4 Ultrasound

Ultrasound-based systems can achieve very high accuracy, down to the sub-millimetre level, in environments with little noise [30]. Their main drawbacks are that ultrasound cannot penetrate walls and has limited coverage, so many fixed units must be installed, and that accuracy degrades significantly in noisy environments [30]. The latter is a serious limitation for the loud industrial settings targeted in this thesis.

2.5.5 Ultra-wideband

Ultra-wideband (UWB) achieves the best accuracy and precision of the common indoor positioning technologies, reaching the centimetre level or better, and its wide signal bandwidth makes it comparatively robust to the multipath reflections that degrade other radio technologies [30]. It is a wireless technology suitable for short-range communication [56]. The signals emitted from UWB devices are narrow pulses, commonly ranging from 0.2 ns to 1.5 ns [56]. This results in a signal that has a wide frequency spectrum [9, 29]. The wide signal, combined with a specific frequency range being assigned to UWB devices, makes UWB a low risk for interfering with other nearby equipment [9, 56]. The short pulses also make UWB systems more resilient towards physical objects disturbing the signals if anything is located between an anchor and the tag [2, 9]. Another characteristic of UWB devices is the potential for low energy consumption due to the short intermittent signals [2, 56].

A UWB positioning system requires a set of anchors and one or more tags [29]. The anchors are stationary devices installed at fixed, known positions in the environment, while a tag is a mobile device attached to the object that is to be tracked. The anchors and the tag exchange UWB signals to estimate the distance between each anchor and the tag. Since the clocks of the devices are not synchronised, this distance is typically obtained through the two-way ranging exchange described in Section 2.4.1.3, which derives the range from the measured round-trip time.

In principle, distances from three anchors are enough to determine the tag's position in 3D space: the three distances give two possible points, and the impossible one (for example a point outside the room or underground) is simply discarded, as described in Section 2.4.1. In practice additional anchors are used, both to resolve this ambiguity reliably and to over-determine the system so that noisy or faulty measurements can be handled; the position is then computed using the multilateration and filtering described in Section 2.4.3. Because the anchor positions are fixed and known while the tag is free to move, such a system continuously reports the tag's position in close to real time and therefore constitutes a RTLS, as defined in Section 2.4. Throughout the remainder of this thesis, the terms *anchor* and *tag* are used in this sense.

3 Related Works

This chapter presents related work on augmented reality, hazard awareness and attention guidance.

3.1 AR and safety

Multiple studies have investigated the use of AR and MR to improve safety in workplace environments such as mining and industrial settings. Akhmetov and Varol [1] mentioned the risk for accidents in an industrial environment and the current use of safety devices within the sector that aim to make the workplace safer. The authors present a new warning system using smart goggles that provides warnings when a user looks away from their task. Experiments in AR were conducted to compare different kinds of AR warnings. Results showed that participants preferred visual warnings over auditory or audiovisual warnings. They also asked the question if AR systems are better than already existing safety measures for increasing the user's attention on their work task [1].

A study from 2016 [26] implies that smart glasses with AR information could have a positive impact on workplace safety and task performance. Interviews conducted as part of the study did, however, reveal that a majority of participants were worried that wearing an HMD could distract from the work tasks. Kim et al. [26] also mentioned existing evidence that AR in an HMD could improve performance for order picking in warehouses.

A 2019 study investigating wearable augmented reality displays (WARD) concludes that the technology improved situational awareness, self-efficacy, and trust, but also increased the operator's experienced workload [42]. A similar result was reached in another study from 2025 [8]. Camacho-Fidalgo et al. [8] conducted a user study that aimed to examine the use of AR and its potential effect on a user's SA. The study was conducted in a room with tasks placed in different locations to simulate an industrial environment. Participants wore an HMD, receiving remote assistance from a test operator in a different room, while navigating the room during the experiment. The authors concluded that participants wearing AR-HMDs noticed fewer hazards in their environment, something that could negatively impact a worker's safety.

Automation bias could potentially be another side effect of using AR cues to alert workers of surrounding hazards. Raikwar et al. [40] conducted a study where the impact of imperfect AR cueing on visual search was investigated. Their results showed that participants' performance was negatively affected by getting inaccurate cues.

Multiple studies have researched the possibility of using AR or MR to increase safety for pedestrians sharing space with heavy equipment and autonomous machines [4, 46]. Baek and Choi [4] developed a warning system for pedestrians when walking around construction and mining sites. The system used visual warnings displayed in smart glasses to warn the wearer when they were close to potential dangers on the site. Bluetooth was used to determine the wearer's proximity to dangers. Results showed that their warning system based on smart glasses caused less stress than a similar system using a smartphone [4].

Another study discussed the use of smart glasses for reading while walking [20]. They examined how text could best be displayed while still enabling participants to be aware of their real-world surroundings. The results showed that displaying a few lines at a time, while allowing the smart glasses wearer to control when new text appeared, was the most effective option for both user safety and overall user experience. Showing too much text at the same time negatively impacted the wearer's SA [20].

3.2 AR and attention guidance

There are also studies and reviews examining AR and safety outside the context of industrial environments. A review study from 2024 discusses the use of AR visualisation to make the user aware of objects outside of their own field of view (FOV). They conclude that there is a gap in research concerning the use of AR to track physical objects [39]. The review also compares the uses of egocentric and exocentric guidance techniques, including a bird's-eye view and a 3D arrow. The authors of the review theorised that an egocentric technique, such as the arrow, might be more suitable for emergency responses.

Another study researched the use of AR in an HMD to support safe navigation for pedestrians [47]. The study compared a context-adaptive navigation system with an always-on system and the authors concluded that context-adaptive user navigation could be useful in industrial environments.

The use of AR often aims to improve the user experience. Displaying additional virtual information within the user's FOV could, however, introduce clutter that increases the cognitive load. Renner and Pfeiffer [41] conducted a study to investigate the usefulness of different visual AR cues when the user is focusing on something in a visually cluttered real-world environment. Cueing workers in a manual assembly line was one potential related scenario mentioned. Six different types of visual cues were used in the study, one of which was an arrow pointing to the next relevant object in a virtual birdhouse building task. Their results showed that the arrow was preferred above the other cues used, and also gave a quicker task completion time [41]. The arrow also led to a quicker detection time when locating objects outside of the participant's FOV [41]. The authors also presented a new type of AR cue called the spherical wave-based guidance (SWAVE) which produced similar results compared to the arrow.

Biswas et al. [5] conducted another study that also compared an arrow design to other types of AR cues when locating objects. This smartphone-based study compared a 3D arrow to Halo3D and 3Dwedge+ as cues for pointing towards objects off-screen. The authors also discussed the optimal placement of the 3D arrow on the smartphone screen to minimise visual clutter for the user [5].

The design of the AR cues used for this thesis were inspired by the designs in Biswas et al. [5] and Tonnis et al. [50]. The latter is a study from 2005, mentioned in the review by Quinn and Gabbard [39]. Their experiment placed the participants in a driving scenario where cues were projected onto the windscreen to indicate hazards outside of the driver's FOV [50]. They compared an exocentric 2D bird's-eye view with an egocentric 3D arrow. Their results showed that the bird's-eye view performed better, but the authors theorised that limitations in the projection technology affected the clarity of the 3D arrow and caused it to appear flatter [50]. This thesis aims to partly revisit this experiment with newer technology, while also extending it by presenting the cues in an HMD.

Weidinger et al. [53] explored the impact of the choice of cue type and cue density on user experience and task performance by conducting a study in immersive VR.

3.3 Real-time locating system

Accurate tracking is required in order to determine when to display a hazard warning. A real-time locating system (RTLS) enables this by monitoring positions of objects, and by extension, their positions relative to the person receiving the warnings. There are several existing localisation technologies for both outdoor and indoor positioning [2, 30]. Ultra-wideband (UWB) is mentioned in several articles as a suitable technology for indoor positioning [2, 9, 29, 56], partly due to its high ranging accuracy and low energy consumption. Li et al. [29] set up a system aiming to track pedestrians in a real-world environment. They installed 21 UWB anchors in the space to enable tracking along the x, y and z axes [29]. Li et al. [30] lists 10 different positioning technologies suitable for indoor environments and discusses different metrics that can be used to decide which technology to use. Bluetooth Low Energy RTLS has also been proposed for tracking hazards on construction sites [21].

3.4 Additional modalities

The user study conducted for this thesis focuses on visual AR cues. Several studies have, however, examined the use of other sensory modalities. Vibrotactile feedback has been researched for both directional guidance [25, 37, 38] and warning signals for hazard awareness [15, 19]. The use of spatial audio for directing attention has also been explored [7, 10, 34, 46, 57].

There have also been studies that either compared different types of modalities [28] or examined combinations of multiple modalities [11, 46]. San Martin and Kildal [46] conducted a user study that compared single modality to audio-visual cues and concluded that the inclusion of auditory signals improved participants' reaction time to potential hazards.

4 Pre-study

This chapter describes the initial research done before implementing a final AR application and conducting the user study.

4.1 Introduction

The work for the thesis has been conducted based on the Double Diamond design process [12] which consists of the four phases discover, define, develop, and deliver. Each phase has its own purpose. The Double Diamond model was defined by the Design Council in 2003 to illustrate the innovation process, see Figure 4.1. The first diamond, consisting of *discover* and *define*, represents the process of researching and understanding a problem. *Develop* and *deliver* in the second diamond then cover the process of further finding and developing a solution to the problem [12].

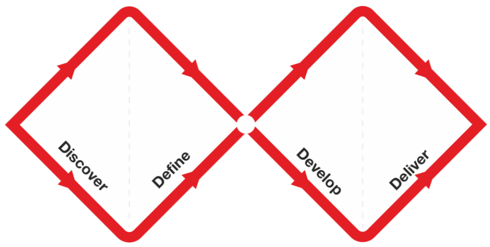


Figure 4.1 The Design Council's visualisation of Double Diamond. Illustrating the phases discover, define, develop, and deliver.

The aim of the thesis was to find a way of supporting and alerting industrial steel mill workers to hazards in their work environment. During the initial research phase, corresponding to discover in the first diamond, emphasis was put on gathering information about different types of existing warning signals. The research focused on compatibility with AR and included signal types such as visual, tactile, and spatial audio. To reduce the scope of the thesis, the user study focused on visual signals. Different technologies for tracking physical objects in a 3D space were also explored.

A prototype of a user study was developed in Unity along with the initial research, see Section 4.2 below. The purpose of the user study was to examine how visual cues can be used to direct a person's attention within their environment. The prototype simulated the experimental setting and helped define the technological needs of a potential real-life study. The insights from the first diamond also led to further investigation of different types of sensory modalities and their usages as signals for attention direction and warnings. As a part of the final phase of *deliver*, a hazard detection and awareness system was implemented as an AR application. The application guided the user's attention towards objects outside their FOV through visual cues indicating their location.

Previous work researched for the thesis had primarily focused on warning participants about virtual objects in VR or AR. An idea emerged of using a real physical object to represent the potential hazard in AR. This approach sought to bring the experimental setup closer to real-life situations where hazards are located in the industrial worker's physical environment. An RTLS would therefore be needed to accurately track the object in the room and identify which direction the cues should point to. The physical hazard would then be the tag in the RTLS. The possibility of using AR to simplify calibration of the RTLS anchors was also explored. The development process leading to the final implementation is illustrated in Figure 4.2.

4.2 First prototype

An initial concept for a user study was developed in Unity alongside the first phase of the research. The aim of the study was to examine the use of visual signals to indicate the presence of hazards in a room. The goal of the prototype was to identify the needs of a real-life study and examine the placement of future AR warning signals within the user's FOV. An initial plan for a real-life study already included the use of an HMD with colour passthrough. The prototype was, however, made in VR and could be tested either with an HMD or in a web browser.

In the scenario used for the prototype participants would be asked to imagine themselves as a steel furnace operator. The scene was set up as a dual-task experiment where their job was to keep the furnace temperature stable (primary task), while also staying alert of potential hazards around them (secondary task). Two different visual cues could be chosen to assist in detecting hazards. There was also an option to run the experiment without any cues.

4.2.1 Primary task

The furnace temperature was shown on a computer screen placed on a desk. The furnace was controlled using the keys 'Q' and 'E'. Pressing 'Q' increased the temperature and pressing 'E' decreased it (Figure 4.3). The temperature gauge on the screen consisted of a small circle moving along a line. A green box around the middle of the line showed the range for an acceptable temperature (Figure 4.3a). The colour of the box changed to red when the circle, and thus the temperature, got too

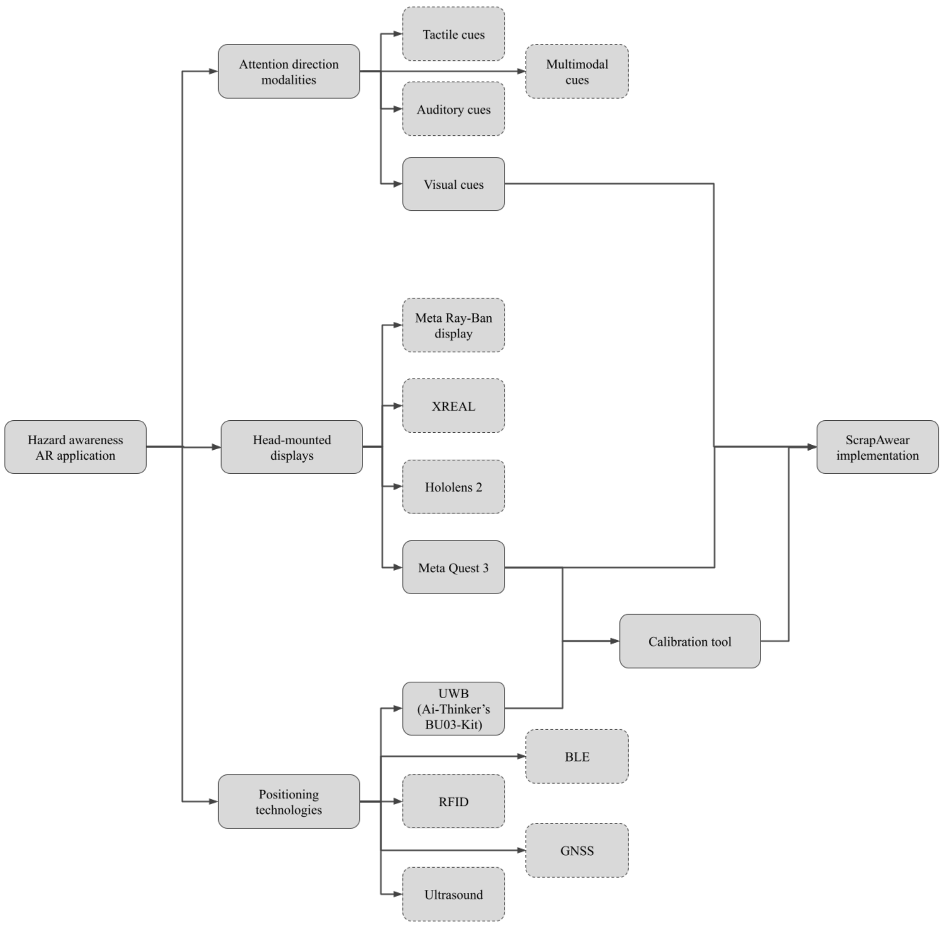
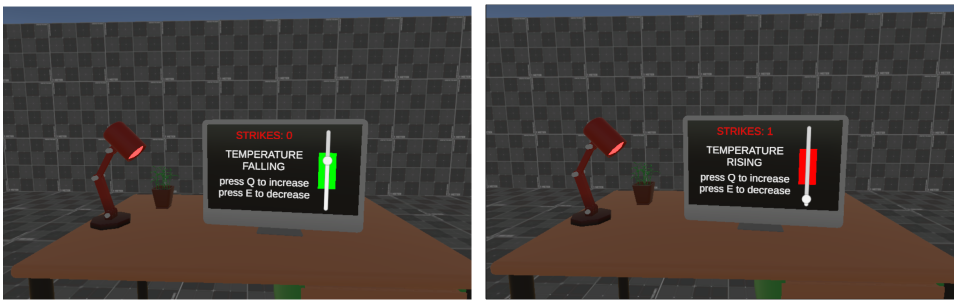


Figure 4.2 ScrapAwear development process.

high or too low (Figure 4.3b). The primary task was to keep the circle within the box. Status text on the screen read either "TEMPERATURE RISING" or "TEMPERATURE FALLING", forcing the circle up towards a higher temperature or down towards a lower one, respectively. The participant would then have to counteract these movements with the keys to keep the circle within the box.

4.2.2 Secondary task

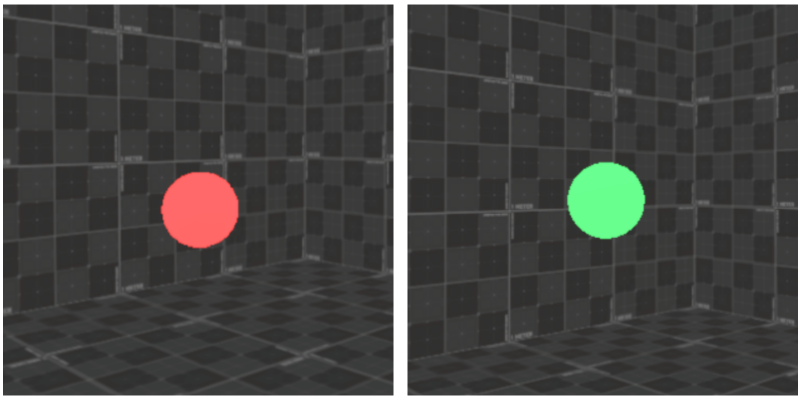
The secondary task of hazard detection was to notice red circular objects that randomly appeared in the room, see Figure 4.4a. Objects were created one at a time, and moved towards the participant's position at the desk. The participant then had to turn and look directly at the object to make it disappear. A created object was first



(a) Temperature falling but stable. **(b) Temperature rising and too low.**

Figure 4.3 Desk setup in the first Unity prototype.

red and would blink, turn green, and then disappear when detected, see Figure 4.4.

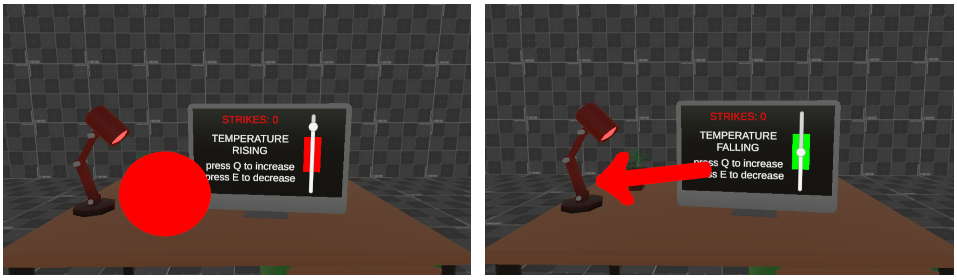


(a) Object red while undetected. **(b) Object green when detected.**

Figure 4.4 The hazard object created and moving towards the user.

4.2.3 Visual warning signals

When starting the prototype, the user could choose between two different visual cues that assisted the detection of the object to varying extents. The cues were placed towards the bottom left of the user's FOV and appeared at the same time as the object was created in the room, signalling that an object was approaching. When the object was detected, both the object and the cue disappeared at the same time. One cue was a red, filled circle that blinked until the object was detected (Figure 4.5a). The circle gave no indication of which direction the object was approaching from. The other visual cue was, however, an arrow that pointed towards the direction of the incoming object (Figure 4.5b).



(a) Circle indicating an object is somewhere in the room.

(b) Pointing towards object

Figure 4.5 Visual cues supporting hazard detection.

4.3 Real-time locating system

Different kinds of technologies were considered for the RTLS, including both ultra-wideband (UWB) and a finished solution from Marvelmind based on ultrasound. UWB was ultimately selected as the base of the RTLS because of the possibility of high precision at a relatively low cost. The system consists of fixed anchors installed in the room that allow tracking of a moveable tag. The exact coordinate placement of the anchors in the room is required for accurate tracking of the tag.

The initial UWB-based RTLS was set up in a small office room. At first a total of five UWB anchors were used. Four anchors were put up in the corners of the room, and the fifth was placed in the middle of one of the walls and connected to the computer running the server. The number of anchors was eventually increased to eight, which improved tag tracking accuracy.

4.4 Head-mounted display

For a real-life deployment in a steel mill setting, Foresight Data Machines expressed a preference for lightweight smart glasses such as the Meta Ray-Ban Display, which would be far more practical for workers to wear over long shifts than a full VR headset. However, at the time of writing, the Meta Ray-Ban Display does not support third-party applications. In a public Q&A, Meta CTO Andrew Bosworth addressed this directly, explaining that “the constraint really is the devices are so tight on kind of compute and thermal space against the battery life. We just don’t feel like we can deliver the user experience that our customers expect. . . ” [6]. Until these hardware constraints are overcome, smart glasses are not a viable target for a prototype that needs to run a custom Unity application.

Other tethered smart glasses such as the XREAL One Pro were also considered. While these devices can display AR content driven by an external host, they lack the persistent spatial anchor infrastructure that this prototype depends on to keep the wearer's head pose aligned with the UWB coordinate system across sessions. This is needed to be able to render the virtual tag at the physical tag's real-world location.

Microsoft HoloLens 2 was also briefly considered, but the Meta Quest 3 was ultimately chosen as the HMD for both the initial prototypes and for the user study. Its built-in depth sensor and inside-out tracking cameras provide robust spatial understanding, and Meta's SDK offers mature support for spatial anchors, which are essential for persisting virtual UWB anchor positions between sessions (described further in Section 5.5.4). The Quest 3 is also significantly less expensive than alternatives such as the Apple Vision Pro, making it a practical choice for a research prototype intended to be replicable.

4.5 Measuring and calibrating in AR

An AR prototype for measuring distance between objects was developed in Unity to evaluate the precision of the virtual measurements. The Unity application displayed the distance between two virtual objects. The measurement was tracked and displayed in the AR environment, see Figure 4.6. The displayed measurements corresponded closely to real-world distances, which led to the idea of using the virtual objects to measure the placements of the physical UWB anchors in the RTLS.

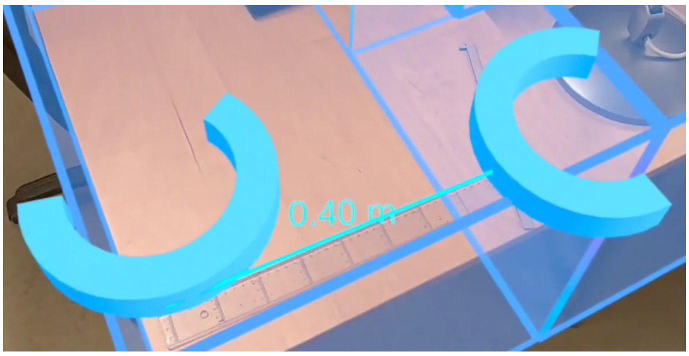


Figure 4.6 Two virtual objects on top of a 40 cm ruler, showing the virtual distance between the centre of the objects.

The authors selected the grab interaction *grab relative to hand* for the virtual menus and anchors. This interaction allowed easy manipulation of the objects from a distance and enabled accurate placement of the virtual anchors. The authors also found that using controllers with the grab interaction provided higher precision than using hand tracking to move virtual objects. Controllers could, however, be less practical in industrial settings where workers need to keep their hands free to perform tasks.

5 ScrapAwear Implementation

This chapter describes the design and implementation of ScrapAwear, the augmented reality and ultra-wideband positioning system built for the user study, covering its hardware, server, Unity application, and supporting web interfaces.

5.1 System overview

The system consists of three main components: a Python server that handles UWB positioning and experiment coordination, a Unity application running on a Meta Quest 3 headset, and web-based interfaces for monitoring and control (Figure 5.1).

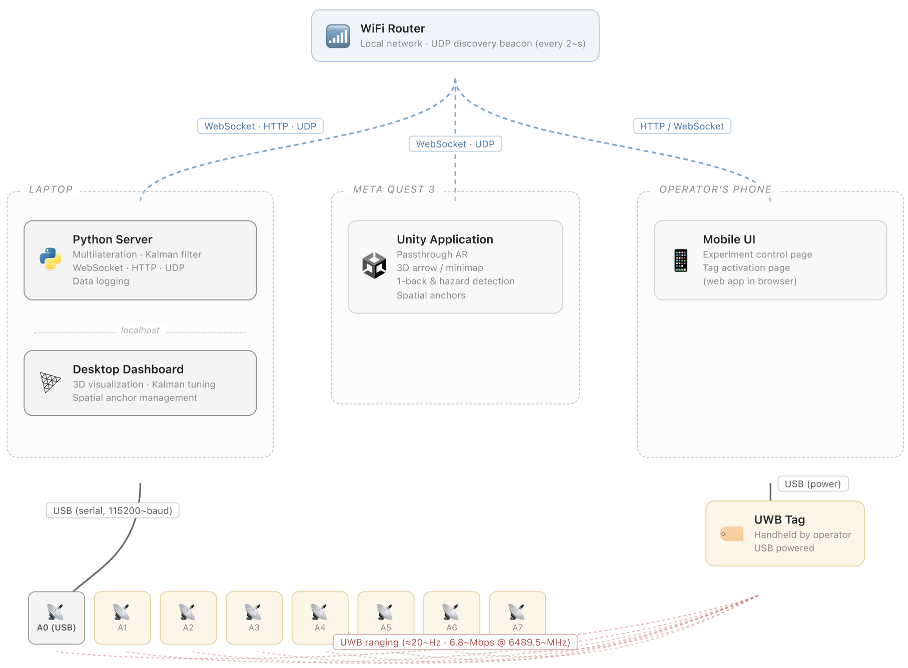


Figure 5.1 Network topology of the experiment, showing all components, the protocols used between them (WebSocket, HTTP, UDP discovery beacon), the wireless Wi-Fi links, the USB connections for serial communication and tag power, and the UWB ranging links between the tag and the anchors.

The positioning server was built on top of an open-source UWB positioning project by DipFlip [14], which provided the core serial communication, multilateration, and Kalman filtering. This was extended with WebSocket-based real-time communication, multi-tag support, spatial anchor integration, experiment task coordination, and web-based control interfaces.

The system operates over a local Wi-Fi network. A laptop runs the Python server, which is connected via USB to one of the UWB anchors for serial communication. The Meta Quest 3 headset connects to the server over Wi-Fi and receives real-time position updates. During test sessions, a mobile phone interface allows the test operator to control tag activation, while a desktop dashboard on the laptop provides real-time monitoring of the positioning system and experiment state. The network topology of the experiment, including the components, the protocols used between them, and the physical (USB and UWB) and wireless (Wi-Fi) links that connect them, is shown in Figure 5.1.

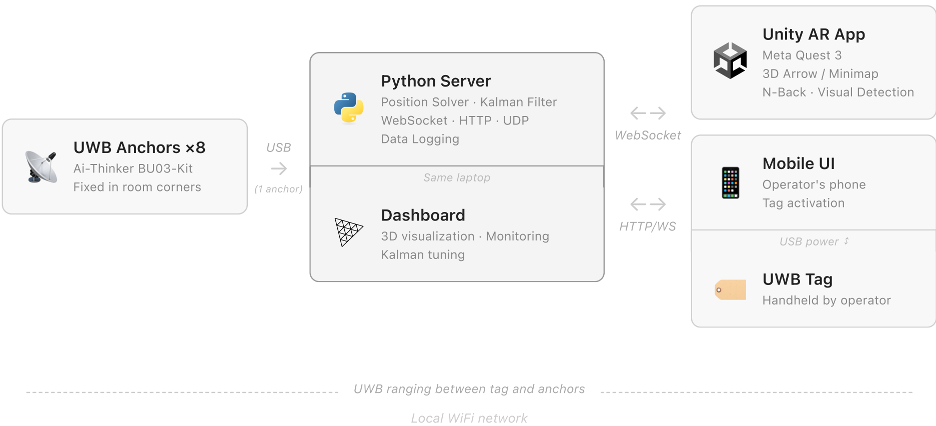


Figure 5.2 System architecture. The Python server receives UWB distance data via USB from one of the anchors and broadcasts positions over Wi-Fi to the headset and web interfaces. The UWB tag is powered by the operator's phone via USB.

5.2 UWB hardware

The RTLS was implemented using Ai-Thinker's BU03-Kit, an inexpensive UWB development board that supports bidirectional ranging with both time difference of arrival (TDoA) and phase difference of arrival (PDoA) capabilities. The modules operate on Channel 5 at 6489.5 MHz with a data rate of 6.8 Mbps. Each module can be configured as either a tag (mobile transmitter) or an anchor (fixed base station). Although the configuration tool exposes device IDs up to 10, the BU03-Kit

Figure 5.3 The experiment room. Three of the four UWB anchors positioned on this wall are circled. The fourth, in the lower left corner, is hidden behind a table.

firmware only reports ranging data for anchor IDs 0–7, limiting the system to eight simultaneous anchors.

The HMD used in the experiment is the Meta Quest 3, running in video passthrough mode. Virtual objects are overlaid onto the camera feed, including the AR navigation cues and a virtual representation of the UWB-tracked tag.

5.2.1 Device configuration

A web-based configuration tool was developed to simplify the setup of UWB modules. The tool is a React application that uses the Web Serial API in Chrome to communicate with the BU03-Kit over USB, allowing configuration of device ID, role (tag or anchor), channel, and data rate. In practice it was used on an Android phone connected to the UWB module via USB On-The-Go (OTG), a USB specification that lets the phone act as a USB host and communicate directly with attached peripherals. This removed the need for a laptop during setup, making it convenient to reconfigure anchors and tags in the field. An AT command console is also available for sending raw AT commands directly to the device.

5.2.2 Development of anchor placement

The current anchor layout was arrived at incrementally during development. Initial tests used five UWB anchors (A0–A4) placed in or near each of the room's four corners, with an additional anchor near the middle of one wall (Figure 5.5). Tracking accuracy with this setup was noticeably worse in height than in the horizontal plane.

Three additional anchors were added near ceiling level, resulting in the eight-anchor configuration used in the user study (Figure 5.6). Adding ceiling-level anchors sig-

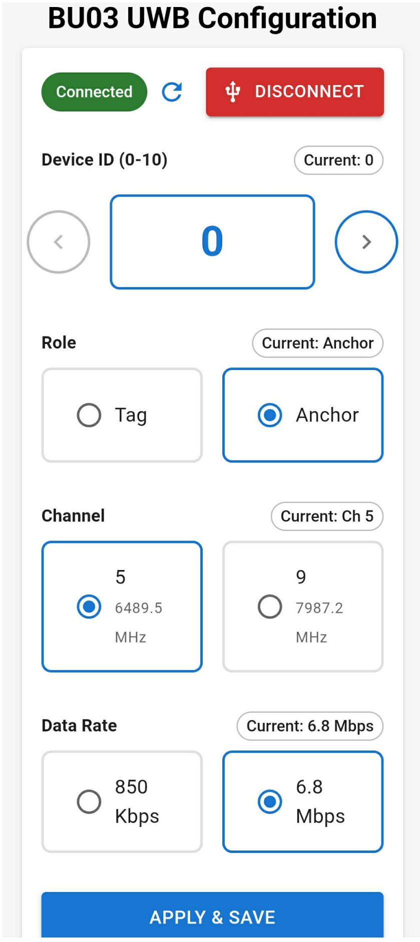


Figure 5.4 The web-based configuration tool, running on a phone connected to a UWB module via USB OTG.

nificantly improved height accuracy.

5.3 Positioning

The positioning pipeline transforms raw UWB distance measurements into smoothed 3D coordinates suitable for real-time AR visualisation (Figure 5.7).

5.3.1 Serial communication

The UWB modules transmit distance measurements as binary packets over a serial connection at 115200 baud. Each packet contains a tag ID and distances to up to eight anchors, encoded as 16-bit little-endian unsigned integers. Values outside

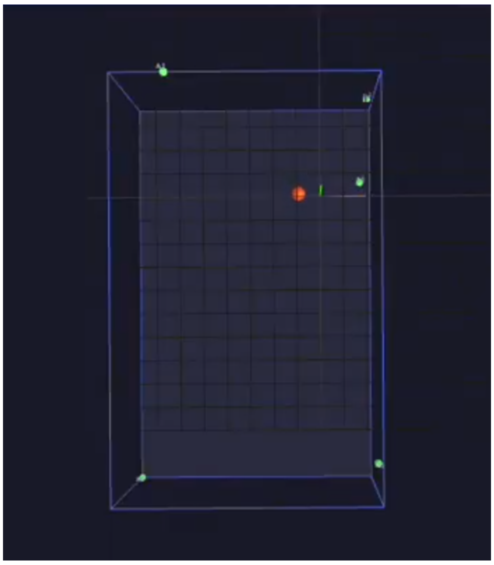


Figure 5.5 Dashboard showing five UWB anchors and one tag in a room.

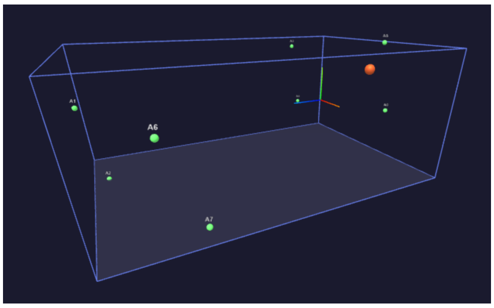


Figure 5.6 Dashboard showing eight UWB anchors (A0–A7) and one tag in a room.

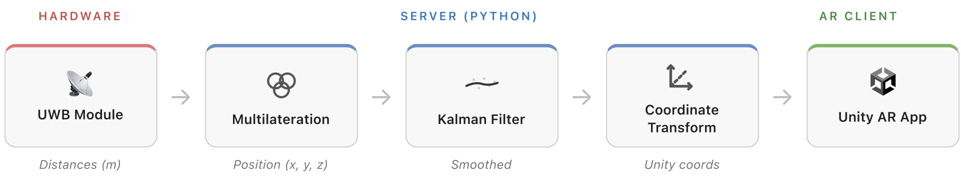


Figure 5.7 The positioning pipeline, from raw UWB distance measurements through multilateration, Kalman filtering, and coordinate transformation to the AR client.

the plausible range of 0.1–10 m are discarded as invalid. Valid measurements are converted to metres and passed through the per-anchor linear calibration model in Equation 5.1 [14], where a is a scale factor and b is an offset in millimetres, fitted during the calibration process described in Section 5.5.4.

$$d_{\text{calibrated}} = a \cdot d_{\text{raw}} + \frac{b}{1000} \tag{5.1}$$

5.3.2 Multilateration

Given calibrated distances from n anchors with known positions, the tag position is estimated using multilateration. Following the least-squares position-location approach [17], the system solves the nonlinear least-squares optimisation problem in Equation 5.2, minimising the sum of squared differences between measured and calculated distances, where $\mathbf{p}$ is the estimated tag position, $\mathbf{a}_i$ is the position of anchor i, and d_i is the calibrated distance to anchor i.

$$\min_{\mathbf{p}} \sum_{i=1}^{n} (\|\mathbf{p} - \mathbf{a}_i\| - d_i)^2 \tag{5.2}$$

A minimum of four anchors is required for a valid 3D solution. When the residual error exceeds 0.5 m, an outlier rejection step based on RANSAC [16] is applied, which tests random subsets of anchors over 10 iterations and selects the solution with the most inliers.

5.3.3 Kalman filtering

The raw multilateration output is smoothed using a linear Kalman filter [24, 54] with a six-dimensional state vector $\mathbf{x} = [x, y, z, v_x, v_y, v_z]^T$ representing position and velocity. The filter uses a constant-velocity motion model with the state transition matrix given in Equation 5.3, where Δt is the time between measurements.

$$\mathbf{F} = \begin{bmatrix} \mathbf{I}_3 & \Delta t \cdot \mathbf{I}_3 \\ \mathbf{0}_3 & \mathbf{I}_3 \end{bmatrix} \tag{5.3}$$

The process noise (default 0.002) and measurement noise (default 0.09) are configurable at runtime through the dashboard interface, allowing real-time tuning during setup.

5.3.4 Coordinate transformation

The server and Unity use different coordinate conventions. The server is Z-up: the x and y axes span the floor plane (right and forward) and z represents height. Unity, by contrast, is Y-up: x is right, y is height, and z is forward. Converting a server position $[x, y, z]$ to Unity therefore requires swapping the y and z axes, producing $[x, z, y]$ in Unity's convention.

In addition to this axis swap, the Unity application uses a spatial anchor as a persistent room reference point. The spatial anchor's position and rotation define a transform that maps between the UWB coordinate system (defined by the physical anchor positions in the room) and Unity's world space. Combined with the axis swap, this transform ensures that virtual objects render at the correct physical locations as the wearer moves around the room.

5.4 Server

The server is implemented in Python as a multi-threaded application with four main components: a serial reader that continuously parses UWB packets, an HTTP server that serves web interfaces and handles REST API requests, a WebSocket server for real-time bidirectional communication with the Unity application and web clients, and a UDP discovery beacon that broadcasts the server's address every two seconds to enable automatic client discovery.

5.4.1 WebSocket communication

The WebSocket server broadcasts position updates to all connected clients at the device's measurement rate (approximately 20 Hz). Communication is bidirectional: the server sends position data, anchor status, and settings updates, while clients send back task results, head pose data, and spatial anchor commands.

5.4.2 HTTP API

The HTTP server provides REST endpoints for system configuration and calibration, including endpoints for reading and updating experiment settings, managing the distance calibration process, persisting anchor positions, and managing the spatial anchor registry.

5.4.3 App settings synchronisation

A shared settings object coordinates the state of the entire system. The settings include the active navigation mode (3D arrow or minimap), experiment name, participant identifier, and flags for calibration mode, tag activation, and experiment state. Settings are persisted to disk and propagated to all connected clients via WebSocket whenever a change occurs. During an active experiment, navigation mode changes are blocked to prevent accidental mid-trial configuration changes.

5.4.4 Data logging

Experiment data is logged so that participant performance and system behaviour can be analysed after each session, both for evaluating the navigation cues against the study's metrics and for verifying that the positioning system performed as expected during each trial. The data is written to CSV files organised by participant and navigation mode. Two types of data are recorded: 1-back task results (trial number, letter, match status, response, reaction time, correctness, block type) and hazard detection events (trial number, activation timestamp, detection time, tag position, camera position, camera forward vector). Data is logged both on the server and locally on the headset to provide redundancy.

5.5 Unity application

The Unity application was developed in Unity 6000.3.7f1 using the Meta XR SDK (version 85.0.0) and the Universal Render Pipeline. The application runs on the Meta Quest 3 in passthrough mode, overlaying virtual elements onto the real environment.

5.5.1 Networking

The application connects to the server using three communication channels. A WebSocket client handles real-time position updates and settings synchronisation, with automatic reconnection on a 2-second interval. A UDP listener receives discovery beacons from the server, enabling automatic connection without manual IP configuration. An HTTP client handles calibration-related API calls.

5.5.2 Spatial anchors

Spatial anchors provide persistence of virtual object positions between sessions. Each virtual UWB anchor is backed by a Meta spatial anchor (`OVRSpatialAnchor`), which maintains its position relative to the physical environment across application restarts. When a spatial anchor's position or rotation changes beyond a small threshold, the updated transform is sent to the server, keeping the UWB coordinate system aligned with the physical room. Virtual anchors are placed directly on top of the physical UWB transceivers, as shown in Figure 5.8.

5.5.3 Anchor management

A virtual menu in the Unity application facilitates the creation and management of UWB anchors. The menu allows the user to create new virtual anchors and assign them an ID, and displays each anchor's distance to the tag (Figure 5.9).

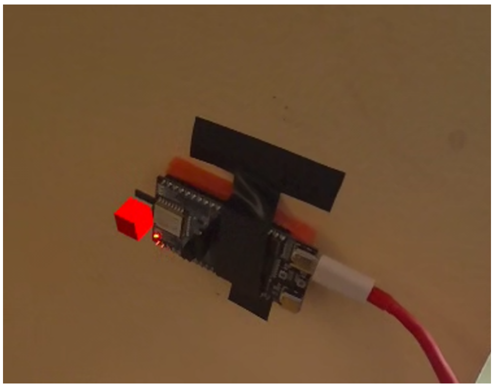


Figure 5.8 UWB transceiver taped to a wall with a virtual anchor, the red cube, placed in AR on top of the physical antenna.

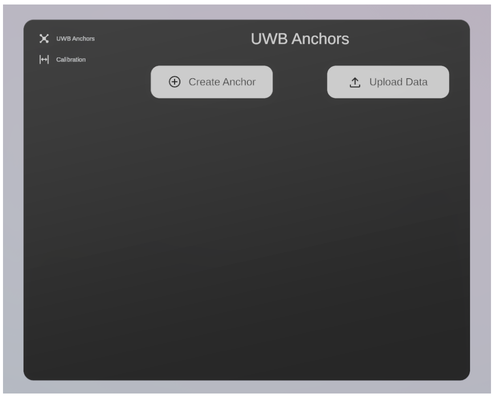


Figure 5.9 Menu panel for adding anchors and uploading data to the server.

5.5.4 Calibration system

Raw UWB distance measurements deviate from the true distance due to per-anchor factors such as antenna mounting, cable lengths, and small differences between hardware units, as well as multipath and signal-propagation effects in the room. Without correction, these biases propagate through the multilateration step and shift the estimated tag position, so each anchor needs an individual calibration before the system can be used. The calibration system in ScrapAwear is implemented as an in-headset tool inside the Unity application that lets the researcher gather the ground-truth distances needed to fit the per-anchor calibration model in Equation 5.1.

The tool requires three different measurements recorded from three different places in the room. The user wearing the headset moves around the room and adds measurements by pressing a button. Each time the button is pressed, the current distance from each anchor to the tip of the user's left index finger is recorded alongside the UWB-measured distances. A line renderer visualises the measurement between the finger and each anchor in real time. After measurements from three locations have been gathered, the server fits a linear calibration model per anchor by comparing the AR-measured true distances with the UWB-reported distances. This approach calibrates all anchors simultaneously using the headset's spatial tracking as ground truth, removing the need for manual measurements with tape measures or laser rangefinders.

Because the calibration relies on the headset's spatial tracking as the distance reference, the accuracy of that reference was first checked against a known length. The AR-measured distance between two virtual markers placed on a 40 cm ruler corresponded closely to the real length (Figure 4.6), supporting the use of AR measurements as ground truth at the scale of the test room.

To quantify the effect of the calibration, the raw UWB range to each anchor was compared with a reference distance for the same anchor, measured by the Unity application using the Meta Quest 3 spatial anchors and treated here as the ground truth. The comparison uses one calibration session of seven reference measurements per anchor. Figure 5.10 plots the raw UWB range against this AR spatial-anchor reference for all eight anchors. The points follow the identity line closely but with a consistent per-anchor spread, indicating that the dominant error is a per-anchor linear bias well suited to the correction in Equation 5.1. Across all anchors the raw ranging error had a mean absolute error (MAE) of 21.7 cm and a root-mean-square error (RMSE) of 26.8 cm, with a near-zero overall bias of +0.3 cm, showing that the deviations are largely per-anchor offsets that cancel in aggregate rather than a single global scale error.

Table 5.1 reports the fitted per-anchor scale a and offset b, the coefficient of determination R^2 of each fit (where 1 indicates a perfect fit), and the ranging error before and after applying the calibration. The linear model fits the measurements well for every anchor (R^2 between 0.955 and 0.997), and applying it reduces the aggregate ranging error from 21.7 cm to 9.8 cm MAE (12.5 cm RMSE), where the ranging error is the difference between the UWB range and the AR spatial-anchor reference. The reduction for each individual anchor is shown in Figure 5.11. These residuals are computed on the calibration measurements themselves and therefore characterise the quality of the fit rather than an independent held-out test, but together with the high R^2 they show that a simple per-anchor linear correction captures most of the systematic ranging error. The ground truth here is the headset's own spatial tracking from a single session rather than an external metrology reference, so the values are indicative of the achievable ranging accuracy rather than a formal accuracy specification. This use of the world-locked AR system to measure and correct the UWB ranging is itself an example of using AR to evaluate a real-time locating system.

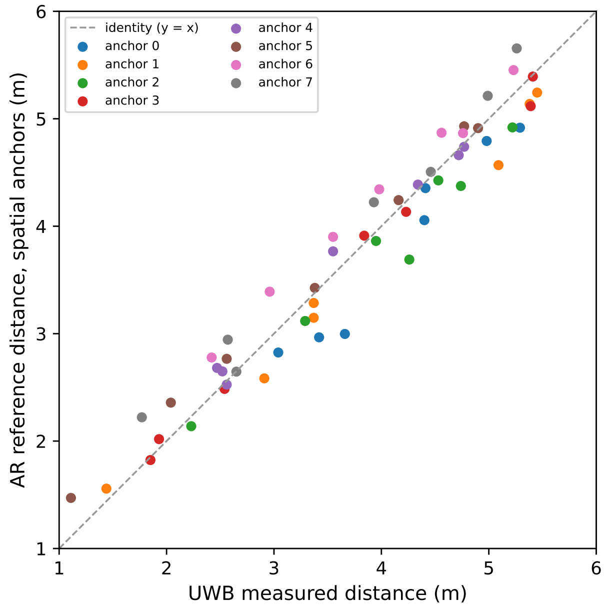


Figure 5.10 **Raw UWB measured distance against the AR reference distance (measured in Unity using the Meta Quest 3 spatial anchors) for all eight anchors, from one calibration session. The dashed line is the identity $y = x$.**

Table 5.1 **Per-anchor UWB ranging calibration from one calibration session (seven reference measurements per anchor). a and b are the fitted scale and offset in Equation 5.1, and MAE is the mean absolute ranging error (UWB range against the AR spatial-anchor reference) before and after calibration.**

Anchor	a	b (mm)	R^2	MAE raw (cm)	MAE cal. (cm)
0	1.064	−594.7	0.955	32.8	15.1
1	0.914	117.8	0.988	24.6	12.9
2	0.917	92.3	0.970	24.2	10.7
3	0.956	112.9	0.995	9.0	8.3
4	0.940	281.5	0.990	10.4	7.1
5	0.921	428.8	0.997	16.9	5.1
6	0.921	615.4	0.994	30.6	5.7
7	0.980	326.5	0.983	25.5	13.6
All				**21.7**	**9.8**

5.6 AR navigation cues

Two AR cues were implemented to guide users' attention toward objects outside their field of view. The cues represent two fundamentally different approaches to spatial guidance: the 3D arrow provides direct egocentric direction ("turn this way"),

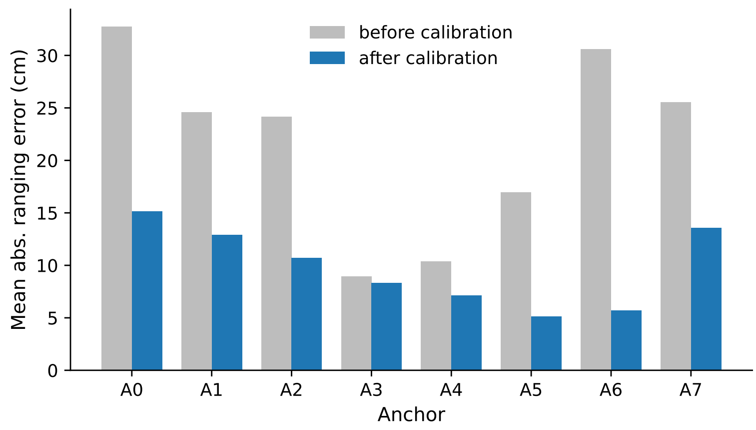


Figure 5.11 Mean absolute UWB ranging error for each anchor before and after applying the per-anchor calibration, measured against the AR spatial-anchor reference from the same calibration session.

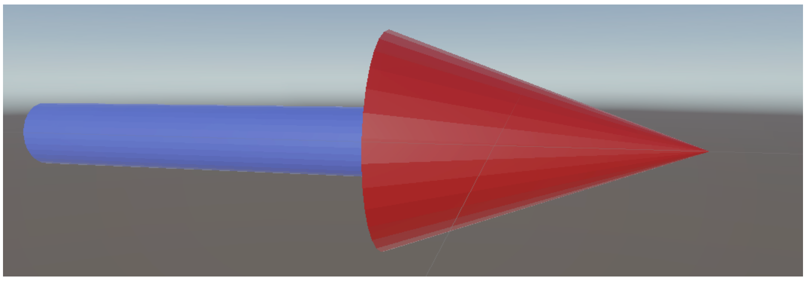
while the minimap provides an allocentric spatial overview ("the object is over there relative to you").

5.6.1 3D arrow

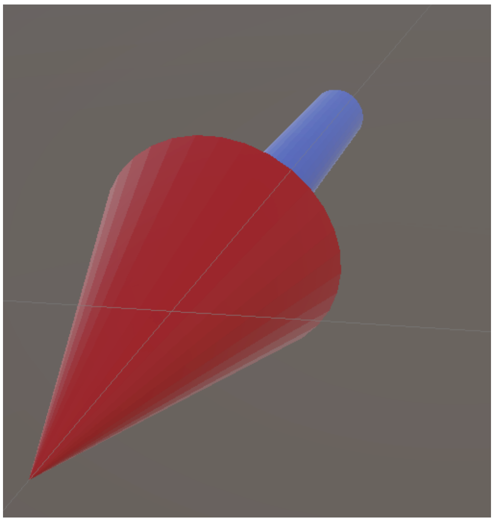
The 3D arrow is implemented as a Unity GameObject that points directly at the target object every frame. Because the arrow rotates in full 3D, it conveys both horizontal direction and vertical offset, pointing downward for low objects and upward for high ones. The arrow is positioned in the user's view, so that seeing the arrow and following its direction is sufficient to locate the target. Figure 5.12 shows the arrow pointing to the right and backwards.

5.6.2 Minimap

The minimap provides a bird's-eye view of the room, displayed as an overlay in the user's view. It operates in egocentric mode, meaning the map rotates with the user's heading so that "up" on the map always corresponds to the direction the user is facing. World positions are projected onto the minimap using a scale factor derived from a configurable world radius (default 5 m). A triangle icon represents the user and a circle represents the target object (Figure 5.13). The minimap displays only horizontal position and provides no vertical (height) information, meaning users must infer whether the target is above or below them from other cues.



(a) Pointing to the right.



(b) Pointing backwards.

Figure 5.12 The 3D arrow cue, seen through the headset's passthrough view, guiding the user towards a target outside their field of view.

5.7 Task implementations

Two tasks were implemented in the Unity application: the 1-back letter task (primary task) and the hazard detection task (secondary task). This section describes their implementation. The procedural details of how participants performed them are covered in the User Study chapter.

5.7.1 1-back task

Letters are drawn from a set of eight visually distinct consonants (C, H, K, L, Q, R, S, T). Each stimulus is shown for a maximum of 2.5 seconds and each trial has a 30% probability of being a match with the previous letter, determined independently. The sequence is structured as an optional practice block of 10 trials followed by an

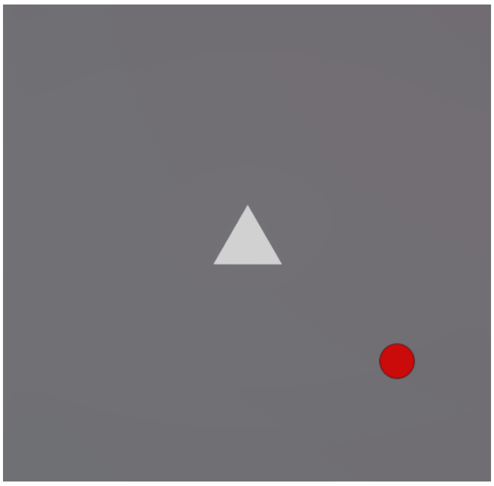


Figure 5.13 Bird's-eye view of the room. The triangle represents the user wearing the HMD and the circle is the moving object in their environment.

experiment block of 50 trials. Responses are captured via the A button on the Meta Quest 3 controller. Visual feedback is rendered by tinting the letter display: green for a correct response and red for an incorrect response or a missed match.

5.7.2 Hazard detection task

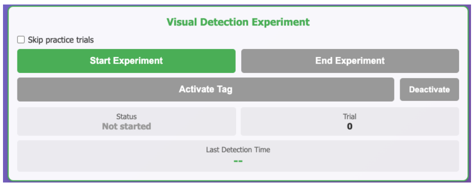
A virtual hazard is rendered at the UWB-tracked tag position. When the task is active the hazard is tinted red. Detection is triggered when the hazard falls within the central 60% of the headset's viewport, at which point it is tinted green and the active AR cue is hidden. Activation timing and spatial placement are configured from the server according to the schedules described in the User Study chapter.

5.8 Web interfaces

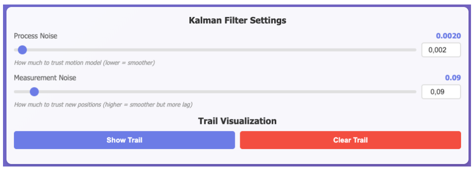
Two web interfaces support the system during operation: a desktop dashboard, used by the researcher to monitor the positioning system and control the experiment, and a mobile interface, used by the tag operator to activate the UWB tag during test sessions.

5.8.1 Desktop dashboard

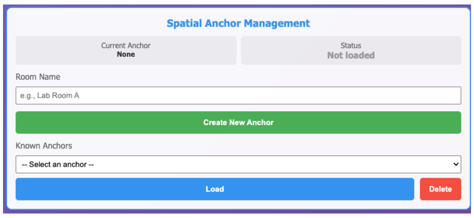
The desktop dashboard provides a real-time overview of the positioning system. A 3D visualisation built with Three.js renders anchor positions, tag trajectories, and connection status. Chart.js graphs display per-anchor distance measurements. The dashboard also includes controls for tuning the Kalman filter parameters (process noise and measurement noise), managing spatial anchors, and controlling the experiment state.



(a) Experiment controls.



(b) Kalman filter tuning.



(c) Spatial anchor management.

(d) Per-anchor distances.

Figure 5.14 Panels of the desktop dashboard, used by the researcher for monitoring the positioning system and controlling the experiment.

A separate panel on the dashboard configures the runtime behaviour of the Unity application (Figure 5.15). In addition to the navigation mode, it controls the in-headset placement of the 1-back letter panel (position and rotation, expressed in the spatial anchor's coordinate frame) and a calibration mode toggle. Calibration mode reveals Unity's main menu and anchor visualisations, allowing the researcher to verify spatial alignment and fine-tune the 1-back panel placement between participants.

5.8.2 Mobile interface

A mobile-optimised web interface was developed for use by the tag operator during sessions. The interface provides two main views: an experiment control page for selecting the navigation mode and starting/stopping the experiment, and a tag activation page with large touch targets for activating and deactivating the UWB tag at the scheduled times.

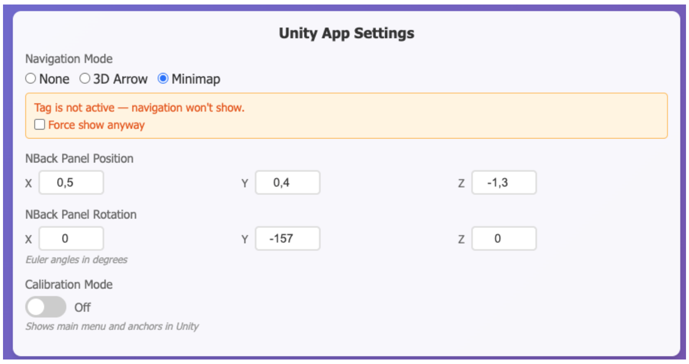


Figure 5.15 **Unity app settings panel, controlling the navigation mode, the position and rotation of the in-headset 1-back letter panel, and a calibration mode toggle.**

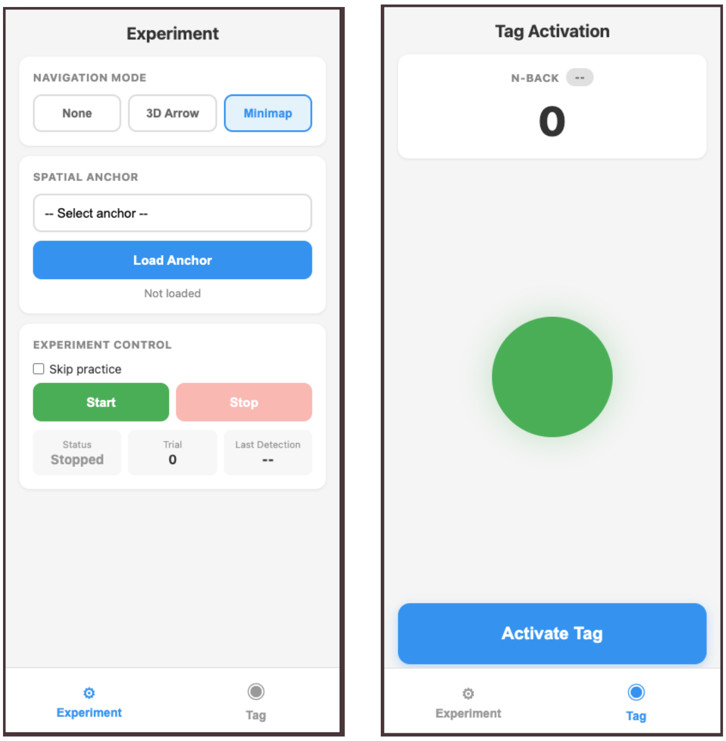


Figure 5.16 **The mobile interface, used by the tag operator during test sessions. Left: experiment control page. Right: tag activation page.**

5.9 Additional capabilities

Beyond the functionality used directly in the user study, the system also supports tracking several UWB tags at once and managing the headset remotely during test sessions.

5.9.1 Multi-tag tracking

The system supports tracking multiple UWB tags simultaneously, as shown in Figure 5.17. This capability was implemented but not included in the user study, as it introduced unresolved design questions, such as how a 3D arrow should indicate multiple targets. Additionally, operating multiple tags during a trial would have required more personnel than were available.

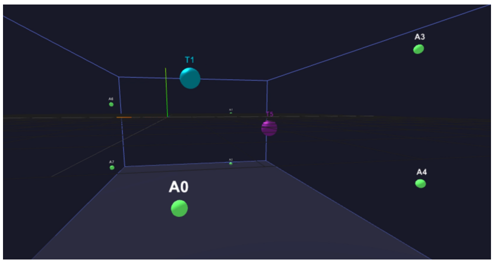


Figure 5.17 Dashboard showing two UWB tags being tracked simultaneously in the room.

5.9.2 Remote headset management

During the user study, the Meta Quest 3 was managed remotely from the researcher's laptop using Meta Quest Developer Hub (MQDH), connected wirelessly via ADB over Wi-Fi. MQDH provides live streaming of the headset's view and a screen recording feature, which was used to capture each test session for later review.

A set of custom ADB shell commands was registered in MQDH to streamline session setup between participants:

- **Start / Stop Unity App**: launches or force-stops the ScrapAwear Unity application.
- **Start / Stop First Hand**: launches or force-stops Meta's *First Hand* demo, which was used as a short AR familiarisation experience at the beginning of each session.
- **Clear First Hand data**: clears the *First Hand* app's persistent data so that each participant experiences the demo from a clean state.
- **Re-grant scene permission**: re-grants the `USE_SCENE` permission that *First Hand* requires, which is revoked when app data is cleared.

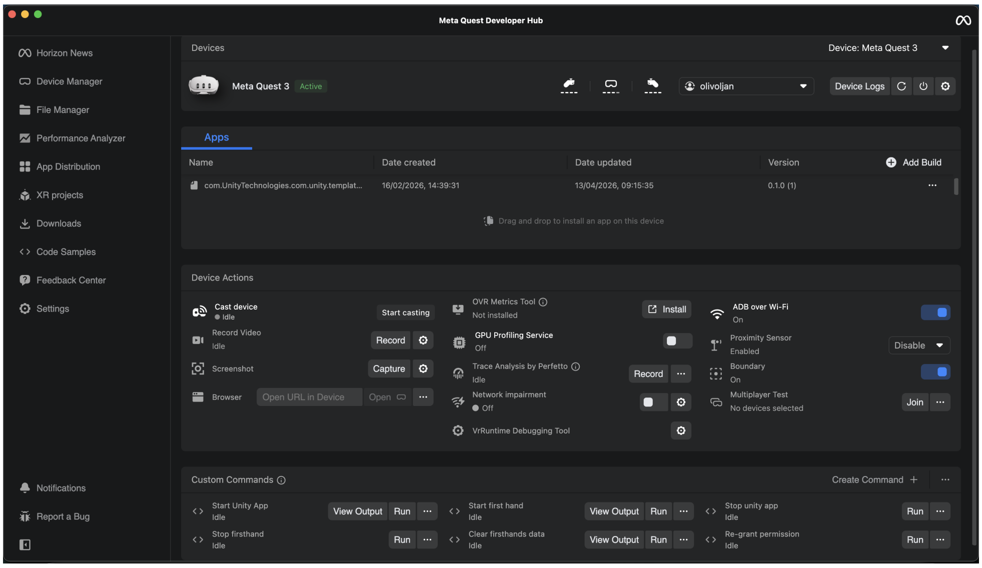


Figure 5.18 Meta Quest Developer Hub, used to stream and record the headset's view and to launch experiment applications via custom ADB commands. The custom commands are visible in the lower panel, and the ADB over Wi-Fi connection is enabled in the Device Actions section.

6 User Study

This section outlines the user study comparing the effectiveness of two different visual cues in augmented reality.

6.1 Introduction

This user study was conducted to investigate whether AR cues could support the detection of physical hazards located outside a person's FOV. Industrial environments require workers to maintain situational awareness while performing their tasks, making hazards outside their FOV a potential concern. The study setup built on the UWB-based RTLS developed in the previous stages of the thesis (see Section 5.3) to track a physical tag representing the hazard. Participants wore an HMD displaying the AR cues and a virtual object that tracked the physical UWB tag (Figure 6.1). The physical tag together with the virtual tag represented a hazard in the room.

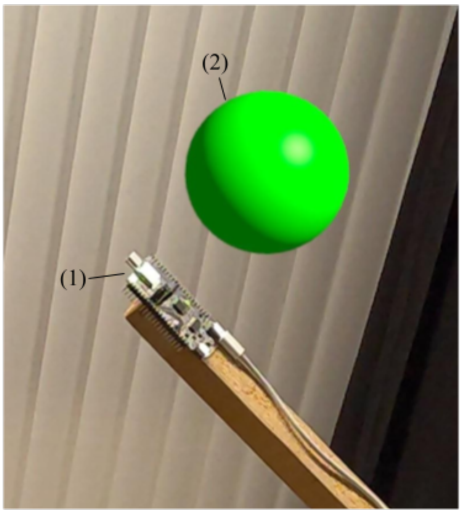


Figure 6.1 View of the tag through the Meta Quest 3, with passthrough. (1) physical UWB tag attached to a stick, tracked by a (2) virtual tag.

The study was conducted as a dual-task experiment where the participants were instructed to complete a 1-back task (primary task) while also detecting objects behind them (secondary task), receiving AR cues as support for finding the objects. Correctness in the primary task and detection time of the object were gathered. The tasks are explained further in the sections below. Three different metrics were collected during the test sessions: accuracy on the 1-back task, hazard detection time measured from tag activation, and subjective workload using the raw NASA Task Load Index (NASA TLX).

The study focused on two different visual cue designs to narrow the scope of the thesis: a 3D arrow and a bird's-eye view minimap (described in Section 5.6). The designs implemented were based on previous research within the field of attention direction. The cues provide the user with different kinds of information. The 3D arrow points towards the object, whereas the minimap shows a view of the room and the object's position within it. The user study used a within-subject design in which all participants completed test rounds using both AR cues.

6.2 Participants

A total of 16 participants (6 female, 10 male), ages 21–26 ($M = 24.4$, $SD = 1.5$), participated in the user study. Participants were recruited through convenience sampling and consisted of friends and acquaintances of the authors. Nine of the participants reported that they had experienced AR with an HMD prior to the study. Eight of the participants reported that they were "very experienced" with technology overall.

6.3 Setup

The UWB-based RTLS was set up in the test room. A total of eight UWB transceivers were installed as physical anchors for the RTLS. Four anchors were put at floor level in corners, and four were put close to the ceiling (see Figure 6.2). The physical UWB anchors were then calibrated using the AR application described in Section 5.5.4.

A black line was put on the floor to show where participants should be standing during the test session. During the test rounds participants wore the Meta Quest 3 and a pair of headphones playing white noise. Two test leaders were present in the room during the whole session. One test leader (tag operator) walked around the room and held a stick with an additional UWB transceiver attached to one end. This transceiver was configured as a tag in the RTLS and tracked by a virtual tag visible through the Meta Quest 3 (Figure 6.1). Another test leader sat at a desk monitoring the session and the results. The room setup can be seen in Figure 6.2. All test sessions were recorded through the Meta Quest 3. The whole session for each participant lasted approximately 30 minutes.

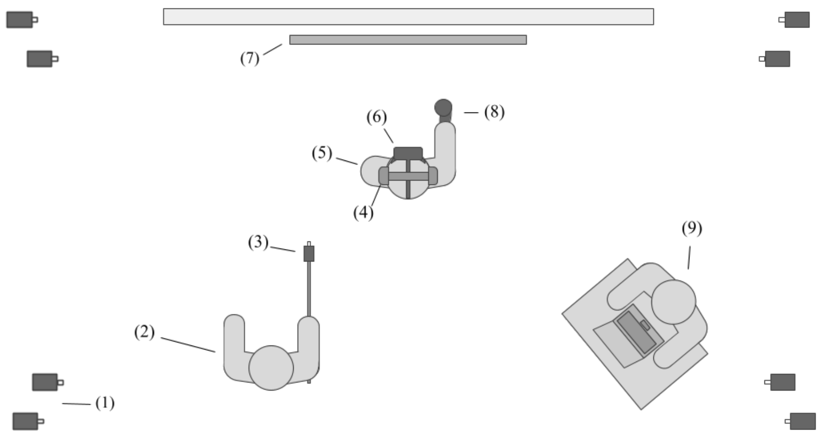


Figure 6.2 The test setup. (1) Two UWB anchors, one close to floor and one close to ceiling, (2) tag operator, (3) UWB tag, (4) headphones, (5) test participant, (6) Meta Quest 3, (7) 1-back task, (8) hand controller, and (9) test leader (at desk)

6.4 Tasks

Each session consisted of two rounds. In each round, participants completed two concurrent tasks: a 50-letter 1-back task as the primary task, and hazard detection as the secondary task. The implementation of the tasks is described in Section 5.7. Participants were instructed to complete tasks both as quickly and accurately as possible.

6.4.1 Primary task

The 1-back task was displayed on a virtual menu placed towards a wall. Participants were instructed to press the A-button on the controller when the letter visible was the same as the one previously shown. They were asked to not press anything if the letter was different from the previous one. Each letter was shown for a maximum of 2.5 seconds. A letter would disappear more quickly if the participant pressed A in response to a correct letter match. The menu showing the 1-back task turned green if the participant pressed A for a correct letter and red if they pressed incorrectly. It also turned red if they failed to press for a letter that would have been correct.

6.4.2 Secondary task

Participants were told that moving hazards would appear behind them, and that they should turn and look at each hazard to detect it. AR assistance was provided in both rounds. The two rounds used different AR cues, and the order was counterbalanced across participants (8 starting with the arrow, 8 with the minimap) to mitigate order effects.

6.5 Procedure

Participants were first asked to fill out a consent form (Figure 6.3). The test leaders then briefly introduced them to the study and gave an overview of the test session. After that, participants completed a short AR demo to familiarise themselves with the Meta Quest 3. This was followed by a 10-letter trial of the 1-back task before the first round started to ensure participants understood the task.

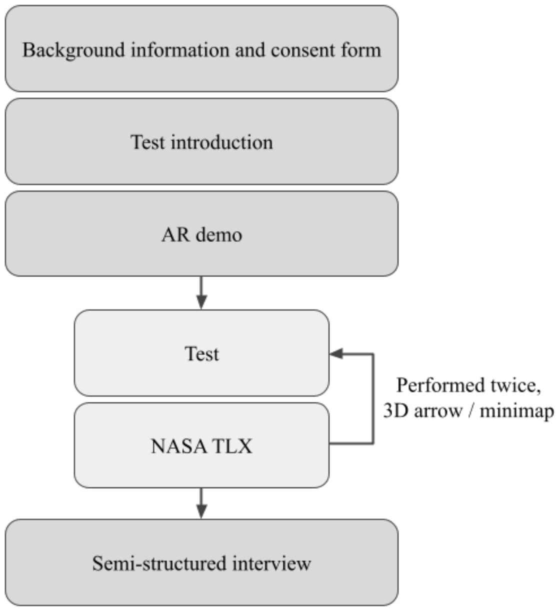


Figure 6.3 Test session for each participant.

The visibility of the virtual object was controlled manually by the operator holding the physical UWB tag. Activation of the virtual tag made the AR cue visible and changed the virtual object's colour to red. The tag was then deactivated when it was viewed by the participant, which also made the AR cue disappear. A test round began with the UWB operator demonstrating the activation of the object and AR cue, allowing the participant to detect and locate the object once.

The virtual tag was then activated a total of six times throughout each test round when a predetermined letter was shown to the participant. Different activation schedules were followed for the two rounds. The schedules, referred to as "A" and "B",

Table 6.1 Tag activation schedules used for the two different test rounds.

Letter	Horizontal	Vertical
6/50	Left	Middle
15/50	Right	Down
27/50	Left	Up
33/50	Middle	Up
39/50	Right	Up
47/50	Middle	Middle

(a) Schedule A, used for the first test round.

Letter	Horizontal	Vertical
4/50	Right	Middle
10/50	Middle	Down
17/50	Right	Down
29/50	Left	Up
35/50	Middle	Up
46/50	Left	Down

(b) Schedule B, used for the second test round.

described when and from which direction the object should move, see Table 6.1a and 6.1b. The first tag activation in schedule A was for example when the sixth letter in the sequence was shown in the 1-back task. Schedule "A" was always used for a participant's first session and schedule "B" was used for the second session, regardless of which AR cue was used. The tag operator walked around the room during the test session, following the schedules. The physical tag, tracked by the virtual tag, was always placed in the starting position before the virtual tag was activated.

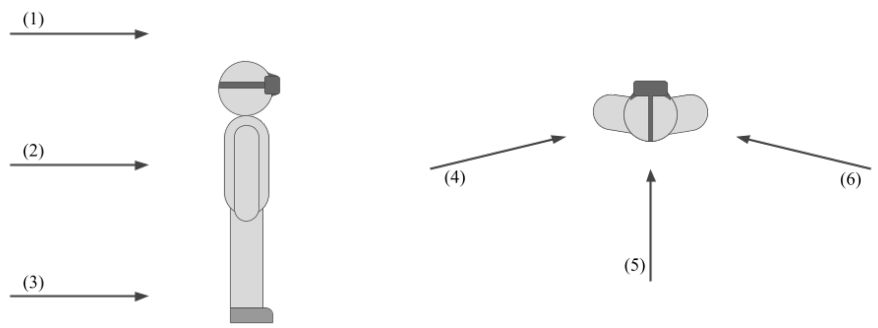


Figure 6.4 Vertical starting positions: (1) up, (2) middle, (3) down. Horizontal starting positions: (4) left, (5) middle, (6) right.

Three different horizontal directions were used: left, middle and right. There were also three different vertical starting points: up, middle, down. Middle in the vertical direction was around waist height for the participant, see Figure 6.4. The virtual object was activated at one of these directional combinations then moved slowly towards the participant while maintaining the tag's vertical position.

After each completed round the participants were asked to fill out the NASA TLX form. When both rounds were completed, a short semi-structured interview was held where the participants were asked about their overall experience.

6.6 Results

All statistical tests use the Wilcoxon signed-rank test with a significance threshold of $p < 0.05$ for paired comparisons unless otherwise noted. The convention for differences is minimap − arrow, so positive values indicate an arrow advantage. One participant (P12) misunderstood the minimap instructions during the test session. Results are reported for all 16 participants, with figures shown both with and without P12 for comparison.

6.6.1 Detection time

The 3D arrow led to significantly faster detection times than the minimap (Figure 6.5). Participants using the arrow detected objects in a mean of 1571 ms ($SD = 400$), compared to 2464 ms ($SD = 1547$) with the minimap. The mean paired difference was 893 ms in favour of the arrow ($p = 0.005$). 13 out of 16 participants were faster with the arrow. Excluding P12, the result remains significant ($p = 0.010$).

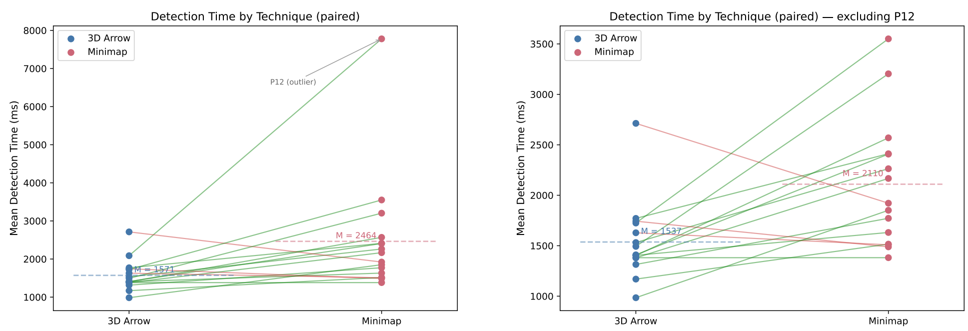


Figure 6.5 Paired detection times per participant. Green lines indicate the arrow was faster; red lines indicate the minimap was faster. Left: all participants ($n = 16$, $p = 0.005$)**. Right: excluding P12** ($n = 15$, $p = 0.010$)**.**

6.6.2 Cognitive load (NASA TLX)

On the NASA TLX, where a higher score indicates greater perceived workload, overall scores were lower with the arrow ($M = 33.6$) than with the minimap ($M = 39.1$), but this difference was not statistically significant ($p = 0.052$). None of the individual TLX subscales reached significance either, see Table 6.2.

Although no individual comparison reached the significance threshold, the direction was consistent: the arrow scored equal or lower on all six subscales (Figure 6.6). 13 out of 16 participants reported lower overall workload with the arrow.

Table 6.2 NASA TLX subscale scores by technique

Subscale	Arrow	Minimap	p
Mental Demand	49.4	56.6	0.091
Physical Demand	16.6	16.6	1.000
Temporal Demand	43.1	49.1	0.050
Performance	23.1	28.4	0.221
Effort	44.4	51.3	0.053
Frustration	25.3	32.4	0.344
Overall	33.6	39.1	0.052

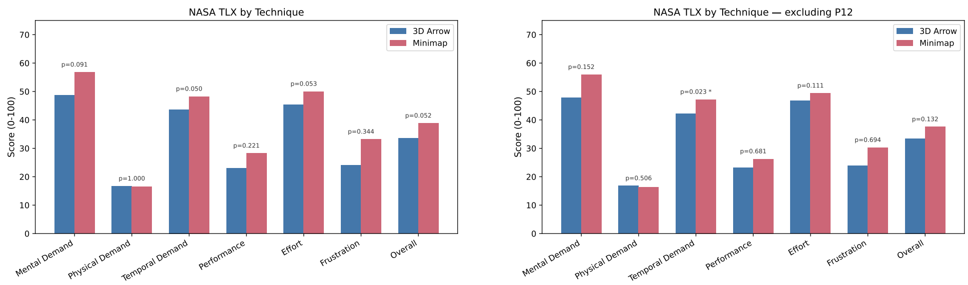


Figure 6.6 NASA TLX subscale scores by technique. Left: all participants ($n = 16$). Right: excluding P12 ($n = 15$).

6.6.3 1-back performance

Accuracy on the 1-back task was near ceiling for both techniques (arrow: 99.3%, minimap: 98.3%), confirming that participants followed the dual-task instructions and suggesting that 1-back may not have been demanding enough to differentiate the conditions.

Reaction time on correct 1-back matches was significantly faster during arrow sessions ($M = 548$ ms) than minimap sessions ($M = 602$ ms), with a mean difference of 54 ms ($p = 0.044$). This result becomes marginal when excluding P12 ($p = 0.055$).

6.6.4 Placement effects

Detection time varied with the spatial placement of the object. The arrow's advantage over the minimap was largest for objects placed low, where the minimap condition showed notably longer detection times (Figure 6.7).

The arrow's advantage was driven primarily by the minimap performing worse for certain placements, rather than the arrow performing better. Interview data suggests this is because the minimap provided no vertical information, making low and high objects harder to locate. One participant (P15) noted: "I think the arrow points in

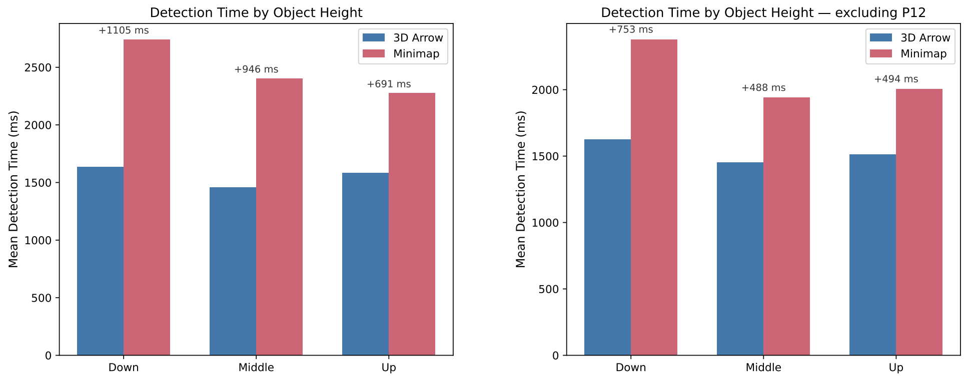


Figure 6.7 Mean detection time by object height. Left: all participants ($n = 16$). Right: excluding P12 ($n = 15$).

multiple dimensions. If the objects had been more spread out vertically, the arrow might have been more helpful."[1]

6.6.5 Order effects

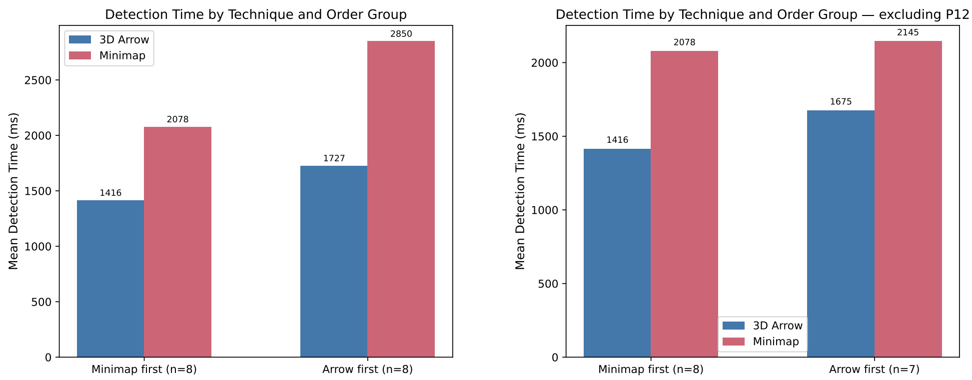


Figure 6.8 Detection time by technique and order group. Left: all participants ($n = 16$). Right: excluding P12 ($n = 15$).

A strong learning effect was observed (Figure 6.8). Regardless of which technique was used, participants were generally faster in round 2 than in round 1. Of the 16 participants, 11 (69%) were faster in their second round, and 11 (69%) perceived their second round as the easiest.

This pattern held across both counterbalancing groups: participants who started with the minimap improved in round 2 with the arrow, and participants who started with

[1] Original: "Jag tror att den här pilen pekar i flera dimensioner. Om föremålen hade varit mer utspridda uppåt, kanske pilen hade varit mer hjälpsam."

the arrow also improved in round 2 with the minimap. The counterbalanced design (Section 6.4) ensures that this learning effect does not systematically favour one technique.

11 out of 16 participants spontaneously mentioned learning or practice effects during the interview.

6.6.6 Interview findings

Semi-structured interviews were conducted at the end of the test session (Figure 6.3). Three main themes emerged from the interviews:

- **Arrow as intuitive and immediate (7/16 participants).** Participants described the arrow as requiring no interpretation. P9 said: "It felt more natural to see the direction rather than having to process: where am I, where is it."[2] P14 reported: "The arrow was much clearer. You saw the arrow and just—now I need to turn around."[3]

- **Minimap requires self-localisation (7/16 participants).** Participants reported that the minimap required a cognitive step of locating themselves on the map before determining where the object was. P4 described the process: "There I am, there it is. In relation to that, it's over there."[4] Two participants (P3, P4) reported abandoning the minimap during the test and searching without it.

- **Dual-task strategies (7/16 prioritised the 1-back task, 5/16 prioritised navigation, 4/16 neither).** Participants developed different strategies for managing the two concurrent tasks. Some focused on the 1-back task and only attended to navigation cues between letters, while others prioritised finding the object quickly. The remaining four reported prioritising neither task in particular, attending to both as the situation demanded. P1 described: "I tried to have a technique where I check which letter comes next and then turn around."[5]

Other notable themes from the interviews included: the arrow creating a sense of urgency (2/16), the arrow being readable in peripheral vision (1/16), and the minimap being faster to parse in a split second for some participants (2/16).

[2] Original: "Det kändes mer naturligt att se riktningen snarare än att behöva bearbeta: där är jag, där är den."

[3] Original: "Pilen var mycket tydligare. Man såg pilen och bara—nu ska jag vända mig om."

[4] Original: "Där är jag, där är den. I förhållande till det så är den där borta."

[5] Original: "Försökte jag ha en teknik att jag kollar vilken bokstav som kommer sen och sen vänder jag mig."

6.6.7 Cross-analysis

To examine whether subjective preferences aligned with objective performance, each participant's stated preference was compared with their actual detection times.

- **Preference distribution.** Of 16 participants, 11 preferred the arrow, 4 preferred the minimap, and 1 had no preference. Preference was typically strong or moderate (13/15 who expressed a preference).
- **Preference matches performance.** 13 out of 15 participants (87%) who expressed a preference actually performed better with their preferred technique. This match rate was significantly above chance (binomial test, $p = 0.007$).
- **Group differences.** Participants who preferred the arrow had a substantially larger arrow advantage in detection time ($M = 1314$ ms) than those who preferred the minimap ($M = -12$ ms; Mann-Whitney U, $p = 0.028$). This pattern held when excluding P12: the arrow-preferrers' advantage reduced to 876 ms but remained significant ($p = 0.038$). Arrow-preferrers also reported a larger workload reduction with the arrow on the NASA TLX (mean TLX difference 8.7 vs 1.0 for minimap-preferrers; $p = 0.045$).

These results suggest that participants' subjective preferences were well-calibrated to their actual performance differences, rather than being driven by novelty or surface-level impressions.

7 Discussion

This chapter interprets the user study results and considers what the findings can and cannot support. It also reviews the limitations of both the study and the implementation, and outlines improvements and directions for future work.

7.1 Process reflection

Developing an AR application that could be used to help industrial workers detect potential hazards was established as the goal early on in the process. During the initial phases of research for the thesis, several studies were found examining adjacent scenarios. Many of these studies directed participants' attention to virtual hazards or objects. The idea of conducting a user study, but with real-life objects (physical objects present in the room rather than purely virtual ones) as hazards, was then established. The imagined system needed for the study required a RTLS to track the hazardous object, an AR application to display warning signals, and a way of combining the data from the RTLS with Unity to direct participants in the correct direction.

The wish to use real-life objects for the user study in this thesis therefore meant that a significant amount of preparatory work was needed before the design of the user study could be finalised. As a result the first test sessions with real participants were conducted later than originally imagined. This meant that there was only time for one iteration of the study. Clearer conclusions could potentially have been reached had the study been repeated based on the findings of this study.

7.2 Interpretation of results

Several aspects of the results deserve closer attention before drawing broader conclusions about visual AR cues for hazard detection.

7.2.1 Implementation versus concepts

The study compared two specific implementations: a single 3D arrow that pointed directly at the active hazard, and a top-down minimap that showed the hazard relative to the participant. The findings therefore support the narrower claim that this arrow outperformed this minimap under the present single-target dual-task setup, rather than a general superiority of arrows over maps. A different minimap design (with vertical information, distance scales, or motion trails) or a different arrow design (without 3D shading, or carrying additional information) could shift the balance. The comparison also reflects a task in which a single hazard had to be located as quickly as possible, which plays directly to the arrow's strength of encoding direction in one step. In a scenario with several simultaneous hazards, or one where the user benefits from maintaining a persistent overview of where hazards are, the minimap's exocentric view could prove more useful. The present result is therefore best read as a starting point for understanding when each cue type is appropriate, rather than a final verdict on arrows versus maps.

7.2.2 Individual differences

The aggregate detection-time advantage was driven primarily by the 11 participants who preferred the arrow. Their mean arrow advantage was 1314 ms, compared to a mean of -12 ms for the 4 minimap-preferrers, who performed roughly equally with both cues. Among the 15 participants who expressed a preference, 13 performed better with their preferred technique (binomial test, $p = 0.007$). Self-reported preference therefore appears to be well-calibrated to actual performance, suggesting that future deployments should consider letting users select their preferred cue rather than imposing a single design.

7.2.3 Vertical guidance

The arrow's lead over the minimap was largest for objects placed below waist height, where mean detection time with the minimap was particularly long. The minimap shows the hazard's bearing on a 2D plane and gives no indication of its height, so participants had to rely on additional visual search to locate low or high placements. P15 articulated this directly: "I think the arrow points in multiple dimensions. If the objects had been more spread out vertically, the arrow might have been more helpful."[1] The minimap's main weakness in this study therefore appears to be its missing vertical channel rather than a fundamental issue with the bird's-eye perspective. This matters for the intended deployment context, where hazards such as overhead cranes approach from above and vehicles or moving stock from around floor level, so a cue that conveys vertical position directly is likely to be particularly valuable for steel mill workers.

[1] Original: "Jag tror att den här pilen pekar i flera dimensioner. Om föremålen hade varit mer utspridda uppåt, kanske pilen hade varit mer hjälpsam."

7.2.4 Qualitative and quantitative convergence

The detection-time results, the consistent direction of the NASA TLX subscales (the arrow scored equal to or lower than the minimap on all six), the faster 1-back reaction times during arrow rounds, and the interview themes all point in the same direction. Participants described the arrow as immediate and direction-revealing ("it felt more natural to see the direction rather than having to process: where am I, where is it"[2]) and the minimap as requiring an extra self-localisation step ("there I am, there it is"[3]). The convergence between subjective experience and objective measures supports the conclusion that the arrow advantage reflects a true difference in cognitive load rather than an artefact of a single measure.

7.3 Study limitations and potential improvements

The design and execution of the study were subject to limitations that should be considered when interpreting the results.

7.3.1 Participants

Several of the limitations of the user study were related to the participant sample. These included a small sample size, limited participant diversity, and a background questionnaire that was too brief.

7.3.1.1 Sample size and statistical power

The relatively low participant count of $n = 16$ contributed to several observed trends (including the NASA TLX results with $p = 0.052$) not reaching statistical significance. A larger sample might have made the trends clearer. The age range in our study was only 21–26, and the majority of participants also had a similar technological background. Future studies could benefit from a broader participant pool with a wider age range and more diverse backgrounds. Participants from the intended target group of steel workers could also be considered. Steel workers were not part of the sample here because participants were recruited through convenience sampling from the authors' acquaintances, and recruiting from an active steel mill was outside the practical scope of this thesis.

Having a larger number of participants could also allow for a between-subject study where each participant only tries one AR cue. This test setup could also allow for

[2] Original: "Det kändes mer naturligt att se riktningen snarare än att behöva bearbeta: där är jag, där är den."

[3] Original: "Där är jag, där är den."

the same tag activation schedule to be used for all tests, and would eliminate the risk for order effects.

7.3.1.2 Height

Participants' heights were not gathered before the test session. The tag operator following the tag schedules took the height into consideration by placing the horizontal "middle" around the participant's waist height. The experience of looking up or down did, however, likely vary between participants of different height. A few participants almost reached the maximum height of the RTLS, which meant a shorter distance to the user and the HMD compared to a shorter participant.

A more comprehensive questionnaire before the test, including a question about participants' heights, could be beneficial to analyse and understand the test results further.

7.3.1.3 Headset-induced discomfort

A few participants reported feelings of nausea or general discomfort while wearing the Meta Quest 3 during the session. The headset's weight, its passthrough latency, and the proximity of the displays add attentional and physical costs that are unlikely to be present with lighter, workplace-appropriate AR glasses. The dual-task results obtained here therefore reflect performance under more demanding wearable conditions than a real deployment with purpose-built hardware would likely impose, which slightly weakens the comparison to a real-world industrial setting.

7.3.2 Tasks

The construction and execution of the tasks for the user study could also be improved. This includes both the clarity of the instructions given to participants and the consistency of the tasks across sessions.

7.3.2.1 1-back task

The responses on the 1-back task show a near 100% accuracy during both test rounds, causing a ceiling effect. This meant that reaction time for detecting the hazardous object was the main measure used to objectively determine overall test round performance. The ceiling effect also suggests that the task was too easy. A more mentally demanding n-back task, such as 2-back, could give a more noticeable result difference between participants and their different test rounds. Requiring participants to press two different buttons, one for a matching letter and one for a non-matching one, could potentially also affect their attention on the 1-back task.

The focus of the user study was, however, to examine the effectiveness of the different AR cues. The purpose of the 1-back task was therefore mainly to occupy the participants and keep their focus away from their surroundings so that they only looked for hazards when an AR cue became visible.

The randomness for the 1-back task could also be seen as a potential concern for the study results. The letters were independently randomised with each new letter having a 30% probability of being a match. This meant that each participant received different amounts of matches per test round, and at different intervals. One participant mentioned in the interview that their task focus during the test rounds differed because of the difference in the letter sequences. They felt that the letters in the first round were almost exclusively different, meaning almost no matches and fewer button presses. This gave them more freedom when trying to locate the objects, something that potentially affected their object detection time. Having predetermined letter sequences assigned to the different steps of the test could minimise the differences between the participants' sessions.

7.3.2.2 Hazard detection

The schedules used for activating the tag were created manually and did not adhere to a distinct pattern. As a result, there was an uneven distribution of the vertical directions used. Results from the user study indicated that the vertical "down" placement had a larger difference in reaction time between the two different AR cues, see Section 6.6.4. If a within-subject design is used in future experiments, more similar tag activation schedules could be implemented to facilitate the comparison between the tests with different AR cues. Participants' height could also be taken into account to make the object's distance from the HMD as similar as possible.

7.3.2.3 Minimap

The minimap cue proved to be more confusing for some of the participants than it was helpful. Two separate participants reported that they stopped using the bird's-eye view cue to determine the location of the object. They instead simply used it as a general indicator that an object was approaching and turned around to manually locate the object without using additional AR support. This could slightly complicate the interpretation of their detection times during the test round using the minimap. Another participant also misunderstood the instructions for the bird's-eye view, creating an outlier in the results.

7.3.3 Order effects

The within-subject design of the study allowed all participants to try both of the AR cue types during the test session. This created a risk for order effects affecting the test results. Attempts were made to counterbalance this by alternating which cue each participant used first. The interview answers from the test sessions revealed that 11 out of the 16 participants reported experiencing some carry-over effect from their first test round that made the second round somewhat easier regardless of which AR cue was used. Alternating the first cue could therefore have balanced out some systematic bias.

The setup of the test session could, however, still create within-subject variability as the different participants have different experiences of the session and respond differently to the order effect.

7.3.4 Task priority

The introduction of the test session always followed the same script. The test leader first mentioned that the participant was going to perform two separate tasks at the same time. The 1-back task was then presented first, followed by the hazard detection task. No emphasis was made by the test leader on either task being a "primary task". Finally participants were instructed to try and complete both tasks as quickly and correctly as possible. A number of participants did, however, mention in the concluding interview that they felt like they prioritised the 1-back task during the session. Future experiments could possibly put more emphasis on prioritising the hazard detection task, and point out the importance of avoiding hazards for their safety.

7.3.5 Virtual and physical tag

In the current test setup the tag operator holding the physical UWB tag is visible for the participant through the Meta Quest 3 regardless of the activation status of the virtual tag. The visual presence of the tag operator could also unintentionally help guide the participants' FOV towards the virtual tag since the operator walked around the room and stood close to the horizontal starting position for each activation. A few participants found it unclear when the hazard was counted as detected as both the physical tag and the tag operator remained in their FOV until they turned back to the 1-back task.

Using only a virtual tag for hazard detection, similar to the prototype in Section 4.2, could potentially create a scenario more similar to a real-life environment. It could also make it clearer for participants when the tag is detected.

The physical tag is also only a proxy for a real hazard. It is a small device carried on a stick and moved at walking pace by an operator, which differs substantially from real industrial hazards such as overhead cranes, on-site vehicles, or moving stock in size, speed, and trajectory. The detection times measured here should therefore be read as a comparison between the two cues under a controlled proxy hazard, rather than as absolute response times to the kinds of hazards the system would ultimately need to surface.

7.3.6 Controlled environment

The tests were conducted within a controlled testing environment that differs in several ways from a real industrial setting, which limits how directly the results transfer. During the sessions the headphones played white noise to mask auditory cues, but white noise is uniform and relatively easy to tune out, whereas industrial noise is typically harsh, dynamic, and more distracting, and could impose a higher cognitive load than was present here. The hazards also only ever appeared outside the participant's field of view, while real hazards can appear both within and outside it, so the study does not capture how the cues perform when a hazard is already partly visible. In addition, the study did not include a no-cue baseline, so it shows which of the two cues performed better but not how much either improves on receiving no AR assistance at all. Further testing could benefit from a user study conducted in an environment more comparable to a real industrial setting, including more realistic noise, physical obstacles, and hazards appearing both inside and outside the field of view, as well as a no-cue baseline condition.

7.4 Future work

The findings suggest several directions for further study, both in refining the cue designs themselves and in extending the evaluation to richer scenarios.

7.4.1 Multiple simultaneous tags

The arrow used in this study points at one target at a time. In a real industrial environment, several hazards may be present and active at the same time. A minimap, or a similar overview cue, is well suited to displaying multiple objects in parallel, while a single arrow would need extending (multiple arrows, switching targets, or a priority rule) to cope with the same scenario. Comparing arrow-style cues with overview cues under varying numbers of concurrent hazards would help identify the point at which an overview becomes more useful than a single direction.

7.4.2 Improved minimap

The clearest weakness identified in this study was that the minimap conveys only horizontal bearing and distance, with no indication of vertical position. Adding a height indicator (for example as a vertical bar, a numeric label, or colour coding by height band) could directly address the placement effect seen in the results. A motion trail showing the hazard's recent path is another simple addition that could give faster intuition about where the hazard will be rather than only where it currently is. Of these, adding a height indicator is the most promising change to test first, since the absence of vertical information was the minimap's main measured weakness in this study.

7.4.3 Simplified 2D arrow

Participant P15 commented that the 3D shading on the arrow felt unnecessary. A flat 2D arrow may retain the same directional immediacy with less interpretive load, and would be easier to render legibly on lower-end AR glasses. A direct comparison between the 3D arrow used here and a flat 2D arrow would clarify whether the visual richness of the 3D arrow contributes to performance or only to aesthetics.

7.4.4 Hybrid cues

The arrow and the minimap address different needs: focused single-target guidance, and overview awareness. A hybrid display that shows both at once, or that switches between them depending on the situation, could in principle inherit the strengths of each. It could equally well double the visual clutter and cognitive load. The trade-off is an open question and a natural follow-up to the present comparison.

7.4.5 Multimodal cues

Both cues used in this study were purely visual. Vibrotactile feedback has been used to support hazard awareness [19] and to convey spatial direction [37], with comparable effectiveness to visual cues in some scenarios [45]. Combining a visual directional cue with a haptic alert could reduce the visual-attention cost of the AR cue and provide a redundant channel for the most critical hazard signals. An evaluation pairing the present arrow with a directional haptic signal would test whether the two channels combine additively or compete for attention.

7.4.6 Field testing and lighter hardware

The user study was carried out in a controlled lab environment using a Meta Quest 3 headset. A natural next step is to evaluate the cues in an environment closer to the intended deployment context: a steel mill or comparable heavy-industry workplace, where overhead cranes, on-site vehicles, and moving stock define the hazards the system is meant to surface. The ambient noise, heat, dust, and visual complexity of such environments differ substantially from a lab and would test how robust the cues remain under realistic conditions. Lighter and more workplace-appropriate AR glasses, suitable for wearing under or alongside standard personal protective equipment (PPE) such as a hard hat, would also reduce a likely barrier to deployment in steel mills, and may change how the cues are perceived if the FOV or rendering fidelity differs substantially from the headset used here.

8 Conclusions

This chapter summarises the main findings, gives direct answers to the research questions stated in the Introduction, and lists the contributions of the thesis.

8.1 Summary of findings

This thesis combined a UWB-based real-time locating system with an AR application for the Meta Quest 3 to support hazard detection for objects outside the user's field of view. A within-subject user study with 16 participants compared two visual AR cues (a 3D arrow and a top-down minimap) in a dual-task setup with a 1-back primary task and hazard detection as the secondary task. Participants detected hazards significantly faster with the arrow ($M = 1571$ ms) than with the minimap ($M = 2464$ ms, $p = 0.005$). Reaction times in the 1-back task were also faster during arrow rounds. The NASA TLX scores for the arrow were equal to or lower than those for the minimap across all six subscales. Participants described the arrow as immediate and direction-revealing, and the minimap as requiring an extra self-localisation step. These interview themes converged with the quantitative measures. As interpreted in Section 7.2, the minimap's disadvantage is attributed mainly to its missing vertical channel, which left participants without a height indication for low and high hazards, whereas the arrow conveyed elevation directly. These findings hold within a single-target, controlled setting, and their transfer to the noisier, multi-hazard conditions of a real steel mill is examined in the discussion.

8.2 Answering the research questions

The thesis posed three research questions, addressed as follows.

RQ1: *How can the world-locking capabilities of the Meta Quest 3 simplify the process of calibrating out-of-view real-time locating systems?*

The implemented AR calibration tool (Section 5.5.4) uses world-locked spatial anchors and the headset's hand tracking as ground truth. The user records three distance measurements from different locations in the room by pointing at each anchor, and a linear calibration model is fitted per anchor by comparing AR-measured distances to UWB-reported distances. This removes the need for manual measurements with tape measures or laser rangefinders, calibrates all anchors simultaneously, and works regardless of whether the anchors are within line of sight. Evaluated against the AR spatial-anchor reference, the fitted per-anchor model reduced the aggregate ranging error from 21.7 cm to 9.8 cm mean absolute error, with each anchor's linear fit explaining almost all of the variation (R^2 between 0.955 and 0.997, Section 5.5.4). The AR-assisted procedure therefore produces an effective calibration, not merely a more convenient one.

RQ2: *How can AR be used to evaluate the effectiveness of real-time locating systems?*

The Meta Quest 3 served as both the deployment surface and the evaluation harness. A virtual tag rendered at the UWB-tracked position made tracking quality (jitter, drift, latency) directly visible to a researcher wearing the headset, supporting qualitative inspection during development. The headset also evaluated the system quantitatively: by treating its world-locked spatial-anchor distances as a reference, the AR application measured the UWB ranging error and drove the calibration reported in Section 5.5.4, using AR to assess and improve the locating system rather than only to display it. For quantitative evaluation of the cues, the same headset captured head pose and detection events during the user study, turning the AR system itself into the measurement instrument that produced the detection-time results reported in this thesis. A caveat is that the headset itself contributes to the experimental conditions: a few participants reported nausea or general discomfort, so cognitive-load measures gathered through the Quest 3 likely overstate the burden that a lighter, workplace-appropriate AR device would impose.

RQ3: *Which AR navigation cue, a 3D arrow or a minimap, better supports the detection of out-of-view physical objects?*

Within the scope of this study, the 3D arrow outperformed the minimap. Detection times were significantly faster with the arrow ($p = 0.005$), 13 of 16 participants were faster with the arrow, NASA TLX scores were consistently equal or lower for the arrow, and the interview themes pointed in the same direction. The result supports the narrower claim that this arrow outperformed this minimap under the present single-target dual-task conditions, rather than a general superiority of arrows over maps.

8.3 Contributions

The main contributions of this thesis are:

- An end-to-end implementation combining a UWB-based real-time locating system (BU03-Kit hardware, positioning server, multi-tag support, Kalman filtering) with a Meta Quest 3 AR application that renders virtual hazards co-located with their physical UWB tags.
- An AR-assisted calibration method for UWB anchors that uses world-locked spatial anchors and the headset's hand tracking as ground truth, removing the need for manual measurement of anchor positions.
- A within-subject user study ($n = 16$) comparing two AR navigation cues for out-of-view hazard detection under dual-task load, with detection time, NASA TLX, 1-back accuracy and reaction time, and semi-structured interviews as data sources.
- Empirical evidence that, within this specific implementation and study setup, a 3D arrow outperformed a top-down minimap on detection time and on every NASA TLX subscale, together with a placement-effect analysis identifying the minimap's missing vertical channel as its main weakness.

References

[1] T. Akhmetov and H. A. Varol. “An augmented reality-based warning system for enhanced safety in industrial settings”. *IEEE Transactions on Industrial Informatics* **19**:7 (2023), pp. 7966–7977. DOI: 10.1109/TII.2022.3216009.

[2] S. M. Asaad and H. S. Maghdid. “A comprehensive review of indoor/outdoor localization solutions in iot era: research challenges and future perspectives”. *Computer Networks* **212** (2022), p. 109041.

[3] R. T. Azuma. “A survey of augmented reality”. *Presence: teleoperators & virtual environments* **6**:4 (1997), pp. 355–385.

[4] J. Baek and Y. Choi. “Smart glasses-based personnel proximity warning system for improving pedestrian safety in construction and mining sites”. *IJERPH* **17**:4 (Feb. 2020), pp. 1–15. DOI: None. URL: https://ideas.repec.org/a/gam/jijerp/v17y2020i4p1422-d323989.html.

[5] N. Biswas, A. Singh, and S. Bhattacharya. “Augmented 3d arrows for visualizing off-screen points of interest without clutter”. *Displays* **79** (2023), p. 102502.

[6] A. Bosworth. *Meta CTO on third-party apps for display glasses*. AMA video, shared via Reddit. Accessed: Apr. 23, 2026. 2025. URL: https://www.reddit.com/r/MetaRayBanDisplay/comments/1p604i8/meta_cto_on_3rd_party_apps_for_display_glasses/.

[7] J. Broderick. *Effect of spatial audio and movement control on user performance, awareness, and preference*. PhD thesis. NUI Galway, 2021.

[8] G. Camacho-Fidalgo, B. Judkins, K. Friederichs, L. Soberanis, V. Hernandez, K. McSweeney, F. Witherden, and E. Rojas-Muñoz. *Analyzing the Impact of Augmented Reality Head-Mounted Displays on Workers’ Safety and Situational Awareness in Hazardous Industrial Settings*. 2025. arXiv: 2503.04075 [cs.HC]. URL: https://arxiv.org/abs/2503.04075.

[9] L. Cheng, Z. Wu, B. Lai, Q. Yang, A. Zhao, and Y. Wang. “Ultra wideband indoor positioning system based on artificial intelligence techniques”. In: *2020 IEEE 21st International conference on information reuse and integration for data science (IRI)*. IEEE. 2020, pp. 438–444.

[10] G. Corrêa De Almeida, V. Costa de Souza, L. G. Da Silveira Júnior, and M. R. Veronez. “Spatial audio in virtual reality: a systematic review”. In: *Proceedings of the 25th Symposium on Virtual and Augmented Reality*. 2023, pp. 264–268.

[11] A. Dechaux, M. Haytam-Mahsoub, M. Kitazaki, J. Lagarde, and G. Ganesh. "Multisensory feedback improves spatially compatible sensori-motor responses". *Scientific Reports* **12**:1 (2022), p. 20253.

[12] Design Council. *The double diamond*. URL: designcouncil.org.uk/our-resources/the-double-diamond/2 (visited on 2026-02-13).

[13] F. van Diggelen. "Indoor gps theory & implementation". In: *2002 IEEE Position Location and Navigation Symposium (IEEE Cat. No. 02CH37284)*. IEEE. 2002, pp. 240–247.

[14] DipFlip. *ultra-wideband-positioning*. https://github.com/DipFlip/ultra-wideband-positioning. Open-source UWB positioning project. Accessed May 18, 2026. 2024.

[15] Y. Fang, J. Lu, N. Bryan-Kinns, J. Ou, P. Lu, and H. Wang. "Vibrotactile warnings design for improving risks awareness in construction environment". *International Journal of Human–Computer Interaction* **41**:9 (2025), pp. 5350–5372.

[16] M. A. Fischler and R. C. Bolles. "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography". *Communications of the ACM* **24**:6 (1981), pp. 381–395.

[17] W. H. Foy. "Position-location solutions by Taylor-series estimation". *IEEE Transactions on Aerospace and Electronic Systems* **12**:2 (1976), pp. 187–194.

[18] M. Gadhvi, M. J. Moore, and M. Waseem. *Physiology, Sensory System*. https://www.ncbi.nlm.nih.gov/books/NBK547656/. Updated May 6, 2023. Accessed Jan. 29, 2026. Treasure Island, FL, 2023.

[19] S. A. Hossain, A. Yin, and T. K. Ferris. "Exploring vibrotactile displays to support hazard awareness in multitasking control tasks for heavy machinery work". *Safety* **11**:1 (2025), p. 26.

[20] Y.-J. Huang, J.-C. Lin, S.-S. Lee, and B.-J. Wu. "Reading and walking with smart glasses: effects of display and control modes on safety". *International Journal of Human–Computer Interaction* **40**:23 (2024), pp. 7875–7891. DOI: 10.1080/10447318.2023.2276529. eprint: https://doi.org/10.1080/10447318.2023.2276529. URL: https://doi.org/10.1080/10447318.2023.2276529.

[21] Y. Huang, A. Hammad, and Z. Zhu. "Providing proximity alerts to workers on construction sites using bluetooth low energy rtls". *Automation in Construction* **132** (2021), p. 103928. ISSN: 0926-5805. DOI: https://doi.org/10.1016/j.autcon.2021.103928. URL: https://www.sciencedirect.com/science/article/pii/S0926580521003794.

[22] *Information technology – real-time locating systems (rtls) – part 2: 2,4 ghz air interface protocol*. Withdrawn; cited for terminology. International Organization for Standardization and International Electrotechnical Commission, 2006. URL: https://www.iso.org/standard/40508.html.

[23] International Federation of Robotics. *Collaborative Robots: How robots work alongside humans*. Accessed: May 20, 2026. 2024.

[24] R. E. Kalman. "A new approach to linear filtering and prediction problems". *Transactions of the ASME–Journal of Basic Engineering* **82**:Series D (1960), pp. 35–45.

[25] A. M. Kappers, J. Bay, and M. A. Plaisier. "Perception of vibratory direction on the back". In: *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*. Springer. 2020, pp. 113–121.

[26] S. Kim, M. A. Nussbaum, and J. L. Gabbard. “Augmented reality “smart glasses” in the workplace: industry perspectives and challenges for worker safety and health”. *IIE Transactions on Occupational Ergonomics and Human Factors* **4**:4 (2016), pp. 253–258. DOI: 10.1080/21577323.2016.1214635. eprint: https://doi.org/10.1080/21577323.2016.1214635. URL: https://doi.org/10.1080/21577323.2016.1214635.

[27] T. Lawal. “Overcoming gnss limitations: real-time locomotive localization algorithms in tunnel-like railway stations” (2026).

[28] D. Lee and S. Choi. “Perceptual alignment of spatial auditory and tactile stimuli for effective directional cueing”. *IEEE Transactions on Visualization and Computer Graphics* (2025).

[29] B. Li, K. Zhao, and E. B. Sandoval. “A uwb-based indoor positioning system employing neural networks”. *Journal of Geovisualization and Spatial Analysis* **4**:2 (2020), p. 18.

[30] C. T. Li, J. C. Cheng, and K. Chen. “Top 10 technologies for indoor positioning on construction sites”. *Automation in Construction* **118** (2020), p. 103309. ISSN: 0926-5805. DOI: https://doi.org/10.1016/j.autcon.2020.103309. URL: https://www.sciencedirect.com/science/article/pii/S0926580519306016.

[31] X. Li, M. Ge, X. Dai, X. Ren, M. Fritsche, J. Wickert, and H. Schuh. “Accuracy and reliability of multi-gnss real-time precise positioning: gps, glonass, beidou, and galileo”. *Journal of geodesy* **89**:6 (2015), pp. 607–635.

[32] T. Margiani, S. Cortesi, M. Keller, C. Vogt, T. Polonelli, and M. Magno. “Angle of arrival and centimeter distance estimation on a smart UWB sensor node”. *IEEE Transactions on Instrumentation and Measurement* **72** (2023), pp. 1–10. DOI: 10.1109/TIM.2023.3282289.

[33] K.-E. Michelsen. “Industry 4.0 in retrospect and in context”. In: *Technical, Economic and Societal Effects of Manufacturing 4.0: Automation, Adaption and Manufacturing in Finland and Beyond*. Springer, 2020, pp. 1–14.

[34] Z. Momynkulov, N. Omarov, and Y. Uxikbayev. “Detection of dangerous situations by sounds in real-time using deep learning”. In: *2024 IEEE 4th International Conference on Smart Information Systems and Technologies (SIST)*. IEEE. 2024, pp. 278–283.

[35] D. Neirynck, E. Luk, and M. McLaughlin. “An alternative double-sided two-way ranging method”. In: *2016 13th workshop on positioning, navigation and communications (WPNC)*. IEEE. 2016, pp. 1–4.

[36] Omron Europe. *A Guide to Collaborative Robot Safety*. White Paper. Accessed: May 20, 2026. 2019. URL: https://files.omron.eu/downloads/latest/publication/en/cobot_safety_expert_article_en.pdf.

[37] M. Pascher, T. Franzen, K. Kronhardt, U. Gruenefeld, S. Schneegass, and J. Gerken. “Haptix: vibrotactile haptic feedback for communication of 3d directional cues”. In: *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI EA ’23. Association for Computing Machinery, Hamburg, Germany, 2023. ISBN: 9781450394222. DOI: 10.1145/3544549.3585601. URL: https://doi.org/10.1145/3544549.3585601.

[38] M. A. Plaisier, C. S. Vleeshouwers, N. Boonstra, Y. Shi, S. J. van der Velden, W. K. Vos, and A. M. Kappers. "Vibrotactile spatial acuity on the back". *Perception* **53**:9 (2024), pp. 619–631.

[39] K. Quinn and J. L. Gabbard. "Augmented reality visualization techniques for attention guidance to out-of-view objects: a systematic review". In: *2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2024, pp. 826–835. DOI: 10.1109/ISMAR62088.2024.00098.

[40] A. Raikwar, D. Mifsud, C. D. Wickens, A. U. Batmaz, A. C. Warden, B. Kelley, B. A. Clegg, and F. R. Ortega. "Beyond the wizard of oz: negative effects of imperfect machine learning to examine the impact of reliability of augmented reality cues on visual search performance". *IEEE Transactions on Visualization and Computer Graphics* **30**:5 (2024), pp. 2662–2670. DOI: 10.1109/TVCG.2024.3372062.

[41] P. Renner and T. Pfeiffer. "Attention guiding techniques using peripheral vision and eye tracking for feedback in augmented-reality-based assistance systems". In: *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. 2017, pp. 186–194. DOI: 10.1109/3DUI.2017.7893338.

[42] A. Rowen, M. Grabowski, and J.-P. Rancy. "Through the looking glass(es): impacts of wearable augmented reality displays on operators in a safety-critical system". *IEEE Transactions on Human-Machine Systems* **49**:6 (2019), pp. 652–660. DOI: 10.1109/THMS.2019.2944384.

[43] *Safety and Health in the Steel Industry: Data Report 2025*. Data report. Accessed: Jan. 27, 2026. World Steel Association, 2025. URL: https://worldsteel.org/safety-and-health/safety-and-health-in-the-steel-industry-data-reports/safety-and-health-in-the-steel-industry-data-report-2025/.

[44] G. Salvendy. *Handbook of human factors and ergonomics*. John Wiley & Sons, 2012. Chap. 19, pp. 553–568.

[45] Z. Samsel, A. Gunia, M. I. Jäger, and J. Schöning. "A comparison of vibrotactile patterns in an early warning system for obstacle detection using a haptic vest". *Applied Ergonomics* **122** (2025), p. 104396. ISSN: 0003-6870. DOI: https://doi.org/10.1016/j.apergo.2024.104396. URL: https://www.sciencedirect.com/science/article/pii/S000368702400173X.

[46] A. San Martin and J. Kildal. "Audio-visual mixed reality representation of hazard zones for safe pedestrian navigation of a space". *Interacting with Computers* **33**:3 (2021), pp. 311–329.

[47] A. Seeliger, R. P. Weibel, and S. Feuerriegel. "Context-adaptive visual cues for safe navigation in augmented reality using machine learning". *International Journal of Human–Computer Interaction* **40**:3 (2024), pp. 761–781. DOI: 10.1080/10447318.2022.2122114. eprint: https://doi.org/10.1080/10447318.2022.2122114. URL: https://doi.org/10.1080/10447318.2022.2122114.

[48] J. de Souza and R. Tartz. "Visual perception and user satisfaction in video see-through head-mounted displays: a mixed-methods evaluation". *Frontiers in Virtual Reality* **5** (2024), p. 1368721. DOI: 10.3389/frvir.2024.1368721. URL: https://www.frontiersin.org/journals/virtual-reality/articles/10.3389/frvir.2024.1368721/full.

[49] The Global Goals. *9 Industry, innovation and infrastructure*. https://globalgoals.org/goals/9-industry-innovation-and-infrastructure/. Accessed: Jan. 27, 2026.

[50] M. Tonnis, C. Sandor, G. Klinker, C. Lange, and H. Bubb. “Experimental evaluation of an augmented reality visualization for directing a car driver’s attention”. In: *Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR’05)*. IEEE. 2005, pp. 56–59.

[51] L. Tremosa. *Beyond ar vs. vr: what is the difference between ar vs. mr vs. vr vs. xr?* Accessed: May 20, 2026. 2025. URL: https://www.interaction-design.org/literature/article/beyond-ar-vs-vr-what-is-the-difference-between-ar-vs-mr-vs-vr-vs-xr.

[52] S. Vi. “Exploring user experience guidelines for designing hmd extended reality applications”. *University of Calgary* (2020). URL: http://hdl.handle.net/1880/111919.

[53] N. Weidinger, T. Schreck, B. H. Thomas, N. ElSayed, and E. Veas. “Investigating the effect of visual cue density on situational awareness during immersive navigation”. In: *2025 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2025, pp. 1333–1343. DOI: 10.1109/ISMAR67309.2025.00138.

[54] G. Welch and G. Bishop. *An introduction to the Kalman filter*. Technical Report TR 95-041. University of North Carolina at Chapel Hill, 2006.

[55] X. Xu, Y. Lu, B. Vogel-Heuser, and L. Wang. “Industry 4.0 and industry 5.0—inception, conception and perception”. *Journal of manufacturing systems* **61** (2021), pp. 530–535.

[56] C. Zheng, Y. Ge, and A. Guo. “Ultra-wideband technology: characteristics, applications and challenges”. *arXiv preprint arXiv:2307.13066* (2023).

[57] Z. Zhu, Y. Zhang, W. Guo, C. Pan, and Z. Zhao. “Asaudio: a survey of advanced spatial audio research”. In: *Proceedings of the 14th International Joint Conference on Natural Language Processing and the 4th Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics*. 2025, pp. 417–442.