CARL
VON
OSSIETZKY
universität OLDENBURG

Faculty II – School of Computing Science, Business
Administration, Economics, and Law

Visual Cues for Locating Out-of-View Objects in Mixed Reality

A dissertation accepted by the School of Computing Science,
Business Administration, Economics and Law
of the Carl von Ossietzky University Oldenburg (Germany)
to obtain the degree and title

Doctor of Natural Sciences (Dr. rer. nat.)

By **Uwe Grünefeld, M.Sc.**
Born on 11th of June 1990, in Leer, Germany

Supervisors:

Dr. Wilko Heuten
Prof. Dr. Susanne Boll-Westermann

Reviewers:

Prof. Dr. Susanne Boll-Westermann
Prof. Dr. Hans Gellersen

Day of Defense:

7th of February 2020

Abstract

In the past decade, Mixed Reality has emerged as a promising technology for supporting users with everyday tasks. It allows one to alter perceived reality by blending it with virtual content. Mixed Reality can thus be used to overlay perceived reality with visual cues that empower users to locate relevant objects in the environment, regardless of whether they are in view. This approach seems auspicious for various scenarios, such as a traffic encounter in which a car driver overlooks a cyclist, or the docking process of large container vessels during which the person in charge must monitor several assisting tugboats. In such situations, visual cues that help users to locate relevant objects out of view can improve situational awareness and help to avoid fatal consequences.

In the presented work, we follow the "Research through Design" methodology to develop and evaluate visual cues in Mixed Reality that empower users to locate out-of-view objects. Initially, we analyzed three different scenarios, conducting an ethnographic study, an accident analysis, and literature reviews. Thereafter, inspired by the different scenarios, we reviewed all relevant background and related work to derive three research questions that are studied in depth in the presented work: (RQ1) To what extent can existing off-screen visualization techniques be adapted to cue the direction to out-of-view objects in Mixed Reality?, (RQ2) How can Mixed Reality devices with small fields-of-view be extended to present directional cues to out-of-view objects?, and (RQ3) In what way can the directions and distances of out-of-view objects be visualized for moving or non-moving objects in Mixed Reality?

Our results show that directional cues presented in the user's periphery are easy to perceive and help the user to locate objects quickly. We showed that using three-dimensional visual cues can result in a lower direction estimation error than using two-dimensional cues. Furthermore, if the used Mixed Reality device suffers from a small field-of-view, radial light displays presented around the screen can be used to cue direction for out-of-view objects. However, sometimes a user must simultaneously locate several out-of-view objects, some of which may be occluded. Directional cues alone are insufficient for such cases, so distance information is required as well. We found that it is possible to convey this information, albeit with increased workload and additional visual clutter. Furthermore, visual cues that help a user to locate out-of-view objects should not disappear when these objects are visible on the screen, since assistance when these objects appear in view improves error rates and overall performance of the visual assistance.

Zusammenfassung

Im letzten Jahrzehnt ist Mixed Reality zu einer vielversprechenden Technologie geworden um Nutzer bei alltäglichen Aufgaben zu unterstützen. Mixed Reality ermöglicht es die wahrgenommene Realität mit virtuellen Inhalten zu verschmelzen. Dadurch kann die vom Nutzer wahrgenommene Realität mit visuellen Hinweisen überlagert werden, die helfen relevante Objekte in der Umgebung zu lokalisieren, selbst wenn diese nicht im Sichtfeld des Nutzers liegen. Dieser Ansatz ist vielversprechend für unterschiedliche Szenarien, zum Beispiel im Straßenverkehr, in welchem ein Autofahrer beim Abbiegen einen Fahrradfahrer übersieht, oder im Fall eines andockenden Containerschiffes, in welchem die verantwortliche Person gleichzeitig mehrere unterstützende Schlepper überwachen muss. In all diesen Szenarien können visuelle Hinweise, die es Nutzern ermöglichen relevante Objekte außerhalb der Sicht zu lokalisieren, helfen, die Situationswahrnehmung zu verbessern und gefährliche Konsequenzen zu vermeiden.

In dieser Arbeit, folgen wir der Research-through-Design Methodik, um visuelle Hinweise in Mixed Reality zu entwickeln und zu evaluieren, die es Nutzern ermöglichen, die Position von Objekten außerhalb der Sicht zu bestimmen. Anfänglich haben wir drei verschiedene Szenarien mit Hilfe einer ethnografischen Studie, einer Unfallanalyse und Literaturrecherchen untersucht. Anschließend, inspiriert von den unterschiedlichen Szenarien, haben wir alle Grundlagen und verwandten Arbeiten analysiert, um drei Forschungsfragen abzuleiten, die in der vorliegenden Arbeit im Detail untersucht werden sollen: (RQ1) In welchem Umfang müssen bestehende Off-Screen-Visualisierungstechniken angepasst werden, um die Richtung zu den Objekten außerhalb der Sicht in Mixed Reality anzuzeigen?, (RQ2) Wie können Mixed Reality-Geräte mit einem kleinen Sichtfeld erweitert werden, um Richtungshinweise auf Objekte außerhalb der Sicht zu präsentieren, und (RQ3) Inwiefern kann die Richtung und Distanz zu statischen oder beweglichen Objekten außerhalb des Sichtfeldes in Mixed Reality dargestellt werden?

Unsere Ergebnisse zeigen, dass visuelle Hinweise in der Peripherie leicht zu erkennen sind und helfen Objekte außerhalb der Sicht schnell zu finden. Wir konnten zeigen, dass drei-dimensionale visuelle Hinweise zu geringeren Richtungsschätzfehlern führen als zwei-dimensionale Hinweise. Außerdem können radiale Lichtdisplays das vorhandene Display erweitern um Richtungsinformationen anzuzeigen, wenn das Sichtfeld des vorhanden Displays problematisch klein ist. Wenn ein Benutzer jedoch mehrere Objekte außerhalb der Sicht loaklisieren soll die ebenfalls verdeckt sind, dann reichen Richtungsinformationen nicht aus und zusätzliche Distanzinformationen werden benötigt. Wir konnten zeigen, dass es möglich ist diese zusätzlichen Informationen anzuzeigen, jedoch mit erhöhter Arbeitslast und zusätzlicher visueller Verdeckung. Darüber hinaus sollten visuelle Hinweise, die dem Benutzer helfen, Objekte außerhalb der Sicht zu lokalisieren, nicht verschwinden wenn diese Objekte auf dem Bildschirm sichtbar sind, um Fehlerraten zu reduzieren und die visuelle Assistenz effektiv zu verbessern.

Acknowledgments

This thesis would not have been possible without many people who supported me during my time as a doctoral student. I would like to express my gratitude to my supervisor **Wilko Heuten** for the great support, useful discussions, remarks, and engagement throughout the learning process. Furthermore, I would like to thank **Susanne Boll** for her helpful advice during the discussions and presentations of this work, and for the opportunity to conduct my research.

My thanks also go to my committee members **Hans Gellersen**, **Marius Brinkmann** and **Martin Fränzle**. Thank you for your time and efforts. Your valuable feedback helped me to further improve this thesis.

A special thank you goes to my colleagues, my collaborators, my students, my friends, and my family. Without the help of every single one of you, this work would not have been possible, and I am very thankful that I received your support and advice throughout the years.

Last but not least, I would like to thank **Dakota Smith**, who helped me to proof-read my thesis and never declined my attempts to discuss the content of this thesis with her.

Contents

1 Introduction

Awareness of objects in the environment is an important ability required for everyday life. Without information about the locations of objects around us, there are many common tasks we cannot accomplish (e.g., safely navigating traffic). To do so, we have different senses that allow us to perceive information not only about ourselves but also about our environments, with each of those senses receiving different types of information. For example, with our eyes we can visually perceive that which we observe, meaning we can see every object that emits or reflects light into our eyes. With our ears we can hear sounds from everywhere around us (provided they are loud enough) and with our skin we can perceive touch. However, when it comes to the importance of each sense in the orchestra of human perception, sense of vision is clearly paramount (and in some cases even dominant [SB14]). About half of the human brain is devoted directly or indirectly to vision [New96]. From this, the human visual system offers a wide range of methods for locating objects in the environment (e.g., stereoscopic vision and depth perception). Furthermore, visual perception offers the highest bandwidth of information, allowing perception of several objects at once. Here, the combination of these attributes suggests that we humans perform well when it comes to visually locating objects in the environment.

However, human visual perception is restricted by a limited field-of-view, which is about 180 degrees horizontally and 90 degrees vertically [PJW73]. Therefore, we can only perceive a part of our environment at once, with everything else out of view. This is a problem that restricts the ability to locate objects in the environment, because one may be unaware of objects outside one's field-of-view. As a result, we cannot locate these out-of-view objects, which may sometimes lead to negative consequences. For example, when driving a car, one may overlook a pedestrian or cyclist and cause serious harm. Therefore, we think supporting people in locating out-of-view objects is a promising approach that could help to avoid a variety of critical situations.

In this thesis, we want to address the problem of the limited human field-of-view. It seems that the apparent solution would be to extend our field-of-view to 360° horizontally and 180° vertically, thereby removing the limit. However, not only is this challenging, but it could also lead to an excess of information that the human brain cannot process efficiently. Therefore, we suggest extending the human field-of-view only toward relevant objects in the environment (e.g., we could extend the visual perception of car drivers to enable them to locate pedestrians or cyclists outside their fields-of-view). Thereby, we could extend the field-of-view, without increasing the required mental resources more than necessary. In order to extend our field-of-view to be able to locate relevant out-of-view objects, we design and develop visual cues that can encode the locations of such objects in our environment. These visual cues are always presented within the field-of-view and are thus always perceivable by the user.

As a result, the main objective that we follow in this thesis is to investigate the following guiding question:

“How can we encode the locations of out-of-view objects with visual cues to empower humans to locate these objects in 3D space?”

Following our main objective, we utilize the periphery of the user to present the visual cues within the human field-of-view. The periphery is an area in our field-of-view located outside the center of our vision (foveal vision) [Kal01]. Compared to our foveal vision, our peripheral vision offers less detailed information. For example, we have problems reading text presented in our periphery [Gol89]. However, utilizing peripheral vision allows us to use the high bandwidth that visual perception offers while at the same time not cluttering the central field-of-view with additional information. When using the periphery, visual cues can be presented radially around the central field-of-view enabling users to then shift their gaze to focus on the visual cue and perceive it in higher level of detail. Furthermore, peripheral vision has a natural role in directing attention, allowing us to arouse a user’s attention with visual cues presented in the periphery and guide them to out-of-view objects.

However, some technology is required to present visual cues in one’s periphery. A promising technology to fulfill that requirement is Mixed Reality. It allows one to alter perceived reality by augmenting it with additional information [MK94]. Thereby, Mixed Reality technology can present visual cues to users that empower them to locate relevant objects in the environment, regardless of whether they are in view. In this work, we focus on two manifestations of Mixed Reality in particular: Augmented Reality (AR) and Virtual Reality (VR). In other words, we investigate different approaches in Mixed Reality that extend our visual perception to enable us to locate out-of-view objects.

1.1 Scenarios

The problem of objects receding from view is relevant in many different scenarios. Whether for aircraft pilots or mariners who must avoid collision, for people playing computer games who want to know where the opponent is, or for service technicians finding machines in an unfamiliar environment, being able to locate surrounding objects is critical. Some solutions have been suggested to address the problem of objects receding from view, but for specific scenarios. In this thesis, we would like to address the problem on a more fundamental level, to gather results that are useful for a wide range of scenarios, even ones that do not yet exist. Still, we need specific scenarios in order to analyze and understand the different contexts in which the problem exists. Furthermore, to ensure that all these scenarios are comparable and share the same problem, we limited ourselves to scenarios in which a user experiences the situation from an egocentric perspective. Therefore, we focus on the perspective of one user that has to locate one or more out-of-view objects. With the help of a literature review, we identified three scenarios which differ greatly from each other and may thus result in very general requirements, leading to a solution that is applicable to a wide range of domains. In the selection of our scenarios, we focused on safety critical situations in which human life is in danger when someone fails to locate the relevant out-of-view object(s). In the following subsections, we describe these scenarios. At the end of this section, we look at all scenarios to abstract the underlying problem.

1.1.1 Traffic Encounter

For many people, navigating traffic is a daily activity. When navigating traffic, there are several options available, such as walking, biking or driving. Nowadays, we have very dense traffic in our cities with many motor vehicles, but also with increasing numbers of cyclists and pedestrians. In Germany, there were 63 million motor vehicles registered in January 2017, meaning about 75% of the population owned a motor vehicle [Kuh17].

With such dense traffic on the roads, there are accidents happening on a daily basis. Worldwide, the total number of road traffic deaths per year has reached about 1.35 million, which equates to about 3700 deaths per day [Org18]. More than half of these deaths occur among vulnerable road users: 28% are motorcyclists and 26% are pedestrians or cyclists. Furthermore, road traffic injury is the leading cause of death for people aged between 5 and 29 years. In addition, traffic-related injuries occur for up to 50 million people annually, some of which are life-altering [Org18]. In a recent study, Kaya et al. examined driving behavior with eye-tracking and found that most drivers turning right do not check for cyclists and pedestrians [KAPD18].

> *"More than half of drivers don't look for cyclists and pedestrians before turning right. [...] Given the minimal personal protection, the severity of these collisions tends to be high."* [KAPD18]

This study [KAPD18] examined drivers who did not look before turning right. However, it is also quite common that even if drivers do look, they still overlook pedestrians and cyclists. This occurs because the field-of-view is restricted by the vehicle and the side mirrors suffer from the so-called *blind spot* (an area close to the vehicle in which nothing can be seen). In German, "blind spot" translates to "death angle," directly implying the potential consequences of this area being unobservable. Failing to look and overlooking before turning are both highly problematic actions, since they can result in horrible accidents and serious injuries, often leaving cyclists and pedestrians the victims of such encounters (see Figure 1.1 - only indicating the possible outcome).

Having said this, not only car drivers are at fault. Similar problems arise for pedestrians and cyclists as well, especially when they are distracted. For exam-

Figure 1.1: Bicycle and helmet lying on the street after an accident between a car and a cyclist (scene posed for picture).

ple, they may be preoccupied with AR games, which have recently gained more popularity (e.g., Pokémon Go[1] or Ingress[2]). Playing AR games on smartphones becomes problematic when it is done while navigating traffic [ALD$^+$16, Eys17]. Interestingly, the simplicity of the game is what makes it so popular [HP18] (besides the technological-improved immersive experience). However, the experience of playing such games on small form-factor devices is like trying to look through a keyhole while concentrating one's focus on a single spot. In this case, increased immersion into the game is not beneficial. Although using a smartphone while operating a car or bicycle is forbidden in most countries, it is still allowed for pedestrians. However, it is highly dangerous for a pedestrian to navigate traffic while texting or playing AR games on a smartphone [NHW08, SBS11]. There are many accident reports that provide evidence for this, but the number of near or minor accidents is likely even higher [Eys17]. In Germany, "Smombie" was the 2015 "youth word of the year." It combines the words "smartphone" and "zombie" to refer to the intensely unaware state of people walking around staring at their phones like zombies[3]. This is made worse by the fact that navigating in traffic is a fundamental requirement of such games.

In summary, thousands of road traffic deaths happen every day, most of which involve so-called *vulnerable road users*. One of the main problems is that in many situations, car drivers as well as vulnerable road users overlook other road users or cannot see one other, resulting in collisions that are often fatal or cause serious injuries. In this scenario, only one out-of-view object is relevant in most situations (e.g., a pedestrian on the sidewalk). However, these objects change their positions over time and therefore, may require the traffic participant to locate these object several times.

---

[1] Pokémon Go. `www.pokemon.com`, last retrieved April 21, 2020

[2] Ingress. `www.ingress.com`, last retrieved April 21, 2020

[3] Smombie. `en.wikipedia.org/wiki/Smartphone_zombie`, last retrieved April 21, 2020

1.1.2 Ship Docking

Another relevant scenario is ship docking. Over the years, container ships have steadily been growing in size [CK00]. This development is mainly caused by increased demand, which requires these container vessels to transport more and more goods. However, the constant lengthening, widening, and deepening of these ships is a huge problem when it comes to safely maneuvering in harbor areas because these areas cannot keep pace in adapting to the growing ships. As a result, maneuvers in harbor areas are the most dangerous [GSB$^+$18].

> *"Most ship collisions [...] occur in harbors, because that is where navigation becomes restricted by other vessels, and man-made structures like jetties, bridges, and piers."* [Fri08]

To ensure safety in harbor maneuvers, all relevant countries in the world are required to have pilots (e.g., in Germany cf. [Haf13]). Pilots are experts for specific areas who support the ship crew with important information about water depths, currents, and maneuvers. Even if the responsibility for the ship remains with the captain, pilots often take over control and give steering commands directly to the helmsman. Therefore, the pilot must be aware of the current situation, including all environmental factors, in order to avoid collisions or grounding. Understanding the situation is of high importance because each error has the potential to cause a collision or grounding [Vei15]. However, accident reports reveal that pilots have problems monitoring their environments successfully. As a result, accidents happen on a regular basis.

To understand the causes of these accidents, we conducted an accident analysis of approximately 1500 accident reports from six different transportation safety authorities [GSB$^+$18]. The most common cause given for container ship accidents in harbor areas was human error. However, further examination of accident reports that cited human error as the main cause revealed that it often came down to a pilot's problematic decision that was based on wrong or missing information.

> *"The pilot could not see the tug from the bridge and assumed that she had been running with the ship stern-to-stern, from which position it would have been relatively easy for the tug to position herself on the ship's port quarter."*
> [MAIB, Report No. 08/1998, Trijine]

> *"The actions of the pilot to reduce the speed of approach of the vessel's bow were unsuccessful because Marineco Toomai was positioned just aft of amidships of Logos II and not on the port quarter as the pilot had assumed."*
> [MAIB, Report No. 01/2008, Logos II]

The maneuvering of large vessels is especially difficult due to the large number of other objects in the harbor environment. Some of these objects must simply be avoided to not risk a collision (e.g., other ships), while others, such as tugboats, assist the vessel during the maneuvering process. Without their help, the container vessel's drive technology cannot create enough force to drive the ship in any direction. From an ethnographic study, we observed the docking process of a large container vessel with two tugboats in the port of Hamburg (see Figure 1.2). We identified the tugboats as the most important objects in the environment and further noted that they are often occluded or out of view. To determine requirements, we interviewed 30 European harbor pilots with 15 years of experience on average. We found that the maximum number of tugboats involved in such a docking process is six. Further, it is relevant to the pilots where these tugs are positioned, how they move, and with how much force they push or pull the container vessel during the docking process (all results and details can be found in our journal paper; see [GSB^{+}18]).

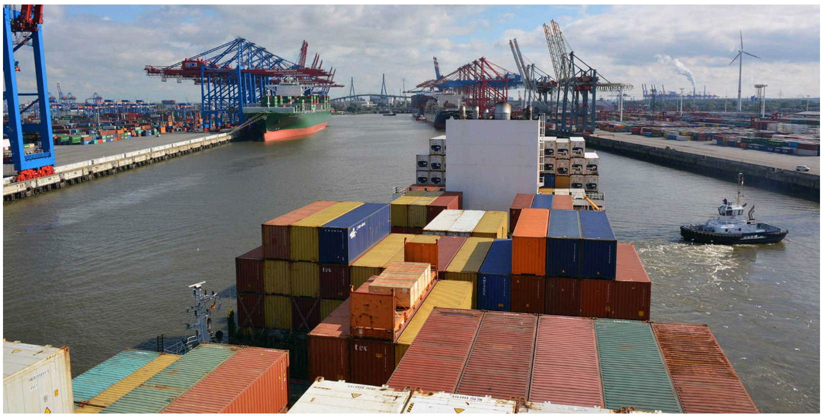


Figure 1.2: Container vessel docking with the assistance of two tugboats in the harbor of Hamburg, Germany (observed in ethnographic study).

During the docking process of large container vessels, pilots must be aware of several objects in their environment, especially the tugboats that assist in the docking process. Therefore, the pilot must observe multiple tugboats while walking over the ship bridge and focusing on the ship's movement in order to ensure safe docking. Here, the tugboats change their positions according to the voice commands given by the pilot.

1.1.3 Fire Fighting

Our last scenario in which objects outside of the user's field-of-view must be located addresses emergency situations. In this scenario, there are circumstances that have caused or will cause harm to people involved. Therefore, the most important task is locating these people as quickly as possible to either remove them from the dangerous area or administer aid for injuries.

> *"During a disaster, users may need to locate unseen patients who are calling for help."* [FMH15]

The workforce that trains for such emergency situations includes firefighters. They are skilled to deal with emergency situations such as burning buildings (e.g., an industrial plant). However, the dense smoke in the building and lack of knowledge about the building structure make it hard to locate people in need for help. In this scenario, a team of two people enters a building and tries to rescue people that need help (see Figure 1.3). For their own security, they attach a rope on one of the two firefighters upon entry to know the way out [K+14]. Two other firefighters secure the rope and monitor the time (for oxygen resources) as well as communicate with the two inside. The two firefighters on the inside of the building must locate people that need help or the source of the smoke, despite very restricted visibility and often having to crawl on the floor for better vision.

Figure 1.3: Firefighter entering a building with restricted visibility due to heavy smoke (picture taken during firefighter training).

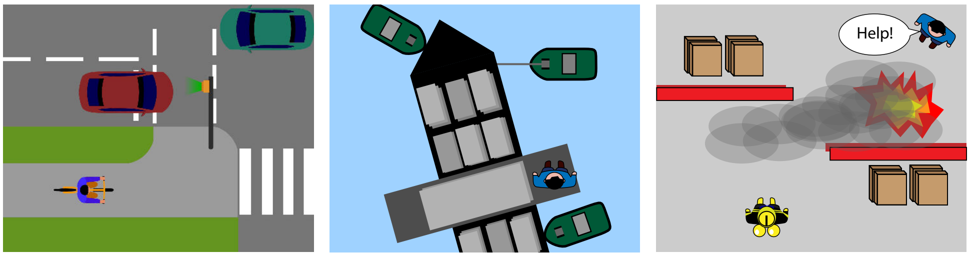
We can summarize this situation as two people walking together through a building with very restricted visibility, while at the same time having to locate many objects (e.g., people in need of help or source of smoke) as quickly as possible. These objects mostly do not change their positions over time. However, even a very close object may be unperceivable if it is occluded by smoke. Virtual Reality has been used to train using simulations of such situations for over a decade now [TSK97, BTG97, ELH19]. Firefighters can learn how to react in certain situations as well as test new head-mounted visualizations that can help in such situations [SH13, AKW+17]. As a result, the transfer of knowledge from the training environment to the real world is rather simple.

Emergency situations are critical situations in which people need to be rescued as quickly as possible. Often, there is more than one way to reach a person in danger but there may not be much knowledge about these paths. In these situations, firefighters must find their way to those in need of help without risking too much injury to themselves. However, Virtual Reality offers the potential to train for these situations and to test out new techniques for assisting firefighters in locating relevant objects.

1.1.4 Problem

All three scenarios share the same problem of one needing to locate unperceived objects outside of view. If one fails to locate the out-of-view objects, it can have fatal consequences. In this thesis, we address this problem by developing visual cues to help one locate these out-of-view objects. Therefore, in this section, we describe the underlying problem that all three scenarios share.

First, each scenario is experienced from an egocentric perspective. For example, in the traffic encounter from the perspective of the car driver, cyclist or pedestrian (see Figure 1.4a), in the ship docking scenario from the perspective of the pilot (see Figure 1.4b), and in the fire fighting scenario from the perspective of the fire fighter (see Figure 1.4c). Figure 1.4 shows all three scenarios.



(a) Traffic encounter. (b) Ship docking. (c) Fire fighting.

Figure 1.4: Schematic representation of the three different scenarios that share the problem of objects receding from view.

All scenarios are either experienced in the real world or in a virtual environment (e.g., to train for the situation). Similarly, out-of-view objects can be either virtual objects (e.g., a representation of a person in need of help or opponents in multi-player games [PKB10]) or real objects in the surrounding environment (e.g., vulnerable road users or points of interest during sightseeing [HB10]). The number of objects can vary from one object (e.g., in the traffic encounter) to up to eight or eleven objects (e.g., in the fire fighting scenario or in gaming). These objects may either be at static positions that do not change (e.g., a person that needs help) or change their positions over time (e.g., a cyclist or tugboat). Furthermore, objects can be distributed on a “ground plane” (e.g., cars and pedestrians on roads or tugboats on water) or in three dimensions (e.g., people in different levels of a building that need to be rescued by a firefighter, aircrafts in an aviation scenarios, or opponents in a computer game). The objects’ distances can change between scenarios as well, with some objects being far away (e.g., tugboats that are up to 500 meters away from the pilot) and others being close to the user (e.g., cyclists or pedestrians relative to a car driver turning right). In some cases, it may also be useful to convey additional information about the out-of-view objects (e.g., their size or type). Still, the object location is the most important piece of information in all scenarios.

1.2 Approach and Methodology

In order to address the problem of objects being hidden from view, we design visualization techniques that assist users in locating these objects. Each visualization technique consists of as many visual cues as there are objects that must be located. Therefore, each out-of-view object is represented by a visual cue that encodes its location. Since the visual cue is an additional element presented in the user's field-of-view that only exists to encode the location of the object out of view, we often refer to the visual cue as a *proxy*. To ensure that the visualization technique does not interfere with other tasks that may need to be carried out, we present the visual cues in the periphery of the user's field-of-view. Thereby, the center of the field-of-view remains uncluttered by excess information. However, a user can shift her or his gaze to perceive the visual cue in more detail (e.g., to be able to locate the out-of-view object more precisely). Furthermore, the periphery can be used as a natural way to visually shift attention towards a specific out-of-view object. For example, when a car driver overlooks a cyclist, the visualization technique can shift the driver's attention to a visual cue in the periphery to make them aware of the cyclist.

In this thesis, our main objective is to develop visualization techniques that help users to locate out-of-view objects. To reach this objective, we followed the "Research through Design" (RtD) method [ZFE07]. The fundamental idea behind this method is to generate new knowledge by studying existing literature and then designing and evaluating research artifacts that represent improved future states [ZF14]. However, to ensure that our approach takes different contexts of use into consideration, we additionally applied selected methods from the human-centered design process. The process is defined in the ISO standard 9241-210:2010 and is basically a cycle of four phases through which each solution must pass and which can be iterated over several times [ISO10]. The four phases are: 1) define the context of use, 2) specify the requirements, 3) design solutions, and 4) evaluate solutions. All four phases benefit from active involvement of users and offer different methods for reaching their objectives. However, since we did not apply the process to develop final products, but rather to investigate our research questions using research prototypes, we selected relevant methods and did not strictly follow the process. For example, since the addressed problem is relevant in many scenarios, we did not focus on specifying one context of use. Instead, we analyzed several scenarios and identified their shared underlying problem.

The results of our analysis of different contexts of use were described in the previous section (see Section 1.1). In summary, we presented three different scenarios that all share the problem of objects being hidden from view. For each scenario, we analyzed and defined the context of use. We did so using ethnographic studies, accident analyses, interviews, and online surveys. Afterwards, we derived a general problem, which will be addressed in this thesis. Here, we abstract from concrete scenarios and investigate how to locate out-of-view objects in different contexts. After explaining the background information and related work, we describe our conceptual design, in which our research questions are defined (see Chapter 4). The research questions presented in this thesis address fundamental aspects of our problem, and the answers to these questions can later be applied to the scenarios that we introduced.

In order to answer our research questions, we used selected methods from the human-centered design process. Each research question focuses on a concrete and fundamental aspect of the problem we want to address. Each research question is addressed in its own chapter, in which the developed prototypes and the study results are described. We used prototyping to produce research artifacts (prototypes), with which we study our ideas for novel visualization techniques. In line with Moggridge, we define a prototype as “a representation of a design, made before the final solution exists” [Mog06]. Since our main objective is to develop visual cues that can be naturally understood by users, we focus more on the cues’ appearances and how users perceive them than on a concrete technical implementation. Therefore, prototyping is a fitting approach to reach that objective, especially low-fidelity (lo-fi) prototyping. This type of prototyping is well suited for quickly producing design solutions that can then be evaluated with users to obtain quick feedback. However, since Mixed Reality does not always offer concrete methods for prototyping for specific use cases, we developed our own lo-fi prototyping tools where necessary. They were built using a combination of digital fabrication technologies, such as 3D printing and laser cutting, and hardware prototyping tools, such as NodeMCU developer boards[4]. Our prototypes are evaluated in user studies in which we often compare different implementations that aim to deliver the same functionality. Thereby, we can receive early feedback that can then be used to further improve the prototypes. Unlike in our analysis of the context of use and requirements, we did not focus on a specific user group in our user studies because we investigated on a fundamental level how visual cues must be designed in order to to be perceived well. Therefore, we ensured our participants were balanced in gender and represented a wide age range to avoid any bias towards a specific user group. All subjects participated voluntarily in our user studies and did not receive any compensation. Participants were informed that they could withdraw from the study at any point in time. We recruited our participants with calls for participation on social media, on the notice boards at our university and research institute, and with flyers distributed in the

---

[4] NodeMCU. `www.nodemcu.com`, last retrieved April 21, 2020

city of Oldenburg. All studies were conducted as laboratory studies, following a within-subject design with repeated measures. Therefore, every participant was required to complete every condition of the study. To avoid any learning effects, we used a counter-balanced design in all our user studies.

The metrics used in the conducted user studies are in line with the existing literature. In all experiments, we used subjective and objective measures to gather participants' opinions about the tested prototypes as well as to collect impartial measurements uninfluenced by participants' opinions. For the subjective measures, we used Likert-items and standardized questionnaires. We used the Likert-items to allow participants to rate statements with 5-point Likert-items (1=strongly disagree, 5=strongly agree). The two standardized questionnaires we used were the System Usability Scale (SUS) questionnaire [Bro96] to measure usability and the NASA Raw TLX (Raw-TLX) questionnaire [Har06] to measure perceived workload. The SUS questionnaire is a "quick and dirty" tool for measuring usability with 10 Likert-items. For each filled out questionnaire, a score between 0 and 100 can be calculated, for which 68 is the average score that represents acceptable usability. The NASA Raw TLX questionnaire is a modified version of the original NASA TLX questionnaire [HS88]. In the modified version, participants must only answer six Likert-items in order for a task load score to be calculated.

For the objective measures, we mainly measured response times, direction estimation error, and distance estimation error for locating out-of-view objects. The differences between groups were assessed using variance tests, while correlations were assessed using correlational tests. All user studies follow a within-subjects design, meaning we always have two or more matched groups. For variance testing, we first assessed the type of data (ratio, interval or ordinal). For ratio or interval data, we used a Shapiro-Wilk test to check for normal distribution of the data. For data not normally distributed, we report median (Md) and interquartile range (IQR). Otherwise, we report mean (M) and standard deviation (SD). Furthermore, we report them when it is general practice (e.g., for reporting the age of participants). To test for significant difference between multiple groups, we used a repeated measures ANOVA if the data was normally distributed. For not normally distributed or ordinal data, we assessed the differences between multiple groups using a Friedman test. To test for significant difference between two groups of normally distributed ratio or interval data, we used a paired t-test. For not normally distributed or ordinal data, we used a Wilcoxon test. The significance level was set as $p<0.05$ for all variance tests. To test for correlation between two variables, we used the Spearman correlation for normally distributed ratio or interval data. Otherwise, we used the Pearson correlation. Data analysis was performed using R[5] (R Studio version 1.1.463 & R version 3.5.3).

---

[5] R project. `www.r-project.org`, last retrieved April 21, 2020

1.3 Thesis Outline

Overall, this thesis consists of eight chapters. Figure 1.5 presents the structure of the thesis. The first three chapters cover the introduction to the problem and the background and related work relevant for the presented research. In chapter 4 the conceptual design and the addressed research questions are described. Chapters 5 to 7 present the user studies that have been conducted to address the foregoing research questions. The last chapter reflects on our findings and highlights potential directions for future work.

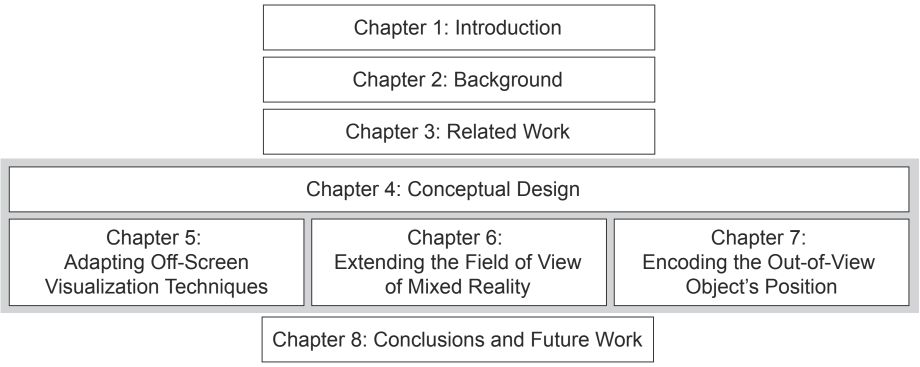


Figure 1.5: Outline of the thesis. Chapters 1, 2, and 3 cover the introduction as well as the background and related work of the thesis. Chapter 4 describes the concept and addressed research questions. Chapters 5 to 7 present user studies that have been conducted to answer the research questions. Chapter 8 provides a discussion of the results and highlights possible directions for future work.

In the first chapter, we introduced the research topic that the thesis addresses and presented three example scenarios that underline the applicability of the conducted research. We continued with an explanation of the used approach and methodology of the thesis. The next chapter describes the background, focusing on human visual perception, cognitive maps, and Mixed Reality. In chapter 3, we present the related work that we identified as relevant during our literature review. We categorize and summarize the related work into three main research areas: 1) previously developed techniques for visualizing off-screen content on small-screen devices (e.g., smartphones), 2) approaches that develop or use peripheral displays, and 3) guidance techniques that help direct attention to relevant objects outside the user's field-of-view.

In chapter 4, we abstract from the scenarios presented in the introduction to define the problem space that we address in this thesis. Afterwards, we describe the different strategies that humans use to locate objects in the environment based on current neuroscientific research of cognitive maps. Thereafter, we examine characteristics of human visual perception to both define the area in which the visual cues will be presented to the user and determine how the visual cues must be designed in order to be perceivable in that area. Afterwards, we discuss the used Mixed Reality technology. We then highlight the challenges of the conducted research. Since we follow the human-centered design process, each of the sections in this chapter will define a certain set of requirements that must be fulfilled. We end the chapter with concrete research questions derived from the problem that we address in this thesis.

Chapter 5 is dedicated to adapting off-screen visualization techniques that point to out-of-view objects. First, we identified three off-screen visualization techniques used for small-screen devices and transferred them to head-mounted Mixed Reality devices. In a user study, we compared all three techniques, showing the individual advantages of each. Afterwards, we focused on the two best-performing techniques and further improved them based on the results from the first user study. Then, we compared the techniques with each other in two user studies with different Mixed Reality devices (Augmented and Virtual Reality). At the end, we took the last of the three techniques from the first user study and tweaked it to further improve it's performance. In another user study, we showed the benefits of the improved technique.

Chapter 6 presents our research on directional cues that are presented on an extended screen. First, we develop a prototyping tool that allows us to test different configurations of display extensions for Augmented and Virtual Reality. Thereafter, we develop a prototype based on our prototyping tool that allows one to cue direction in Virtual Reality and evaluate the usefulness of doing so. Afterwards, we transfer the results of the study to a prototype that aims to support cueing direction to out-of-view objects in Augmented Reality. We finish the chapter with another prototype that focuses on an usecase with pedestrians.

Chapter 7 covers all our work related to encoding the positions of out-of-view objects in Mixed Reality. First, we develop a prototyping tool that allows one to quickly test visual cues for optical see-through Augmented Reality. Afterwards, we create three variants of a new visualization technique to encode the positions of out-of-view objects and test it with our prototyping tool. Thereafter, we take the best variant, implement it for video see-through AR, and compare it to other out-of-view object visualization techniques. Then, we improve the technique further to allow it to be used in Virtual and optical see-through Augmented Reality, as well as to optimize it for devices with smaller fields-of-view. We conclude the chapter with another study in which we test our technique against an existing technique from gaming for out-of-view objects that change position over time.

In the last chapter, chapter 8, we discuss the research that has been conducted in the scope of this thesis. First, we briefly summarize all foregoing chapters. We then discuss the contributions made to each research question. Thereafter, we examine the scenarios presented in the introduction and give recommendations which of the developed techniques could be used to address that scenario and why. We end the thesis with an outlook on future work.

1.4 Publications

Excerpts of this work have been published in peer-reviewed scientific conferences and journals. In the following we list all core publications, ordered by their publication date in descending order.

- Gruenefeld, Uwe ; Prädel, Lars ; Heuten, Wilko: Improving Search Time Performance for Locating Out-of-View Objects in Augmented Reality. In: *Proceedings of Mensch Und Computer 2019*. New York, NY, USA : ACM, 2019 (MuC'19). – ISBN 978–1–4503–7198–8, 481–485
- Gruenefeld, Uwe ; Koethe, Ilja ; Lange, Daniel ; Weiss, Sebastian ; Heuten, Wilko: Comparing Techniques for Visualizing Moving Out-of-View Objects in Head-mounted Virtual Reality. In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2019. – ISSN 2642–5246, S. 742–746
- Gruenefeld, Uwe ; Stratmann, Tim C. ; Jung, Jinki ; Lee, Hyeopwoo ; Choi, Jeehye ; Nanda, Abhilasha ; Heuten, Wilko: Guiding Smombies: Augmenting Peripheral Vision with Low-Cost Glasses to Shift the Attention of Smartphone Users. In: *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 2018, S. 127–131
- Gruenefeld, Uwe ; Stratmann, Tim C. ; Brueck, Yvonne ; Hahn, Axel ; Boll, Susanne ; Heuten, Wilko: Investigations on Container Ship Berthing from the Pilot's Perspective: Accident Analysis, Ethnographic Study, and Online Survey. In: *TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation* 12 (2018), September, Nr. 3, 492-498. `http://dx.doi.org/10.12716/1001.12.03.07`. – DOI 10.12716/1001.12.03.07
- Gruenefeld, Uwe ; Ali, Abdallah E. ; Boll, Susanne ; Heuten, Wilko: Beyond Halo and Wedge: Visualizing Out-of-view Objects on Head-mounted Virtual and Augmented Reality Devices. In: *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*. New York, NY, USA : ACM, 2018 (MobileHCI '18). – ISBN 978–1–4503–5898–9, 40:1–40:11
- Gruenefeld, Uwe ; Stratmann, Tim C. ; Ali, Abdallah E. ; Boll, Susanne ; Heuten, Wilko: RadialLight: Exploring Radial Peripheral LEDs for Directional Cues in Head-mounted Displays. In: *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*. New York, NY, USA : ACM, 2018 (MobileHCI '18). – ISBN 978–1–4503–5898–9, 39:1–39:6
- Gruenefeld, Uwe ; Stratmann, Tim C. ; Prädel, Lars ; Heuten, Wilko: MonoculAR: A Radial Light Display to Point Towards Out-of-view Objects on Augmented Reality Devices. In: *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*

Adjunct. New York, NY, USA : ACM, 2018 (MobileHCI '18). – ISBN 978–1–4503–5941–2, 16–22

- Gruenefeld, Uwe ; Lange, Daniel ; Hammer, Lasse ; Boll, Susanne ; Heuten, Wilko: FlyingARrow: Pointing Towards Out-of-View Objects on Augmented Reality Devices. In: *Proceedings of the 7th ACM International Symposium on Pervasive Displays.* New York, NY, USA : ACM, 2018 (PerDis '18). – ISBN 978–1–4503–5765–4, 20:1–20:6
- Gruenefeld, Uwe ; Hsiao, Dana ; Heuten, Wilko: EyeSeeX: Visualization of Out-of-View Objects on Small Field-of-View Augmented and Virtual Reality Devices. In: *Proceedings of the 7th ACM International Symposium on Pervasive Displays.* New York, NY, USA : ACM, 2018 (PerDis '18). – ISBN 978–1–4503–5765–4, 26:1–26:2
- Gruenefeld, Uwe ; Ennenga, Dag ; Ali, Abdallah E. ; Heuten, Wilko ; Boll, Susanne: EyeSee360: Designing a Visualization Technique for Out-of-view Objects in Head-mounted Augmented Reality. In: *Proceedings of the 5th Symposium on Spatial User Interaction.* New York, NY, USA : ACM, 2017 (SUI '17). – ISBN 978–1–4503–5486–8, 109–118
- Gruenefeld, Uwe ; Hsiao, Dana ; Heuten, Wilko ; Boll, Susanne: EyeSee: Beyond Reality with Microsoft HoloLens. In: *Proceedings of the 5th Symposium on Spatial User Interaction.* New York, NY, USA : ACM, 2017 (SUI '17). – ISBN 978–1–4503–5486–8, 148–148
- Gruenefeld, Uwe ; Stratmann, Tim C. ; Heuten, Wilko ; Boll, Susanne: PeriMR: A Prototyping Tool for Head-mounted Peripheral Light Displays in Mixed Reality. In: *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services.* New York, NY, USA : ACM, 2017 (MobileHCI '17). – ISBN 978–1–4503–5075–4, 51:1–51:6
- Gruenefeld, Uwe ; Ali, Abdallah E. ; Heuten, Wilko ; Boll, Susanne: Visualizing Out-of-view Objects in Head-mounted Augmented Reality. In: *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services.* New York, NY, USA : ACM, 2017 (MobileHCI '17). – ISBN 978–1–4503–5075–4, 81:1–81:7

Further publications on related topics that may also have contributed to the idea and outcome of this thesis have been published by the author:

- Gruenefeld, Uwe ; Prädel, Lars ; Heuten, Wilko: Locating Nearby Physical Objects in Augmented Reality. In: *International Conference on Mobile and Ubiquitous Multimedia.* New York, NY, USA : ACM, 2019 (MUM '19). – ISBN 978–1–4503–7624–2, 1:1-1:10

- Jung, Jinki ; Lee, Hyeopwoo ; Choi, Jeehye ; Nanda, Abhilasha ; Gruenefeld, Uwe ; Stratmann, Tim C. ; Heuten, Wilko: Ensuring Safety in Augmented Reality from Trade-off Between Immersion and Situation Awareness. In: *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2018. – ISSN 1554–7868, S. 70–79
- Gruenefeld, Uwe ; Bargen, Rieke von ; Heuten, Wilko: Identification of Out-of-View Objects in Virtual Reality. In: *Proceedings of the Symposium on Spatial User Interaction*. New York, NY, USA : ACM, 2018 (SUI '18). – ISBN 978–1–4503–5708–1, 182–182
- Gruenefeld, Uwe ; Löcken, Andreas ; Brueck, Yvonne ; Boll, Susanne ; Heuten, Wilko: Where to Look: Exploring Peripheral Cues for Shifting Attention to Spatially Distributed Out-of-View Objects. In: *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. New York, NY, USA : ACM, 2018 (AutomotiveUI '18). – ISBN 978–1–4503–5946–7, 221–228
- Stratmann, Tim C. ; Löcken, Andreas ; Gruenefeld, Uwe ; Heuten, Wilko ; Boll, Susanne: Exploring Vibrotactile and Peripheral Cues for Spatial Attention Guidance. In: *Proceedings of the 7th ACM International Symposium on Pervasive Displays*. New York, NY, USA : ACM, 2018 (PerDis '18). – ISBN 978–1–4503–5765–4, 9:1–9:8
- Stratmann, Tim C. ; Gruenefeld, Uwe ; Boll, Susanne: EyeMR: Low-cost Eye-tracking for Rapid-prototyping in Head-mounted Mixed Reality. In: *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*. New York, NY, USA : ACM, 2018 (ETRA '18). – ISBN 978–1–4503–5706–7, 90:1–90:2
- Löcken, Andreas ; Blum, Sarah ; Stratmann, Tim C. ; Gruenefeld, Uwe ; Heuten, Wilko ; Boll, Susanne ; Par, Steven van d.: Effects of Location and Fade-in Time of (Audio-)Visual Cues on Response Times and Success-rates in a Dual-task Experiment. In: *Proceedings of the ACM Symposium on Applied Perception*. New York, NY, USA : ACM, 2017 (SAP '17). – ISBN 978–1–4503–5148–5, 12:1–12:4

Reading Hint: Although this thesis represents the work of one PhD candidate, it will use the word "we". This was done to avoid passive constructions, which are more difficult to read. Further, the work conducted as part of this thesis would not have been possible without the help of numerous people. Using "we" instead of "I" is meant to acknowledge all of these people and their contributions. Still, this thesis contains only research that was originally planned and implemented by the author.

2 Background

This chapter reviews the background knowledge fundamental for developing visual cues, which help one to locate out-of-view objects in Mixed Reality. First, it describes human visual perception and highlights the characteristics that must be considered when developing the visual cues to ensure they are well perceivable (see Section 2.1). Second, it reviews cognitive maps, which are the basis for human spatial understanding of where objects in the environment are located (see Section 2.2). Third, it analyzes Mixed Reality to define concrete technologies that will be used to present the visual cues to the user (see Section 2.3).

2.1 Human Visual Perception

Most of the information that is recorded by receptors in our central nervous system is visual information. Therefore, it becomes clear that the eye is one of the most essential organs in the human body.

> *"[...] half of the human brain is devoted directly or indirectly to vision",* said Professor Mriganka Sur of MIT's Department of Brain and Cognitive Sciences (cf. [New96]).

Only on closer inspection does one become aware of the variety of properties of human visual perception. The human eyes help one to recognize different forms and colors, but also assist in estimating locations of objects in the environment on a higher level. However, the visual characteristics, such as color perception or sharpness of sight, are not constant within our field of vision. Thus, this section reviews the fundamentals of the human eye. It describes the monocular and binocular fields-of-view as well as foveal and peripheral vision.

2.1.1 Fundamentals of the Human Eye

The operating principle of the human eye is to perceive light that is emitted or reflected by objects in the environment. As a result, the human eye is a passive component. It functions as an observer that does not send out any sight rays [Kal01]. Instead, it processes incoming light that enters through an opening in the iris called the pupil (see Figure 2.1). The iris muscle changes the size of the pupil, thereby controlling the amount of light that can enter the eye. Interestingly, the irises of both eyes work coupled: if one iris closes due to excess exposure (high luminance), then the iris of the other eye also reacts. The remaining light is bundled by the lens and projected onto the retina, a surface that covers most of the inside of the human eye and consists of over 100 million receptors that react to the incoming light [BABB09].

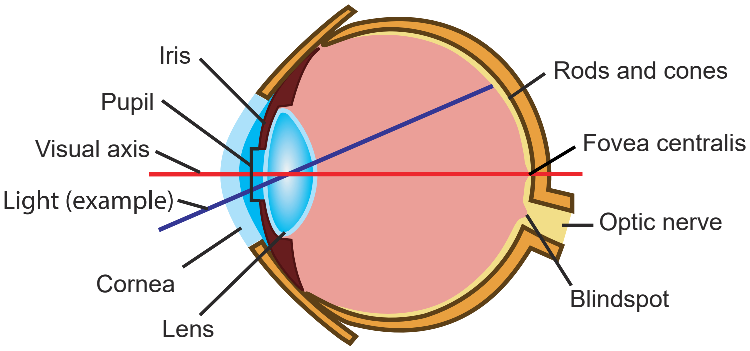


Figure 2.1: Horizontal section through the human eye (cf. [CCK⁺89, Kal01]).

The resulting picture that appears on the retina, coming from the bundled light of the lens, is rotated 180° and has certain wavelength filtered out of the light that entered the lens. Further, the lens of the human eye is adjustable to allow us to focus both distant and near objects on the retina. The point where the visual axis of the optical system hits the retina is called the *fovea centralis* or the *yellow spot* (see Figure 2.1).

While the fovea centralis offers the sharpest vision, the point where the optic nerve leaves the human eye offers no vision at all, and is therefore referred to as the blind spot. However, even though this point is very close to the center of the human visual field, it does not negatively affect visual perception because the missing information resulting from the blind spot of one eye is replaced in our brain by the visual information from the other eye [Kal01].

The structure of the retina is shown in Figure 2.2. The retina offers two types of sensors to perceive the incoming light: cones and rods. Cones are responsible for *photopic vision* (color vision in bright light), and rods offer *scotopic vision* (black-and-white vision in dim light). However, to reach these photoreceptors, light must first pass between the ganglion cells layer, which lies just interior to the photoreceptor cell layer. The axons of the ganglion cells bundle to form the optic nerve, which connects the eye with the brain (see Figure 2.1). The structure of the retina seems strange because light must pass the neuron layer before reaching the photoreceptors. Thereby, the general performance of our eyes is reduced. However, this evolutionary design flaw has only limited influence on the human visual perception characteristics. More important for the characteristics of human vision is the influence of cones and rods.

Cones and rods share the same basic structure, but differ in their photosensitivity. A single photon can trigger an electrical response in a rod, while dozens to hundreds of photons are required for a cone to react. In high brightness, the rods are saturated with light and therefore do not contribute to visual perception in daylight. During twilight and night there is not enough light to

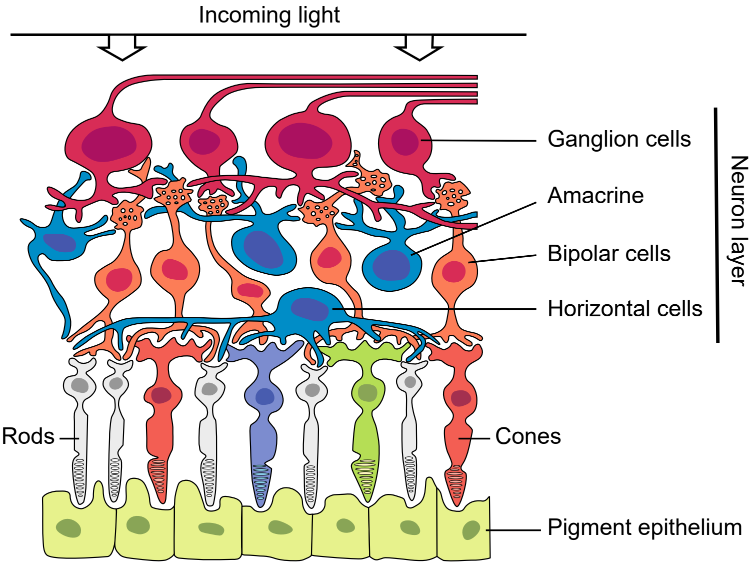


Figure 2.2: Due to the structure of the retina, light must first pass through the neuron layer to reach the photoreceptors (cf. [Kal01]).

cause the cones to react, so the activation of only the rods leads to black and white vision.

Three different types of cones (S-/M-/L- cones) react to short, medium, and long wavelengths of light, thus recognizing different colors. The absorption curves of the cone types still overlap, but one type alone is not sufficient to provide enough information about all color types. The decrease in the number of cones towards the edge of the eye leads to worse color vision in the periphery. In addition, the number of cones for recognizing red and green is five times higher than the number of cones for recognizing blue [Gol89]. Due to this fact, the human eye can perceive representations in red over yellow to green in much more detail than blue.

That the distribution of rods and cones over the retina is uneven was already known in 1935 by Sterberg [Ost35]. Figure 2.3 shows the density of cones and rods on the retina of an eye. The fovea centralis contains the maximum number of cones. The number of cones decreases sharply the further one moves away from the fovea. On the other hand, there are no rods in the fovea, so there is a blind spot in the scotopic vision. For this reason, stars in the night sky can be observed better when one does not look directly at them, but instead a little bit to the side [Kal01]. Towards the edge of the retina, cones and rods mix until only rods remain at the edge. As mentioned earlier, the point of sharpest vision is the fovea centralis. This fact is based on several human eye characteristics. First, the neuron layer with the ganglion cells of the retina is pushed to the side at this point, so the light rays can directly reach the cones. In addition to the cones' high density at the fovea, they are also strongly connected to the central nervous system in the fovea centralis (because almost each cone has it's own ganglion cell connected). At all other points, cones and rods share this connection with others of their kind. These facts combined with the circumstance that the visual axis of the eye directly points to the fovea centralis lead to the increased sharpness in the center of our vision. Therefore, when focusing on objects, the human eyes align in a way that projects the relevant object directly onto the fovea centralis.

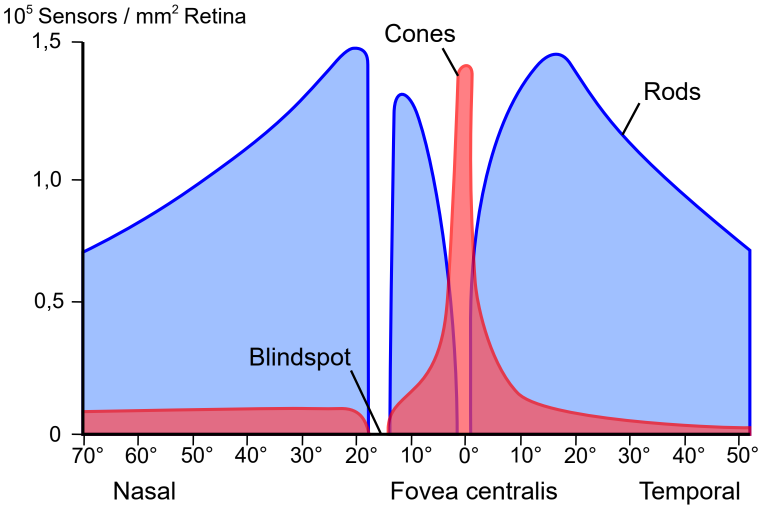
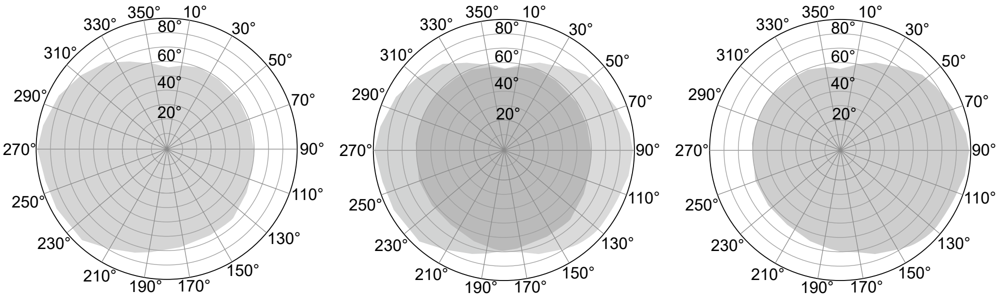


Figure 2.3: Illustration of the distribution of cones and rods on the retina of a human eye for the horizontal field-of-view (cf. [Kal01]).

2.1.2 Monocular and Binocular Fields-of-View

The monocular field of vision refers to the field-of-view of a single human eye and has a width of about 140° horizontally. The binocular field-of-view refers to the field-of-view of both eyes. However, it is not twice as large as the monocular field of vision because there is an overlapping area in the field-of-view that is perceived by both eyes simultaneously [PJW73]. As mentioned before, this is the reason why the blind spot of the eye does not have a negative effect on the perception characteristics of the human eye. The horizontal field-of-view is about 180° while the vertical field-of-view is about 130°, 60° up and 70° down [PJW73, Gol89]. The monocular fields-of-view and how they overlap to form the binocular field-of-view can be seen in Figure 2.4.

The distance between the eyes of approximately seven centimeters results in different image positions on the retinas of each eye, known as *binocular disparity*. To explain the term, binocular disparity means that an object near our eyes can project different images to each eye, thereby, unoccluded objects near our eyes can be seen differently by them. For example, from a finger that we hold in



(a) Left monocular vision. (b) Binocular vision. (c) Right monocular vision.

Figure 2.4: Monocular and binocular fields-of-view [PJW73, Kal01, SRJ11].

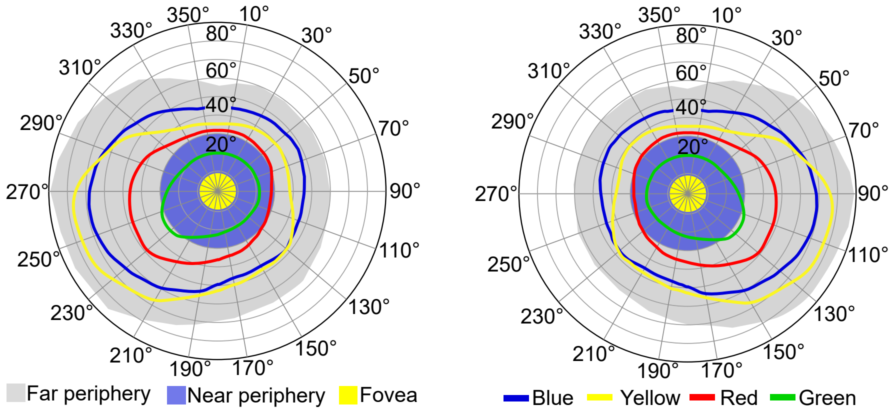
front of our face, our left eye may see parts of the finger nail, while our right eye may see the finger but not its nail. Our brain processes the individual images of our eyes via binocular fusion in such a way that a spatial impression of depth is created, which is called *stereoscopic vision*. With greater distance from objects, binocular disparity decreases. Depth perception from six meters resembles one-eyed vision. At this point, other characteristics are used to infer distance, such as relative size, shade, overlapping, and perspective [Kal01].

We must differentiate the visual field from the field-of-view. The visual field is measured without eye movement [Gol89]. Thus, the focus stays unchanged. For the field-of-view, however, we also consider eye movement, whereby the field-of-view can be shifted a maximum of 60° to the sides and 40° up and down [PJW73]. Any further displacement of our field-of-view is only possible by moving the head. We call the extended visual field that includes eye movements field-of-view or field of vision [Kal01].

2.1.3 Foveal and Peripheral Perception

The human visual field can be classified into two areas with different perception characteristics: foveal and peripheral perception. Here, we further distinguish between far and near periphery because in the near periphery almost all colors can still be perceived, including colors with higher resolution such as red, while the far periphery contains areas that have no color vision at all (see Figure 2.5). However, the term peripheral vision refers to both near and far periphery at the same time. For foveal vision, the literature further classifies three subareas: fovea, para-fovea and peri-fovea [Kal01].

Perception in the periphery can be differentiated from foveal perception by various factors (e.g., color, shape and depth perception [Kal01]). This is due to the distribution of the photoreceptors on the human retina (see Figure 2.3). This distribution applies to both photopic and scotopic vision. When looking at Figure 2.3, it becomes clear how rapidly the sharpness of our vision decreases



(a) Foveal and peripheral perception.

(b) Color perception.

Figure 2.5: Human visual perception [PJW73, Kal01]. *Best seen in color.*

as one approaches the edge of the eye. The cones are mostly distributed in the foveal area of the visual field. However, the human eye has 20 times more rods than cones. However, the rods are distributed over a larger area and often the output of several rods is connected to one ganglion cell.

In the work of Strasburger et al., the authors summarize the various strands of research on peripheral vision and relate them to theories of form perception [SRJ11]. A central topic of their paper is the recognition of characters in peripheral vision, with two different level of contrast, and the impact of surrounding contours, known as crowding. Further, the recognition of more complex stimuli, such as textures, faces, and scenes, reveals the substantial impact of mid-level visual and cognitive factors. They report that peripheral vision is limited regarding pattern categorization by a distinctly lower representational complexity and processing speed. However, utilizing peripheral vision allows us to use the high bandwidth that visual perception offers while at the same time not cluttering the central visual field with additional information.

2.2 Cognitive Maps

To understand how humans can locate objects in their environment or know where they currently are in this environment, it is fundamental to understand cognitive mapping and the research behind it.

> *"The cognitive map is how we know where we are"*, said John O' Keefe when explaining his Nobel Prize-awarded work[1].

---

[1] Interview with John O' Keefe. `www.youtube.com/watch?v=favAUcplFKI`, last retrieved April 21, 2020

In general, *cognitive maps* (sometimes referred to as *mental maps*) can be classified as a research field within spatial cognition, which basically contains all concepts that contribute to human spatial understanding [New02]. The field itself is an interdisciplinary field that can be found in environmental psychology but also in computer science (e.g., to program robots that are able to navigate in complex environments) or in social sciences (e.g., to understand how different cultures express spatial knowledge). In this thesis, when talking about cognitive maps we refer to a specific research field that is mainly investigated by researchers from psychology and neuroscience.

The idea of cognitive maps emerged with observations that Tolman made in 1948 [Tol48]. While observing rats in an experiment, he discovered that "something like a field map gets established in the rat's brain". Further, he highlighted that such maps come in multiple scales. For example, there are narrow strip maps and broad comprehensive maps. The idea that there are different grains of mapping was later extended by Montello [Mon93], who proposed that the human brain distinguishes four main categories of space: 1) figurale space, which can be compared to table-top space, 2) vista space, which is the space that can be perceived visually from a single location, 3) environmental space, which cannot be perceived from a single location (e.g., city), and 4) geographical space, which encompasses things such as countries and continents.

To describe what a cognitive map is, one can say that they are understood as a complex interplay between various types of cells that help to answer questions such as, "How far are two objects from each other?", or "How can one get from A to B?", or "What does the layout of a particular area or building look like?" A cognitive map is defined as a mental representation of the environment that surrounds us [Tve00]. For example, when living in a city, one has a cognitive map of that city to be able to successfully navigate from one place to another within this city. However, there is a level of complexity required before the human brain creates a cognitive map. In previous studies, it was found that cognitive maps represent areas that are larger than what can be seen with a single glance [Wag06]. For example, when one is in the library and looks at a specific book on the shelf, no cognitive map is required. However, when one is trying to find a book on a shelf that contains multiple areas (e.g., different research fields) or when one is asked to navigate through the library, a cognitive map is required. Therefore, in order to comprehend spatial structures like the spatial relationship between books on a shelf or the layout of the library, the human brain generates cognitive maps by combining information pieces that were collected from several views [She03]. As a result of this, cognitive maps become more detailed over time through addition of newly perceived information.

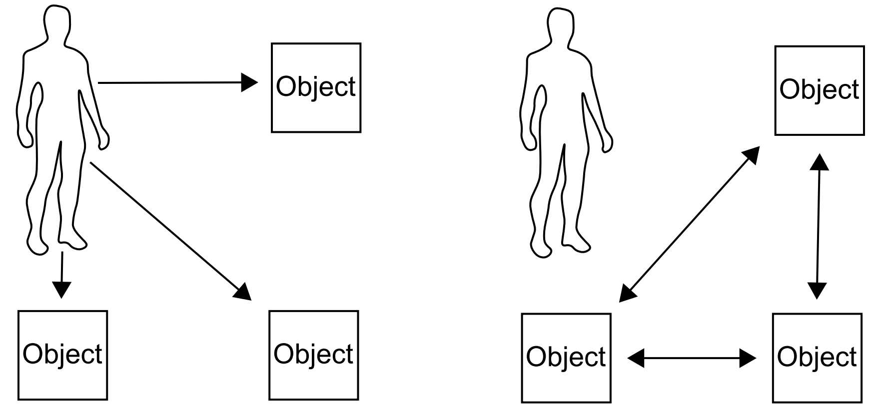
In the last decades, the field of neuroscience has gained a lot of knowledge regarding which parts of the brain are involved in creating cognitive maps and scientists have been able to specify which cells are active during these processes. In an experiment conducted by O'Keefe and Dostrovsky, the researchers found so-called *place cells* in the hippocampus of the brain that fire when one reaches a certain location, indicating cognitive maps are formed and located in the hippocampus of the human brain [OD71]. In a later study, it was found that the main motor responsible for creating those maps is curiosity [ON78]. The cognitive map for all kinds of environments is always a combination of different locations that are connected by certain rules that describe the direction and distances among them [ON78]. This has the advantage that when using a cognitive map to navigate to a specific location, one can find all possible routes to that location. After O'Keefe and Dostrovsky had discovered place cells, Taube et al. discovered *head-directional cells*, which are active depending on the direction in which a person currently looks [TMR90]. These head-directional cells are not influenced in any way by where a person is: only the head direction has an influence on their behavior. More recently, Hafting et al. [HFM$^+$05] discovered *grid cells* that fire when certain locations are visited after each other in a regular triangular way. The authors suggest that this may contribute to an understanding of self-motion. In summary, researchers from the field of neuroscience were able to identify cells responsible for the understanding of locations in space called place cells [OD71], cells responsible for understanding the directions in which the head faces called head-directional cells [TMR90], and cells that help one to understand how one moves through space called grid cells [HFM$^+$05]. However, to locate objects in our environment, additional cells are required. In 2009, Lever et al. found so-called *boundary vector cells* that help to encode the distance or direction to objects regarding the self [LBJ$^+$09]. Some years later, Knierim et al. discovered cells that indicate distance and direction in relation to another object [KND14]. With these cells, humans have two different ways to locate objects, relative to themselves or relative to other objects (more details in the following Subsection 2.2.1).

Furthermore, in previous studies, researchers have found evidence that there are differences in humans' spatial cognition abilities (e.g., [SM04]). They found that when dealing with spatial information, the performance is "dependent on specific characteristics of the learning situation." Furthermore, it seems that the differences in spatial cognition of individuals are correlated with their results on spatial tasks (e.g., for large scale environments) [FS06]. In general, studies showed that performance on spatial tasks is influenced by the way this information is stored in the human brain. For example, when people try to estimate the distance between two locations (distance estimation), they tend to shift these locations closer together if they belonged to the same semantic cluster (e.g., two places in the same part of the cities may be stated to be closer together than they actually are) [HM86]. In the case of direction estimation, when estimating the directional relationship between two spatial objects (e.g., two streets that intersect each other), people tend to describe the derived angles with a tendency towards simplified angles (e.g., in the experiment participants shifted the angles towards 90° [MB83]).

In this thesis, we investigate how to visually assist users who must locate objects outside of their fields-of-view. We understand the term locate as supporting users' cognitive processes with visual cues to help them understand object locations (cf. [Kit94]). However, in order to encode and communicate a location, it must be embedded into a stable reference system [HT66].

2.2.1 Frames of Reference

Spatial relationships are described using a *frame of reference*. Characteristics of a frame of reference include: origin of the coordinate system, orientation, and handedness (relation between axes). As seen in the previous section, the human brain has different types of cells that support two different ways of locating surrounding objects: relative to themselves [LBJ+09] or relative to other objects [KND14]. Decades before these cells were discovered, these two frames of reference had been investigated in various studies [How93]. They have been shown to be consistent with recent discoveries in neuroscience. The first frame of reference is called *egocentric* and refers to locating objects relative to the self. The second frame of reference is called *allocentric* (or also called *exocentric*) and refers to locating objects relative to each other. Both frames of reference are illustrated in Figure 2.6.



(a) Egocentric frame of reference. (b) Allocentric frame of reference.

Figure 2.6: Two different frames of reference (cf. [How93]).

The more often-used approach to locate objects in the environment independent of species and culture is to locate them in reference to the self [Haz83]. Furthermore, all scenarios that we introduced in the beginning of the thesis are experienced from an egocentric perspective (see Section 1.1). Therefore, in the following we will focus on the egocentric frame of reference, in which one's body is normally the center [Roe59]. However, there are four different egocentric frames of reference, ranging from a very specific point in our body to larger body parts: the station-point frame is associated with the nodal point of the eye, the retinocentric frame is associated with the retina, the head-centric frame is associated with the head, and the body-centric frame (torsocentric) is associated with the torso [How93]. In this thesis, we understand the egocentric perspective as the head-centric frame, which is commonly used. However, the exact reference point that humans use to relate to from an egocentric perspective is not clear and differs between studies [LH11].

In the past, several studies have been conducted to compare the egocentric and allocentric frames of reference. For example, in 1994, Milgram and Kishino gave a taxonomy of mixed reality presentation schemes ranging from egocentric to allocentric, suggesting the use of egocentric visualizations for local guidance [MK94]. One year later, Barfield et al. collected more evidence that local guidance is best supported by egocentric information [BRFI95]. In summary, we can say that the egocentric perspective is beneficial in most scenarios and has certain advantages over an allocentric view. For example, an egocentric view should be chosen when information has a relevant relationship with the position of the observer (e.g., if an object needs to be located in reference to the self).

2.3 Mixed Reality

In the history of computer technology, the ways in which humans interacted with computers fundamentally changed about every twenty years [Ben18]. In the start of the sixties, command line interfaces were used, with the keyboard as the input device. To interact with these interfaces, users were required to know some basic commands while fully relying on textual representations of the interface state. Twenty years later, in the early eighties, graphical user interfaces that used a mouse as the input device became more common. Graphical user interfaces have the advantage over command line interfaces that they do not require a vast amount of knowledge to be operated. Here, the mapping of a hardware device (mouse) to a pointer on the screen made it easier for novices to interact with computer interfaces. However, users cannot directly touch a button presented on the screen using a mouse. Therefore, twenty years later, natural user interfaces started to gain popularity. Here, users can directly interact with content presented on the screen, either with a pen or directly with the finger. However, in all these decades, users were restricted to a screen that functioned as a small window to the digital world. Nowadays, twenty years later, Mixed Reality (MR) interfaces allow one to jump through the window directly into the digital world, enabling novel ways of interaction. We therefore think that Mixed Reality technology has the potential to fundamentally change the way we interact with the digital world.

Nowadays, Mixed Reality devices are most known in two unique representations: Augmented Reality (AR) and Virtual Reality (VR). Both share the ability to alter humans perception of the world. Human perception is altered by either adding virtual objects to their existing environment (AR) or by creating a complete virtual environment (VR). Combined with a tracked head-mounted device (HMD), VR and AR allow one to naturally change their view based on head movement, even while mobile.

The term *Mixed Reality* was introduced by Milgram et al. in the mid-nineties [MK94]. Mixed Reality can be understood as a superordinate that contains Augmented Reality and Virtual Reality. Milgram et al. describe Mixed Reality as the "Reality-virtuality continuum", a continuous scale between virtuality and reality (see Figure 2.7). Everything between virtuality and reality is called Mixed Reality. In the paper by Milgram et al. [MK94], Virtual Reality was not directly called Mixed Reality. It was rather referred to as virtuality on the continuum. However, we argue that every virtuality that can be experienced by users needs some kind of manifestation in the real world. Otherwise, users are not able to interact with and experience this virtuality. For example, in a typical VR experience, physical objects such as the VR headset or controller are required. Furthermore, Microsoft, a pioneer in the area, refers to their own VR headsets as immersive Mixed Reality headsets[2]. For these reasons, we consider Virtual Reality to be part of Mixed Reality in this thesis.

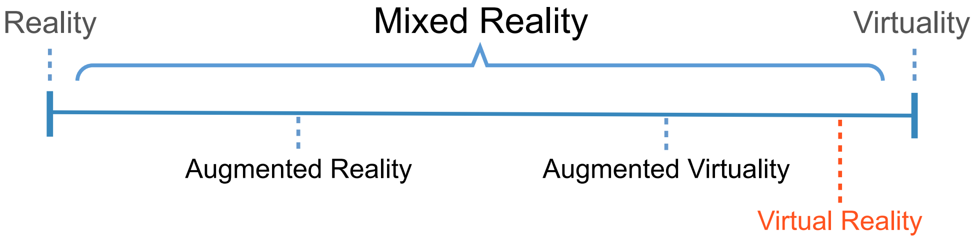


Figure 2.7: Reality-virtuality continuum by Milgram et al. [MK94] with Virtual Reality added to the continuum.

In this section, we will give a brief overview of the history of Augmented and Virtual Reality technologies and define the terms. Further, we will highlight the characteristics of both technologies. We will conclude this section with advantages of head-mounted devices, a form factor of both Augmented and Virtual Reality.

---

[2] Immersive Mixed Reality Headsets. `docs.microsoft.com/en-us/windows/mixed-reality/immersive-headset-hardware-details`, last retrieved April 21, 2020

2.3.1 Virtual Reality

The idea of being immersed into a virtual world has existed for many decades now and has been highlighted by well-known representations, such as the Holodeck in StarTrek in 1974[3]. The fundamental idea was first summarized by Ivan Sutherland in 1965 and was described as the ultimate display.

> *"The ultimate display would, of course, be a room within which the computer can control the existence of matter. A chair displayed in such a room would be good enough to sit in. Handcuffs displayed in such a room would be confining, and a bullet displayed in such a room would be fatal. With appropriate programming such a display could literally be the Wonderland into which Alice walked."* Ivan Sutherland (cf. [Sut65]).

However, the history of Virtual Reality goes way back to the middle of the 19th century, decades before Sutherland described his idea of the ultimate display. Back then, stereo viewers for still imagery were invented by Holmes and Brewster. These stereoscopes used two lenses with two slightly shifted photographs to create a stereoscopic effect. A century later, the first 3D cinemas introduced stereoscopic vision to a broader audience. However, more important for the development of current VR devices were the new optic lenses for larger fields-of-view introduced in the late seventies. These optics were called large expanse, extra perspective optics (LEEP Optics) and were created by Eric Howlett. With these lenses 90° direct field-of-view and 140° corneal field-of-view were possible. In the eighties, VR went beyond stimulating the human senses and started to focus on new input modalities, such as bringing the user's hands into the experience. The Electronic Visualization Laboratory developed the first data glove called Sayre Glove [SZ94]. It was combined with a LEEP-based head-mounted VR device and built upon fiber optic bend sensors to detect gestures performed by the user. In the nineties, the Cave Automatic Virtual Environment (CAVE) was presented by Cruz-Neira et al. [CNSD$^+$92]. CAVE consisted of three to six walls with stereo video projection that track the user's view port with infrared markers. Despite all the contributions to the field, the buzz around Virtual Reality technology began to fade, starting in the mid-ninties and lasting for more than a decade. However, with the introduction of the first smartphone (the iPhone[4]) in 2007, the slowed development of VR technology started to take off again. The reason for this is that the components used in modern smartphones, such as small displays, have drastically improved in resolution over the years. Nowadays, smartphones typically have a resolution around 300 pixels per inch. Another aspect is that smartphones came with improved sensors for detecting acceleration and orientation. Both technologies were developed for better smartphones but delivered exactly the components that were also required for a better VR experience. In

---

[3] Holodeck. `en.wikipedia.org/wiki/Holodeck`, last retrieved April 21, 2020

[4] iPhone. `en.wikipedia.org/wiki/IPhone`, last retrieved April 21, 2020

2012, Palmer Luckey initiated a Kickstarter campaign[5] to fund the development of a consumer VR headset. This was only possible because of the dropping hardware prices and hardware being available from the smartphone market. The campaign was successful and raised almost 2.5 million US dollars from around 10.000 backers. Later, in 2017, Google introduced Cardboard[6]. The device is basically a fold-out cardboard viewer that utilizes the user's smartphone to allow simple Virtual and video see-through Augmented Reality experiences (the latter will be explained in Subsection 2.3.2).

Different definitions for *Virtual Reality* exists. In 1992, Greenbaum defined it as “an alternate world filled with computer-generated images that respond to human movements” [Gre92]. In the same year, Coates defined it as “electronic simulations of environments experienced via head mounted eye goggles and wired clothing enabling the end user to interact in realistic three-dimensional situations” [Coa92]. As mentioned by Steuer, both definitions limit Virtual Reality to specific hardware, such as head-mounted eye-goggles, or to specific interactions, such as human movements [Ste95]. Therefore, we suggest using the definition by Pimental and Teixeira, which does not require specific hardware or concrete interaction, but rather focuses on immersion as the key aspect of Virtual Reality.

> *“An interactive, immersive experience generated by a computer”,* Pimental and Teixeira (cf. [PT93]).

This definition highlights the importance of immersion for a technology to be called Virtual Reality. Immersion is the extent to which the used technology delivers a vivid illusion of reality to the senses of a human being [PPPW97]. When users are immersed into Virtual Reality, they reach a state of consciousness that is referred to as presence and which can be assessed with the presence questionnaire [WS98]. Presence is a psychological sense of being in the virtual environment. When talking about Virtual Reality, it is mostly understood as a visual experience. Therefore, three key characteristics are required: 1) a 3D stereoscopic display, 2) a wide field-of-view, and 3) low-latency head tracking[7]. When these three key characteristics are fulfilled, the VR device provides a compelling immersive experience to the user. However, when the head tracking has a higher latency or when presented movement and perceived movement do not match (e.g., on a VR roller coaster) users can suffer from *motion* or *simulator sickness.* Simulator sickness can be assessed with the simulator sickness questionnaire [KLBL93] or the VR-specialized cybersickness questionnaire [AWM05]. Furthermore, currently available Virtual Reality hardware can be operated in two modes: seated or room scale mode. In the seated mode, users often have only

---

[5] Oculus Rift VR Kickstarter campaign. `www.kickstarter.com/projects/1523379957/oculus-rift-step-into-the-game`, last retrieved April 21, 2020

[6] Cardboard. `en.wikipedia.org/wiki/Google_Cardboard`, last retrieved April 21, 2020

[7] Mark Billinghurst - Introduction to Virtual Reality. `www.slideshare.net/marknb00/lecture1-introduction-to-virtual-reality`, last retrieved April 21, 2020

three degrees of freedom, meaning that only the head rotation influences the VR experience but head movement has no effect. In the room scale mode, users have all six degrees of freedom, where head rotation and movement are tracked by the Virtual Reality system.

Nowadays, there are four promising application domains in Virtual Reality that we would like to highlight. The first is training and education [Pso95]. Virtual Reality allows us to immerse users into realistic looking environments while at the same time avoiding any risk of injury that may come with some real-world environments (e.g., firefighter training with smoke and fire in the environment). Further, it can offer training possibilities in environments that are hard to simulate (e.g., speaking in front of a crowd) or very expensive to operate (e.g., docking a container ship in a harbor). Virtual Reality can also help one to relearn motor skills after a stroke [HKBL07]. The second application domain of VR is gaming and entertainment. Nowadays, there are more than fifty different VR games in the Steam online store that users actively play[8]. The third application domain is telepresence. Here, an expert or operator can be at a different physical location than the problem or task is. For example, in a manufacturing plant, when there is an issue with one of the machines, an expert from the producer of that machine can virtually visit the plant and discuss the problem in front of a digital representation with other workers on site. The last and fourth application domain for VR is to test certain Augmented Reality applications. Various studies have shown the usefulness and validity of testing AR applications in VR [EBM$^+$97, LBBH10, RGCH16]. The advantage here is that, with no real world influences, VR can generate a clean experimental setup for investigating specific influencing factors. In addition, the virtual and real worlds can be presented with the same resolution and brightness to the user, making the digital content harder to distinguish from the real world. Further, testing with VR headsets is helpful when testing how well AR content is designed for upcoming AR hardware with fidelity higher than that of current AR devices.

2.3.2 Augmented Reality

Augmented Reality (AR) allows one to overlay digital content onto the real world in order to alter the perception of it. This digital content can be explored by different kinds of devices (e.g., smartphones or AR glasses). Nowadays, smartphones are widely available and used for exploring AR content, while AR glasses can thus far only be found on the enterprise market (e.g., Microsoft Hololens, MagicLeap). There have been some studies by which the degree of fidelity (e.g., rendering quality [KSF10, SJS14], refresh rate [LBS$^+$16], and registration accuracy [SFS$^+$07, MSS11], etc.) was not high enough for the user to feel immersed using previous AR devices. Thus, the presence of the real world overshadowed the virtual. However, the recent emergence of increasingly capable mobile hardware

---

[8] Steam VR games. `store.steampowered.com/vr`, last retrieved April 21, 2020

and the optimization of tracking solutions have enormously enhanced the fidelity of AR content. This boost in performance has enabled a variety of new applications, location-based games being one of them (e.g., Pokémon Go[9] or Ingress[10]).

Many would agree that Augmented Reality was born in 1968. Three years after Sutherland introduced his idea of the ultimate display, he developed the first head-mounted display, which was not actually worn on the head of the user [Sut68]. The was device called "The Sword of Damocles" and was so heavy that it had to be mounted on the ceiling (see Figure 2.8). The user's environment was extended with a simple wireframe of a cube, presented by a see-through display.

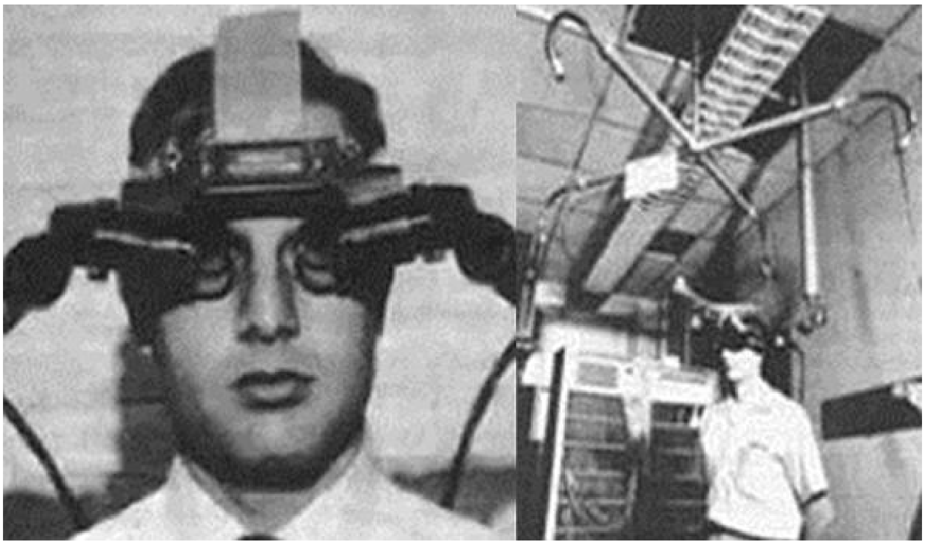


Figure 2.8: Ivan Sutherland's "The Sword of Damocles". Image source: [Sut68]

The term Augmented Reality was first introduced in 1992 by Tom Caudall and David Mizell [CM92]. They developed a head-up display for Boeing that was to be used in aircraft construction. One of the first academic papers mentioning Augmented Reality (KARMA - Knowledge-Based Augmented Reality [FMS93]) was published by Feiner et al. in 1993. His research group developed the first mobile augmented reality system (MARS) in 1997 [BCL+15]. The system consists of a head-worn display and a backpack carried on the back. GPS was used to display information about buildings on the Columbia University campus in New York [FMHW97]. It took some time before these research prototypes found their way into commercial products. They first came in the form of smart glasses, such as the Google Glass[11] released in 2013. In more recent iterations, products such as the Microsoft Hololens[12] offered even more pervasive experiences by allowing

---

[9] Pokémon Go. `www.pokemon.com`, last retrieved April 21, 2020

[10] Ingress. `www.ingress.com`, last retrieved April 21, 2020

[11] Google Glass. `en.wikipedia.org/wiki/Google_Glass`, last retrieved April 21, 2020

[12] Microsoft Hololens. `en.wikipedia.org/wiki/Microsoft_HoloLens`, last retrieved April 21, 2020

users to pin digital content onto real world objects. The Hololens is considered to be the first fully AR-capable headset on the market with respect to the definition of Augmented Reality.

Ronald Azuma was the first to define the *Augmented Reality* system (AR system) in a survey paper in which he also summarized the state-of-the-art in Augmented Reality in 1997 [Azu97]. His definition of an AR system contained three characteristic properties: 1) AR combines real and virtual elements (users can see both at the same time), 2) AR is interactive in realtime (users can interact with virtual content), and 3) AR is registered in 3D (virtual objects have their fixed reference points in space). According to this definition, the first consumer products such as the Google Glass are not AR devices. These devices are therefore considered to be smart glasses only.

Augmented Reality describes a human-machine interaction in which the environment or reality of the user is enhanced by digital information. However, this computer-supported information does not have to be presented via head-mounted device, it can be presented at many different positions. A classification of the display placement of head-mounted devices compared to other augmented reality devices can be seen in Figure 2.9.

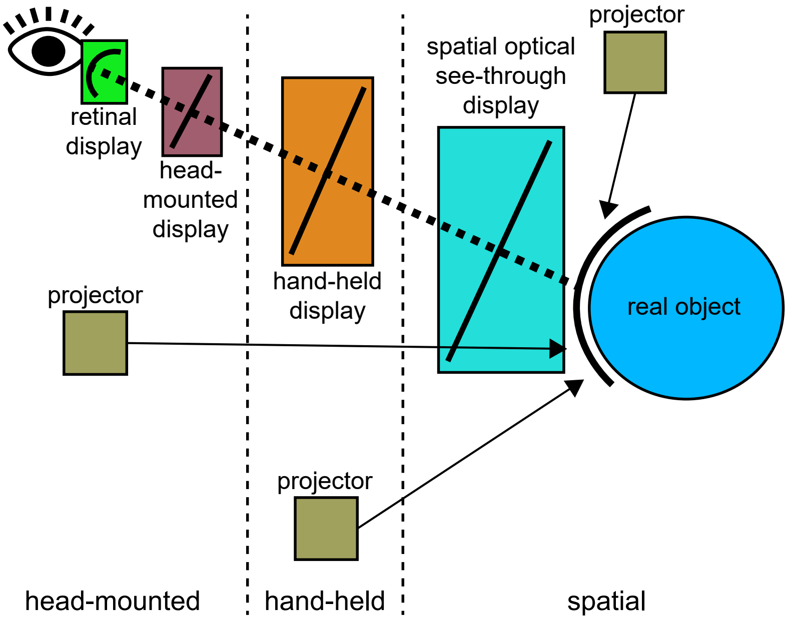


Figure 2.9: Overview of device technologies and their display placement in relation to the real object and the observer (cf. [BR06]).

A more detailed description of head-mounted devices and their advantages can be found in the next subsection (see Subsection 2.3.3). For Augmented Reality there are two types of head-mounted devices: video and optical see-through AR devices. An optical see-through AR device uses a transparent display to present the virtual content to the user, while a video see-through AR device uses a non-transparent display that has the video feed of a camera streamed onto it and thereby, creates the see-through experience. The resulting effect is referred to as the *window effect* (also called the *magic lens metaphor*) [BSP^{+}93, ZSd13]. To our knowledge, for all current optical see-through AR devices, the real world overshadows the virtual. Furthermore, the technology used in optical see-through devices does not currently allow wider fields-of-view. Therefore, video see-through AR devices are useful when prototyping AR experiences that would be limited by the technical constraints of current optical see-through devices. However, as with Virtual Reality, video see-through AR devices can cause motion sickness.

Content in AR can be presented using different frames of reference [FMHS93]. It can be 1) screen-stabilized, which means content is always in the same screen position (e.g., for text, menus or status information), 2) body-stabilized, which means content is always in the same position around the body (e.g., for tool palettes or virtual watches), 3) world-stabilized, which means content is always in the same place in the real world (e.g., labels added to physical objects or added virtual objects such as new furniture), or 4) tracker-stabilized, which means content is bound to a tracker target or marker (e.g., to allow the position of the virtual object to be changed by moving the tracker).

In the last years, there has been a lot of improvement in the field of Augmented Reality. Recent advances in AR technology allow us to use it for a variety of applications. We want to highlight five promising application domains. First, Augmented Reality can be used for gaming and entertainment, where users can have a large screen everywhere in their flat to watch Netflix or to play multiplayer games in large areas [HSHS05, OBS16]. Second, it can be used in education to improve learning results [KS02, DDM09]. Third, it can support users during navigation tasks, with arrows projected directly onto the street or for unobtrusive navigation for pedestrians [RC17]. Fourth, it can enable new ways to collaborate, where a person from another location can join with a virtual avatar to work on a problem [RGK^{+}16, LKN^{+}16]. Fifth, AR has great potential to support users during tasks on site (e.g., driving, ship docking or fire fighting). Here, it can help users to identify relevant objects faster (e.g. in an assembly task [TOBM03] or for order picking [SK08]).

2.3.3 Head-mounted Displays

A *head-mounted display*, HMD for short, is a visual output device worn on the head that can display images directly in front of the user's eyes. Both Augmented and Virtual Realities can be experienced in a head-mounted device. A head-mounted device has several advantages over other form factors. In both AR and VR, virtual content can be placed all around the user's head and can be explored in a natural way by looking around. A head-mounted device normally includes sensors to detect head movement and can adjust the view port to the user accordingly. This is a very natural way of interacting with the surrounding virtual content and is referred to as the *AR browsing technique* [FM14].

Head-mounted devices can come as standalone devices for AR as well as for VR, which allow users to stay mobile and walk around in their environments. Further, due to the device's being head-mounted, the hands of users remain free for other tasks. This is a very important advantage in several scenarios, such as fire fighting or driving. In some scenarios, users already wear helmets with visors (e.g., firefighter), and in those cases helmets can be extended to show the virtual content. Furthermore, head-mounted devices have various applications (e.g. in aviation, engineering or video games) and have been proven to be more useful than other form factor devices for many scenarios [CR06].

3 Related Work

This chapter reviews recent approaches that are related to our problem of visually encoding spatial information about out-of-view objects in Mixed Reality. Specifically, it identifies and reviews three approaches that serve as the foundation and inspiration for solving the problem of objects receding from view. The first approach focuses on off-screen visualization techniques that address the related problem of content receding from view on small-screen devices (see Section 3.1). The second approach focuses on extending the field-of-view of current head-mounted devices to present additional information to the user (see Section 3.2). The third and last section focuses on techniques that guide the user's attention to objects in the environment (see Section 3.3). Each of these approaches will be presented in a dedicated section, where each section concludes with potential strategies of how to utilize the presented related work to address the problem of objects receding from users' fields-of-view.

3.1 Off-Screen Visualization Techniques

Just as objects recede from users' fields-of-view, content can also recede from small-screen devices such as smartphones or tablets. Thereby, users may overlook certain information because it is not shown on the screen and without any assistance it is a hard task for users to locate off-screen content.

> *"The limited viewport size of mobile devices requires that users continuously acquire information that lies beyond the edge of the screen",* (cf. [EACI11]).

To address this problem, researchers have developed various off-screen visualization techniques capable of visualizing locations of off-screen objects on small-screen devices. These techniques can be classified into three main approaches: Overview+detail, Focus+context, and Contextual views [GBGI08, CKB09]. In the *Overview+detail* approach, a miniature map of the surroundings is shown in addition to the detailed view. For example, while navigating traffic, the current location of the user is presented on the screen with a high zoom level to ensure clarity of the street names, and at the same time a miniature map is presented to the user on the bottom right to show the progress of the overall route. Unlike the Overview+detail approach, both *Focus+context* and *Contextual views*, overlay the current detailed view (focus) with context information along the borders of the screen. For the navigation example, both would still result in the current location of the user being presented on the screen with a high zoom level. However, instead of a miniature map in the corner, the start and end point of the route would be visualized along the borders of the screen. The transition between focus and context is how the approaches differ. In the Focus+context approach, the transition is soft (e.g. fisheye-views that convey a distorted view [SB94]), while for Contextual views the transition is hard (e.g. arrows pointing into off-screen space [BCG06]). In the navigation example a soft transition would mean that the rest of the route gets compressed into the space along the borders of the screen. A hard transition would mean that an arrow for example points in the direction of the start and end point, possibly with additional labels to quantify distance. Contextual views are derived from the Focus+context approach and often referred to as Focus+context approaches as well [GI07].

In this section, we distinguish between visualization techniques based on the dimensions they support. For example, for map applications, off-screen techniques that can point into 2D space are sufficient. In this case, the content can be understood as a (large) 2D plane where the screen shows only a (small) snippet of that plane to the user. Here, off-screen content can be to the left, right, top or bottom of the current snippet shown on the screen. We refer to techniques that only support the 2D space as 2D off-screen visualization techniques (see Subsection 3.1.1). If the content is not limited to a 2D plane but instead can be all around the user (e.g., an Augmented Reality app that adds digital information to paintings in a museum), then we refer to these techniques as 3D off-screen visualization techniques (see Subsection 3.1.2). Additionally, we investigate off-screen techniques that are used in computer games (see Subsection 3.1.3). Those approaches have not been evaluated scientifically. However, they can serve as a source of inspiration when developing new visualization techniques for head-mounted devices. A list of all off-screen visualization techniques that are discussed in this section is shown in Table 3.1.

3.1.1 2D Off-Screen Visualization Techniques

To overcome the limitations of smaller screen sizes and to enable larger content areas, graphical interfaces that support zooming, paging, scrolling, or panning have been proposed in the past [BH94, Joh95]. While in general they may be a valid strategy for presenting more information to the user, these interfaces introduce new problems, such as content receding from view [EACI11]. To address this problem and to support users in understanding where they are in the greater picture, the Overview+detail approach has been investigated in several studies (e.g., [BWI90, NS00, HF01, CKB09]). An example of a well-known Overview+detail interface is the online map service Google Maps[1]. Here, an inset in the bottom right corner allows the user to understand the context of the detailed region. However, a disadvantage of the Overview+detail approach is the cognitive load required to mentally integrate all views into one overall understanding of the scene [CKB09, TSL$^+$11]. Contextual information along the borders however, as in the Focus+context and Contextual views approaches is more in line with the human frame of reference [JHPR11]. However, compared to Overview+detail, Focus+context and Contextual views are not able to keep the proportions of 3D

---

[1] Google Maps. `www.google.com/maps`, last retrieved April 21, 2020

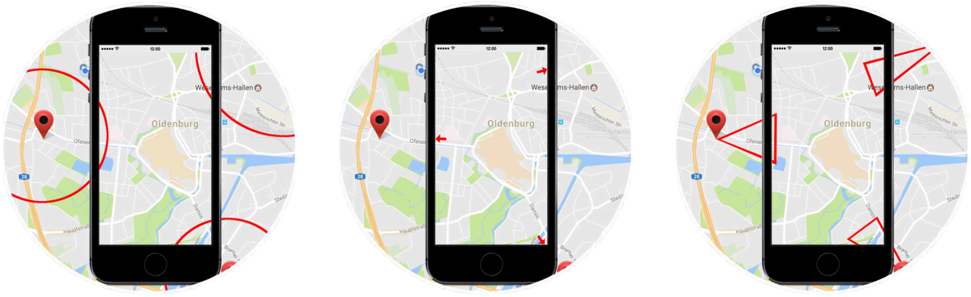
Table 3.1: Comparison of presented off-screen visualization techniques. The visualization techniques that are used in computer games are not listed here.

Name	**Dimension**	**Approach**
Map Windows [BWI90]	2D	Overview+detail
Snap-together [NS00]	2D	Overview+detail
Fish-eye View [CKB09]	2D	Focus+context
Continuous Fisheye [SZG$^+$96]	2D	Focus+context
Fishnet [BLH04]	2D	Focus+context
City Lights [ZMG$^+$03]	2D	Contextual Views
EdgeRadar [GI07]	2D	Contextual Views
Halo [BR03]	2D	Contextual Views
Arrows [BCG06, HPB10]	2D	Contextual Views
Wedge [GBGI08]	2D	Contextual Views
Worlds in Miniature [SCP95, BHF02]	3D	Overview+detail
3D Mobile Maps [OEN09]	3D	Overview+detail
3D-Arrows [CB04, BC07, SHB10]	3D	Contextual Views
SidebARs [SH13]	3D	Contextual Views
Aroundplot [JHPR11]	3D	Contextual Views
3D-Halo Circle [TSL$^+$11]	3D	Contextual Views
3D-Halo Projection [TSL$^+$11]	3D	Contextual Views
Halo3D [PMN18]	3D	Contextual Views
Halo360 [PMN19]	3D	Contextual Views

space since they compress all information along the borders [JFR17] (e.g., two objects that are off-screen to the right have to be projected along the right screen border, regardless of how far away from each other they are). Therefore, it is especially hard to preserve topological information for off-screen objects when using these approaches.

A different approach is Focus+context. Here, fisheye views are the most prominent example of off-screen visualization techniques that utilize this approach (e.g., [SZG$^+$96, BLH04, CKB09]). To understand how they work, it helps to imagine a magnifying glass lying on top of a map. The position where the magnifying glass lies is zoomed in while the rest of the map is shown without any magnification. As with the Overview+detail approach, Focus+context focuses on presenting all available context information. However, when users are required to quickly and accurately locate off-screen objects on small-screen devices, Contextual views was found to perform best [IGY06, BCG06, BC11].

Contextual Views is derived from the Focus+detail approach and is able to present contextual information without distortion by conveying only information about relevant objects. Additionally, it overlays only the screen borders with contextual information, leaving the focus uncluttered by information [BCG06, GBGI08, CKB09]. One of the first Contextual views was presented by Zellweger et al. [ZMG$^+$03], who provided contextual information along the borders. However, in this instance, users found it difficult to guess the actual positions of the off-screen objects. Therefore, Halo was suggested as an improvement [BR03]. This technique uses circles drawn with their centers around the off-screen objects, which cuts the border of the screen slightly (see Figure 3.1a). Even though only an arc of the full circle is visible on the screen, users are still able to mentally complete the shape to locate the center of the Halo. This effect is referred to as *amodal completion* or *amodal perception* [Mic91, EZ93]. However, a problem of Halo is cluttering, which is the accumulation of many Halos in corners. Therefore, one year later, Gustafson et al. [GI07] presented EdgeRadar. Their results show that users can locate off-screen objects with their technique more accurately than with Halo. Further, several studies compared Halo with Arrow approaches [BCG06, HPB10], where Arrows with fixed sizes performed worse and scaled arrows performed better. The number of visible objects also has a high impact on the performance. An example of Arrows is shown in Figure 3.1b. Another new technique that tries to avoid cluttering is Wedge [GBGI08]. This technique also utilizes amodal completion, relying on users to mentally complete the shape that is only partially shown. Instead of circles like the ones used for Halo, they propose using isosceles triangles, which use less space (see Figure 3.1c). Here, the tip of the triangle points to the off-screen object while the two remaining corners appear on screen. Gustafson et el. showed that users were significantly more accurate in locating off-screen objects with Wedge than with Halo [GBGI08]. However, none of these techniques have been investigated in head-mounted devices, while they look promising to be able to cue direction to out-of-view objects as well.



(a) Halo [BR03]. (b) Arrows [BCG06]. (c) Wedge [GBGI08].

Figure 3.1: Three different 2D off-screen visualization techniques.

3.1.2 3D Off-Screen Visualization Techniques

In many cases, the information presented to the user is not restricted to the two dimensions of the display. In mobile Augmented Reality, the real environment of the user is annotated (e.g., [WDH09]). In virtual environments (e.g., [CN08]), a third dimension is used to present the information to users. In such cases, the problem of locating off-screen objects becomes even harder to solve because information can potentially be behind the user. A first technique to address this problem in virtual environments is Worlds in Miniature. The technique was suggested in the mid-nineties by Stoakley et al. [SCP95]. Worlds in Miniature follows the Overview+detail approach and presents a representation of the user's environment (world) in the center of the screen. The same approach has also been investigated in mobile Augmented Reality [BHF02, OEN09]. However, presenting an overview of the scene requires additional screen space, which is fairly limited on small-screen devices. Therefore, it may be advantageous to present the contextual information along the borders of the screen.

To locate off-screen objects in three-dimensional space, different Arrow-based techniques have been proposed [CB04, BC07, SHB10]. All these techniques use three-dimensional arrows that point in the directions of the off-screen objects. Furthermore, all techniques are inspired by the Contextual Views approach. However, the contextual information is not presented along the borders of the screen; instead, the 3D Arrows are always presented on the ground in front of the user. Thereby, visual clutter is introduced to the center of the screen. In both of their papers, Burigat and Chittaro investigated the technique for navigation purposes in virtual environments on a desktop computer, showing that 3D Arrows outperform 2D techniques [CB04, BC07]. Schinke et al. applied the technique to hand-held Augmented Reality for visualizing points of interest (e.g., to support tourists). They compared 3D Arrows with an Overview+detail approach and found that using 3D Arrows results in faster, more precise responses. This highlights the advantages of Contextual Views over Overview+detail approaches, even though the information was not presented along the borders of the screen.

A different approach is AroundPlot from Jo et al. [JHPR11]. This technique is based on EdgeRadar [GI07] and presents off-screen objects as dots along the border of the screen. Basically, a mapping from 3D spherical coordinates to a 2D orthogonal fisheye is used to compress the information about off-screen objects onto the user's screen. To adjust to a user's current region of interest, the technique uses dynamic magnification, meaning the area along the borders of the screen can increase or decrease to show information in higher resolution based on the phone's rotation. The problem of AroundPlot is the same corner-density problem found with Halo [BR03] and EdgeRadar [GI07]. Further, the positive and negative effects of dynamic magnification must be examined in further studies. Another approach is presented by Siu and Herskovic [SH13], who propose SidebARs for improving awareness of off-screen objects. Like AroundPlot, it presents contextual information along the borders of the screen. However, this technique uses icons instead of dots to represent the off-screen objects and the surrounding objects are presented either on the left or right border of the screen. The technique is designed for hand-held devices and was evaluated with firefighters, who found it to be a promising technique for visualizing surrounding off-screen objects. However, the technique can be problematic when objects appear on a different height level compared to the user's current position. In addition, the technique does not convey any distance information about the off-screen objects.

Different approaches have investigated ways of transferring existing off-screen visualization techniques to 3D space in order to convey the exact location of and distance to an off-screen object. Therefore, Trapp et al. proposed 3D Halo Circle and 3D Halo Projection [TSL^{+}11] and implemented them to locate points of interest in virtual environments on hand-held devices. Both techniques are based on Halo [BR03] and draw either multiple circles around a building in parallel to the ground, reaching partially into the user's screen (3D Halo Circle), or a single circle around a building in parallel to the user's screen, again partially reaching into that screen (3D Halo Projection). Here, only 3D Halo Circle can visualize the 3D distance to the off-screen object, while adding a lot of visual clutter to the screen. As with all proposed techniques based on 3D Arrows [CB04, BC07, SHB10], the contextual information is presented utilizing the complete screen of the user. More recent work has explored ways of transferring existing off-screen visualization techniques to mobile Augmented Reality applications. Halo3D and Halo360 by Perea et al. are two examples [PMN18, PMN19]. However, their work is limited to hand-held devices and its visualizations of off-screen objects in front of the user are indistinguishable from those off-screen objects behind the user. This is a problem for scenarios in which off-screen objects are distributed 360° around the user (e.g., sightseeing or gaming). Still, it is quite useful for their described scenario to visualize points of interests in industrial facilities where off-screen objects are mostly located in front of the user.

3.1.3 Off-screen Techniques in Computer Games

Aside from in research, off-screen visualizations are frequently used in computer games. An early example of such a visualization technique can be found in the 2D game Tecmo Bowl (1987)[2]. Here, simplified arrows are used to point to the off-screen content, where off-screen content is either on the left or the right side of the screen. Similarly, in modern games, arrows are often used to point to off-screen content. For example, in Rocket League (2015)[3], an auto ball game, an 3D arrow is used for pointing towards the ball. When comparing several computer games, one can observe that Contextual views is used when off-screen content must be located precisely and quickly. However, besides arrows, there are also more complex visualizations used when not only the direction is required but also the distance, when the topological information between multiple off-screen objects must be preserved, or when users are required to understand the movement of such off-screen objects over time. Therefore, in 3D games like X-

---

[2] Tecmo Bowl. `en.wikipedia.org/wiki/Tecmo_Bowl`, last retrieved April 21, 2020

[3] Rocket League. `www.rocketleague.com`, last retrieved April 21, 2020

Wing (1993)[4] or the newer Eve: Valkyrie (2016)[5], a radar-like visualization is used that gives the user an overview of the surrounding area. These more complex visualizations follow the Overview+detail approach. However, since the computer games themselves do not offer any systematic user evaluations, we consider them more as a source of inspiration for our visualization techniques than as baseline measures. Further, some techniques used in games have been studied in previous research and have therefore been listed in previous sections.

3.1.4 Summary

In this section, we presented several approaches and techniques for visualizing the locations of off-screen objects in two and three-dimensional space. We saw that the problem of objects receding from view on small-screen devices is very similar to the problem of objects receding from a user's field-of-view. Therefore, it may be promising to use existing off-screen visualization techniques in head-mounted devices to locate out-of-view objects. However, all proposed off-screen visualization techniques are evaluated using small-screen devices. None of them have been adapted for the visualization of out-of-view objects on head-mounted devices. Therefore, it seems promising to adapt and test these techniques for head-mounted devices. To do so, following the Overview+detail approach may be helpful in situations in which users have to understand the locations of several out-of-view objects in relation to one other, or when those objects are not at fixed positions. However, a disadvantage of this approach is the cognitive load required to mentally integrate the overview with the current view of the user [CKB09, TSL^{+}11]. Since the Focus+context and Contextual Views approaches are more in line with the human frame of reference [JHPR11], they may have the advantage that users are able to locate out-of-view objects faster and more precisely than with the Overview+detail approach. Further, the Contextual Views and Focus+context approaches applied to head-mounted devices give the advantage that the user's focus remains undisturbed because the context information is presented along the borders. This is especially important when users must perform others tasks besides locating out-of-view objects that require visual focus. However, for the proposed visualization techniques in 3D space, the advantage of an uncluttered center of the screen is not there anymore, since most techniques use it to present information to the user (e.g., 3D Arrows [CB04, BC07, SHB10] or 3D Halo [TSL^{+}11]). Therefore, it may be a better approach to adapt 2D off-screen visualization techniques to head-mounted devices, in order to better adapt them to the needs of such devices.

---

[4] X-Wing. `en.wikipedia.org/wiki/Star_Wars:_X-Wing`, last retrieved April 21, 2020

[5] Eve: Valkyrie. `en.wikipedia.org/wiki/Eve:_Valkyrie`, last retrieved April 21, 2020

3.2 Head-mounted Peripheral Displays

In the past, different kinds of peripheral displays have been developed. Some of them were created with the aim of extending the field-of-view of current head-mounted displays, others for presenting information (e.g., notifications) to a user. As we saw in the previous section, our human visual system perceives information presented in the periphery in a different way compared to information that appears in the foveal area. These differences must be taken into account when developing peripheral displays or presenting information in the periphery.

3.2.1 Extending the Field of View of Head-mounted Devices

Like the problem of objects receding from our human field-of-view, content can also recede from head-mounted devices that offer too-small fields-of-view. To overcome this problem, different technical approaches have been investigated in the past. However, simply extending the fields-of-view of current devices does not make sense, because we perceive information presented in the periphery with lower resolution. An overview of all techniques that are discussed in this subsection can be found in Table 3.2.

Table 3.2: Comparison of proposed techniques for extending the field-of-view of head-mounted Augmented and Virtual Reality devices.

Name	**Category**	**Technology**
HMD with two LCDs [BJK05]	Regular glasses	LCD displays
Smart Glasses [NK16]	Smart glasses	LED array
Light Modulators [MF13]	AR	Light Modulators
Pinlight Display [MLR+14]	AR	LCD + Point Lights
SparseLightVR/AR [XB16]	AR + VR	LED array
Compressed Fisheye [OWK+14]	VR	Lenses
Peripheral Blurred Images [YM16]	VR	Lenses
Wide FOV Lenses [RTH16]	VR	Lenses

In previous work, most approaches that aim to extend the field-of-view focus on one of the following three device categories: 1) optical see-through Augmented Reality, 2) video see-through Augmented Reality, or 3) head-mounted Virtual Reality. In this subsection, we will mainly focus on technical approaches that extend the fields-of-view of head-mounted devices towards the human field-of-view of about 180° horizontally. It should be mentioned that there are approaches that aim to extend the field-of-view even further (e.g., with additional cameras that capture the scene behind the user and present it as a picture in picture mode on the head-mounted device [ALM+12, ALMM14, FHNI14]). These approaches do not encode the locations of specific out-of-view objects, but rather present the user with additional views. This requires them to mentally integrate the views and thus has a negative impact on workload. Furthermore, due to the location of the camera, these views often have a spatial offset compared to the human eyes due to the location of the camera that users perceive as unnatural [ALM+12, LKKP19].

To extend the field-of-view of Augmented Reality devices is technically more challenging because the used display technology must ensure that users can still perceive their environments. As discussed earlier (see Section 2.3), there are two approaches to achieve this: optical or video see-through Augmented Reality. However, optical see-through AR devices are especially challenging because the display technology has to be transparent enough to allow for an optical see-through experience. Therefore, most approaches do not simply extend the field-of-view by attaching additional displays. They rather aim to develop a new form of optical see-through device that allows a wide field-of-view from the beginning. Here, two solutions have been proposed. One solution is to use stacked spatial light modulator layers that are positioned closer than the typical eye accommodation distance to achieve a wider field-of-view [MF13]. Another solution uses semi-transparent LCD panels with point light sources behind them to project the information from the LCD panels directly in the user's eyes [MLR+14]. Both solutions are promising. However, further research is required to bring these technologies to the market.

A different approach of extending the field-of-view of optical see-through AR devices is to use non-transparent display technology in the periphery. The idea is that users do not always need to rely on perceiving their environment in their periphery, so in these cases foveal vision is sufficient. An example of this was suggested by Xiao et al. [XB16]. They investigated low-resolution LED arrays surrounding a central high-resolution display for Augmented and Virtual Reality devices. In a user study they showed that sparse peripheral displays are useful in conveying peripheral information and that they improve situational awareness.

Video see-through devices are very similar to Virtual Reality devices in that video see-through is mainly based on VR devices with additional cameras to capture the environment. Here, different approaches have been suggested for extending the fields-of-view of such devices. For example, Orlosky et al. [OWK+14] present a method for extending the limited FOV of HMDs using a fisheye view that compresses the peripheral view. They found that users could detect 62.2% of objects distributed in 180° space, while they could detect 89.7% with the naked eye. This, however, works for environments in 180° on a smaller FOV, and has a negative effect on perception of detected objects since smaller objects can disappear due to the compression. A different method was proposed by Yamada and Manabe [YM16]. They use two different lenses with different magnifications. Here, the second lens is attached around the first lens and uses much higher compression to present more information. However, while their prototype was usable for extending the FOV, two levels of magnification means the foveal FOV is clear while the periphery is milky. This lack of detail can lead to one overlooking information. In the long term, the goal is to develop so-called *Frezel* lenses that support the human field-of-view. An example is presented by Rakkolainen et al., whose optical lenses can cover a user's complete FOV [RTH16]. Their paper is a work in progress and display functionality is not yet implemented. It is more a technical concept to describe how to utilize the human FOV in a HMD. However, VR devices with fields-of-view like the human field-of-view have been suggested recently (e.g., Primax 5K XR[6] with a 200° field-of-view).

Besides extending the field-of-view, different approaches exist that focus on a peripheral display only. Instead of extending an existing screen, they aim to deliver information to the user independent of any other display. In that sense, they do not literally extend the field-of-view, but rather present contextual information to the user. Such approaches can be referred to as regular or smart glasses. An example for this is the early work presented by Beak et al. [BJK05]. They showed that even low-resolution displays can feasibly to extend the field-of-view, when being attached on the left and right sides of regular glasses. Another example is the peripheral vision glasses prototype presented by Nakuo and Kunze [NK16]. Here, instead of LCD displays, the authors use LED arrays that support higher brightness and allow display of low-resolution patterns in the periphery.

3.2.2 Utilizing the Human Visual Periphery

In the foregoing subsection, we summarized different technical approaches for extending the field-of-view of head-mounted devices. In this subsection, we focus on the different application domains of such peripheral displays and the effects on humans regarding physiology and visual perception. An overview of all investigated research directions presented in this subsection can be found in Table 3.3.

Table 3.3: Comparison of investigated research directions for peripheral displays.

Name	**Category**	**Description**
EyeQ [CIP+06]	Notification	Unobtrusive notifications
NotifEye [LV14]	Notification	Social media notifications
"Smart" Ski Helmet [NFEL17]	Notification	Warnings for skiers
Multimodal Alarms [CHB17]	Notification	Intensive Care Unit Alarms
SparseLightVR/AR [XB16]	Physiology	Simulator sickness
Optic Flow [BWB+13]	Physiology	Self-motion
AmbiGlasses [PHF+12]	Perception	Direction accuracy
Visual Language [LDRR+16]	Perception	Pattern perceptibly

We identified the three most relevant research directions for peripheral displays: 1) notifications, warnings and alarms, 2) effects of peripheral displays on human physiology, and 3) perception characteristics of peripheral displays.

---

[6] Primax 5k XR. `www.pimax.com`, last retrieved April 21, 2020

In the beginning of this century, researchers started to explore the presentation of information and notifications with peripheral displays [DGYG+04, KM06]. However, these peripheral displays are not understood as head-mounted displays. They are rather understood as ambient displays, physical devices placed in the user's environment. Here, one could argue that those devices are not necessarily peripheral since the user can directly look at them. Therefore, it has been suggested to have those displays mounted on the head in some form, to ensure they stay in the user's periphery [CIP+06, LV14]. An early example of such a peripheral display is Eye-q, which was developed by Constanza et al. [CIP+06]. The idea was to enable subtle, discreet and unobtrusive notification to the user. The advantage over audio notifications is in having non-disruptive and distraction-free stimuli, both for nearby people and the user. Further, they showed in an experiment that the notifications could be designed to meet specific levels of visibility and disruption for the wearer. However, although Eye-q can notify the user, it does not support interactions with notifications and cannot shift attention towards elements other than the notifications themselves. For discrete interaction with eyewear in public, NotifEye was introduced by Lucero et al. [LV14]. They investigated social network notifications and possible user interaction with these notifications in depth. Peripheral displays were also investigated in specific applications, such as alarms in critical care units [CHB17] and warnings for skiers [NFEL17], which shows their usefulness in real environments. Although notifications do not necessarily have a spatial attribute, they show that the peripheral vision can be utilized to successfully show perceivable information to a user.

Peripheral displays also have an influence on the user's physiology. For example, it is well-known that head-mounted devices with a mismatch between the presented motion of the virtual content and the perceived motion potentially lead to simulator sickness [GB08]. Interestingly, in the paper from Xiao et al. [XB16], the authors showed that peripheral LED arrays can create a more immersive experience and help to reduce motion sickness. Furthermore, it has been shown that perception of self-motion can be altered by peripheral displays [BWB+13]. Here, Bruder et al. found that peripheral displays have the potential to make a user perceive self-motion as faster or slower than it really is.

When using head-mounted peripheral displays (e.g., to cue direction to objects outside the user's field-of-view), it is important to know how accurately users perceive the presented information. Therefore, AmbiGlasses were introduced by Poppinga et al. [PHF+12]. They are a prototype that uses twelve LEDs to illuminate the periphery of the user's field-of-view. In their user study, they showed that participants were able to locate the correct LED with 71% accuracy and estimate the rough location of the LED with 92% accuracy. However, the positioning of the LEDs in the AmbiGlasses prototype was chosen based on the glasses to which they were attached. Here, the oval shape of those glasses may have negatively influenced the perception accuracy of those LEDs. Besides single LEDs, Luyten et al. developed an easily perceivable visual language for near-eye out-of-focus displays [LDRR+16]. Their approach is based on the idea that motion in the periphery can be perceived especially well. They found that having simple shapes and a small set of colors is important for improving perception and comprehension of what is being shown on such displays. Further, their findings showed that a usable visual language can be developed by making clever use of orientation and meaningful motion.

3.2.3 Summary

To summarize, different technologies have been proposed to extend the fields-of-view of head-mounted Augmented or Virtual Reality devices. Because of the biological characteristics of the human visual field, no high resolution is required. Thus, LED arrays seem to be a simple and low-cost solution to augment human vision. Further, LED arrays can be used for Augmented as well as Virtual Reality devices and have the potential to counter motion sickness problems [XB16]. Most importantly, they can be utilized to encode information with spatial attributes (e.g., direction) and may therefore be a good approach for visually cueing to out-of-view objects. To our knowledge, related work has only investigated disadvantageous form factors that do not allow one to cue in any direction [PHF+12]. Therefore, we think it would be a good idea to investigate the suitability of peripheral displays for pointing to objects receding from view.

3.3 Attention Guidance

When a user fails to perceive certain out-of-view objects, one may want to guide the attention of the user to those objects in order to avoid any negative consequences, depending on the scenario.

> *“Of all things that your eyes see at any instant, you are conscious of only those few to which you direct your attention”,* (cf. [HTP07]).

Changing where users direct their visual attention is possible using visual cues. According to Posner and Petersen [PP90], there are three phases necessary for shifting the user’s attention: 1) disengage the current target, 2) shift attention between stimuli, and 3) engage the new target. In this section, we summarize visual and non-visual guiding techniques. A list of all these techniques can be found in Table 3.4.

3.3.1 Visual Cues

Different visual cues have been developed to guide users to out-of-view objects or to address related problems. In this subsection, we structure the related work

Table 3.4: Comparison of different techniques to guide to out-of-view objects.

Name	**Modality**	**Cue location**
Shift Driver's Attention [TSLB05]	Visual	Head-up Display
Auto Pilot [LCH$^+$17]	Visual	Smartphone
Camera Pose [SRK$^+$10]	Visual	Smartphone
Vibrating Icons [MJUO17]	Visual	Smartphone
Smartwatch [Lyo16]	Visual	Smartwatch
On-body Cues [HLSH09]	Visual	On-body locations
Visual Effects [DGD17]	Visual	Head-mounted VR
Attention Funnel [BTOX06]	Visual	Head-mounted AR
ParaFrustum [SEO$^+$14]	Visual	Head-mounted AR
Support Order Picking [SK08]	Visual	Head-mounted AR
AR assistance systems [RP17]	Visual	Head-mounted AR
Cue Control [BMNN19]	Audio	Head-mounted VR
(Audio-) Visual Cues [LBS$^+$17]	Audio + Visual	Environment
HapticBelt [PKB10]	Haptic	Belt
HapticHead [KR17]	Haptic	Head-mounted device
Sensory Augmentation [KKP16]	Haptic + Audio	Head-mounted device
Haptic + Visual Cues [SLG$^+$18]	Haptic + Visual	Head-mounted device

with regard to their placement. The visual cues range from cues that are placed in the environment and are perceivable by multiple users to cues that are presented on head-mounted devices directly to one specific user.

An early work on visually guiding attention, was presented by Tönnis et al., who investigated how to direct a car driver's attention with a head-up display placed at a fixed position in front of the driver [TSLB05]. In their work, they differed between targets in the driver's frame of reference and targets in an exocentric frame of reference. As with head-up displays, visual cues for shifting the user's attention have also been investigated with smartphones. For example, Lin et al. [LCH$^+$17] investigated guiding gaze in 360° videos on smartphones. They presented two approaches for guiding attention in 360° videos: Auto Pilot (bringing the target to the viewer) and Visual Guidance (indicating the direction of the target). They showed that if increased head movement is necessary (e.g., when following a sports video), users preferred Auto Pilot. Furthermore, users found it frustrating to shift to a target that was already gone or a part of a scene that had already taken place (e.g., a tackling in soccer). This highlights the need for accurate visualization of out-of-view objects. Besides applications in the field of Virtual Reality, there have been other approaches for guiding attention in mobile Augmented Reality. For example, approaches have been developed to help users to inspect a scene from a specific camera pose [SRK$^+$10] or to guide the user's attention to specific points of interests outside of the smartphone's current field-of-view [MJUO17].

Furthermore, researchers have investigated different body locations for presenting visual cues to the user. For example, Lyons investigated different visual parameters for drawing the user's attention to information on a wrist-worn smartwatch [Lyo16]. The author found statistically significant differences for size and frequency, which were positively correlated with length of reaction time. Other work compared various on-body positions for placement of visual cues. For example, Harrison et al. investigated the use of wearable visual cues on seven different body locations between the shoulders and feet and measured the respective reaction times [HLSH09]. They measured average reaction times over 15 seconds for all investigated body locations. However, they found that the response times were faster when a user observed the state change of the light. Here it may make sense to place those visual cues close to the human visual field to ensure that the state change is perceived in any situation.

Therefore, different visual cues have been investigated for head-mounted devices, some of them specifically for Virtual Reality devices. For example, in the work "Attention guidance for immersive video content in head-mounted displays" by Daneau et al. [DGD17], the authors investigated four visual effects to implicitly drive the user's attention. All effects relied on manipulation of the displayed content (e.g., black-out or grey-scale every part except the part to which the attention should be guided). Further, the effects can only guide attention to parts that are already visible. Their results show that implicit attentional guidance remains challenging. Others have been investigating cues for head-mounted Augmented Reality devices.

Augmented Reality allows one to overlay digital content on the real world, in order to alter perception of it [Azu97]. In the last decades, researchers have focused on improving tracking, interaction, and display technologies [ZDB08] to create more immersive experiences. However, due to the low degree of fidelity, which is influenced by rendering quality [SJS14] and refresh rate [LBS$^+$16], users can still distinguish between digital content and the real world [PS09]. While this inhibits full immersion, it is helpful for shifting the attention of the user to physical objects in the environment. One technique that makes use of visual cues in head-mounted AR to shift the user's attention is Attention Funnel [BTOX06]. This technique uses several boxes in a row, starting at the viewers' head position and ending at the object's position. Here, the user is guided to the location of the out-of-view object. This is similar to ParaFrustum [SEO$^+$14], where users are guided to a specific pose. However, both approaches add a lot of visual information to the screen and the location of the object out-of-view can only be perceived after some initial movements of the user. In order to draw attention to an object with the help of an Augmented Reality application and thus support picking, Schwerdtfeger and Klinker evaluated various Augmented Reality visualizations. The different visualizations were arrow-based (based on [FMS93]), frame-based, and tunnel-based (based on [BTOX06]) visualizations. In the arrow-based visualization, an arrow points to the object (here a tray from which the user should take an object). The visualization is complemented by a compass-like arrow that is displayed about 30 cm in front of the user and points to the next relevant object. In frame-based visualization, the respective box is marked by a rectangular frame. This is also supplemented by the meta-visualization of the compass arrow. In tunnel-based visualization, a tunnel of rectangles points to the box from which a part is to be picked [SK08]. Similarly, Renner and Pfeiffer investigated different peripheral and in-view attention guidance techniques for augmented reality applications [RP17]. They specifically examined attentional guiding in assembly and order-picking tasks. Therefore, they investigated several methods, such as Attention Funnel and 2D arrow, and compared them to their newly-developed method called spherical wave-based guidance (swave). All techniques utilized the current user's gaze with eye-tracking. Altogether, the arrow-based guidance technique turned out to be fastest and best-rated by the study participants. This supports the conclusion that visual guidance is most effective in the periphery of the user and that simple techniques such as Arrows seem to be promising for guiding user attention.

3.3.2 Non-Visual Cues

Guiding users to out-of-view objects cannot only be achieved with visual cues. Different modalities, such as auditory and haptic cues, have been investigated for use in attentional guidance. For example, in the work by Bala et al. [BMNN19], the authors investigated auditory cues for guiding user attention in a 360° setup. They showed that cueing was most effective when using only music or a combination of music and diegetic effects. In our previous work, we looked at visual and auditory cues [LBS^{+}17]. Here, we showed that adding a sound cue results in faster response times. Furthermore, participants reacted more quickly to LEDs that faded in over time. Since audio is already integrated in all head-mounted Virtual and Augmented Reality devices, it requires no additional effort to be implemented. Therefore, we think that in some cases it may be beneficial to add auditory cues for locating objects out of view.

Besides using auditory cues, several approaches have suggested the use of haptic cues at different positions. For example, cues could be placed on a tactile torso to present the location of several people around [PKB10]. Here, the authors found that all investigated vibration patterns are useful for encoding direction towards another person while playing a computer game. Wearing the belt resulted in improved the situational awareness. Another work was presented by Kaul and Rohs, who looked into tactile cues presented to the user's head [KR17]. The authors implemented a vibrotactile grid around the head for 3D guidance in VR and AR. Their results indicate that visual feedback works with less error and higher reaction time, but the tactile cues may still be useful in situations in which visual feedback cannot be used (e.g., for people with low or no vision). Other work investigated the combination of haptic cues with other modalities. For example, in our previous work we combined haptic and visual cues [SLG^{+}18]. Here, we found that tactile cues led to faster arousal times than visual cues, whereas the attention shift speed for visual cues was faster than it was for tactile cues. Different works have compared the haptic modality to others. For example, in the work “Head-Mounted Sensory Augmentation Device: Comparing Haptic and Audio Modality,” the authors Kerdegarhi et al. showed that the haptic modality leads to significantly lower route deviation in navigation when compared to audio feedback [KKP16].

From related work, we know about the existence of cross-modal effects [HS08]. These effects have implications for multi-modal cueing to guide attention. Many of those effects are related to the combination of visual and auditory stimuli (e.g., the *ventriloquism effect* [Ber99] or the *McGurk effect* [MM76]). However, most of these effects can be avoided when both cues are presented at the same location and the transported information does not differ between the cues [DS00].

3.3.3 Summary

Several visual cues have been designed to guide the attention of the user to out-of-view objects. Here, it is important that users can perceive the cues. Therefore, visual cues presented close to the eyes of users are the most effective ones. Furthermore, visual cues have been compared to haptic and auditory cues and have been found to be the most accurate in terms of encoding 3D location. However, it may be useful to combine visual cues with auditory ones, since there are promising relevant results and the hardware is already there for most head-mounted Augmented and Virtual Reality devices. Furthermore, it has been shown that the combination of a visual cue with a cue of a different modality can improve response time. However, when guiding the attention of a user to an out-of-view object, one can only cue the location of a single object at a time. This may make sense in situations in which we are certain that this specific out-of-view object is the one to which we should direct our attention (e.g., to avoid a possible collision while driving a car). However, in most situations, missing information or overly complex environments make it hard to decide if the user’s attention should be guided to a specific object. Here, it can even make a situation worse if the attention is guided to an object that is currently less important than another one. Therefore, guidance techniques to out-of-view objects only make sense when the application domain is well understood and it is clear that the user is required to quickly locate one specific out-of-view object.

4 Conceptual Design

In this chapter, we describe the conceptual design of this thesis. In the first section, we define the problem space that we want to address, including a definition of out-of-view objects (see Section 4.1). Afterwards, in the next two sections, we describe two different strategies for locating these objects and identify the area in the human field-of-view that we utilize to present the visual cues to users (see Section 4.2 and 4.3). In the next section, we take a closer look at head-mounted Mixed Reality, the technology used to display the visual cues (see Section 4.4). At the end of this chapter, we list the challenges of this thesis and describe the research questions that will be addressed in this thesis (see Section 4.5 and 4.6).

4.1 Problem Space

In the introduction, we described different scenarios in which failure to perceive out-of-view objects can have fatal consequences (see Section 1.1). However, instead of narrowing down the scope of this thesis to objects receding from view in one specific scenario, we rather want to investigate the fundamental problem that these scenarios share and thereby contribute to solving the problem of objects receding from view for a broader range of scenarios. Therefore, in this section, we will first define the fundamental problem of objects receding from view and then describe the problem space and the included properties that will be the subject of the research in this thesis.

The problem that all described scenarios share is the restricted human field-of-view. Due to biological factors, humans can only perceive parts of their environments at once, with everything else hidden out of view. The human field-of-view extends to about 180° horizontally and to about 130° vertically [PJW73, Gol89]. Everything outside this field-of-view is impossible for one to visually perceive with the naked eye. In other words, objects in the environment can be within (*in-view objects*) or outside the human field-of-view (*out-of-view objects*). However, objects are not designated as in-view or out-of-view objects; rather, they can switch between being in-view or out-of-view. This can happen when the object changes its position or the human observer moves (e.g., turning of head, body or both).

When analyzing different scenarios in the introduction, we found that objects outside our fields-of-view can have fatal consequences if not perceived. The fundamental problem that these scenarios share is that users cannot locate objects outside their fields-of-view and may even be unaware of them. Thereby, an out-of-view object can cause a serious incident, simply because its location is unknown to an observer (e.g., a cyclist or pedestrian not visible to a car driver that turns). For this reason, we want to develop and evaluate visual cues that help users to locate these out-of-view objects and prevent fatal consequences that may have happened otherwise.

Therefore, we first want to describe the problem space and the properties that should be investigated. Out-of-view objects can have multiple properties, with different combinations of properties relevant in different scenarios. However, we argue that locating these objects and the properties that come with this task are shared among all scenarios, and are therefore the most critical to examine. Furthermore, these properties have the strongest benefit in terms of *situation awareness*. Following Endsley's definition, this requires "the perception of elements relevant in the current situation" as the first of three levels that must be reached in order to attain situation awareness [End95]. In other words, without the ability to perceive elements relevant in the current situation (level 1: perception), it is impossible to comprehend the situation (level 2: comprehension) and successfully predict the outcome of any action (level 3: prediction). As a result, we focus on locating out-of-view objects in this thesis, but this requires certain characteristics to be fulfilled. For instance, some out-of-view objects are located on the same height level as the user (e.g., different road users in a traffic encounter). In other scenarios, these objects can be anywhere in 3D space (e.g., in aviation scenarios or in a burning building where people that need help are on different floors). Furthermore, depending on the situation, there can be one relevant out-of-view object or multiple such objects. For example, in the traffic encounter, there is usually just one important out-of-view object (because usually only one "object" has the right of way). In the ship docking scenario, however, up to six tugboats may be involved. However, it is unlikely that all six tugboats are outside the pilot's field-of-view and six is the reported upper limit. Therefore, in most situations, five out-of-view objects may be the more realistic maximum number in this scenario. Furthermore, in the firefighting scenario, it is possible that even more objects may be involved; eight or more objects at the same time is not unusual. Aside from the number of relevant objects, it is also important to consider if the objects stay at the same location or not. Here, the out-of-view objects can stay at the same locations and not move over time (e.g., in the firefighting scenario) or they can change their locations (e.g., in the traffic encounter, ship docking, and aviation scenarios).

There are other properties that may be relevant in specific scenarios, such as size, form, type, or name of relevant out-of-view objects. However, we will focus on developing visual cues that help in locating one or more relevant out-of-view objects at fixed or changing locations. We choose this focus because, based on our analyzed scenarios, we think the properties related to locating these objects are the most relevant ones that the scenarios share and which need to be investigated first. However, in future work, these visual cues can be advanced further to encode additional properties that may be required to successfully support more scenarios.

4.2 Strategies for Locating Out-of-View Objects

In this section, we describe and analyze different strategies that humans can use to locate out-of-view objects. In the background chapter, we identified two fundamental frames of reference that humans use to locate objects in their environments (see Section 2.2). There is the egocentric frame of reference, which refers to locating objects relative to the self, and the allocentric frame of reference (or exocentric frame of reference), which refers to locating objects relative to one other. However, when users must locate objects in their environments quickly, an egocentric visualization may be the better choice because it visualizes the location of an object relative to the user, considering head and body rotation as well as the user's current position. This differs from an allocentric visualization, in which users must first identify their current position and orientation (e.g., in cases in which the visualization does not automatically adapt to the user's orientation). We think using an egocentric visualization has certain advantages over an allocentric visualization for locating out-of-view objects when users are part of the environment. This is backed up by related work. Milgram and Kishino also suggested using an egocentric visualization for local guidance [MK94], while Barfield proved with experimental results that local guidance is supported best by egocentric information [BRFI95]. Furthermore, the more often used approach for locating objects in the environment independent of species and culture is to locate them with regard to the self [Haz83]. All scenarios that we introduced in the beginning are experienced from an egocentric perspective and contain the user as part of the environment (see Section 1.1).

A key characteristic of a frame of reference is the origin of the coordinate system. For an egocentric frame of reference, the origin of the coordinate system can be located at different locations and therefore, different egocentric frames of reference exist. Howard describes four principal egocentric frames of reference: 1) body-centric frame (torsocentric) associated with the torso, 2) head-centric frame associated with the head, 3) retinocentric frame associated with the retina, and 4) station-point frame associated with the nodal point of the eye [How93]. However, humans have a natural reference point as the origin of their egocentric coordinate system that is somewhere in the brain, but the exact location is still unclear today [LH11]. Therefore, when we talk about the egocentric frame of reference, we refer to the head-centric frame of reference. Unlike the body-centric frame of reference, it considers head rotation. It also does not have two possible origins, unlike the retinocentric and station-point frames of reference, which have one for each eye.

To define different strategies for locating out-of-view objects from an egocentric perspective, we first need to understand how the human brain locates objects. As described in the section about cognitive maps in the background (see Section 2.2), Lever et al. discovered so-called boundary vector cells that help to encode the distance or direction to objects with regard to the self [LBJ$^+$09]. Several decades earlier, Howard said that the fundamental geometrical concepts for localizing objects are “direction and distance in relation to the location and orientation of the reference point” [HT66]. Therefore, to locate an object, a user needs to know the direction and distance relative to oneself (see Figure 4.1).

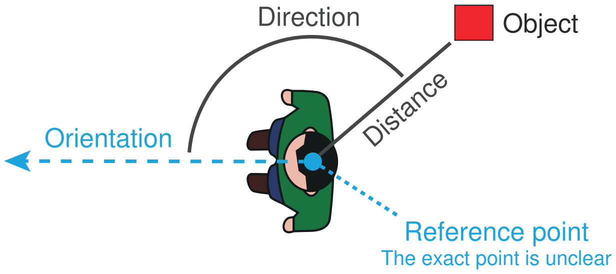
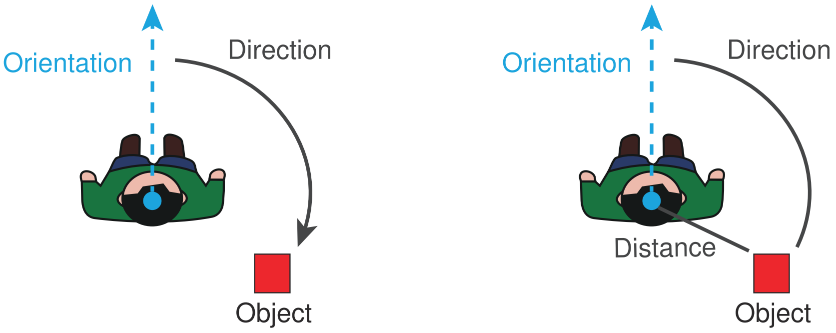


Figure 4.1: To locate an out-of-view object, one needs to know the direction and distance to that object relative to oneself (the exact reference point that humans use to refer to themselves is not clear and differs between studies [LH11]).

Based on the fact that humans utilize direction and distance to understand the location of an object, we can derive two different visualization strategies to help them locate out-of-view objects: 1) we can visualize the direction of the out-of-view object and require the user to turn their head in that direction to perceive the distance to the object by looking at it (see Figure 4.2a), and 2) we can visualize the direction of and distance to the out-of-view object, making it unnecessary for the user to turn their head to locate the object (see Figure 4.2b).



(a) Visualizing direction. (b) Visualizing direction and distance.

Figure 4.2: Two different strategies for locating out-of-view objects that are investigated in this thesis.

Both strategies have their advantages and disadvantages that make each of them useful for a certain selection of scenarios. For example, if users want to locate multiple out-of-view objects, it makes sense to visualize their directions and distances because direction only would result in higher physical demand to locate these objects due to the head movement necessary to perceive the distance information. Furthermore, the distance information often cannot be perceived because the out-of-view objects are also occluded and thus the distance must be visualized. Examples are the ship docking scenario in which containers block the view of the tugboats, or the firefighting scenario in which dense smoke renders everything visually unperceivable. However, visualizing direction alone can be sufficient for multiple objects too if it is enough to know from which direction the out-of-view objects are coming, or if only one object that is not occluded needs to be located. An example of the latter is the traffic encounter, in which typically a maximum of one out-of-view object needs to be located at a single time.

4.3 Peripheral Visualization

Human visual perception dominates the other human senses [SB14] and offers a wide range of methods that help in locating objects in the environment (e.g., stereoscopic vision or perception of depth cues). Furthermore, vision offers the highest bandwidth of information compared to the other senses [New96]. For these reasons, we utilize human visual perception to present visual cues to the user for locating out-of-view objects. The visual cues are presented on a display mounted on the head of the user and thereby allow control of where in the human field-of-view the information is presented (see Section 4.4).

The human field-of-view covers about 180° horizontally and about 130° vertically [PJW73, Gol89]. However, only a small area in the center of the human field-of-view (fovea) can be perceived very sharply (see Section 2.1), while visual acuity declines towards the border of the field-of-view. Everything outside of our fovea is referred to as periphery. The periphery of our vision perceives information with a lower resolution and has the main purpose of guiding our visual attention. Therefore, when visual cues are presented in the periphery, a user can shift their gaze to perceive a visual cue in more detail at any time. Furthermore, it allows foveal vision to stay uncluttered by information when users rotate their heads to look in a different direction instead of shifting their gaze to the periphery. In other words, the visualization and its visual cues should be in the peripheral field-of-view when the user is looking straight ahead. This gives users the control to decide when to focus on a visual cue and when not to do so. The visual cues should also not be so far away that gazing at them is uncomfortable, but also not so close that it is impossible to focus on the environment instead. A factor influencing this is the field-of-view of the used head-mounted device, which may be restricted by a too-small display (see Section 4.4). The visualization area in which we present our visual cues is shown in Figure 4.3. We think that the radial visualization area (blue) is the most suitable area for displaying information to the user. It considers human color perception, leaves the foveal area uncluttered, and requires equally long eye movements (saccades) from the center in each direction to perceive the visual cues in more detail. Furthermore, it lies within binocular vision, allowing movement of both eyes to any point in that visualization area.

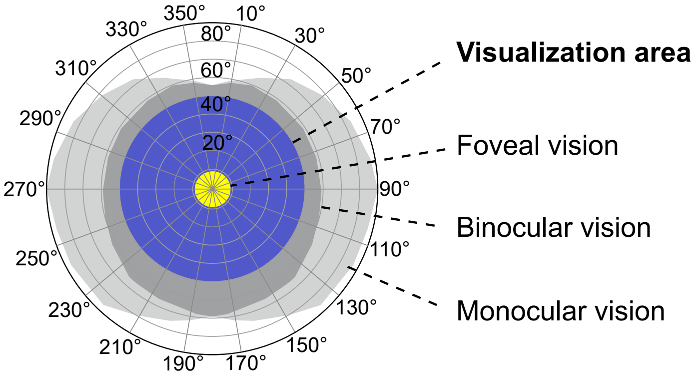


Figure 4.3: The peripheral visualization area (blue; radial area from 8° to 50°) in which the visual cues are presented to the user. *Best seen in color.*

4.4 Head-mounted Mixed Reality

In recent years, Augmented Reality (AR) and Virtual Reality (VR) technologies have experienced a sustained upswing (e.g., in navigation [OPF$^+$16], in gaming [OBS16], or for localization [HJDB16]). The fundamental idea of these technologies is to alternate perceived reality by augmenting or virtualizing it. Experienced in a head-mounted device, users can use such technologies hands-free and while mobile. This has advantages in many spatial working environments where machines must be operated by hand (e.g. emergency rooms [SH13]) or in situations in which the user is moving. Augmented and Virtual Reality can both be summarized by the term Mixed Reality (see Section 2.3) and the technology is promising for presenting visual cues in the periphery of the user.

However, current Mixed Reality devices suffer from limited fields-of-view, which amplifies the problem of objects receding from view. Here, optical see-through Augmented Reality devices often have very limited fields-of-view (e.g., the Microsoft Hololens[1] version 1 has a field-of-view of 30° horizontally and 17° vertically), while Virtual Reality devices have larger fields-of-view (e.g., the Oculus Rift[2] and HTC Vive[3] have fields-of-view between 90° and 110° diagonally). However, compared to the human field-of-view, it is still smaller, and therefore even more objects recede from view (see Figure 4.4).

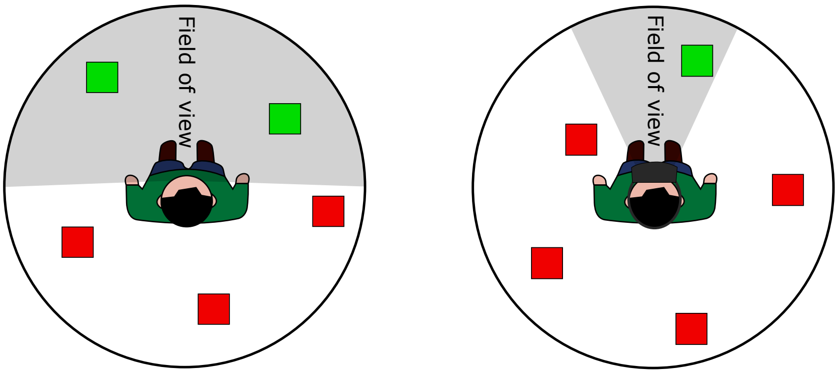
For Augmented Reality, out-of-view objects can be either virtual objects (e.g., opponents or enemies in multi-player games [PKB10]) or real objects in the sur-

---

[1] Microsoft Hololens. `www.microsoft.com/en-us/hololens`, last retrieved April 21, 2020

[2] Oculus Rift. `www.oculus.com/rift`, last retrieved April 21, 2020

[3] HTC Vive. `www.vive.com/us`, last retrieved April 21, 2020



(a) Natural field-of-view [Kal01]. (b) Device field-of-view.

Figure 4.4: Natural and device field-of-view in comparison. *Best seen in color.*

rounding environment (e.g., points of interest during sightseeing [HB10]). For Virtual Reality, however, only virtual objects exist (or at least are presented by the display of the VR headset). Therefore, in Augmented Reality, two fields-of-view exist: the device field-of-view and the human field-of-view. For virtual objects, we will always use the device field-of-view as reference because otherwise virtual objects may be theoretically in the human field-of-view but not yet visible on the screen. They could then be classified as in-view objects and therefore not be presented in the visualization technique. For real objects, we will use the human field-of-view.

For our implementations, we develop prototypes with Arduino[4] (hardware), the Google Cardboard[5] (video see-through Augmented and Virtual Reality), the Microsoft Hololens (Augmented Reality), the Oculus Rift (Virtual Reality), and the HTC Vive (Virtual Reality) since these devices were state-of-the-art at the time of use. Our visualization techniques are developed using Unity3D[6], as it supports both Augmented and Virtual Reality devices. For tracking we used marker detection with Vuforia[7]. AR applications for the Hololens were developed using the Windows Mixed Reality Toolkit[8] and VR applications were developed using SteamVR[9].

---

[4] Arduino. `www.arduino.cc`, last retrieved April 21, 2020

[5] Google Cardboard. `www.arvr.google.com/cardboard`, last retrieved April 21, 2020.

[6] Unity3D. `www.unity3d.com`, last retrieved April 21, 2020

[7] Vuforia. `www.developer.vuforia.com`, last retrieved April 21, 2020

[8] Windows Mixed Reality Toolkit. `www.github.com/microsoft/MixedRealityToolkit-Unity`, last retrieved April 21, 2020

[9] SteamVR Unity3D Plugin. `www.valvesoftware.github.io/steamvr_unity_plugin`, last retrieved April 21, 2020

4.5 Challenges

Developing visual cues in Mixed Reality that help in locating out-of-view objects is challenging. In this section, we identified three major challenges that must be addressed to develop these peripheral cues: 1) perceivable visual cues, 2) limited fields-of-view, and 3) mental demand.

4.5.1 Perceivable Visual Cues

Our approach utilizes human peripheral vision to present visual cues to the user that help in locating out-of-view objects. However, human visual perception is limited in the periphery and certain characteristics must be considered when designing perceivable visual cues. For example, one must consider that peripheral vision is less sharp, limits perception of forms and shapes, and cannot perceive color very well (see Section 2.1). To avoid having unperceivable visual cues, we should ensure that we use a good color contrast and that the visual cues are large enough. Here, a visual cue presented farther away from the center of our vision must be larger to remain perceivable. Furthermore, we exclude night vision from our design process because it is fundamentally different due to the distribution of the rods and cones and the fact that the rods in the periphery take over under low light at night. Another challenge could be with 3D visual cues, in that the three dimensions might not be perceived well, because the shape is too far in the periphery and too close to the eyes, for the three dimensions of the visual cue to be perceived. These aspects need to be kept in mind when designing the visual cues.

4.5.2 Limited Fields-of-View

Building Mixed Reality devices is technologically challenging, which explains why all current existing head-mounted Augmented and Virtual Reality devices suffer from having too-small fields-of-view. For comparison, the field-of-view of the Hololens version 1 is more than ten times smaller than the human field-of-view [XB16]. Therefore, it may not always be possible to present peripheral visual cues with the same distance to the center of the field-of-view, because the device field-of-view simply does not allow it. In such cases we have two options: 1) we can present the visual cues closer to the center of the visual field, or 2) we extend the field-of-view of the Mixed Reality device with peripheral light displays. In the first case, we may add more visual clutter to the screen and should therefore measure in how far this affects the user. In the second case, we may suffer from lower display resolution on the extended screen and must therefore design simpler cues to help in locating out-of-view objects. However, we can also replace the Mixed Reality device with another one that has a larger field-of-view. For example, we can use Virtual Reality or video see-through Augmented Reality to design a visualization technique for optical see-through Augmented Reality, and thereby test for upcoming Augmented Reality devices that are not yet available.

4.5.3 Mental Capacities

The last major challenge of the topic addressed in this thesis is the mental demand required to understand the visual cues and locate the out-of-view objects. How much information about out-of-view objects can be processed? Can one comprehend the direction and distance to multiple out-of-view objects all at once? It could also be that visualizing both direction and distance is already too much information. Therefore, we use a NASA Raw TLX questionnaire in most of our studies to quantify the mental demand of the task [Har06]. Thereby, we can analyze whether there are enough mental capacities available to present even more information about out-of-view objects if necessary, such as form, size, and type. Furthermore, we explained that an egocentric visualization works best in terms of quickly locating objects out of view. However, when these objects change position and we visualize the direction and distance to these objects, then an allocentric visualization may be more suitable for understanding the movement of these objects outside of the user's field of view.

4.6 Research Questions

The analysis of the related work, the different strategies for locating out-of-view objects, and the technological constraints led to three key research questions. In the following, we list the research questions addressed in this thesis and describe our general approach to solving them. Each research question will be addressed in a following chapter.

RQ1 To what extent can existing off-screen visualization techniques be adapted to cue the direction to out-of-view objects in Mixed Reality?

RQ2 How can Mixed Reality devices with small fields-of-view be extended to present directional cues to out-of-view objects?

RQ3 In what way can the directions and distances of out-of-view objects be visualized for moving or non-moving objects in Mixed Reality?

4.6.1 RQ1: To what extent can existing off-screen visualization techniques be adapted to cue the direction to out-of-view objects in Mixed Reality?

With the first research question, we investigate possible solutions for visual cues that present directions to out-of-view objects (strategy: visualizing direction; see Figure 4.2a). Interestingly, just as objects recede from view, content can also recede from small-screen devices such as smartphones. Therefore, researchers have developed various off-screen visualization techniques to show location of off-screen content (see Section 3.1). In this research question, we want to evaluate in how far existing off-screen visualization techniques can be adapted to head-mounted devices. Therefore, we will select promising off-screen visualization techniques and transform them from 2D techniques to 3D techniques.

4.6.2 RQ2: How can Mixed Reality devices with small fields-of-view be extended to present directional cues to out-of-view objects?

While visualization of out-of-view objects on head-mounted Mixed Reality devices is possible, it adds problematic visual clutter to the screen because current Mixed Reality devices have narrow fields-of-view. Therefore, it can make sense to extend the screen with peripheral light displays that present additional information to the user. In this research question, we want to investigate how we can extend the fields-of- view of Mixed Reality devices in order to present directional cues that point to out-of-view objects. Therefore, we develop radial light displays (based on LEDs) that we attach to Mixed Reality devices and use to investigate how suitable this approach is for cueing direction.

4.6.3 RQ3: In what way can the directions and distances of out-of-view objects be visualized for moving or non-moving objects in Mixed Reality?

In our last research question, we want to investigate more complex visualization techniques. Here, we build upon our second strategy for locating out-of-view objects (strategy: visualizing direction and distance; see Figure 4.2b). To present the direction and distance to the user, we use the display of the head-mounted device. However, since we need to visualize multiple out-of-view objects along with additional information such as distance, we need to develop visual cues that use less space than the 3D visual cues from our first research question. Here we ask, in what way can one visualize the direction of and distance to non-moving or moving objects? Therefore, we develop a novel visualization technique called EyeSee360, which is capable of visualizing direction and distance for multiple out-of-view objects simultaneously, and compare it with other techniques using both static and moving out-of-view objects.

5 Adapting Off-Screen Visualization Techniques

Just as objects recede from our fields-of-view, content recedes from small-screen hand-held devices. For example, when zooming in on a map app such as Google Maps[1] to see a specific part of the map in more detail, other parts of the map and relevant point of interest recede from the smartphone screen into off-screen space. To address the problem of content receding from view on devices with smaller screens such as smartphones, researchers have suggested different techniques, as reported in Section 3.1. For example, some of these techniques use visual cues to point into the direction of relevant off-screen content. These techniques have proven to be useful for smartphones and therefore may be promising when applied to head-mounted displays. Therefore, in this chapter, we want to explore in how far these existing techniques can be adapted to visually cue into the direction of relevant objects out of the user's field-of-view using head-mounted Mixed Reality:

RQ1: To what extent can existing off-screen visualization techniques be adapted to cue the direction to out-of-view objects in Mixed Reality?

To address this research question, we started by selecting promising off-screen visualization techniques from related work that were developed to point to off-screen objects on small-screen devices. To enhance existing Augmented Reality (AR) devices with this capability, we adapted well-known 2D off-screen object visualization techniques (Arrow, Halo, Wedge) and applied them to AR to cue direction to out-of-view objects. As the name suggests, Arrow uses an arrow pointing into the direction of the off-screen content. Halo and Wedge use well-known shapes that are only partially visible on the screen; however, humans are able to locate off-screen content by mentally completing these shapes. To transfer these techniques to AR, we created a projection to translate 3D coordinates of relevant out-of-view objects to 2D coordinates on an orthogonal projection plane in front of the user. Thereby, we could apply the selected off-screen visualization technique to the 2D orthogonal projection plane. Since we did not know how well these techniques would perform, we restricted ourselves to 90 degrees in front of the user (see Section 5.1).

Thereafter, we selected the two best-performing techniques from the first user study (Halo and Wedge) and further improved them to work with out-of-view objects spatially distributed all around the user. Here, we extended the support to both Augmented and Virtual Reality to see in how far this choice affects the performance of both techniques. Instead of applying the techniques as 2D visual cues to a 2D plane in front of the user, we extended both techniques to use 3D visual cues that bend around the user's head to point into the directions of the relevant out-of-view objects, while preserving most of the original characteristics of both techniques. Afterwards, we compared both techniques in two user studies to evaluate them, one in Augmented and one in Virtual Reality (see Section 5.2).

---

[1] Google Maps. `www.google.com/maps`, last retrieved April 21, 2020

In the last part of this chapter, we focus on improving Arrow, the visualization technique from the first part of this chapter that did not perform as well as Halo or Wedge. Our main goal was to investigate if the fundamental concept of amodal completion used in Halo and Wedge can be transferred to Arrow. Furthermore, we wanted to know in how far the technique can utilize 3D visual cues to cue direction to spatially distributed out-of-view objects in 3D space. We hypothesized that the concept of mentally completing only partially visible shapes could be applied to uniform movements as well, in the sense that a user who sees only part of the uniform movement, can mentally complete the movement out of view. To investigate this, we evaluate our improved Arrow technique in a user study in Augmented Reality in which we compare it to another technique capable of cueing direction to out-of-view objects (see Section 5.3).

5.1 Comparing Off-Screen Visualization Techniques

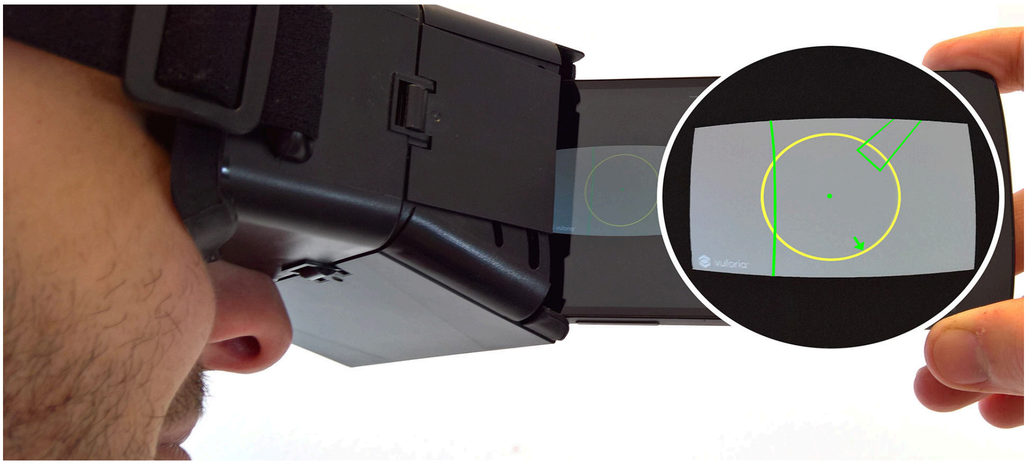


Figure 5.1: Adapted 2D off-screen visualization techniques (Arrow, Halo, and Wedge) in head-mounted Augmented Reality.

In the past, different off-screen visualization techniques have been proposed to address the problem of content receding from view on devices with small screens. These different techniques can be classified into three main approaches: Overview+detail, Focus+context, and Contextual views. In our analysis of related work (see Section 3.1), we described that Contextual views was identified as performing best [IGY06, BCG06, BC11]. Furthermore, the other two approaches suffer from crucial disadvantages that make each a poorer selection for the visualization of out-of-view objects. For example, the Overview+detail approach results in a higher cognitive load due to the mental workload required to integrate all views [CKB09, TSL^{+}11], while the Focus+context approach presents contextual information with distortion [SZG^{+}96, BLH04, CKB09], which may be less suitable for a display presented close to the eyes. Therefore, we selected different off-screen visualization techniques from the Contextual views approach and applied them to AR for cueing direction to out-of-view objects. We employed a projection plane orthogonal to the user's line-of-sight and in the user's view frustum and utilized the well-known 2D off-screen visualization techniques of Arrow [BCG06], Halo [BR03], and Wedge [GBGI08] to aid in visualizing out-of-view objects. To make the 2D visualizations applicable in 3D space, we used a two-step projection to translate 3D coordinates to 2D coordinates on the orthogonal projection plane (see Subsection 5.1.2). To compare the performances of the different visualizations in Augmented Reality, we conducted a comparative user study.

Here, our research contributions include:

1. An adaptation of three 2D off-screen visualization techniques (Arrow, Halo, and Wedge) for head-mounted AR that serves as a baseline for future work.
2. An evaluation of the adapted techniques for visualization of out-of-view objects in head-mounted Augmented Reality.

The work presented in this section was published as a late-breaking work at the MobileHCI conference in 2017 [GAHB17].

5.1.1 Approach

We restricted ourselves to 90 degrees of 3D space in front of the user. Therefore, we avoided out-of-view objects behind the user, as this makes the adaptation more complex. For example, an object that is exactly 180 degrees behind the user can be represented ambiguously on both the left and right sides of the viewing plane. Furthermore, as a first step we wanted to evaluate whether the projection plane in the user's view frustum was a feasible option for encoding direction information. In a first comparative study, we implemented the three selected visualization techniques (Arrow, Halo, and Wedge) in video see-through Augmented Reality. With video see-through AR, a camera image is looped onto a screen directly in front of the user's eyes. This was implemented with Vuforia[2] in Unity3D[3] using Google Cardboard[4]. As an additional minor contribution, we were able to evaluate the feasibility of Google Cardboard as a cheap and fast development platform for out-of-view Augmented Reality visualization techniques. The Google Cardboard that was used in the study can be seen in Figure 5.4.

---

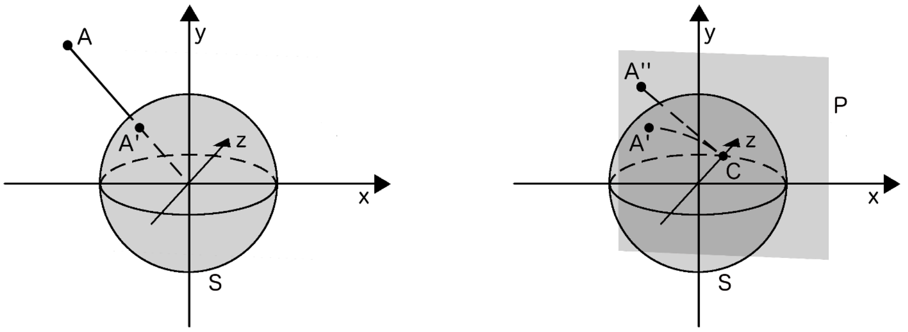
[2] Vuforia. `www.vuforia.com`, last retrieved April 21, 2020

[3] Unity3D. `www.unity3d.com`, last retrieved April 21, 2020

[4] Google Cardboard. `arvr.google.com/cardboard`, last retrieved April 21, 2020

5.1.2 Projection from 3D to 2D

A key aspect of our approach is the projection of out-of-view objects in 3D space onto a 2D plane (projection plane) and further applying 2D techniques that indicate direction towards the out-of-view objects. This direction information should consider the human frame-of-reference, meaning that objects behind the user should be indicated by turning the head to the side and not upwards. Additionally, the mapping needs to be proportional for all directions. The proportional projection method we apply here draws on the Mercator projection[5].



(a) First Step: Projection of 3D coordinate A on sphere S.

(b) Second Step: Project sphere coordinate A' on plane P.

Figure 5.2: Example of proportional projection that draws on the Mercator projection to map 3D out-of-view objects onto an 2D orthogonal plane.

An example mapping is shown in Figure 5.2. The user's head serves as the origin for the coordinate system, looking into the direction of the positive z-axis. First, we map point A, which represents an out-of-view object, onto sphere S. The mapped point A' is the intersection point of the line segment between point A and the origin point of sphere S (see Figure 5.2a). Second, we map point A' onto plane P (see Figure 5.2b), which is placed orthogonal to the x and y axes at contact point C. Then we calculate the shortest line on the surface of sphere S with monotonic y-values between C and A'. Finally, we map the line on the plane while keeping the length information and the same angle from C. Point A'' at the end of this line is the 3D to 2D mapped point.

5.1.3 Implementation

Our implementations of the 2D off-screen visualization techniques (Arrow, Halo, and Wedge) adapted to head-mounted Augmented Reality are shown in Figure 5.3. Each visualization technique uses one visual cue to encode one out-of-

---

[5] Mercator Projection. `en.wikipedia.org/wiki/Mercator_projection`, last retrieved April 21, 2020

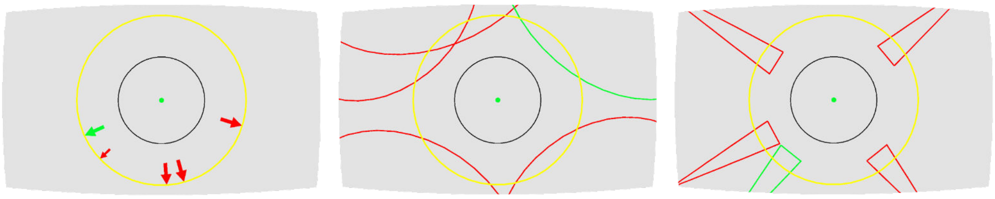
view object. A brief overview of the implementation details for each visualization technique is provided below:

Arrow points towards the out-of-view object's location. The arrow itself scales depending on how far it is from the object: a bigger arrow is used to indicate a closer distance and a smaller arrow is used for a location further away [BCG06].

Halo surrounds the out-of-view objects with circles. The center of each circle is exactly at the out-of-view object's position. The circles are just large enough to be on-screen [BR03].

Wedge uses isosceles triangles to represent the position of each out-of-view object. The tip of the triangle is at the object's position. They make room for one other to avoid cluttering [GBGI08].

Instead of using the border of the screen to show contextual information, we used a circle (seen as a yellow circle in Figure 5.3). This is more in line with the peripheral capabilities of the human vision system [SRJ11, LL09]. Moreover, with a circular border, the orthographic projection and the radial projection are equivalent.



(a) Arrow [BCG06]. (b) Halo [BR03]. (c) Wedge [GBGI08].

Figure 5.3: Examples of the adapted 2D off-screen visualization techniques (Arrow, Halo, Wedge).

A green dot in the center of the screen served as a cursor for communicating the user's guessed position for the out-of-view object. The dot was controlled by head movement and could be confirmed with a Bluetooth remote control. Additionally, we used a black circle fixed in the environment around the green cursor to control the visibility of the visualization. As long as the pointer remains within the black circle, the visualization is there. If the pointer leaves the black circle, the visualization disappears. We had to limit the visibility to avoid simple approximation approaches by the participant while estimating directions to the out-of-view objects. The visual cues of the visualization techniques are shown in red by default and green when highlighted. Here, highlighted means that the task asks the user to focus on this specific out-of-view object and its visual cue.

5.1.4 Study Design

To evaluate the performance of the adapted visualization techniques (Arrow, Halo, and Wedge), we conducted a within-subjects controlled laboratory study. Our study had two independent variables: visualization with three levels (Arrow vs. Halo vs. Wedge) and number of objects with three levels (one vs. five vs. eight). We varied the number of shown out-of-view objects both because Halo suffers from on-screen cluttering (cf., [BCG06, HPB10]) and because during pilot tests we found that eight objects was the threshold for cluttering. Furthermore, these numbers of objects are specified in our requirements. We used quantitative methods to evaluate user performance, where our dependent variable was direction error. Additionally, we gave participants the SUS questionnaire [Bro96] in order to gain insight into the perceived usability.

The direction error here is the angular error, which is the angle between the user's assessment of the out-of-view object's position and the correct position in 3D space. We did not measure task completion time since we were primarily concerned with direction error in this preliminary study. We were also unsure of how well Google Cardboard would perform with these visualization techniques because preliminary trials revealed simulator sickness effects during the task that measured completion time.

For this study, we derived the following sub-question from our first research question: *(RQ1a) Which of the 2D adapted visualization techniques (Arrow, Halo, and Wedge) performs best with respect to direction accuracy and perceived usability of various out-of-view objects?*

Since Wedge outperformed both Halo and Arrow in prior research [GBGI08] using a smaller number of objects, we hypothesized that:

H_2 *Wedge would result in better user performance (i.e., lower angular direction error) than both Halo and Arrow.*

5.1.5 Procedure

Participants were first given a demo of the Google Cardboard device (see Figure 5.4) where they could test out the different visualization techniques. The within-subjects study was divided into three blocks, each testing one visualization technique. We counter-balanced the blocks across all participants. Each visualization technique was tested with one, five and eight out-of-view objects. Each number of out-of-view objects was tested five times (which was deemed sufficient from pilot testing). The number of out-of-view objects in every run was randomized. Given the foregoing, we had 3 (blocks) x 3 (number of objects) x 5 (iterations) resulting in 45 runs per participant.

(a) Person wearing Google Cardboard. (b) Cardboard (plastic) without phone.

Figure 5.4: The Google Cardboard used in the study.

In each run, we randomly selected one out-of-view object and highlighted its visual cue in green (see Figure 5.3). Then, the participant had to guess the position of the out-of-view object without seeing it. A green cursor controlled by a wireless remote allowed each participant to select the object's position. To avoid getting the exact position of an out-of-view object through head movement, the visualization technique was only visible in a small area directly in front of the participant. Moving the green cursor out of the black circle disabled the visualization technique, so the participant had to guess the out-of-view object's position by the affordances the technique offered. After each block, participants had to fill out a SUS questionnaire [Bro96] about the technique in that block. At the end of the study, participants filled out a general information form, which included items such as age and gender. In addition, they rated their experience with head-mounted devices on a 5-point Likert-scale, where 1 is strongly disagree and 5 is strongly agree. They were also asked if they showed any signs of simulator sickness). Overall, each participant took approximately 45 minutes to finish the experiment. During our study, we did not consider eye movement, therefore we told the participants to keep looking at the green dot in the center of the screen.

5.1.6 Participants

We recruited 22 participants[6] (7 female), aged between 20 and 38 years (M=25.5, SD=3.7). None of them suffered from color vision impairment. All had normal or corrected-to-normal vision. Ten of the participants already had experience with head-mounted displays.

---

[6] For mean effect sizes of (f=0.2), at least 390 data points are necessary, wherein with the planned study design this makes for testing at least 9 participants. We calculated this value with G*Power under two-way ANOVA (α=0.05 and 1-β=0.9). We based it on three techniques with three different number of objects for each, which makes nine in total. The numerator df is (3-1)*(3-1)=4.

5.1.7 Results

Direction Error We consider the effects of each of the two factors, visualization and number of objects, on direction error. The mean errors are Arrow=6.53°, Halo=3.87° and Wedge=4.47°. Our data does not follow a normal distribution (Shapiro-Wilk-Test ($p<0.001$)), and thereafter we compared more than two matched groups using the non-parametric Friedman test. The Friedman test revealed a significant effect of visualization on direction error ($\chi^2(2)=159.88$, $p<0.001$, N=22). A post-hoc test using Wilcoxon Signed-rank with Bonferroni correction showed significant differences between the three groups (see Table 5.1).

Table 5.1: Pairwise comparison of visualization techniques.

Comparison	**p-value**	**ϕ-value**
Halo vs. Arrow	<0.001	0.46
Wedge vs. Arrow	<0.001	0.35
Wedge vs. Halo	<0.01	0.20

Next, we ran a Friedman test to evaluate whether there was a significant effect of number of objects on direction error. First, we looked at interactions between number of objects and visualization. Here, we did not find a significant effect ($\chi^2(2)=3.08$, $p=0.21$, N=22). Then, we looked into the number of objects for each visualization separately. Here again, there were no significant effects for Arrow ($\chi^2(2)=2.17$, $p=0.34$, N=22), Halo ($\chi^2(2)=0.67$, $p=0.71$, N=22), or Wedge ($\chi^2(2)=4.74$, $p=0.09$, N=22). The mean direction errors for all combinations of visualization and numbers of objects are shown as a boxplot in Figure 5.5.

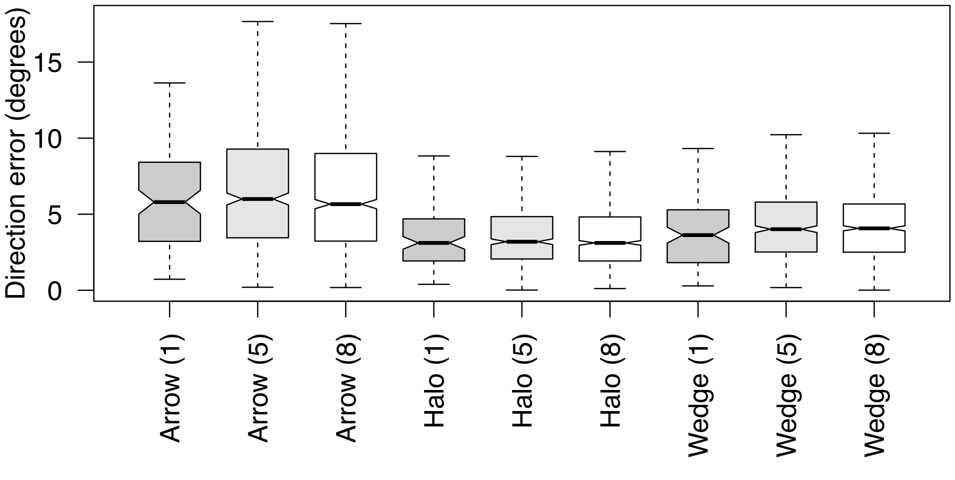


Figure 5.5: Boxplot of median direction errors for the tested visualization techniques and numbers of objects (shown within the parentheses).

Perceived Usability With respect to perceived usability, Wedge (70) was deemed on average to be most usable, where it just passes the accepted SUS literature threshold score of 68. Halo (66) was a runner up, and Arrow (61) was perceived to have poorer usability.

5.1.8 Discussion

Our initial results indicate differences between the established techniques that we adapted for head-mounted AR. In line with our hypothesis H_1, we saw that Halo and Wedge performed better with respect to direction error than did Arrow. However, we expected Wedge to outperform Halo, due to Halo's known cluttering effects [GBGI08]. Therefore, our hypothesis H_1 did not hold. We think this may be due to the fact that, instead of using a rectangular border as is done for small-screen devices, we used a circular border in line with human perception characteristics. Therefore, the corner density problem of Halo did not affect the performance [GBGI08]. However, with respect to perceived usability, Wedge performed slightly better than Halo.

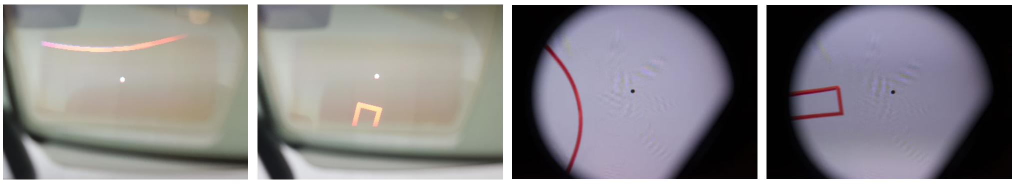
Moreover, we observed that direction error increased when the angle between the user's line of sight and the out-of-view object increased. In other words, the error increases when users have to turn their heads more. Here, the use of an orthogonal plane in front of the user is a problem for higher degree values because the transferred techniques only indicate head movements towards the off-screen objects. In other words, even an object presented behind the user would be indicated using a visual cue pointing to the right or left side. For this reason, we restricted the out-of-view objects to an area 90° in front of the user. To summarize, our approach is feasible for visualizing the positions of objects in 3D space 90° in front of the user, but for 180° or even 360° we have to further adapt Arrow, Halo, and Wedge or come up with novel visual cues.

Additionally, we observed how video-see-through AR in platforms such as Google Cardboard perform. We discovered that one advantage of our approach was that participants did not suffer from simulator sickness, according to their reporting in the general information form at the end of the study. However, Participants did face a slightly delayed picture of their environment, for which they compensated by moving their heads more slowly. Therefore, video-see-through might not be well-suited for tasks that involve fast movement, such as measuring the time to search for an out-of-view object. However, from a development perspective, the Google Cardboard platform was simple to use for video see-through AR. With Unity3D and Vuforia, there exists a beginner-friendly development environments for prototyping new Augmented and Virtual Reality experiences that are cost-effective and very simple to set up.

5.1.9 Conclusion

In this section, we compared three off-screen visualization techniques for head-mounted Augmented Reality with respect to their performance for visualizing the direction of out-of-view objects. We found that Halo objectively performed best, while Wedge subjectively performed best. In the next section, our goal is to expand our work in the 3D space to 180° or 360°. In this endeavor, we will investigate if a curved projection plane performs better in terms of minimizing direction error. Furthermore, we would like to adapt the visualization techniques for optical-see-through devices (e.g., Hololens) and investigate if the concept of amodal completion used for Halo and Wedge is applicable to Arrow as well.

5.2 Beyond Halo and Wedge: 3D Visual Cues



(a) HaloAR. (b) WedgeAR. (c) HaloVR. (d) WedgeVR.

Figure 5.6: Implementation of Halo and Wedge visualization techniques for Augmented and Virtual Reality.

In the last section, we adapted three different off-screen visualization techniques to address the problem of objects being hidden from view. Each of those visualization techniques used a different type of visual cue to encode direction to out-of-view objects. However, we applied these techniques onto an orthogonal plane in front of the user. As a result, with increasing directional angles towards the out-of-view objects, the techniques performed worse. We believe that this is due to the used projection, which maps 3D objects onto a 2D plane. Therefore, we think that instead of transforming the 3D out-of-view objects to 2D positions on a plane in front of the user, we should rather transform the used 2D visual cues into 3D space. Furthermore, we selected Halo and Wedge as the two best-performing techniques from our last study.

We developed four new visualization techniques (HaloVR, WedgeVR, HaloAR and WedgeAR) that are inspired by 2D mobile off-screen visualization techniques [BR03, GBGI08] and applied them to Virtual and Augmented Reality. In Section 5.1, we showed that these off-screen visualization techniques can be used to point towards out-of-view objects in AR. However, the use of a 2D overlay is not extendable for pointing to out-of-view objects distributed 360° around the user. Therefore, we developed our visual cues as 3D objects that point in the directions of out-of-view objects. We limited our visual cues to visualize only the directions towards out-of-view objects, which is sufficient for bringing out-of-view objects into the user's field-of-view (e.g., showing a relevant point-of-interest during sightseeing) and therefore, well-suited as a first approach towards visualizing out-of-view objects in Virtual and Augmented Reality.

To evaluate the performance of our different visualization techniques (HaloVR, WedgeVR, HaloAR and WedgeAR), we conducted two user studies. The first user study was done in VR using the Oculus Rift and the second study was done in AR using the Microsoft Hololens.

Here, our research contributions include:

1. Development of four 3D out-of-view visualization techniques inspired by 2D off-screen visualization techniques for head-mounted VR (HaloVR, WedgeVR) and AR (HaloAR, WedgeAR).
2. An evaluation and thereby comparison of the techniques developed (HaloVR, WedgeVR, HaloAR, and WedgeAR) for visualizing out-of-view objects in head-mounted VR and AR.

The work presented in this section was published as a full paper at the MobileHCI conference in 2018 [GABH18].

5.2.1 Approach

We explore the visualization of out-of-view objects 360° around the user. Our techniques utilize the periphery of the user to point towards out-of-view objects. However, our techniques are not constrained to a specific location in the periphery, so they can be shown in the near or far periphery. This makes sense if the head-mounted device has a small field-of-view, as it brings all displayed content closer to foveal vision, if one is looking straight ahead (e.g., for the Hololens). As mentioned earlier, our work draws on Section 5.1, which shows that out-of-view object visualizations can be successfully adapted from 2D off-screen visualization techniques. We wanted to explore how well our 3D visualization techniques can work across a) narrow FOV ranges and b) visual representations in AR and VR. Furthermore, we needed to evaluate our techniques for both technologies (VR, AR) because they have different influencing factors (e.g., different lighting conditions). For VR, we used the Oculus Rift because the display resolution and FOV are the current state of the art[7]. The same applies to the Hololens in AR, with the advantage of highly accurate placement of 3D objects in real environments. We address each of the Virtual and Augmented Reality technologies in different parts, where the first part describes our user study in Virtual Reality and the second part does the same for Augmented Reality.

---

[7] It was state-of-the-art at the time we conducted this research.

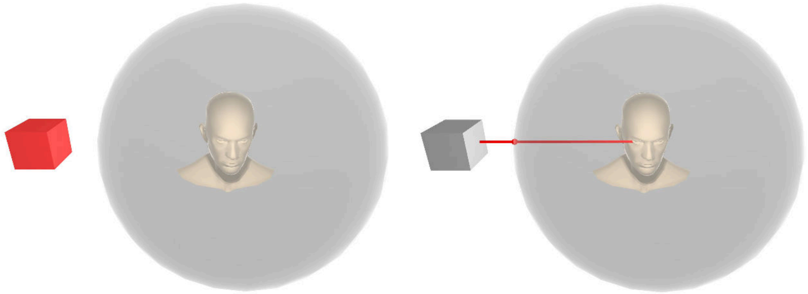
5.2.2 Designing Out-of-View Visualization Techniques

A key aspect of our approach is the use of two off-screen visualization techniques (Halo [BR03] and Wedge [GBGI08]) as inspiration for our developed 3D visualization techniques. Both techniques make use of well-known simple shapes that users can mentally complete even when only part of the shape is visible, a process known as *amodal completion* [Mic91]. We choose Halo and Wedge because both techniques rely on the Contextual views approach [BCG06]. Contextual views mostly represents only objects of interest and does so using simple shapes pointing to the object [GBGI08]. These visualization techniques and their visual cues were described in Subsection 5.1.3.

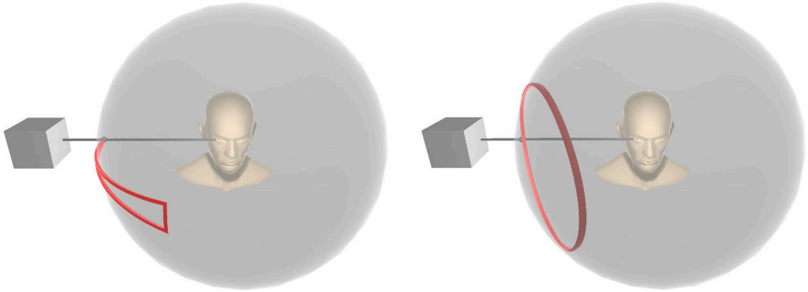
To visualize the direction towards out-of-view objects with our techniques, we first need to obtain that direction from the object's position in 3D space. Therefore, we project the positions of out-of-view objects onto an imaginary sphere around the user's head. This projection is done by drawing an imaginary line between the user's head and the position of the out-of-view object. The point where this line intersects with the imaginary sphere is the normalized vector representing the 3D direction towards the out-of-view object (see Figure 5.7a and Figure 5.7b). Next, we obtain a normalized vector for each out-of-view object representing its 3D direction. Each normalized vector is located on the imaginary unit sphere around the user's head. As a second step, we project our techniques (HaloVR, WedgeVR, HaloAR and WedgeAR) onto the inner surface of the imaginary sphere around the user's head (see Figure 5.7c and Figure 5.7d). We accomplish this by spanning our visual cues between the intersection of the sphere (see Figure 5.7b) and the user's line of sight. The sphere we used has a virtual radius of one meter.

To fully utilize the existing space around the user, we show the visual cues for out-of-view objects in all directions (including above the user's head) with a constant distance to the user's line of sight. While the distance to the user's line of sight can be adapted based on need (near periphery and small field-of-view, far periphery and large field-of-view), the distance is always kept the same. For this approach, we derive encodings that are based on human perception characteristics [Kal01, SRJ11] (e.g., color perception and resolution) to ensure that every out-of-view object is adequately and equally perceived. Showing the visual cues to out-of-view objects in all directions has the further advantage that it provides the user with a better understanding of the direction of an out-of-view object even if they do not understand the amount of head movement required to find it. In addition to the original Halo [BR03] and Wedge [GBGI08] implementations, we needed to adapt our techniques for VR and AR views:

All techniques. We added transparency to all techniques. Using this approach, we ensure no content is occluded by our visual cues, and it further helps to reduce clutter in cases of overlapping visual cues.



(a) Out-of-view object and user. (b) Intersection on sphere.



(c) WedgeVR and WedgeAR. (d) HaloVR and HaloAR.

Figure 5.7: Projection of out-of-view object onto sphere with applied visualization techniques (WedgeVR, WedgeAR, HaloVR, and HaloAR).

HaloVR, HaloAR. For HaloVR and HaloAR, we needed to change the way the techniques pointed towards out-of-view objects. Originally, each object was in the center of a Halo, but applying these techniques to the inner surface of a sphere would limit Halo to only show objects 90° away from the user's line of sight. To overcome this limitation, we span HaloVR and HaloAR in between the user's line of sight and the direction towards the out-of-view object (cf. Figure 5.7b). The techniques are shown in Figure 5.6c and Figure 5.6a.

WedgeVR, WedgeAR. For WedgeVR and WedgeAR, we needed to remove the ability to make space for other visual cues. There are two reasons for this. First, this approach avoids Wedges jumping during head movements. In pilot testing, users reported losing track of specific out-of-view objects when Wedge jumping was involved. The wedges jumping due to the original approach of having them make space for other Wedges. This is because they were applied onto the inner surface of a sphere, which is a double-curved geometry. Making space for other wedges combined with head movement requires wedges to jump. In the original work, this did not happen because the Wedges were applied to a 2D plane. The second reason for removing the ability to make space is to ensure that all visual cues are always displayed with the same distance towards the user's line of sight and can be perceived equally well. The techniques are shown in Figure 5.6d and Figure 5.6b.

We implemented our techniques using the Unity3D game engine. To ensure smooth surfaces and 3D shapes, we generated the meshes for all techniques during run-time. Furthermore, we allowed adaption of the number of polygons used for each technique to ensure good rendering performance on different platforms. Additionally, we added a directional light source to simulate natural light effects like shadows and reflections. We used for each platform the same "intrusion" level into the user's field-of-view (5° Hololens, 25° Oculus). More platform specific implementation details are described in the corresponding parts. All four techniques (HaloVR, HaloAR, WedgeVR and WedgeAR) are available as Unity packages under MIT License on GitHub[8].

5.2.3 Part I: Evaluating HaloVR and WedgeVR

In this section, we provide an evaluation of the performance of our designed techniques HaloVR and WedgeVR in Virtual Reality.

5.2.3.1 Apparatus

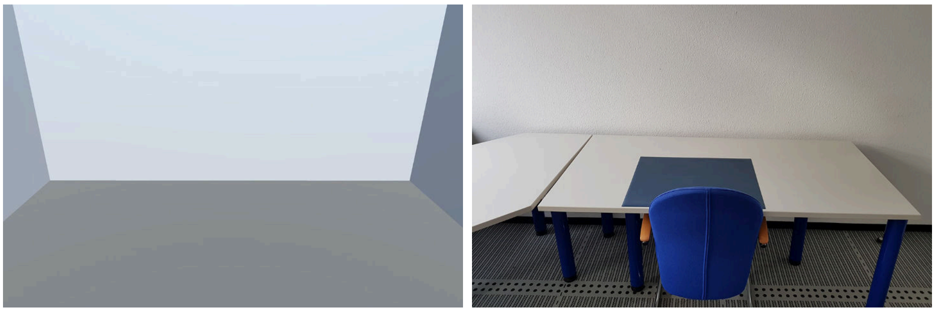
For testing our techniques in VR, we used the Oculus Rift because the display resolution and FOV are the current state of the art. The FOV of the Oculus Rift is about 90 degrees. This provides a large design space for applying our techniques. Therefore, we decided to place the visual cues in the far periphery of the user. This helps to avoid clutter and object occlusion. Further, it serves to maintain an immersive experience. Also, as shown in Section 5.1, the periphery of the user can be used for out-of-view visualization and provides a better user experience. During pilot testing, we observed that an area of 50° in the center of the screen should remain unaffected by our visualization techniques. This allows the user to focus on their primary task while still being able to sufficiently observe moving out-of-view objects in the far periphery as a secondary task. Since we are interested in how well these techniques perform, we needed to first test them in environments without other influencing variables. Therefore, we ran our study in an empty space (see Figure 5.8a). We used the Oculus Touch controller as an input device for our study.

5.2.3.2 Study design

To evaluate the performance of our techniques (HaloVR and WedgeVR), we conducted a within-subjects controlled laboratory study in VR with the Oculus Rift. Our study had two independent variables: visualization with two levels (HaloVR vs. WedgeVR) and number of objects with three levels (one vs. five. vs. eight). We varied the number of shown out-of-view objects to investigate

---

[8] GitHub OutOfView-Project. `www.github.com/UweGruenefeld/OutOfView`, last retrieved April 21, 2020



(a) Virtual world environment in VR. (b) Real world environment in AR.

Figure 5.8: Environment in Virtual and Augmented Reality.

the threshold for a maximum number of out-of-view objects and because of prior work that showed Halo suffers from on-screen cluttering (cf., [BCG06, HPB10]). Besides that, the same numbers of out-of-view objects (one vs. five vs. eight) were investigated in Section 5.1.

We used quantitative methods to evaluate user performance, where our dependent variables were search time, object selection accuracy and direction error. Search time is the time users need to locate and select an out-of-view object in the scene, while object selection accuracy specifies the number of objects users selected correctly. The direction error here is the angular error, which is the angle between the user's assessment of the out-of-view object's position and its correct position in 3D space. Additionally, we gave participants the SUS [Bro96] and RAW-TLX [Har06] questionnaires in order to gain insight into perceived usability and workload. To gain further insight into the user experience of all techniques, we additionally provided participants with Likert-scale items that asked about their performance during the study.

For this study, we derived the following sub-question from our first research question: *(RQ1b) Which of the visualization techniques (HaloVR, WedgeVR) performs best with respect to search time, object selection accuracy, and direction error for different numbers of out-of-view objects in Virtual Reality?* We posit the following hypotheses:

H_2 *A higher number of out-of-view objects results in worse search time performance of the tested visualization technique.*

H_3 *Based on the findings in Section 5.1 where Halo outperformed Wedge with respect to direction error, we hypothesize HaloVR will result in lower direction error as WedgeVR.*

H_4 *Both techniques (HaloVR and WedgeVR) result in acceptable usability.*

5.2.3.3 Procedure

Our study was divided into two tasks: a search task and a direction estimation task. Both tasks were divided into two blocks (four blocks in total), where each block tested one technique (HaloVR or WedgeVR). We counter-balanced each block across all participants. Each technique was tested with one, five, and eight out-of-view objects (see Subsubsection 5.2.3.2 for an explanation). Each number of out-of-view objects was tested five times. The number of objects tested in a trial was selected randomly at run time. Objects were randomly distributed in 3D space as out-of-view objects. We stored the seeds of the position generation to test the same positions for each technique. However, by randomly shuffling the seeds from the previous technique we ensured that participants could not recognize the positions of the out-of-view objects tested with the previous technique.

Overall, we tested four blocks with three different numbers of objects for five iterations, resulting in 60 trials per participant. After each block of the search task, we asked participants to fill out SUS and NASA Raw-TLX questionnaires. At the end of all blocks, participants were asked to fill out our subjective and demographic questionnaires.

Task A: Search Time

Each of the two blocks of this task started with a practice trial and an explanation of the visualization and the task the user had to perform. In each run of this task, a randomly chosen visual cue (HaloVR or WedgeVR) was highlighted in red and the participant had to find the represented out-of-view object with the support of the visual cue by selecting it with a cursor and pressing a button on the controller. The cursor was controlled by head movement.

Task B: Direction Estimation

Each of the two blocks of this task started with a practice trial and an explanation of the visualization and the task the user had to perform. In each run of this task, a randomly chosen visual cue (HaloVR or WedgeVR) was highlighted in red and the participant had to estimate the direction of the represented out-of-view object by selecting the estimated direction with a cursor and pressing a button on the controller. The cursor was controlled by head movement. The visualization technique was only visible when the user gazed into the starting direction. Any kind of head movement towards the out-of-view object disabled the visualization technique (i.e., made it disappear). Each participant took approximately 40 minutes to finish the experiment.

5.2.3.4 Participants

We recruited 16 participants (6 female), aged between 21 and 54 (M=30.06, SD=7.72). None of them suffered from color vision impairment. All had normal or corrected-to-normal vision. The participants did not receive any compensation.

5.2.3.5 Results

Search Task

For the search task, we consider the effects of the two factors (visualization, number of objects) on search time and object selection accuracy (where object selection accuracy refers to whether an object was found during the trial). The mean search times for the visualization techniques are: HaloVR=2.26s and WedgeVR=2.24s. The total number of incorrectly selected objects are: HaloVR (232/240=96.7% accuracy) and WedgeVR (234/240=97.5% accuracy). The times for the search task are compared in Figure 5.9.

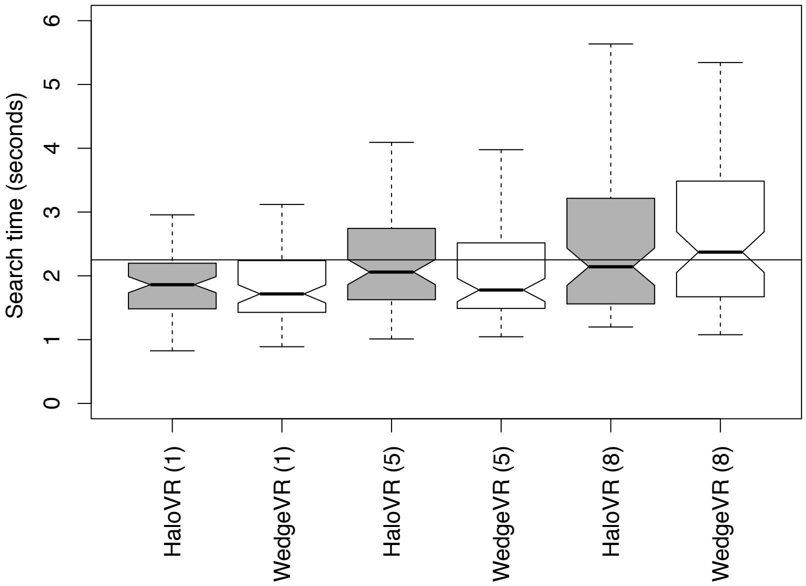


Figure 5.9: Boxplot of median search times of visualization techniques and number of objects (line indicates overall mean search time and number in parentheses indicates the number of objects tested).

A Shapiro-Wilk-Test showed that our data is not normally distributed ($p < 0.001$). As we compare two matched groups within subjects, we directly performed a Wilcoxon Signed-rank test. Here we found no significant effect of visualization technique on search time (W=14404, Z=-0.052, p=0.959, $\phi<0.01$). This indicates that HaloVR and WedgeVR do not sufficiently differ with respect to search time.

For a more detailed analysis, we compared the different combinations of visualization techniques and numbers of objects (HaloVR(1)=1.96s, HaloVR(5)=2.31s, HaloVR(8)=2.52s, WedgeVR(1)=1.87s, WedgeVR(5)=2.18s, and WedgeVR(8)= 2.66s). Since our data is not normally distributed and we compare six matched groups within subjects, we ran a Friedman test, which revealed a significant effect of different combinations of visualization techniques and numbers of objects on search time ($\chi^2(5)=37.15$, $p<0.001$, N=16). A post-hoc test using Wilcoxon Signed-rank with Bonferroni correction showed significant differences between some conditions, shown in Table 5.2. We can conclude that the search time for one object is smaller than for multiple objects. Furthermore, for WedgeVR the search time proportionally increases, where WedgeVR(1) < WedgeVR(5) < WedgeVR(8).

Table 5.2: Pairwise comparison of technique combinations for search time performance (number in parentheses indicates number of objects tested).

Combination	**p-value**	**ϕ-value**
HaloVR(1) vs. WedgeVR(1)	1	undef.
HaloVR(5) vs. WedgeVR(5)	0.240	0.14
HaloVR(8) vs. WedgeVR(8)	0.200	0.10
HaloVR(1) vs. HaloVR(5)	0.003	0.23
HaloVR(1) vs. HaloVR(8)	<0.001	0.32
HaloVR(5) vs. HaloVR(8)	0.307	0.08
WedgeVR(1) vs. WedgeVR(5)	0.007	0.21
WedgeVR(1) vs. WedgeVR(8)	<0.001	0.40
WedgeVR(5) vs. WedgeVR(8)	<0.001	0.28

Estimation Task

We consider the effects of the two factors (visualization, number of objects) on direction error. The mean errors for the visualizations are: HaloVR=23.02° and WedgeVR=20.67°. The direction errors are compared in Figure 5.10. A Shapiro-Wilk-Test showed that our data is not normally distributed ($p<0.001$). As we compare two matched groups within subjects, we performed a Wilcoxon Signed-rank test. Here we found no significant effect of visualization technique on direction error (W=13165, Z=-1.203, p=0.230, ϕ=0.05). This indicates that HaloVR and WedgeVR do not significantly differ with respect to estimation accuracy.

Furthermore, we tested the Pearson's product moment correlation coefficient between direction error and angle towards out-of-view object. The direction errors for HaloVR and WedgeVR can be seen in Figure 5.11. The correlations are: HaloVR=0.66 (t(238)=13.68, $p<0.001$) and WedgeVR=0.71 (t(238)=15.54, $p<0.001$). Our results indicate that there is a medium-strong correlation between

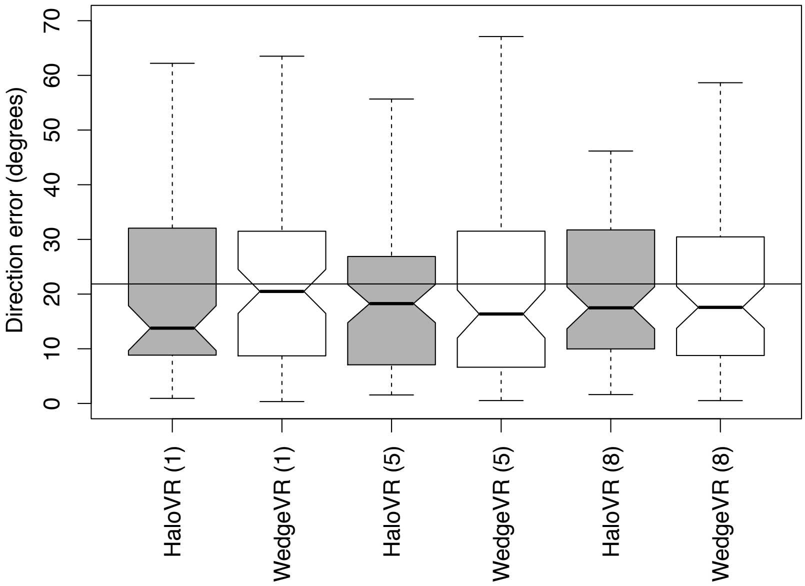


Figure 5.10: Direction error of visualization techniques (line indicates mean direction error and number in parentheses indicates the number of objects tested).

direction error and angle towards out-of-view object for both techniques. This shows that a higher angle results in a higher direction error.

NASA RAW-TLX

For NASA Raw-TLX [Har06] scores, HaloVR scored 29.44 and WedgeVR scored 23.88. Both values indicate a low workload, with WedgeVR having a slightly lower workload than HaloVR.

System Usability Scale

For SUS scores, HaloVR scored 87 and WedgeVR scored 92, both of which are over the threshold for acceptable usability [Bro96]. This shows that both techniques are usable for VR.

Likert-scale Questionnaire

At the end of the study, we asked participants to answer four questions with 5-point Likert items. Participants stated they were able to easily find the out-of-view objects with HaloVR (Md=4, IQR=1) as well as with WedgeVR (Md=5, IQR=1). Furthermore, they stated they were able to correctly estimate the positions of out-of-view objects with HaloVR (Md=3, IQR=1.25) and with WedgeVR (Md=3.5, IQR=1). Overall, nine participants preferred WedgeVR while seven

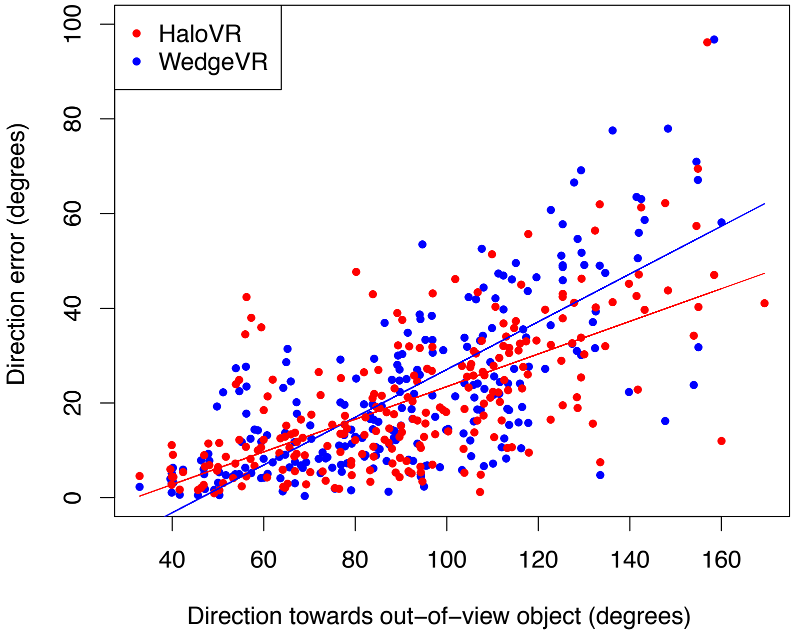


Figure 5.11: Correlations between direction error and angle towards out-of-view object for HaloVR and WedgeVR. *Best seen in color.*

preferred HaloVR. Furthermore, three participants stated that it was harder to estimate the position than it was to search for the object. However, all three of those participants stated that WedgeVR better supports position estimation than HaloVR.

5.2.3.6 Discussion

Number of Objects

The number of out-of-view objects visualized simultaneously had a significant effect on search time performance, while it had no effect on direction error. This can be explained by some participants needing more time when distraction is added by multiple visual cues. Therefore, we suggest using as few visual cues simultaneously as possible. Since we found a significant effect of number of objects on search time performance, we can accept our hypothesis H_2.

Comparison of Techniques

Both techniques performed well with respect to search time, object selection accuracy and direction error. In both tasks there were no significant differences between HaloVR and WedgeVR. Based on these findings, we cannot accept our hypothesis H_3 because we did not find a significant effect for direction error.

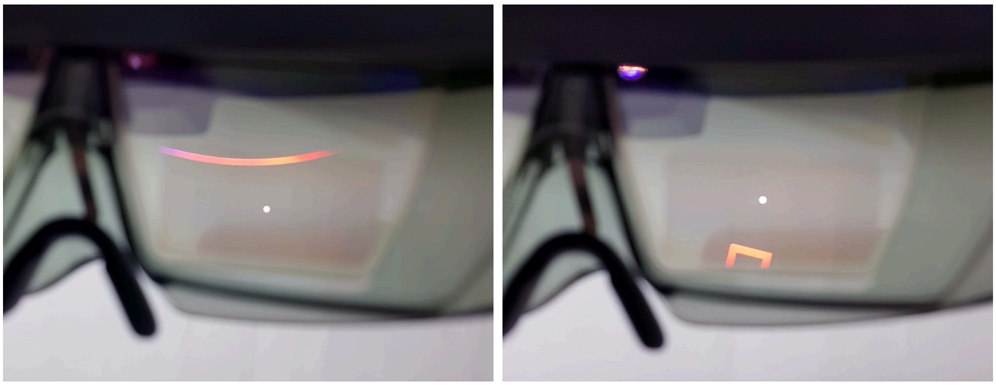
Both techniques had acceptable usability (HaloVR=87, WedgeVR=92) and low workload (HaloVR=29.44, WedgeVR=23.88); therefore, we can accept H_4.

Clutter

Four out of 16 participants stated that they preferred WedgeVR for the study because it resulted in less clutter to their view. They further liked the aspect that WedgeVR always points directly to the out-of-view object.

5.2.4 Part II: Evaluating HaloAR and WedgeAR

Here, we evaluate the performance of our designed techniques (HaloAR and WedgeAR) for foveal visualization of out-of-view objects in Augmented Reality. The two techniques are shown in Figure 5.12.



(a) HaloAR on Hololens. (b) WedgeAR on Hololens.

Figure 5.12: Tested techniques in Augmented Reality.

5.2.4.1 Apparatus

For AR, we used the Microsoft Hololens because the placement of 3D objects in the real environment and the display technology used are currently the state of the art for HMD AR devices. Further, the limited field-of-view is quite challenging and is helpful for testing how well our techniques generalize. The field-of-view of the Microsoft Hololens is about 30 degrees. In comparison with VR, this reduces the available space for our visual cues by a factor of three. On devices with such small fields-of-view, showing our visualizations in the far periphery is not possible. Therefore, we moved them to the near periphery when one is looking straight ahead, leaving only an area of 10 degrees unaffected. Further, we used a Hololens clicker instead of the Oculus Rift controller used in the previous study.

5.2.4.2 Study Design

To evaluate the performance of our designed techniques (HaloAR and WedgeAR) in AR, we conducted a second within-subjects controlled laboratory study with the Microsoft Hololens. Our study in Augmented Reality was based on the same study design as our previous VR study.

For this study, we derived the following sub-question from our first research question: *(RQ1c) Which visualization technique (HaloAR, WedgeAR) performs best with respect to search time, object selection accuracy, and direction error for different numbers of out-of-view objects in Augmented Reality?* We posit the following hypotheses:

H_5 *A higher number of objects results in worse search time performance.*

H_6 *Based on findings in Section 5.1 where Halo outperformed Wedge with respect to direction error, we hypothesize that HaloAR outperforms WedgeAR.*

H_7 *We expect the smaller Hololens FOV to negatively affect performance for search time, object selection accuracy, and direction error even when the visualizations are presented in the near periphery instead of the far periphery.*

H_8 *Both techniques (HaloAR and WedgeAR) result in acceptable usability.*

5.2.4.3 Procedure

Our study was divided into two tasks: a search task and a direction estimation task. Both tasks were similar to the tasks used for the VR study. We tested four blocks with three different numbers of objects for five iterations, resulting in 60 trials per participant. Overall, each participant took approximately 40 minutes to complete the study.

5.2.4.4 Participants

We recruited 16 participants (7 female), aged between 20 and 56 (M=30.63, SD=10.34). None of them suffered from color vision impairment. All had normal or corrected-to-normal vision. The participants did not receive any compensation.

5.2.4.5 Results

Search Task

For the search task, we consider the effects of the two factors (visualization and number of objects) on search time and object selection accuracy (where object selection accuracy refers to whether an object was found during the trial). The mean search times for the visualization techniques are: HaloAR=3.84s and WedgeAR=4.00s. The total numbers of incorrectly selected objects are: HaloAR

(235/240=98%) and WedgeAR (231/240=96.3%). The search times are compared in Figure 5.13.

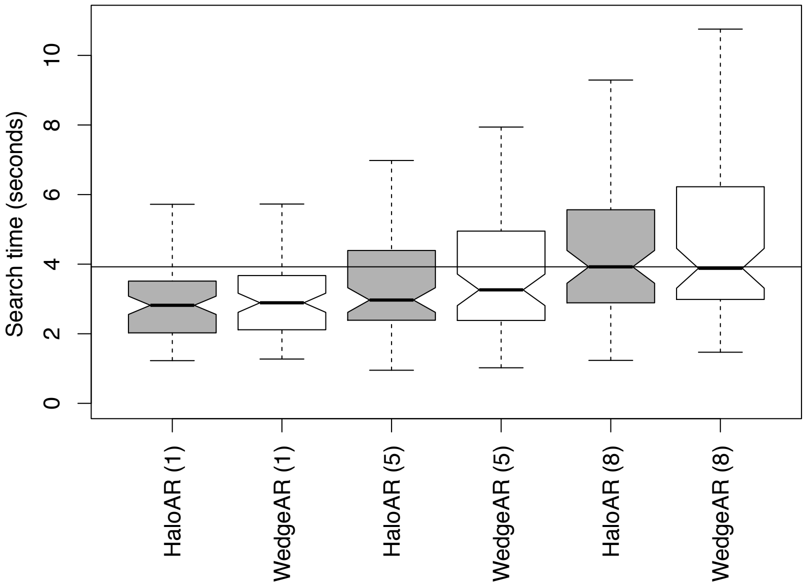


Figure 5.13: Boxplot of median search times of visualization techniques and number of objects (line indicates mean search time and number in parentheses indicates the number of objects tested).

A Shapiro-Wilk-Test showed that our data is not normally distributed ($p< 0.001$). As we compare two matched groups within subjects, we performed a Wilcoxon Signed-rank test. Here we found no significant effect of visualization technique on search time (W=13161, Z=-1.2065, p=0.228, ϕ=0.06). This indicates that HaloAR and WedgeAR do not significantly differ with respect to search time.

For a more detailed analysis, we compared the different combinations of visualization techniques and numbers of objects (HaloAR(1)=3.05s, HaloAR(5)=3.57s, HaloAR(8) =4.91s, WedgeAR(1)=3.29s, WedgeAR(5)=3.97s, and WedgeAR(8)= 4.75s). Since our data is not normally distributed and we compare six matched groups withi-subjects, we ran a Friedman test, which revealed a significant effect of different combinations (visualization, number of objects) on search time ($\chi^2(5)$=68.08, $p<0.001$, N=16). A post-hoc test using Wilcoxon Signed-rank with Bonferroni correction showed significant differences between all conditions, shown in Table 5.3. We can conclude that the number of objects has a significant effect on search time in Augmented Reality.

Table 5.3: Pairwise comparison of combinations in terms of search time performance (number in parentheses indicates the number of objects tested).

Combination	**P-value**	**ϕ-value**
HaloAR(1) vs. WedgeAR(1)	0.869	0.13
HaloAR(5) vs. WedgeAR(5)	0.399	0.07
HaloAR(8) vs. WedgeAR(8)	1	undef.
HaloAR(1) vs. HaloAR(5)	0.025	0.18
HaloAR(1) vs. HaloAR(8)	<0.001	0.46
HaloAR(5) vs. HaloAR(8)	<0.001	0.31
WedgeAR(1) vs. WedgeAR(5)	<0.001	0.27
WedgeAR(1) vs. WedgeAR(8)	<0.001	0.39
WedgeAR(5) vs. WedgeAR(8)	=0.001	0.25

Estimation Task

We consider the effects of the two factors (visualization and number of objects) on mean direction error. The mean errors for the visualization techniques are: HaloAR=29.79°and WedgeAR=36.03°. The direction errors are compared in Figure 5.14. A Shapiro-Wilk-Test showed that our data is not normally distributed ($p<0.001$). As we compare two matched groups within subjects, we performed a Wilcoxon Signed-rank test. Here we found a significant effect of visualization technique on direction error ($W=11531$, $Z=-2.720$, $p=0.006$, $\phi=0.12$). This provides evidence that HaloAR results in significantly better performance than WedgeAR with respect to estimation accuracy. However, here we found no significant differences between the six groups (HaloAR(1), HaloAR(5), HaloAR(8), WedgeAR(1), WedgeAR(5), WedgeAR(8)) when conducting a Friedman test ($\chi^2(5)=6.91$, $p=0.228$, $N=16$).

Furthermore, we tested the Pearson's product moment correlation coefficient between direction error and angle towards out-of-view object. The direction errors of HaloAR and WedgeAR can be seen in Figure 5.15. The correlations are HaloAR=0.294 ($t(238)=4.75$, $p<0.001$) and WedgeAR=0.372 ($t(238)=6.18$, $p<0.001$). Our results indicate that there is a weak correlation between direction error and angle towards out-of-view object for both techniques. This means a higher angle is less likely to result in a higher direction error than a lower angle.

NASA RAW-TLX

For NASA Raw-TLX scores, HaloAR scored 21.75 and WedgeAR scored 24.81 [Har06]. The values do not differ much, but both indicate a lower workload.

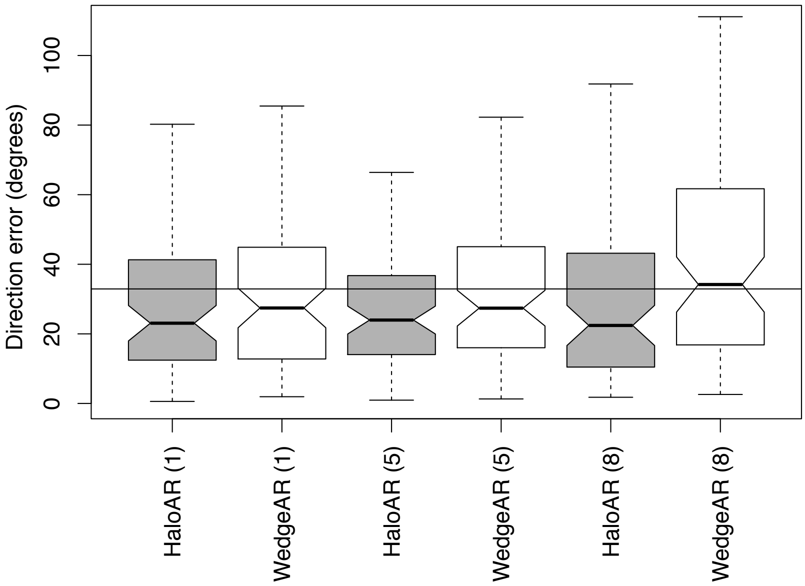


Figure 5.14: Boxplot of median direction error of visualization techniques (line indicates mean direction error and number in parentheses indicates the number of objects tested).

System Usability Scale

HaloAR scored 85 and WedgeAR scored 82 on the SUS, which is above the threshold for acceptable usability [Bro96]. According to the SUS scores, we find that both techniques are usable for AR HMDs.

Likert-scale Questionnaire

At the end of the study, we asked the participants to answer four questions with 5-point Likert-scale items. Participants stated they were able to easily find the out-of-view objects with HaloAR (Md=4, IQR=2) as well as with WedgeAR (Md=4, IQR=1). Furthermore, they said they were able to easily estimate the positions of out-of-view objects with HaloAR (Md=3, IQR=1.5) and WedgeAR (Md=3, IQR=2). Overall, six participants preferred WedgeAR while ten preferred HaloAR.

5.2.4.6 Discussion

Number of Objects

The number of out-of-view objects visualized simultaneously had a significant effect on search time performance, while it had no effect on direction error. This can be explained by some participants needing more time when distraction is

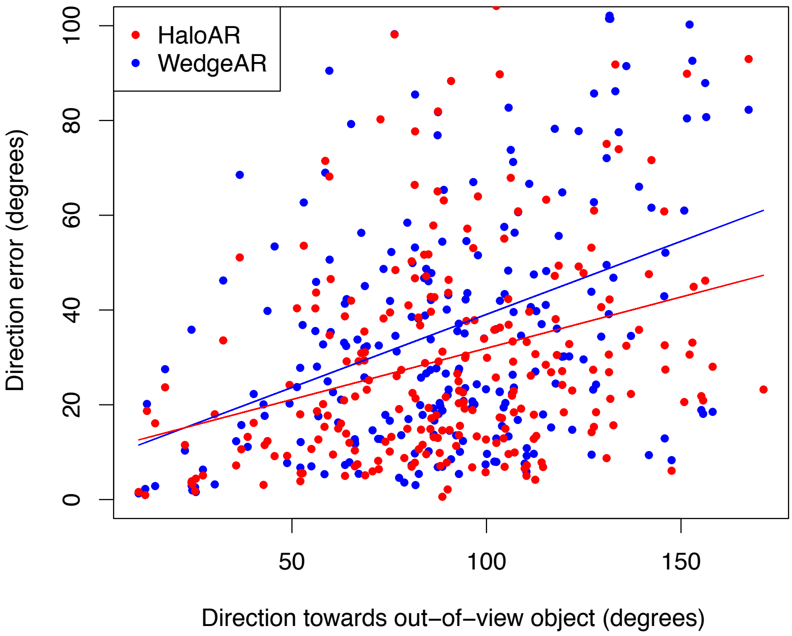


Figure 5.15: Correlations between direction error and angle towards out-of-view object for HaloAR and WedgeAR. *Best seen in color.*

added by multiple visual cues. Therefore, we suggest using as few visual cues simultaneously as possible. Given our results, we accept hypothesis H_5.

Comparison of Techniques

Both techniques performed well with respect to search time, object selection accuracy and direction error. We found a significant effect of visualization technique on direction error, but no significant effect of visualization technique on search time or object selection accuracy. To interpret this, the near peripheral presentation of HaloAR may have been easier to understand than WedgeAR and therefore may have led to lower direction estimation errors. Here, we must reject our hypothesis H_6.

Furthermore, both techniques demonstrated acceptable usability (HaloAR=85 and WedgeAR=82) and low workload (HaloAR=21.75 and WedgeVR=24.81), therefore we can accept H_8.

Small Field-of-View

The small field-of-view of the Hololens allows us to use only a small number of visual cues simultaneously. This is due to frequent overlapping of multiple visual cues, which adds clutter to the screen even with our added transparency. This is supported by the significant increases in search time of 1.38 seconds for

HaloAR(5) to HaloAR(8) and 0.78 seconds for WedgeAR(5) to WedgeAR(8). However, this affects HaloAR more strongly than it does WedgeAR. Here, we can accept our hypothesis H_7.

5.2.5 Implications

Advantages of Head-mounted Devices

Since our techniques are inspired by off-screen visualization techniques [BR03, GBGI08], they can be perceived as similar and therefore familiar to users. In this regard, our techniques also overlay the viewing frustum of the user. Combined with an HMD, our techniques offer a constant flow of information regarding out-of-view objects for presentation in both the far periphery and on-demand guidance for near peripheral presentation under more limited FOVs, such as those of currently available AR devices.

Reducing the Number of Visual Cues

How many out-of-view objects should be visualized is dependent on the task. For estimating the positions of out-of-view objects, the number of objects has no influence on user performance for up to eight objects. However, search time is negatively affected by a higher number of visual cues across both VR and AR. Therefore, we recommend reducing the number of visual cues that are visible at the same time during a search task. This can be achieved by successively guiding the user from one out-of-view object to the next in a sequential manner, or by showing only relevant (e.g., determined by usecase) out-of-view objects at a given time.

Far Peripheral vs. Near Peripheral Visualization

To adapt our techniques across different FOVs, we followed two different approaches for presenting our visual cues to the user. For small FOV devices where no far peripheral vision is available, we used a near-peripheral presentation. For larger FOV devices, we presented the visual cues in the far periphery. Based on our experiments, we recommend varying the distance of visual cues towards the user's line of sight between 8° and 30°. A lower angular distance results in more clutter since all visual cues are overlay one other independently of the direction to which they are pointing. A higher angular distance will result in worse perception of shapes and a less accurate estimation of position for out-of-view objects (cf., [Kal01, SRJ11]).

Guidance Towards Out-of-View Objects

Since we implemented out-of-view visualization techniques, we assumed that no visualization is necessary when the objects are visible on screen. However, during pilot tests, some participants stated that our artificial out-of-view objects all look the same, and they were therefore unable to decide which object they were seeking for when multiple out-of-view objects were closer together. To solve this problem, we decided to have 5° at the border of the screen where the visual cue remains active. However, it was still problematic for participants to distinguish between objects close together. This resulted in 14 wrongly selected objects in VR and 14 wrongly selected objects in AR. As an outcome of this, we recommend not hiding the visualization when it is unclear which object the user is seeking or replacing it with an in-view visual cue.

Performance of Visualization Techniques

Both experiments in VR and AR showed that our techniques are perceived as usable for seeking and selecting out-of-view objects. This is further supported by an average search time of 2.25 seconds in VR and 3.92 seconds in AR. However, our quantitative results showed that HaloAR performs significantly better with regard to direction error than WedgeAR (HaloAR=29.79°, WedgeAR=36.03°). Nevertheless, the correlation between the direction towards out-of-view objects and the direction error is stronger for VR (HaloVR=0.66, WedgeVR=0.71) than for AR (HaloAR=0.29, WedgeAR=0.37), which indicates that users were better at estimating smaller angles towards out-of-view objects in VR than in AR. To avoid different graphic performances between the platforms, we reduced the number of rendered polygons to a level that is supported by the Hololens and used the same number of polygons for the Oculus.

5.2.6 Limitations

Reduced 3D Perception

Since our visualizations (per individual device) are always positioned with the same distance to the user's line of sight, it is not possible to look at the 3D visual cue from different angles, making it harder to perceive the full volumetric shape of the visual cues. This limitation is part of our approach for positioning our visual cues at the same position in the user's periphery and not in the environment. However, by attaching the visualizations to the head of the user, we can ensure that the visual cues are always visible and do not disappear from view.

Ecological Validity

We tested both visualization techniques in VR as well as in AR in a controlled lab study. To measure user performance under these settings, we needed to control the environment as much as possible. However, this limits our understanding of how such techniques can be used across real applications and usecases (e.g., representing objects or characters in Virtual Reality games). Nevertheless, our work invites such ecological testing of out-of-view visualization techniques as a future research agenda.

5.2.7 Conclusion

In this section, we developed HaloAR, HaloVR, WedgeAR, and WedgeVR to visualize the positions of out-of-view objects. Our findings showed that all techniques are usable under constraints of their specific technology (VR or AR). We showed two different approaches of using our techniques (far and near peripheral visualizations). Furthermore, we found that the limited FOV of current AR devices has a negative impact on user performance. Our work opens avenues for further investigating out-of view-object visualization techniques, where we believe ecological testing and lowering direction estimation error will improve the adoption of such visualization approaches.

5.3 FlyingARrow: Flying Towards Out-of-View Objects

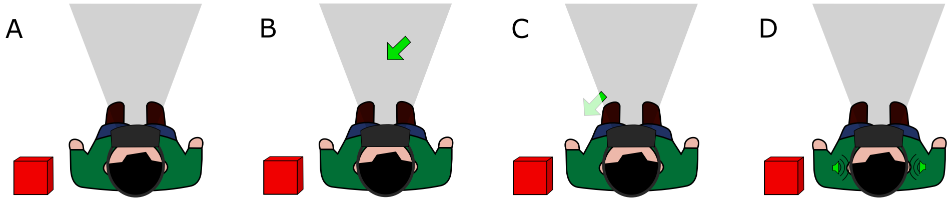


Figure 5.16: Visualization technique FlyingARrow (green) as image sequence (A-D), out-of-view object (red), and field-of-view (gray).

In our first study, we compared three different visualization techniques with one other, finding that the technique Arrow resulted in poorer performance than did Halo and Wedge. We think this may be because Halo and Wedge use a concept known as amodal completion, which refers to the idea that humans can mentally complete a shape even if only a part of it is visible. This concept is not part of Arrow in the original sense. Therefore, in this section, we aim at transferring this idea to Arrow. However, instead of mentally completing a shape, we hypothesize that humans may also be able to mentally complete the trajectory of an object traveling with the same speed in a fixed direction. Therefore, we combine the visual cue of an out-of-view object with an acoustic signal to point towards objects in 3D space. As shown by previous work, this combination can lead to faster reaction times [LBS+17]. Further, the combination of both modalities helps to reduce the visual information load [Wic08] and clutter on the screen. We use a 3D arrow that flies from the user's line of sight towards the position of the out-of-view object and returns an acoustic signal on it. The audio signal does not play from the 3D position of the object, but on both speakers in order to indicate that the 3D arrow has reached the out-of-view object (see Figure 5.16). Thereby, we can focus on the visual aspect of this technique.

To evaluate the performance of FlyingARrow, we conducted a user study in Augmented Reality comparing it against EyeSee360 using the Microsoft Hololens. The development of EyeSee360 will be described later in this thesis (see Section 7.1). However, we will describe the fundamental idea of the technique in a dedicated subsection (see Subsection 5.3.2). The reason for breaking out of our linear storyline is that we investigated all three research questions in parallel and to the time we started to explore FlyingARrow, EyeSee360 had been proven to be the best-performing technique.

Here, our research contributions include:

- A multi-modal visualization technique for pointing to out-of-view objects on small screen Augmented Reality devices.
- Comparison of FlyingARrow to the later-described radar-like visualization technique EyeSee360.

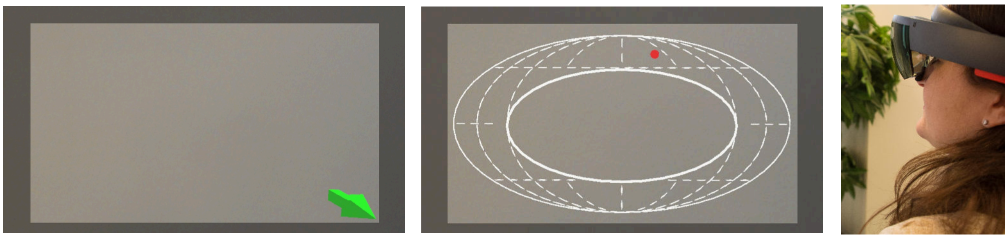
The work presented in this section was published as a full paper at the PerDis conference in 2018 [GLH+18].

5.3.1 Approach

We explore how to guide users toward out-of-view objects 360° around the user. Our technique FlyingARrow uses a combination of visual and auditory cues to point toward out-of-view objects. For guiding toward out-of-view objects the 3D distance between the out-of-view object and the user is not relevant. Therefore, we stick with encoding the 3D directions toward these objects. Our work draws on Section 5.1 and Section 5.2, which show that out-of-view object visualization can be successfully adapted from 2D off-screen visualization techniques to head-mounted Augmented Reality. In a user study, we compare how well our technique FlyingARrow performs in comparison to EyeSee360.

5.3.2 EyeSee360

EyeSee360 (see Section 7.1) is a visualization technique that allows a user to know the direction and distance of out-of-view objects. Figure 5.17b shows how EyeSee360 looks on the Hololens. This grid system compresses 3D position information onto a single 2D plane. The inner ellipse of EyeSee360 represents the



(a) FlyingARrow on Hololens. (b) EyeSee360 on Hololens. (c) Participant.

Figure 5.17: FlyingARrow and EyeSee360 on Microsoft Hololens. In a) and b) the lighter area indicates the Hololens screen.

FOV of the current user, and the area outside the ellipse is outside the user's view. Each dotted line represents a 45° section of the user's view. The horizontal line expresses the altitude of the object. The vertical curved lines represent the direction to the object. For example, the red dot at the upper right part of EyeSee360 represents an object that is almost 45° to the right and more than 45° up. The color encodes the distance from close (red) to far away (blue). More detailed information can be found in the chapter that describes the visualization technique EyeSee360 (see Section 7.1).

5.3.3 FlyingARrow

Our developed technique FlyingARrow combines the visual representation of an out-of-view object with an acoustic signal to point toward the object in 3D space. The combination of both modalities helps to reduce the visual information load [Wic08] and clutter on the screen. As with amodal completion or amodal perception [Mic91, EZ93], where users can mentally complete simple shapes even when only part of the shape is visible, we hypothesize that users can mentally complete a uniform movement when only one part of that movement is visible. Therefore, we use a 3D arrow (cf. Figure 5.17a) that flies with uniform movement from the user's line of sight toward the position of the out-of-view object and returns an acoustic signal on it (see Figure 5.16). We hypothesize that the user can mentally complete the out-of-view movement. We decided to use another modality to inform the user when the 3D Arrow reaches the out-of-view object. We chose sound because all current devices are equipped with audio and because previous work showed that audiovisual cues are useful for reducing reaction times [LBS+17]. The 3D direction of the out-of-view object is encoded by the direction in which the 3D Arrow points, and the distance is encoded by the flight duration of the arrow. We implemented FlyingARrow in the 3D game engine Unity3D. The source code is available as an Open Source project on GitHub[9]

---

[9] Github OutOfView-Project. `https://github.com/UweGruenefeld/OutOfView`, last retrieved April 21, 2020

and supports various Augmented Reality as well as Virtual Reality devices (e.g., Microsoft Hololens, HTC Vive or Oculus Rift).

5.3.3.1 Identifying parameters of FlyingARrow

We identified various parameters to adjust our technique FlyingARrow to small field-of-view devices. Therefore, we derived fitting settings of these parameters from related work or with pilot testing.

Speed of 3D Arrow is connected to flight duration because there is a limited time period in which humans can memorize perceptions. This period is stated by prior work as being up to three seconds (also called the three-seconds-phenomenon) [Pöp02]. Therefore, we adapt the speed of the 3D-Arrow so that it is able to fly from the user's line of sight to the out-of-view object within three seconds.

Size of 3D Arrow is chosen based on findings in pilot testing that investigated the smallest arrow size for which users still could perceive the shape and the direction. Since the field-of-view of the Hololens is rather small, the 3D Arrow had to use at least one-twentieth of the screen space to be easily perceivable by the participants (cp. Figure 5.17a).

Sound variants are useful with regard to the amount of information that can be encoded. Using simple sound coming from both speakers of the Hololens, we could advance to 3D sound coming from the direction of the out-of-view object. Further, we thought of using periodic repeated sounds at one-second intervals to assist in mentally understanding trajectory. However, since none of these worked best during pilot testing, we used the simplest possible sound pattern in which a single sound is given on the out-of-view object from both speakers, without information.

5.3.4 Study Design

To evaluate the performance of our novel visualization technique FlyingARrow, we conducted a within-subjects controlled laboratory study in Augmented Reality with the Microsoft Hololens. Our study's only independent variable was visualization with two levels (FlyingARrow vs. EyeSee360). We used quantitative methods to evaluate user performance, taking search time, search error, and direction error as our dependent variables. Search time is measured as the time a user needs to locate and select an out-of-view object in the scene, while search error is specified as the number of objects a user wrongly selects. The direction error here is the angular error, which is the angle between the user's assessment of the out-of-view object's position and its correct position in 3D space.

For this study, we derived the following sub-question from our first research question: *(RQ1d) Does FlyingARrow perform better than EyeSee360 on a small field-of-view Augmented Reality device with respect to search time, search error, direction accuracy, perceived usability, and workload?*

H_9 We expect FlyingARrow to result in lower search time than EyeSee360.

H_{10} Based on previous work, we hypothesize that uniform movement used in FlyingARrow can be completed in a mentally similar way as amodal completion, and therefore lead to better direction estimation accuracy than EyeSee360.

H_{11} We expect FlyingARrow to have a lower workload than EyeSee360.

5.3.5 Procedure

The within-subjects study was divided into two tasks: a search task and a direction estimation task. Both tasks were divided into two blocks, with each block testing one technique (FlyingARrow, EyeSee360). We counter-balanced the two blocks across all participants. The out-of-view objects were randomly distributed in 3D space. We stored the seeds of the position generation to test the same positions for each technique. However, by randomly picking the order, we ensured that participants would not recognize a previous pattern of positions from the foregoing technique.

At the end of the experiment, we asked participants to fill out a SUS questionnaire and a RAW-TLX questionnaire for each technique. Further, participants were asked to fill out our individual subjective questionnaire and a demographic questionnaire. Overall, each participant took approximately 40 minutes to finish the experiment.

5.3.5.1 Task A: Search Time

Each block of this task started with three test trials (not included in results), along with an explanation of the visualization technique and the task to achieve. In each run of this task, the participant had to find the out-of-view object with the support of the technique by selecting it with a cursor and a remote control. Each block was tested 10 times.

5.3.5.2 Task B: Direction Estimation

Each block of this task started with three test trials (not included in results) and an explanation of the visualization and the task to achieve. In each run of this task, the participant had to estimate the position of a randomly placed out-of-view object. Each block was tested 10 times.

5.3.6 Participants

We recruited 12 participants (5 female), aged between 20 and 54 (M=27, SD= 8.96). None of them suffered from color vision impairment. All had normal or corrected-to-normal vision.

5.3.7 Results

5.3.7.1 Search Task

For the search task, we consider the effects of one factor (visualization) on search time and object selection accuracy (where object selection accuracy means an object was not found during the trial). The mean search times for the visualization techniques are: FlyingARrow=6.24s and EyeSee360=6.01s. The total numbers of wrongly selected objects are: FlyingARrow (40/120=33.3% search error) and EyeSee360 (19/120=15.9% search error). The search times are compared in Figure 5.18.

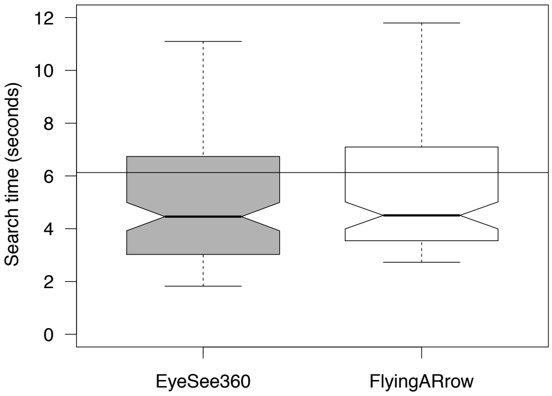


Figure 5.18: Boxplot of median search times of the different visualization techniques (line indicates mean search time).

A Shapiro-Wilk-Test showed that our data is not normally distributed ($p < 0.001$). As we compare two matched groups within subjects, we directly performed a Wilcoxon Signed-rank test. Here we found no significant effect of visualization on search time (W=3312, Z=-0.832, p=0.407, ϕ=0.05). This indicates that FlyingARrow and EyeSee360 do not significantly differ with respect to search time.

5.3.7.2 Estimation Task

We consider the effects of one factor (visualization) on mean direction error. The mean errors for the visualization techniques are: FlyingARrow=33.52° and

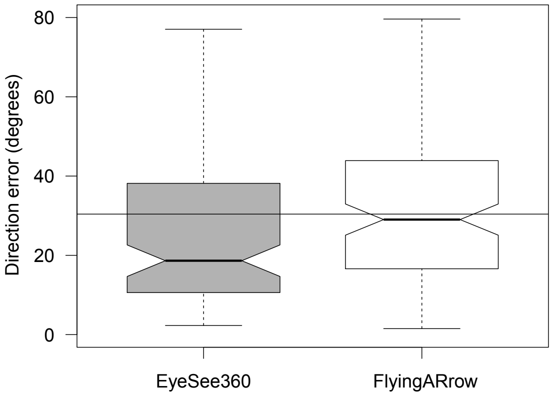


Figure 5.19: Boxplot of median direction errors of the different visualization techniques (line indicates mean direction error).

EyeSee360=27.28°. The direction errors are compared in Figure 5.19. A Shapiro-Wilk-Test showed that our data is not normally distributed ($p<0.001$). As we compare two matched groups within subjects, we performed a Wilcoxon Signed-rank test. Here we found no significant effect of visualization on direction error ($W=2972$, $Z=-1.723$, $p=0.085$, $\phi=0.11$). This indicates that FlyingARrow and EyeSee360 do not significantly differ with respect to estimation accuracy.

5.3.7.3 RAW-TLX

For NASA Raw-TLX [Har06] scores, FlyingARrow scored 39.24 and EyeSee360 scored 46.74. Both values indicate an acceptable workload, though FlyingARrow has a slightly lower workload than EyeSee360.

5.3.7.4 System Usability Scale

For SUS scores, FlyingARrow scored 68 and EyeSee360 scored 51. Therefore, FlyingARrow is over the threshold for acceptable usability, while EyeSee360 is not [Bro96]. In sum, only FlyingARrow is usable on small field-of-view AR devices.

5.3.7.5 Questionnaire

At the end of the study, we asked participants to answer four questions with 5-point Likert items. Participants stated that they were able to easily find the out-of-view objects with FlyingARrow (Md=4, IQR=1.25), while they were neutral for EyeSee360 (Md=3, IQR=1). Furthermore, they stated that they were able to correctly estimate the positions of out-of-view objects with FlyingARrow (Md=4, IQR=1.25), but not with EyeSee360 (Md=2.5, IQR=1.25). Overall, seven participants preferred FlyingARrow while five preferred EyeSee360.

5.3.8 Discussion

5.3.8.1 Advantages of Head-mounted Devices

Since our technique is inspired by off-screen visualization techniques [BR03, GBGI08], it can be perceived as similar and therefore familiar to users. The simply shape of our 3D Arrow is especially easy to understand. Combined with a head-mounted device, our technique can provide on-demand guidance for foveal presentation under more limited FOVs. Showing the visualization technique on-demand is especially helpful for avoiding clutter on devices with limited fields-of-view (e.g., Hololens).

5.3.8.2 FlyingARrow vs. EyeSee360

Our main goal was to transfer Arrow to head-mounted AR devices and to improve usability as well as direction accuracy of the technique. Compared to EyeSee360, we reached that goal. However, FlyingARrow resulted in decreased performance in terms of search time, object selection accuracy, and direction estimation. We believe this is mostly due to lack of understanding of mentally complete uniform movement and may be improved in future work. However, a SUS-score of 68 for FlyingARrow does not indicate good usability. We argue that this score is influenced by the Hololens device. Three participants stated that the Hololens itself felt uncomfortable to wear and that it negatively influenced their rating of both techniques.

5.3.8.3 Amodal Completion for Direction Estimation

Besides the comparison to EyeSee360, our user study showed that users were able to estimate the direction of out-of-view objects or locate them in 3D space. From this, we can assume that mentally complete uniform movement works. However, future work is needed to improve the technique.

5.3.8.4 Multimodal Technique

We showed that pointing to out-of-view objects can be done by splitting to multiple modalities. However, for future work, we suggest testing different modalities such as tactile feedback and investigating redundant encoding of direction to out-of-view objects.

5.3.8.5 Limitations

For each out-of-view object, our technique used a visual cue flying from the user's line of sight toward the out-of-view object, but only once. Therefore, it was difficult for participants to locate an out-of-view object when the visual cue was already gone. In order to overcome this limitation, we suggest either letting 3D Arrows repeatedly fly toward the out-of-view object or keeping 3D Arrows in-view by sticking them at the screen border. Further, future work is needed to evaluate the performance of FlyingARrow in more realistic scenarios (e.g., gaming). We imagine that FlyingARrow may retain advantages over EyeSee360 because of its reduced visual clutter.

5.3.9 Conclusion

In this section, we compared our novel visualization technique FlyingARrow with the previous technique EyeSee360. We showed that perceived usability and workload was lower for our technique FlyingARrow. Additionally, it reduced the amount of clutter added to the user' s screen. However, EyeSee360 objectively performed best regarding direction estimation, object selection accuracy, and search time. However, we showed the potential of using FlyingARrow and mental completion of uniform movement. To improve the direction estimation and search time performances, future work that further explores combinations of different modalities for pointing toward out-of-view objects is required.

5.4 Summary

In this chapter, we adapted existing off-screen visualization techniques to cue direction to out-of-view objects. First, we selected three promising off-screen visualization techniques from related work (Arrow, Halo, and Wedge) and applied them to AR to cue direction to out-of-view objects. Since we did not know how well these techniques would perform, we restricted ourselves to 90 degrees in front of the user and compared the adapted techniques in a first user study (see Section 5.1). We found that Halo objectively performed best while Wedge subjectively performed best.

Thereafter, we selected these two best-performing techniques and further improved them to work for out-of-view objects spatially distributed all around the user. Further, we extended the support to both, Augmented and Virtual Reality. Afterwards, we compared both techniques in two user studies: one in Augmented and one in Virtual Reality (see Section 5.2). While our techniques resulted in overall high usability, we found the choice of AR or VR impacts mean search time (VR: 2.25s, AR: 3.92s) and mean direction estimation error (VR: 21.85°, AR: 32.91°). Moreover, while adding more out-of-view objects significantly affects search time across VR and AR, direction estimation performance remains unaffected.

In the last part of this chapter we improved Arrow, the visualization technique from the first user study that performed worse than the other two techniques, to cue the direction of out-of-view objects spatially distributed in 3D space all around the user. We compared the improved FlyingARrow technique with EyeSee360 in a user study in Augmented Reality (see Section 5.3). We found that perceived usability and workload were better for FlyingARrow than for EyeSee360. However, FlyingARrow performed slightly worse in terms of direction estimation, object selection accuracy, and search time.

6 Extending the Field-of-View of Mixed Reality

Augmented and Virtual Reality are currently experiencing a second spring, triggered by affordable do-it-yourself solutions like Google Cardboard[1] and built upon by Virtual Reality headsets like the Oculus Rift[2] or the HTC Vive[3]. However, compared to the human field-of-view, the field-of-view of current Augmented or Virtual Reality devices is several times smaller, resulting in less immersive experiences and underutilized human visual capabilities. Current head-mounted VR devices are limited to around 110° horizontally, and AR devices to around 40° horizontally. This means that in the user's periphery, either parts of the visual scene are missing in VR or no virtual content is visible in AR. This is partly due to technical limitations, where extending the FOV of such devices requires more pixels to calculate, emits more heat, and results in lower comfort due to increased weight. Importantly, this restricted FOV limits the immersive potential of these systems. This is especially true for VR scenarios, which often rely on user awareness of the positions of out-of-view objects that lie outside the restricted FOV (e.g., opponents in a multi-player game). This leaves visual information-processing capabilities of users underutilized. Furthermore, the experience of users is less immersive because of an abruptly ending display. Therefore, different kinds of light displays have been suggested, with promising results. For example, Xiao et al. showed that simulator sickness can be reduced with peripheral light displays [XB16]. Additionally, these displays can increase situational awareness by displaying additional information. Therefore, we ask:

RQ2: How can Mixed Reality devices with small fields-of-view be extended to present directional cues to out-of-view objects?

To address this research question, we started by developing a low-cost prototyping tool for developing head-mounted peripheral light displays based on LEDs. We call this prototyping tool PeriMR. This tool is built on top of Google Cardboard, to support both video see-through Augmented Reality and Virtual Reality, and allows researchers to develop new peripheral head-mounted light displays for Mixed Reality (see Section 6.1).

Thereafter, we investigated how well humans can perceive radial peripheral light displays that use multicolor LEDs and how well these displays perform with respect to cueing direction to out-of-view objects in Virtual Reality. We started with Virtual Reality since it offers clean lab conditions, such as no influence by external light. To investigate radial peripheral displays, we developed RadialLight, a low-cost radial peripheral display that augments existing head-mounted devices, implemented as 18 radially positioned LEDs around each eye to cue direction. RadialLight is built using the prototyping tool PeriMR. We first in-

---

[1] Google Cardboard. `arvr.google.com/cardboard`, last retrieved April 21, 2020

[2] Oculus Rift. `www.oculus.com/rift`, last retrieved April 21, 2020

[3] HTC Vive. `www.vive.com`, last retrieved April 21, 2020

vestigated direction estimation accuracy of multi-colored cues presented to one versus two eyes. We then evaluated direction estimation accuracy and search time performance for locating out-of-view objects in two representative 360° video VR scenarios (see Section 6.2).

Thereafter, we transferred our results to Augmented Reality and developed MonoculAR, a radial light display presented to one eye. We developed two different versions of MonoculAR. Both consist of twelve radially positioned LEDs, but one presents them as off-screen LEDs and the other presents them as virtual LEDs on the screen. In a user study, we compared both systems to evaluate which of the two approaches results in shorter search time for locating out-of-view objects in Augmented Reality (see Section 6.3).

In the last part of this chapter, we developed a peripheral light display that works as low-cost standalone glasses and allows guiding of the user's attention to hazardous objects in the environment. Here, we looked at the concrete scenario of pedestrians navigating traffic while being distracted by their smartphones and risking collisions with other traffic participants. We developed our prototype with peripheral displays and technically improved it with the experience of five usability experts. Afterwards, we conducted an experiment on a treadmill to evaluate the effectiveness of collision warnings in our prototype. During the experiment, we compared three different light stimuli (instant, pulsing and moving) in terms of response time, error rate, and subjective feedback (see Section 6.4).

6.1 PeriMR: Prototyping Peripheral Light Displays

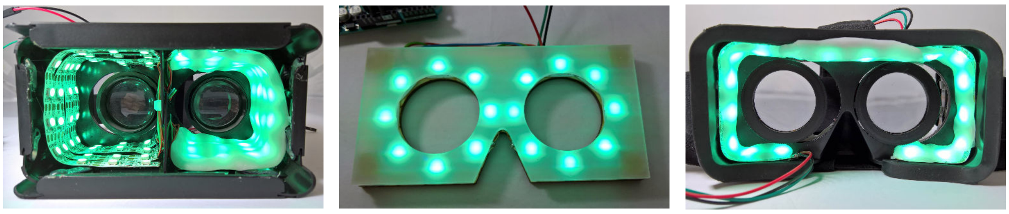

(a) Behind lenses. (b) In front of lenses. (c) Sidebars.

Figure 6.1: Prototypes for testing the design space.

Mixed Reality devices with small fields-of-view can be extended to cover a larger area of the human field-of-view. This is helpful for many usecases, such as to reduce simulator sickness, and may also be useful for displaying additional information about relevant objects in the environment to the user. However, to investigate this, we first need a tool that assists in prototyping different ways of extending the screen. Different prototyping tools have been suggested. For example, PapAR from Lauber et al. transfers the well-known method of paper-prototyping to AR [LBB14]. Two layers are used. One is a normal sheet of paper on which the environment is drawn. Another is a transparent foil that shows the overlaying digital information. However, since peripheral displays must consider the human perception characteristics, this approach is not viable for research on information presentation in the periphery. A similar problem arises from other prototyping approaches that focus on interface elements (e.g., [SC12]). To our knowledge there is no research that addresses prototyping for head-mounted peripheral light displays for Mixed Reality. Therefore, we developed a low-cost prototyping tool for developing head-mounted peripheral light displays that is easy to extend. Our prototyping tool is built on top of Google Cardboard to support video see-through Augmented and Virtual Reality. First, we tested the design space for such a prototyping tool (see Figure 6.1). Then, we improved the Google Cardboard platform to support generic peripheral light displays. These displays can be attached to the cardboard by sliding them into a drawer. Additionally, we added an affordable microcontroller board with Wi-Fi and a battery to allow usage in a mobile context. The prototyping tool can be built using a laser cutter.

Here, our research contributions include:

- A low-cost and highly adjustable prototyping tool for peripheral light displays in Augmented and Virtual Reality.

The work presented in this section was published as a demo paper at the MobileHCI conference in 2017 [GSHB17].

6.1.1 Requirements for VR and AR Prototyping

As the first step, we identified three different characteristics that must be fulfilled by our prototyping tool:

Extendability Our prototyping tool should support video-see-through Augmented and Virtual Reality devices. Further, it should support the use of different lenses and should allow LED placement on various positions.

Flexibility The prototyping tool should be reconfigurable to support different design patterns. Reconfiguration should be possible within a user study.

Low-cost The prototyping tool should be low-cost and publicly available.

6.1.2 Concept of PeriMR

As a first step, we investigated the design space for peripheral LEDs on head-mounted video see-through Augmented and Virtual Reality devices. We tested three positions: between lenses and smartphone display (behind lenses), between eyes and lenses (in front of lenses) and attached to the frame (sidebars). In addition to positioning, we investigated different kinds of diffusors: Gorilla Plastic[4], milky plexiglass, and without diffusors. The prototypes are shown in Figure 6.1.

The outcome of the pre-testing with the developed prototypes was that the best positions for the LEDs were in front of the lenses or on the sidebars. For the sidebars, no diffusor was needed. For in front of the lenses, semi-transparent plexiglass worked best. Behind the lenses also worked to some extent, but due to reflections in the lenses, only low-resolution LED arrays make sense here. Since the testing of behind the lenses needed further exploration, we left this position out of the scope of our design space for this version.

As mentioned already, our prototyping tool was developed based on the official Google Cardboard template. We modified the template using SketchUp[5]. During two iterations, we changed the following points to fit the template to our needs:

1. We changed the proportions and increased the space before the lenses to allow more LEDs and displays to be applied in this space. Further, it allows the participant to wear glasses along with PeriMR.
2. We made the plane including the lenses detachable by adding a drawer with a slide-in. The plane can be changed, thereby allowing support for different LEDs and display patterns. Furthermore, different sizes of lenses can be used by this approach. An example is described in the next section.
3. Additionally, we created two easy-to-attach sidebars, one for each side. These can be used for testing with LED strips on them.

After cutting all parts we assembled the prototyping tool and then added a NodeMCU developer board[6] that is programmed with Arduino and has a low-cost Wi-Fi board attached and one Li-Po battery. These components are lightweight, affordable and allow mobile usage of the prototyping tool.

The resulting 3D model is shown in Figure 6.2: Figure 6.2a) shows the assembled PeriMR, Figure 6.2b) shows the component parts, and Figure 6.2c) shows the replaceable parts.

6.1.3 Prototyping with PeriMR

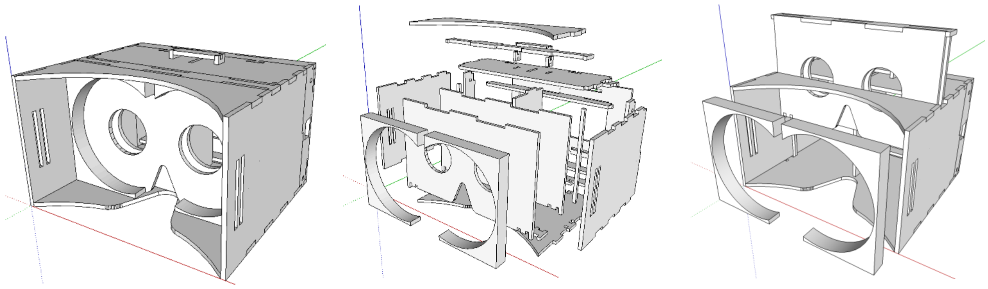
Key to our prototyping tool is the drawer for replaceable slide-ins. For demonstration, we created an example slide-in shown in Figure 6.3. This example uses a radial projection of LEDs in two rows around the center. We used a semi-transparent plexiglass as a diffusor. Creating new slide-ins is one way to build

---

[4] Gorilla Plastic. `www.gorilla-plastic.com`, last retrieved April 21, 2020

[5] SketchUp. `www.sketchup.com`, last retrieved April 21, 2020

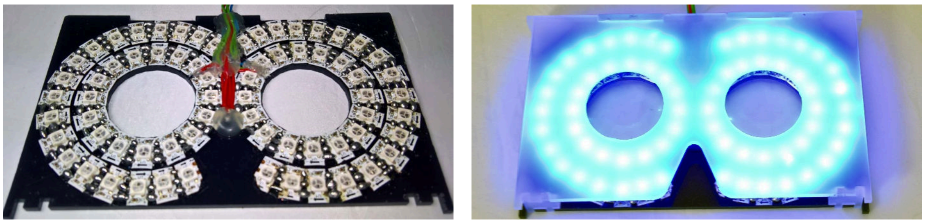
[6] NodeMCU. `en.wikipedia.org/wiki/NodeMCU`, last retrieved April 21, 2020



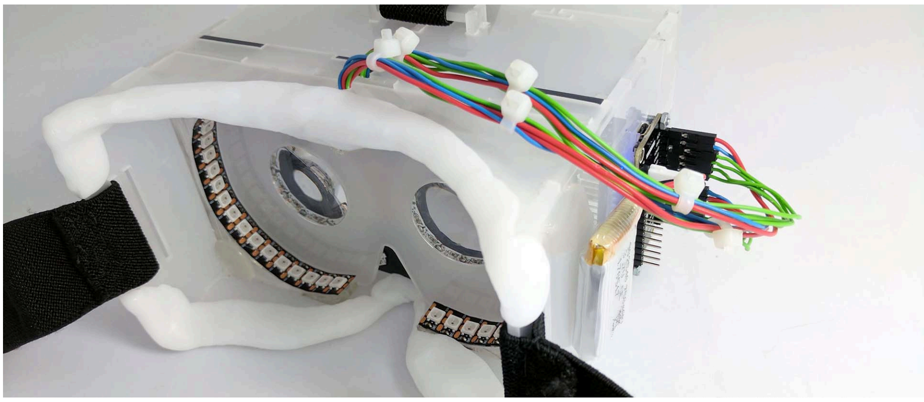
(a) Assembled. (b) Component parts. (c) Replaceable parts.

Figure 6.2: Different views of the 3D model of PeriMR.

prototypes with PeriMR. Another way is to add new detachable sidebars (see Figure 6.2c). Both ways can be used by altering the template[7], which is publicly available on GitHub.



(a) Circuits. (b) Completed slide-in.



(c) Completed prototype which was build using PeriMR.

Figure 6.3: Example of a prototype build with PeriMR.

6.1.4 Limitations

For our prototyping tool we see four main limitations:

---

[7] PeriMR Template. `www.github.com/UweGruenefeld/PeriMR`, last retrieved April 21, 2020

1. It builds on Google Cardboard and uses a smartphone for the Augmented and Virtual Reality experience. Therefore, the graphic power for the main experience is rather limited, but it is acceptable since the peripheral light displays are the focus of the investigation.
2. We recommend that a laser cutter is used to build the prototype. Building it with other tools may take longer and is not as accurate.
3. The attachment of the sidebars is not perfect yet. One must glue them to the frame to fix them for usage. This means they cannot be replaced during studies. In a new iteration, the sidebars should also be detachable as well.
4. Our prototyping tool does not support the use of curved light displays.

6.1.5 Conclusion

In this section, we presented an affordable and simple-to-build prototyping tool for peripheral light displays. The tool is built upon Google Cardboard and adds an easy-to-use mechanism for testing different setups. However, in the next step, we want to use our prototyping tool PeriMR to develop light display prototypes and evaluate these prototypes with users.

6.2 RadialLightVR: Investigating Radial Displays in Virtual Reality

(a) Frontal view. (b) Zoom in view. (c) Study participant.

Figure 6.4: The developed RadialLight prototype.

To support directional cueing of out-of-view objects in Augmented and Virtual Reality, we build on prior work (cf., Xiao and Benko's SparseLight [XB16]) and propose RadialLight, a radial peripheral light display (see Figure 6.4). Compared to SparseLight, RadialLight consists of equally distributed radial LEDs, making full use of the 360° space available. RadialLight is implemented as a proof-of-concept prototype to aid directional cueing (with relative direction mapping to select objects). This helps to avoid rendering the full environment, and can aid in drawing attention to objects in the user's periphery. Moreover, since this display can be lightweight and inexpensive to construct, it can easily be modified to augment existing HMDs. Furthermore, showing the directions of out-of-view objects is one common example of how one could use radial light displays. In RadialLight, we ensured that each LED can be perceived with the same accuracy regardless of position. This is especially important in critical situations (e.g., when pointing in the direction of approaching danger).

In this section, we explore how accurately participants can perceive radial peripheral light displays that make use of different colors to visualize directional information (e.g., the direction of an out-of-view object). Furthermore, we investigate the effect that different background scenarios have on direction estimation, where we introduce a search task to test the ecological validity of our results. In our experiments, we show that participants could not distinguish between LED cues presented to one or both eyes simultaneously, participants estimated LED cue direction within a maximum of 11.8° average deviation, and out-of-view objects in less-distracting scenarios (ship bridge) were selected more quickly but with similar directional accuracy. Our findings support prior work showing that peripheral displays can be useful in expanding the FOV in HMDs [XB16].

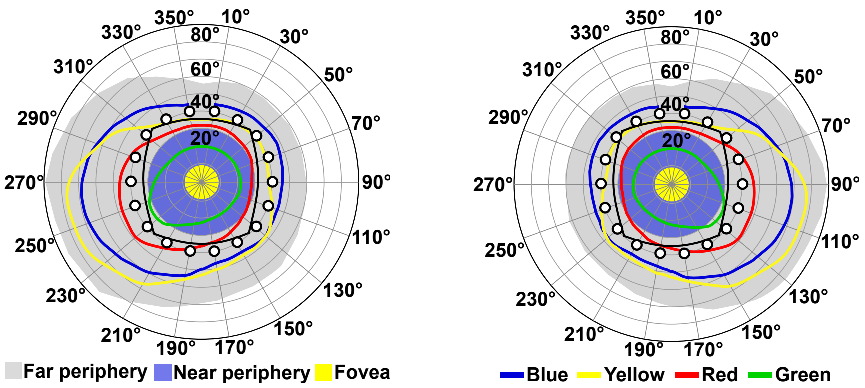
Here, our research contributions include:

- An empirical evaluation that shows the effectiveness (in terms of directional cue accuracy and out-of-view object search time) of our system in 360° VR scenarios.
- We introduce RadialLight, a low cost radial peripheral display that augments existing HMDs, implemented as 18 LEDs radially positioned around each eye to cue direction.

The work presented in this section was published as a full paper at the MobileHCI conference in 2018 [GSA^{+}18].

6.2.1 RadialLight System

RadialLight (see Figure 6.4) was built using the prototyping tool PeriMR (cf. Section 6.1). It is based on prior work from Xiao et al. [XB16] and uses the Google Cardboard platform, which combines a smartphone with cut cardboard to create a VR and video see-through AR device. We modified Google Cardboard to include laser-cut plexiglass as the diffusor and to ensure a more solid foundation. We added 18 radially positioned and individually addressable RGB LEDs (WS2812B) around each eye to cue direction towards out-of-view objects. To control LEDs, we used a NodeMCU developer board (ESP8266) that is programmed with Arduino and has a low-cost Wi-Fi board attached that serves as a Wi-Fi access point. The board is powered by a Li-Po battery (3.7V). We developed a REST-API to directly change LEDs over Wi-Fi from a Google Pixel XL smartphone via Web Requests. As such, RadialLight is a standalone headset that does not require connection to any external device.



(a) Left eye.

(b) Right eye.

Figure 6.5: RadialLight's LED placement relative to human FOV. Black line: RadialLight's smartphone FOV. Color lines: human color perception. Circles: LEDs. *Best seen in color.*

RadialLight's FOV with LEDs around each eye is shown in Figure 6.5. The human monocular FOV [PJW73] showing foveal, near and far peripheral vision for each eye is shown in yellow, blue, and gray, respectively. Given related work [Kal01, SRJ11], we used colors that are perceivable outside the smartphone's FOV: yellow, blue and white. We adapted our LED placement to fit the nearest perceivable color, e.g., yellow. For our experiment, we used these colors (blue, yellow, white), which map to one, two, or all RGB channels, respectively. To ensure optimal viewing, we placed LEDs in radial formation around the users' eyes to ensure every LED would be perceived with similar perceptual characteristics, given the decreasing level of detail a human eye perceives with increasing radial distance [SRJ11].

To guide head movement using directional cues, we followed our previous approach (see Section 5.2) in which an imaginary sphere around the user's head is used to map all out-of-view objects onto the sphere and as a result have a normalized vector into the direction of the out-of-view object. If we consider the user's line of sight also as a point on that sphere, an imaginary line on the surface of the sphere between these two points indicates the head movement necessary to see the object. Here, head movements are limited to 90° vertically and 180° horizontally. Therefore, we restricted the FOV to 180° in front of the user, wherein each shortest path on the sphere becomes the optimal head movement towards the object. An advantage of our approach is that it easily extends to 360° around the user, since cues are independent of head movements. Practically, this means there are no differences between 90°, 145°, or even 180°. However, since we cue direction towards out-of-view objects, we suggest cueing only one object at a time. Otherwise, most of the time many LEDs would be lit (at least those near the horizon, since even virtual environments tend to place objects along a ground plane) and it would be unclear where the user should look.

6.2.2 Study I: Directional Accuracy and LED Color

We first investigate directional accuracy across colors in RadialLight to find the average direction deviation error across LED positions. We tested LED color differences using either a background (white) or no background (black space). Second, we investigated user performance when cues were presented to one versus both eyes.

6.2.2.1 Study Design

To evaluate directional accuracy across LED colors, we ran a lab-based user study in an empty white-walled room with darkened windows (to avoid effects of different light conditions). Furthermore, RadialLight was designed to minimize incoming light (by placing material around the frames). Our experiment consisted of a 3 colors (blue vs. yellow vs. white) x 2 backgrounds (no background vs. white light) repeated measures design, where we measured cue directional accuracy (our dependent variable) as well as subjective Likert-scale measurements. Furthermore, we tested differences between switching on LEDs to both eyes versus one eye.

For this study, we derived the following sub-question from our second research question: *(RQ2a) How do radial peripheral displays that use multicolor LEDs affect direction estimation performance?* Given that this was an exploratory study, we did not posit specific hypotheses. However, we expected high performance in direction estimation given the radial nature of the two LED ring displays, where each set of LEDs has an equal distance to the eye. Furthermore, we explored different LED colors to determine the best-perceived and most-preferred colors for inclusion in our following study (see Subsection 6.2.3). Also, participants were asked to state whether they saw LEDs on one versus both eyes after each trial. We did not examine search time for the first study since our focus here is on the impact of different colors and directional cues of radial lights rather than the performance of RadialLight in specific usecases, such as searching for out-of-view objects.

6.2.2.2 Procedure

For both our experiments, we ran pretests to determine optimal color luminance, choice of input device, and HMD calibration. To ensure LEDs were perceivable but not too bright compared to the backgrounds tested, we relied on the unicolor model to provide the same luminance across all colors [MKS$^+$05]. For the input device, we tested a modified numeric pad, scroll wheel, and jog dial. Jog dial was most suitable, as it felt natural to indicate direction without requiring one to look at a screen. To ensure all participants had the headset centered, we placed a black circular pointer on the lens and smartphone display for calibration purposes (see Figure 6.6).

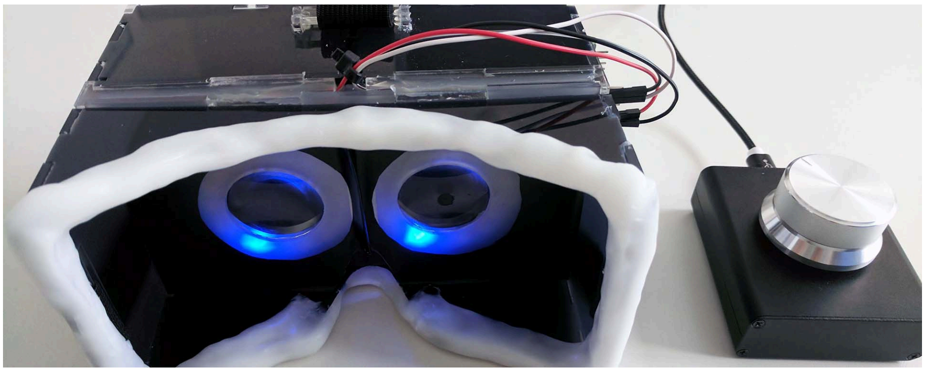


Figure 6.6: Study apparatus of first study (brightness of LEDs increased for better visibility).

Participants signed a consent form, then sat near a desk where they wore our RadialLight prototype (see Figure 6.4c). After a short study introduction, participants calibrated the HMD and underwent a tutorial during which each of the six conditions (blocks) was presented. Thereafter, we tested each condition in one block, resulting in six total blocks. For all conditions, we tested all 18 LEDs on both eyes once (1944 trials), and four runs per condition when the LEDs were on one eye only (two left; two right) (432 trials total). This was done based on pretests that showed participants do not perceive such differences. Presentation order was randomized. For each trial, one direction cue is shown to participants for 5 seconds, where the cue duration was empirically determined to be suitable. Thereafter, the cue is disabled and the participant has to specify the cue direction with a jog dial (see Figure 6.6) that controls an arrow on the screen. Experiment sessions lasted approximately 30 minutes.

6.2.2.3 Participants

We had 18 volunteer participants[8] (9 female), aged between 21-53 years (M= 29.06, SD=4.43). All participants had normal or corrected-to-normal vision and none had any color vision impairment.

6.2.2.4 Results

We investigated differences in direction estimation accuracy across LED colors, background conditions, and presentation on one vs two eyes. As a Shapiro-Wilk test showed that our data is not normally distributed ($p<0.001$), we conducted Wilcoxon rank-sum tests to check for significant effects for each of our independent variables and their conditions on direction estimation accuracy.

---

[8] Based on a Latin-square design with six conditions. For mean effect sizes of ($f=0.20$), at least 164 observations are necessary, which requires testing at least 9 participants.

LED Color

The average deviation (error) for direction estimation was 10.16° (SD=7.79°). A pairwise Wilcoxon test with Holm-Bonferroni adjustments revealed that the average deviation for blue LEDs was significantly lower than that for yellow LEDs (W=82368, Z=-3.37, $p<0.01$, ϕ=0.12). The median error for both conditions was the same (Md=10°, IQR=10°) and the means differed by 1.36°. These show that, while all colors performed well in cueing direction, use of the blue LED led to the lowest error.

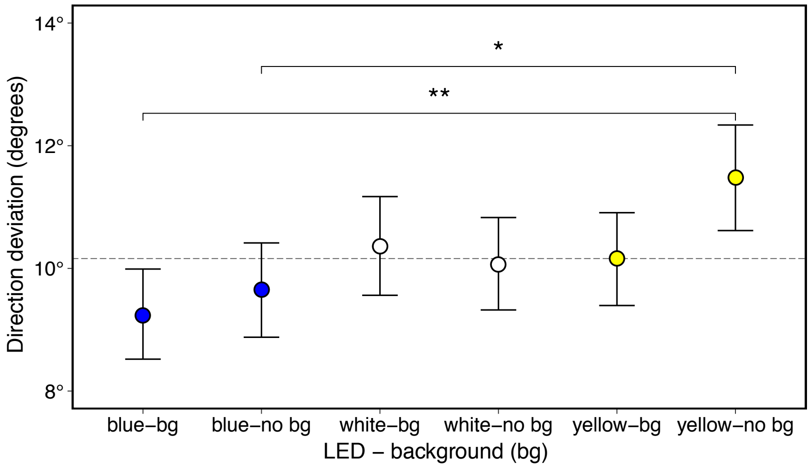


Figure 6.7: Average error and 95% confidence intervals for direction estimation accuracy (dotted line marks the average error for all conditions).

Background

A Wilcoxon Rank Sum test did not reveal any significant differences between background (M=9.92°, SD=7.63°) and no background (M=10.40°, SD=7.95°) across all colors (W=235500, Z=1.00, p=0.27, ϕ=0.03), which shows that a plain background did not influence users' direction estimation accuracy.

LED Color vs. Background

A pairwise Wilcoxon test with Holm-Bonferroni adjustments showed significant differences in direction estimation accuracy between blue-background and yellow-no background (W=19240, Z=-3.89, $p<0.01$, ϕ=0.10). Furthermore, it showed significant differences in direction estimation accuracy between the condition blue-nobackground and yellow-nobackground (W=20510, Z=-3.09, $p<0.05$, ϕ= 0.08). Figure 6.7 shows the average deviation and 95% confidence intervals of input error (degrees) for LED colors across background conditions. This supports our finding that blue LEDs outperform other colors irrespective of background.

One vs. Two Eyes

We looked into differences in average direction deviation for cues presented to one (M=10.66°, SD=8.36°) or two eyes (M=10.05°, SD=7.66°) simultaneously. To handle sample imbalances for LEDs shown on one versus two eyes, we downsampled the two eyes group (N=1944) to the one eye group sample size (N=432). We used random downsampling without replacement and tested for differences (Wilcoxon signed-rank test) in direction error across 1000 sampling runs (seeds)[9]. To combine probabilities, we used Fisher's method[10] [KM02]. When p-values tend to be small, the test statistic X^2 will be large, which suggests the null hypotheses are not true for every test. Here again, we found no statistically significant difference between one versus two eyes (χ^2(1, N=2,000)=2038.26, p=0.27). Furthermore, only three participants noticed differences between eye conditions, where they stated they saw the LED light only on one eye (18/432, 0.04%), of which they were correct in 50% of the cases. These findings indicate that LEDs being shown on one vs two eyes does not affect direction estimation accuracy.

6.2.3 Study II: Cue Direction and Search in 360° Video

To ensure external validity, we tested RadialLight in two 360° video scenarios (car cockpit and ship bridge) using blue LED color for directional cues. We measured directional accuracy and cue search time for out-of-view objects. The scenarios can be seen in Figure 6.8. The ship bridge consisted of a tugboat and surrounding water. Most of this video was simply blue without much to see. The car cockpit, on the other hand, included oncoming vehicles as well as pedestrians walking on the left and right sidewalks.

6.2.3.1 Study Design

Given that we observed no significant differences between LED colors in the first study, we chose blue LEDs to represent directional cues since they showed the lowest error on average. Like our first study, our second study was designed as a lab study and took place in the same environment. Our experiment consisted of two tasks and followed a repeated-measures design with 360° video as the independent variable with two levels (car cockpit vs. ship bridge). Both scenarios provided a first person experience of either a ship or a car moving (the latter with pedestrians and activity on the road). The first task was direction estimation, where our dependent variable was the user's estimation of object's direction based on the cue. The second task was an out-of-view object search task, where our dependent variable was search time performance.

---

[9] Since p-value combination under Fisher's method follows a X^2-square distribution, we needed a minimum of 220 runs to achieve 0,95 power and 0.3 effect size under α=0.05.

[10] This is a common method used for aggregating probabilities, however we tested other methods (e.g., voting) and results did not differ.

For this study, we derived the following sub-question from our second research question: *(RQ2b) How well does RadialLight perform with respect to cue directional accuracy and out-of-view object search time performance in 360° video scenarios?* Given our exploratory work, we did not posit hypotheses. However, we expected that the car cockpit scenario would be more distracting because of the pedestrians and oncoming vehicles and would therefore result in lower performance across each task than would the ship bridge scenario.

(a) 360° video ship bridge snapshot. (b) 360° video car cockpit snapshot.

Figure 6.8: The two tested 360° video scenarios.

Task 1: Direction Estimation

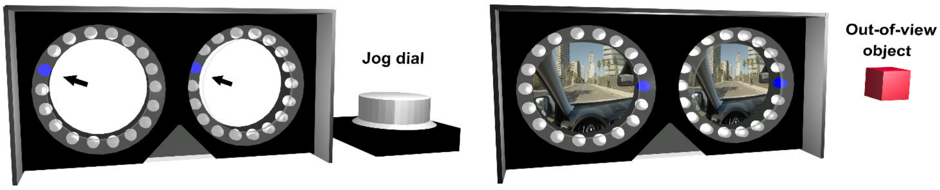
The first task was like study I's estimation task, but here we changed the levels of our independent variable to ship bridge and car cockpit 360° videos (see Figure 6.8). We measured the angle deviation between the LED position and the user's subjective assessment.

Task 2: Out-of-View Object Search

Here, we randomly distributed virtual objects (occupying 5° of 20° directional view) in positions exactly 90° out of view. Each LED indicates one object. We influenced the same independent variable as in the first task and measured the time it took for the participant to locate the object.

6.2.3.2 Procedure

The procedure of this study was identical to that of study I, except for the following: here we showed only blue LEDs for directional cues and let users experience our two 360° video scenarios (car cockpit and ship bridge) instead of plain backgrounds. After each cue presentation, the participant indicated direction using the jog dial (see Figure 6.9a). The search task followed a similar procedure, but in this case the user only had to turn their head to locate the cued out-of-view object (Figure 6.9b), at which point the object would disappear. Participants had a cursor representing their gaze, which they used to select the cued out-of-view objects. Search time was stopped when the participant successfully selected the out-of-view object. Then, the participant turned to face the front and begin the next trial.



(a) Arrow manipulated with jog dial. (b) Out-of-view object on right side.

Figure 6.9: Study setup; a) Direction estimation task with background condition, b) Out-of-view object search task with car cockpit condition.

6.2.3.3 Participants

In this study, we had 12 volunteer participants[11] (4 female), aged between 21-38 years (M=26.75, SD=4.43). All participants had normal or corrected-to-normal vision and no color vision impairment.

6.2.3.4 Results

We investigated the effect of different 360° video VR scenarios (car cockpit and ship bridge) as backgrounds on ensuring ecological validity of cue direction estimation accuracy. Furthermore, we measured search time performance. A Shapiro-Wilk test showed that our data on direction estimation accuracy (p<0.001) and search time performance (p<0.001) are not normally distributed. Thereafter, we conducted Wilcoxon rank-sum tests to check for significant effects of our independent variables on direction estimation accuracy and search time performance.

360° VR Scenario: Direction Estimation

The average deviation for direction estimation was 11.8° (SD=10.58°) for the car cockpit scenario and 11.2° (SD=9.48°) for the ship bridge scenario. We found no significant effect of scenario on direction estimation accuracy using a Wilcoxon rank-sum test (W=12746, Z=0.70, p=0.45, ϕ=0.03), which shows that cue direction is identifiable regardless of the tested scenario.

---

[11] For mean effect sizes of (f=0.20), at least 164 observations are necessary, which requires testing at least 9 participants. We calculated this value with G*Power under Wilcoxon signed-rank test (α=0.20 and $1-\beta$=0.80).

360° VR Scenario: Search Time

There was a significant effect of scenario on search time performance (W=20042, Z=2.06, $p<0.05$, $\phi=0.09$). This shows that a more distracting environment, such as a first person experience in a moving car, slows users in finding out-of-view objects. Figure 6.10 shows the average direction deviations and 95% confidence intervals for the car cockpit (M=3.67°, SD=1.38°) and ship docking (M=3.34°, SD=0.78°) scenarios.

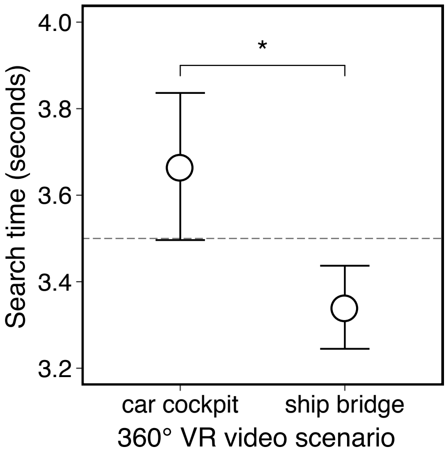


Figure 6.10: Median selection time performance and 95% confidence intervals. The dashed line marks the mean selection time performance for both conditions.

6.2.4 Implications

Our work has implications for building peripheral HMDs and for designing peripheral directional cues:

Radial Peripheral LEDs Suitable for Directional Cueing

From both studies, we observed that every direction was perceived with nearly the same average angle deviation (e.g., no significant differences were observed), which indicates that radial LED placement for encoding direction towards out-of-view objects is a suitable approach.

LED Color Does Not Strongly Affect Performance

While we found in study I that blue LED color resulted in lower direction estimation error than the other tested colors (white, yellow), the highest average error observed is still low (<12°). This indicates that choice of peripheral LED color does not strongly affect user performance for direction estimation tasks.

Radial Monocular Display Sufficient for Binocular Vision

Participants did not recognize the difference between LEDs being shown to one eye versus both eyes, as evidenced by the fact that participants could not identify the trials in which the cue was presented to one eye. Furthermore, direction accuracy using RadialLight was not significantly affected. This highlights that having a single monocular display with a single LED switched on is sufficient for peripheral direction cueing, which helps to reduce the power consumption and cost of such displays.

360° VR Scenario Complexity Increases Object Search Time

While average direction deviation differences between the ship bridge (11.2°) and car cockpit (11.8°) scenarios showed no significant effects, search time performance for locating out-of-view objects was affected. While we only tested two 360° video scenarios, this effect on search time may be more pronounced across more distracting scenarios or with increasing user engagement (e.g., in a ship monitoring situation).

Limitations

Since we place LEDs in the periphery of users, light reflections experienced when wearing RadialLight should be avoided. Although we took measures to avoid this (e.g. by using black tape around the LEDs; see Figure 6.6), participants reported that some reflections occurred (1.3% of cases, N=2914). However, this did not affect estimation performance. We did not test multiple LED color combinations within a task; however, based on related work (e.g., [Kal01]) we believe participants can easily distinguish between colors due to radial LED placement. Finally, we only investigated 360° video scenarios, and not more engaging VR scenarios (e.g., games where the user is involved). While this was beyond the scope of our work, we suspect user performance will generally drop as a function of engagement.

6.2.5 Conclusion

We introduced RadialLight and explored LED-based directional cueing for locating out-of-view objects. Evaluating RadialLight in two user studies, our findings highlight the usefulness of directional cueing in such peripheral displays and in expanding the FOV in HMDs (cf., [XB16]). While we evaluated our system in 360° video VR scenarios, we believe our results are more generally applicable to AR, VR, and mixed reality environments. In the next section, we want to compare our approach to one that uses on-screen visual cues, removing the need for extra hardware to be added to the HMD. Further, we want to transfer our peripheral display to Augmented Reality.

6.3 MonoculAR: Single Radial Display for Augmented Reality

(a) MonoculAR from a user's perspective.

(b) View from the side.

Figure 6.11: MonoculAR attached to Microsoft HoloLens.

In this section, we explore the utility of augmenting head-mounted Augmented Reality devices with MonoculAR, a peripheral light display, to point towards out-of-view objects. MonoculAR is designed as twelve radially positioned light cues with two implementations: 1) on-screen virtual light cues and 2) off-screen LEDs attached around the device's display. We evaluate the performance of both implementations in a controlled user study with two different tasks. Users must search for specific out-of-view objects in both tasks, but in one task only one out-of-view object is presented at a time, while in the other task multiple out-of-view objects are visible simultaneously. In this section, we want to evaluate whether virtual light cues presented on the screen can eliminate the need to add extra hardware, since each occupies only a few pixels of the screen. Furthermore, it is relevant how well these two implementations perform compared to each other.

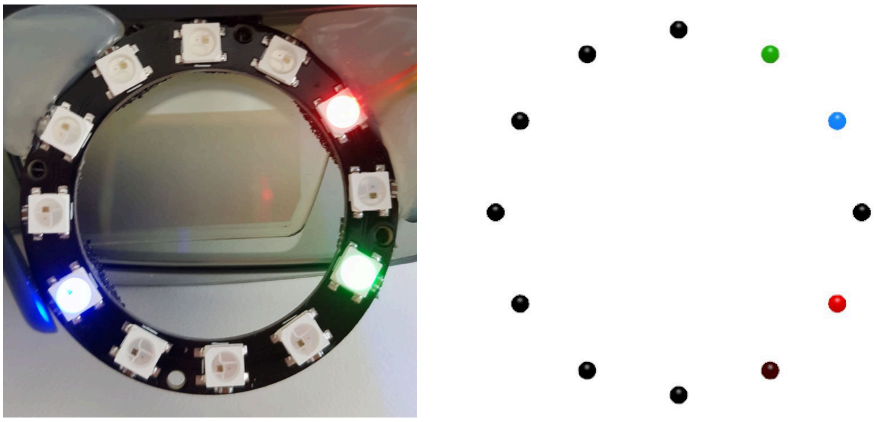
Here, our research contributions include:

- Two implementations (virtual, physical) of radially positioned light cues.
- A comparative evaluation of these two implementations for pointing towards out-of-view objects on Augmented Reality devices.

The work presented in this section was published as a poster paper at the MobileHCI conference in 2018 [GSPH18].

6.3.1 MonoculAR System

MonoculAR is inspired by the SparseLight system presented by Xiao and Benko [XB16]. However, it differs in various aspects. Instead of augmenting both eyes, MonoculAR augments only one eye with twelve radially positioned light cues (see Figure 6.12a). This makes MonoculAR a low-cost solution that can be easily attached to existing hardware. Furthermore, we discuss a second implementation of our proposed system that uses on-screen virtual light cues (see Figure 6.12b). Both implementations aim to help users locate out-of-view content.



(a) Off-screen LEDs. (b) On-screen virtual light cues.

Figure 6.12: Two different implementations of MonoculAR.

6.3.1.1 Directional Cues

To cue directions towards out-of-view objects, MonoculAR maps out-of-view objects onto an imaginary sphere around the user's head onto a point A to remain with their direction information (see Section 5.1). The current user's head pose is then also mapped to an according point B on that sphere. Next, the shortest path[12] on the sphere between the user's head pose B and the out-of-view object A is drawn. The exit angle of that path at point B is then used to determine which LED on the ring is used to point towards the out-of-view object. Thereby, users can simply turn their heads to the direction of the illuminated LED to find the according out-of-view object. The color used is based on the color of the out-of-view object. If one visual cue is used to represent multiple out-of-view objects, then the objects' colors are combined.

6.3.1.2 Hardware Implementation

The hardware of MonoculAR was built by combining an LED ring with a WiFi-supported microcontroller. We added twelve radially positioned and individually addressable RGB LEDs (WS2812B) around one eye to cue direction towards out-of-view objects. To control LEDs, we used a NodeMCU developer board (ESP8266) programmed with Arduino, which serves as a Wi-Fi access point. The board is powered by a Li-Po battery (3.7V). We developed a REST-API to directly change LEDs over Wi-Fi via Web Requests. As such, MonoculAR is a standalone headset that does not require connection to any external device. The brightness of the LEDs was adjusted to match the brightness of the HoloLens display by matching the average brightness values of the colors.

---

[12] Great Circle Distance. `www.wikipedia.org/wiki/Great-circle_distance`, last retrieved April 21, 2020

6.3.1.3 Software Implementation

We implemented MonoculAR in the 3D game engine Unity3D. The on-screen virtual light cues are implemented as light-emitting spheres that appear black when no direction is cued. The refresh rate of these on-screen virtual lights is synced with the refresh rate of the hardware LEDs to make them comparable. Furthermore, the virtual light cues are placed towards the border of the Hololens display to be as close as possible to the positions of the hardware lights.

6.3.2 Experiment

To compare both implementations of MonoculAR, we conducted a controlled user study with our prototype.

6.3.2.1 Study Design

To evaluate the performance of each implementation of MonoculAR, we conducted a within-subjects controlled laboratory study in Augmented Reality using the Microsoft Hololens. Our study's only independent variable was implementation, with two levels (on-screen vs. off-screen). We used quantitative methods to evaluate user performance, taking search time and search error as our dependent variables. Search time is measured as the time a user needs to locate and select an out-of-view object in the scene, while search error is specified as the number of objects a user wrongly selected.

For this study, we derived the following sub-question from our second research question: *(RQ2c) Does the off-screen implementation of MonoculAR perform better than the on-screen implementation on a small field-of-view Augmented Reality device with respect to search time, search error, and perceived usability?* We posit the following hypotheses:

H_{12} We expect the on-screen implementation to result in lower search time than the off-screen implementation (because the on-screen visualization is closer to the user's focus and therefore may be more perceivable).

H_{13} We expect the off-screen implementation to be subjectively perceived best (because the user's screen does not become visually cluttered by the cues).

6.3.2.2 Procedure

The within-subjects study was divided into two search tasks. Each task was divided into two blocks, with each block testing one implementation (off-screen vs. on-screen). We counter-balanced the two tasks and two blocks across all participants. The out-of-view objects were randomly distributed in 3D space. However, we avoided spawning objects visible on the screen. We stored the seeds of the position generation for each task to test the same positions for each implementation. However, by choosing the order at random, we ensured that participants would not recognize a previous pattern of positions from the foregoing implementation.

After all blocks, we asked participants to fill out our individual subjective questionnaire and a demographic questionnaire. Overall, each participant took approximately 30 minutes to finish the experiment.

Task A: Search Task (Single Object)

In this task, participants had to search for an out-of-view object that was the only existing object in the environment. The object was a white cube, so the visual cue shown to the participants was also white. To start a trial of this task, participants had to focus on a point directly in front of them. When a participant clicked to indicate their readiness, the virtual cube appeared in the 360° space around and a white LED cued its direction. Upon finding the out-of-view object, participants had to select it with the cursor, at which point the time was stopped. Each implementation was tested in ten trials. There was one additional training trial in the beginning, but this is excluded from the results.

Task B: Search Task (Multiple Objects)

In this task, participants had to search for an out-of-view object that was presented along with two other objects in the environment. In each trial there were three cubes: one red, one blue, and one green. To start a trial of this task, the participant had to focus on a point directly in front of them. When the participant clicked to indicate their readiness, the three virtual cubes appeared in the 360° space around them. Next, a random color was chosen from the three options and the participant was asked to locate the cube of that color. All three cubes were visualized on MonoculAR during this task (see Figure 6.12). When the participant found the out-of-view object, they had to select it with their cursor. The time was then stopped. Each implementation was tested in ten trials. There was again one additional trial in the beginning for training, which is excluded from the results.

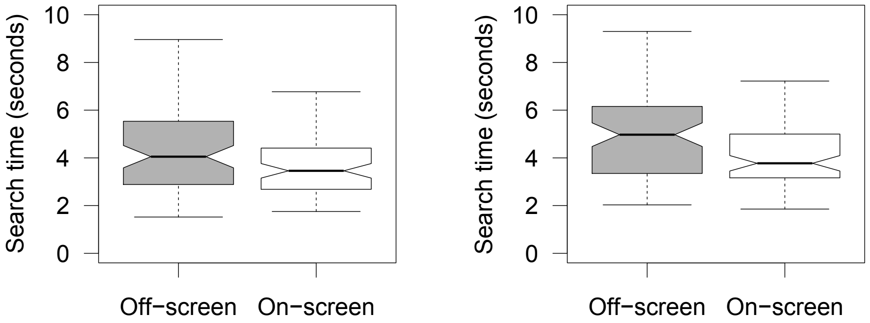
6.3.2.3 Participants

We had 8 volunteer participants (4 female), aged between 23-31 years (M=26.25, SD=2.49). All participants had normal vision apart from one, who had corrected-to-normal vision. None suffered from color vision impairment. We asked participants to answer two 5-point Likert-items to rate their experience with head-mounted devices and the Microsoft Hololens. Participants stated that they had more experience with head-mounted devices (Md=4, IQR=1.25), while they had less experience with the HoloLens (Md=2.5, IQR=1.25).

6.3.2.4 Results

Task A: Search Task (Single Object)

For the search task with a single object shown, we consider the effects of one factor (implementation) on search time. The mean search times for both implementations are: on-screen=3.78s and off-screen=4.33s. The search times are compared in Figure 6.13a. A Shapiro-Wilk-Test showed that our data is not normally distributed ($p<0.001$). As we compare two matched groups within subjects, we directly performed a Wilcoxon Signed-rank test. Here we found only minimal evidence for an effect of implementation on search time ($W=1230$, $Z=-1.87$, $p=0.060$, $\phi=0.15$).



(a) Task A: Single object. (b) Task B: Multiple objects.

Figure 6.13: Boxplots of median search times (in seconds) of both implementations (off-screen vs. on-screen) for both tasks.

Task B: Search Task (Multiple Objects)

For the search task with multiple objects, we consider the effects of one factor (implementation) on search time and object selection accuracy (where object selection accuracy refers to the number of wrongly selected objects per trial). The mean search times for the implementations are: on-screen=4.54s and off-screen=5.06s. The total number of wrongly selected objects is zero for both implementations. The search times are compared in Figure 6.13b. A Shapiro-Wilk-Test showed that our data is not normally distributed ($p<0.001$). As we compare two matched groups within subjects, we directly performed a Wilcoxon Signed-rank test. Here we found a significant effect of implementation on search time ($W=1114$, $Z=-2.43$, $p=0.015$, $\phi=0.19$).

Comparison Between Search Tasks

To compare the two different tasks, we consider the effects of one factor (task) on search time. The mean search times for the tasks are: Task A=4.05s and Task

B=4.80s. A Shapiro-Wilk-Test showed that our data is not normally distributed ($p < 0.001$). As we compare two matched groups within subjects, we directly performed a Wilcoxon Signed-rank test. Here we found a significant effect of implementation on search time (W=4555, Z=-3.21, p=0.001, ϕ =0.18).

Questionnaire

At the end of the study, we asked participants to answer four questions with 5-point Likert items. Participants stated that they could easily find the one out-of-view object with both on-screen MonoculAR (Md=5, IQR=0.25) and off-screen MonoculAR (Md=4, IQR=1). In the second task in which several objects were presented, participants stated that they could easily find the out-of-view object with on-screen MonoculAR (Md=4, IQR=1.25) but were neutral for off-screen MonoculAR (Md=3, IQR=0.5). Overall, seven participants preferred on-screen MonoculAR while only one preferred off-screen MonoculAR.

6.3.3 Discussion

6.3.3.1 On-Screen vs. Off-Screen Implementation

For both tasks, the implementation that used on-screen virtual light cues worked better than the one that used off-screen LEDs. While these results were only significant for the task in which multiple objects were shown, we believe that significance could also be achieved for the single-object search task by using a larger sample size. Here, we accept our hypothesis H_{12}. We argue that on-screen visualization outperforms off-screen visualization due to the proximity of the light cues to the user's focus. Thereby, the light cues of the on-screen visualization can be perceived more easily and participants can better distinguish between the colors [Kal01, SRJ11]. Participants subjectively rated the implementations in line with our quantitative results. Therefore, we cannot accept our hypothesis H_{13}. In our hypothesis, we argued that the on-screen visualization would clutter the screen and therefore be subjectively perceived as worse. However, our study was designed as a controlled lab study, so no other visual content was visible on the screen except the out-of-view objects. Therefore, visual clutter did not affect the results of our study.

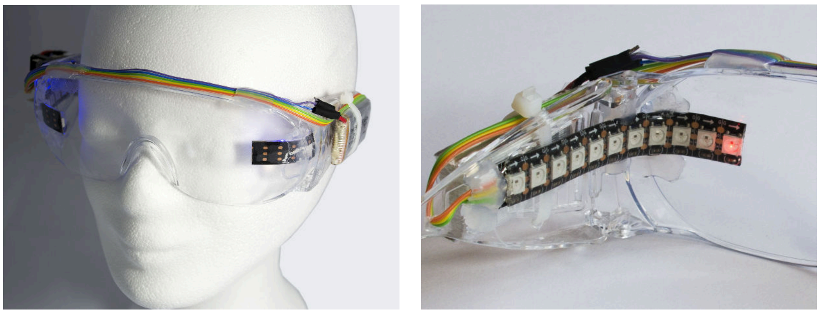
6.3.3.2 Multiple Out-of-View Objects

Our results show that having more out-of-view objects leads to higher search times. We think this is due to overlapping of cues (e.g., when one out of twelve LEDs must indicate the directions towards two out-of-view objects). Furthermore, color perception may have an effect here because some colors can be more easily distinguished than others [Kal01]. We suggest, if possible, cueing only cue one out-of-view object at a time.

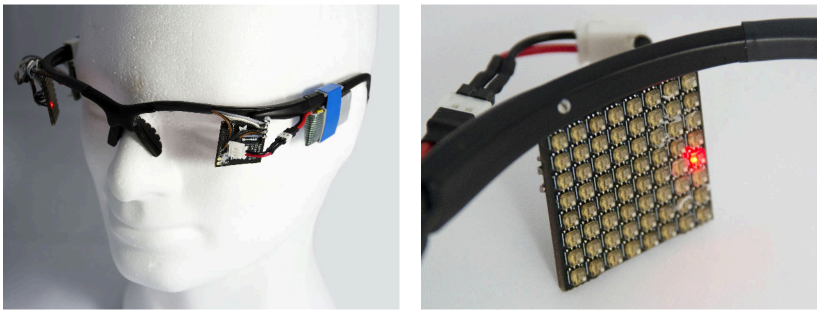
6.3.4 Conclusion

In this section, we presented MonoculAR, a technique for pointing to out-of-view objects on Augmented Reality devices. The technique supports two different implementations: 1) on-screen and 2) off-screen visualization. In a first study, we explored search time performance of both implementations. We showed that the on-screen variant is preferred by users and results in shorter search times. Moreover, our results suggest that visualizing only one out-of-view object at a time improves search time performance. Future work is required to investigate additional influencing factors, such as visual clutter.

6.4 Guiding Smombies: Low-Cost Glasses for Guiding Attention



(a) First iteration of prototype. (b) Display of first prototype.



(c) Second iteration of prototype. (d) Display of second prototype.

Figure 6.14: Two iterations of our low-cost glasses with peripheral vision for guiding attention.

Augmented Reality (AR) allows one to overlay digital content onto the real world, in order to alter perception of it. This digital content can be explored using various devices (e.g., AR glasses or smartphones). Nowadays, smartphones are widely available and often used for exploring AR content. There have been some studies in which the degree of fidelity (e.g., rendering quality [KSF10, SJS14], refresh rate [LBS^{+}16], registration accuracy [SFS^{+}07, MSS11], etc.) was not high enough for the user to feel immersed. Thus, the presence of the real world overshadowed the virtual. However, the recent emergence of increasingly capable mobile hardware and the optimization of tracking solutions have enormously enhanced the fidelity of Augmented Reality content.

This is especially interesting with regard to AR games, which have recently spiked in popularity (e.g., Pokémon Go[13] or Ingress[14]). Playing AR games on smartphones becomes problematic when it is done while navigating traffic (e.g., as a pedestrian [Eys17] or as a driver [ALD$^+$16]). Interestingly, the simplicity of the game is what makes it so popular [HP18] (besides the technically improved immersive experience of course). However, playing such games on small form-factor devices makes the experience like that of looking through a keyhole while concentrating one's focus on a single spot. In this case, higher immersion in the game is not beneficial. Although using a smartphone while driving a car or riding a bike is forbidden in most countries, it is still allowed for pedestrians. However, the combination of playing AR games or otherwise using a smartphone and navigating traffic is highly dangerous. This is especially true when the user walks while using a smartphone [SBS11, NHW08]. There are several accident reports that provide evidence for this, while the number of near or minor accidents is likely even higher [Eys17]. In Germany, "Smombie" was the 2015 "youth word of the year." It combines the words "smartphone" and "zombie" to refer to the intensely unaware state of people walking around staring at their phones like zombies[15]. This is worsened by the fact that navigating in traffic is a fundamental requirement of such games.

In this section, we developed low-cost peripheral AR glasses to support pedestrians in critical traffic encounters and to evaluate different visual cues for shifting user attention (see Figure 6.14). Our first iteration was based on safety glasses with attached LED strips. Thereafter, we developed a second iteration of our prototype based on feedback from five expert interviews. Then we conducted an experiment on a treadmill to evaluate the suitability of three different types of visual cues (instant, pulsing, moving) (see Figure 6.15). Participants were given a smartphone game with an N-back task in order to engage them in a task and simulate workload.

Here, our research contributions include:

1. A novel display which augments peripheral vision with warning information to draw the user's visual attention towards potential hazards.
2. An evaluation of three different visual cues for guiding user attention.

The work presented in this section was published as a poster paper at the ISMAR conference in 2018 [GSJ$^+$18].

---

[13] Pokémon. `www.pokemon.com`, last retrieved April 21, 2020

[14] Ingress. `www.ingress.com`, last retrieved April 21, 2020

[15] Smartphone zombie. `en.wikipedia.org/wiki/Smartphone_zombie`, last retrieved April 21, 2020

6.4.1 Approach

To support pedestrians in critical traffic encounters, we aim to augment the user's peripheral vision with low-cost glasses. In our approach, we limited ourselves to the most frequent traffic encounter, in which a car is approaching from either the left or right side of the user. For our proposed solution, we followed related work providing evidence for 1) good perception of peripheral displays, 2) motion as a well-perceived cue in the periphery, and 3) on-body visual cues successfully guiding user attention. We implemented motion with a combination of LEDs positioned in such a way that they are unaffected by different head poses typical for smartphone users. We identified three different visual cues as possible candidates for guiding a user's attention. The first visual cue "instant" is our baseline condition, while the second "moving" and third stimuli "pulsing" are based on prior work that concluded that motion is perceived well in the periphery [Lyo16, LDRR$^+$16, LBS$^+$17] (cf. Figure 6.15). To test the different conditions, we developed a prototype with peripheral displays like the smart glasses from Nakao and Kunze [NNCK16]. Unlike their approach, our displays use multicolor LEDs and are rotated towards the user for better perceptibility. As a first step towards the development of well-perceivable moving cues in the periphery, we tested a high-density LED strip attached to a pair of safety glasses. We discussed the prototype with usability experts and received recommendations for a technically improved version. As a second step, we built an improved version of our prototype and evaluated its performance for displaying three different visual cues in a controlled user study.

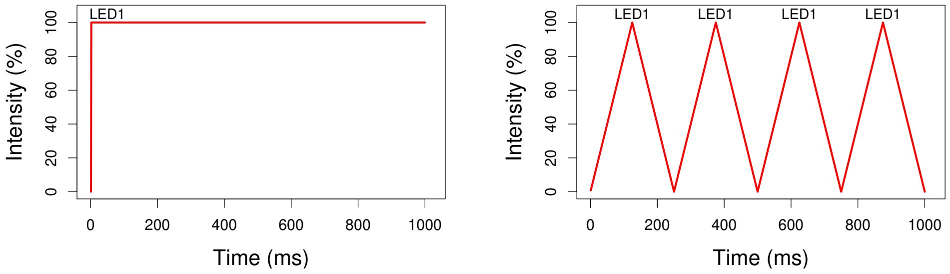
6.4.2 LED Strip Prototype

We started to investigate whether animated cues are suitable for shifting attention towards either the left or right side of the user with a low fidelity prototype. This first prototype was based on the combination of safety glasses and LED strips. We used LED strips with the highest density available for consumers (144 LEDs per meter). They consisted of RGB LEDs of the type WS2812B[16]. The LED strips were positioned to be in the periphery of the user [Kal01]. We added a NodeMCU developer board programmed in Arduino with a low-cost Wi-Fi board attached and one Li-Po battery. These components are lightweight, affordable, and allow mobile usage of the LED strip prototype. Further, we set up a web interface to manipulate the stimuli shown by the LED strips. The source code is available under MIT License on Github[17]. The LED strip prototype can be seen in Figure 6.14a and 6.14b.

---

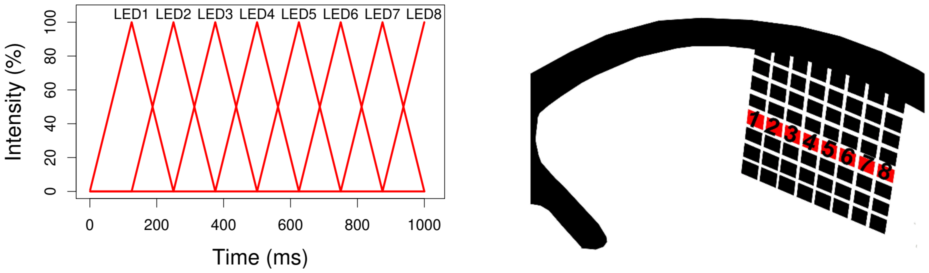
[16] Adafruit LED WS2812B. `cdn-shop.adafruit.com/datasheets/WS2812B.pdf`, last retrieved April 21, 2020

[17] Github GuidingSmombies-Project. `www.github.com/UweGruenefeld/GuidingSmombies`, last retrieved April 21, 2020



(a) Stimulus instant.

(b) Stimulus pulsing.



(c) Stimulus moving.

(d) LED order (right side).

Figure 6.15: Evaluated visual cues for guiding attention.

6.4.3 Interview with Usability Experts

In order to gain early insights for our developed prototype, we conducted interviews with usability experts. In these interviews, we wanted to explore whether our lo-fi prototype is suitable for showing animated cues for guiding user attention. We further wanted to explore fitting parameters for our visual cues (e.g., speed of "moving" and "pulsing" stimuli).

6.4.3.1 Procedure

We decided to do the interviews about our first prototype with each usability expert separately in order to avoid any influence they could have on one another. The interview consisted of two parts. In the first part, we introduced the expert to the problem we wanted to solve and our approach. The expert was then asked to give comments on the idea and our prototype. In the second part, the expert could try out the prototype and test the different visual cues. They could also change the different parameters (e.g., speed, movement direction, color, or brightness). Each pretest lasted about 20 minutes.

6.4.3.2 Participants

We interviewed 5 experts (2 female), aged between 26 and 35 (M=30, SD=3.56). All of them had at least three years of experience in human-computer interaction.

6.4.3.3 Results

Animated Light Stimuli

For the "moving" visual cues, we asked if the cue should move towards the hazard or away from it. However, compared to other domains in which users preferred the cue moving away (e.g., driving [BCHB16, SGT09]), here the experts agreed on the visual cue moving towards the hazard. One expert stated that he perceived the cue as if his head were being dragged towards the hazard.

LED Placement

All experts mentioned several aspects regarding the LEDs and their placement. All agreed that the distance between the LEDs was too big for visual cues to be perceived as motion in the periphery. Further, it became clear that the LED strip used for the first iteration of our prototype did not need to extend very far into the periphery, as the outer LEDs could not be perceived. Furthermore, two experts had the problem that the LED strips were not in line with their eyes. Therefore, they perceived the stimuli as too high or too low.

Usability

Three experts stated that no lenses are necessary for the functionality of the first prototype. Therefore, they suggested that we remove them, to both decrease the weight and avoid possible reflection of the visual cues. It was also mentioned that, due to clutter, the user loses the ability to perceive peripheral information of the environment to some extent.

6.4.4 LED Matrix Prototype

To overcome the limitations of the first prototype, we decided to use an LED matrix instead of an LED strip and to remove the lenses from the frame (see Figure 6.14c). By taking this approach, we could increase the LED density to 200 LEDs per meter per row (4 LEDs per square centimeter). Further, with an LED matrix, it is possible to adjust the visual cue to the user's line of sight by picking a different row on the matrix for cue presentation. However, with the bigger dimensions of the matrix, the problem of too much clutter in the periphery remains. The type of LED was the same as in the LED strip prototype.

6.4.5 Experimental Evaluation

To evaluate the performance of animated visual cues in our low-cost peripheral vision glasses (second prototype cf. Figure 6.14c), we conducted a within-subjects controlled laboratory study in Augmented Reality with an Android smartphone.

6.4.5.1 Study Design

Our study's only independent variable was peripheral visual cue, with three levels (instant vs. pulsing vs. moving), where "instant" is the baseline condition. We used quantitative methods to evaluate user performance, taking response time and error rate as our dependent variables. Response time is measured as the time from the presentation of a visual cue on the peripheral vision glasses to the participant's verbal identification of a letter appearing on a display laterally behind them. We recorded this response in two ways: 1) the director of the study pressed a button, and 2) voice activation stopped a timer. The error rate is specified as the percentage of stimuli to which a user wrongly reacted.

For this study, we derived the following sub-question from our second research question: *(RQ2d) In how far can peripheral vision be augmented to visually guide the attention of pedestrians in critical traffic encounters?* Here, we posit the following hypotheses:

H_{14} We expect the "moving" cue to result in faster responses than "instant."

H_{15} We hypothesize that the cue "instant" results in the highest error rate.

6.4.5.2 Apparatus

Our apparatus consisted of a treadmill and two displays positioned laterally behind the participant, one on the left side and one on the right. The displays were placed 135° to the left and right to prevent participants from perceiving changes on the display. Here, we decided to move even further into the periphery, as specified in our requirement, because we hypothesized that moving light stimuli would be most effective here.

6.4.5.3 Participants

We recruited 8 participants (2 female), aged between 22 and 31 (M=26, SD= 3.25). None of them suffered from color vision impairment. All had normal or corrected-to-normal vision. Participants with corrected vision wore contact lenses.

6.4.5.4 Procedure

In the beginning, the experimenter explained the procedure and asked the participant to sign a form of consent. Then the participant put on the peripheral vision glasses and went on the treadmill (see Figure 6.16). A safety clip attached to the participant's clothing ensured an emergency stop of the treadmill in the event that a participant tripped. In addition, there was an emergency stop button always within reach of the participant. To ensure that the participant could always press this button, we did not use any additional hand-held device other than the smartphone. The speed of the treadmill was fixed at 2.5 km/h.

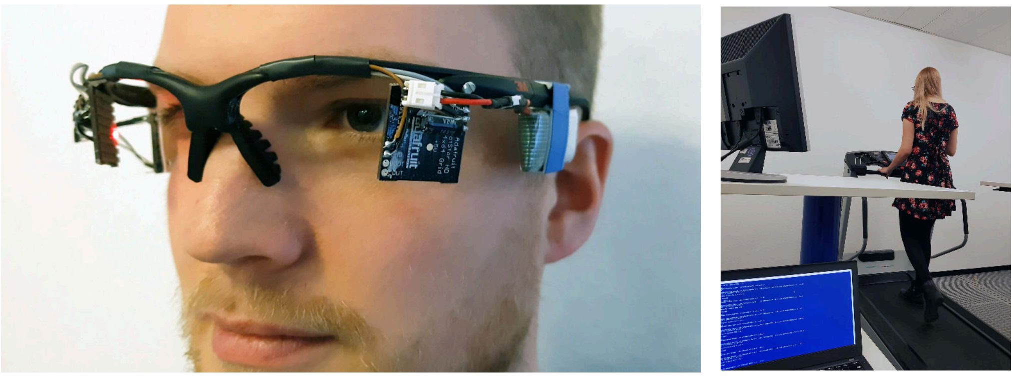

(a) The prototype used in the experiment.
(b) Study participant.

Figure 6.16: The apparatus of the experiment.

Participants held a smartphone in their dominant hand and played a two-position-back memory game (primary task). The other hand was free to hold on to the treadmill or press the stop button. Every 20-40 seconds, the peripheral vision glasses displayed a randomized visual cue to the left or right side. Simultaneously, the respective display showed a random uppercase letter for a duration of five seconds. Participants were instructed to immediately react to the stimulus and read the letter out loud (secondary task). There were 30 stimuli in total and all stimuli were represented equally.

In the end, the participants were asked to answer a questionnaire, rating the three stimuli on a five-point Likert scale for the statements: "I could see the light stimulus very quickly," "I felt the direction indicated by the light stimulus as intuitively understandable," and "I found the light stimulus alarming." Further, we asked the participants to pick their favorite stimulus and collected demographic data. Experiment sessions lasted approximately 30 minutes.

6.4.5.5 Results

Error Rate

We consider the effects of three different conditions (instant, pulsing, and moving) on error rate (where error rate refers to how frequently participants did not react to a given visual cue or stated the wrong letter). Participants were correct in 100% of the trials independent of condition type, resulting in a 0% error rate in all conditions.

Response Time

We consider the effects of three different conditions (instant, blinking, and moving) on participants' response times. We measured response time using only the key press of the study director. We discarded the vocal response times due to the frequent noise (e.g., squeaking of shoes on the treadmill) that resulted in many false positives. However, there was a significant positive correlation between key press response time and vocal response time (Spearman ρ=0.46, p<0.001). The mean response times for the different conditions were: instant=2.29s (SD=0.48s), pulsing=2.28s (SD=0.53s), and moving=2.15s (SD=0.54s). A Shapiro-Wilk-Test showed that our data is not normally distributed (p<0.001). We therefore ran a Friedman test, which revealed a significant effect of condition type on response time ($\chi^2(2)$=6.93, p=0.031, N=8). A post-hoc test using Wilcoxon Signed-rank with Holm-Bonferroni correction showed significant differences between some conditions (see Table 6.1). The response times are compared in Figure 6.17. To conclude, "moving" is significantly faster than both "instant" and "pulsing", which did not significantly differ from each other.

Table 6.1: Pairwise comparison of different conditions in the study.

Comparison	**p-value**	**ϕ-value**
instant vs. pulsing	0.640	0.04
instant vs. moving	0.032	0.19
pulsing vs. moving	0.032	0.20

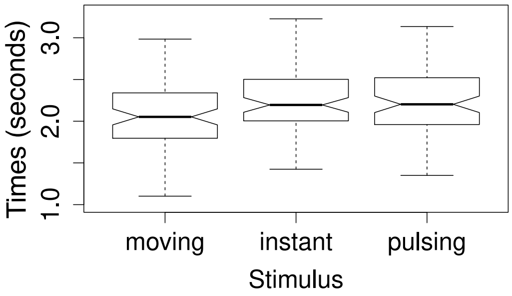


Figure 6.17: Response times of different visual cues.

Subjective Questionnaire

At the end of the study, we asked participants to answer three questions for each condition (instant, pulsing, and moving). The questions were five-point Likert-items. Participants stated that they could see the visual cues very quickly for "instant" (Md=3.5, IQR=2.25), "pulsing" (Md=4, IQR=1.25), and "moving" (Md=5, IQR=0). They stated that they felt the direction indicated by the visual cue was intuitively understandable for all cues equally (Md=5, IQR=0). Furthermore, they stated that they found the visual cues alarming for "instant"

(Md=2.5, IQR=1.5), “pulsing” (Md=4, IQR=1.25), and “moving” (Md=3.5, IQR=1). Overall, seven participants preferred “moving” while one participant preferred “instant”.

6.4.6 Discussion

6.4.6.1 Advantages of Peripheral Visual Cues

Visual cues presented in the periphery using a head-mounted device offer the possibility of shifting the user’s attention on-demand without cluttering the main visual field. Our results showed that visual cues are easily perceivable in the periphery. In our hypothesis H_{15}, we expected the “instant” stimulus to result in a higher error rate. However, participants did not make any errors. Therefore, we cannot accept our hypothesis H_{15}.

6.4.6.2 Perception of Motion

Based on prior work, we hypothesized that motion, specifically motion over position, would result in faster responses H_{14} (cf. [SRJ11, LDRR+16, LBS+17]). Therefore, we investigated two different visual cues that change over time (where “pulsing” changes its intensity and “moving” changes its position). Our results showed that “moving” resulted in significantly faster response times than “pulsing” or “instant”. Therefore, we can accept our hypothesis H_{14}.

6.4.6.3 Further Applications

Besides using our developed prototype for guiding a user’s attention towards approaching cars or any other critical objects in the environment, we can also imagine using the peripheral displays for less critical scenarios (e.g., notifications or navigation). This is highlighted by previous work that already demonstrated the usefulness of such head-mounted displays for unobtrusive notifications [CIP+06], alarms [CHB17], warnings [NFEL17], or navigation tasks [PHF+12].

6.4.7 Conclusion

In this section, we evaluated three different visual cues for guiding attention of walking smartphone users to the left or right side to locate potential danger. The visual cues were presented in the periphery of the user using our peripheral vision prototype, which was developed in two iterations. Our results showed that visual cues are well suited for guiding the attention of a smartphone user and that the “moving” cue results in a significantly faster response. In later research, we combined our prototype with image processing and machine learning for detection of approaching cars in traffic situations [JLC+18]. In the future, it may be useful to evaluate our motion cue under more realistic circumstances and to combine our prototype with existing AR glasses.

6.5 Summary

In this chapter, we investigated in how far the field-of-view of Mixed Reality devices can be extended to present directional cues to out-of-view objects. First, we developed a low-cost prototyping tool for developing peripheral light displays. Then we used that tool to build RadialLight, a prototype using 18 radially positioned LEDs. In a first study, we evaluated the directional accuracy with which participants could perceive different LED colors presented to one or both eyes. In a second study, we introduced two 360° videos and investigated their influence on direction accuracy and search time performance. Key findings were that participants could not distinguish between LED cues presented to one vs. both eyes simultaneously, participants estimated LED cue direction within a maximum 11.8° average deviation, and out-of-view objects in less distracting scenarios were selected more quickly.

Thereafter, we transferred our results to Augmented Reality and built MonoculAR, a one-eye radial display for cueing direction to out-of-view objects. In a user study, we compared two implementations of the system: on-screen virtual LEDs and off-screen LEDs. We showed that the on-screen variant is preferred by users and results in shorter search times. Moreover, our results suggest visualizing only one out-of-view object at a time can improve search time performance.

In the last section of this chapter, we developed low-cost glasses that help guide the attention of smartphone users in hazardous situations. In a user study, we compared three visual light cues in terms of their resulting average response times and error rates. Overall, we could show that all light stimuli were suitable for guiding user attention (100% correct). However, moving light resulted in significantly faster responses and was subjectively perceived best.

7 Encoding the Out-of-View Object's Position

In the previous two chapters, we developed different visual cues for directing attention to out-of-view objects, with cues presented either on the screen (see Chapter 5) or on LEDs radially positioned around the screen (see Chapter 6). However, many situations require not only the directions to the out-of-view objects, but also the distances to them. For example, when an object is both occluded and out of view, a user looking in the right direction may still be unable to locate the object because they cannot see it. Situations in which monitoring the out-of-view objects is only a secondary task and users cannot turn their heads multiple times to locate all out-of-view objects are also problematic, especially in situations in which many out-of-view objects must be located (e.g., ship docking [GSB+18]). Therefore, a visualization technique for displaying directions and distances of multiple objects at the same time is required. In our previous chapter (see Chapter 5), we developed visual cues that cue direction to out-of-view objects. Here, it may be good to extend these cues to support visualizing distance information as well. Nevertheless, these cues use 3D shapes that add visual clutter to the screen and are not well suited for small field-of-view devices (e.g., Hololens) or the visualization of many objects at the same time. Therefore, in this chapter, we develop a new visualization technique that uses visual cues to encode positions (directions and distances) of out-of-view objects. However, when encoding the positions of out-of-view objects, we also have to consider out-of-view objects that change their positions over time (e.g., the tugboats during the ship docking process [GSB+18]). Therefore, we ask:

RQ3: In what way can the directions and distances of out-of-view objects be visualized for moving or non-moving objects in Mixed Reality?

To answer this research question, we start by developing a visualization technique for encoding the positions of non-moving out-of-view objects. Before, we used Google Cardboard for prototyping new visualization techniques (see Chapter 5). However, developing early prototypes with Google Cardboard has some limitations (e.g., it takes a while to implement changes). Therefore, we developed a prototyping tool that allows us to quickly test and refine design ideas for head-mounted Augmented Reality. Our prototyping tool can be used before prototyping with another platform such as Google Cardboard to quickly collect user feedback. Thereafter, we developed a novel visualization technique called EyeSee360, which encodes the directions and distances of multiple out-of-view objects at the same time. We tested three different variants of EyeSee360 in a user study using the developed prototyping tool. Afterwards, we selected the best variant and compared it to our adapted off-screen visualizations (Arrow, Halo, and Wedge) in a second user study (see Section 5.1) using video see-through AR (Google Cardboard). Furthermore, in a second part of that user study, we investigate how well EyeSee360 encodes the distances to out-of-view objects and how fast users can locate out-of-view objects with the technique (see Section 7.1).

Thereafter, we transfer EyeSee360 to optical see-through Augmented Reality and compare all different variants of EyeSee360 to one other in terms of search time performance using the Hololens. We did this because current optical see-through AR devices suffer from small fields-of-view and because the variant that was preferred in the previous section adds visual clutter and thereby may negatively affect search time performance. Furthermore, the different variants have only been compared in our prototyping tool in terms of direction estimation accuracy, not search time performance. Therefore, we conducted a user study with the Microsoft Hololens to evaluate which variant of EyeSee360 results in the shortest search times (see Section 7.2).

In the last part of this chapter, we investigate in what way direction and distance of a moving out-of-view object can be visualized most effectively. Here, we hypothesize that Contextual views approaches such as EyeSee360 may perform worse because they present the positions of out-of-view objects relative to the user. While this kind of visualization may be helpful for quickly locating an object or knowing it's position, it is unclear how beneficial it would be when objects change position over time. Therefore, we selected 3D Radar (well-known from computer game implementations) and compared it to our developed visualization technique EyeSee360. In this regard, we conducted a user study in which we distinguished between three types of movement (linear, distance, and orbital movement). However, since current Augmented Reality devices suffer from very limited fields-of-view, we conducted this study in Virtual Reality (see Section 7.3).

7.1 EyeSee360: Direction and Distance of Out-of-View Objects

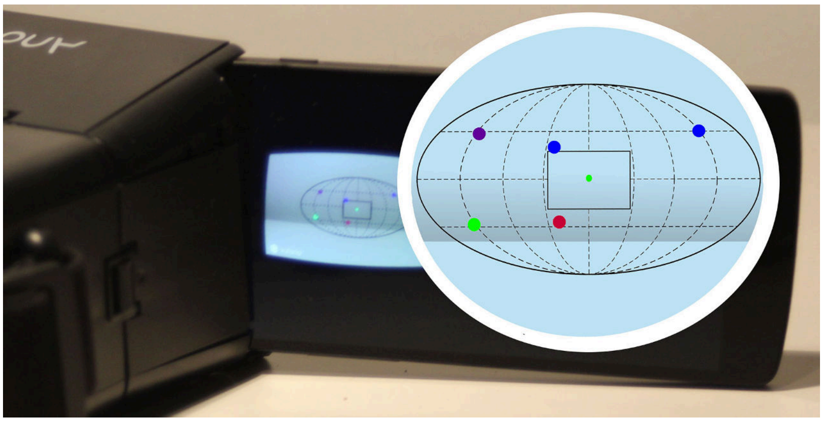


Figure 7.1: Google Cardboard with zoom-in on our developed EyeSee360 visualization technique.

In our previous chapter, we adapted Arrow, Halo, and Wedge to head-mounted Augmented Reality (see Chapter 5). The adapted techniques utilize 3D shapes as visual cues and our results showed that all these techniques are applicable for head-mounted devices. However, they only encode direction to out-of-view objects and the 3D shapes add visual clutter to the screen. Furthermore, they are not well-suited for small field-of-view devices (e.g., Hololens) or the simultaneous visualization of many objects. Therefore, to visualize multiple out-of-view objects at the same time without adding too much clutter, we propose a novel visualization technique called EyeSee360. The technique uses a radar-like visualization to display out-of-view objects. We compare it to the adaptations of well-known 2D techniques (Arrow, Halo, and Wedge) proposed in Section 5.1. Because these adaptations can only visualize direction towards out-of-view objects, we can only use them to evaluate the directional encoding of EyeSee360. We had to evaluate EyeSee360 against visualization techniques that only encode direction because, to our knowledge, there is no egocentric visualization technique that addresses visualization of both direction and distance for out-of-view objects at the same time. Therefore, we aim to design, implement, and evaluate a novel visualization technique for out-of-view objects in head-mounted Augmented Reality.

Here, our research contributions include:

- A lo-fi head-mounted prototyping tool that allows quick testing and refining of design ideas for Augmented Reality.
- The design, implementation, and evaluation of our novel out-of-view object visualization technique EyeSee360 and comparison to three adapted 2D off-screen visualization techniques (Arrow, Halo, and Wedge).

The work presented in this section was published as a full paper at the SUI symposium in 2017 and received an honorable mention for best paper [GEA+17].

7.1.1 Approach

To address the problem of out-of-view objects in 3D space, we divide the problem into two sub-problems: visualizing the direction of an out-of-view object and visualizing the distance to the object. This makes sense because many use-cases require only a visualization of the direction information (e.g. monitoring tasks). Therefore, we are able to evaluate the direction and distance aspects of our technique independent of each other. As a first step, we attempt to encode direction information. To that end, we drew on prior work on mobile off-screen visualization techniques (see Section 3.1).

For EyeSee360, we use the Contextual views approach, which overlays the screen border with contextual information. In other words, the information overlays are not centered on the screen, but are displayed peripherally along the outer part of the screen. Thereby, the user's direct (foveal) line of sight remains uncluttered [LL09]. Furthermore, following this approach, visualizations use representations off-screen objects that are placed in the same directions as their corresponding off-screen objects and are thereby in line with the human frame-of-reference [JHPR11].

In previous work, it was shown that moving objects can be tracked more accurately with EdgeRadar than with Halo [GI07]. This is relevant since EdgeRadar [GI07] served as inspiration for our technique. We evaluate EyeSee360 against our previously developed Contextual views techniques (Arrow, Halo, and Wedge) which we adapted for head-mounted Augmented Reality (see Section 5.1). The reason for testing Against Arrow, Halo, and Wedge is that we investigated all three research questions in parallel and to the time we started to explore EyeSee360, these techniques were the only ones we had developed so far.

7.1.2 Part I: EyeSee360 Concept Validation

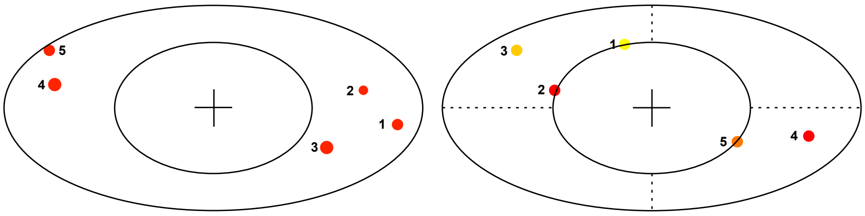
As a first step, we explore the design and development of our EyeSee360 technique for visualizing out-of-view objects distributed 360° around the user.

7.1.2.1 Concept Definition

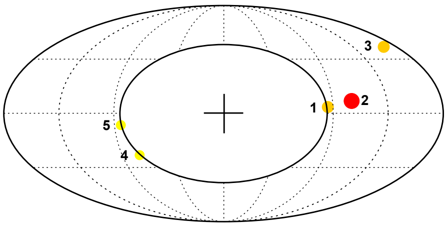
We created three different variants of EyeSee360 for the encoding of direction information to offer three different levels of visual assistance. In the peripheral field-of-view, an inner ellipse and an outer ellipse are drawn. The inner ellipse is the representation of the field-of-view and is sized so as not to occlude the user's focus. Objects within the inner ellipse are not considered to be out of view. The outer ellipse represents the 180° line (behind the user) from top (90°) to bottom (90°). We chose an ellipse because the binocular nature of human vision means that it is wider that it is high (see Subsection 2.1.2). Furthermore, we need to represent up to 180° behind the user and only up to 90° to the top.

Visual cues that represent out-of-view objects are displayed as dots between the inner and outer ellipses. The position of a dot represents the position of its corresponding out-of-view object. Left and right are shown up to 180°, while up and down are shown up to 90°. This makes it possible to visualize any position for an out-of-view object in 3D space. The three different variants that we explored for directional encoding are as follows: 1) only the inner and outer ellipses are visible to the user (no helplines; see Figure 7.2a), 2) the ellipses are shown along with two helplines on the horizontal and vertical axes (basic helplines; see Figure 7.2b), and 3) the ellipses are shown with helplines on both axes as well as on each 45° step (all helplines; see Figure 7.2c). We have three variants for distance encoding as well: 1) size, 2) color, and 3) a combination of size and color (see Figure 7.2). The fixation cross seen in Figure 7.2 is always centered in the user's field-of-view.

To represent distance to an object, the following visual cue characteristics can be considered: size, color, brightness, and animation. Brightness was excluded



(a) No helplines with size. (b) Basic helplines with color.



(c) All helplines with combination of size and color.

Figure 7.2: Three different variants of EyeSee360.

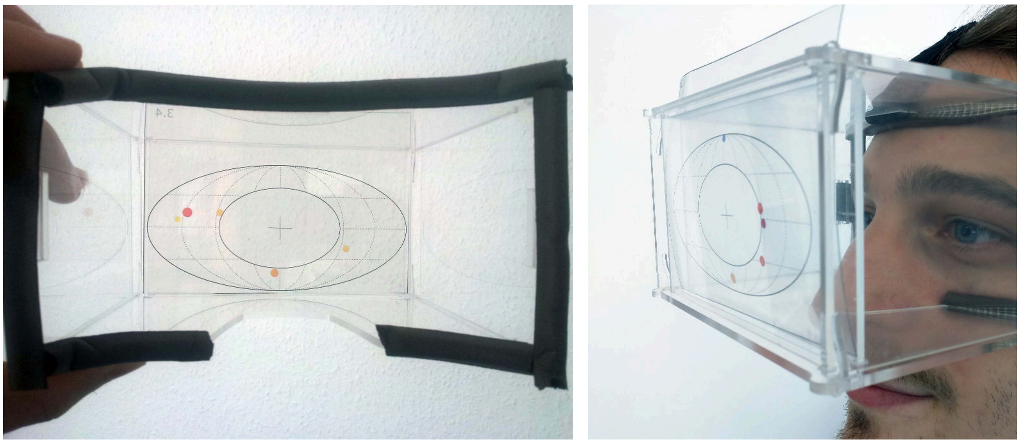
from the selection because it would cause cues for objects with low alpha values to appear faded. Also, since brightness is influenced by the environment, unintentional changes in perception could occur. We could also animate objects (e.g. make them blink) with a high frequency. However, animation in the periphery may immediately captures attention and should therefore only be used for guiding attention, not for continuous visualization of information. The two remaining cue characteristics, size and color, have been selected as possible candidates for representing distance. Colors are already important when it comes to attention (e.g., for traffic lights). On the other hand, we have a natural association between distance and size. The farther away something is, the smaller it looks. However, a combination of both forms is also possible. Here, it is important that the two colors are clearly distinguishable on every step of their color gradient. Therefore, we chose red and yellow for our experiment. Red is used for closer objects because we consider them to be more important, while yellow is used for objects farther away. Further, it is important that the different sizes of the distance encodings are clearly distinguishable. However, a limit of the size must be considered. Very large visual cues for the out-of-view objects lead to more frequent overlapping. Size, color, and a combination of both size and color for distance representation are therefore compared in the following concept validation study.

7.1.2.2 Concept Validation Study

To validate the concept of EyeSee360, we ran a concept validation study. Our goal was to evaluate whether users can determine the positions of out-of-view objects with EyeSee360. To test our concept, we developed a prototyping tool that allows researchers to rapidly test non-changing head-mounted see-through views without implementing them. With this tool, we could use simple slides to obtain users' impressions. Here, we needed to develop our own prototyping tool because previous tools, such as the PapAR tool by Lauber et al. [LBB14], do not consider use with head-mounted devices or focus on the interface elements [SC12] instead of visualizing information. Our prototyping tool does not limit the human field-of-view thanks to transparent materials.

7.1.2.3 Designing the Lo-fi Prototyping Tool

We laser cut our prototyping tool (see Figure 7.3). We used Plexiglass manufactured glasses that allow insertion of film slides (known as transparencies or viewfoil) in the front. The user is then able to explore static see-through layers and their effects on the perceived environment. It is even possible to have more than one layer. Since no lenses are used, care must be taken that human perception does not allow any content to be focused directly in front of the user. Accordingly, a suitable distance between the film and the human eye must be used. It should be noted that this distance must always be greater with increasing biological age. For example, for persons over 70 years old, if the distance reaches 70 cm, it would no longer be suitable since such glasses would be too bulky and heavy.



(a) Prototyping tool form a user's perspective.

(b) Side view.

Figure 7.3: Person wearing our lo-fi AR prototyping tool.

7.1.2.4 Study Design

To evaluate the performances of the three variants of directional and distance encoding described above, we conducted a comparative user study. The user study was designed as a lab study and took place in an empty office room with white walls and darkened windows to avoid effects of different light conditions. We lit the room with artificial light (around 600 lux). We used quantitative methods to objectively evaluate performance as well as SUS questionnaires to gain insights into the perceived usability of each variant. For this validation study, we fixed several parameters: 1) all out-of-view objects were homogeneously distributed in 3D space to fit all directions and distances equally, 2) the positions of the objects and the user's viewing angles to them did not change throughout the study, and 3) the number of displayed objects was fixed at five. We investigated user performance (accuracy) and subjective variable perception for object direction and distance.

In this study, we have two independent variables: assistance with three levels (no helplines vs. basic helplines vs. all helplines) and attribute with three levels (size vs. color vs. combination). Since we could not separate direction information from distance information, it was not necessary to test out all 9 possible combinations. Instead, the three versions of direction and distance encoding were combined. This resulted in three overall visualizations: no helplines with size, basic helplines with color, and all helplines with combination. (see Figure 7.2).

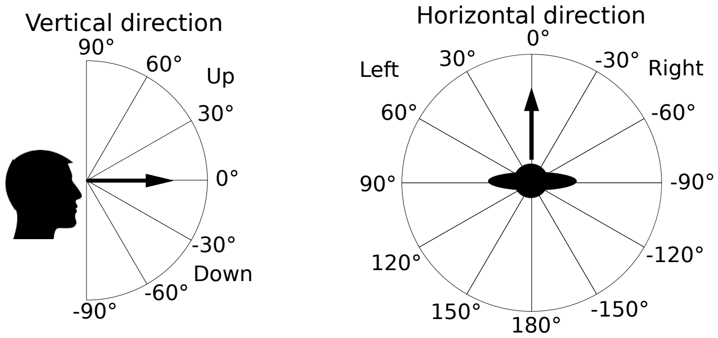
We used a questionnaire to measure variable perception in the periphery. We asked participants how many object representations they perceived per run if they focused on a cross in the center of the slides. This was our perceptibility dependent variable. Performance (for direction and distance) was measured through paper-based responses, where participants had to indicate on sheets of paper with diagrams. First, we asked for vertical direction towards the object, second for horizontal direction towards the object, and third for distance to the object. Vertical and horizontal directions were measured by binning responses into 30° range buckets (see Figure 7.4a and 7.4b). For example, a participant could say the vertical direction is between 30° and 60° and the horizontal direction is between 150° and 180°. Distance was measured in four buckets: very near, near, far, very far (see Figure 7.4c). Before the test session, each participant was given a tutorial to become familiar with these distance buckets.

For this study, we derived the following sub-question from our third research question: *(RQ3a) Which EyeSee360 visualization concept performs best with respect to accuracy for direction and distance towards out-of-view objects and can be perceived well in the periphery?* Our dependent variables were: perceptibility (counting the number of visual cues), horizontal direction, vertical direction, and perceived distance. We posit the following hypotheses:

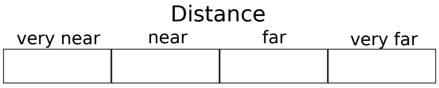
H_{16} We hypothesize that the direction encoding with all helplines results in better user performance than the one with no helplines.

H_{17} We think using the combination of size and color for distance encoding results in better user performance than using either of them individually.

To avoid learning effects or fatigue, we counterbalanced our independent variables. Using a Latin square design, we arrived at 3 rows for the study. For each



(a) Vertical direction form. (b) Horizontal direction form.



(c) Distance form.

Figure 7.4: Paper-based forms for collecting user input.

visualization combination, 10 slides were prepared. This equals a total number of 30 slides. Since our rapid prototyping approach requires changing of the slides, we limited each condition to only 4 out of the 10 possible slides (which are variations of the condition). The 4 slides used were chosen at random for each condition.

7.1.2.5 Procedure

At the start of the study, participants received an introduction to out-of-view objects and were given a demo where they could look at the different variants of EyeSee360. All possible variants of the visualizations were explained in detail using the three slides in Figure 7.2. As already mentioned, the order of the visualizations was counterbalanced. The three different visualizations were then applied successively to the AR prototype during the main part of the study. First, participants were asked how many visual cues they could see while focusing on the cross in the center of the slide. The maximum number of visual cues that can be seen on each slide is five. Next, the experimenter, went through each numbered visual cue with the participant, asking for the three necessary values: vertical direction, horizontal direction, and distance. As mentioned earlier, the used forms can be seen in Figure 7.4. After all visualizations had been completed, participants filled out two questionnaires: one was a SUS questionnaire and the other consisted of four questions concerning the implementation of the visualizations. At the end of the study, participants filled out a personal information form (which included typical items such as age and gender as well as experience with out-of-view object visualizations rated on a 5-point Likert-scale, where 1 is strongly disagree). Each session lasted approximately 60-75 minutes.

7.1.2.6 Participants

We recruited 19 participants[1] (8 female), aged between 20 and 61 years (M=28.3, SD=11). None suffered from color vision impairment, and all had normal or corrected-to-normal vision. Eleven had no experience with out-of-view object visualizations and eight were somewhat familiar with such visualizations (Md=1, IQR=1-2).

7.1.2.7 Results

Perceptibility

We consider the effects of attributes used (size, color, or combination of both) on object perceptibility. On each slide, five out-of-view objects were shown to the participant. To measure the perception performance, the user was asked to focus on a cross in the center of the slide and state the number of visible representations for out-of-view objects. This helped us to figure out whether a representation was too small or a color could not be perceived. The mean number of perceived objects (max=5) was 4.91 for no helplines with size, 4.63 for basic helplines with color, and 4.63 for all helplines with the combination of size and color. Normality here was not assumed because the Shapiro-Wilk test was significant ($p<0.001$). We therefore ran a Friedman test, which revealed a significant effect of different encodings on perception error ($\chi^2(2)=21.24$, $p < 0.001$, N=19). A post-hoc test using Wilcoxon Signed-rank with Bonferroni correction showed significant differences between some conditions, as shown in Table 7.1.

Table 7.1: Pairwise comparison of perceptibility conditions.

Comparison	**p-value**	**ϕ-value**
No helplines with size vs. basic helplines with color	<0.001	0.33
No helplines with size vs. all helplines with both	<0.001	0.36
Basic helplines with color vs. all helplines with both	0.85	0.03

Direction Accuracy

We consider the effects of three different variants of our EyeSee360 technique (no helplines with size, basic helplines with color, and all helplines with combination) on vertical and horizontal direction accuracy. It is important to know that we divided the different degree values into buckets, where each bucket represents

---

[1] To obtain sufficient power for our data, we need 18 participants in our study. We calculated this value with G*Power under two-way ANOVA (α=0.05 and $1-\beta$=0.8) based on the three different variants of visualization techniques. For the perception in the periphery we get 216 data points because we get this based on the slides and not on every object (f=0.22). For direction and distance towards out-of-view object we get 1080 data points and we are able to show small effect sizes of (f=0.1).

a 30° angle. These buckets were used to simplify participant input entry. For example, if the out-of-view object's position is between 0° and 30° horizontally and the user thinks it is located between 60° and 90°, then the error will be 2 because the guessed class is two points away from the correct one.

We had 6 buckets for vertical range (90° up, 90° down) and 12 buckets for horizontal range (360° surround). The mean errors for vertical direction are: no helplines with size=0.19, basic helplines with color=0.15, and all helplines with combination=0.12 (max. error possible is 5). The mean errors for horizontal direction are: no helplines with size=0.63, basic helplines with color=0.51, and all helplines with combination=0.28 (max. error possible is 6). For vertical direction error, a Shapiro-Wilk-Test showed that our data is not normally distributed ($p<0.001$), and thereafter we ran a Friedman test that revealed no significant effect of different encodings on vertical direction error ($\chi^2(2)=2.98$, $p=0.225$, $N=19$). For horizontal direction error, a Shapiro-Wilk-test showed that our data is not normally distributed ($p<0.001$), and thereafter we ran a Friedman test that revealed a significant effect of different encodings on horizontal direction error ($\chi^2(2)=48.27$, $p<0.001$, $N=19$). A post-hoc test using Wilcoxon Signed-rank with Bonferroni correction showed significant differences for some conditions (see Table 7.2).

Table 7.2: Pairwise comparison of help variations for horizontal direction.

Comparison	**P-value**	**ϕ-value**
No helplines with size vs. basic helplines with color	0.076	0.06
No helplines with size vs. all helplines with both	< 0.001	0.24
Basic helplines with color vs. all helplines with both	< 0.001	0.19

Here, the pairwise comparisons show that the lower mean error for horizontal direction for all helplines with combination is significant compared to no helplines with size and basic helplines with color. Figure 7.5 shows the horizontal and vertical direction errors for all three variants combined. Clearly, most errors fall under error class 1, which means the erroneously chosen buckets were usually right next to the correct buckets.

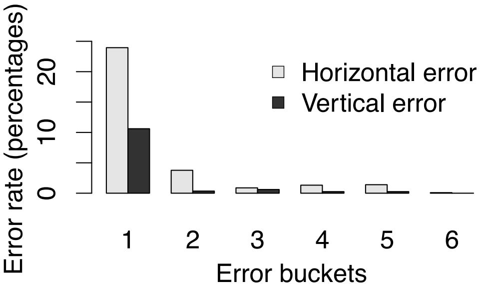


Figure 7.5: Barplot showing rates of horizontal and vertical direction errors.

Distance Accuracy

We investigated the effects of three different attributes (no helplines with size, basic helplines with color, all helplines with combination) on distance accuracy. We again created a bucket-based encoding for distance: very near, near, far, very far. The mean errors for distance are: 0.15 for no helplines with size, 0.07 for basic helplines with color, and 0.09 for all helplines with combination (max. error possible is 3). A Shapiro-Wilk-test showed that the data are not normally distributed ($p<0.001$) and a Friedman test revealed a significant effect of the encoding types on distance ($\chi^2(2)=20.34$, $p<0.001$, $N=19$). A post-hoc test using Wilcoxon Signed-rank with Bonferroni correction showed significant differences for some conditions (see Table 7.3).

Table 7.3: Pairwise comparisons of help variations for distance.

Comparison	**p-value**	**ϕ-value**
No helplines with size vs. basic helplines with color	< 0.001	0.14
No helplines with size vs. all helplines with both	< 0.001	0.12
Basic helplines with color vs. all helplines with both	0.618	0.02

System Usability Scale

For this concept validation study, our rapid prototype of EyeSee360 scored 71 on the SUS, which is just over the threshold of 68 for acceptable usability [Bro96].

User Feedback

To gain more specific feedback, we gave participants four 5-point Likert-scale questions (1=strongly disagree, 5=strongly agree). Participants stated that they could easily distinguish between the colors used for distance encoding (Md=4, IQR=4-5) and could easily distinguish between the sizes used for distance encoding (Md=4, IQR=3-4). In addition they found the inner ellipse of the visualization confusing (Md=2, IQR=1-2), but thought the helplines in the visualization were helpful for understanding the cued direction (Md=5, IQR=4-5). Participants reported that the color yellow is not easily perceived, especially when it is shown on the periphery. This is in line with our findings regarding the perceptibility of representations in our early prototype that use the color yellow. Another point reported by users was that two very close out-of-view objects cannot be encoded by our visualization because one can occlude the other.

7.1.2.8 Discussion

While validating our lo-fi EyeSee360 prototype, we discovered a number of interesting findings. First, using all helplines was deemed most useful by participants and resulted in the lowest error for direction estimation, which fed into later design iterations. Therefore, we can accept our hypothesis H_{16}. Second, color as well as the combination of color and size were overall the best attributes for distance encoding, as shown by our quantitative results. However, since there was no significant difference between color and color with size, we cannot accept our hypothesis H_{17}. Finally, the SUS scores indicated acceptable perceived usability for this lo-fi prototype, which provided further support for implementing our final EyeSee360 prototype.

7.1.3 Part II: EyeSee360 Evaluation Study

Given the promising results of the concept validation study, we implemented the best-performing version of our novel EyeSee360 visualization technique. This was done for two reasons. First, we are now able to evaluate this technique in an actual implemented prototype, which increases ecological validity. Second, we can now compare our new visualization technique with the adapted versions of three well-known 2D off-screen visualization techniques (Arrow, Halo and Wedge) for encoding direction towards out-of-view objects (see Section 5.1). We chose to compare EyeSee360 to Arrow, Halo and Wedge from our first study because, at the time of this study, these techniques were the only ones that we had developed and tested so far. However, since most of the techniques in this thesis have been tested with comparable search and direction estimation tasks, we still can compare the results of the different techniques to each other. In the second part of this study, we evaluated the distance encoding of EyeSee360. We further investigate the subjective workloads incurred by all tested visualization techniques and assess the usability of each with the SUS questionnaire. To evaluate the scalability of EyeSee360, we tested it with different numbers of objects (three vs. five vs. eleven).

7.1.3.1 Implementation

All visualizations were implemented for Google Cardboard. Thereby, we tested video see-through Augmented Reality. Vuforia was additionally used for the implementation to keep the out-of-view objects at fixed positions in the environment. We used the Vuforia environment tracking based on the gyroscope sensor of the smartphone. The Google cardboard we used for the evaluation had a field-of-view of 45° horizontally and 30° vertically. Our development was done in Unity. Our implementation of EyeSee360 supports various devices (e.g., Google Cardboard, Microsoft Hololens, Oculus Rift, etc.) and is published as open source code on Github[2].

Our validation study illustrated that a visualization with 45° helplines is best suited for showing direction to out-of-view objects (variant 3: all helplines with

---

[2] Github OutOfView-Project. `www.github.com/UweGruenefeld/OutOfView`, last retrieved April 21, 2020

combination; see Figure 7.2c). The representations of distance using colors or color and size were the best performing variants. We chose to represent distance with color only to avoid having too many conditions and because this resulted in the smallest mean error. However, we found that showing yellow in the periphery was not very perceptible. For this reason, a color gradient from blue to red was chosen based on the cold and warm metaphor used, for example, in heatmaps[3] [Har00]. Here, red stands for very close and blue for far away.

In addition, the EyeSee360 visualization must now represent objects dynamically through their visual cues. This means that the image has to change depending on the user's viewing direction. Furthermore, the previous visualization must be adapted to the smartphone. Here, the field-of-view was represented by an inner ellipse. Due to the video see-through variant, however, the camera image is looped through the device and output on the screen. Since this output no longer has a round shape, but a rectangular shape, the inner area of the visualization must be adapted to also become rectangular. In other words, the looped camera image is the new field-of-view and, since this image is rectangular, the focus area needs to be rectangular too. A further adaptation is that this inner area has to move when the user looks up or down; that is, it must move in the correct vertical direction as the user adjusts the field-of-view. In these cases, the inner visualization moves along.

7.1.3.2 Study Design

The second study was also designed as a lab study and took place in the same empty office room with the same light conditions. The study is split into two parts, where the first aim is to compare the direction encoding of EyeSee360 with Arrow, Halo, and Wedge (part 1), and the second is to evaluate the distance encoding for EyeSee360 (part 2). Furthermore, we measure the SUS scores and the subjective workload for the different visualization techniques with the NASA Raw-TLX [Har06]. For this study, we had to fix three parameters: 1) all out-of-view objects were randomly distributed in 3D space with equal possibilities for every possible direction and distance (except the 3D space within the user's view), 2) distance was ranged between 0% and 100% with a one meter distance at 100%, and 3) the positions of the out-of-view objects were world-fixed, but not ego-fixed (or user-fixed). In other words, these objects are non-moving and stay in the same positions over time.

In the first part of the study, we investigate whether the dependent variable direction accuracy is influenced by the independent variables visualization (Arrow vs. Halo vs. Wedge vs. EyeSee360), environment (180° vs. 360°), and number of objects (three vs. five vs. eleven). This repeated-measures within-subjects factorial design results in 24 different conditions. Additionally, we investigated whether the dependent variable's usability and workload are influenced by the

---

[3] Heatmap. `en.wikipedia.org/wiki/Heat_map`, last retrieved April 21, 2020

independent variable visualization (Arrow vs. Halo vs. Wedge vs. EyeSee360). In the second part of the study, we investigate if the dependent variables distance accuracy and search time are influenced by the independent variables visualization (EyeSee360), environment (360°) and number of objects (three vs. five vs. eleven). This repeated-measures within-subjects factorial design results in 3 different conditions.

For this study, we derived the following sub-question from our third research question: *(RQ3b) Which visualization (Arrow, Halo, Wedge, EyeSee360) for head-mounted AR results in the best user performance with respect to direction, usability, and workload?* We posit the following hypotheses:

H_{18} EyeSee360 performs better than the Arrow, Halo, and Wedge visualizations for estimating direction for all environments (180°, 360°).

H_{19} The measured workload for EyeSee360 is higher than for the visualizations Arrow, Halo, and Wedge.

7.1.3.3 Procedure

At the beginning of the study, participants got an introduction to out-of-view objects and were given a demo where they could test the different visualization techniques. Then participants did part 1 (see Section 7.1.3.3) and part 2 (see Section 7.1.3.3) of the study. At the end of the study, participants filled out a personal information form that included typical items such as age and gender but also asked them to rate their experience with out-of-view object visualization techniques and head-mounted devices on a 5-point Likert-scale, where 1 is strongly disagree and 5 is strongly agree). Each session lasted approximately 45-60 minutes.

Part 1: Direction Estimation

In the first part, participants had to locate out-of-view objects in two different environments (180° vs. 360°). These ranges are 180° ahead of the user or 360° around the user. The orders of the visualizations, numbers of objects, and environments were generated according to a Latin square design and were then randomized. The out-of-view objects were created randomly in the corresponding area of the environment (180° vs. 360°). As mentioned earlier, none of the objects occupied the user's viewport. Each combination of visualization, number of objects, and environment was tested in 3 iterations. In each iteration, three representations were chosen at random and were successively highlighted in green. When a representation was highlighted, the participant had to guess the position of the corresponding object without being able to see it.

To accomplish this, the participant had a green cursor in the center of the screen and a remote controller to confirm the current cursor position as the out-of-view object's position. The cursor could be moved via head movement. To stop participants from being able to obtain the exact position of an out-of-view object in a given technique through head movement, the visualization technique was only visible in a small area directly in front of the participant. Moving the green cursor out of a black circle disabled the visualization technique, at which point the participant had to guess the out-of-view object's position by the affordances the technique offered (see Figure 7.6). The direction error is measured as the angle between the position specified by the participant and the actual position of the object. We had to change the input method from verbal reporting in the concept evaluation study to digital input to enable participants to specify the positions of out-of-view objects with higher accuracy. Further, in pilot trials, participants stated that writing on a paper while looking through a video-see-through device feels uncomfortable.

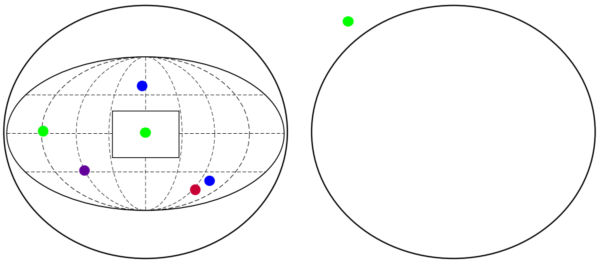


Figure 7.6: User input during the direction estimation task. Left: green cursor is inside the black circle and EyeSee360 is visible. Right: green cursor is outside the black circle and EyeSee360 is not visible.

After each visualization, the participants had to fill out SUS [Bro96] and NASA-TLX [HS88] questionnaires about the used visualization technique.

Part 2: Distance Estimation and Search task

In the second part, we evaluated the distance encoding for EyeSee360. Additionally, participants were asked to search for the out-of-view object. The different numbers of objects were presented in a random order. Out-of-view objects were created randomly in 360° around the user. Again, none of the objects occupied the user's viewport. Each number of objects was tested in two iterations. In each iteration, three representations were chosen randomly and the participant had to complete the following two tasks for each chosen representation. In the first task, the chosen representation is highlighted and the user must estimate the distance towards the represented out-of-view object by moving a three-dimensional object on an axis extending from the user into the viewing space. The object is moved with a remote control. Then, as a secondary task, the highlighting disappears and the user must locate the represented object. This is accomplished by using the green cursor in the center of the screen to select one of the visible out-of-view objects and confirm it with the remote control.

7.1.3.4 Participants

We recruited 16 participants[4] (8 female), aged between 20 and 63 years (M=30.6, SD=12.7) Most participants did not have much experience with visualizations of out-of-view objects (e.g., from video games) (Md=1, IQR=1-2), nor with head-mounted devices (e.g., AR or VR) (Md=1, IQR=1-1).

7.1.3.5 Results

Direction Accuracy

We consider the effects of the three factors (visualization, number of objects, environment) on mean direction error. The mean errors for the visualization techniques are: Arrow=31.23°, Halo=29.74°, Wedge=28.52°, and EyeSee360=21.25°. The direction errors are compared in Figure 7.7.

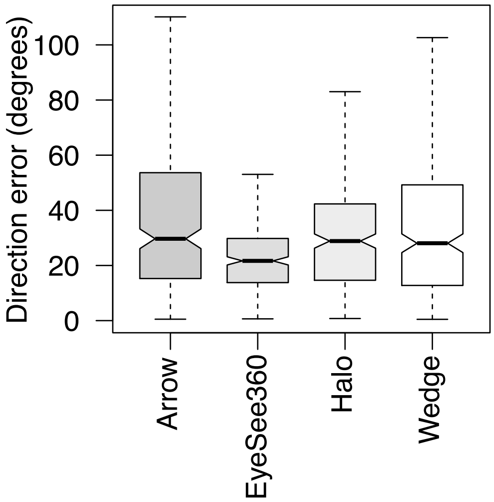


Figure 7.7: Boxplot of median direction errors for visualizations.

A Shapiro-Wilk-Test showed that our data is not normally distributed ($p < 0.001$), and thereafter we ran a Friedman test that revealed a significant effect of visualization technique on direction error ($\chi^2(3)=27.55$, $p<0.001$, N=16). A post-hoc test using Wilcoxon Signed-rank with Bonferroni correction showed significant differences between the four groups (see Table 7.4). EyeSee360 has a significantly lower direction error than Arrow or Halo. Furthermore, in Figure 7.8, it is clear that the deviation from the median is smallest for EyeSee360.

---

[4] The number of participants was calculated with G*Power for ANOVA. For direction we have 24 combinations. We assumed an alpha of 0.05 and a power of 0.8 (α=0.05, 1-β=0.80). We need at least 14 participants to measure mean effect sizes (f=0.25). For distance we again assumed an alpha of 0.05 and a power of 0.8 (α=0.05, 1-β=0.80). With 14 participants we can also measure mean effect sizes of (f=0.25).

Table 7.4: Pairwise comparisons of visualization techniques.

Comparison	**p-value**	**ϕ-value**
Halo vs. Arrow	0.723	0.01
Wedge vs. Arrow	0.053	0.06
EyeSee360 vs. Arrow	<0.001	0.20
Wedge vs. Halo	0.280	0.03
EyeSee360 vs. Halo	<0.001	0.20
EyeSee360 vs. Wedge	<0.001	0.15

Environment

The mean direction errors for environment are 24.06% for 180° and 31.31% for 360°. As we compare two matched groups within subjects, we directly performed a Wilcoxon Signed-rank test. Here we found a significant effect of environment (W=241170, Z=-8.05, $p<0.001$, ϕ=0.17). We further investigated effects of environment on each technique. We did a Wilcoxon Signed-rank test for Arrow (W=12644, Z=-5.77, $p<0.001$, ϕ=0.17), Halo (W=17866, Z=-2.08, $p<0.05$, ϕ=0.12), Wedge (W=14258, Z=-4.63, $p<0.001$, ϕ=0.14), and EyeSee360 (W=15216, Z=-3.95, $p<0.001$, ϕ=0.12). These were all shown to be significant.

Number of Objects

Further, we investigated if the number of objects has a significant effect on direction error. We compare three matched groups within subjects with a non-parametric Friedman test, which revealed no significant effect of number of objects on direction error ($\chi^2(2)=3.2604$, p=0.20, N=16).

Subjective Workload

According to hypothesis H_{19}, we expected that EyeSee360 would have a higher subjective workload than Arrow, Halo, or Wedge. Here, we compared the four matched groups (workload scores: Arrow=47.6, Halo=42.5, Wedge=44, EyeSee360=46.6) within subjects using a non-parametric Friedman test. The Friedman test revealed no significant effect of visualization on Nasa Raw-TLX score ($\chi^2(3)=0.43$, p=0.93, N=16). This means we can not accept or reject the null hypothesis of H_{19} that there is no difference in overall workload across techniques.

Distance Accuracy

We compare three matched groups (3, 5, or 11 objects) within subjects with a non-parametric Friedman test, which revealed no significant effect of number of objects on distance error in EyeSee360 ($\chi^2(2)=1.8788$, p=0.39, N=16). The distance was measured in a range between 0% and 100%.

Search Time

Additionally, we investigated search time for EyeSee360. In 297 runs, participants found 251 out-of-view objects correctly (success rate of 84.5%). Furthermore, the time needed to locate the out-of-view objects is important, where the mean time to find an object was 10.4 seconds. The mean search time for all correctly located objects was 9.8 seconds. Both values are quite high due to the time required for head rotation, selecting the out-of-view object, and confirming the input. As a next step, we divided the rotation necessary to find the object into buckets (containing 30° each). In Figure 7.8, it is shown that the time needed to find an out-of-view object increases for higher angles, except for 91° to 120°.

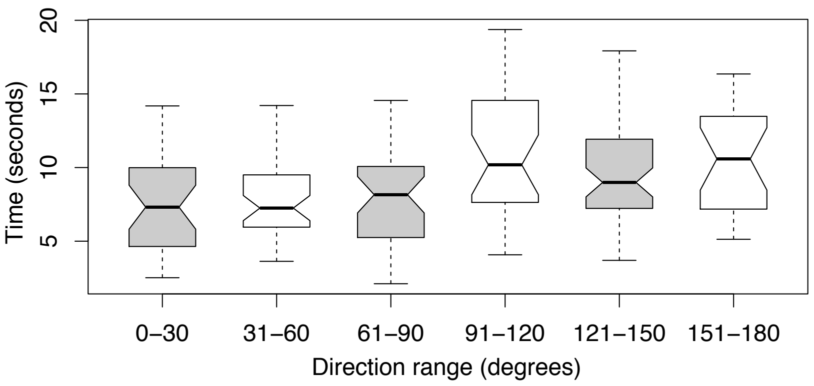


Figure 7.8: Median time for horizontal direction buckets.

System Usability Scale

To evaluate the perceived usability of EyeSee360, the SUS questionnaire was used. The result for the final prototype of EyeSee360 (68) is above Halo (66) and under Wedge (70). This is also under the SUS score for the rapid prototype version (71). This is likely due to the fact that we used video see-through AR for the final prototype. Latency with looping the camera to the screen can cause simulator sickness and decreased usability.

Learning effects

The number of trials and the time needed by participants are insufficient for investigating learning effects within our data. Therefore, to gain more specific feedback, we gave participants two 5-point Likert-scale questions (1=strongly disagree, 5=strongly agree) to investigate this. Participants felt they were able to understand the positions of out-of-view objects visualized with EyeSee360 more quickly over time (Md=4, IQR=3-4), as well as more precisely (Md=4, IQR=4-4).

7.1.3.6 Discussion

In our final evaluation study of EyeSee360, we have a number of noteworthy findings. First, it appears that EyeSee360 had the lowest error for direction estimation, in contrast to the adapted 2D techniques. A pairwise comparison revealed that EyeSee performs significantly better than Arrow and Halo, but not significantly better than Wedge. Therefore, we cannot accept or decline hypothesis H_{18}. Second, the 180° environment setting resulted in better direction accuracy than the 360° condition, which fits our expectation that estimating direction for a larger spatial area is more difficult than for a narrower range. The number of objects did not have any significant influence.

For workload, we also did not find any differences between the tested techniques. Therefore, we cannot accept or reject hypothesis H_{19}. However, we do observe a positive correlation between search task time and direction angle. This can have detrimental effects on workload in a real-life scenario. With respect to distance estimation, we found that there were no differences across the different numbers of objects; however, the distance error was in the 5-10% range (M=9.34, SD=10.26) for all tested conditions. Finally, it seems participants feel they get better at the task over time, despite the fact that this was not explicitly tested.

7.1.4 Implications

Advantages of AR

Basically, EyeSee360 was inspired by 2D off-screen visualization techniques such as EdgeRadar [GI07] and is therefore somewhat similar to these techniques and familiar. For that reason, EyeSee360 uses a plane projection in the view frustum of the user. Combined with a head-mounted device, our technique offers a constant flow of information regarding out-of-view objects in the periphery.

From Video to Optical See-Through AR or VR

Another issue that arises in our work is whether our findings can transfer to other head-mounted AR devices (e.g., optical see-through). To use our technique with an optical see-through device, it may be necessary to measure the user's facial field, if the technique is intended to be used to locate physical objects in the environment. By contrast, for devices with video-see-through technology (as we have done), the facial field is determined by the camera lens used and is therefore easier to determine. While in this study we were concerned with video see-through AR technology, we can see the potential of extending this work to other AR and especially Virtual Reality (VR) environments (since everything is rendered digitally in such an environment).

Usefulness of Lo-fi Prototyping Tool

From our first study, we developed a low-fidelity prototyping tool to quickly test design ideas for out-of-view objects for head-mounted AR. We argue for the usefulness of this approach, as it saves both development time and allows designers to test and iterate quickly based on user feedback. This is especially applicable for a visual domain such as out-of-view objects, where the exact parameters for size, color and form can vary even between users (e.g., color blind individuals).

Ecological Validity

It is important to reflect on whether our developed techniques can be used in a real-world scenarios, such as ship docking [OLL15] or gaming environments for social awareness [PKB10]. While this was out of our current scope, the ultimate test of how well our EyeSee360 system can support users would require longitudinal in-the-wild testing, wherein we can gain greater insight into the interactions between learning effects, errors, and the specific contexts in which a head-mounted AR device is used (e.g., while mobile).

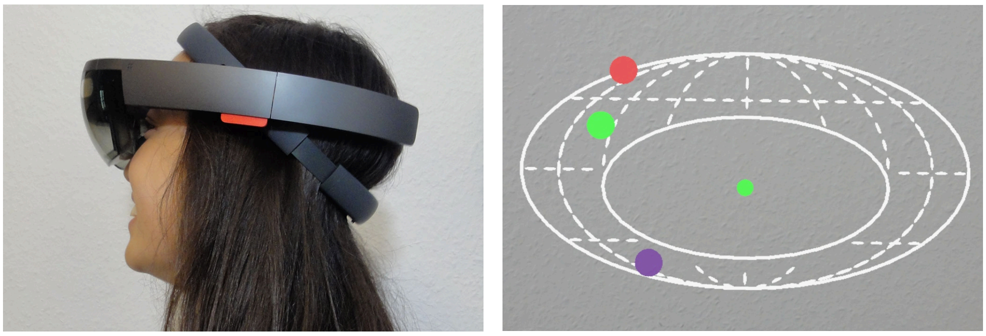
Study Limitations

We did not measure task completion times for direction accuracy in the second study comparing Arrow, Halo, Wedge, and EyeSee360. It was not deemed relevant for answering our initial research question and we wanted to focus on whether users are able to identify the object, not whether they can do so quickly. This can be addressed by future work when the time it takes to locate out-of-view objects is critical. Further, one could argue that the user having to rotate their head to identify the direction negatively affects direction estimation. However, alternative modes, such as having the user click on a 3D sphere or point in a direction, have similar limitations.

7.1.5 Conclusion

In this section, we studied the problem of visualizing out-of-view objects in head-mounted AR. Therefore, through a concept validation study and a comparative evaluation study, we proposed our EyeSee360 technique. This was shown to outperform the 2D adapted techniques from previous work. Further work should focus on reducing direction and distance errors. Additionally, further comparative studies are needed to evaluate EyeSee360 against other techniques such as AroundPlot[JHPR11]. Together, our methods and findings provide the groundwork on which future research on out-of-view object visualization in head-mounted AR can build, particularly for optical see-through (e.g., Hololens) or Virtual Reality devices.

7.2 EyeSee360 for Optical See-Through Augmented Reality



(a) Person wearing Hololens.

(b) EyeSee within Hololens.

Figure 7.9: Implementation of EyeSee360 on the Hololens which shows the human field-of-view and three real-world out-of-view objects.

Knowing the positions of out-of-view objects is a problem in many different environments (e.g., ship docking [OLL15] or gaming [PKB10]). A solution for this problem is our visualization technique EyeSee360, proposed in the previous section (see Section 7.1). When possible, EyeSee360 is shown in the user's periphery, thereby leaving the user's focus uncluttered. We evaluated EyeSee360 for a video-see-through device. However, video-see-through decreases the natural human field-of-view and is known for causing simulator sickness. Optical-see-through devices, on the other hand, leave the human field-of-view unchanged and do not cause sickness. However, with the limited fields-of-view of current AR devices (e.g., Microsoft Hololens[5]), virtual content still recedes from the user's view. This is problematic because it turns locating virtual objects into an exhausting and frustrating task.

To address the problem of virtual objects receding from view, different visualization techniques have been developed [BTOX06, JHPR11, SH13]. One of those techniques is EyeSee360 from the previous section (see Section 7.1). This technique allows one to visualize multiple out-of-view objects at the same time (it has been tested with up to eleven objects without showing any performance decrease in participants). Furthermore, EyeSee360 has been compared to most of the state-of-the-art techniques and was found to result in the best search time performance, lowest workload, and highest usability [BSEN18]. However, especially on small field-of-view devices, EyeSee360 as it is currently used introduces a lot of visual clutter to the screen in the form of additional information (assistance). We hypothesize that this additional information interferes with perception of other content and may lead to less optimal search time performance. Further, in the original work, three different variants of EyeSee360 (see Section 7.1) were pro-

---

[5] The Microsoft Hololens version 1 has a field-of-view of 30° horizontal and 17.5° vertical. `en.wikipedia.org/wiki/Microsoft_HoloLens`, last retrieved April 21, 2020

posed, each offering different levels of assistance information, but they were never compared to one other in terms of search time performance.

In this work, we aim to improve the search time performance for locating virtual out-of-view objects in Augmented Reality. We compare three variants of EyeSee360 with different levels of visual information (assistance) in a laboratory user study. Thereby, we want to investigate the influence of more information (assistance) but also more visual clutter on search time performance for state-of-the-art AR hardware (Hololens).

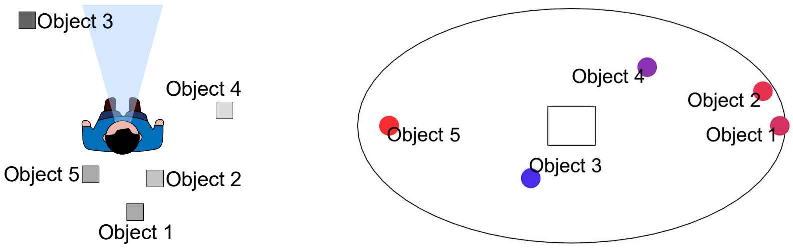
Here, our research contributions include:

- We add to a better understanding of the effects of visual clutter and assistance on search time for locating objects out of view.
- A comparison of three different variants of EyeSee360 with regard to best search time performance.

The work presented in this section was published as a short paper at the MuC conference in 2019 and received an honorable mention for best short paper [GPH19a].

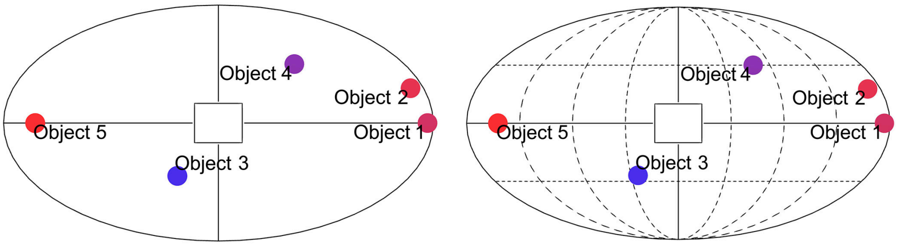
7.2.1 EyeSee360

In this work, we aim to improve search time performance of the out-of-view visualization technique EyeSee360 (see Section 7.1). We chose EyeSee360 because it supports multiple out-of-view objects and encodes their directions and distances relative to the user in head-mounted Augmented Reality devices. Furthermore, compared to other techniques, EyeSee360 shows the lowest direction estimation error and the lowest search times for objects distributed 360° around the user (cf. Section 5.3, Section 7.1 and [BSEN18]). EyeSee360 concentrates information about out-of-view objects onto a grid system in the user's periphery (see Figure 7.10). This grid system compresses 3D position information onto a single 2D plane. The inner rectangle of EyeSee360 represents the FOV of the user, and the area outside the rectangle represents the area outside of the user's view. However, EyeSee360 supports three different variants. The different variants of EyeSee360 offer different levels of assistance for locating objects out of view. The first variant offers no assistance, the second variant shows vertical and horizontal axis lines, and the third variant shows additional dotted lines representing 45° sections of the user's view (see Figure 7.10). The three variants of EyeSee360 have not been evaluated against one other in terms of search time performance. However, we hypothesize that the variant with all helplines offers too much unnecessary information and adds visual clutter to the screen, especially on devices with smaller fields-of-view. We think this results in higher search times and therefore, we will compare the three variants in a user study.



(a) Example scene.

(b) No helplines.



(c) Basic helplines.

(d) All helplines.

Figure 7.10: Out-of-view visualization techniques for virtual objects in example scene. (a) illustration explaining relative positions (bird's eye view); (b-d) user views through Hololens without built-in RGB camera image for better readability (each visualization fits the device screen 1:1).

7.2.2 User Study

In this study, we compare the different variants of EyeSee360 using the Hololens.

7.2.2.1 Study Design

To evaluate the performances of different variants of EyeSee360 for visualization of out-of-view objects, we conducted a within-subjects controlled laboratory study in Augmented Reality with the Microsoft HoloLens. During the study, participants were asked to search for objects out of view. Our independent variable was assistance with three levels (no helplines vs. basic helplines vs. all helplines). We used quantitative methods to evaluate user performance, taking search time and subjective Likert-items as our dependent variables.

For this study, we derived the following sub-question from our third research question: *(RQ3c) In how far does the visual clutter added by the different variants of EyeSee360 influence search time performance?* We posit the following hypothesis:

H_{20} We expect the EyeSee360 variants with no helplines and basic helplines to result in faster search time performance because they add less visual clutter and are therefore less distracting.

7.2.2.2 Apparatus

We set up an empty office space (3 x 3 meter) with darkened windows and an artificial light source to control the brightness throughout the experiment (around 500 lux). Our experiment and all variants of EyeSee360 are implemented in Unity3D[6], a 3D game development platform, and the Microsoft Hololens, a head-mounted Augmented Reality device.

7.2.2.3 Implementation

The inner ellipse of EyeSee360 represents the user's field-of-view. Since the Hololens is an optical see-through-device, the field-of-view is nearly the same as the human field-of-view. Our adaption of the inner ellipse can be seen in Figure 7.9b. Further, we had to change the color of the ellipses and helplines from black to white because black is not very visible on optical-see-through devices. One of the constraints of the Hololens is the limited display, which has a field-of-view that extends only 30° horizontally and 17.5° vertically. Therefore, EyeSee360 cannot be shown in the user's periphery.

7.2.2.4 Procedure

At the start of the study, participants received an introduction to the Hololens. After, we started testing the different variants of EyeSee360. Each of the three variants of EyeSee360 was tested in one block. All blocks were counter-balanced, using a balanced Latin square design. For each block, we had two test trials and ten measured trials. In each trial, five virtual objects were randomly placed 360 degrees around the user but not in view (see Figure 7.10a for an example). Each virtual object was assigned a label, starting with "Object" plus a random number from one to five. On-screen text informed the participant which of the five objects to seek. The participant had to search for the virtual object by moving their head into the direction of the object. When the virtual object appeared in view, the trial was successfully finished. We stored the ten randomly generated positions of the first block and used them in different orders for the other two conditions to ensure the search times would be comparable. After each block, participants were asked to answer two questions regarding performance and distraction. At the end, participants were asked for their favorite variant and filled out a demographic questionnaire. Each participant took approximately 25 minutes to finish the experiment.

7.2.2.5 Participants

We recruited 12 volunteer participants (5 female), aged between 25 and 54 years (M=35.75, SD=10.38). None suffered from color vision impairment, 8 had nor-

---

[6] Unity3D. `www.unity3d.com`, last retrieved April 21, 2020

mal vision, and 4 had corrected-to-normal vision. We asked the participants to rate their experience with Augmented Reality on a 5-point Likert-scale. The participants stated that they had had limited experience (Md=2, IQR=1.5).

7.2.2.6 Results

Search Time Performance

For the first task, we consider the effects of one factor (assistance) on search time to locate out-of-view objects. The mean search times for assistance are: no helplines=4.51s, basic helplines=4.89s, and all helplines=6.02s. The search times are compared in Figure 7.11.

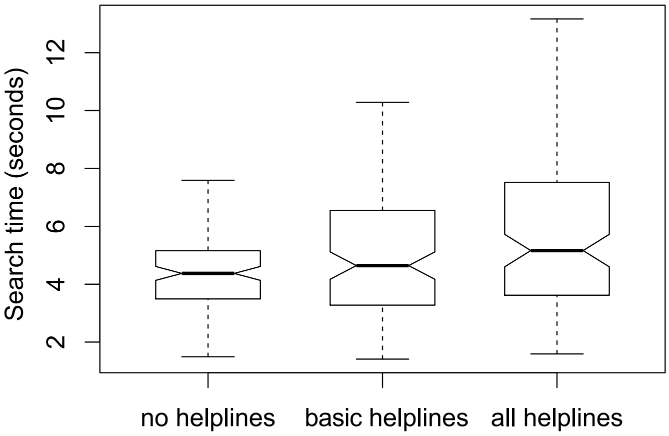


Figure 7.11: Boxplot of search times for different variants of EyeSee360 for out-of-view object visualization.

A Shapiro-Wilk-Test showed that our search time data is not normally distributed (p<0.001), and thereafter we ran a Friedman test that revealed a significant effect of assistance on search time ($\chi^2(2)=7.27$, p=0.026, N=12). A post-hoc test using Wilcoxon Signed-rank with Bonferroni-Holm correction showed significant differences between some of the conditions (see Table 7.5). All helplines has significantly higher search times than zero helplines or no helplines.

Table 7.5: Pairwise comparisons of Variants of EyeSee360 (ϕ-values report the calculated effect sizes).

Comparison	**p-value**	**ϕ-value**
No helplines vs. basic helplines	0.142	0.10
No helplines vs. all helplines	<0.001	0.26
Basic helplines vs. all helplines	0.001	0.21

Likert-scale Questionnaire

After each condition, we asked the participants to answer two questions with 5-point Likert-scale items (1=strongly disagree, 5=strongly agree). The results are shown in Table 7.6 and Figure 7.12. Participants stated that they could quickly locate the out-of-view object for all variants: no helplines=4 (IQR=1), basic helplines=4 (IQR=0), and all helplines=4 (IQR=1). Further, participants stated that they did not get distracted by the visualization for most of the variants: no helplines=1 (IQR=1.25), basic helplines=2 (IQR=1.25), and all helplines=3 (IQR=2.25).

Table 7.6: Results from 5-point Likert-item questionnaires.

Condition	**Performance**		**Distraction**	
	Md	IQR	Md	IQR
No helplines	4	1	1	1.25
Basic helplines	4	0	2	1.25
All helplines	4	1	3	2.25

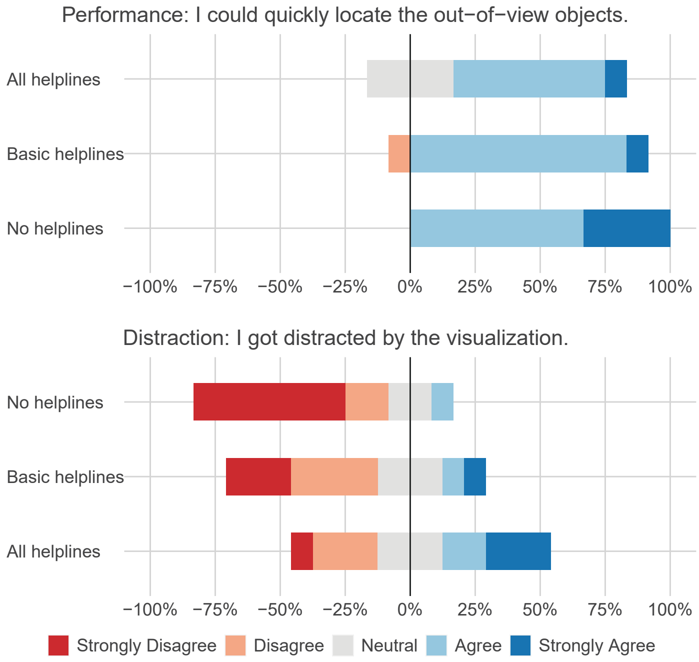


Figure 7.12: Results from 5-point Likert-item questionnaires. *Best seen in color.*

At the end of the experiment, we asked participants for their favorite variant. Nine participants stated that they preferred no helplines and three stated that they preferred basic helplines.

7.2.2.7 Discussion

Search Time Performance

We observed better search times for EyeSee360 for optical see-through AR than with the previously used video see-through AR (see Section 7.1). However, we think this is mostly due to the fact that no click was required to prove that the out-of-view object had been found. From our results, we saw that no helplines and zero helplines worked significantly faster than all helplines. Therefore, we can accept our hypothesis H_{20}. We found no significant difference between no helplines and zero helplines. However, the subjective results show that participants preferred no helplines and stated that this variant has the least visual clutter.

Direction Estimation

In some situations, it may be enough to understand the location of a virtual object without needing to look at the object itself (e.g., when looking for some free space to place new virtual objects). We argue that, in those situations, the variant with all helplines may be more efficient because it gives additional information about the exact locations of those objects.

Visualization on Demand

Our results show that the EyeSee360 variant with the least visual clutter performs best in terms of search time performance. However, for small field-of-view HMDs, there is still a lot of visual information overlaying virtual and real objects, which can interfere when users interact with it. We recommend using EyeSee360 on demand. Like a task manager that gives one an overview of all active tasks, the visualization could work as a location manager that gives one an overview of the locations of all virtual objects.

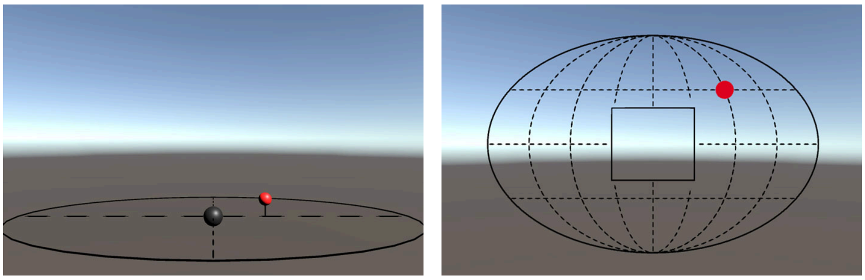
Limitations

For small field-of-view HMDs, there is still a lot of visual information overlaying virtual and real objects, which can interfere when users interact with it. For example, EyeSee360 encodes the precise location of an out-of-view object, although the general direction in which a user needs to turn his head might be sufficient.

7.2.3 Conclusion

In this section, we compared three different levels of assistance for EyeSee360 in terms of search time performance. Our results show that the variants with less visual clutter (no helplines, basic helplines) perform significantly better than the variant with the most visual clutter (all helplines). Even though there were no performance differences between the variants no helplines and basic helplines, participants preferred the variant with no helplines in most cases. For the best experience, we recommend showing the visualization technique only when users need to search for out-of-view objects.

7.3 Comparing Techniques for Moving Out-of-View Objects



(a) 3D Radar.

(b) EyeSee360.

Figure 7.13: Selected visualization techniques. The red sphere represents the same out-of-view object in both techniques.

Over the past few years, head-mounted Virtual Reality (VR) devices have been steadily advancing from a technological point of view. These advances in VR technology allow it to be used for a variety of applications (e.g., training [JPD$^+$18], simulation [LCRR18], or gaming [RGR17]). However, in many of these applications, the limited field-of-view (FOV) of the VR device leads to objects receding from view (e.g., opponents in computer games). This is a problem because users cannot perceive these out-of-view objects, and therefore have no information about their positions or movements (e.g., to avoid accidents during ship docking a pilot has to keep track of multiple potentially occluded tugboats while assisting a vessel at the same time [GSB$^+$18]). Furthermore, due to the restricted human visual range [Kal01], increasing the FOV of VR devices [OWK$^+$14, XB16] would not solve this problem. In previous work, different techniques have been proposed to visualize the positions of out-of-view objects on head-mounted devices (e.g., 3D Radar [BSEN18] or EyeSee360 (see Section 7.1)). Thus far, all developed techniques have been evaluated with out-of-view objects located at fixed positions. However, visualizing only static out-of-view objects is insufficient for many scenarios because the movement of these objects can be crucial. For example, when a user plays a VR spaceship game in which they must avoid being attacked by an opponent, knowledge of both position and movement becomes critical for a successful counterattack. Besides in computer games, moving out-of-view objects may be relevant in traffic scenarios (e.g., when determining whether it is safe to overtake [LHB15]), or in monitoring tasks (e.g., when assessing the position and movement of tugboats during the docking process of container vessels [GSB$^+$18]).

In this section, we compare the two well-known visualization approaches (Overview+detail and Focus+context (Contextual views)) in a user study to evaluate their performances for the visualization of moving out-of-view objects. For each approach, we selected a representative visualization technique from prior work: 1) 3D Radar [BSEN18] for Overview+detail, and EyeSee360 (see Section 7.1) for Focus+context. To gain novel insights into how well these techniques visualize moving out-of-view objects, we compare their performances for three different kinds of movements derived from various use cases. Therefore, we conducted a laboratory user study with 15 participants in VR. We measured movement estimation error, assessed usability with the System Usability Scale (SUS) [Bro96], and evaluated subjective performance with an individual questionnaire.

Here, our research contributions include:

- A comparison of two techniques 1) EyeSee360, and 2) Radar3D for moving out-of-view objects in Virtual Reality.

The work presented in this section was published as a short paper at the IEEE VR conference in 2019 [GKL+19].

7.3.1 Visualizing Moving Out-of-View Objects

To investigate how to best visualize moving out-of-view objects, we first identified two relevant visualization approaches from related work: Overview+detail and Focus+context (including the later-derived Contextual views approaches). Both approaches are suitable for visualizing out-of-view objects; however, in previous work, the Focus+context approach worked best (cf. Section 7.1, [BSEN18]). However, all out-of-view objects tested in previous work were always located at fixed positions and not moving in 3D space. For moving out-of-view objects, we hypothesize that the compression of information along the borders in the Focus+context approaches not only leads to losing the proportions of 3D space, but also makes it harder to understand movement. Therefore, we expect the Overview+detail approach to work best for visualizing moving out-of-view objects. To investigate this, we selected a representative technique for each approach: 1) Radar3D for Overview+detail, and 2) EyeSee360 for Focus+context. We then compared them in a user study.

7.3.1.1 3D Radar Technique

3D Radar is a visualization technique that is frequently used in various computer games (e.g., Elite Dangerous). Our implementation is based on previous work [BSEN18]. Figure 7.13a shows how 3D Radar looks. 3D Radar uses a sphere in its center to represent the user. A circle around this sphere represents zero on the y-axis. The zeros on the x- and z-axes are represented by two dotted lines.

To avoid cluttering the foveal vision of the user, we moved the technique from the center of the screen to the bottom of the screen, directly in front of the user. Each out-of-view object is represented by a sphere (called a proxy). Each proxy is placed relative to the sphere in the center where it is placed relative to the user in the real world. If the object is higher or lower than the user, a line is drawn between the proxy and ground circle to show where on the circle the object would be without height.

7.3.1.2 EyeSee360 Technique

EyeSee360 is a technique for visualizing the 3D positions of out-of-view objects in the user's periphery (cf. see Section 7.1). Figure 7.13b shows how EyeSee360 looks. EyeSee360 concentrates information about out-of-view objects on a grid system located in the user's periphery. This grid system compresses 3D position information on a single 2D plane. The inner rectangle of EyeSee360 represents the FOV of the current user, and the area outside the rectangle represents the area outside of the user's view. More information on EyeSee360 can be found in Section 7.1. Here we used the variant of EyeSee360 with all helplines because it offers the highest amount of visual assistance and because we develop it for Virtual Reality, which offers a larger field-of-view compared to the previously tested AR devices.

7.3.1.3 Investigated Types of Movement

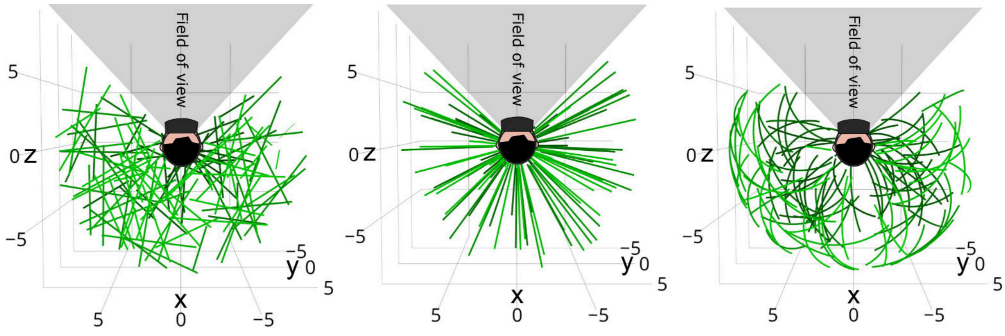
In our comparative user study, we want to compare different types of movement. These movement types are derived from various scenarios. We distinguish three kinds of movement:

(M1) Linear movement from any point A to any point B.

(M2) Distance movement towards or away from the user.

(M3) Orbital movement around the user.

In general, the movement of out-of-view objects can be described by two points: a start point A and an end point B of the movement. The movement between these two points can then be a linear or non-linear movement. As a first general type of movement, we consider the linear movement between any point A and any point B *(M1)*. This kind of movement is for example relevant for determining whether an out-of-view object will cross the user's path (e.g., to know whether it is safe to overtake). The next type of movement *(M2)* is a special case of *(M1)*. It describes objects moving towards or away from the user. These movements are especially critical in computer games in which a user wants to know if an opponent is moving towards or away from them. Besides linearly moving out-of-view objects, we also consider orbital movement around the user *(M3)*. This kind of movement is relevant to the user because a change in direction requires a different head movement for localizing an out-of-view object. It is also relevant in monitoring tasks (e.g., to assess the movement of tugboats around a larger container vessel).



(a) Linear *(M1)*. (b) Distance *(M2)*. (c) Orbital *(M3)*.

Figure 7.14: Randomly generated movements tested in the user study (color gradient from light green (y=10) to dark green (y=-10)). Each line represents a tested movement. The origin of the coordinate system (0,0,0) was calibrated to the head of the participants. *Best seen in color.*

7.3.2 Experiment

7.3.2.1 Study Design

To evaluate the performance of each visualization strategy for moving out-of-view objects, we conducted a within-subjects controlled laboratory study in VR. We investigate whether the dependent variables, movement estimation error and estimation error of start and end positions, are influenced by the independent variables, visualization (3D Radar vs. EyeSee360) and type of movement (*M1* vs. *M2* vs. *M3*). We calculated the movement estimation error as the angle between the movement vector (from start to end position) and the user's assessment of this movement vector (assessed start to end position). The estimation error of the start position is calculated as the angle between the start position of the movement and the user's assessment of the position. The estimation error of the end position is calculated in similar way. For all trials, we set the number of objects to one at a time, the movement duration to five seconds, and the movement length to eight meters. Our repeated-measures within-subjects factorial design results in six different conditions.

For this study, we derived the following sub-question from our third research question: *(RQ3d) Which approach (Focus+context, Overview+detail) works best for encoding direction and distance for moving out-of-view objects?* We posit the following hypotheses:

H_{21} We expect 3D Radar to perform better than EyeSee360 with regard to movement estimation error for all movements because the Overview+detail approach is able to keep the proportions of 3D space, therefore allowing the user to better assess the movement.

H_{22} We expect that 3D Radar will be subjectively perceived as best.

7.3.2.2 Procedure

The study was divided into two counter-balanced blocks, with each block testing one visualization technique (3D Radar and EyeSee360). In each block, we tested three different types of movement (*M1* vs. *M2* vs. *M3*) with seven iterations[7]. The tested types of movement were randomly generated and are fully visualized in Figure 7.14. The randomly generated movements of the first block were stored and tested in a randomized order for the second block. This ensured that we tested the same movements for both techniques.

All movements we tested have the same length (eight meters) and are within a sphere S with a diameter of ten meters (one meter in our VR environment represents one meter in the real world). For each tested movement, we randomly selected a start position A on the surface of the sphere S. The end position B was then selected depending on the type of movement. For linear movement *(M1)*, we randomly selected a point B with a distance of eight meters to point A. For distance movement *(M2)*, point B is the result of point A divided by five. Since point A has a distance of ten meters, point B has a distance of two meters from the user, and the resulting movement has a length of eight meters. For orbital movement *(M3)*, we randomly selected a point B on sphere S with the great-circle distance of eight meters on sphere S. For each movement, we checked if all points of that movement were within sphere S, not closer than two meters from the user, and outside of the user's FOV. If this was not the case, we generated a new random movement. Further, we randomly selected 50% of the movements and inverted their directions.

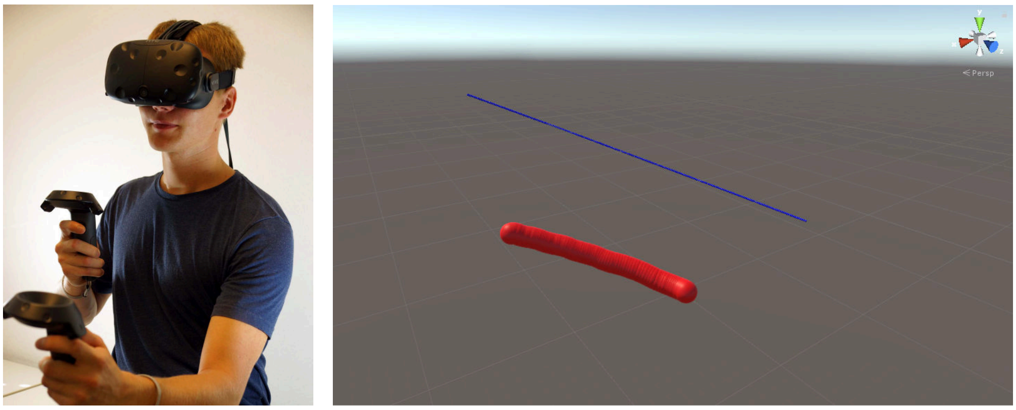
We started our experiment with a short introduction to out-of-view objects and VR. Afterwards, participants started with the two blocks. Each block started with three test trials (not included in results), along with an explanation of the visualization technique and the task to achieve. After each block, we asked participants to fill out a System Usability Scale (SUS) questionnaire [Bro96], which measures usability. Further, at the end of the experiment, participants were asked to fill out our individual subjective questionnaire and a demographic questionnaire. Overall, each participant took approximately 45 minutes to finish.

In each iteration, the user had to focus on a static point in front of them. Then the out-of-view object moved from the generated start position to the generated end position (always five seconds). The moving out-of-view object was invisible

---

[7] This number was derived from pretesting.

and was only visualized with a proxy in the current tested visualization technique. After the movement was over, the visualization technique disappeared, and the user was asked to draw the movement they saw with a tracked controller from start to end position. A dotted line indicated the input to the user. It was possible to reenter the perceived movement. Further, participants were allowed to input relative positions (e.g., the point (10,0,0) could be entered as (1,0,0) if all other points of that movement were scaled accordingly).



(a) VR user. (b) Example user input (red); object movement (blue).

Figure 7.15: User input for experiment in Virtual Reality.

7.3.2.3 Implementation

Both visualization techniques are implemented in Unity3D, a 3D game development platform, and the HTC Vive HMD, a head-mounted VR device. We adjusted both techniques to the maximum distance (ten meters) of the tested movements. For 3D Radar we adjusted the black circle to the maximum distance, and in EyeSee360 we encoded the distances of out-of-view objects with the visual cue color, from blue (maximum distance) to red (minimum distance).

7.3.2.4 Participants

We recruited 15 participants (6 female), aged between 18 and 35 (M=23, SD= 4.36). None of them suffered from color vision impairment. All had normal or corrected-to-normal vision. Participants with corrected vision wore contact lenses.

7.3.2.5 Results

For the experiment, we consider the effects of the two factors (visualization, type of movement) on movement estimation error and estimation error of start and end positions.

Estimation Error of Start and End Positions

The mean estimation errors of start positions for the visualization techniques are: 3D Radar=39.9° and EyeSee360=33.6°. Normality here was not assumed because the Shapiro-Wilk test was significant ($p<0.001$). A Wilcoxon Signed-rank with Holm-Bonferroni correction revealed a significant difference between the two visualization techniques ($W=20353$, $Z=-2.80$, $p=0.005$, $\phi=0.16$).

The mean estimation errors of end position for the different techniques are: 3D Radar=37.2° and EyeSee360=37.4°. Normality was not assumed because the Shapiro-Wilk test was significant ($p<0.001$). A Wilcoxon Signed-rank with Holm-Bonferroni correction revealed no significant difference between the two visualization techniques ($W=25766$, $Z=0.55$, $p=0.587$, $\phi=0.03$).

Movement Estimation Error

The mean movement estimation errors for the visualization techniques are: 3D Radar=30.7° and EyeSee360=46.1°. Normality here was not assumed because the Shapiro-Wilk test was significant ($p<0.001$). A Wilcoxon Signed-rank with Holm-Bonferroni correction showed a significant effect of visualization technique on movement estimation error ($W=33788$, $Z=5.50$, $p<0.001$, $\phi=0.31$).

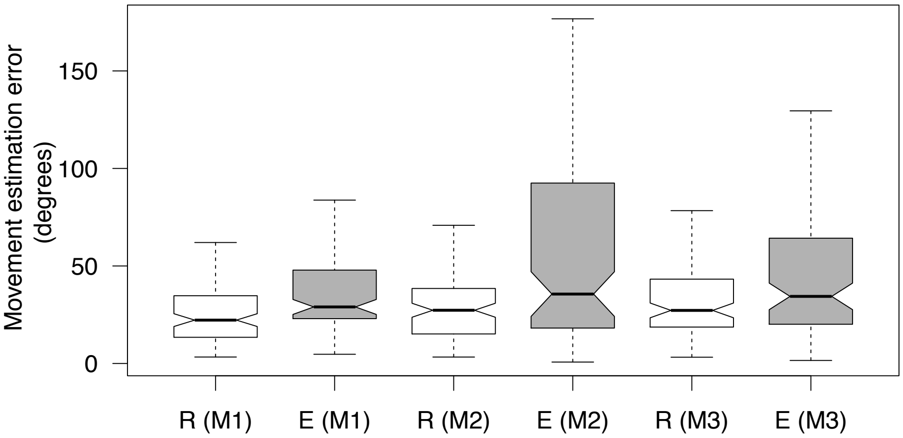


Figure 7.16: Boxplot of movement estimation error (R=3D Radar, E=EyeSee360; value in parentheses indicates the tested movement type).

To further understand the movement estimation errors for the visualization techniques, we compared the mean errors for the different types of movement (see Figure 7.16). We ran a Friedman test that revealed a significant effect of visualization technique and movement type on movement estimation error ($\chi^2(5)=23.275$, $p<0.001$, $N=15$). A post-hoc test using Wilcoxon Signed-rank with Holm-Bonferroni correction showed significant differences between all conditions (see Table 7.7).

Table 7.7: Pairwise comparison for different movement types.

3D Radar	**EyeSee360**	**Type**	**P-value**	**ϕ-value**
27.1°	37.9°	*M1*	<0.001	0.26
29.0°	53.5°	*M2*	<0.001	0.28
36.1°	47.1°	*M3*	0.050	0.14

System Usability Scale

For SUS scores, EyeSee360 scored 47.5. This score is clearly below acceptable usability and far from the SUS score reported in the previous sections (which was 68 for head-mounted AR; see Section 7.1). This may be because in these scenarios the out-of-view objects were moving, and EyeSee360 did not seem to be able to visualize these moving objects. However, 3D Radar scored 71.2, which is over the threshold for acceptable usability. Our results show that 3D Radar is usable for moving out-of-view objects in head-mounted Virtual Reality, while EyeSee360 is not.

Subjective Questionnaire

At the end of the study, we asked participants to answer four questions with 5-point Likert items. Participants stated that they were able to exactly determine the movements of out-of-view objects with 3D Radar (Md=4, IQR=0), but not with EyeSee360 (Md=2, IQR=1.5). Furthermore, they stated that they were able to determine the movement quickly with 3D Radar (Md=4, IQR=0), but not with EyeSee360 (Md=2, IQR=1). Overall, 13 participants preferred 3D Radar, while only two preferred EyeSee360.

7.3.3 Discussion and Limitations

In the following subsection, we discuss the results and limitations of our conducted experiment in Virtual Reality.

Overview+detail vs. Focus+context

Focus+context approaches do not have the ability to keep the proportions of 3D space, since they compress all information along the borders of the screen. In our results, we showed that EyeSee360 performed worse than 3D Radar with regard to movement estimation error. We think that this is mainly due to the Focus+context approach that is used for EyeSee360. We predicted this outcome in hypothesis H_{21}, and therefore we can accept our hypothesis H_{21}. However, the Overview+detail approach has disadvantages when it comes to quickly locating out-of-view objects (e.g., a visual cue in EyeSee360 already encodes in which direction the user should move their head to locate the out-of-view object).

Linear Movement

Almost all participants (n=12) perceived the movement of linearly moving objects as orbital in EyeSee360, while the participants perceived it as linear with 3D Radar. Although, some participants had difficulties drawing straight lines, the lines of trials with 3D Radar were less curved. We think this is due to the mapping of 3D direction information on a 2D plane in EyeSee360.

Distance Movement

The encoding of distance with color was problematic in EyeSee360 (cp. Figure 7.14b (M2)). Here, users often mixed up the direction of the movement (i.e., perceived objects as if they were approaching when they were actually moving away, or vice versa) resulting in a high movement estimation error up to 180°. The problem is that color has no spatial attributes, and therefore encoding 3D distance with color is problematic. We think that changing the proxy size depending on the distance to the out-of-view object may improve the perception of distance in EyeSee360.

Orbital Movement

In our study, we investigated different types of movement. Interestingly, 3D Radar performed better than EyeSee360 for all three types of movement. However, both visualization techniques had problems with visualization of orbital movement. Here, the orbital movement could be better reproduced by all participants with 3D Radar. With EyeSee360, it seemed to depend strongly on how intuitively the color coding and movement was perceived. We think that orbital movement was problematic for participants in general because it was harder to draw that movement with the tracked controller.

Ecological Validity

In our study, we evaluated both techniques in a simple “ground and sky” scene. We think this is a suitable approach to gather first insights into how well moving objects out of view are perceived in the tested techniques. However, future work should evaluate the two techniques in more realistic scenes.

Different Parameters

We focused on comparing two different approaches for visualizing moving objects out of view. Therefore, we reduced the complexity of our user study by reducing the number of independent variables in our design. However, future work should investigate how users can perceive the movements of multiple out-of-view objects. Further, movements with different durations and lengths can be tested.

7.3.4 Conclusion

In this section, we compared two visualization approaches, Overview+detail and Focus+context, for moving out-of-view objects in a user study. We selected one representative visualization technique for each approach: 1) 3D Radar for Overview+detail and 2) EyeSee360 for Focus+context. Our results show that 3D Radar objectively and subjectively works best for moving out-of-view objects. Furthermore, 3D Radar can encode more information, such as the orientation or size of an object out of view. In future work, we want to test both techniques in more realistic scenarios (e.g., VR games).

7.4 Summary

In this chapter, we developed a novel visualization technique for encoding direction and distance for multiple objects out-of-view. First, we proposed a prototyping tool that allows quick testing and refining of design ideas for head-mounted Augmented Reality. Thereafter, we developed three different variants of EyeSee360 and tested them with our prototyping tool in a user study. Then, we selected the best performing variant and compared it to our adapted off-screen visualization techniques with regard to direction estimation accuracy. In our results, we could highlight the usefulness of our prototyping tool and found that EyeSee360 results in the lowest error for direction estimation of objects hidden from view.

Afterwards, we transferred EyeSee360 from video to optical see-through AR and compared all three variants of EyeSee360 in terms of search time performance using the Hololens. Our results show that variants of EyeSee360 with less assistance result in shorter search times and are preferred by users over variants of EyeSee360 that offer more visual assistance.

In the last part of this chapter, we investigated the visualization of moving out-of-view objects. Here, we compared EyeSee360 with another technique from computer games called 3D Radar for three different types of movement. We found that using 3D Radar resulted in a significantly lower movement estimation error and higher usability, as measured by the system usability scale [Bro96]. Furthermore, 3D Radar was preferred by 13 out of 15 participants for visualization of moving out-of-view objects.

8 Conclusions and Outlook for Future Work

In this chapter, we conclude the work described in all foregoing chapters. We start with a brief synopsis of the work carried out in this thesis (see Section 8.1). After, we highlight the contributions this work makes to the research questions described in Chapter 4. Here, we summarize our contributions and then, we provide detailed answers to the research questions (see Section 8.2). Thereafter, we give recommendations for the scenarios described in the introduction (see Section 8.3). We finish with an outlook for future work and predict when and in which application field we expect the used technology to flourish (see Section 8.4).

8.1 Synopsis

In this thesis, we followed the human-centered design approach to develop and evaluate visual cues in Mixed Reality that assist users with locating out-of-view objects. To explore this, we identified three clearly distinguishable scenarios that all share the problem of objects receding from view. For each of these scenarios, we conducted a context of use analysis consisting of an ethnographic study, accident analysis, and literature reviews. Thereafter, we abstracted from these concrete scenarios and defined the underlying problem. Then, we derived three research questions, based on background knowledge and related work. Thereafter, we studied each of the derived research questions in depth in the presented work. The research questions focus on the following three aspects: 1) To what extent can existing off-screen visualization techniques be adapted to cue the direction to out-of-view objects in Mixed Reality?, 2) How can Mixed Reality devices with small fields-of-view be extended to present directional cues to out-of-view objects?, and 3) In what way can the directions and distances of out-of-view objects be visualized for moving or non-moving objects in Mixed Reality? Our results show that directional cues to out-of-view objects are easy to perceive and help the user to locate these objects. However, when a user needs more information, such as the position or movement of out-of-view objects, it is possible to convey that information, albeit with increased workload and additional visual clutter. Furthermore, visualizing objects only when they are outside of the user's field-of-view is insufficient, since assistance when these objects appear in view or are occluded improves error rates and overall performance of the visual assistance.

8.2 Contributions to the Research Questions

In the following, we will highlight our contributions to the research questions addressed in this thesis. First, we will summarize the core contributions to each research question. After, we provide detailed answers that aim to answer the research questions.

8.2.1 RQ1: To what extent can existing off-screen visualization techniques be adapted to cue the direction to out-of-view objects in Mixed Reality?

As a first step to answer this research question, we suggested projecting the relevant out-of-view objects from 3D space onto a 2D plane in front of the user. Thereby, we were able to apply three well-known 2D off-screen visualization techniques (Arrow, Halo, and Wedge) to these projected out-of-view objects using that 2D plane. This is somewhat similar to how these techniques visualize off-screen objects on a smartphone display. Our results indicate differences among the established techniques that we adapted for head-mounted Augmented Reality. We found that Halo resulted in the lowest error for direction estimation, while Wedge was subjectively perceived as best. Moreover, we observed that with our projection of out-of-view objects on a 2D plane, direction error increased when the angle between the user's line of sight and the out-of-view object increased. For this reason, the adapted visualization techniques are less feasible for scenarios in which the out-of-view objects are distributed 360° around the user.

Therefore, in our second step, we selected the two best visualization techniques (Halo and Wedge) from our previous study and further improved them to support out-of-view objects spatially distributed 360° around the user. For this purpose, we removed the projection of out-of-view objects on a 2D plane in front of the user and instead changed the 2D visual cues (Halo and Wedge) to 3D visual cues that point into the directions of out-of-view objects. In total, we developed four out-of-view visualization techniques, two for Virtual Reality (HaloVR, WedgeVR) and two for Augmented Reality (HaloAR, WedgeAR), and conducted two user studies in which we tested our techniques in Virtual and Augmented Reality. Since our techniques are inspired by off-screen visualization techniques [BR03, GBGI08], they can be perceived in a similar way and feel familiar to users. While our techniques resulted in overall high usability, we found the choice of AR or VR impacts mean search time (VR: 2.25s, AR: 3.92s) and mean direction estimation error (VR: 21.85°, AR: 32.91°). Moreover, while adding more out-of-view objects significantly affects search time across VR and AR, direction estimation performance remains unaffected. Furthermore, since we implemented out-of-view visualization techniques, we assumed that no visualization is necessary when the objects are visible on the screen. However, some participants stated during pilot tests that our artificial out-of-view objects all look the same and that this made them unable to decide for which object they were searching for when multiple out-of-view objects were closer together. To solve this problem, we decided to have 5° at the border of the screen where the visual cue remains active. However, it was still difficult for participants to distinguish between objects close together.

As our third and last step, we developed the multi-modal visualization technique FlyingARrow, which is based on Arrow, to point to out-of-view objects on small-screen devices. We compared FlyingARrow with EyeSee360 and found that it resulted in higher usability and lower workload. However, FlyingARrow performed slightly worse with respect to search time and direction error. Nevertheless, we showed that humans are to some extent able to mentally continue a uniform motion over a short period of time.

To summarize, Halo and Wedge are both suitable for cueing direction to out-of-view objects. Here, the number of visual cues presented at the same time (number of out-of-view objects represented on the screen) had no effect on how well participants were able to estimate the directions of out-of-view objects. However, reducing the number of visual cues that are present at the same time improves search time performance. Furthermore, on devices with small fields-of-view, our developed technique FlyingARrow was proven to be useful for encoding the location of a single out-of-view object.

8.2.2 RQ2: How can Mixed Reality devices with small fields-of-view be extended to present directional cues to out-of-view objects?

To answer this research question, we first developed a low-cost and highly adjustable prototyping tool (PeriMR), which allows one to create different kinds of peripheral light displays to extend the field-of-view of current Augmented and Virtual Reality devices. It builds on Google Cardboard and uses a smartphone for the Virtual and Augmented Reality experiences.

Afterwards, we developed RadialLight using the PeriMR prototyping tool. RadialLight is a low-cost radial peripheral display that augments existing HMDs, implemented as 18 radially positioned LEDs around each eye to cue direction. We show the effectiveness of RadialLight in terms of cue directional accuracy and out-of-view object search time in two user studies for VR. Key findings show that participants could not distinguish between LED cues presented to one or both eyes simultaneously and that participants estimated LED cue direction within a maximum 11.8° average deviation. However, while average direction deviation differences between the ship bridge (11.2°) and car cockpit (11.8°) scenarios showed no significant effects, search time performance for locating out-of-view objects was affected. Therefore, we conclude that out-of-view objects in less distracting scenarios are selected more quickly with a radial light display.

After we showed that Virtual Reality devices can be extended to present directional cues, we developed a similar technique for Augmented Reality. Instead of using a radial light display for each eye, we decided to use a single radial light display, which we call MonoculAR. We propose two implementations of MonoculAR: 1) additional LEDs attached to the device, or 2) virtual LEDs presented on the screen of the device. We compared the performances of these two implementations in a user study, discovering that participants find out-of-view objects more quickly when the light cues are presented on the screen and closer to the center of their field-of-view. Moreover, our results suggest visualizing only one out-of-view object at a time to improve search times.

As the last step, we developed a novel display which augments peripheral vision with warning information to draw the user's visual attention towards potential hazards. Here, we wanted to see how useful peripheral light displays are for guiding the attention. Therefore, in a user study, we compared three different visual cues with regard to response times and error rates. Overall, we could show that all light stimuli were suitable for shifting the users' attention (100% correct). However, a motion light cue resulted in significantly shorter response times and was subjectively perceived as best.

To summarize, we were able to show that radial peripheral LEDs are suitable for directional cueing, while the LED color does not strongly affect performance. However, visual cues presented closer to the foveal vision are more effective. Furthermore, radial monocular displays seem to be sufficient for guiding to out-of-view objects. Additionally, we found visual cues that incorporate motion presented in the periphery to be highly effective for guiding the user's attention.

8.2.3 RQ3: In what way can the directions and distances of out-of-view objects be visualized for moving or non-moving objects in Mixed Reality?

To address our research question, we started by developing a lo-fi head-mounted prototyping tool that allows quick testing and refining of design ideas. Thereafter, we designed and implemented a novel out-of-view object visualization technique (EyeSee360) that allows one to visualize the directions and distances of out-of-view objects. We evaluated the technique in two studies. In the first study, we tested three different variants of the technique with our lo-fi prototyping tool. We argue for the usefulness of this approach, as it both saves development time and allows designers to test and iterate quickly based on user feedback. In the second study, we compared the best-working variant of EyeSee360 from our first study to our three adapted 2D off-screen visualization techniques (Arrow, Halo, and Wedge) for non-moving out-of-view objects. Since EyeSee360 was inspired by 2D off-screen visualization techniques such as EdgeRadar [GI07], it is somewhat like these techniques. In our results, we found that using EyeSee360 results in the lowest error for direction estimation of out-of-view objects. Furthermore, participants were able to estimate the relative distances of out-of-view objects.

Although EyeSee360 was designed to visualize both direction and distance for out-of-view objects, we further wanted to investigate which variant of EyeSee360 works best for search tasks. In our previous work, we tested EyeSee360 for video see-through Augmented Reality. Here, we wanted to evaluate how well the technique works for optical see-through Augmented Reality. Therefore, we conducted a user study in which we compared all three variants of EyeSee360 with regard to best search time performance. From our results, we saw that using no helplines or basic helplines resulted in participants finding objects significantly more quickly than using all helplines. However, we argue that, in direction estimation tasks, the variant with all helplines would be more efficient because it gives additional information about the exact locations of objects. For small field-of-view HMDs, there is still a lot of visual information overlaying virtual and real objects, which can interfere when users interact with it. Therefore, we recommend using EyeSee360 on demand on such a device.

Recently, Bork et al. [BSEN18] compared six different visualization techniques for effective guidance towards out-of-view objects. They suggested two new Overview+detail techniques (3D Radar and Mirror Ball) and evaluated them against four existing Focus+context techniques: 3D Arrows [SHB10], Aroundplot [JHPR11], EyeSee360 [GHHB17] and sidebARs [SH13]. They found significantly lower completion times and better usability when using EyeSee360. Further, 3D Radar was the best-performing Overview+detail technique. However, they again only evaluated the guidance to non-moving out-of-view objects with fixed positions in 3D space. Therefore, in our last step, we took the best-performing Focus+context (Contextual views) technique, EyeSee360, and compared it to the best-performing Overview+detail technique, 3D Radar, for moving out-of-view objects in Virtual Reality. In our results, we showed that EyeSee360 performed worse than 3D Radar with regard to movement estimation error. We think this is because Focus+context approaches are unable to maintain the proportions of 3D space, since they compress all information along the borders of the screen. This is especially problematic when one has to estimate how much an out-of-view object moves over time. Almost all participants perceived the movement of linearly-moving objects as orbital in EyeSee360, while they perceived it as linear with 3D Radar. This indicates that the encoding of distance with color was problematic in EyeSee360.

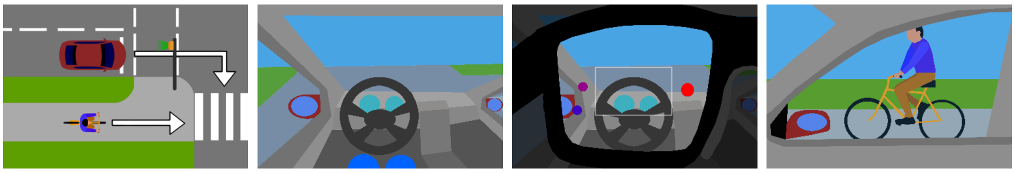
To summarize, our results showed that EyeSee360 performs best for encoding direction and distance to non-moving out-of-view objects. Here, color encoding can be used to encode relative distance. However, this color encoding was shown to be problematic with regard to moving out-of-view objects. Here, 3D Radar performed significantly better for all tested types of movement. This highlights the fact that an allocentric visualization techniques performs best for moving out-of-view objects.

8.3 Recommendations for the Scenarios

Aside from our research questions, our introduction included an explanation of three different scenarios in which out-of-view objects are an important problem (see Chapter 1). In this section, we discuss, based on our results, which kind of visualization technique is best suited for each scenario.

8.3.1 Traffic Encounter

This scenario consists of several stakeholders (e.g., car drivers, cyclists, pedestrians). However, the requirements for these stakeholders overlap somewhat. For example, there is often a single out-of-view object relevant at a time. Furthermore, the most important part is to quickly locate the out-of-view objects, whereas the accuracy with which users perceive the out-of-view objects is less important. Additionally, it is important that the center of the field-of-view stays uncluttered by information. Further, it makes sense to use an Augmented Reality solution since the users are in a real environment. With more advanced AR devices in the future, it may also be suitable to use a visualization technique presented on the AR glasses themselves (see Figure 8.1).



(a) Traffic situation. (b) Driver view. (c) Visual cue. (d) Cyclist.

Figure 8.1: An example to demonstrate our approach of using visual cues to guide attention presented in [GLB$^+$18].

As described in our other work [GLB$^+$18], to address the problem of unperceived critical out-of-view objects one could use two different strategies: 1) one could bring the information about the critical object to the user's attention or 2) bring the user's attention to the critical object. However, bringing the information to the user's attention would probably interfere with the user's primary task (e.g., driving a car) because information would overlay the user's focus. On the other hand, guiding the user's attention to the critical out-of-view object can be time consuming (e.g., if the object is behind the user). Therefore, in our paper [GLB$^+$18], we suggest a third strategy: 3) we guide the user's attention to visual cues of critical out-of-view objects in the user's periphery. Thereby, we do not overlay the user's focus and we can reduce the time required to locate objects. Furthermore, we continuously show all relevant out-of-view objects in the user's periphery and highlight visual cues that represent critical objects in their visual presentation. This enables users to be aware of their surroundings in general, not only in dangerous situations (e.g., to continuously monitor traffic in the periphery during automated driving to increase trust [WSFR17]).

However, without relying on more advanced future AR devices, we think that a peripheral display alone could be a good solution in this scenario as well. For example, a peripheral LED display that presents visual cues that guide in the directions of the out-of-view objects would be useful here (cf. Section 6.4). The most important point for any kind of solution is that it does not clutter the foveal vision, since this is required for the primary task (e.g., driving).

8.3.2 Ship Docking

In our ethnographic study (see Section 1.1), we identified that the ship docking process often involves multiple tugboats that assist in the process and are out of the pilot's field-of-view. Furthermore, these tugboats are often occluded by the ship hull or container, meaning they may not be visible even when the pilot looks into their direction. However, the pilot is mostly focused on the movement of the container ships and only monitors the tugboats as a secondary task. We believe that an Overview+detail approach is best suited for this scenario because our results showed that this approach works well for moving out-of-view objects (see Section 7.3). Therefore, we suggest using 3D Radar for this scenario. Since 3D Radar presents the situation as a miniature version of the situation from an overview perspective, the technique is additionally able to encode more information about the tugboats, such as orientation, form, or size.

To evaluate different human-computer interfaces for different maritime scenarios, we developed the portable ship bridge simulator Matjes [SGS$^+$18]. However, from our ethnographic study and interviews with harbor pilots [GSB$^+$18], we learned that pilots use the full extent of the ship bridge to walk to different positions for the best possible view. This is problematic because it requires a large setup (depending on the vessel around 70 meters in width) and corresponding visualization of the harbor area. Therefore, we would recommend an evaluation using a Virtual Reality setup, in which pilots relocate by using virtual movement.

8.3.3 Fire Fighting

The last of our three scenarios from the introduction is fire fighting. Here, we focused more on the training aspect than firefighting under real circumstances. In this scenario, the firefighter has to locate several relevant objects in a building that is on fire (e.g., people that need to be rescued or a fuse box to turn off electricity). Depending on the scenario, one or multiple objects must be located. However, smoke may reduce the sight of the firefighter to effective blindness and no navigation information is available. In this scenario, the developed visualization technique EyeSee360 may be best suited because the objects do not change their positions (at least they are not constantly moving). Furthermore, EyeSee360 encodes the relative distances to the out-of-view objects with color, allowing fire fighters to quickly assess which objects are closer quickly.

8.4 Future Work

In the past decade, Mixed Reality technology has been growing rapidly. We observed an expanding consumer market for Virtual Reality headsets (e.g., Oculus, HTC, Microsoft, Samsung, Dell, etc.). Furthermore, the Augmented Reality market is developing as well, especially the business to business market, with different prototypes of AR glasses already available today (e.g., Epson, Microsoft, Magic Leap, etc.). However, compared to VR, the AR market is not growing as strongly, with some companies also going bankrupt (e.g., Meta, DAQRI). Therefore, we think that VR for training and education is already here, while we expect AR for business and industrial scenarios to take some time to develop (e.g., AR guidance for pilot's during the ship docking process or AR in manufacturing plants). Here, we think it will take several more years before Augmented Reality headsets and glasses finally hit the consumer market, and even more time before these devices can be used by pedestrians, cyclists or car drivers.

Apart from the results presented in this thesis, we already conducted some research studies to investigate some possible directions for future research. In the following, we would like to highlight the four most promising approaches:

1. We think that it may be helpful to show additional information on demand to reduce visual clutter, especially on small field-of-view devices. The fundamental idea here is to use eye tracking to measure the user's gaze and to show additional information about the out-of-view object when the user looks at the visual cue that represents that object. Here, we first developed EyeMR, a tool for rapid prototyping with eye tracking in Mixed Reality [SGB18]. Thereafter, we used the prototyping tool to develop an EyeSee360 variant that shows the name of an out-of-view object, when one looks at the visual cue that represents the object [GBH18].

2. It may be interesting to further investigate how the developed visualization techniques can be used to guide the attention of the user to an out-of-view object. We showed that this is feasible for car drivers [GLB+18] and for pedestrians [GSJ+18]. However, this may be relevant in more scenarios and it would be interesting to see if users interpret the visualization technique well enough that looking at the visual cue is all that is required to safely locate the represented out-of-view object. Furthermore, it may be promising to combine the visual cues with cues from another modality as we did with auditory cues [LBS+17, GLH+18] or haptic cues [SLG+18].

3. Building on the previous point, it may also be interesting to see if the visualization of out-of-view objects has a positive effect on the situation awareness of users in different situations. We conducted a first user study that showed promising results for increasing the situational awareness of pedestrians that navigate traffic while on their smartphones [JLC+18].

4. The last point we want to make is that in-view visual cues are very useful. From our results, we saw that visualization of out-of-view objects may not be enough because guidance may also be required when these objects appear in the field-of-view. Therefore, we investigated a scenario in which nearby

objects had to be located and used it to compare in-view visualization, out-of-view visualization, and their combination to a printed map. We found that, in certain scenarios in which a smaller number of objects at fixed positions needs to be located, in-view visualization techniques are more useful than visualization techniques for out-of-view objects [GPH19b]. However, additional research is required that focuses more on larger numbers of out-of-view objects in locations not known to the user is required.

In our work, we developed and evaluated visual cues that assist users in locating out-of-view objects. However, we focused strongly on fundamental research that is not focused on a specific scenario but rather looks at different types of the problem in a more general context. Therefore, the ecological validity of our techniques need to be investigated further. We suggest that future work should focus on specific use cases and build upon the findings presented in this thesis to select promising visualization techniques for specific usecases and compare them to one another.

8.5 Concluding Remarks

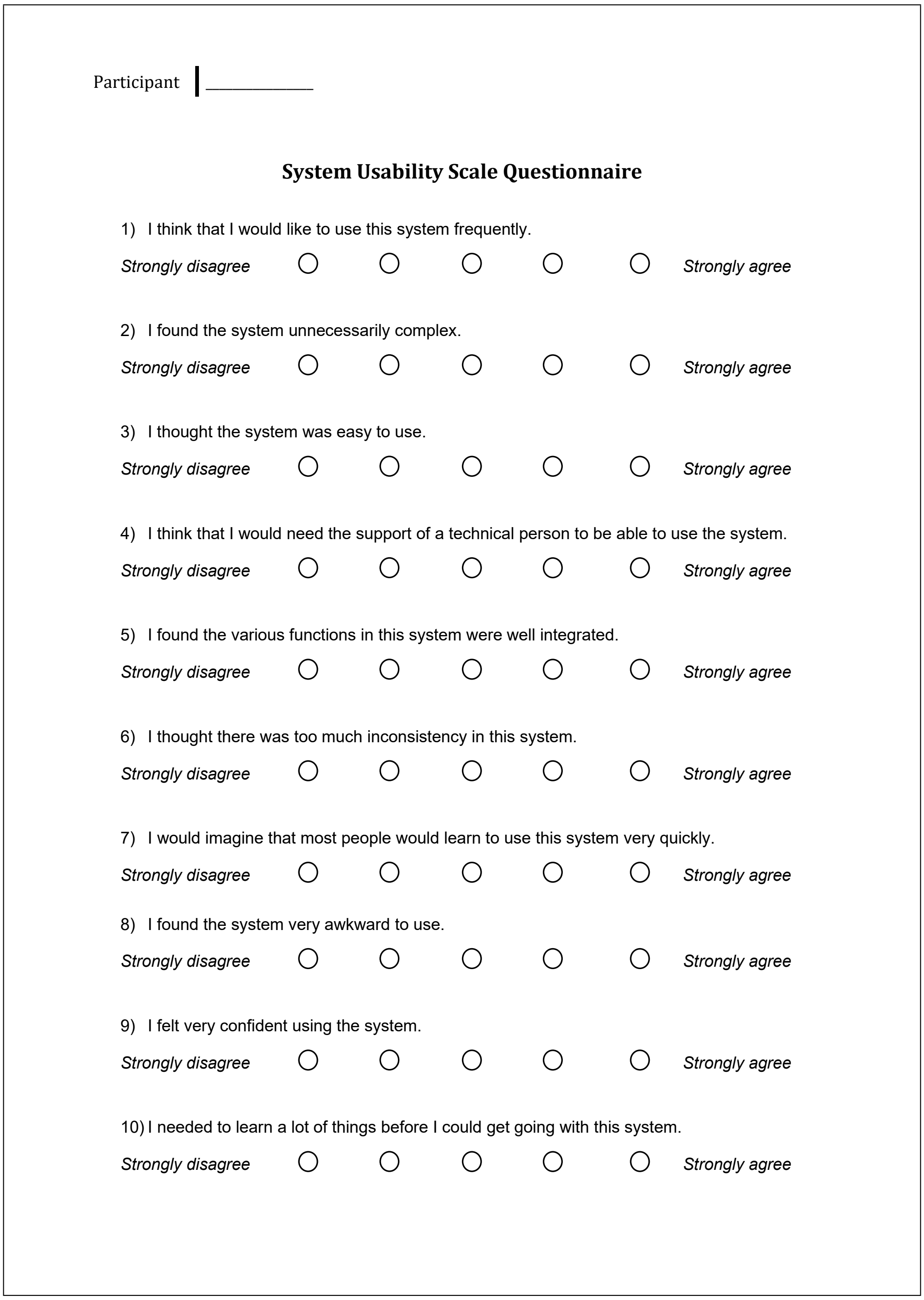
In this thesis, we tested all developed visualization techniques in Virtual Reality as well as in Augmented Reality in controlled laboratory studies. This allowed us to reduce the number of influencing factors and to test the techniques under optimal circumstances. However, this limits our understanding of how such techniques can be used in various real-life scenarios. Nevertheless, our work invites such ecological testing of out-of-view visualization techniques as a future research agenda. Therefore, our work opens up avenues for further investigation of out-of-view object visualization techniques, where we believe ecological testing and lowering direction estimation error will improve the adoption of such visualization approaches.

Furthermore, all visualization techniques were developed with Unity3D, enabling support for all major Augmented and Virtual Reality platforms. All visualization techniques that were developed in this thesis are available for free on GitHub under the MIT open source license[1]. We did this to enable researchers and developer all over the world to test the proposed visualization techniques and further improve them or evaluate them for specific scenarios.

---

[1] GitHub OutOfView-Project. `www.github.com/UweGruenefeld/OutOfView`, last retrieved April 21, 2020

A System Usability Scale Questionnaire



Participant | ____________

System Usability Scale Questionnaire

1) I think that I would like to use this system frequently.

Strongly disagree ○ ○ ○ ○ ○ *Strongly agree*

2) I found the system unnecessarily complex.

Strongly disagree ○ ○ ○ ○ ○ *Strongly agree*

3) I thought the system was easy to use.

Strongly disagree ○ ○ ○ ○ ○ *Strongly agree*

4) I think that I would need the support of a technical person to be able to use the system.

Strongly disagree ○ ○ ○ ○ ○ *Strongly agree*

5) I found the various functions in this system were well integrated.

Strongly disagree ○ ○ ○ ○ ○ *Strongly agree*

6) I thought there was too much inconsistency in this system.

Strongly disagree ○ ○ ○ ○ ○ *Strongly agree*

7) I would imagine that most people would learn to use this system very quickly.

Strongly disagree ○ ○ ○ ○ ○ *Strongly agree*

8) I found the system very awkward to use.

Strongly disagree ○ ○ ○ ○ ○ *Strongly agree*

9) I felt very confident using the system.

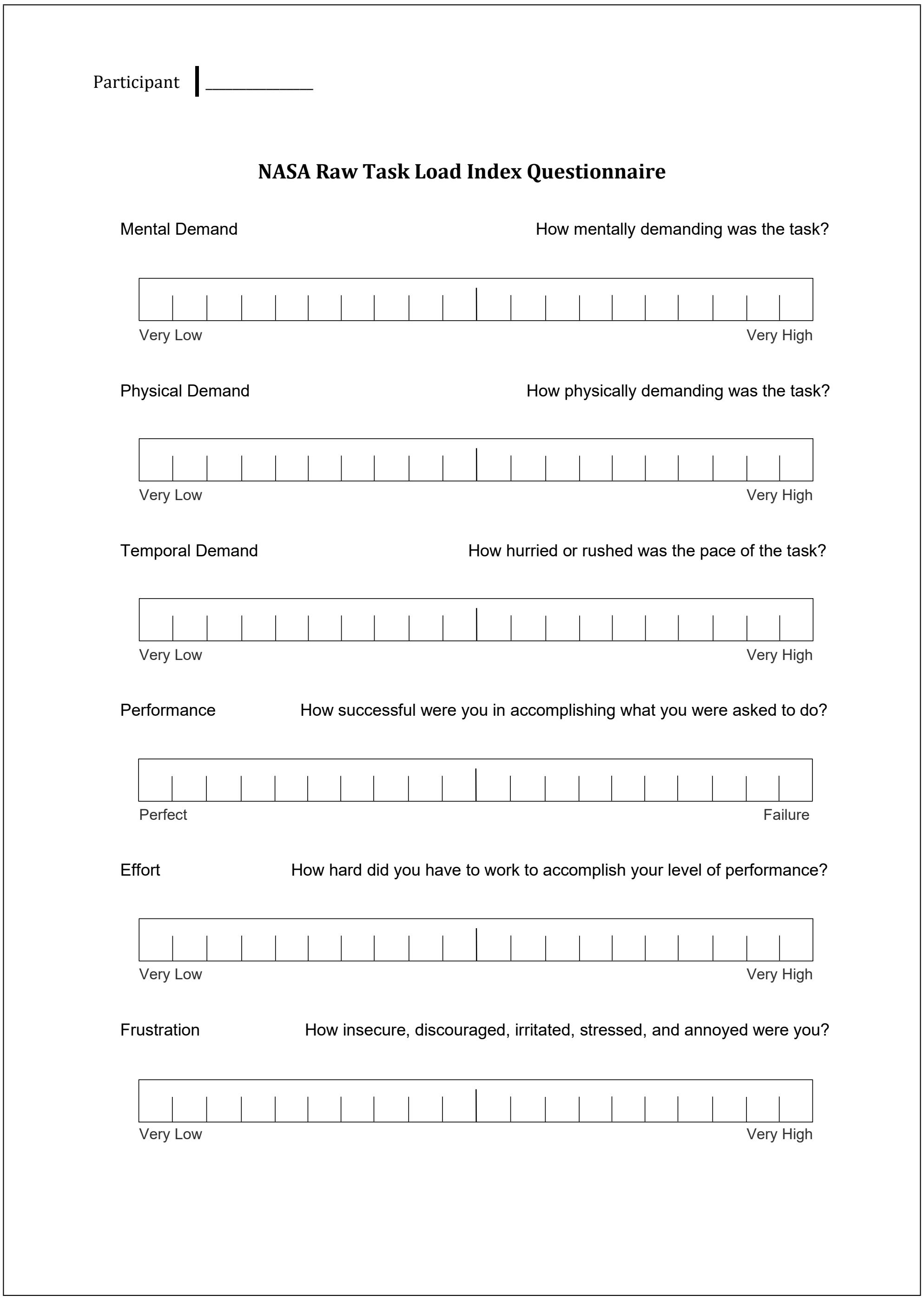
Strongly disagree ○ ○ ○ ○ ○ *Strongly agree*

10) I needed to learn a lot of things before I could get going with this system.

Strongly disagree ○ ○ ○ ○ ○ *Strongly agree*

Figure A.1: System Usability Scale Questionnaire [Bro96].

B NASA Raw Task Load Index Questionnaire



Participant ____________

NASA Raw Task Load Index Questionnaire

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low — Very High

Figure B.1: NASA Raw Task Load Index Questionnaire [Har06].

Figures

Tables

Bibliography

[AKW$^+$17] Abdelrahman, Yomna ; Knierim, Pascal ; Wozniak, Pawel W. ; Henze, Niels ; Schmidt, Albrecht: See Through the Fire: Evaluating the Augmentation of Visual Perception of Firefighters Using Depth and Thermal Cameras. In: *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*. New York, NY, USA : ACM, 2017 (UbiComp '17). – ISBN 978–1–4503–5190–4, 693–696

[ALD$^+$16] Ayers, John W. ; Leas, Eric C. ; Dredze, Mark ; Allem, Jon-Patrick ; Grabowski, Jurek G. ; Hill, Linda: Pokémon Go - A New Distraction for Drivers and Pedestrians. In: *JAMA Internal Medicine* 176 (2016), Nr. 12, 1865-1866. `http://dx.doi.org/10.1001/jamainternmed.2016.6274`. – DOI 10.1001/jamainternmed.2016.6274

[ALM$^+$12] Ardouin, Jérôme ; Lécuyer, Anatole ; Marchal, Maud ; Riant, Clément ; Marchand, Eric: FlyVIZ: A Novel Display Device to Provide Humans with 360° Vision by Coupling Catadioptric Camera with Hmd. In: *Proceedings of the 18th ACM Symposium on Virtual Reality Software and Technology*. New York, NY, USA : ACM, 2012 (VRST '12). – ISBN 978–1–4503–1469–5, 41–44

[ALMM14] Ardouin, Jérôme ; Lécuyer, Anatole ; Marchal, Maud ; Marchand, Eric: Stereoscopic rendering of virtual environments with wide Field-of-Views up to 360°. In: *2014 IEEE Virtual Reality (VR)*, 2014. – ISSN 2375–5334, S. 3–8

[AWM05] Ames, Shelly L. ; Wolffsohn, James S. ; Mcbrien, Neville A.: The development of a symptom questionnaire for assessing virtual reality viewing using a head-mounted display. In: *Optometry and Vision Science* 82 (2005), Nr. 3, S. 168–176. `http://dx.doi.org/10.1097/01.opx.0000156307.95086.6`. – DOI 10.1097/01.opx.0000156307.95086.6

[Azu97] Azuma, Ronald T.: A Survey of Augmented Reality. In: *Presence: Teleoperators and Virtual Environments* 6 (1997), Nr. 4, 355-385. `http://dx.doi.org/10.1162/pres.1997.6.4.355`. – DOI 10.1162/pres.1997.6.4.355

[BABB09] Barraco, Rosita ; Adorno, Dominique P. ; Bellomonte, Leonardo ; Brai, Maria: A study of the human rod and cone electroretinogram a-wave component. In: *Journal of Statistical Mechanics: Theory and Experiment* 2009 (2009), mar, Nr. 03, S. 1–

11. http://dx.doi.org/10.1088/1742-5468/2009/03/p03007. – DOI 10.1088/1742–5468/2009/03/p03007

[BC07] Burigat, Stefano ; Chittaro, Luca: Navigation in 3D virtual environments: Effects of user experience and location-pointing navigation aids. In: *International Journal of Human-Computer Studies* 65 (2007), Nr. 11, 945 - 958. http://dx.doi.org/https://doi.org/10.1016/j.ijhcs.2007.07.003. – DOI https://doi.org/10.1016/j.ijhcs.2007.07.003. – ISSN 1071–5819

[BC11] Burigat, Stefano ; Chittaro, Luca: Visualizing references to off-screen content on mobile devices: A comparison of Arrows, Wedge, and Overview + Detail. In: *Interacting with Computers* 23 (2011), 02, Nr. 2, 156-166. http://dx.doi.org/10.1016/j.intcom.2011.02.005. – DOI 10.1016/j.intcom.2011.02.005. – ISSN 0953–5438

[BCG06] Burigat, Stefano ; Chittaro, Luca ; Gabrielli, Silvia: Visualizing Locations of Off-screen Objects on Mobile Devices: A Comparative Evaluation of Three Approaches. In: *Proceedings of the 8th Conference on Human-computer Interaction with Mobile Devices and Services*. New York, NY, USA : ACM, 2006 (MobileHCI '06). – ISBN 1–59593–390–5, 239–246

[BCHB16] Borojeni, Shadan S. ; Chuang, Lewis ; Heuten, Wilko ; Boll, Susanne: Assisting Drivers with Ambient Take-Over Requests in Highly Automated Driving. In: *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. New York, NY, USA : ACM, 2016 (Automotive'UI 16). – ISBN 978–1–4503–4533–0, 237–244

[BCL+15] Billinghurst, Mark ; Clark, Adrian ; Lee, Gun u. a.: A survey of augmented reality. In: *Foundations and Trends in Human–Computer Interaction* 8 (2015), Nr. 2-3, 73–272. http://dx.doi.org/10.1561/1100000049

[Ben18] Benko, Hrvoje: *ISMAR 2018: Keynote*. https://www.ismar2018.org/keynote_benko/, 2018. – Last retrieved 2/9/2019

[Ber99] Bertelson, Paul: Ventriloquism: A case of crossmodal perceptual grouping. Version: 1999. http://dx.doi.org/10.1016/S0166-4115(99)80034-X. In: Aschersleben, Gisa (Hrsg.) ; Bachmann, Talis (Hrsg.) ; Müsseler, Jochen (Hrsg.): *Cognitive Contributions to the Perception of Spatial and Temporal Events* Bd. 129. North-Holland, 1999. – DOI 10.1016/S0166–4115(99)80034–X. – ISSN 0166–4115, Kapitel 14, 347 - 362

[BH94] Bederson, Benjamin B. ; Hollan, James D.: Pad++: A Zooming Graphical Interface for Exploring Alternate Interface Physics. In:

Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology. New York, NY, USA : ACM, 1994 (UIST '94). – ISBN 0–89791–657–3, 17–26

[BHF02] Bell, Blaine ; Höllerer, Tobias ; Feiner, Steven: An Annotated Situation-awareness Aid for Augmented Reality. In: *Proceedings of the 15th Annual ACM Symposium on User Interface Software and Technology.* New York, NY, USA : ACM, 2002 (UIST '02). – ISBN 1–58113–488–6, 213–216

[BJK05] Baek, Jin-uk ; Jung, Jaehoon ; Kim, Gerard J.: Head Mounted Display with Peripheral Vision. In: *Proceedings of the 2005 International Conference on Augmented Tele-existence.* New York, NY, USA : ACM, 2005 (ICAT '05). – ISBN 0–473–10657–4, 282–282

[BLH04] Baudisch, Patrick ; Lee, Bongshin ; Hanna, Libby: Fishnet, a Fisheye Web Browser with Search Term Popouts: A Comparative Evaluation with Overview and Linear View. In: *Proceedings of the Working Conference on Advanced Visual Interfaces.* New York, NY, USA : ACM, 2004 (AVI '04). – ISBN 1–58113–867–9, 133–140

[BMNN19] Bala, Paulo ; Masu, Raul ; Nisi, Valentina ; Nunes, Nuno: "When the Elephant Trumps": A Comparative Study on Spatial Audio for Orientation in 360º Videos. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* New York, NY, USA : ACM, 2019 (CHI '19). – ISBN 978–1–4503–5970–2, 695:1–695:13

[BR03] Baudisch, Patrick ; Rosenholtz, Ruth: Halo: A Technique for Visualizing Off-screen Objects. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* New York, NY, USA : ACM, 2003 (CHI '03). – ISBN 1–58113–630–7, 481–488

[BR06] Bimber, Oliver ; Raskar, Ramesh: Modern Approaches to Augmented Reality. In: *ACM SIGGRAPH 2006 Courses.* New York, NY, USA : Association for Computing Machinery, 2006 (SIGGRAPH '06). – ISBN 1–59593–364–6, 1–86

[BRFI95] Barfield, Woodrow ; Rosenberg, Craig ; Furness III, Thomass A.: Situation Awareness as a Function of Frame of Reference, Computer-Graphics Eyepoint Elevation, and Geometric Field of View. In: *The International Journal of Aviation Psychology* 5 (1995), Nr. 3, 233-256. `http://dx.doi.org/10.1207/s15327108ijap0503_1`. – DOI 10.1207/s15327108ijap0503_1. – PMID: 11541915

[Bro96] Brooke, John: SUS: A 'quick and dirty' usability scale. In: Jordan, Patrick W. (Hrsg.) ; Thomas, Bruce (Hrsg.) ; Weerdmeester, Bernhard A. (Hrsg.) ; McClelland, Ian L. (Hrsg.): *Usability Evaluation in Industry.* London, UK : CRC Press, 1996. – ISBN 978–0–7484–0460–5, Kapitel 21, S. 189–194

[BSEN18] Bork, Felix ; Schnelzer, Christian ; Eck, Ulrich ; Navab, Nassir: Towards Efficient Visual Guidance in Limited Field-of-View Head-Mounted Displays. In: *IEEE Transactions on Visualization and Computer Graphics* 24 (2018), Nov, Nr. 11, S. 2983–2992. `http://dx.doi.org/10.1109/TVCG.2018.2868584`. – DOI 10.1109/TVCG.2018.2868584. – ISSN 1077–2626

[BSP+93] Bier, Eric A. ; Stone, Maureen C. ; Pier, Ken ; Buxton, William ; DeRose, Tony D.: Toolglass and Magic Lenses: The See-through Interface. In: *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques.* New York, NY, USA : ACM, 1993 (SIGGRAPH '93). – ISBN 0–89791–601–8, 73–80

[BTG97] Bliss, James P. ; Tidwell, Philip D. ; Guest, Michael A.: The Effectiveness of Virtual Reality for Administering Spatial Navigation Training to Firefighters. In: *Presence: Teleoperators and Virtual Environments* 6 (1997), Nr. 1, 73-86. `http://dx.doi.org/10.1162/pres.1997.6.1.73`. – DOI 10.1162/pres.1997.6.1.73

[BTOX06] Biocca, Frank ; Tang, Arthur ; Owen, Charles ; Xiao, Fan: Attention Funnel: Omnidirectional 3D Cursor for Mobile Augmented Reality Platforms. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* New York, NY, USA : ACM, 2006 (CHI '06). – ISBN 1–59593–372–7, 1115–1122

[BWB+13] Bruder, Gerd ; Wieland, Phil ; Bolte, Benjamin ; Lappe, Markus ; Steinicke, Frank: Going with the flow: Modifying self-motion perception with computer-mediated optic flow. In: *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2013, S. 67–74

[BWI90] Beard, David V. ; Walker II, John Q.: Navigational techniques to improve the display of large two-dimensional spaces. In: *Behaviour & Information Technology* 9 (1990), Nr. 6, 451-466. `http://dx.doi.org/10.1080/01449299008924259`. – DOI 10.1080/01449299008924259

[CB04] Chittaro, Luca ; Burigat, Stefano: 3D Location-pointing As a Navigation Aid in Virtual Environments. In: *Proceedings of the Working Conference on Advanced Visual Interfaces.* New York, NY, USA : ACM, 2004 (AVI '04). – ISBN 1–58113–867–9, 267–274

[CCK$^+$89] Chaudhuri, Subhasis ; Chatterjee, Shankar ; Katz, Norman ; Nelson, Mark ; Goldbaum, Michael: Detection of blood vessels in retinal images using two-dimensional matched filters. In: *IEEE Transactions on Medical Imaging* 8 (1989), Sep., Nr. 3, S. 263–269. `http://dx.doi.org/10.1109/42.34715`. – DOI 10.1109/42.34715. – ISSN 0278–0062

[CHB17] Cobus, Vanessa ; Heuten, Wilko ; Boll, Susanne: Multimodal Head-mounted Display for Multimodal Alarms in Intensive Care Units. In: *Proceedings of the 6th ACM International Symposium on Pervasive Displays.* New York, NY, USA : ACM, 2017 (PerDis '17). – ISBN 978–1–4503–5045–7, 26:1–26:2

[CIP$^+$06] Costanza, Enrico ; Inverso, Samuel A. ; Pavlov, Elan ; Allen, Rebecca ; Maes, Pattie: Eye-q: Eyeglass Peripheral Display for Subtle Intimate Notifications. In: *Proceedings of the 8th Conference on Human-computer Interaction with Mobile Devices and Services.* New York, NY, USA : ACM, 2006 (MobileHCI '06). – ISBN 1–59593–390–5, 211–218

[CK00] Cullinane, Kevin ; Khanna, Mahim: Economies of scale in large containerships: optimal size and geographical implications. In: *Journal of Transport Geography* 8 (2000), Nr. 3, 181 - 195. `http://dx.doi.org/https://doi.org/10.1016/S0966-6923(00)00010-7`. – DOI https://doi.org/10.1016/S0966–6923(00)00010–7. – ISSN 0966–6923

[CKB09] Cockburn, Andy ; Karlson, Amy ; Bederson, Benjamin B.: A Review of Overview+Detail, Zooming, and Focus+Context Interfaces. In: *ACM Comput. Surv.* 41 (2009), Januar, Nr. 1, 2:1–2:31. `http://dx.doi.org/10.1145/1456650.1456652`. – DOI 10.1145/1456650.1456652. – ISSN 0360–0300

[CM92] Caudell, Thomas P. ; Mizell, David W.: Augmented reality: an application of heads-up display technology to manual manufacturing processes. In: *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences* Bd. ii, 1992, S. 659–669 vol.2

[CN08] Chittaro, Luca ; Nadalutti, Daniele: Presenting Evacuation Instructions on Mobile Devices by Means of Location-aware 3D Virtual Environments. In: *Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services.* New York, NY, USA : ACM, 2008 (MobileHCI '08). – ISBN 978–1–59593–952–4, 395–398

[CNSD$^+$92] Cruz-Neira, Carolina ; Sandin, Daniel J. ; DeFanti, Thomas A. ; Kenyon, Robert V. ; Hart, John C.: The CAVE: Audio Visual

Experience Automatic Virtual Environment. In: *Commun. ACM* 35 (1992), Juni, Nr. 6, 64–72. `http://dx.doi.org/10.1145/129888.129892`. – DOI 10.1145/129888.129892. – ISSN 0001–0782

[Coa92] Coates, George: Program from Invisible Site-a virtual sho, a multimedia performance work presented by George Coates Performance Works. In: *San Francisco, CA* (1992)

[CR06] Cakmakci, Ozan ; Rolland, Jannick: Head-worn displays: a review. In: *Journal of Display Technology* 2 (2006), Sep., Nr. 3, S. 199–216. `http://dx.doi.org/10.1109/JDT.2006.879846`. – DOI 10.1109/JDT.2006.879846. – ISSN 1551–319X

[DDM09] Dunleavy, Matt ; Dede, Chris ; Mitchell, Rebecca: Affordances and Limitations of Immersive Participatory Augmented Reality Simulations for Teaching and Learning. In: *Journal of Science Education and Technology* 18 (2009), Feb, Nr. 1, 7–22. `http://dx.doi.org/10.1007/s10956-008-9119-1`. – DOI 10.1007/s10956–008–9119–1. – ISSN 1573–1839

[DGD17] Danieau, Fabien ; Guillo, Antoine ; Doré, Renaud: Attention guidance for immersive video content in head-mounted displays. In: *2017 IEEE Virtual Reality (VR)*, 2017. – ISSN 2375–5334, S. 205–206

[DGYG$^+$04] De Guzman, Edward S. ; Yau, Margaret ; Gagliano, Anthony ; Park, Austin ; Dey, Anind K.: Exploring the Design and Use of Peripheral Displays of Awareness Information. In: *CHI '04 Extended Abstracts on Human Factors in Computing Systems*. New York, NY, USA : ACM, 2004 (CHI EA '04). – ISBN 1–58113–703–6, 1247–1250

[DS00] Driver, Jon ; Spence, Charles: Multisensory perception: Beyond modularity and convergence. In: *Current Biology* 10 (2000), Nr. 20, R731 - R735. `http://dx.doi.org/10.1016/S0960-9822(00)00740-5`. – DOI 10.1016/S0960–9822(00)00740–5. – ISSN 0960–9822

[EACI11] Ens, Barrett ; Ahlström, David ; Cockburn, Andy ; Irani, Pourang: Characterizing User Performance with Assisted Direct Off-screen Pointing. In: *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*. New York, NY, USA : ACM, 2011 (MobileHCI '11). – ISBN 978–1–4503–0541–9, 485–494

[EBM$^+$97] Ellis, Stephen R. ; Breant, Francois ; Menges, Brian ; Jacoby, Richard ; Adelstein, Bernard D.: Factors influencing operator interaction with virtual objects viewed via head-mounted see-through displays: viewing conditions and rendering latency. In: *Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality*, 1997. – ISSN 1087–8270, S. 138–145

[ELH19] Engelbrecht, Hendrik ; Lindeman, Robert W. ; Hoermann, Simon: A SWOT Analysis of the Field of Virtual Reality for Firefighter Training. In: *Frontiers in Robotics and AI* 6 (2019), 101. `http://dx.doi.org/10.3389/frobt.2019.00101`. – DOI 10.3389/frobt.2019.00101. – ISSN 2296–9144

[End95] Endsley, Mica R.: Toward a Theory of Situation Awareness in Dynamic Systems. In: *Human Factors* 37 (1995), Nr. 1, 32-64. `http://dx.doi.org/10.1518/001872095779049543`. – DOI 10.1518/001872095779049543

[Eys17] Eysenbach, Gunther: Pedestrian Inattention Blindness While Playing Pokémon Go as an Emerging Health-Risk Behavior: A Case Report. In: *J Med Internet Res* 19 (2017), April, Nr. 4, S. 1865–1866. `http://dx.doi.org/10.2196/jmir.6596`. – DOI 10.2196/jmir.6596

[EZ93] Elder, James ; Zucker, Steven: The effect of contour closure on the rapid discrimination of two-dimensional shapes. In: *Vision Research* 33 (1993), Nr. 7, 981 - 991. `http://dx.doi.org/https://doi.org/10.1016/0042-6989(93)90080-G`. – DOI https://doi.org/10.1016/0042–6989(93)90080–G. – ISSN 0042–6989

[FHNI14] Fan, Kevin ; Huber, Jochen ; Nanayakkara, Suranga ; Inami, Masahiko: SpiderVision: Extending the Human Field of View for Augmented Awareness. In: *Proceedings of the 5th Augmented Human International Conference*. New York, NY, USA : ACM, 2014 (AH '14). – ISBN 978–1–4503–2761–9, 49:1–49:8

[FM14] Fedosov, Anton ; Misslinger, Stefan: Location Based Experience Design for Mobile Augmented Reality. In: *Proceedings of the 2014 ACM SIGCHI Symposium on Engineering Interactive Computing Systems*. New York, NY, USA : ACM, 2014 (EICS '14). – ISBN 978–1–4503–2725–1, 185–188

[FMH15] Farra, Sharon L. ; Miller, Elaine T. ; Hodgson, Eric: Virtual reality disaster training: Translation to practice. In: *Nurse Education in Practice* 15 (2015), Nr. 1, 53 - 57. `http://dx.doi.org/https://doi.org/10.1016/j.nepr.2013.08.017`. – DOI https://doi.org/10.1016/j.nepr.2013.08.017. – ISSN 1471–5953. – Special Section: Disaster Management Nursing Education

[FMHS93] Feiner, Steven ; MacIntyre, Blair ; Haupt, Marcus ; Solomon, Eliot: Windows on the World: 2D Windows for 3D Augmented

Reality. In: *Proceedings of the 6th Annual ACM Symposium on User Interface Software and Technology.* New York, NY, USA : ACM, 1993 (UIST '93). – ISBN 0–89791–628–X, 145–155

[FMHW97] Feiner, Steven ; MacIntyre, Blair ; Höllerer, Tobias ; Webster, Anthony: A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. In: *Personal Technologies* 1 (1997), Dec, Nr. 4, 208–217. `http://dx.doi.org/10.1007/BF01682023`. – DOI 10.1007/BF01682023. – ISSN 1617–4917

[FMS93] Feiner, Steven ; Macintyre, Blair ; Seligmann, Dorée: Knowledge-based augmented reality. In: *Communications of the ACM* 36 (1993), Nr. 7, S. 53–62

[Fri08] Frittelli, John: Ship navigation in harbors: Safety issues. In: *CRS Report for the Congress* Congressional Research Service, the Library of Congress, 2008, S. 20

[FS06] Fields, Alexa W. ; Shelton, Amy L.: Individual skill differences and large-scale environmental learning. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 32 (2006), Nr. 3, S. 506–515. `http://dx.doi.org/10.1037/0278-7393.32.3.506`. – DOI 10.1037/0278–7393.32.3.506. – ISSN 1939–1285(Electronic),0278–7393(Print)

[GABH18] Gruenefeld, Uwe ; Ali, Abdallah E. ; Boll, Susanne ; Heuten, Wilko: Beyond Halo and Wedge: Visualizing Out-of-view Objects on Head-mounted Virtual and Augmented Reality Devices. In: *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services.* New York, NY, USA : ACM, 2018 (MobileHCI '18). – ISBN 978–1–4503–5898–9, 40:1–40:11

[GAHB17] Gruenefeld, Uwe ; Ali, Abdallah E. ; Heuten, Wilko ; Boll, Susanne: Visualizing Out-of-view Objects in Head-mounted Augmented Reality. In: *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services.* New York, NY, USA : ACM, 2017 (MobileHCI '17). – ISBN 978–1–4503–5075–4, 81:1–81:7

[GB08] Groen, Eric L. ; Bos, Jelte E.: Simulator Sickness Depends on Frequency of the Simulator Motion Mismatch: An Observation. In: *Presence: Teleoperators and Virtual Environments* 17 (2008), Nr. 6, 584-593. `http://dx.doi.org/10.1162/pres.17.6.584`. – DOI 10.1162/pres.17.6.584

[GBGI08] Gustafson, Sean ; Baudisch, Patrick ; Gutwin, Carl ; Irani, Pourang: Wedge: Clutter-free Visualization of Off-screen Locations. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA : ACM, 2008 (CHI '08). – ISBN 978–1–60558–011–1, 787–796

[GBH18] Gruenefeld, Uwe ; Bargen, Rieke von ; Heuten, Wilko: Identification of Out-of-View Objects in Virtual Reality. In: *Proceedings of the Symposium on Spatial User Interaction*. New York, NY, USA : ACM, 2018 (SUI '18). – ISBN 978–1–4503–5708–1, 182–182

[GEA+17] Gruenefeld, Uwe ; Ennenga, Dag ; Ali, Abdallah E. ; Heuten, Wilko ; Boll, Susanne: EyeSee360: Designing a Visualization Technique for Out-of-view Objects in Head-mounted Augmented Reality. In: *Proceedings of the 5th Symposium on Spatial User Interaction*. New York, NY, USA : ACM, 2017 (SUI '17). – ISBN 978–1–4503–5486–8, 109–118

[GHH18] Gruenefeld, Uwe ; Hsiao, Dana ; Heuten, Wilko: EyeSeeX: Visualization of Out-of-View Objects on Small Field-of-View Augmented and Virtual Reality Devices. In: *Proceedings of the 7th ACM International Symposium on Pervasive Displays*. New York, NY, USA : ACM, 2018 (PerDis '18). – ISBN 978–1–4503–5765–4, 26:1–26:2

[GHHB17] Gruenefeld, Uwe ; Hsiao, Dana ; Heuten, Wilko ; Boll, Susanne: EyeSee: Beyond Reality with Microsoft HoloLens. In: *Proceedings of the 5th Symposium on Spatial User Interaction*. New York, NY, USA : ACM, 2017 (SUI '17). – ISBN 978–1–4503–5486–8, 148–148

[GI07] Gustafson, Sean G. ; Irani, Pourang P.: Comparing Visualizations for Tracking Off-screen Moving Targets. In: *CHI '07 Extended Abstracts on Human Factors in Computing Systems*. New York, NY, USA : ACM, 2007 (CHI EA '07). – ISBN 978–1–59593–642–4, 2399–2404

[GKL+19] Gruenefeld, Uwe ; Koethe, Ilja ; Lange, Daniel ; Weiss, Sebastian ; Heuten, Wilko: Comparing Techniques for Visualizing Moving Out-of-View Objects in Head-mounted Virtual Reality. In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2019. – ISSN 2642–5246, S. 742–746

[GLB+18] Gruenefeld, Uwe ; Löcken, Andreas ; Brueck, Yvonne ; Boll, Susanne ; Heuten, Wilko: Where to Look: Exploring Peripheral Cues for Shifting Attention to Spatially Distributed Out-of-View Objects. In: *Proceedings of the 10th International Conference on*

Automotive User Interfaces and Interactive Vehicular Applications. New York, NY, USA : ACM, 2018 (AutomotiveUI '18). – ISBN 978–1–4503–5946–7, 221–228

[GLH+18] Gruenefeld, Uwe ; Lange, Daniel ; Hammer, Lasse ; Boll, Susanne ; Heuten, Wilko: FlyingARrow: Pointing Towards Out-of-View Objects on Augmented Reality Devices. In: *Proceedings of the 7th ACM International Symposium on Pervasive Displays.* New York, NY, USA : ACM, 2018 (PerDis '18). – ISBN 978–1–4503–5765–4, 20:1–20:6

[Gol89] Goldstein, Bruce E.: *Sensation and perception, 3rd ed.* Belmont, CA, US : Wadsworth/Thomson Learning, 1989 (Sensation and perception, 3rd ed.). – ISBN 0–534–09672–7 (Hardcover)

[GPH19a] Gruenefeld, Uwe ; Prädel, Lars ; Heuten, Wilko: Improving Search Time Performance for Locating Out-of-View Objects in Augmented Reality. In: *Proceedings of Mensch Und Computer 2019.* New York, NY, USA : ACM, 2019 (MuC'19). – ISBN 978–1–4503–7198–8, 481–485

[GPH19b] Gruenefeld, Uwe ; Prädel, Lars ; Heuten, Wilko: Locating Nearby Physical Objects in Augmented Reality. In: *International Conference on Mobile and Ubiquitous Multimedia.* New York, NY, USA : ACM, 2019 (MUM '19). – ISBN 978–1–4503–7624–2, 1:1-1:10

[Gre92] Greenbaum, Paul: The lawnmower man. In: *Film and video* 9 (1992), Nr. 3, S. 58–62

[GSA+18] Gruenefeld, Uwe ; Stratmann, Tim C. ; Ali, Abdallah E. ; Boll, Susanne ; Heuten, Wilko: RadialLight: Exploring Radial Peripheral LEDs for Directional Cues in Head-mounted Displays. In: *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services.* New York, NY, USA : ACM, 2018 (MobileHCI '18). – ISBN 978–1–4503–5898–9, 39:1–39:6

[GSB+18] Gruenefeld, Uwe ; Stratmann, Tim C. ; Brueck, Yvonne ; Hahn, Axel ; Boll, Susanne ; Heuten, Wilko: Investigations on Container Ship Berthing from the Pilot's Perspective: Accident Analysis, Ethnographic Study, and Online Survey. In: *TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation* 12 (2018), September, Nr. 3, 492-498. `http://dx.doi.org/10.12716/1001.12.03.07`. – DOI 10.12716/1001.12.03.07

[GSHB17] Gruenefeld, Uwe ; Stratmann, Tim C. ; Heuten, Wilko ; Boll, Susanne: PeriMR: A Prototyping Tool for Head-mounted

Peripheral Light Displays in Mixed Reality. In: *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services.* New York, NY, USA : ACM, 2017 (MobileHCI '17). – ISBN 978–1–4503–5075–4, 51:1–51:6

[GSJ+18] GRUENEFELD, Uwe ; STRATMANN, Tim C. ; JUNG, Jinki ; LEE, Hyeopwoo ; CHOI, Jeehye ; NANDA, Abhilasha ; HEUTEN, Wilko: Guiding Smombies: Augmenting Peripheral Vision with Low-Cost Glasses to Shift the Attention of Smartphone Users. In: *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 2018, S. 127–131

[GSPH18] GRUENEFELD, Uwe ; STRATMANN, Tim C. ; PRÄDEL, Lars ; HEUTEN, Wilko: MonoculAR: A Radial Light Display to Point Towards Out-of-view Objects on Augmented Reality Devices. In: *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct.* New York, NY, USA : ACM, 2018 (MobileHCI '18). – ISBN 978–1–4503–5941–2, 16–22

[Haf13] HAFENLOTSENBRÜDERSCHAFT: *Hafenlotsenordnung.* `http://www.hamburg-pilot.de/Hafenlotsordnung_2013_Stand_07052013.pdf`, 2013. – Letzter Zugriff am 07. November 2019. §5 Absatz 1 Nummer 1 HmbGVBl (Hamburgisches Gesetz- und Verordnungsblatt)

[Har00] HARDIN, Larry: Red and yellow, green and blue, warm and cool: explaining colour appearance. In: *Journal of Consciousness Studies* 7 (2000), Nr. 8-9, 113-122. `http://www.ingentaconnect.com/content/imp/jcs/2000/00000007/F0020008/1046`. – ISSN 1355–8250

[Har06] HART, Sandra G.: Nasa-Task Load Index (NASA-TLX); 20 Years Later. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50 (2006), Nr. 9, 904-908. `http://dx.doi.org/10.1177/154193120605000909`. – DOI 10.1177/154193120605000909

[Haz83] *Kapitel* 1. In: HAZEN, Nancy L.: *Spatial Orientation: A Comparative Approach.* Boston, MA : Springer US, 1983. – ISBN 978–1–4615–9325–6, 3–37

[HB10] HENZE, Niels ; BOLL, Susanne: Evaluation of an Off-screen Visualization for Magic Lens and Dynamic Peephole Interfaces. In: *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services.* New York, NY, USA : ACM, 2010 (MobileHCI '10). – ISBN 978–1–60558–835–3, 191–194

[HF01] Hornbæk, Kasper ; Frøkjær, Erik: Reading of Electronic Documents: The Usability of Linear, Fisheye, and Overview+Detail Interfaces. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA : ACM, 2001 (CHI '01). – ISBN 1–58113–327–8, 293–300

[HFM+05] Hafting, Torkel ; Fyhn, Marianne ; Molden, Sturla ; Moser, May-Britt ; Moser, Edvard I.: Microstructure of a spatial map in the entorhinal cortex. In: *Nature* 436 (2005), August, Nr. 7052, 801–806. `http://dx.doi.org/10.1038/nature03721`. – DOI 10.1038/nature03721. – ISSN 1476–4687

[HJDB16] Heller, Florian ; Jevanesan, Jayan ; Dietrich, Pascal ; Borchers, Jan: Where Are We?: Evaluating the Current Rendering Fidelity of Mobile Audio Augmented Reality Systems. In: *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. New York, NY, USA : ACM, 2016 (MobileHCI '16). – ISBN 978–1–4503–4408–1, 278–282

[HKBL07] Henderson, Amy ; Korner-Bitensky, Nicol ; Levin, Mindy: Virtual Reality in Stroke Rehabilitation: A Systematic Review of its Effectiveness for Upper Limb Motor Recovery. In: *Topics in Stroke Rehabilitation* 14 (2007), Nr. 2, 52-61. `http://dx.doi.org/10.1310/tsr1402-52`. – DOI 10.1310/tsr1402–52. – PMID: 17517575

[HLSH09] Harrison, Chris ; Lim, Brian Y. ; Shick, Aubrey ; Hudson, Scott E.: Where to Locate Wearable Displays?: Reaction Time Performance of Visual Alerts from Tip to Toe. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA : ACM, 2009 (CHI '09). – ISBN 978–1–60558–246–7, 941–944

[HM86] Hirtle, Stephen C. ; Mascolo, Michael F.: Effect of semantic clustering on the memory of spatial locations. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 12 (1986), Nr. 2, S. 182–189. `http://dx.doi.org/10.1037/0278-7393.12.2.182`. – DOI 10.1037/0278–7393.12.2.182. – ISSN 1939–1285(Electronic),0278–7393(Print)

[How93] *Kapitel* 22. In: Howard, Ian P.: *Spatial Vision within Egocentric and Exocentric Frames of Reference*. USA : Taylor & Francis, Inc., 1993. – ISBN 0748400826, S. 338–358

[HP18] Harborth, David ; Pape, Sebastian: Exploring the Hype: Investigating Technology Acceptance Factors of Pokémon Go. In: *2017*

IEEE International Symposium on Mixed and Augmented Reality (ISMAR) Bd. 00, 2018, 155-168

[HPB10] Henze, Niels ; Poppinga, Benjamin ; Boll, Susanne: Experiments in the Wild: Public Evaluation of Off-screen Visualizations in the Android Market. In: *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries.* New York, NY, USA : ACM, 2010 (NordiCHI '10). – ISBN 978–1–60558–934–3, 675–678

[HS88] Hart, Sandra G. ; Staveland, Lowell E.: Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In: *Advances in Psychology* 52 (1988), 139 - 183. `http://dx.doi.org/http://dx.doi.org/10.1016/S0166-4115(08)62386-9`. – DOI http://dx.doi.org/10.1016/S0166–4115(08)62386–9. – ISSN 0166–4115

[HS08] Ho, Cristy ; Spence, Charles: *The multisensory driver: Implications for ergonomic car interface design.* CRC Press, 2008. – 1–143 S. – ISBN 978–1–3155–5542–3

[HSHS05] Hughes, Charles E. ; Stapleton, Christopher B. ; Hughes, Darin E. ; Smith, Eileen M.: Mixed reality in education, entertainment, and training. In: *IEEE Computer Graphics and Applications* 25 (2005), Nov, Nr. 6, S. 24–30. `http://dx.doi.org/10.1109/MCG.2005.139`. – DOI 10.1109/MCG.2005.139. – ISSN 0272–1716

[HT66] Howard, Ian P. ; Templeton, William B.: *Human spatial orientation.* Oxford, England : John Wiley & Sons, 1966 (Human spatial orientation.)

[HTP07] Huang, Liqiang ; Treisman, Anne ; Pashler, Harold: Characterizing the limits of human visual awareness. In: *Science* 317 (2007), Nr. 5839, S. 823–825

[IGY06] Irani, Pourang ; Gutwin, Carl ; Yang, Xing D.: Improving selection of off-screen targets with hopping. In: *Proceedings of the SIGCHI conference on Human Factors in computing systems* ACM, 2006, S. 299–308

[ISO10] ISO: *Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems.* `https://www.iso.org/standard/52075.html`. Version: March 2010. – ISO 9241-210:2010

[JFR17] Jäckle, Dominik ; Fuchs, Johannes ; Reiterer, Harald: Topology-Preserving Off-screen Visualization: Effects of Projection Strategy and Intrusion Adaption. In: *arXiv e-prints* (2017), Juni

[JHPR11] Jo, Hyungeun ; Hwang, Sungjae ; Park, Hyunwoo ; Ryu, Jung hee: Aroundplot: Focus+context interface for off-screen objects in 3D environments. In: *Computers & Graphics* 35 (2011), Nr. 4, 841 - 853. `http://dx.doi.org/https://doi.org/10.1016/j.cag.2011.04.005`. – DOI https://doi.org/10.1016/j.cag.2011.04.005. – ISSN 0097–8493. – Semantic 3D Media and Content

[JLC$^+$18] Jung, Jinki ; Lee, Hyeopwoo ; Choi, Jeehye ; Nanda, Abhilasha ; Gruenefeld, Uwe ; Stratmann, Tim C. ; Heuten, Wilko: Ensuring Safety in Augmented Reality from Trade-off Between Immersion and Situation Awareness. In: *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2018. – ISSN 1554–7868, S. 70–79

[Joh95] Johnson, Jeff A.: A Comparison of User Interfaces for Panning on a Touch-controlled Display. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* New York, NY, USA : ACM Press/Addison-Wesley Publishing Co., 1995 (CHI '95). – ISBN 0–201–84705–1, 218–225

[JPD$^+$18] John, Nigel W. ; Pop, Serban R. ; Day, Thomas W. ; Ritsos, Panagiotis D. ; Headleand, Christopher J.: The Implementation and Validation of a Virtual Environment for Training Powered Wheelchair Manoeuvres. In: *IEEE Transactions on Visualization and Computer Graphics* 24 (2018), May, Nr. 5, S. 1867–1878. `http://dx.doi.org/10.1109/TVCG.2017.2700273`. – DOI 10.1109/TVCG.2017.2700273. – ISSN 1077–2626

[K$^+$14] Klaene, Bernard J. u. a.: *Structural firefighting.* Jones & Bartlett Publishers, 2014. – ISBN 978–1–4496–4239–6

[Kal01] Kalat, James W.: *Biological psychology, 7th ed.* Belmont, CA, US : Wadsworth/Thomson Learning, 2001 (Biological psychology, 7th ed.). – ISBN 0–534–51400–6 (Hardcover)

[KAPD18] Kaya, Nazli E. ; Ayas, Suzan ; Ponnambalam, Canmanie T. ; Donmez, Birsen: Visual attention failures during turns at intersections: An on-road study. In: *Victoria* 1 (2018)

[Kit94] Kitchin, Robert M.: Cognitive maps: What are they and why study them? In: *Journal of Environmental Psychology* 14 (1994), Nr. 1, 1 - 19. `http://dx.doi.org/https://doi.org/10.1016/S0272-4944(05)80194-X`. – DOI https://doi.org/10.1016/S0272–4944(05)80194–X. – ISSN 0272–4944

[KKP16] Kerdegari, Hamideh ; Kim, Yeongmi ; Prescott, Tony J.: Head-Mounted Sensory Augmentation Device: Comparing Haptic and Audio Modality. In: Lepora, Nathan F. (Hrsg.) ; Mura, Anna (Hrsg.) ; Mangan, Michael (Hrsg.) ; Verschure, Paul F. (Hrsg.) ; Desmulliez, Marc (Hrsg.) ; Prescott, Tony J. (Hrsg.): *Biomimetic and Biohybrid Systems*. Cham : Springer International Publishing, 2016, S. 107–118

[KLBL93] Kennedy, Robert S. ; Lane, Norman E. ; Berbaum, Kevin S. ; Lilienthal, Michael G.: Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. In: *The International Journal of Aviation Psychology* 3 (1993), Nr. 3, 203-220. `http://dx.doi.org/10.1207/s15327108ijap0303_3`. – DOI 10.1207/s15327108ijap0303_3

[KM02] Kost, James T. ; McDermott, Michael P.: Combining dependent P-values. In: *Statistics & Probability Letters* 60 (2002), Nr. 2, S. 183 – 190. – ISSN 0167–7152

[KM06] Kooi, Frank L. ; Mosch, Marcel: Peripheral Motion Displays: Tapping the Potential of the Visual Periphery. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50 (2006), Nr. 16, 1604-1608. `http://dx.doi.org/10.1177/154193120605001619`. – DOI 10.1177/154193120605001619

[KND14] Knierim, James J. ; Neunuebel, Joshua P. ; Deshmukh, Sachin S.: Functional correlates of the lateral and medial entorhinal cortex: objects, path integration and local–global reference frames. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 369 (2014), Nr. 1635, 20130369. `http://dx.doi.org/10.1098/rstb.2013.0369`. – DOI 10.1098/rstb.2013.0369

[KR17] Kaul, Oliver B. ; Rohs, Michael: HapticHead: A Spherical Vibrotactile Grid Around the Head for 3D Guidance in Virtual and Augmented Reality. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA : ACM, 2017 (CHI '17). – ISBN 978–1–4503–4655–9, 3729–3740

[KS02] Kaufmann, Hannes ; Schmalstieg, Dieter: Mathematics and Geometry Education with Collaborative Augmented Reality. In: *ACM SIGGRAPH 2002 Conference Abstracts and Applications*. New York, NY, USA : ACM, 2002 (SIGGRAPH '02). – ISBN 1–58113–525–4, 37–41

[KSF10] Kruijff, Ernst ; Swan, J E. ; Feiner, Steven: Perceptual issues in augmented reality revisited. In: *Mixed and Augmented Reality*

(ISMAR), 2010 9th IEEE International Symposium on IEEE, 2010, S. 3–12

[Kuh17] Kuhnimhof, Tobias: *Car ownership and usage trends in Germany–Response to the Commission on Travel Demand's Call for Evidence: Understanding Travel Demand.* 2017

[LBB14] Lauber, Felix ; Böttcher, Claudius ; Butz, Andreas: PapAR: Paper Prototyping for Augmented Reality. In: *Adjunct Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications.* New York, NY, USA : ACM, 2014 (AutomotiveUI '14). – ISBN 978–1–4503–0725–3, 1–6

[LBBH10] Lee, Cha ; Bonebrake, Scott ; Bowman, Doug A. ; Höllerer, Tobias: The role of latency in the validity of AR simulation. In: *2010 IEEE Virtual Reality Conference (VR)*, 2010. – ISSN 1087–8270, S. 11–18

[LBJ+09] Lever, Colin ; Burton, Stephen ; Jeewajee, Ali ; O'Keefe, John ; Burgess, Neil: Boundary vector cells in the subiculum of the hippocampal formation. In: *The Journal of Neuroscience* 29 (2009), Nr. 31, S. 9771–9777. `http://dx.doi.org/10.1523/JNEUROSCI.1319-09.2009`. – DOI 10.1523/JNEUROSCI.1319–09.2009. – ISSN 1529–2401(Electronic), 0270–6474(Print)

[LBS+16] Lincoln, Peter ; Blate, Alex ; Singh, Montek ; Whitted, Turner ; State, Andrei ; Lastra, Anselmo ; Fuchs, Henry: From Motion to Photons in 80 Microseconds: Towards Minimal Latency for Virtual and Augmented Reality. In: *IEEE Transactions on Visualization and Computer Graphics* 22 (2016), April, Nr. 4, S. 1367–1376. `http://dx.doi.org/10.1109/TVCG.2016.2518038`. – DOI 10.1109/TVCG.2016.2518038. – ISSN 1077–2626

[LBS+17] Löcken, Andreas ; Blum, Sarah ; Stratmann, Tim C. ; Gruenefeld, Uwe ; Heuten, Wilko ; Boll, Susanne ; Par, Steven van d.: Effects of Location and Fade-in Time of (Audio-)Visual Cues on Response Times and Success-rates in a Dual-task Experiment. In: *Proceedings of the ACM Symposium on Applied Perception.* New York, NY, USA : ACM, 2017 (SAP '17). – ISBN 978–1–4503–5148–5, 12:1–12:4

[LCH+17] Lin, Yen-Chen ; Chang, Yung-Ju ; Hu, Hou-Ning ; Cheng, Hsien-Tzu ; Huang, Chi-Wen ; Sun, Min: Tell Me Where to Look: Investigating Ways for Assisting Focus in 360 Video. In: *Proc. CHI '17.* New York, NY, USA : ACM, 2017. – ISBN 978–1–4503–4655–9, 2535–2545

[LCRR18] Loreto, Cédric D. ; Chardonnet, Jean-Rémy ; Ryard, Julien ; Rousseau, Alain: WoaH: A Virtual Reality Work-at-Height Simulator. In: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2018, S. 281–288

[LDRR$^+$16] Luyten, Kris ; Degraen, Donald ; Rovelo Ruiz, Gustavo ; Coppers, Sven ; Vanacken, Davy: Hidden in Plain Sight: An Exploration of a Visual Language for Near-Eye Out-of-Focus Displays in the Peripheral View. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA : ACM, 2016 (CHI '16). – ISBN 978–1–4503–3362–7, 487–497

[LH11] Limanowski, Jakub ; Hecht, Heiko: Where do we stand on locating the self? In: *Psychology* 2 (2011), Nr. 04, S. 312

[LHB15] Löcken, Andreas ; Heuten, Wilko ; Boll, Susanne: Supporting Lane Change Decisions with Ambient Light. In: *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. New York, NY, USA : ACM, 2015 (AutomotiveUI '15). – ISBN 978–1–4503–3736–6, 204–211

[LKKP19] Liang, Feng ; Kevin, Stevanus ; Kunze, Kai ; Pai, Yun S.: PanoFlex: Adaptive Panoramic Vision to Accommodate 360° Field-of-View for Humans. In: *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology*. New York, NY, USA : ACM, 2019 (VRST '19), 1121-1122

[LKN$^+$16] Lee, Ji-Yong ; Kwon, Joung-Huem ; Nam, Sang-Hun ; Lee, Joong-Jae ; You, Bum-Jae: Coexistent Space: Collaborative Interaction in Shared 3D Space. In: *Proceedings of the 2016 Symposium on Spatial User Interaction*. New York, NY, USA : ACM, 2016 (SUI '16). – ISBN 978–1–4503–4068–7, 175–175

[LL09] Larson, Adam M. ; Loschky, Lester C.: The contributions of central versus peripheral vision to scene gist recognition. In: *Journal of Vision* 9 (2009), Nr. 10, 6. `http://dx.doi.org/10.1167/9.10.6`. – DOI 10.1167/9.10.6

[LV14] Lucero, Andrés ; Vetek, Akos: NotifEye: Using Interactive Glasses to Deal with Notifications While Walking in Public. In: *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology*. New York, NY, USA : ACM, 2014 (ACE '14). – ISBN 978–1–4503–2945–3, 17:1–17:10

[Lyo16] Lyons, Kent: Visual Parameters Impacting Reaction Times on Smartwatches. In: *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*.

New York, NY, USA : ACM, 2016 (MobileHCI '16). – ISBN 978–1–4503–4408–1, 190–194

[MB83] Moar, Ian ; Bower, Gordon H.: Inconsistency in spatial knowledge. In: *Memory & Cognition* 11 (1983), Mar, Nr. 2, 107–113. `http://dx.doi.org/10.3758/BF03213464`. – DOI 10.3758/BF03213464. – ISSN 1532–5946

[MF13] Maimone, Andrew ; Fuchs, Henry: Computational augmented reality eyeglasses. In: *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2013, S. 29–38

[Mic91] Michotte, Albert, and Thines, Georges, and Butterworth, George and Costall, Alan: *Michotte's experimental phenomenology of perception / edited by Georges Thines, Alan Costall, George Butterworth*. L. Erlbaum Associates Hillsdale, N.J, 1991. – xi, 258 p. : S. – ISBN 0898596068

[MJUO17] Matsuzoe, Shizuko ; Jiang, Shan ; Ueki, Miwa ; Okabayashi, Keiju: Intuitive Visualization Method for Locating Off-screen Objects Inspired by Motion Perception in Peripheral Vision. In: *Proceedings of the 8th Augmented Human International Conference*. New York, NY, USA : ACM, 2017 (AH '17). – ISBN 978–1–4503–4835–5, 29:1–29:4

[MK94] Milgram, Paul ; Kishino, Fumio: A taxonomy of mixed reality visual displays. In: *IEICE TRANSACTIONS on Information and Systems* 77 (1994), Nr. 12, S. 1321–1329

[MKS+05] Murray, Ian J. ; Kulikowski, Janus J. ; Stanikunas, Rytis ; Vaitkevičius, Henrikas ; Daugirdiene, Avsra: Colour Matching of Isoluminant Samples and Backgrounds: A Dimming Effect. In: *Perception* 34 (2005), Nr. 8, 927-932. `http://dx.doi.org/10.1068/p5301`. – DOI 10.1068/p5301. – PMID: 16178145

[MLR+14] Maimone, Andrew ; Lanman, Douglas ; Rathinavel, Kishore ; Keller, Kurtis ; Luebke, David ; Fuchs, Henry: Pinlight Displays: Wide Field of View Augmented Reality Eyeglasses Using Defocused Point Light Sources. In: *ACM SIGGRAPH 2014 Emerging Technologies*. New York, NY, USA : ACM, 2014 (SIGGRAPH '14). – ISBN 978–1–4503–2961–3, 20:1–20:1

[MM76] McGurk, Harry ; MacDonald, John: Hearing lips and seeing voices. In: *Nature* 264 (1976), Dec, Nr. 5588, 746–748. `http://dx.doi.org/10.1038/264746a0`. – DOI 10.1038/264746a0. – ISSN 1476–4687

[Mog06] Moggridge, Bill: *Designing Interactions*. The MIT Press, 2006

[Mon93] Montello, Daniel R.: Scale and multiple psychologies of space. In: Frank, Andrew U. (Hrsg.) ; Campari, Irene (Hrsg.): *Spatial Information Theory A Theoretical Basis for GIS*. Berlin, Heidelberg : Springer Berlin Heidelberg, 1993. – ISBN 978–3–540–47966–6, S. 312–321

[MSS11] Mulloni, Alessandro ; Seichter, Hartmut ; Schmalstieg, Dieter: Handheld augmented reality indoor navigation with activity-based instructions. In: *Proceedings of the 13th international conference on human computer interaction with mobile devices and services* ACM, 2011, S. 211–220

[New96] News, MIT: *Research sheds light on visual processing.* `http://news.mit.edu/1996/visualprocessing`, December 1996. – last retrieved 2019-11-03

[New02] *Kapitel* 4. In: Newcombe, Nora S.: *Spatial Cognition.* American Cancer Society, 2002. – ISBN 9780471214427, 113–163

[NFEL17] Niforatos, Evangelos ; Fedosov, Anton ; Elhart, Ivan ; Langheinrich, Marc: Augmenting Skiers' Peripheral Perception. In: *Proceedings of the 2017 ACM International Symposium on Wearable Computers.* New York, NY, USA : ACM, 2017 (ISWC '17). – ISBN 978–1–4503–5188–1, 114–121

[NHW08] Nasar, Jack ; Hecht, Peter ; Wener, Richard: Mobile telephones, distracted attention, and pedestrian safety. In: *Accident Analysis & Prevention* 40 (2008), Nr. 1, 69 - 75. `http://dx.doi.org/https://doi.org/10.1016/j.aap.2007.04.005`. – DOI https://doi.org/10.1016/j.aap.2007.04.005. – ISSN 0001–4575

[NK16] Nakuo, Takuro ; Kunze, Kai: Smart Glasses with a Peripheral Vision Display. In: *Proc. UbiComp '16.* New York, NY, USA : ACM, 2016. – ISBN 978–1–4503–4462–3, 341–344

[NNCK16] Nakao, Takuro ; Nakatani, Masashi ; Chan, Liwei ; Kunze, Kai: Smart Glasses with a Peripheral Vision Display. In: *Proceedings of the 2016 Virtual Reality International Conference.* New York, NY, USA : ACM, 2016 (VRIC '16). – ISBN 978–1–4503–4180–6, 16:1–16:3

[NS00] North, Chris ; Shneiderman, Ben: Snap-together Visualization: A User Interface for Coordinating Visualizations via Relational Schemata. In: *Proceedings of the Working Conference on Advanced Visual Interfaces.* New York, NY, USA : ACM, 2000 (AVI '00). – ISBN 1–58113–252–2, 128–135

[OBS16] Oppermann, Leif ; Blum, Lisa ; Shekow, Marius: Playing on AREEF: Evaluation of an Underwater Augmented Reality Game for Kids. In: *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services.* New York, NY, USA : ACM, 2016 (MobileHCI '16). – ISBN 978–1–4503–4408–1, 330–340

[OD71] O'Keefe, John ; Dostrovsky, Jonathan: The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat. In: *Brain Research* 34 (1971), S. 171–175. `http://dx.doi.org/10.1016/0006-8993(71)90358-1`. – DOI 10.1016/0006–8993(71)90358–1. – ISSN 1872–6240(Electronic), 0006–8993(Print)

[OEN09] Oulasvirta, Antti ; Estlander, Sara ; Nurminen, Antti: Embodied Interaction with a 3D Versus 2D Mobile Map. In: *Personal Ubiquitous Comput.* 13 (2009), Mai, Nr. 4, 303–320. `http://dx.doi.org/10.1007/s00779-008-0209-0`. – DOI 10.1007/s00779–008–0209–0. – ISSN 1617–4909

[OLL15] Ostendorp, Marie-Christin ; Lenk, Jan C. ; Lüdtke, Andreas: Smart Glasses to Support Maritime Pilots in Harbor Maneuvers. In: *Procedia Manufacturing* 3 (2015), 2840 - 2847. `http://dx.doi.org/http://dx.doi.org/10.1016/j.promfg.2015.07.775`. – DOI http://dx.doi.org/10.1016/j.promfg.2015.07.775. – ISSN 2351–9789

[ON78] O'Keefe, John ; Nadel, Lynn: *The hippocampus as a cognitive map.* Oxford: Clarendon Press, 1978

[OPF+16] Outram, Benjamin I. ; Pai, Yun S. ; Fan, Kevin ; Minamizawa, Kouta ; Kunze, Kai: AnyOrbit: Fluid 6DOF Spatial Navigation of Virtual Environments Using Orbital Motion. In: *Proceedings of the 2016 Symposium on Spatial User Interaction.* New York, NY, USA : ACM, 2016 (SUI '16). – ISBN 978–1–4503–4068–7, 199–199

[Org18] Organization, World H.: *Global status report on road safety 2018.* Geneva, Switzerland : World Health Organization, 2018

[Ost35] Osterberg, Gustav: Topography of the layer of rods and cones in the human retina. In: *Acta* (1935), S. 11–97

[OWK+14] Orlosky, Jason ; Wu, Qifan ; Kiyokawa, Kiyoshi ; Takemura, Haruo ; Nitschke, Christian: Fisheye Vision: Peripheral Spatial Compression for Improved Field of View in Head Mounted Displays. In: *Proc. SUI '14.* New York, NY, USA : ACM, 2014. – ISBN 978–1–4503–2820–3, 54–61

[PHF$^+$12] Poppinga, Benjamin ; Henze, Niels ; Fortmann, Jutta ; Heuten, Wilko ; Boll, Susanne: AmbiGlasses-Information in the Periphery of the Visual Field. In: *Mensch & Computer*, 2012, S. 153–162

[PJW73] Parker Jr, James F. ; West, Vita R.: Bioastronautics Data Book: NASA SP-3006. In: *NASA Special Publication* 3006 (1973)

[PKB10] Pielot, Martin ; Krull, Oliver ; Boll, Susanne: Where is My Team: Supporting Situation Awareness with Tactile Displays. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA : ACM, 2010 (CHI '10). – ISBN 978–1–60558–929–9, 1705–1714

[PMN18] Perea, Patrick ; Morand, Denis ; Nigay, Laurence: Halo3D: A Technique for Visualizing Off-screen Points of Interest in Mobile Augmented Reality. In: *Proceedings of the 29th Conference on L'Interaction Homme-Machine*. New York, NY, USA : ACM, 2018 (IHM '17). – ISBN 978–1–4503–5109–6, 43–51

[PMN19] Perea, Patrick ; Morand, Denis ; Nigay, Laurence: Spotlight on Off-Screen Points of Interest in Handheld Augmented Reality: Halo-based Techniques. In: *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces*. New York, NY, USA : ACM, 2019 (ISS '19). – ISBN 978–1–4503–6891–9, 43–54

[Pöp02] *Kapitel* 5. In: Pöppel, Ernst: *Three Seconds: A Temporal Platform for Conscious Activities*. Berlin, Heidelberg : Springer Berlin Heidelberg, 2002. – ISBN 978–3–642–50023–7, 73–79

[PP90] Posner, Michael I. ; Petersen, Steven E.: The Attention System of the Human Brain. In: *Annual Review of Neuroscience* 13 (1990), Nr. 1, 25–42. `http://dx.doi.org/10.1146/annurev.ne.13.030190.000325`. – DOI 10.1146/annurev.ne.13.030190.000325

[PPPW97] Pausch, Randy ; Pausch, Randy ; Proffitt, Dennis ; Williams, George: Quantifying Immersion in Virtual Reality. In: *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*. New York, NY, USA : ACM Press/Addison-Wesley Publishing Co., 1997 (SIGGRAPH '97). – ISBN 0–89791–896–7, 13–18

[PS09] Petersen, Nils ; Stricker, Didier: Continuous Natural User Interface: Reducing the Gap Between Real and Digital World. In: *Proceedings of the 2009 8th IEEE International Symposium on Mixed and Augmented Reality*. Washington, DC, USA : IEEE Computer Society, 2009 (ISMAR '09). – ISBN 978–1–4244–5390–0, 23–26

[Pso95] Psotka, Joseph: Immersive training systems: Virtual reality and education and training. In: *Instructional Science* 23 (1995), Nov, Nr. 5, 405–431. `http://dx.doi.org/10.1007/BF00896880`. – DOI 10.1007/BF00896880. – ISSN 1573–1952

[PT93] Pimentel, Ken ; Teixeira, Kevin: Virtual reality through the new looking glass. (1993)

[RC17] Rehman, Umair ; Cao, Shi: Augmented-Reality-Based Indoor Navigation: A Comparative Analysis of Handheld Devices Versus Google Glass. In: *IEEE Transactions on Human-Machine Systems* 47 (2017), Feb, Nr. 1, S. 140–151. `http://dx.doi.org/10.1109/THMS.2016.2620106`. – DOI 10.1109/THMS.2016.2620106. – ISSN 2168–2291

[RGCH16] Ren, Donghao ; Goldschwendt, Tibor ; Chang, Yunsuk ; Höllerer, Tobias: Evaluating wide-field-of-view augmented reality with mixed reality simulation. In: *2016 IEEE Virtual Reality (VR)*, 2016. – ISSN 2375–5334, S. 93–102

[RGK$^+$16] Rudi, David ; Giannopoulos, Ioannis ; Kiefer, Peter ; Peier, Christian ; Raubal, Martin: Interacting with Maps on Optical Head-Mounted Displays. In: *Proceedings of the 2016 Symposium on Spatial User Interaction*. New York, NY, USA : ACM, 2016 (SUI '16). – ISBN 978–1–4503–4068–7, 3–12

[RGR17] Rietzler, Michael ; Geiselhart, Florian ; Rukzio, Enrico: The Matrix Has You: Realizing Slow Motion in Full-body Virtual Reality. In: *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*. New York, NY, USA : ACM, 2017 (VRST '17). – ISBN 978–1–4503–5548–3, 2:1–2:10

[Roe59] Roelofs, Otto C.: Considerations on the visual egocentre. In: *Acta Psychologica, Amsterdam* 16 (1959), S. 226–234. `http://dx.doi.org/10.1016/0001-6918(59)90096-4`. – DOI 10.1016/0001–6918(59)90096–4

[RP17] Renner, Patrick ; Pfeiffer, Thies: 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, S. 186–194

[RTH16] Rakkolainen, Ismo ; Turk, Matthew ; Höllerer, Tobias: A Compact, wide-FOV Optical Design for Head-mounted Displays. In: *Proceedings of the 22Nd ACM Conference on Virtual Reality Software and Technology*. New York, NY, USA : ACM, 2016 (VRST '16). – ISBN 978–1–4503–4491–3, 293–294

[SB94] SARKAR, Manojit ; BROWN, Marc H.: Graphical fisheye views. In: *Communications of the ACM* 37 (1994), Nr. 12, S. 73–83

[SB14] STOKES, Dustin ; BIGGS, Stephen: The Dominance of the Visual. In: STOKES, D. (Hrsg.) ; MATTHEN, M. (Hrsg.) ; BIGGS, S. (Hrsg.): *Perception and its Modalities.* Oxford University Press, 2014, S. 5

[SBS11] STAVRINOS, Despina ; BYINGTON, Katherine W. ; SCHWEBEL, David C.: Distracted walking: Cell phones increase injury risk for college pedestrians. In: *Journal of Safety Research* 42 (2011), Nr. 2, 101 - 107. `http://dx.doi.org/https://doi.org/10.1016/j.jsr.2011.01.004`. – DOI https://doi.org/10.1016/j.jsr.2011.01.004. – ISSN 0022–4375

[SC12] SÁ, Marco de ; CHURCHILL, Elizabeth: Mobile Augmented Reality: Exploring Design and Prototyping Techniques. In: *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services.* New York, NY, USA : ACM, 2012 (MobileHCI '12). – ISBN 978–1–4503–1105–2, 221–230

[SCP95] STOAKLEY, Richard ; CONWAY, Matthew J. ; PAUSCH, Randy: Virtual Reality on a WIM: Interactive Worlds in Miniature. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* New York, NY, USA : ACM Press/Addison-Wesley Publishing Co., 1995 (CHI '95). – ISBN 0–201–84705–1, 265–272

[SEO+14] SUKAN, Mengu ; ELVEZIO, Carmine ; ODA, Ohan ; FEINER, Steven ; TVERSKY, Barbara: ParaFrustum: Visualization Techniques for Guiding a User to a Constrained Set of Viewing Positions and Orientations. In: *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology.* New York, NY, USA : ACM, 2014 (UIST '14). – ISBN 978–1–4503–3069–5, 331–340

[SFS+07] SCHICHO, Kurt ; FIGL, Michael ; SEEMANN, Rudolf ; DONAT, Markus ; PRETTERKLIEBER, Michael L. ; BIRKFELLNER, Wolfgang ; REICHWEIN, Astrid ; WANSCHITZ, Felix ; KAINBERGER, Franz ; BERGMANN, Helmar ; WAGNER, Arne ; EWERS, Rolf: Comparison of laser surface scanning and fiducial marker–based registration in frameless stereotaxy. In: *Journal of Neurosurgery* 106 (2007), Nr. 4, 704-709. `http://dx.doi.org/10.3171/jns.2007.106.4.704`. – DOI 10.3171/jns.2007.106.4.704. – PMID: 17432726

[SGB18] STRATMANN, Tim C. ; GRUENEFELD, Uwe ; BOLL, Susanne: EyeMR: Low-cost Eye-tracking for Rapid-prototyping in Head-mounted Mixed Reality. In: *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications.* New York, NY,

USA : ACM, 2018 (ETRA '18). – ISBN 978–1–4503–5706–7, 90:1–90:2

[SGS+18] Stratmann, Tim C. ; Gruenefeld, Uwe ; Stratmann, Julia ; Schweigert, Sören ; Hahn, Axel ; Boll, Susanne: Mobile Bridge-A Portable Design Simulator for Ship Bridge Interfaces. In: *TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation* 12 (2018), Nr. 4

[SGT09] Straughn, Stacie M. ; Gray, Rob ; Tan, Hong Z.: To Go or Not to Go: Stimulus-Response Compatibility for Tactile and Auditory Pedestrian Collision Warnings. In: *IEEE Transactions on Haptics* 2 (2009), April, Nr. 2, S. 111–117. `http://dx.doi.org/10.1109/TOH.2009.15`. – DOI 10.1109/TOH.2009.15. – ISSN 1939–1412

[SH13] Siu, Teresa ; Herskovic, Valeria: SidebARs: Improving Awareness of Off-screen Elements in Mobile Augmented Reality. In: *Proceedings of the 2013 Chilean Conference on Human - Computer Interaction*. New York, NY, USA : ACM, 2013 (ChileCHI '13). – ISBN 978–1–4503–2200–3, 36–41

[SHB10] Schinke, Torben ; Henze, Niels ; Boll, Susanne: Visualization of Off-screen Objects in Mobile Augmented Reality. In: *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*. New York, NY, USA : ACM, 2010 (MobileHCI '10). – ISBN 978–1–60558–835–3, 313–316

[She03] Shelton, Amy L.: Putting spatial memories into perspective: Brain and behavioral evidence for representational differences. In: *Human spatial memory: Remembering where*. Lawrence Erlbaum Associates, 2003, S. 309–327

[SJS14] Steptoe, William ; Julier, Simon ; Steed, Anthony: Presence and discernability in conventional and non-photorealistic immersive augmented reality. In: *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2014, S. 213–218

[SK08] Schwerdtfeger, Bjorn ; Klinker, Gudrun: Supporting Order Picking with Augmented Reality. In: *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*. Washington, DC, USA : IEEE Computer Society, 2008 (ISMAR '08). – ISBN 978–1–4244–2840–3, 91–94

[SLG+18] Stratmann, Tim C. ; Löcken, Andreas ; Gruenefeld, Uwe ; Heuten, Wilko ; Boll, Susanne: Exploring Vibrotactile and Peripheral Cues for Spatial Attention Guidance. In: *Proceedings of the 7th ACM International Symposium on Pervasive Displays*. New

York, NY, USA : ACM, 2018 (PerDis '18). – ISBN 978–1–4503–5765–4, 9:1–9:8

[SM04] Shelton, Amy L. ; McNamara, Timothy P.: Orientation and Perspective Dependence in Route and Survey Learning. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30 (2004), Nr. 1, S. 158–170. `http://dx.doi.org/10.1037/0278-7393.30.1.158`. – DOI 10.1037/0278–7393.30.1.158. – ISSN 1939–1285(Electronic), 0278–7393(Print)

[SRJ11] Strasburger, Hans ; Rentschler, Ingo ; Jüttner, Martin: Peripheral vision and pattern recognition: A review. In: *Journal of Vision* 11 (2011), Nr. 5, 13. `http://dx.doi.org/10.1167/11.5.13`. – DOI 10.1167/11.5.13

[SRK+10] Shingu, Jun ; Rieffel, Eleanor ; Kimber, Don ; Vaughan, Jim ; Qvarfordt, Pernilla ; Tuite, Kathleen: Camera pose navigation using Augmented Reality. In: *2010 IEEE International Symposium on Mixed and Augmented Reality*, 2010, S. 271–272

[Ste95] Steuer, Jonathan: Communication in the Age of Virtual Reality. Version: 1995. `http://dl.acm.org/citation.cfm?id=207922.207925`. In: Biocca, Frank (Hrsg.) ; Levy, Mark R. (Hrsg.): *Social Responses to Communication Technology*. Hillsdale, NJ, USA : L. Erlbaum Associates Inc., 1995. – ISBN 0–8058–1550–3, Kapitel Defining Virtual Reality: Dimensions Determining Telepresence, 33–56

[Sut65] Sutherland, Ivan E.: The ultimate display. In: *Multimedia: From Wagner to virtual reality* (1965), S. 506–508

[Sut68] Sutherland, Ivan E.: A Head-mounted Three Dimensional Display. In: *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I*. New York, NY, USA : ACM, 1968 (AFIPS '68 (Fall, part I)), 757–764

[SZ94] Sturman, David J. ; Zeltzer, David: A survey of glove-based input. In: *IEEE Computer Graphics and Applications* 14 (1994), Jan, Nr. 1, S. 30–39. `http://dx.doi.org/10.1109/38.250916`. – DOI 10.1109/38.250916. – ISSN 0272–1716

[SZG+96] Schaffer, Doug ; Zuo, Zhengping ; Greenberg, Saul ; Bartram, Lyn ; Dill, John ; Dubs, Shelli ; Roseman, Mark: Navigating Hierarchically Clustered Networks Through Fisheye and Full-zoom Methods. In: *ACM Trans. Comput.-Hum. Interact.* 3 (1996), Juni, Nr. 2, 162–188. `http://dx.doi.org/10.1145/230562.230577`. – DOI 10.1145/230562.230577. – ISSN 1073–0516

[TMR90] Taube, Jeffrey S. ; Muller, Robert U. ; Ranck, James B.: Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. In: *Journal of Neuroscience* 10 (1990), Nr. 2, 420–435. `http://dx.doi.org/10.1523/JNEUROSCI.10-02-00420.1990`. – DOI 10.1523/JNEUROSCI.10–02–00420.1990. – ISSN 0270–6474

[TOBM03] Tang, Arthur ; Owen, Charles ; Biocca, Frank ; Mou, Weimin: Comparative Effectiveness of Augmented Reality in Object Assembly. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA : ACM, 2003 (CHI '03). – ISBN 1–58113–630–7, 73–80

[Tol48] Tolman, Edward C.: Cognitive maps in rats and men. In: *Psychological Review* 55 (1948), Nr. 4, S. 189–208. `http://dx.doi.org/10.1037/h0061626`. – DOI 10.1037/h0061626. – ISSN 1939–1471(Electronic),0033–295X(Print)

[TSK97] Tate, David L. ; Sibert, Linda ; King, Tony: Virtual environments for shipboard firefighting training. In: *Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality*, 1997, S. 61–68

[TSL+11] Trapp, Matthias ; Schneider, Lars ; Lehmann, Christine ; Holz, Norman ; Döllner, Jürgen: Strategies for visualising 3D points-of-interest on mobile devices. In: *Journal of Location Based Services* 5 (2011), Nr. 2, 79-99. `http://dx.doi.org/10.1080/17489725.2011.579579`. – DOI 10.1080/17489725.2011.579579

[TSLB05] Tonnis, Marcus ; Sandor, Christian ; Lange, Christian ; Bubb, Heiner: Experimental Evaluation of an Augmented Reality Visualization for Directing a Car Driver's Attention. In: *Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality*. Washington, DC, USA : IEEE Computer Society, 2005 (ISMAR '05). – ISBN 0–7695–2459–1, 56–59

[Tve00] Tversky, Barbara: Levels and structure of spatial knowledge. In: *Cognitive mapping: Past, present and future* (2000), S. 24–43

[Vei15] Veit, Sven-Michael: *Die Elbe ist ein enges Revier*. `http://www.taz.de/!879062/`, 2015. – Letzter Zugriff am 11. Juni 2017

[Wag06] Wagner, Mark: *The geometries of visual space*. Mahwah, NJ, US : Lawrence Erlbaum Associates Publishers, 2006 (The geometries of visual space.). – ISBN 0–8058–5253–0 (Paperback)

[WDH09] Wither, Jason ; DiVerdi, Stephen ; Höllerer, Tobias: Annotation in outdoor augmented reality. In: *Computers & Graphics* 33 (2009), Nr. 6, 679 - 689. `http://dx.doi.org/https://doi.org/10.1016/j.cag.2009.06.001`. – DOI https://doi.org/10.1016/j.cag.2009.06.001. – ISSN 0097–8493

[Wic08] Wickens, Christopher D.: Multiple Resources and Mental Workload. In: *Human Factors* 50 (2008), Nr. 3, 449-455. `http://dx.doi.org/10.1518/001872008X288394`. – DOI 10.1518/001872008X288394. – PMID: 18689052

[WS98] Witmer, Bob G. ; Singer, Michael J.: Measuring Presence in Virtual Environments: A Presence Questionnaire. In: *Presence: Teleoperators and Virtual Environments* 7 (1998), Nr. 3, 225-240. `http://dx.doi.org/10.1162/105474698565686`. – DOI 10.1162/105474698565686

[WSFR17] Wintersberger, Philipp ; Sawitzky, Tamara von ; Frison, Anna-Katharina ; Riener, Andreas: Traffic Augmentation As a Means to Increase Trust in Automated Driving Systems. In: *Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter*. New York, NY, USA : ACM, 2017 (CHItaly '17). – ISBN 978–1–4503–5237–6, 17:1–17:7

[XB16] Xiao, Robert ; Benko, Hrvoje: Augmenting the Field-of-View of Head-Mounted Displays with Sparse Peripheral Displays. In: *Proc. CHI '16*. New York, NY, USA : ACM, 2016. – ISBN 978–1–4503–3362–7, 1221–1232

[YM16] Yamada, Wataru ; Manabe, Hiroyuki: Expanding the Field-of-View of Head-Mounted Displays with Peripheral Blurred Images. In: *Proc. UIST '16 Adjunct*. New York, NY, USA : ACM, 2016. – ISBN 978–1–4503–4531–6, 141–142

[ZDB08] Zhou, Feng ; Duh, Henry Been-Lirn ; Billinghurst, Mark: Trends in Augmented Reality Tracking, Interaction and Display: A Review of Ten Years of ISMAR. In: *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*. Washington, DC, USA : IEEE Computer Society, 2008 (ISMAR '08). – ISBN 978–1–4244–2840–3, 193–202

[ZF14] *Kapitel* 8. In: Zimmerman, John ; Forlizzi, Jodi: *Research Through Design in HCI*. New York, NY : Springer New York, 2014. – ISBN 978–1–4939–0378–8, 167–189

[ZFE07] Zimmerman, John ; Forlizzi, Jodi ; Evenson, Shelley: Research through Design as a Method for Interaction Design Research in

HCI. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA : Association for Computing Machinery, 2007 (CHI '07). – ISBN 9781595935939, 493–502

[ZMG+03] Zellweger, Polle T. ; Mackinlay, Jock D. ; Good, Lance ; Stefik, Mark ; Baudisch, Patrick: City Lights: Contextual Views in Minimal Space. In: *CHI '03 Extended Abstracts on Human Factors in Computing Systems*. New York, NY, USA : ACM, 2003 (CHI EA '03). – ISBN 1–58113–637–4, 838–839

[ZSd13] Zhang, Emile ; Saito, Hideo ; de Sorbier, François: From smartphone to virtual window. In: *2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, 2013, S. 1–6