Interaction with mobile augmented reality

An exploratory study using design research to investigate mobile and handheld augmented reality

Tor Gjøsæter

Dissertation for the degree philosophiae doctor (PhD)
at the University of Bergen

2015

Dissertation date: June 2nd 2015

Year: 2015

Title: Interaction with mobile augmented reality

An exploratory study using design research to investigate mobile and handheld augmented reality

Author: Tor Gjøsæter

Print: AIT OSLO AS / University of Bergen

Acknowledgements

Firstly, a sincere thank you to Professor Weiqin Chen for her dedicated support throughout this process. Without her guidance, this thesis would never have happened.

My colleagues at the department: Frode Guribye, who deserves a thank you for teaching me the extracurricular aspects of being an academic, Lars Nyre for his inspiring words and Kristine Jørgensen for the fruitful collaboration.

The team at Bryggen Research, including Aleksander Krzywinski, Jørund Fjøsne and Martin Lie. Thank you for your constant encouragement and for bothering to listen to my arguments. My associates in the Augmented Reality Research Group: Jonas Waage, Ingerid Rødseth, David Burns, and Jan Andre Johnsen. Thank you for an exciting time developing and releasing ARad.

My sister, Trude Meling, for going the extra mile and coming with me to India to keep me company while I presented a paper. My parents, Geir Olav Meling and Audny Gjøsæter, for not asking, "How is the thesis work coming along?" over the last year, and just silently trusting it would eventually come to pass.

The in-laws, Lars Magne Halleraker and Heidi Irene Vespestad, for their support. Lastly, but most importantly, I would like to thank Helga Marie Halleraker, my superior half. There comes a point in every thesis when you need someone to be there with you for the paper rejections, the acceptances, the fallbacks, the grind, and the eventual triumph. Thank you for the small things that make life worth living.

Bømlo, November 2014

Abstract

Mobile augmented reality (MAR) has matured significantly over the last two decades, from the conceptual idea of an untethered handheld augmented reality experience in the mid-90s, to prototype technological demonstrations in the early 00s, to actual practical applications in use on smartphones everywhere today. Until recently, the field has been concerned with solving important technical obstacles particular to augmented reality (AR) systems, as well creating systems to augment the world in usable and entertaining ways.

To further the understanding of MAR specifically, the concept needs to be seen in its own context through a clear conceptual model and an exploration of how the underlying technology supports the interplay between user and device. MAR systems provide an entirely new point of entry to content such images, videos and 3D information, both in a physical and a technological sense. What content is suitable for this interactive platform and how does the creation of meaning take place in this context?

To investigate this type of technology, the specific traits of the system need to be taken into account when gathering data. Similarly, the dissemination of findings about this very visual platform requires a visual language to effectively convey meaning. There are few general design guidelines for MAR that are founded on human computer interaction (HCI) theory and empirical data. Thus, additional domain-specific guidelines must be composed for those designing MAR systems.

To approach these issues, this thesis has adopted the design science research (DSR) framework, which provides a set of guidelines for conducting science and research on designed artifacts. This framework also has a strong focus on solving real-world problems and acknowledges artifacts as contributions in themselves. Its guidelines have been used to formulate the problem statements, directing the research design and evaluation as well as the dissemination of findings and results. In conjunction with the overarching guidelines, Think-Aloud (THA) and video recording have been

used for evaluation, while user-centered design has served as a foundation specifying the design of the artifacts. A qualitative analysis of the findings has been performed based on theories from HCI.

The result of this work is a tangible application that is freely available on the App Store for iOS. Additionally, we present a conceptual model describing handheld mobile augmented reality and an approach to using THA and video recording for evaluation and analysis of MAR systems. A novel approach for illustrative dissemination of findings using the empirical data is described, and general guidelines for MAR systems are presented. Finally, this thesis serves as a guide for conducting similar design science research on AR technology.

The contribution of these results is a further understanding of the conceptualization, design, evaluation and dissemination of MAR.

List of publications

Paper 1 - T. Gjøsæter, "Computer Supported Collaborative Design Using Augmented Reality," presented at the SOCINFO '09: Proceedings of the 2009 International Workshop on Social Informatics, IEEE Computer Society, 2009.

Paper 2 - T. Gjøsater, "A taxonomy of handheld augmented reality applications," presented at the 2012 4th International Conference on Intelligent Human Computer Interaction (IHCI), 2012, pp. 1–6.

Paper 3 - T. Gjøsæter and K. Jørgensen, "Combining Think Aloud and Comic Strip Illustration in the Study of Augmented Reality Games" presented at the NOKOBIT 2012, 2012, pp. 77–90.

Paper 4 - T. Gjøsæter, "Affordances in Mobile Augmented Reality Applications," International Journal of Interactive Mobile Technologies (iJIM) Vol 8, No 4, 2014 pp. 45-55.

When referencing the papers in this thesis, they are referenced by these numbers.

Contents

Table of acronyms

1080p	High definition video with a resolution of 1920x1080
AR	Augmented Reality
CAR	Contact-lens augmented reality
CCS	Charge-coupled Device
CPU	Central Processing Unit
CSCW	Computer Supported Collaborative Work
CV	Computer Vision
DR	Diminished Reality
DSR	Design Science Research
EEG	Electroencephalography
FTP	File Transfer Protocol
GPS	Global Positioning System
GUI	Graphical User Interface
h264	A video codec
HCI	Human Computer Interaction
HMAR	Handheld Mobile Augmented Reality
HMD	Head Mounted Display
HMPD	Head-Mounted Projective Display
IEEE	Institute of Electrical and Electronics Engineers
iOS	A mobile operating system
IS	Information Systems
ISMAR	International Symposium on Mixed and Augmented Reality
IT	Information Technology
MAR	Mobile augmented reality
MPEG	A video format
OLED	Organic Light-emitting Diode
PC	Personal Computer
SDK	Systems Development Kit
SVN	Apache Subversion

THA	Think aloud
TV	Television
UCD	User Centered Design
UI	User Interface
USE questionnaire	Usefulness, Satisfaction, and Ease of use
VG	Verdens Gang
VR	Virtual Reality
WIMP	Windows, Icons, Menus, Pointer

1. Introduction

How we think of and use mobile devices has changed drastically since the release of the first mobile phones. First used as devices for making calls on the go, they have evolved to support a wide range of tasks. Texting, photography, listening to music, web-browsing, and other functionalities traditionally found in dedicated devices have been incorporated into what we now know as smartphones.

With the advancement of smartphones to include better cameras, more powerful CPUs, and easy distribution for new applications (app stores), mobile augmented reality (MAR) has unfolded from proof-of-concept prototypes in R&D projects in the early 2000s to usable applications available to everyone with a smartphone.

AR-related technologies that support entertainment, browsing of location-based content, and the identification and augmentation of real-world objects, are now available to users. In the area of entertainment, we find games that merge real-world and interactive virtual game elements. Another example can be seen in mapping applications that overlay the world with relevant metadata in 3D and in real time, applications promising to bring “print to life” (e.g., “Layar-Augmented Reality”) and applications aimed at aiding design by providing real-time visualization of 3D models in the real world. Schmalstieg and Wagner (2007) developed an augmented reality game for educational use called “Expedition Schatzsuche,” which runs on a Windows CE device. A system that translates text and overlays that translation in the real world using a Nokia N900 was described in a 2011 paper by Fragoso, Gauglitz, Zamora, Kleban, and Turk. MapLens (Morrison et al., 2009) provides the user with extra metadata for printed maps.

A number of researchers in the field of AR and MAR (De Sà & Churchill, 2013; Swan & Gabbard, 2005; Zhou, Duh & Billinghurst, 2008) have been calling for the research community to describe and analyze user interactions with *real* MAR systems, beyond informal user testing. This thesis aims to answer this call by presenting a design-research and development project on the topic of AR on mobile devices.

The project has been a collaboration between media outlets, technology providers, and researchers, resulting in several tangible artifacts. The design of these artifacts has been evaluated and analyzed to describe and explain how the user experiences the MAR interaction paradigm.

1.1 Augmented Reality

Mobile augmented reality (MAR) or handheld mobile augmented reality (HMAR) falls within the boundaries of augmented reality (AR) research. The initial description and idea behind AR can be attributed to Sutherland's 1965 essay on the topic. A descriptive technical definition for AR is provided by Ronald Azuma (1997). This definition is a basis for discussing AR in precise terms.

Azuma's definition clearly states that technology that augments reality must have the following three properties:

1. It should combine the real and the virtual.
2. The augmentations should be interactive in real time.
3. They should be registered in three dimensions.

In contrast to virtual reality (VR), AR does not replace the real world with a simulated world (Steuer, 1992). Rather, it seeks to combine the real with the virtual, with the virtual (augmentations) being interactive in real time and in three dimensions. Feiner (2002) attributes the coining of the term “augmented reality” to researchers at Boeing who were working with assembly line optimization. The researchers presented their assembly line AR as follows:

> *This technology is used to “augment” the visual field of the user with the information necessary in the performance of the current*

> *task, and therefore we refer to the technology as "augmented reality" (AR).* (Caudell & Mizell, 1992)

Up to that point, augmented reality as a field of research belonged to a few privileged institutions and corporations. The head-mounted displays and computers capable of delivering the graphics needed to create augmented reality applications were expensive; furthermore, the computer vision and tracking algorithms remained largely closed-source and unavailable beyond the research labs pioneering the field.

Azuma (1997) identifies six applications for AR technology based on work performed up to this time: medical visualization, maintenance and repair, annotation, robot path planning, entertainment, and military navigation and targeting. With the release of the ARToolKit (Lamb, 2003) website containing an open-source computer vision-based tracker, it became possible for researchers worldwide to create their own content and build upon the tools provided by Lamb. With the porting of ARToolKit onto a Windows CE phone in 2003 (Wagner & Schmalstieg, 2003), the stage was set for the advent of handheld mobile augmented reality (HMAR).

With the recent proliferation of mobile devices incorporating more processing power and features, including GPS, accelerometers, gyroscopes, and advanced computer vision algorithms, AR has become both feasible and affordable, leading to its widespread adoption.

1.2 HMAR, MAR and AR

Prior to any further discussion of augmented reality on handheld and mobile devices, it will be beneficial to provide a clear definition of the technology. The two acronyms, MAR and HMAR, represent different subsets of AR. MAR may be understood as any type of standalone mobile augmented reality running on a mobile device while still keeping true to Azuma's (1997) overall definition (Huang, Hui, Peylo, & Chatzopoulos, 2013). MAR may be based on head mounted display (HMD-AR), handheld (HMAR), contact lenses (CAR; see Anthony, 2013) or any type of technology a user can carry with them to visually augment the world. In a recent extensive study, Huang, Hui, Peylo, and Chatzopoulos (2013) categorize the computing platforms for AR systems as being either tethered to notebooks (see Piekarski & Thomas, 2002) carried in backpacks, running and viewed on PDAs (see Pasman & Woodward, 2003), tablet computers (see Ferdinand, Müller & Ritschel, 2005), Ultra Mobile PCs (see Kang et al., 2008), mobile phones (see Mohring, Lessig, & Bimber, 2004) and AR-glasses.

HMAR is a subset of MAR specifically referring to handheld devices that can display AR content (Figure 1).

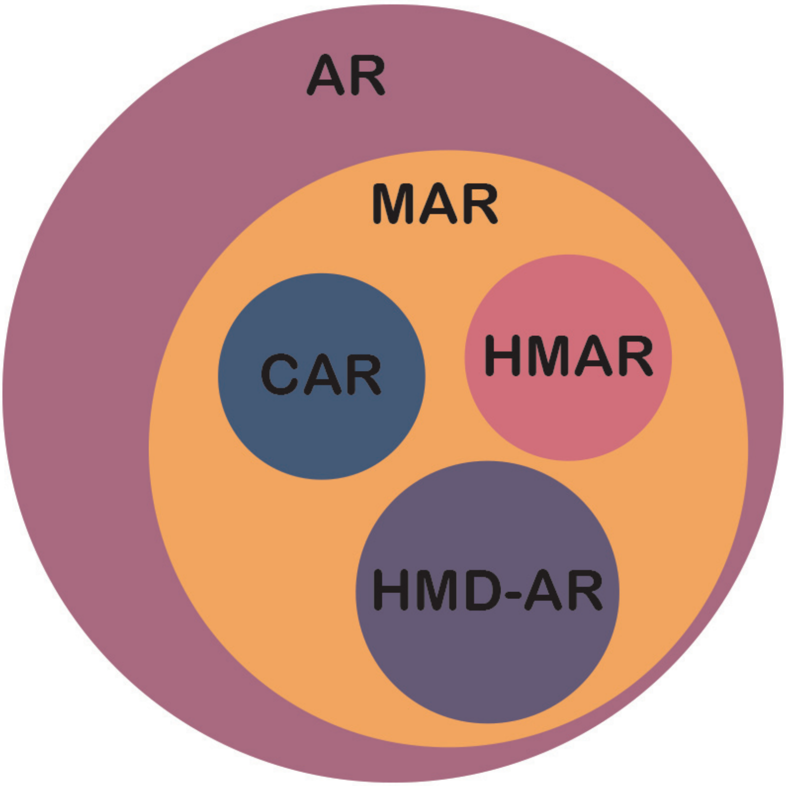

Figure 1: AR encompasses a broad spectrum of systems. Within the MAR domain are systems such as CAR, HMAR and HMD-based AR.

The research community has yet to establish a standardized set of acronyms, and there exist competing proposals for certain acronyms. In this thesis and context, MAR and HMAR will effectively refer to the same technology. Effort has been made to use the HMAR acronym to specifically refer to AR on handheld mobile devices.

1.3 Background and previous research

Zhou et al. (2008) pointed out in their survey of augmented reality that AR applications to that point had largely taken the form of technology demonstrations. Thus, they presented a call for research on usable applications outside the research setting.

The technological demonstrations that created the intial outline for HMAR can be traced back to Rekimoto's description of a handheld AR system for collaborative design (Rekimoto, 1996) called TransVision. Rekimoto argued that head-mounted displays isolate users from the real world and that a handheld system would allow for a more natural interaction in which the users could observe and use body language while designing.

Following in Rekimoto's footsteps, the next big technological demonstration came when Wagner and Schmalstieg created the first standalone handheld mobile AR system (Wagner & Schmalstieg, 2003), while Möhring, Lessig, and Bimber (2004) presented the first video see-through AR system on a consumer cell phone. "Standalone" refers to an untethered device that provides an augmented reality overlay of the real world. A range of technological demonstrations have since been provided to exemplify the potential use of HMAR. These include adaptations of technology demonstrations from the general field of AR into HMAR, and novel ideas applied to HMAR. Several commercial and open source systems development kits (SDKs) and applications have also come about (Huang et al., 2013).

Several technical demonstrations led to the conceptualization and design of ARad, a prototype application developed by HIT Lab Nz in 2007 for creating augmented reality advertisements. According to Schmalstieg, Langlotz, and Billinghurst (2011), the development of ARad inspired further investigation on the topic. A simple tracking application with embedded 3D content was made available on the Symbian platform. When the user launched the application and pointed it toward a marker, different zoo animals appeared. Schmalstieg et al. (2011) note that in “all cases the most challenging aspects have been the content creation and application distribution, not the application programming” (p. 26).

In 2005, Swan and Gabbard (2005) pointed out the scarcity of user studies in the field of AR. Dünser, Grasset, and Billinghurst (2008) concluded[1] in a survey paper that only about 10% of AR-related papers between 1992 and 2007 included any type of user evaluation. Furthermore, the papers looked at early prototypes, largely focusing on cognition-, perception-, or task-related performance issues. De Sà and Churchill’s (2013) recent review of mobile augmented reality (MAR) user studies concludes that a gap in the knowledge remains. Few extensive user studies on AR in general can be found, and even fewer user studies focus on the explicit MAR domain, while still fewer focus on the handheld platform. De Sà and Churchill (2013) gather that “despite the appeal and the growing number of services and applications, very few guidelines, design techniques and evaluation methods have been presented in the existing literature” (p. 160).

In their review, De Sà and Churchill (2013) categorize user and design studies within the domain of mobile augmented reality. In a table showing these different user and design studies (p. 145), the authors present, how the studies gathered data, their design approach, the type of system used, and so forth. Of the ten presented studies, two used video recordings as a data gathering technique. Another study (Damala, Cubaud, Bationo, & Houlier, 2008) did not make use of the video data, but hoped to analyze the video data in the future. In two different works, Morrison and colleagues (Morrison et al., 2009; Morrison, Mulloni, Lemmelä, & Oulasvirta, 2011) performed an extensive video-based study, thereby illustrating the potential of using video to

[1] http://www.hitlabnz.org/

analyze collaborative aspects of mobile augmented reality applications. None of the studies reported by De Sà and Churchill mentioned affordance or current guidelines related to the affordances of the applications. In regard to the application domain, none of the studies directly investigated marker-based systems for content remediation. The prototypes and applications listed by De Sà and Churchill are all indeed prototypes and are not available to the general public.

With regard to conceptual knowledge about HMAR systems, some work has been done. In this thesis, it is argued that there is a difference between an HMD-based mobile AR system and a handheld mobile AR system. No literature exemplifying how HMAR systems differ from other AR systems, however, was found. Some studies have used Milgram's continuum (Milgram, Takemura, Utsumi, & Kishino, 1994) to group HMAR with other AR and VR systems. The simplified representation presented in Figure 2 is used to illustrate the continuum visually.

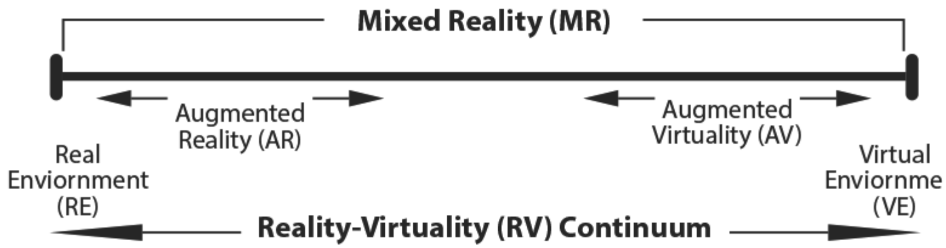

Figure 2: Milgram's Reality Virtuality (RV) continuum allows the definition of mixed reality environments that use a display to show real and virtual objects together. The continuum moves from entirely virtual on the right extreme to entirely real on the left extreme.

Milgram's continuum advances from the "real environment," consisting of real objects, through an increasingly virtual system to an entirely "virtual environment." Mixed reality in this context becomes an "environment as one in which real-world and virtual-world objects are presented together within a single display" (Milgram et al., 1994, p. 283)

Rosenblum, Feiner, Julier, and Swan (2012) disavow HMAR entirely, believing it to be a useless field within AR: “We believe that if AR is to realize its full potential, hand-held form factors, despite much of the hype they are receiving now, simply are not adequate” (p. 445). This statement comes in sharp contrast to Zhou et al. (2008), who note that “handheld displays are a good alternative to HMD and HMPD systems for AR applications, particularly because they are minimally intrusive, socially acceptable, readily available and highly mobile” (p. 198).

Olsson and Salo (2012) classify the currently existing MAR applications as either “AR-browsers” or “image recognition-based” AR applications. Browser-type applications typically augment the world through a magic lens using GPS as a frame of reference to generate the virtual world, while image recognition-based systems use computer vision capabilities to track real-world objects. While this categorization may be useful in Olsson and Salo’s study, it could create confusion when studying applications that allow a user to browse a newspaper with image recognized markers. Hence, a conceptual model that takes into account the overall traits of MAR applications would be beneficial to the field.

The current state of research into MAR can be summarized as having a technical focus, with a small but gradually increasing emphasis on the human factors of AR. The technology focus stems from many complex issues still requiring solutions. The International Symposium on Mixed and Augmented Reality (ISMAR) is the main conference the state of the art of technical AR research can be found. Papers presented at the 2014 conference give some indications of where the research focus is aimed. Computer vision remains a complex and intricate computer science issue that AR relies heavily on. Further improvements in the technical areas crucial for AR, such as rendering (Rohmer, Buschel, Dachselt & Grosch, 2014), reconstruction and fusion (Foxlin, Calloway & Zhang, 2014), tracking (Zheng, Schmalstieg & Welch, 2014) and user interfaces (Piumsomboon et al., 2014), are clearly needed. However, papers focusing on human factors, theory, and evaluation (Marzo, Bossavit & Hachet, 2014; Shilkrot, Montfort & Maes, 2014) have recently been called for and included in proceedings at ISMAR.

Studies of HMAR that go beyond the technical aspects are crucial for several reasons. An understanding of what makes a cogent user experience in the HMAR domain is important to those who design with the user in mind. There lies great potential in HMAR systems to create new tools for doing serious work as well as systems for entertainment. Researchers, designers, and technical and commercial interests seek to create novel and better HMAR systems. Thus, guidelines for design grounded in empirical data, conceptual models providing a common and coherent vocabulary and understanding, and a methodology that would allow anyone to research and distribute their findings about HMAR are needed.

1.4 Meaning-making and experience

In this thesis, we propose an answer for how to design HMAR systems that facilitate meaning-making. According to Harrison, Tatar and Sengers (2007), three different approaches to the creation of meaning can be found in human computer interaction (HCI) theory. Using a model of paradigms, the authors characterize HCI research as having three waves. The first paradigm is described as “interaction as a form of man-machine coupling” (p. 3) inspired by engineering and ergonomics with an overarching goal of optimizing the interaction between man and machine. The second paradigm “is organized around a central metaphor of mind and computer as coupled information processors.” (p. 3), where the aim of the research is to understand how information is communicated, flows and is efficiently transformed. The authors go on to define the principles behind the *third paradigm* in contrast to the others with regards to construction of meaning:

> *The first paradigm tends to take a pragmatic approach to meaning, ignoring it unless it causes a problem, while the second interprets meaning in terms of information flows. The third paradigm, in contrast, sees meaning and meaning construction as a central focus.*
> (Harrison et al., 2007, p. 5)

Meaning-making is, as they argue, “irreducibly connected to the viewpoints, interactions, histories, and local resources available to those making sense of the interface and therefore to some extent beyond the reach of formalization” (p. 6). Understanding context hence becomes crucial in this third paradigm and cannot be disregarded as simply “those non-technological factors that affect the use of the technology” (p. 6). The authors term the paradigm “*situated perspectives*” and it is worth noting that they do not try to renounce other paradigms in HCI with this framing. Rather, the framing can be used as a tool for understanding the different approaches to HCI theory, what questions the different paradigms seek to answer, and what methods may be fruitful in seeking those answers. They conclude by stating, “it would probably be unwise to attempt to uncover the rich appropriations of a situated technology with an objective laboratory test.” (p. 18).

Meaning-making is frequently used in relation with computer supported collaborative learning. FitzGerald et al. (2012) emphasize the physical aspect of AR in learning, stating that MAR “enables us to integrate real-world experience and meaning within specific physical contexts” (p. 2). Furthermore, they argue that “any material object we interact with is artificial, in the sense that our perception of that object is shaped by culture and history” (p. 3). This is a central point in this thesis; we must seek to understand the meaning made by the artifact through its interaction within a context, as well as grounded in culture.

As the field is moving towards this third paradigm, we may look to the thoughts put forth to the HCI community by McCarthy and Wright (2004) in their book “Technology as experience” to elaborate the topic further. While not explicitly declaring their perspective as equivalent to the third paradigm, they perceive that “although HCI has its roots in laboratory subjects such as psychology and computer science, it has in recent times been strongly influenced by concerns for experience.” (loc. 2466).

McCarthy and Wright (2004) put forth a concept they call the threads of experience, where the idea is to “help us think more clearly about technology as experience.” The idea can be summarized as having four threads: the sensual, the emotional, the compositional and the spatio-temporal. The sensual thread is concerned with the "sensory engagement with a situation, which orients us to the concrete, palpable, and visceral character of experience" (loc. 1072). They sensual thread is the "sense or meaning immediately available in a situation" (loc. 1177), while the emotional thread is concerned with "sense or meaning ascribed to an object or person because of the values, goals, and desires we have" (loc 1177). The compositional thread "is concerned with relationships between the parts and the whole of an experience" (loc. 1177). Rogers, Sharp and Preece (2011) frame this thread as the "narrative part of an experience as it unfolds, and the way a person makes sense of them" (p.152). The final thread is the spatio-temporal thread, which Rogers, Sharp and Preece (2011) summarize as referring to "the space and time in which our experiences take place and their effect upon those experiences" (p. 152). According to Rogers et al. (2011), McCarthy and Wright´s (2004) framework can "aid thinking about the whole experience of a technology rather than as fragmented aspects, e.g. its usability, its marketability, or utility" (p. 152).

The major point McCarthy and Wright (2004) try to make is about the "danger that the cultural analysis being developed in HCI and CSCW in the name of practice theory, activity theory, or user experience can be used in a meaningless manner, separate from people fearing, hoping, imagining, revolting, and consoling" (loc. 2474). Hence, we should aim to consider these aspects when trying to understand what constitutes a meaningful experience of technology. In this thesis, interaction with technology is recognized as rooted in the cultural and contextual, and as having an emotional ingredient.

1.5 Problem area and research question

The topic of this thesis is the interaction and interfaces in mobile augmented reality. Based on the studies of Zhou et al. (2008), Swan and Gabbard (2005), and de Sà and Churchill (2013), as well as their calls for research into the applicable human factors of AR technology, an overarching research question has been formulated.

How can interaction with HMAR be designed to facilitate meaning-making?

To answer this question, several problems need to be investigated. In this thesis, design science research (DSR) is used to frame the research. The DSR approach to science provides guidelines and a framework for structuring research coherently. Research questions in design science should originate from a clearly defined problem area or areas.

Problem 1. How to conceptualize HMAR

AR is a fragmented field with competing technologies that overlap and weave together, so a clear definition of HMAR as a specific subset of MAR needs to be specified to facilitate discussion and dissemination of results. The aim of this research problem is to identify and define, in concise terms, what HMAR is. A clear description of the different technical and contextual components that facilitate the experience of HMAR needs to be conceptualized.

Problem 2. What content to interact with on HMAR systems

What content do users think will work on the HMAR platform? Different content types need to be evaluated, contrasted, and analyzed to provide an understanding of what content should be made available on an HMAR system. In light of this, the question becomes what content makes sense to communicate through HMAR.

Problem 3. How to research and present findings on (H)MAR

How can we capture data about and analyze the use of MAR technology and disseminate the findings in an appropriate way? MAR systems have specific traits that need to be taken into consideration when gathering data for evaluation. Similarly, conveying and illustrating findings on this platform come with challenges, so a visual language is needed to allow the reader to assess the interpreted data.

Problem 4. Design guidelines are needed for (H)MAR systems

Designing for a relatively new technology such as MAR presents a challenge. General design guidelines based on HCI theories have generally been lacking for MAR. Additional guidelines thus need to be formulated for MAR application developers, so that they can create higher quality solutions.

These problem areas are the motivation behind this research. The goal of the research is to shed light on these four problems and to answer the overarching research question by integrating the findings and analyses from my published papers into this thesis.

1.6 Thesis structure

So far, the reader has been introduced to the field of research and the problem areas addressed in this thesis.

Chapter 2 will present the methodology used in shaping the overall research and the design efforts described in the thesis. In this chapter, a detailed overview of the approaches to data collection and evaluation is put forward.

The following chapter, Chapter 3, provides an overview of the theoretical principles from HCI that underpin the thesis.

This is followed by Chapter 4, which recounts and explains the construction efforts undertaken during this PhD project. It discusses the individual prototypes created during the thesis work.

Chapter 5 provides a summary of the published articles that form the basis for the thesis. The chapter shows the progression of academic work, from an initial publication on the topic of mobile AR, then on to a conceptual work aimed at framing HMAR, followed by the development of tools for evaluation and dissemination, and finally, a publication combining the tools to formulate design guidelines for the field.

In the final chapters, the tangible results from the academic publications, the artifacts, and the overall findings of the thesis are presented and discussed.

2. Method

This thesis makes use of different methodologies and theories in the study of MAR. This section of the thesis presents the following:

- The research framework used for this project.
- A discussion of the evaluation methodology, data collection techniques and analysis used for this thesis.

2.1 Research framework

Design is a central activity in many applied disciplines. As Cross (2001) states, "*The first half of the twentieth century had seen the rapid growth of scientific underpinnings in many types of design - e.g. materials science, engineering science, building science, behavioural science*" (p. 3).

This thesis uses design science research (DSR), as described by Hevner, March, Park, and Ram (2004), as a framework to guide the research process using information systems. The set of guidelines proposed by Hevner et al. has been used to plan and execute the research and development activities done for this PhD project. In this chapter, the motivation and rationale for using DSR is presented, along with an overview of DSR and its guidelines. Alternative approaches for framing the research are also briefly discussed.

2.1.1 Design science research

DSR, according to Hevner et al (2003), must be a process that can be roughly outlined as follows:

- A problem area is identified and formulated.
- An artifact is designed with the purpose of solving parts of or the entirety of a problem.
- The created artifact is evaluated using suitable methods.
- Knowledge gained from designing and evaluating the artifact should then be disseminated.

Hevner provides guidelines to support and guide these core activities. In the following sections, a description of how the guidelines have been used during this PhD project is presented.

Hevner stresses the idea that cutting-edge information systems (IS) enable organizations to change the way they do business, and that this is one reason for doing IS research. The application described in this thesis was, to some degree, aligned with the motives of Verdens Gang (VG). VG—a Norwegian media organization—was a collaborative partner in developing the first iterations of our application (Figure 3). VG saw the potential of AR on mobile devices for increasing ad revenue through engaging augmented advertisements and richer editorial content in their newspapers.

Figure 3: ARad augmenting a printed advertisement in a nationally circulated newspaper. The mobile display in figure shows an animated troll that allows for user interaction.

Hevner states that the goal of design science research is utility. During this study, we have developed an application that provides utility in the form of augmented content for printed media, and which is available on the App Store.

2.1.2 Guidelines

To articulate the different stages of design science, Hevner formulated seven guidelines that support the design science process. Below, details of how each guideline has been addressed in this study are discussed.

Guideline 1: Design as an artifact—Create an innovative, purposeful artifact.

This study is concerned with ARad, an application mediating augmented reality content in newspapers, which was created to facilitate the investigation of HMAR technology. The artifact and its development are detailed in the published papers included in Section 4.3, “Development and design of ARad.”

Guideline 2: Problem relevance—Specify a problem domain

The study aims to understand how to provide a compelling user experience on an MAR platform. The questions guiding this research can be found in Section 1.5 “Problem areas and research question.” A review of previous research can be found in Section 1.3, “Background and previous research.”

Guideline 3: Design evaluation

The utility, quality, and efficacy of the design artifact is documented and discussed in two publications (Papers 3 & 4). We evaluate the design using established evaluation methodologies within the HCI domain (described in detail in Section 2, “Method”). It is worth noting that this thesis is concerned with the usability and user experience of the application. Hence, the methodology—and necessarily, the form of the reported design evaluation—are reflected in this choice.

Guideline 4: Research contribution

Hevner states that contributions from design science can include a design artifact, formalisms, and ontologies or methodologies. In this case, the research contribution is the design artifact itself, which is freely available. The studies also provide conceptual, descriptive, and prescriptive knowledge about HMAR applications. For Hevner, some of this knowledge would simply be categorized as formalisms and ontologies. Hence, a taxonomy was created to enable us to categorize and discuss the nature of HMAR (Paper 2). The third contribution is in methodologies. Two papers (Papers 3 & 4) are dedicated to evaluation methods for the visually dependent AR applications and how to disseminate such knowledge. A summary of the concrete results and contributions of this thesis can be found in Section 6.2, “Results.”

Guideline 5: Research rigor

The methods used for evaluation in this thesis are common within the field of HCI. The methodology behind the design of artifacts is described in the Chapter 4, “Construction.” The design phase of ARad and the artifact itself are described in Section 4.3, “Development and design of ARad.” A current version of the application is readily available for examination in the App Store.

While focusing on the behavioral sciences as the finders of truth in IS research, Hevner and Chatterjee (2010) note that the "… rich phenomena that emerge from the interaction of people, organizations, and technology may need to be qualitatively assessed to yield an understanding of the phenomena adequate for theory development or problem solving" (loc. 7357; see also Klein & Myers, 1999).

The qualitative methods used to gather data are described in detail in Section 2.2, “Data gathering.”

Guideline 6: Design as a search process

The iterative nature of the design process must be taken into account. Iterations of the technological prototypes (in Section 4.3, “Development and design of ARad” and in the Appendix Section 9.1, “Documentation of development”), and proposed taxonomies and ideas are documented in Section 4.5 "Taxonomy development" and Appendix Section 9.4, “Evolving graphics.” It is common to present a final “good” version when submitting an article for publication. While these ideas and versions are naturally the most refined, the process leading to the final ideas can be interesting to look at as well. Documenting the process provides an impression of the iterations leading to the final artifact.

Guideline 7: Communication of research

Some of the published material presented in this thesis is targeted toward the AR academic field. However, one article in particular (Paper 2) is targeted toward a broader audience. In another paper (Paper 3), particular effort was put into discussing how to disseminate the research findings (Paper 3).

Thinking about the knowledge attained through design science as conceptual, descriptive, and prescriptive (Iivari, 2007), in conjunction with Hevner’s guidelines, provides a suitable framework for this thesis. Since the premise of this research was driven by the need to solve a problem using a tangible IT artifact, Hevner’s framework provides the conceptual tools for shaping the research as well as a solid foundation for discussing the different parts of the research project.

2.1.3 Alternative frameworks

> *Everywhere, our knowledge is incomplete and problems are waiting to be solved. We address the void in our knowledge and those unresolved problems by asking relevant questions and seeking answers to them. The role of research is to provide a method for obtaining those answers by inquiringly studying the evidence within the parameters of the scientific method.* (Hevner & Chatterjee, 2010, loc. 747)

DSR is a design-focused framework and methodology that provides the tools to perform research on design in information systems using a scientific approach.

Explicitly design-focused frameworks exist for engineering, architecture, computer science, software engineering, media, art, and so forth. A very closely related approach to design can be seen for software design.

> *Software design sits at the crossroads of all the computer disciplines: hardware and software engineering, programming, human factors research, ergonomics. It is the study of the intersection of human, machine, and the various interfaces—physical, sensory, psychological—that connect them.* (Winograd, Bennet, De Young & Hartfield, 1996)

In their landmark book, Winograd et al. (1996) provide a rationale for designing software with all these factors in mind. However, the reasoning is fragmented, and no overarching framework exists within their methodology. Winograd et al. 's approach to software design fits very well with some research, and different scientific or commercially-oriented software development methodologies can be formulated to focus on all or some the aspects mentioned. In contrast, the DSR framework provides strict guidelines, which ensure a scientific and utilitarian approach from the outset. Winograd et al.'s (1996) software design is flexible, with a range of entry points, be they exploratory, conceptual, or more problem-oriented. DSR, on the other hand, focuses on the creation itself, with the goal of solving a clearly defined problem. Additionally, DSR provides concrete guidelines to ensure that all aspects of the design process are at least considered when researching the design.

2.1.4 Why DSR?

> *The fundamental principle of design science research is that knowledge and understanding of a design problem and its solution are acquired in the building and application of an artifact*. (Hevner & Chatterjee, 2010, loc. 794)

DSR was chosen as a research framework because it acknowledges the artifact itself as a contribution. Simon (1996) defines the artificial, from which an artifact comes, “in as neutral a sense as possible, as meaning man-made as opposed to natural” (p. 4).

This project was planned, from its inception, to revolve around the design of some sort of mobile AR application. When the problem area was identified, the process of building and evaluating an artifact was set in motion. In this DSR project, the problem revolves around creating real user experiences of MAR—having to overcome all sorts of technical hurdles in doing so—to be able to examine and evaluate the end result and contribute to the understanding of the MAR phenomenon.

We believe that artifacts and their descriptions are, in themselves, interesting for the HCI community. This belief has been strengthened by the fact that conferences such as NordiCHI and CHI have recently started seeking submissions for “design cases” and explicitly request descriptive papers detailing the design process of concrete artifacts.

Moreover, DSR provides a coherent framework that gives direction for what steps need to be taken to ensure the quality and contribution of the created artifact(s).

2.2 Data gathering

The data used in the included articles consist mostly of qualitative data gathered during multicamera, video-recorded think-aloud (THA) sessions. This section will argue why this was a suitable approach to collect data about the use of MAR technology.

The THA method for evaluation has a strong history in HCI. Initially developed and used by Ericsson and Simon in 1980, the technique has gradually been adapted for research in HCI. It is a qualitative method that enables a deeper look at IT artifacts, in contrast with the more formal approaches of heuristic evaluation and cognitive walkthrough.

It can be argued that THA is a technique that allows investigation beyond the scope of predefined tasks or potential usability pitfalls. Heuristic evaluation and cognitive walkthrough excel at these points, and are efficient and highly usable in a business environment. They allow assessment of the general quality and effectiveness of an IT artifact. However, the methods do little to shed light on the thought processes behind the users' choices. They do not capture the minute details of interaction. We believe that revealing and analyzing the details of the interaction allows for a better understanding of MAR applications. The details offer insight into the affordances of interface elements and how the user experience can be articulated.

2.2.1 Think aloud

The think-aloud technique is described by Dix, Finlay, Abowd, and Beale (2003) as being "easy to perform and [having] the advantage of simplicity." Dix et al.'s relaxed advocacy for this simple method is discussed in detail by Nielsen, Clemmensen, and Yssing (2002), who refute, to some extent, its simplicity and discuss the consequences for an informal approach to THA.

Preece, Rogers, and Sharp (2007) and Hoonhout (2008) attribute the development of the think-aloud protocol to Ericsson and Simon (1980), while Nielsen et al. (2002), refer to Karl Duncker (1945) as originally describing the outline of the method in use today. Nonetheless, THA has been further developed by other researchers, such as Bennerstedt & Ivarsson (2010) and Plowman & Stephen (2008).

Think-aloud is a well-known method in human computer interaction (HCI) research. Recently, it became a popular tool, both for investigating overall user experience and for pinpointing usability problems. In this thesis, we differentiate between issues related to usability and user experience. Usability problems are typically identifiable and reproducible short periods of problematic sequences of interaction when using an IT artifact. User experience should be viewed as the overall impression held by a user after a longer sequence of interaction, or after the entire experience of the artifact.

Two common approaches to think-aloud are concurrent and retrospective sessions (Ericsson & Simon, 1980). In concurrent think-aloud, the users engage with the object and verbalize their thoughts during its use; what they are doing and what they are trying to do should be vocalized. This approach is suitable for discovering usability issues with the applications. Retrospective think-aloud, on the other hand, seeks to collect a more general view of the artifact post-session, where the researcher might show the recording of the artifact being used to the users themselves.

Some previous studies have used think-aloud as an evaluation tool for augmented reality books and as an observational evaluation technique. Dünser and Hornecker (2007) employed think-aloud to evaluate how young children interact with augmented reality books. Liarokapis and colleagues (Liarokapis, Macan & Malone, 2009a; Liarokapis, Macan, Malone, Rebolledo-Mendez & de Freitas, 2009b) used think-aloud to evaluate and discuss implementation of different interaction methods in AR games.

Nielsen et al. (2002) assert the importance of addressing the potential issues related with the think-aloud technique before, during, and after a session. Think-aloud sessions may take different forms in relation to the topic under investigation. To ensure opportune data during the THA sessions, we addressed the issues and weaknesses of THA described in the literature, specifically those discussed by Nielsen et al. (2002), Hoonhout (2008), Dix et al. (2003), and Preece et al. (2007). Below, we will present the considerations made when performing the THA sessions.

The quality of the recording impacts the analysis of the gathered data. Our approach to this is described in Section 2.2.2, “Video recording.”

Another perhaps equally important task when initiating a think-aloud session is the pre-session instruction and the selection of participants. If the users are unfamiliar with the technology, they may find the session intimidating. Hoonhout (2008) argues for using the application’s target group when performing a think-aloud session. However, doing so is not always feasible, especially in the case of a younger target group for games, since the researcher might want users who are able to withstand the cognitive load of a think-aloud session, who have familiarity with the technology, and who will be available for testing.

Before starting the recording of the session, we informed the users of our research interest and how the THA session would be structured. We showed them how the equipment worked and what would be recorded from the session. By law, we also informed the users of their rights with regards to usage of the data we would gather. Any questions the users might have were addressed, and they then signed a contract that described their rights when participating. This meant that at any point, they could withdraw their participation from the project and have the recorded video of them deleted. This part of the think-aloud session served an additional purpose of easing the participants into the session and reducing nervousness and unease. It is expected that some people may feel uneasy about being recorded on video, and a vital part of a think-aloud is to reduce these feelings prior to initiating the session.

Most AR games are physical by nature. This means that the user must move around and interact with markers or interact with a handheld device that displays the augmentation. For the study of EyePet, we needed to make a tradeoff in how comfortable the users would be in relation to letting them experience the game as close to its intended context of use as possible. EyePet is a game that is meant to be played on the floor in front of a TV. The EyePet was modified for the recording session by moving the TV higher up on the wall and having the participants use a table to interact with the augmentation rather than crawling on the floor. This tradeoff was done mainly based on our own experience with other unrecorded pilot studies. We believed that having the participants crawl around on the floor during the evaluation would negatively impact the study, as they would be highly physically uncomfortable in an already mentally taxing situation.

In addition, one needs to be aware that AR games are notoriously sensitive to tracking. This means that the room where the evaluations were performed needed to be correctly lit, and the markers we used needed to be printed on the correct paper.

As described in Papers 3 and 4, we deliberately performed concurrent think-aloud along with a retrospective think-aloud and an interview session that served as a debriefing in which the users could discuss their overall experience of the games. According to Hoonhout (2008), a retrospective session serves as a setting to discuss the user experience in general. Users are removed from the artifact in question, and in some studies, they watch a video of their own actions and explain their choices.

The strongest argument for using think-aloud in this thesis was its coupling with video recording. Effort went into assuring a synchronized multicamera recording of entire sessions, where both the frame of the applications were rendered and the users' interactions were recorded with a handheld or tripod-mounted camera. This effort allowed us to create figures and visualizations tightly connected to the video data during post-processing.

Augmented reality is, in essence, a highly visual medium that is tightly connected to the context of its use. AR seeks to augment the world of the user through different display approaches. We believe it is material to see how the user interacts with the world where the device lives and where the augmentation lives on the AR device. Therefore, we chose to record the context of use from a different perspective, using a handheld camera while also capturing video from the devices' augmented displays. Our THA sessions may seem to contrast to typical sessions conducted in front of a PC monitor. AR users are forced to use their body to a much greater degree. We find this relation to physical artifacts to be of importance when investigating the use of AR. A significant strength in our method is the extensive high quality recording of this visual and physical medium.

2.2.2 Video recording

Video-based qualitative research has, in recent years, gained an increasing popularity in fields such as computer supported collaborative work (CSCW) and HCI (Heath, Hindmarsh & Luff, 2010). A key advantage to this kind of method is that it can help capture "aspects of social activities in real-time: talk, visible conduct, and the use of tools, technologies, objects and artifacts" (Heath et al., 2010, pp. 5–6). It is also a promising method for studying interaction with context-aware handheld mobile devices (Schmidt, 2013). Nevertheless, certain methodological challenges emerge when studying such devices.

Different approaches to capturing AR on handheld and static devices were employed in this thesis. The approaches were developed to enable the capture of the user's interaction with the environment and with the augmentations appearing on the different devices. In this thesis, we investigate several distinct types of AR interfaces, which require different recording equipment. Depending on the number of users performing the evaluation and the environmental concerns, we needed to consider the comfort for the user, the lighting, and other environmental factors.

Video recording has been employed in the AR field to some degree, mostly to demonstrate the utility of systems. My first trials using video recordings of AR interaction date back to my master thesis (Gjøsæter, 2008) and Paper 1. These evaluations employed video screen capture of the content appearing in a head mounted display (HMD) worn by users, along with a recording from a handheld camera, which captured the users' movements in the evaluation context. We found handheld cameras confer some obvious benefits and some less obvious pitfalls.

The benefit of using a handheld camera is that it enables a directed view of the interaction with paper markers and the environment. When using an HMD-based AR, this may be an appropriate approach, since the users in our evaluations were moving around a lot. However, using a handheld camera does not easily allow quantitative analysis of the data captured. The evaluation may also be impacted, as the handheld camera can interfere with the natural interaction otherwise taking place.

Tripod with fixed-screen capture

This approach is suitable for any AR interface using a fixed camera position. The setup requires a device for capturing the interaction with real world, and a device for capturing the display showing the AR interface.. In our case, we used a recorder that captured the signal going to the fixed display (Figure 4).

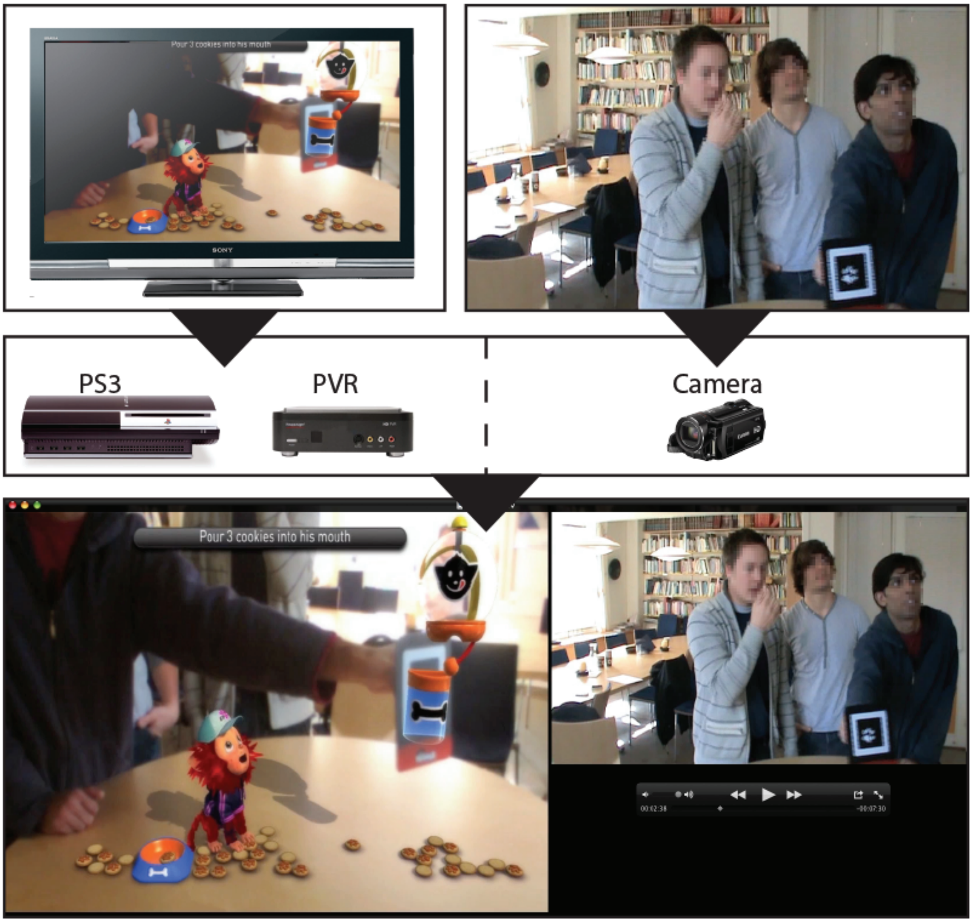

Figure 4: Video screen capture using a recording device connected to a PS3,along with a video of the user interaction using a camera mounted on a tripod.

Tripod with handheld screen capture

Since handheld devices need to be mobile, we used a camera balanced and rigged at an acceptable range to capture the video data. In this way, we could capture the users' hand and finger interactions with the device as well as the augmentations shown on the display of the mobile device. (Figure 5).

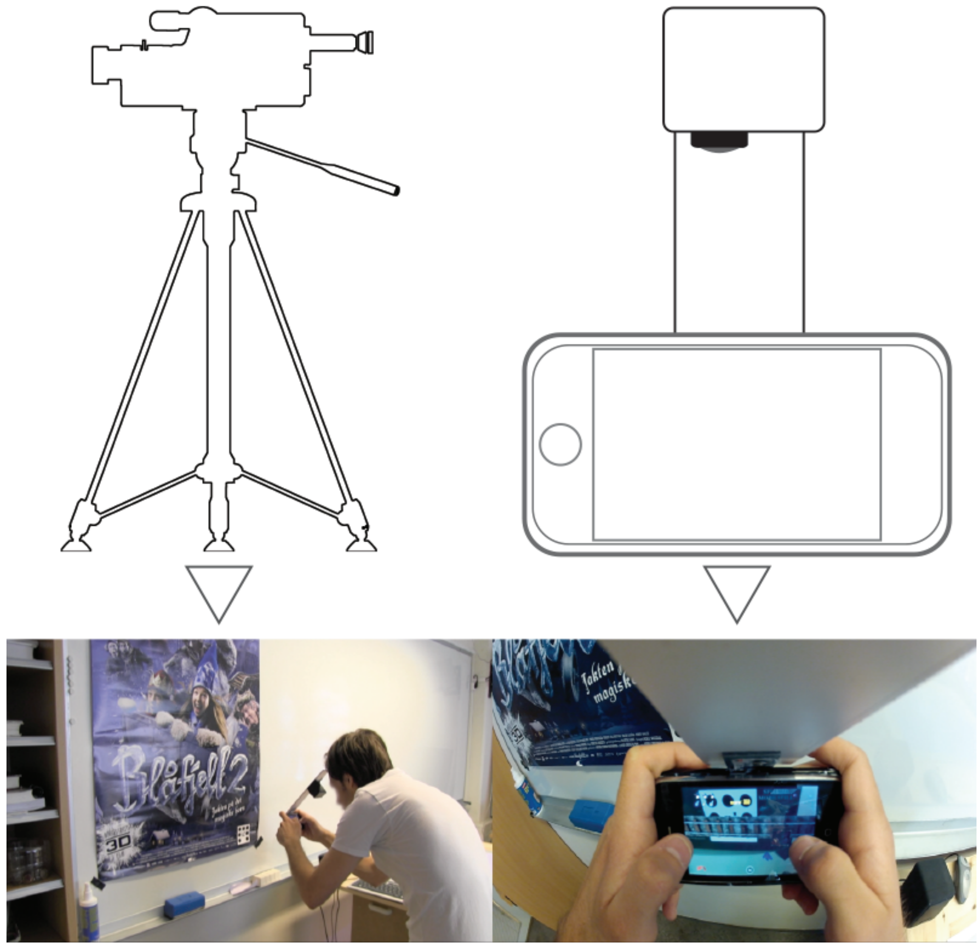

Figure 5: Recording of a handheld device screen using a small camera rigged to capture the screen as well as the user's hand and finger interactions. A camera mounted on a tripod records the user's movement and interaction with the environment.

2.2.3 Quantitative data

We also performed a very small-scale quantitative study after the retrospective debriefing in the think-aloud evaluation of ARad. This data did not find an outlet in a suitable publication. The sample size was small, thus making it difficult to generalize from the data.

To collect quantitative data, the respondents answered relevant questions from a USE questionnaire (Lund, 2001; Tullis & Albert, 2008). The users rated statements on a 7-point Likert scale from “Strongly agree” to “Strongly disagree.” Questions from the “Satisfaction” part of this general questionnaire were presented to the users. Such a questionnaire can provide some indication about the level of satisfaction experienced when interacting with the application.

2.3 Analysis

In *Video in Qualitative Research*, Heath, Hindmarsh, and Luff (2010) describe their approach to reviewing and analyzing video data. They suggest a three-step approach when beginning review of video data.

First, a preliminary review is done with cataloguing of data. The cataloguing should be a simple classification of data, without any in-depth analysis. In the evaluation of ARad, the first cataloguing was simply digitizing the two camera angle recordings and relating them to the different participants in the study (Figure 6).

Figure 6: The raw data from tripod and mounted camera catalogued by respondent in a folder structure. The raw data consisted of video from a tripod mounted camera, as well as a small camera mounted to the device to capture the user’s interaction with it.

Secondly, Heath et al. (2010) suggest performing a substantive review of the data to identify fragments of interesting interactions for further analysis. These reviews are performed “to find further instances of events or phenomena, so as to enable comparison and to delineate aspects of interactional organization" (Heath et al., 2010, loc. 1516).

In this thesis, a categorization was performed to compare interaction with different types of AR content. This is shown in figure 7.

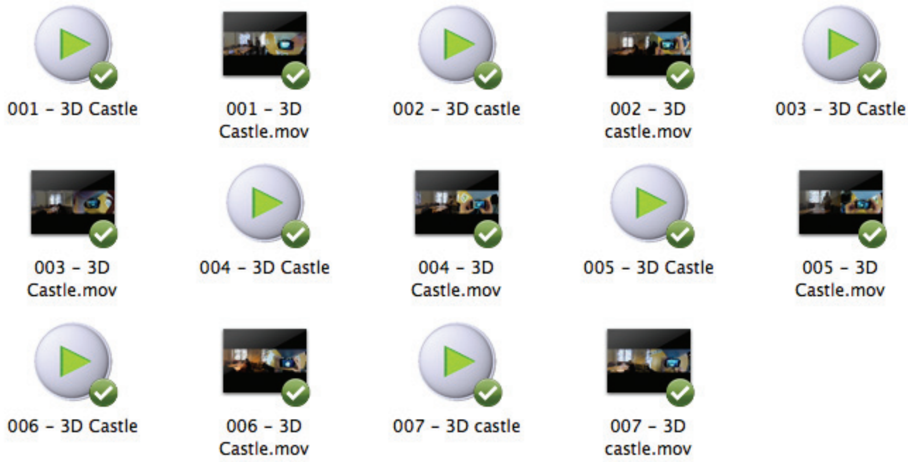

Figure 7: Catalogued instances of interaction with the 3D castle. The figure shows files containing transcripts of these instances from multiple camera video recordings. The numbers in the filenames correspond to the different evaluators.

Heath et al. (2010) further suggest an analytic review of the corpus to identify candidate fragments for transcription and more detailed analysis. To support this process, the candidate fragments were extracted from the catalogued data using Final Cut Pro. The candidate fragments were then transcribed using InqScribe , a tool for video transcription that allows linking with time-codes and export of individual frames for further processing (Figure 8). Heath et al. (2010) use Jefferson´s (1984) transcription system in their work. In this thesis, the features of the tools used (InqScribe and Final Cut Pro) provided the foundation for creating usable and interactive transcripts for use in collaboration with others.

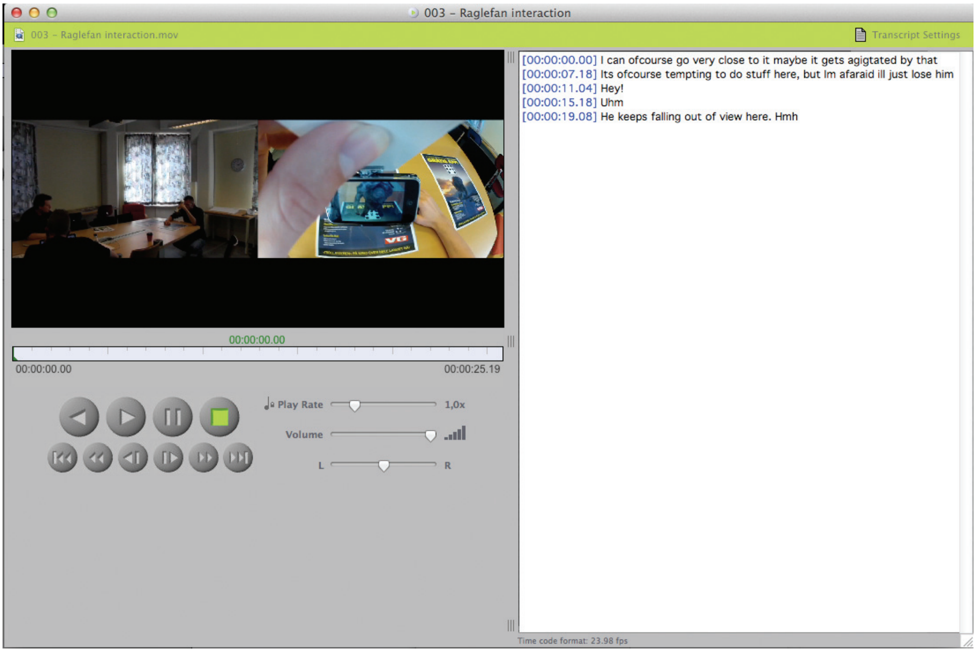

Figure 8: The transcription tool InqScribe in use. InqScribe provides support for video transcription, and has features for hands-free control of video, insertion of time-codes and so forth.

Heath et al. (2010) note that "if you are presenting a particularly fine aspect of conduct it is worth considering how this can best be done for a reader to see the necessary detail" (loc. 2677). Through analysis of the transcriptions and the observable interactions on video, interesting frames were broken down into a comic-strip format to exemplify and visualize the details of interaction (Figure 9).

Figure 9: The transcribed fragment shown in Figure 8 developed into a comic strip.

In the case of Figure 9, it illustrates how one can use comic-strip format a tool for communicating findings in video recordings. Figure 9 illustrates what the user is saying, and how the hands are moved while interacting with the system. In the bottom frame we can see the user trying to interact with the augmentation by moving a finger to touch it.

Corpus

Four instances of video-based studies on the following four artifacts were performed during this thesis work. Greater detail regarding the artifacts can be found in Chapter 4, “Construction.”

1. FurnitAR evaluation (Paper 1)

This artifact was studied using video captured from the HMD. A photo-camera was used during the evaluation to document interesting aspects of the sessions. Two furniture design teams with a total of five participants divided between the two teams engaged with the system for up to twenty minutes each.

2. ArtAR evaluation (Paper 2)

Four evaluations were done in addition to a pilot study. One of the sessions had two users. A total of four dual camera recordings (handheld and screen-capture) were made, ranging from twenty to thirty minutes in duration with an additional debriefing session.

3. EyePet evaluation (Paper 3)

The EyePet data was captured from a stationary screen and tripod (shown in Figure 4 setup. Two sessions took place; the firsts was a pilot test with one user. The evaluation itself was a group of three persons interacting with the game for about an hour, with a debriefing session afterwards.

4. ARad evaluation (Paper 3 & 4)

Data was gathered from one pilot study and seven live sessions. Video was recorded in 1080p from two sources (Figure 5). The participants interacted with the system for 20-30 minutes each. A retrospective debriefing was also recorded. This evaluation was the most extensive of the four and generated a very large dataset.

In summary, the corpus is extensive and varied. The approach to data collection was iterative, with successive improvements made regarding the technical aspects of video capture and post processing. As technology has advanced, multi-camera syncing has become easy, and editing of very large source videos has become feasible through higher available computing power. Importing, editing and exporting tasks that might have been overnight batch jobs in 2009, can now be done faster than real-time on modern systems.

Cultural analytics - Quantitative analysis

Using stationary cameras, both for capturing the interaction of the users and for visualization of the augmentations, enables interesting approaches to analyzing the data. An experiment performed at CalIT2, using their HiperWall, explored this issue. A side effect of using video recordings is the inherent workload required to categorize and transcribe the data. By employing different techniques, one can sort the data and look for patterns. For instance, in the case of the EyePet data, we were able to create an interactive visualization of the entire dataset.

To do this, some straightforward analysis techniques were used to sort the data according to user movements and the actions taking place on screen. This led to the discovery of a simple pattern in the game’s content. The interaction in a typical mini-game in EyePet was first characterized by a very frantic phase showing lots of movement on the user’s part. This was followed by a sequence of dramatic change in the augmentation interface, where the users did not move at all. What we could see from the data was that the users were rewarded with interactive graphics of exploding color whenever a task was completed. Examples of working with video analysis in this manner are provided in the Appendix, Section 9.3.

3. Theoretical principles

This chapter outlines the underlying theories from HCI and new media that were used to frame our approach to MAR systems. The outline focuses on theoretical perspectives of design, post-wimp interfaces, affordance in HCI and remediation from the new media perspective.

3.1 Design

Design is a considerable part of this thesis, and understanding of what entails good design is important in creating usable artifacts. Principles for good design exist, and they can be used to guide the design process. In his book, *The Design of Everyday Things*, Norman (1998) argues for creating artifacts that are intuitive and easy to use. Furthermore he proposes a set of design principles—which are now familiar in the field of HCI—namely, the use of conceptual models, feedback, constraints and affordance to make the design usable, and human friendly.

Others have built upon Norman's work, refined it for use in HCI specifically. Sharp, Rogers and Preece's (2002) *Interaction design - beyond human computer interaction* provides a more formal overview of interaction design and the principles behind it, in an HCI setting.

The work of Sharp et al. (2002) and Norman (1988) place a strong focus on the human aspect of design. This is because artifacts with interfaces available to humans will inevitably be used by humans. Awareness of different approaches to design and evaluation of those designs through observation of those using the technology is thus important in understanding the interaction taking place. In fact, the iterative approach to the design cycle has become the de facto method for design (see Hevner et al., 2004; Sharp et al., 2002). Recently, a shift has occurred towards a focus on designing user experiences, in which the experience of technology is designed to create meaningful and personal experiences (see Hassenzahl, 2013). Hassenzahl (2013) states that "experience emerges from the integration of perception, action, motivation, and cognition into an inseparable, meaningful whole." Designing with all of these facets in mind may seem daunting; however, to achieve meaningful interaction with the artifact, these aspects must be considered.

The theoretical foundation laid by Sharp et al. (2002), Norman (1988), and Hassenzahl (Hassenzahl, 2013) provides an overarching approach to design. This approach includes a mindfulness of concrete human limitations, an understanding of the human aspect beyond cognitive limitations, and employing good practice in the design process.

For this research Hevner's (2004; 2010) approach to design science research was used in combination with the design *knowledge* provided by the above mentioned authors.

3.2 Post-wimp interfaces

Jacob et al. (2008) states that we "are in the midst of an explosion of emerging human-computer interaction techniques that redefine our understanding of both computers and interaction" (p. 1). A great influx of new and easily transformable technology is becoming increasingly available to HCI professionals. A plethora of high-powered handheld devices, programmable micro controllers and a wide range of sensors communicating via uncomplicated protocols have become commonplace and affordable. One example is the Arduino platform, with its wide range of shields enabling wireless sensing, robot control in combination with sensing gyroscopes, temperature sensors, and so forth, which has enabled new ways of interacting with technology.

The field of HCI is moving beyond WIMP (Windows, Icons, Menus, Pointer) interaction into the post-wimp era. As noted by van Dam in his 1997 paper, WIMP does not utilize the full spectrum of interaction possibilities, as it does not take advantage of "speech, hearing, and touch" (p. 3). Beyond this, we now are venturing into interaction via gestures in the air (see Hürst & Wezel, 2012) and brain computer interfaces using EEG (see Wolpaw, Birbaumer, McFarland, Pfurtscheller & Vaughan, 2002), to mention two aspects beyond speech, hearing and touch. These interaction metaphors are still on an individual level, although some attention has recently been directed towards crowd or group interactions. In a study of football crowds, Reeves, Sherwood and Brown (2010) uncovered potential sensitivities a designer should take into account when developing technology for these settings.

Nonetheless, it is important to note that the WIMP metaphor is highly incorporated into most cultures. We see this in the new touch smartphones, where although they have subtracted the mouse-pointer from the equation, remnants of WIMP still remain with the icons, pop-up widgets similar to menus, and so forth. While van Dam (1997) states that “a large percentage of our neurons lie in the visual cortex and vision is our highest-bandwidth information channel,” it is worth noting that the intangible and tangible interaction possibilities offered through the *human presence* still has great potential for new interaction metaphors. New technology is enabling the HCI field to move beyond the WIMP paradigm.

3.3 Affordance in HCI

In this thesis, we are seeking to further the understanding of interaction with mobile augmented reality. The concept of affordance in HCI has, in this sense, provided the foundation for investigating how MAR applications afford interaction with their users.

To provide meaningful knowledge about affordances, we have adopted the sociocultural perspective of Kaptelinin and Nardi (2012). This view, and its borrowing of a “web of mediators” from Bødker and Andersen (2005), allows us to look at affordances in MAR with respect to the cultural aspect. The framework proposed by Kaptelinen and Nardi looks at affordances from an individual perspective with regards to culture, going beyond the scope set by Gibson (1977) and later, Norman (1988).

Previous experience with different content types must be taken into account when designing user experiences for MAR. Doing so requires an understanding of the combination of known usability conventions and conventions of esthetics directly related to the content being presented.

The framework proposed by Kaptelinin and Nardi (2012) allows us to look at affordances from an individual perspective with regard to culture. According to Kaptelinin and Nardi's literature review, "A growing number of studies in HCI and related areas call for re-defining the notion of affordances to include social and cultural aspects of human interaction with the world" (Kaptelinin & Nardi, 2012, p. 970). Kaptelinin and Nardi's approach is called the mediated action perspective, and "is concerned with how humans act in their cultural environments, rather than with how animals act in their natural habitats" (Kaptelinin & Nardi, 2012, p. 971).

If we allow the discussion about affordances in HCI to include these aspects, we can study IT artifacts with the cultural aspect in mind. Such an approach is fruitful for this study, since we are creating an artifact by relying on existing conventions from a wide range of interaction and content paradigms. These include the WIMP paradigm, AR interaction techniques, existing media content, and smartphone conventions.

Bødker and Anderson (2005) identify auxiliary affordances that take into account "complex relations within webs of mediation." This is understandable, as there is often a need to perform indirectly or directly related actions to achieve an outcome. The authors identify maintenance affordances, aggregation affordances, and learning affordances as examples of affordances that technology may employ to achieve its intended purpose:

- *Learning affordances*: What steps the users go through to learn how to interact with an AR application, and how we might improve these affordances.

- *Maintenance affordances*: What needs to be taken care of to allow applications to function as effective mediators (maintenance).

- *Aggregation affordances*: How the application intertwines with analog and digital artifacts to achieve its outcome. Aggregation affordances illustrate the fact that some applications must be combined with other artifacts to achieve their intended purpose. AR's character is to use the environment to provide augmentations, and in this regard, we aim to identify the essential aggregation affordances of an MAR application.

In addition, we propose a domain specific affordance—*remediation affordance*. This affordance encapsulates the affordances associated with content, and specifically, how we re-communicate the existing affordances of content through an interface. In our case, remediation affordance deals with bringing the content from a known interface to a different medium.

3.4 Re-interfacing

Bolter and Grusin (2000) see augmented reality as a hyper-mediated visual space and "the insistence that everything that technology can present must be presented at one time—this is the logic of hypermediacy" (p. 269). This definition is fitting, since augmented reality seeks to enable us to see the world how we want it to be, with just the right amount of information. Macintyre and colleagues (2001) argue strongly for AR being a new media experience, mainly because its "fluid blend of the physical and the virtual, and the inevitable tension between them, offers rich dramatic possibilities that are impossible in any other medium" (p. 2).

AR technology promotes the ideal of a pure and fully transparent user experience, where the interaction with the medium is transparent to the user, and only the experience of the content remains. Although some transparency is achievable on a handheld device, Rosenblum et al. (2012) believe that "if AR is to realize its full potential, hand-held form factors, despite much of the hype they are receiving now, simply are not adequate" (p. 445). We agree with Rosenblum to some degree. The handheld form factor can never create the same immersion that a totally transparent AR solution blending the real and virtual seamlessly can. On the other hand, even with all its hype, MAR will have utility until this paradigm shifts.

Pavlik and Bridges (2013) provide an overview of the general emergence of mobile AR in journalism. Content creators for the applications they identify (Aurasma, Layar, Wikitude World Browser, Junaio, and Blippar) may benefit from a further understanding of how mediators unknowingly force the continuation of old practices when communicating content through new media.

Regarding interactions, Bolter, Engberg, and Macintyre (2013) argue that the field of media studies is not responsible for providing design guidelines for the HCI community. However, as content and interaction are becoming increasingly intertwined, one might look to polyesthetics (Engberg, 2012) to shed light on aspects of new media. Polyesthetics suggests that media studies should be concerned with contributing knowledge about the esthetics of content, since the interaction between content and interface creates the user experience.

Content in direct relation to the interface provides the complete user experience. Poor content in an application that otherwise has an excellent user interface will deliver a poor user experience, and vice versa. This is of paramount importance in any MAR application, as the content to some extent becomes the user interface itself. To experience the entirety of the 3D content in a correct manner, as the author intended, requires significant effort from the user in the same manner that reading a text does. Additionally, it requires thoughtfulness on the parts of the author and designer regarding how meaning can be conveyed through this interaction paradigm. As put forth by van Dam (1997), “Point and click, the hallmark of WIMP GUIs, has become part of modern culture” (p. 54). While this de facto standard of interaction is safe and is certainly something that users can effortlessly understand, MAR’s qualities allow novel approaches to interaction. When investigating gesture-based interaction in MAR applications, Hürst and Wezel (2012) note that common “interaction concepts for such applications are often limited to pure 2D pointing and clicking on the device’s touch screen” (p. 1).

In the spirit of remediation, we propose the term “re-interfacing” to capture this idea. When re-interfacing traditional interaction metaphors through a new interface, the re-interfaced system may not effectively use the capabilities inherent to the new interface.

4. Construction

This section will describe the theories applied when constructing the different artifacts. The development of artifacts as applications is described in the following sections; the development of the taxonomy found in Paper 2 is described in Section 4.5, “Taxonomy development.”

The development for the different artifacts can be seen as iterative and user centered. In this section, we will detail the development of ARad, as this is the artifact that should be of the most interest to the research community and that is publicly available. FurnitAR (Paper 1) and ArtAR (Paper 2) did not have a long iterative lifecycles with incremental improvement, and the applications existed only in an experimental setting. However, we try to provide sufficient detail to understand the design of these applications as well, in the following section.

4.1 Early MAR prototypes

In addition to ARad, in Paper 2 we mention the evaluation of an application named ArtAR; details of the development of that application can be found in Tone Nordbø’s master’s thesis (see Nordbø, 2011). While the application evaluation was a joint effort, the design of the artifact in question was done mainly by Nordbø. Lastly, in Paper 1, an artifact named FurnitAR is described. Figure 10 below categorizes the different applications based on what type of MAR system they are.

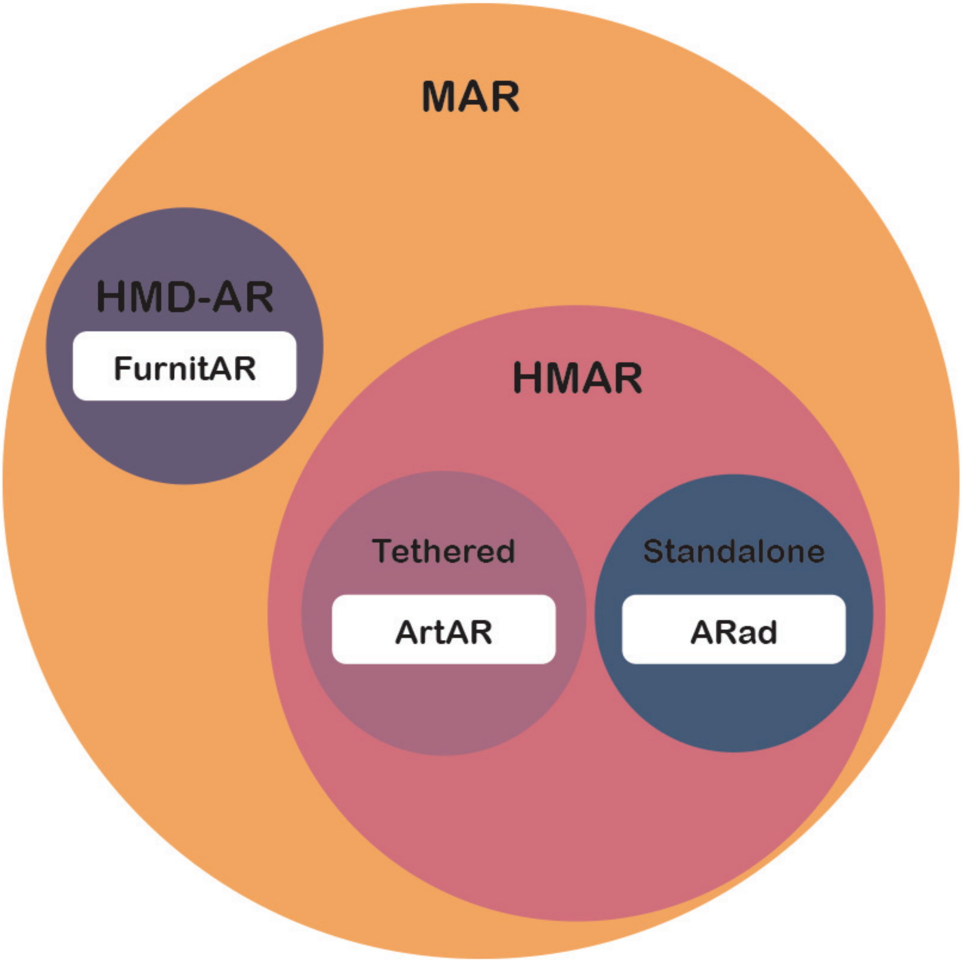

Figure 10: The different applications evaluated during this thesis work. Two of the systems are handheld AR systems, one standalone and the other tethered, and the third system is a tethered, HMD-based AR application.

Figure 10 shows how the technologies supporting the different systems differ in configuration. In FurnitAR, an HMD tethered to a laptop with a long cable was used to allow the users mobility when investigating digital furniture prototypes. The ArtAR prototype (seen in Figure 11) was a tethered HMAR prototype. Similarly to FurnitAR, the software was run on a laptop, while providing augmentation on a handheld device.

Figure 11: The tethered HMAR system used for evaluation of ArtAR. A tethered handheld display with a mounted camera was connected to a laptop running the system.

ARad was the last system created during this thesis work and is a standalone HMAR system that runs on a modern smartphone.

4.2 User-centered development

A user-centered design (UCD) approach was employed in the development of all applications. As this is a design research study, we are interested in how the users interact with and experience the artifacts we design. Hence, it is appropriate to choose UCD as the overarching development philosophy.

We involved users at different stages of development for all applications to help inform design choices. UCD involves the philosophy of creating technology that is adapted to users, rather than forcing users to adapt to new technology. In 2010, UCD was formalized into an ISO standard (DISISO, 2009). The standard describes six principles that will ensure a user-centered design process:

1. The design is based upon an explicit understanding of users, tasks and environments.
2. Users are involved throughout design and development.
3. The design is driven and refined by user-centered evaluation.

4. The process is iterative.
5. The design addresses the whole user experience.
6. The design team possesses multidisciplinary skills and perspectives.

"There is a spectrum of ways in which users are involved in UCD but the important concept is that users are involved one way or another." (Abras, Maloney-Krichmar & Preece, 2004)

Across the different projects, users were involved during predevelopment, early prototyping, and formal testing of the resulting application. However, user involvement during gathering of requirements for the system differed somewhat between applications. In any case, real users were always involved in the testing and evaluation of the prototypes and finished applications.

FurnitAR went from an idea to prototype with user input only regarding technical requirements. These requirements revolved around support for different 3D formats used by the furniture designers. Subsequently, the design was evaluated to verify and understand how furniture prototyping with AR could be done.

ArtAR was designed based on requirements provided by an artist who wanted to explore AR as an artistic tool. An interview detailed in Nordbø's thesis (Nordbø, 2011) provides design rationale for and idea behind the development of this artifact.

ARad had a very iterative lifecycle, with input from collaborators and from formal and informal user studies. The inception of ARad came from the problem of creating an augmented reality viewer for different types of content in a newspaper context. The design stemmed from work with experienced AR developers, Norwegian and French 3D artists, and the newspaper outlet interested in trying out the technology.

4.3 Development and design of ARad

The knowledge gained from developing and evaluating other AR applications led to the development of ARad. In the following sections, we document the different development stages of ARad.

It is worth noting that the process was iterative. Some of the features we implemented in the early version were removed and later reintroduced as we tested their intended functionality.

Development of ARad started in 2009. A range of systems development kits (SDKs) were tested to identify a suitable environment for developing a mobile application that could provide dynamic AR content for print media. As our research interest was in the area of user experience, it was decided to leverage existing technology to develop the application. Thus, rather than trying to develop a more efficient or completely novel tracking algorithm, the focus was aimed at creating a usable application for distributing and interacting with AR content. The SDK we used was developed by a French firm, Int13. It was easy to use and allowed for rapid prototyping and deployment on several platforms.

The SDK, coupled with the existing technology, forced us to be creative to meet the functional and nonfunctional requirements for each subsequent iteration. Early prototypes were developed for the Symbian and Windows Mobile platforms while we waited for iOS support in the SDK. When the SDK finally supported all three target platforms in the fall of 2010, the first version of ARad was released for Symbian, Windows Mobile, and iOS simultaneously.

There were five subsequent versions of ARad: a development and prototype version used to illustrate the concepts to VG; the first public release with offline content; an online version of the same application; an upgrade of the content system; and finally, an upgrade to the 3D engine itself.

4.3.1 Development version

The development version served as a proof of concept technology to gain interest from VG for developing AR to support their newspaper. In one iteration, we showed how one might use the AR technology to advertise using 3D models of cars (Figure 12) using printed markers (Figure 13) in a newspaper.

Figure 12: How a 3D car model might look using MAR with a newspaper. This prototype was used an illustrative example to introduce HMAR to VG.

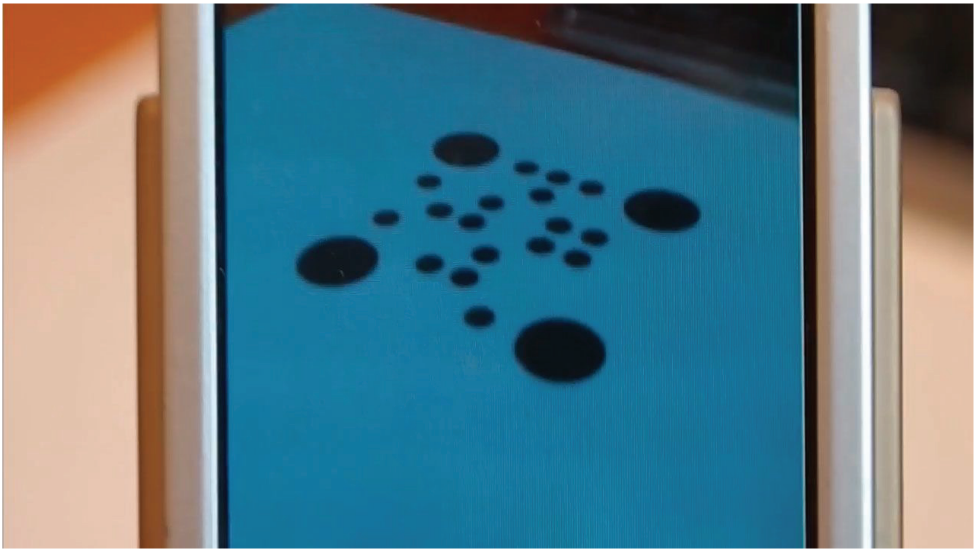

Figure 13: The esthetic of the marker in the prototype version. Further improvements in the SDK changed how the markers looked in the subsequent versions.

The development version was coded in Lua and had no formal requirements. It was used as a tool for introducing and discussing the use of AR in relation to print media from VG. Buchenau and Suri (2000) strongly argue for using experience prototyping when a designer wants to communicate an idea or a technological concept. A very high fidelity prototype (Figure 12) was used to "demonstrate context and to identify issues and design opportunities" (p. 425). At the time (in 2009), AR was not a widely understood, used, or recognized technology, and thus a hands-on approach was needed to illustrate the technology's potential.

4.3.2 First release

A first release for Symbian, Windows Mobile, and iOS was delivered in the fall of 2010. This version did not support the downloading of new content and was shipped with preloaded 3D models and animations. The content consisted of a Raglefant and a Jotne (both interactive trolls) on a marker (Figure 14). In addition, a small car chase game was included to provide a more interactive experience (Figure 15).

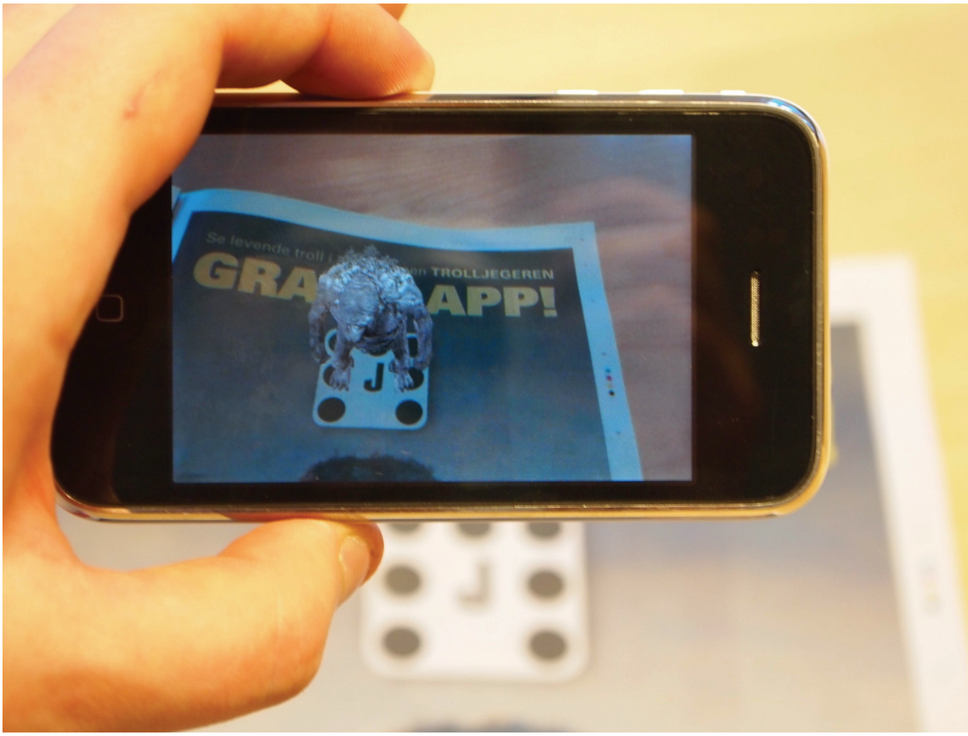

Figure 14: First version of ARad. This version supported two types of trolls and a car chase game. In this figure, a jotne (en. Jötunn) troll is standing on a marker, breathing.

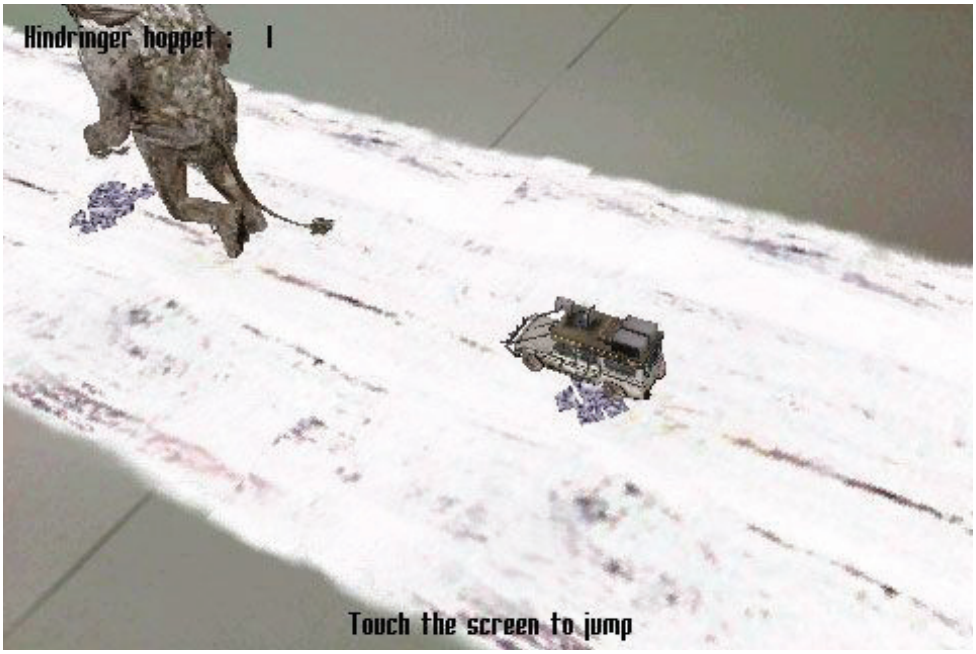

Figure 15: Early version of car chase game. The game revolves around avoiding obstacles while being chased by a jotne.

This version was meant as a proof of concept for AR on a real-world stage. Our collaborators wanted to see whether the technology worked across different platforms, whether the system for distribution was functional, and so forth. The development process was hectic and required several rounds of content modifications provided by the 3D content creators. Nonfunctional and functional requirements were not set outright. However, several functional limitations in the SDK placed a specific limit on the amount of polygons and the texture size of the 3D content. Additionally, this prototype was written with a wrapper script (Lua); the features of the scripting language limited the functionality in the game. A custom method needed to be implemented by the SDK providers to allow the Raglefant to look at the camera.

4.3.3 Second release

This release allowed the downloading of new content to the device on demand. This meant that we had to implement a GUI to handle the loading of content from different content providers, and an infrastructure to distribute markers was implemented (Figure 16).

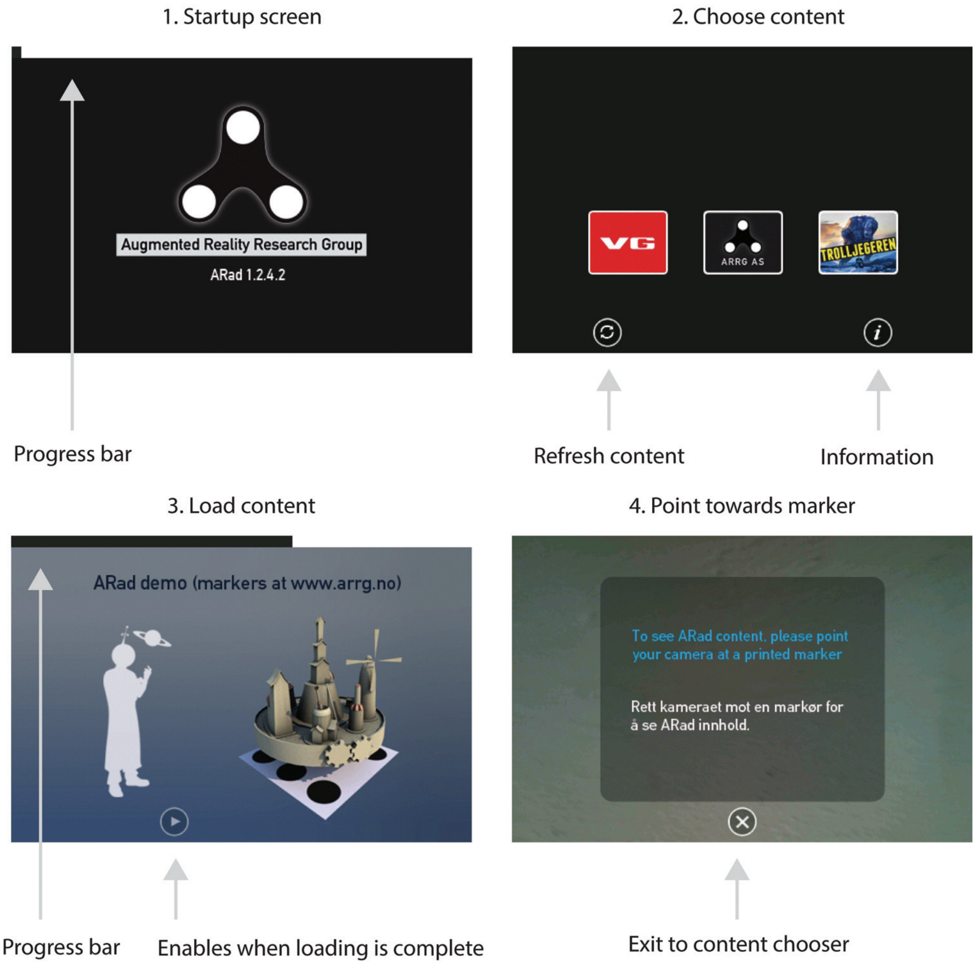

Figure 16: Interaction flow in the application. First, the user starts up the application (1. Startup screen). Second, the different icons for the content providers are presented (2. Choose content). After the user chooses a content provider, the content loads into memory while the user watches a loading screen (3. Load content). When loading is complete, the user can press play to look for markers.

Figure 17: An image slideshow on a marker. The user can use the buttons in the lower corners to change the image showing on the marker.

At this stage, the development moved from an informal prototype approach to active engagement in developing the functionality to augment newspaper content.

System requirements:

- The user should be able to select content from a content provider.
- The content providers should be able distribute a range of content such as image slideshows (Figure 17*)*, video sequences with audio, static 3D models with audio, and interactive, animated 3D models with audio.

Implementation:

Since the LUA scripting language used in the first release did not support the functional requirements for the next release, we implemented it using the C libraries underlying the LUA script language.

The system works by first downloading an index .pak file containing icons and an XML-like file specifying the different content providers. These content providers link to further .pak files containing the actual content. The content is loaded and parsed when the user selects a content provider.

```
typedef struct
{
    CHAR * imagePtr;
    CHAR * model;
    CHAR * author;
    CHAR * caption;

}SLIDESHOW_IMAGE;

typedef struct
{
    CHAR * marker; //Denne mÅ vekk
    U32 loadedImages;
    SLIDESHOW_IMAGE images[16];
    U32 currentImage;

}SLIDESHOW;
```

Figure 18: Using structs to organize different content types. The example above shows a slideshow (SLIDESHOW) containing a set of slideshow images (SLIDESHOW_IMAGE).

Figure 18 illustrates how structs are used to link up the internal structure of ARad in relation to how content types are represent in code.

4.3.4 Third release

The third release of ARad received a major content update as well as bug fixes. At this stage, it had become clear what we wanted to investigate, namely, the user experience of different types of content. Thus, content was created and the application was adapted to work with the different types of content. Additionally, we added some custom game content (Figure 19) as well as support for movies in h.264 format.

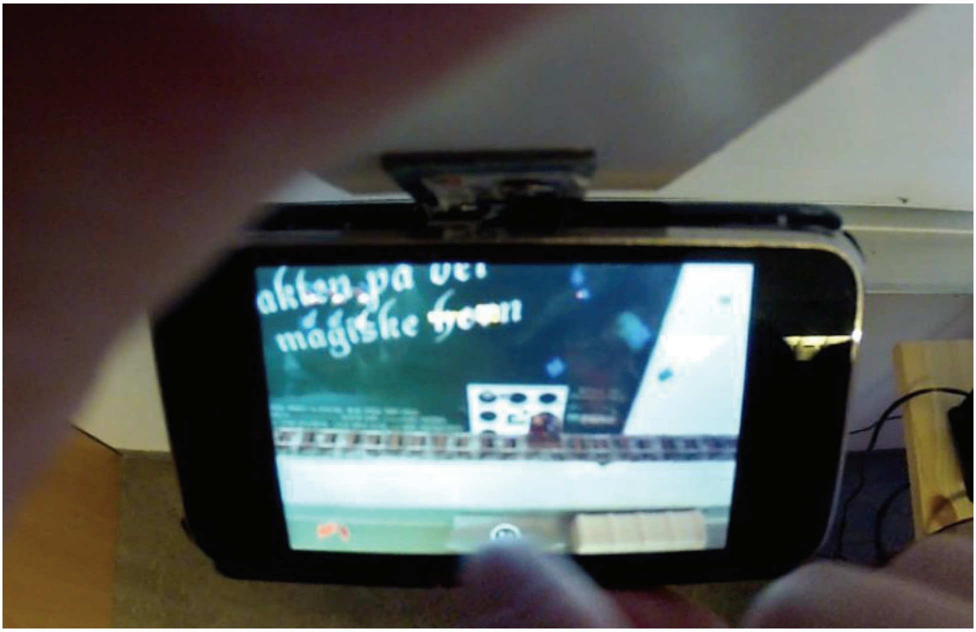

Figure 19: A new mini-game was added in which the player maneuvers a carriage to collect points.

The new game content consisted of an interactive mini-game based on a children's movie, in which the player maneuvers an AR carriage to capture different elements to achieve a high score.

4.3.5 Fourth release

In the fourth release, the 3D and tracking engine received major updates, and these were brought into the application and put on the App Store. The move from software-rendered surfaces to an OpenGL implementation made a huge impact on performance as well as on the user perception of the 3D models themselves.

The five development stages encapsulate the development of the application’s technical aspects. As mentioned previously, there already exist quite a few papers describing different methods of tracking, implementation of 3D engines, and so forth. However, as this thesis is focused on the methodological aspects of evaluating these systems, the user experience of these systems, and conceptual knowledge about HMAR, we do not focus on the technical aspects of ARad in any great detail.

Table 1: A comparison of the different versions of ARad

	Purpose	**Content**	**Rendering**	**Publicly available**	**Video**	**Supported Platforms**
Development Version	Technology demonstration	Pre packed	Software based	No	Low resolution MPEG	Symbian
First release	Feasibility demonstration	Pre packed	Software based	Yes	Low resolution MPEG	Symbian, Windows Mobile, iOS
Second release	Support dynamic content	Downloadable	Software based	Yes	Higher resolution MPEG	iOS
Third release	Content update	Downloadable	Software based	Yes	Higher resolution MPEG	iOS
Fourth release	System update	Downloadable	Hardware accelerated	Yes	h264, MPEG	iOS

Table 1 shows the characteristics of the different versions. From the first release onward, the system was available to the public and in use on different platforms. The table shows that the second release supported downloadable content through an FTP infrastructure set up to provide the content. In the third release, a major overhaul of the content and the architecture supporting downloadable content was done. The fourth release presented hardware accelerated 3D content, as well as native (and arguably better looking) .h264 video content. It is worth noting that we initially wanted ARad to run on Symbian, Windows Mobile and iOS simultaneously. However, the effort required to maintain and distribute to three different platforms was not feasible and the least used platforms were discontinued.

4.4 Design as a search process

Through development, we came to agree with Hevner’s statement that design should be a searching process. While the four release versions of ARad were represented in four identifiable versions in the App Store, some features did not make it into the publicly available versions. The development process was interesting in itself, as a search process for learning more about standalone AR on mobile devices.

Regretfully, the early development of ARad was not documented in a particularly extensive manner outside of an SVN repository. However, some documentation did take place, and images with short textual description are included in the Appendix as a historical document.

4.5 Taxonomy development

The taxonomy we present in this thesis is constructed to serve as a tool for discussing MAR in general. As discussed briefly in Section 1.3 “Background and previous research,” there have been endeavors to frame AR in different contexts.

Milgram’s continuum (Milgram et al., 1994) is often used to place AR applications on a spectrum between a real and virtual environment. This may provide a notion of where an application fits within the field of mixed reality, but the application’s focus may be lost in a discussion about its level of ‘AR-ness.’ Thus, while the continuum is useful to discuss the general qualities of a system, it is limited to a broad view.

Olsson and Salo (2012) simplify the categorization of MAR into two types: “AR-browsers” or “image recognition-based” AR applications. Zhou et al.’s (2008) five categories for AR research in general are (1) Tracking techniques, (2) Interaction techniques, (3) Calibration and registration, (4) AR applications, and (5) Display techniques. Dubois, Gray, and Nigay’s (2002) model for framing AR development provides broad conceptual knowledge that can be used for the purposes of development.

The above-mentioned concepts and categories can be used in the broad categorization and framing of AR artifacts. However, they are not suited for illustrating and discussing the identifiable parts of an MAR system in particular.

A rough outline of the idea behind the taxonomy was proposed at a workshop at Chalmers. The proposal was grounded in a literature review of publications covering MAR, HMAR, and AR (see Paper 2). Figure 20 shows the first draft of the visual illustration:

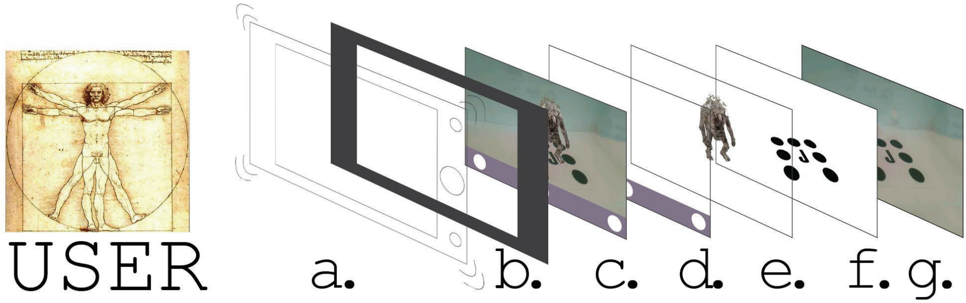

Figure 20: An early idea for a taxonomy of HMAR interfaces. The model has several layers to describe the HMAR interface in great detail.

In Figure 20, the layers are categorized as follows:

User: The user of the AR application

a) The top level of the AR interface; gives audible or tactile feedback

b) The ergonomics of the device performing AR

c) The combined expression of the AR application

d) The screen interface element

e) The tracked augmentation

f) The subtracted AR field of view (CV)

g) The world where AR lives

It became apparent that c) was a redundant layer that caused confusion. The model was refined to that shown in Figure 24, which provides a more coherent visual representation of a taxonomy of HMAR interfaces. The final proposed model is discussed in depth in Section 6.1.1, “The conceptualization of HMAR.”

5. Summary of articles

The four published articles can be seen as addressing different problem areas. When referencing the papers in this thesis, they are referenced by number.

Paper 1: Computer Supported Collaborative Design using Augmented Reality

Paper 2: A taxonomy of handheld augmented reality applications

Paper 3: Combining Think Aloud and Comic Strip Illustration in the Study of Augmented Reality Games

Paper 4: Affordances in mobile augmented reality applications

These four papers follow a natural progression. Paper 1 can be seen as an early attempt to design and evaluate an AR artifact. After this work was published, it became clear that a framework for discussing AR and HMAR was needed to clearly define what HMAR is. Paper 2 attempts to provide a clear definition of HMAR, what it consists of, and how to identify interesting areas for further exploration. In Paper 3, a rigorous effort is made to create a toolset for evaluating and presenting findings from highly visual mediums such as HMAR and AR. Finally, Paper 4 utilizes the tools and experience from the previous publications to contribute concrete guidelines for the design of AR artifacts.

5.1 Computer supported collaborative design using augmented reality

This article is a usability study for a prototype MAR system called FurnitAR (Figure 21). The system allows for digital, collaborative, and rapid prototyping of furniture early in the design process. A usability study was performed using video recordings and think-aloud sessions. We found that the lack of the physical aspect a designer gets from using cardboard prototypes hurt the user experience. However, the efficiency and agility achieved by prototyping in AR was impressive.

Figure 21: A screen capture from a head mounted display during evaluation. The screen capture shows a collection of virtual furniture. © 2009 IEEE

5.2 A taxonomy of handheld augmented reality applications

During my research into handheld augmented reality, it became clear that a model of HMAR systems would be beneficial, as this would enable discussion of the different conceptual components of HMAR systems. The proposed taxonomy (Figure 24) breaks down a typical HMAR application into distinct parts. Layers in the taxonomy allow researchers and professionals within the field to discuss the different parts and relate their research contributions in the field of HMAR to tangible parts of the system. A literature review was performed to assist in making the categorizations.

5.3 Combining think-aloud and comic strip illustration in the study of augmented reality games

This paper describes an approach for evaluating highly visual IT systems and disseminating the findings. With the advent of augmented reality games on handheld and entertainment system platforms, we saw a need to develop a methodology for capturing data as well as disseminating these findings. The methodology for capturing the visual interaction involves video recorded from multiple angles combined with the established usability evaluation methodology, think-aloud (THA). The recorded data is presented in a comic strip format (Figure 22) for easy visualization and dissemination.

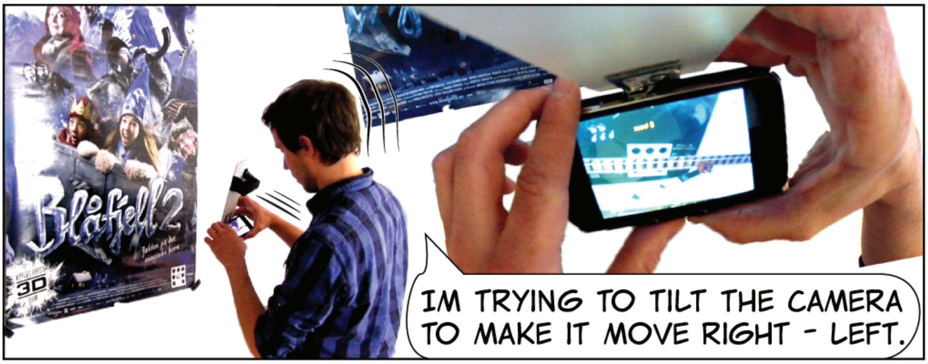

Figure 22: The comic strip format presented in the paper. The format is useful for conveying findings from video data in text.

5.4 Affordances in mobile augmented reality applications

While evaluating the IT artifact created to visualize the content of the HMAR application, it became clear that we were engaging in remediation. This paper describes how we remediated existing content and 3D content (Figure 23) into a new media paradigm—HMAR. It describes how the remediation was performed in the ARad application and how affordances that were directly related to the remediation of traditional AR content were presented . Additionally, we describe affordances related to maintenance, aggregation, and learning, and design guidelines for this particular kind of system are presented to the reader. Finally, we argue that AR content more or less becomes the interface. Hence, it is crucial to acknowledge the esthetics of AR content in relation to the user interface when designing HMAR systems.

Figure 23: A figure illustrating a finding from the paper. The user moves around the marker using his body to interact with the content.

6. Results, findings and contribution

In this section, the results are presented, categorized, and discussed in relation to the associated problem area defined in the introduction. The different artifacts are discussed in terms of the research question and the design research methodological approach. A categorization of the results regarding the production of knowledge expected from DSR studies is presented. Finally, a discussion of the results and a critical view of the research are put forward.

6.1 Problem area

As stated in the introduction of this thesis, we seek to answer four problems directly related to mobile augmented reality interfaces. In this section, these problems are discussed by referencing the published papers included with this thesis.

6.1.1 The conceptualization of HMAR

Practical handheld mobile augmented reality systems should be standalone (untethered) systems providing some sort of augmentation of the real world. In Paper 2, a taxonomy is presented that encompasses the general elements in an HMAR system (Figure 24).

The taxonomy breaks down the interconnected part that constitute a whole HMAR system. The model enables discussion of how these parts work together to create the entirety of the user experience.

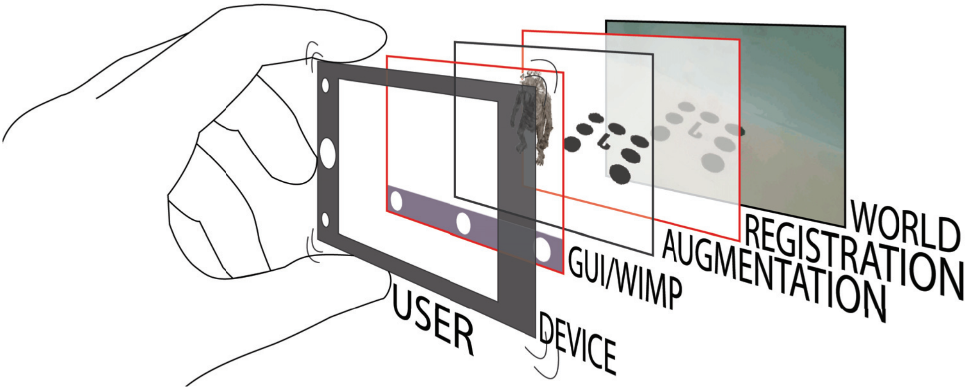

Figure 24: The taxonomy of HMAR presented in Paper 2. The "world" layer shows a simple marker on a white A4 sheet of paper. No additional graphics surround the marker. The "registration" layer is illustrated using the marker subtracted from the background; this is a familiar reference in the AR research community. A 3D model of a troll figure projected in 3D represents the "augmentation" layer. Some simple buttons and a frame represent the "GUI/WIMP" layer in the model. The illustration shows the device with a display, including buttons and vibration. A hand symbolizes the "user" input and interaction with the HMAR device and interface. The hand is simply an illustrative example; it is meant to express the entire spectrum of interaction possibilities available to the user (vision, touch, speech, and so forth). © 2012 IEEE

The taxonomy breaks down the HMAR system into the following parts: world, registration, augmentation, GUI/WIMP, device, and user. This allows us to answer the question of what an HMAR system is. An HMAR system is a handheld device that uses sensors to register the world. Beyond marker-based or natural feature based tracking, Huang (2013) describes other approaches to sensing in MAR. Pose tracking for AR can be performed using wireless networks and can be inertial by means of accelerometers or gyroscopes, magnetic via magnetometers, electromagnetic by, for instance, GPS, or ultrasonic. This tracking from sensor data is used to overlay the world with augmentations in 3D and to provide an interface that allows the user to interact with these augmentations. Interaction happens either through the physical device itself, or more commonly, through a custom GUI or classical WIMP metaphor.

One example of the utility of the taxonomy is shown when discussing HMAR. By using the taxonomy, one can discuss how the different parts affect user experience. The relationship between the separate parts creates the HMAR user experience. Different researchers are interested in distinct parts of the HMAR system. One example is the research on the Vesp'r system (Kruijff & Veas, 2007). The main focus in this research is not on improving registration or augmentation; the aim is to evaluate the physical ergonomics of a device. Registration, augmentation, and other aspects are only briefly discussed. How the device feels and can be manipulated in the users' hands is the topic of interest. In contrast, a study by Klein and Murray (2008) has a narrow focus on compositing for small cameras, looking directly to how the augmentation appears on a screen. Different technical areas of video streams, shaders, etc., are explored to create an augmentation that blends in with the camera stream capturing the world.

During the thesis work, the taxonomy was used when disseminating knowledge to laymen, students, and other researchers. The taxonomy can be used as an introductory tool to illustrate the different technical aspects of an HMAR system or to provide a topic for discussions.

The taxonomy has seen real-world use in disseminating the findings from ARad development in a commercial and research context. From a commercial point of view, the taxonomy has been used to explain why some functionality cannot be implemented in a reasonable timeframe. In the case of ARad, we had very limited access to the camera vision module in the SDK. When the client wanted markers to look different, the taxonomy was used to explain how the markers in the world relate to the registration. The SDK allowed great freedom in creating and modifying augmentations as well as the GUI/WIMP. However, the look and feel of markers was a given, and they could not be easily modified.

6.1.2 What is meaningful content for HMAR

During the thesis work, several prototypes for experiencing AR content were created. These prototypes have been evaluated, and new prototypes have been built iteratively. During a conversation with a colleague, the term "remediation" appeared in the context of discussing a prototype. A book by Bolter and Grusin (2000) was suggested to enable an understanding of why these prototypes had the content they did. In this book, *Remediation - Understanding New Media*, the concept of remediation is proposed. The book puts forth the idea that when progressing from one media type (e.g., radio) to the next (e.g., TV) the old media brings with it its traits, thus shaping how the new content appears. As an example, early TV broadcasting often took the shape of radio, but with the addition of the features of the new media. For example, early broadcast television had a tendency to feature already well-known radio personalities in front of the camera, albeit not in a radio studio but in a more esthetically pleasing setting. These personalities would have conversations with guests, very much as they did on radio, while features of the new media were added in, such as a visually pleasing intro sequence.

This tendency became apparent in the functional requirements for the subsequent versions of ARad. The media institution we were collaborating with required that the visual content they already possessed—images and video—should be presented in ARad.

In Paper 4, a study of different types of content in ARad was performed. The evaluators tried out three different content types: 2D static image slideshows, 2D video (with audio), and 3D content. The study found that some content types may be more suitable for the MAR platform than others. The users found their experience of the image content to be the least interesting of the three types. The users liked seeing videos on the markers, while the content they found most engaging was the 3D content.

In light of the remediation concept, the argument we put forward as to what content should be presented on a MAR platform is to use AR for what it is good at. AR as a visual tool excels at augmenting the world.

In the case of displaying images, it can be done in a very satisfactory way using existing print techniques. In our case, the image content suffered from having lower resolution than normal printed images.

Videos, on the other hand, were found by the users to be interesting and engaging. Embedding videos physically within print media is, thus far, an impractical endeavor; it can be done using bendable OLED (organic light-emitting diode) screens (Cheon et al., 2006), but would be very costly on a large scale. Using AR to provide users with video snippets was demonstrated by our study to be feasible and enjoyable.

The content the users found most interesting and enjoyable, however, was the 3D content. This content type allows new ways of engaging with print media. It can provide information not accessible through 2D means.

To provide meaning in this context, we must present the users with meaningful content. Meaningful content on HMAR is content that is made richer by the possibility of interaction of the system. Content that does not make use of these possibilities, but rather tries to poorly mimic existing conventions of remediation, may be less meaningful than content that enables rich interaction.

6.1.3 Performing research on MAR

In this thesis, we propose video as an invaluable tool for gaining an understanding of MAR applications. Several characteristics of MAR applications make video suitable to use as a data gathering tool. The model described in Paper 2 can be used to illustrate how video can provide insight into interactions with an MAR system. Depending on the area of interest within MAR, observing and analyzing how the user interacts with and understands the world, the device, the augmentations, the registration, and the GUI in an AR context may be of interest.

Using video capture allows minimal interference with the users' interactions and allows detailed analysis of interaction post-evaluation. However, to capture satisfactory data requires technical and methodological considerations. Paper 3 and the Section 2.2 "Data gathering" describe in detail how we performed the evaluations.

The following steps, which are presented in Paper 3, describe an approach to multicamera recording for researching AR.

Pre-session:

- Prepare and test recording equipment (video streams and audio)
- Manage lighting and physical constraints to minimize impact on session
- Hold pilot sessions to identify topics of interest for subsequent session
- Prepare questions or questionnaire for retrospective session

Session:

- Inform users about recording equipment
- Inform users of research interest to prevent off-topic data
- Encourage users to use Think Aloud and not to wander off topic

Post-session:

- Debrief the participants
- Label and archive data

Postproduction:

- Combine video streams
- Identify interesting sequences
- Transcribe sequences in detail

Presentation:

- Analyze transcribed sequences for frames of reference
- Stylize, crop, and enhance points of interest

Analysis:

- Apply analytical method as appropriate to material and interest

This approach covers the key steps for performing data gathering for AR system evaluation. The list covers aspects of pre-session tasks, for instance, managing lighting, an important aspect when dealing with computer vision (CV) systems. A pilot study will likely uncover weaknesses regarding lighting, recording equipment and so forth. Furthermore, the pilot reveals points to consider during the session itself. Post session, the researchers should perform a debriefing of the participants and rapidly label and archive the data. The labeling and archiving of data can, in some instances, be done between sessions if the data is store digitally. This approach generates a significant amount of data, and if the session needs to be paused, many small files will be generated. In the postproduction step, the researchers first need to combine the video-streams and start identifying candidate fragments for presentation and analysis.

Communication of research

The seventh guideline for design research is named, “Communication of research.” A crucial part of any research is communicating findings and providing insightful analysis. Sequential art (i.e. comic strip style) has been used extensively to communicate findings from video-based studies in the Department of Education at the University of Gothenburg. Jonas Ivarsson (2010) has developed a range of different styles for communicating the interactions and verbalizations of students and teachers using different tools in an educational context. He uses conversation analysis on video data to uncover, describe, and analyze interesting findings.

The language of comics has been adapted to a certain degree, using wording like “panels,” “frame,” and “speech bubbles.” We were inspired by the cartoon style representations in Ivarsson’s work within architectural education (Ivarsson, 2010) and virtual game worlds (Bennerstedt & Ivarsson, 2010). Another approach is to present the actions of participants and to embed frames from the video in a transcription (see Murphy, Ivarsson, & Lymer, 2012).

In Paper 3, we argue that comic strip illustration as a tool for communicating findings is fruitful. The inspiration for this approach came from the work done in Paper 1. When presenting this paper in an academic context, video was used to illustrate the prototype and evaluation. The comic strip approach attempts to bridge the gap between live video and still images, allowing readers to get an impression, on paper, of the physical interaction that occurred.

Will Eisner (1990) does an extensive walkthrough of usable techniques for creating these effects in *Comics and Sequential Art*. A frame can be stylized to convey as much information as possible without the presence of distracting elements that confuse the viewer. Figures 25 and 26 illustrate the concept we propose. In Figure 25, a screenshot from a video sequence shows how the user interacts with the device.

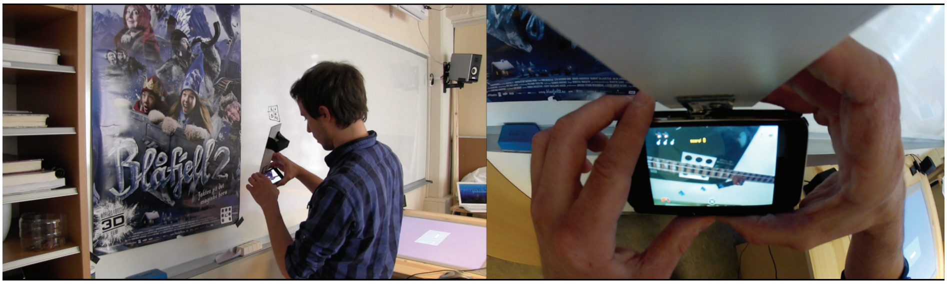

Figure 25: Raw video data, as it is often presented in papers.

Figure 26 attempts to communicate more relevant information by removing superfluous elements and highlighting the important aspects of interaction.

Figure 26: Modified video data using comic strip principles to convey the interaction.

While Figure 25 illustrates the context reasonably well, it does not give the reader information about how the user moves and interacts with the technology. Figure 26, on the other hand, uses speech bubbles to convey the speech related to the action, and uses streaks to illustrate movements and interactions.

This approach may be used to illustrate findings for mimetic games or similar handheld and static-display AR applications.

6.1.4 Design guidelines for MAR

Designing for MAR presents a significant challenge. In this section, we will provide design guidelines for different aspects of MAR applications. In Paper 4, we present design guidelines using a framework for affordances developed by Kaptelinin and Nardi (2012). Using the categories of their framework, the paper offers guidelines for learning, maintenance, aggregation, and remediation.

Learning: What steps the users go through to learn how to interact with an AR application, and how we can design to allow faster and better learning.

Maintenance: What needs to be taken care of (maintained) to allow applications to function effectively.

Aggregation: How the application is intertwined with analog and digital artifacts to achieve its purpose.

Remediation: Specifically, how we re-communicate content from existing interfaces in a new interface. In the case of MAR, the guidelines reflect the care needed to effectively use the MAR interface to bridge the gap between known and new ways of interacting with content.

By analyzing the findings in Paper 4, we present the following guidelines in Table 2:

Table 2: Set of design guidelines

1. Clearly show what objects the application is referring to in the physical world to support learning.
2. Represent the internal state insofar that users may learn how the computer vision algorithm performs satisfactorily.
3. Inform users of actions needed to adequately maintain an environment suitable for augmenting.
4. Afford the relationship the device has to the trackable (e.g a marker).
5. Allow easy transition between known interaction paradigms and the augmented interaction paradigm of remediated content
6. 3D content should convey its meaning at natural viewing angles
7. Provide the same entry point for interaction even though the content is different
8. Humanoid characters that allow interaction should responsive to the user

The guidelines we present in Table 2 are discussed below in relation learning, maintenance, aggregation and remediation.

Learning

1. Clearly show what objects the application is referring to in the physical world to support learning.

This can be achieved by using visual clues that make clear how the software understands the world. New users should be able to learn and understand how the software works with the real world and objects in the world.

2. Represent the internal state insofar that users may learn how the computer vision algorithm performs satisfactorily.

Designers should signal the inner workings of the tracking algorithm. While the technology is improving rapidly, it still relies on the user operating it optimally to achieve a satisfactory user experience. Display cues or other types of feedback to guide the user in making optimal use of the application are necessary.

Maintenance

3. Inform users of actions needed to adequately maintain an environment suitable for augmenting.

AR applications using CV are very sensitive to lighting conditions. Hence, the users need to be made aware of how to maintain suitable conditions for the AR application to operate in. This may be achieved by introducing functions to the hardware such as external sensors for light and providing environmental information to enable users to improve the performance of the application. Similarly, AR applications that use, for instance, GPS sensor data to create augmentations need to make the user aware of how to improve signal quality.

Aggregation

4. Afford the relationship the device has to the trackable.

Guidelines related to aggregation illustrate the fact that some applications must be combined with other artifacts to achieve their intended purpose. Visual clues (in the real world as well on the device) can be used to make the users direct the device toward the object of interest. Bear in mind that textual clues may not be perceived.

Remediation

5. Allow easy transition between known interaction paradigms and the augmented interaction paradigm of remediated content.

Users will try to view content in ways that they are familiar with in other media. This behavior is learned from previous interaction with media. Users may intuitively swipe or use other gestures learned from other types of media, and this should be taken into account when designing interfaces for remediated content.

6. 3D content should convey its meaning at natural viewing angles.

Although 2D content can easily be viewed from a bird's eye view, this is not the case for 3D content. 3D models are often designed for viewing from a slightly front-angled perspective. As the user will, in most cases point, the device towards the marker, the content needs to give meaning from this angle. Hence, 3D content needs to be designed explicitly to support viewing from a bird's eye view.

7. Provide the same entry point for interaction even though the content is different.

We found that our effort to align the interaction more closely with that of QR-code interaction confused the users rather than improving the user experience. Specifically, we allowed the user to pick up content from markers and view it in full screen, unaugmented and unrelated to the world. This caused confusion among the users.

8. Humanoid characters that allow interaction should be responsive to the user

Users may become uncomfortable or provoked if humanoid characters that afford interaction are unresponsive (see Wagner, Billinghurst, & Schmalstieg, 2006).

Modern systems design requires people from a wide range of disciplines to work together (Borchers, 2001). In the development of ARad, people from media, research, development, and 3D authoring have collaborated to create this artifact. According to Borchers (2001), "Concrete design guidelines can serve as vocabulary among design teams" (p. 366). They also provide support in the design process when creating content and interfaces for MAR applications.

6.2 Results

The individual results from this thesis are summarized as follows:

a) A tangible application freely available on the App Store for iOS.

b) A taxonomy for describing and framing research on handheld AR systems (conceptual knowledge).

c) A methodology for studying handheld AR systems and static AR systems using think-aloud and video recordings.

d) An approach to illustrating findings from highly visual applications based on recorded video data.

e) A paper demonstrating the application of the taxonomy and the approach to think-aloud in addition to the comic strip visualization.

f) Descriptive and prescriptive knowledge about affordances in remediating HMAR applications.

g) A design science research blueprint (this thesis) for conducting research on MAR systems.

These results are not only useful separately in the MAR field, but also as a whole. The application, taxonomy, methodology, approach to visualization, and resulting guidelines follow a natural progression, from the formulation of problem areas through design, evaluation, and analysis, with the dissemination of conceptual, prescriptive, and descriptive knowledge. As separate parts, the results may serve as tools in different contexts, as described below, or as an overall process blueprint for doing design science in MAR.

a) In accordance with design science, this thesis provides something tangible for the business and academic communities alike. The application, freely available in the App Store, provides an entry point for laymen wishing to explore HMAR. The experience of augmented reality can occasionally be difficult to explain in words, as it is a highly interactive and visual medium. Our application enables users to engage with the technology firsthand by downloading a freely available copy in the App Store.

b) The taxonomy described in Paper 2 is straightforward and can be applied in the commercial as well as academic context. The taxonomy reveals the different parts that constitute handheld mobile AR systems. Handheld MAR systems exhibit different properties than general MAR applications, and thus we found it useful to create a distinct description of the HMAR domain. For commercial actors, this description can be used to categorize and explain HMAR applications using a shared vocabulary. In academia, it can be useful to pinpoint research interests in specific areas of HMAR. We find that using the taxonomy enables discussion about the parts that make up HMAR.

c) Our methodology for recording and analyzing augmented reality and similar applications is described in detail in Paper 3. Different setups for experiencing AR are explained and illustrated. The methodology we employ to gather data from these systems is discussed, namely, the video-based data gathering techniques in combination with the think-aloud protocol. This methodology is decidedly a fitting approach to study both static display augmented reality and handheld mobile augmented reality. Paper 1 can be seen as an early exercise in using video data to describe MAR.

d) The approach to visualizing sequences of interaction presented in Paper 3 allows researchers and practitioners to provide in-depth descriptions of important interactions. Examples and rationales for use are presented to the reader. We find that this detailed, visual description of interaction allows us to communicate important aspects of interaction with AR applications, which would otherwise be difficult to convey with text or static screenshots alone.

e) Paper 4 utilizes the methodology to explore the affordances of remediating HMAR applications, providing researchers with a framework for conducting similar research. We diligently apply the approach to data gathering and visualization found in Paper 3 to explore the phenomenon of remediation in HMAR.

f) A set of affordances describing the nature of an HMAR application is provided in Paper 4. These affordances provide a starting point for explicit guidelines relevant to remediation of HMAR applications. The affordances we describe—aggregation, maintenance, learning, and remediation—provide general guidelines for HMAR applications.

g) This thesis as a whole can be used as a coherent document describing the components of a design research project for MAR.

6.3 Contribution of new knowledge to the field

Figure 27 illustrates how the published papers presented in this thesis answer the research questions and how they provide new conceptual, prescriptive, and descriptive knowledge.

Iivari (2007) goes to considerable lengths discussing design science in his essay, “A Paradigmatic Analysis of Information Systems as a Design Science.” He derives three types of knowledge discernible from design science:

1.) Conceptual knowledge

Conceptual knowledge can provide frameworks and scaffolds for discussing IT artifacts. For instance, taxonomies can provide an overview of an application domain. As such, “concepts and conceptual frameworks at this level aim at identifying essences in the research territory and their relationships” (Iivari, 2007, p. 46).

2.) Descriptive knowledge

The observed facts, on a descriptive level, are aimed at “describing, understanding and explaining how things are” (Iivari, 2007, p. 46). In this thesis, we provide descriptive knowledge in the form of observational data that is then analyzed. This data originates from several video-recorded user studies.

3.) Prescriptive knowledge

Knowledge at the “prescriptive level is interested in how things could be and how to achieve the specified ends in an effective manner” (Iivari, 2007, p. 46). We provide prescriptive advice in the form of affordances that can be formulated into design guidelines.

This three-level structure can be used to organize knowledge about design science (Figure 27).

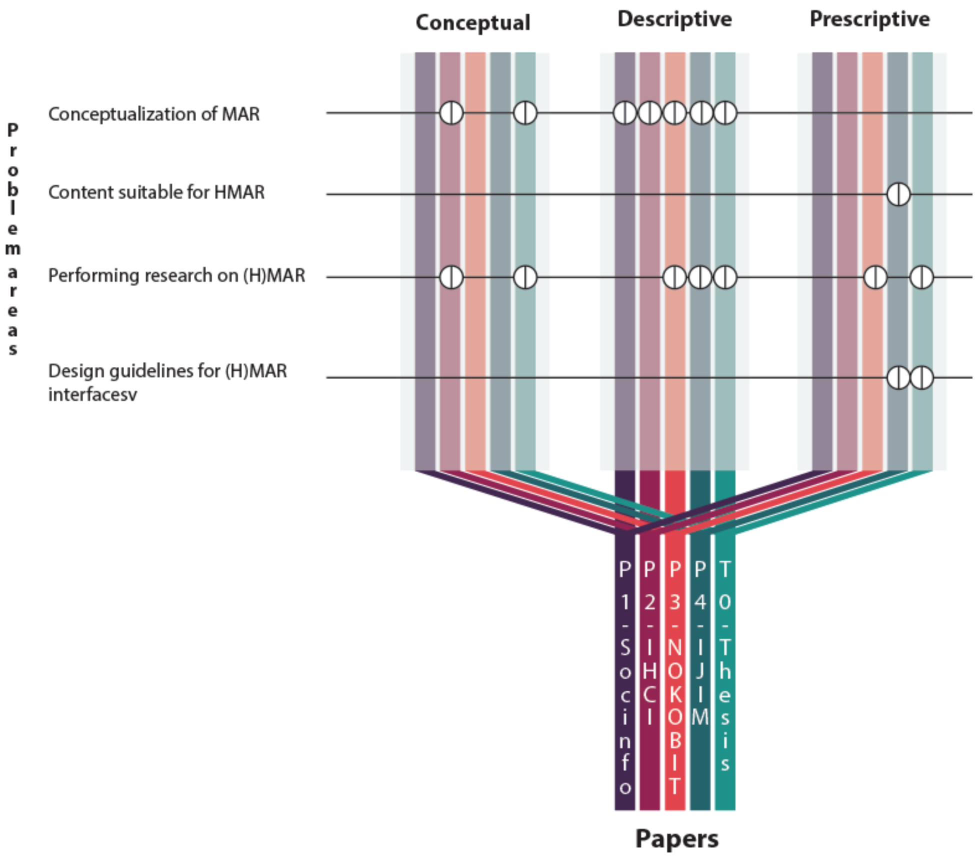

Figure 27: A graphic illustration of the topics covered by the different papers.

Paper 1 was the first article published on the topic of MAR and specifically addressed HMD-based AR. It illustrates our first thinking about using video for data gathering and subsequent analysis. The paper is very *descriptive* in its extensive detailing of the system and its use. The descriptive knowledge we present in this paper is aimed at practitioners in the fields of CSCW, AR, and furniture design who are interested in seeing a description of how AR can be used to support collaborative design activities.

Paper 2 provides *conceptual* knowledge about HMAR by presenting a taxonomy of a typical HMAR system. The intent of the taxonomy is to provide a coherent but conceptual model of what form HMAR systems take. The taxonomy can be used in a teaching context to explain the internal relationships in such systems, during or before development, to frame different technical aspects of HMAR.

Paper 3 provides *descriptive*, *conceptual*, and *prescriptive* knowledge in the form of a design tool for presenting findings and an approach to researching a specific type of system. Researchers or students interested in evaluating systems similar to the ones we describe in this paper may find the presented approach a good starting point. Similarly, those with an interest in visual communication of video-based research may find the technique describing comic-strip visualization helpful.

Paper 4 provides *descriptive* knowledge about the IT artifact and observations of its use. In addition, *prescriptive* knowledge about design in the form of affordances and guidelines is presented. The prescriptive and descriptive knowledge presented in this paper is aimed towards researchers, content creators and developers interested in design advice in this area of MAR.

This thesis bridges the gap between the different papers and presents the collection as a coherent whole. It gathers the different artifacts and sets them in perspective to each other under the DSR umbrella. The thesis allows a more detailed account of the design and evaluation; hence, researchers or designers interested in the details of DSR may find the thesis to be of interest. The development process is documented in more detail, with a description of the iterative steps showing the progression from prototype to final version.

7. Discussion

This section will discuss the results, theoretical insights and implications of this thesis, and provide a critical view of the methodology and research as well as some thoughts on the ethical implications involved.

7.1 Problem area

In this section, the different problem areas will be discussed in relation to existing literature in the field.

Conceptualization of HMAR

Milgram´s (1994) continuum is often cited when describing AR. The continuum´s strength is that it allows a broad discussion of AR, without necessarily tying it to a specific AR technology. In their discussion of AR taxonomies and definitions, Wu, Lee, Chang, and Liang (2013) come to the conclusion with regards to conceptualization of AR that "for educators and designers, defining AR in a broad sense would be more productive because such a definition suggests that AR could be created and implemented by varied technologies, such as desktop computers, handheld devices, head-mounted displays and so on" (Wu et al., 2013, p. 4). This is true; however, these varied technologies carry with them traits that should be understood prior to any design undertaking. The conceptual model presented in this thesis allows designers to understand HMAR on a narrow conceptual level modeled after its traits.

Other researchers in field of MAR have sought to categorize MAR and HMAR to describe the characteristics of the technology. This undertaking is needed, as Huang, Hui, Peulo and Chatzopoulos (2013) argue that there "are several survey papers on AR, but none is dedicated to Mobile Augmented Reality (MAR)" (p. 1). The conceptual model presented in this thesis is intended to bridge the gap between broad conceptual models (Milgram et al., 1994), survey papers (Huang et al., 2013) and sub-classification (Olsson & Salo, 2012) of MAR technology, and to provide a specific and visual model of HMAR.

Meaningful content on the MAR platform

In this thesis, it is argued that content and interface, particularly on the MAR platform, should *not* be seen separate entities serving different purposes. Rather, the content becomes the user interface. One should seek to understand the entire user experience of MAR applications when presenting content. Bolter and Grusin (2000) and Macintyre and colleagues (2001) argue that AR brings a new media experience with great potential to offer rich user experiences of content in new ways. Pavlik and Bridges (2013) discuss the emergence of MAR in journalism and how different applications support this activity. On one hand, researchers from the field of media studies argue that the their field should not responsible for providing design advice for the HCI community:

> "*We believe that media studies can be productive, although not directly by providing design rules or guidelines. Instead, the study of media history and the computer's place in that history can offer a vocabulary that helps designers reflect creatively on classes of problems and their solutions.*" (Bolter et al., 2013, p. 38).

On the other hand, they propose that "media studies can provide a vocabulary for describing interfaces and applications and a way of thinking about how we interact with them" (Bolter et al., 2013, p. 44). This concept is framed as *polyesthetics* (Engberg, 2012). It carries with it a strong focus on esthetic value: "Although there may be no such thing as a completely wrong design aesthetic, there are more or less appropriate paths for various technologies and media forms" (Bolter et al., 2013, p. 44). Hence, this thesis suggests that the HCI community should appropriate polyesthetics when considering design. On an MAR platform, the content becomes the interface. In the case of ARad, the user is interacting directly with remediated content. Designing in a vacuum with little regard or understanding of the "more or less appropriate paths" for creating coherent media experiences will impact the interface, as they are inevitably intertwined.

In this thesis, Bolter and Grusin's (2000) thoughts on remediation has been used as a foundation for discussing the perception of the content we have evaluated and analyzed. Cross-discipline research can be fruitful, and has been helpful in understanding MAR in this thesis. Similarly, cross-discipline interaction may be useful for other research areas. AR systems providing content in other contexts may benefit from a similar approach to polyesthetics, such as in the fields of history, arts, industrial design and so forth.

An approach to framing interaction as a component of meaning-making can be found in a study on art by Otitoju and Harris (2008). They invoke McLuhan (1967) in their introduction, asking the following: "As Marshall McLuhan said, 'the medium is the message.' How is a medium changed by making it interactive?" (p. 193). The authors define interaction as "a situation whereby a person, a system, a story, an exhibit, a device, an architectural piece engages us in an emotional, physical, or intellectual dialogue with itself" (p.193). The study concludes that "interaction is a component of meaning-making" (p. 201).

While evaluating ARad, the users engaged with the system and were asked during the THA sessions how they experienced the different content types. As the findings show, they found the interactive 3D content the most enjoyable. McLuhan (1967) states that "the medium is the message," meaning that form the medium takes is inserted into the message and thus affects how the message is perceived. In this regard, we must acknowledge that for meaning-making to occur in the interaction with a system, that system must be designed to utilize the best qualities of the medium.

Otitoju and Harris (2008) look at meaning-making, both individually and in pairs. They find that "discussion enriches the process of meaning-making" (p. 201). The system evaluated in this thesis is intended for single users, where the meaning-making occurs in the interaction between the content/interface and the user.

In the field of game design, Jørgensen (2013) argues that “designing game user interfaces cannot be reduced to design of icons and menus. It is also about creating a gameworld environment that supports certain gameplay activities, and for this reason it is hard to separate gameworld and interface in digital games” (p. 19). Similarly, in a MAR application, there is no clear barrier between content as we define it, and the interface.

Performing research on MAR

Formal evaluations of AR are few and far between. As Dünser, Grasset and Billinghurst (2008) note, “One reason for the lack of user evaluations in AR could be a lack of education on how to evaluate AR experiences, how to properly design experiments, choose the appropriate methods, apply empirical methods, and analyze the results” (p. 2). In their survey, Dünser et al., (2008) found that of the publications on AR and user evaluations, 75 use objective measurements, 29 use subjective measurements, 9 apply qualitative analysis, 7 use any form of usability evaluation techniques, while 41 have informal evaluations. It is worth noting that this review was done on the broad term of AR; for MAR, the selection gets narrower. In a more recent survey of MAR specifically, De Sà and Churchill (2013) conclude similarly that few evaluation methods for MAR exist.

The overarching theme of this thesis is a description of how to use DSR broadly to direct MAR research, and more specifically, how to use think-aloud in conjunction with video recording as a method to conduct research. In this thesis, we propose the approach of combining think-aloud with video recording as a viable method to investigate MAR. The step-by-step approach presented herein can serve as a detailed guide for undertaking similar studies on MAR systems.

Think-aloud has been used previously to evaluate AR and MAR systems (Dünser & Hornecker, 2007; Liarokapis, Macan, & Malone, 2009a; see Liarokapis, Macan, Malone, Rebolledo-Mendez, & de Freitas, 2009b; Morrison et al., 2009; 2011). Roto et al. (2004) did an extensive and novel study applying video recording to the gathering of data about mobile handheld device use. Video has great potential in capturing the minute interactions that take place with AR. However, using multi-camera recording in combination with think-aloud is not a “catch all” method. It has drawbacks, particularly with regard to the giant datasets often generated, and how to effectively make use of this data. Video requires a mindful approach when gathering and analyzing data (see Heath et al., 2010). This is in contrast to think-aloud, which is claimed to be a forgiving method to use (see Dix, Finlay, Abowd, and Beale, 2003).

Video recording can also have secondary uses with regards to communication of research findings, as we will discuss in the following section.

Communicating research

> *In drawing upon video, geographers have struggled with how to bring video into their texts, relying on either transcripts that lose much of the looks of the original events or sequences of frame-grabs that lose the words.* (Laurier, 2014, p. 2).

This quote from Laurier (2014) captures the problem we faced after having gathered immense amounts of video data. Communication of research findings is an integral part of DSR. Thus, this thesis presents a comic strip format for communicating findings, as useful tool for effectively conveying our results in text. Inspiration behind this format came from work of Ivarsson (2010), Bennerstedt and Ivarsson (2010) and Goodwin (2009), and their approaches to the communication of transcripts. Laurier (2014) states that “it is still early days in figuring out how to combine or translate existing transcription conventions and time-series images with or into those of the comic strip, the diagram or indeed the map” (p. 15). In this thesis, we propose our approach as a contribution to the convention of transcript presentation. Similar Lauer (2014), we find problems with preserving video recordings ‘as is’ in their graphical presentation. While the comic strip format provides visual techniques for explicating movement, for instance, it may require significant abstraction or deviation from the original frame grabs to emphasize simple events such as the movement of a finger. Furthermore, the technique is time consuming and requires familiarity with graphical tools such as Illustrator or

Photoshop, as well as knowledge of how to communicate effectively using comic strip conventions (see Eisner, 1990). Additionally, publishers themselves may be a hurdle to the use of this technique (see Heath et al., 2010, loc. 2741). Limitations on the number, size, color and resolution or images may also be barriers for effectively communicating the meaning.

Design guidelines

In a review of the literature on the topic of HCI principles for AR systems design, Dünser et al. (2007) argue that "clearly there is a need for more HCI and usability research in the field of Augmented Reality" (p. 1). They further argue that there is a fundamental difference between interacting with AR and traditional GUI-based interfaces, and this must potentially be taken into consideration when developing guidelines. In concluding their discussion on the matter, they note that it "may be difficult to develop more specific guidelines that would accommodate all AR system designers" (p. 4). General design guidelines for AR systems, covering such diverse system types as HMAR, MAR, CAR, and so forth, each having their own specific traits, potential and drawbacks, is a daunting task. In this thesis, we propose guidelines specific to HMAR. While the guidelines offered regarding learnability, aggregation, maintenance and remediation may have generalizable characteristics, they are formulated for a HMAR context and should thus be considered in that context. Further research is needed to verify the usefulness of the guidelines in different contexts.

In De Sà and Churchill's (2013) review of MAR specifically, the authors similarly find few studies providing design advice related to MAR. They note the works of Damala, Cubaud, Bationo and Houlier (2008); Morrison, Mulloni, Lemmelä and Oulasvirta (2011); Avery, Thomas and Piekarski (2008); Schinke, Henze and Boll (2010); and You et al (2008) as providing design advice through formal evaluations of MAR. While these studies do indeed provide insight into MAR and its qualities, they do not all provide concrete guidelines. A significant effort is required by the reader to distill the analysis presented in these studies into concise guidelines applicable to a design process.

However, there exist some studies that have formulated more or less concrete guidelines and have presented them to the reader. Henrysson, Billinghurst and Ollila (2005), for example, offer concrete guidelines in relation to the type of game, performance considerations, screen real-estate, and consideration of good tangible metaphors. Wetzel, McCall, Braun and Broll (2008) contribute an extensive table of design guidelines for MAR games. Their study is grounded in game theory and the design guidelines are formulated very specifically towards game experience design.

This thesis contributes a set of concrete design guidelines founded in theories from both HCI and media studies. The guidelines were formulated through an extensive evaluation of a specific HMAR application. They contribute to the arguably limited collection of existing design guidelines for HMAR.

Research question

Sharp, Rogers and Preece (2002) define interaction design as “designing interactive products to support people in their everyday and working lives” (p. 6). Furthermore, they go on to explain that a central theme of interaction design is a highly iterative, cross-discipline process, rooted to some extent in theory but which also relies heavily on good practice in the creation of usable products.

The research question put forth in this thesis is directly related to interaction design and is formulated as follows:

How can interaction with HMAR be designed to facilitate meaning-making?

To answer this research question, we have addressed four problem areas integral to the question. Firstly, we described a conceptual model of HMAR. This conceptual model is needed to understand the specifics of what constitutes an HMAR system—a clear prerequisite to considering what types of interaction are feasible to design. We propose that interaction on an HMAR system happens across a spectrum of integrated parts. What we design in the world affects how we think of registration, augmentation, and so forth. Similarly, the device capabilities impact the design of interfaces, the resolution of camera stream impacts registration, and the combined look of rendered augmentations in relation to the physical world.

Secondly, we have looked at how users interact with content on the HMAR system. This thesis proposes that content and interface cannot be looked at as separate parts. To facilitate meaning-making in these kinds of systems, the content needs to be designed with interaction in mind.

Thirdly, in order to design, a clearly defined framework is needed to guide the process. Consequently, we describe a methodology for approaching the design of HMAR systems using a combination of DSR, think-aloud combined with video recording, and comic-strip illustration to evaluate and disseminate knowledge about HMAR.

Fourth, to guide others interested in designing for HMAR, a set of guidelines outlining different aspects of HMAR interaction has been proposed. These guidelines propose specific design suggestions related to learnability, maintenance, aggregation and remediation in HMAR systems.

Thus, the final answer to the research question is as follows:

Meaningful interaction with HMAR can be designed via rigorous and iterative design process, supported by design guidelines for HMAR, where the designers are mindful of the necessary integration between content and interface.

This overall finding is directly related to Sharp et al.'s (2002) definition of interaction design. This thesis has formulated an iterative process aimed at enabling cross-discipline cooperation, rooted in theories from HCI and media studies, resulting in guidelines that support good practice when designing HMAR artifacts.

7.2 Theoretical insights and implications

The intended theoretical implications of this thesis involve three parts.

Firstly, by introducing the taxonomy, we hope to spur a discussion about what an HMAR application is. The taxonomy is grounded in the literature review as well as in our own interpretation of what HMAR entails. The taxonomy may not, however, predict the future of the field. Novel interfaces are bound to surface, as augmented reality increasingly is becoming a more widespread phenomenon. The taxonomy accounts for the present technology and leaves some room for future innovation in CV algorithms and device capabilities. However, as our insight into HMAR is limited to the current technology, it is hard to argue without speculating about the future form of HMAR. One might be inclined to think of future interfaces embedded in the hand itself or of virtual handhelds projected into reality by contact lenses. The human hand has been holding tools since the dawn of humanity, and it is hard to envision a future where humans disregard handheld tools.

A handheld form factor may not be embraced in the future, at least not in the form we propose in the taxonomy. Authorities on AR (e.g., Rosenblum) shun the handheld form factor in favor of HMD-based AR. Thus, the lifespan and impact of the taxonomy for handheld mobile AR may be short-lived indeed. Nonetheless, it will at least provide a descriptive historical document that illustrates the technology as the beginning of the 21st century.

Secondly, the intended theoretical implications for data gathering and the visualization methodologies for visual AR games are twofold. One aspect we hope to promote is the detailed methodological approach to performing research on MAR systems, which is comprehensible and should be useful to both bachelor’s and master’s students. These students could apply the methodology to existing games or applications or to those they create themselves, or. The other facet is the idea is to take considerable care when presenting findings. The comic strip approach is an improvement on the simple presentation of transcribed text because it allows a more direct view of the interaction and the visual and textual data. This is particularly important for highly visual AR software, as body movement and the world itself are the context for the findings. In addition, by using this technique, readers can get a better sense of how the system actually looked. If, for instance, the ARad application becomes obsolete at some point and is removed from the App Store, the cartoon strip visualization has preserved interaction with the technology in a historical document.

Thirdly, one of the main points of discussion in Paper 4 is the idea that content and interface in visual applications are intertwined. Regular WIMP-based user interfaces have a clear-cut distinction between content (text, images, 3D models) and UI WIMP elements, but we argue that this is not the case in AR interfaces. The content is intertwined with the interface itself to a greater degree in AR applications. Hence, we subscribe to the idea of polyesthetics, where the content itself should be considered as including user interface elements rather than solely the content.

7.3 Method

The social setting present in the EyePet session might have contributed to more careful wording by the participants. This is because some participants do not want to appear foolish or unknowing about what is taking place in the game, leading to a reluctance to express doubt. Such an attitude carries over to the retrospective debriefing session in which they reflect upon their actions and provide even more careful verbalizations of their interaction. It is worth noting that in the EyePet study, we had participants of three different nationalities and that the session was performed in English, the second language of all three.

Some of the content in both the EyePet and ARad games had a young target audience (early childhood to preadolescence). This implies that the user interface should at least be easy to use, even if the content might not be the most appealing to adult users.

The rig we created to investigate ARad impacts the ergonomics of wielding a mobile device. However, tradeoffs are needed to provide robust and compelling video data.

Seeing and recording the user interacting with the interface is key to analyzing video-based concurrent THA sessions.

7.4 Critical view of own research

The video-recorded data bears some inherent problems from the Think-Aloud method. A commonly cited problem with the THA is that a user may not be able to properly think aloud or that the context may impact their performance. In Paper 3, we conducted a study using adults to investigate the EyePet game. Preferably, we should have included children in the study. However, rules and regulations in Norway present obstacles for using children in studies that include video recordings. While the use of adults rather than children may have led to different results, the data obtained from the adults did provide interesting findings and enabled us to develop the methodology we set out to refine.

In regards to video data, it has proven quite time-consuming to analyze every interesting interaction aspect. More effort could arguably have been spent analyzing, choosing, and presenting interesting data points. However, to respect the length limitations of journals, the most illustrative examples had to be presented in a coherent manner.

Real world application

Zhou et al. (2008) called for practical and real-world use of AR. In the attempt to remediate AR for newspapers, we have indeed created something commercially viable. However, our take on the issue is not necessarily the most original. As time has passed during this research, several arguably more user-friendly and feature-rich applications have surfaced. However, we are one of the few groups in the field doing research into the usability and user experience of AR applications. As Swan notes (2005), informal usability studies, often quantitative with a small sample size, dominate the publications in AR. While papers describing novel technology are contributions to the field, they do little to further the understanding of how users interact with AR applications. While Rosenblum (2012) argues that handheld AR is inferior to its HMD siblings—which we believe, at least for the time being—handheld mobile AR is the most user-accessible technology by which to experience AR.

Acronyms

Throughout the thesis work, it has been difficult to maintain a coherent use of acronyms for AR, HAR, HMAR, and MAR technology. To appease the reviewers of the different papers, HMAR has sometimes been changed to HAR or MAR, hence contributing to the existing confusion on the use of acronyms within this field. More effort could have been spent to assure the precise use of the acronyms in the published works.

7.5 Ethical implications

Our research into AR, we would argue, has few clear and significant ethical ramifications. However, it is worth noting that the tools we are creating are attempting to change the user experience of printed media. In that sense, the technology we have created may afford some change in the way users consume and perceive content. HMAR in printed media is not yet pervasive. However, if it becomes a commonplace occurrence in printed media, it may alter the way news is consumed.

The ethics of AR in general can and should be discussed. Augmented reality technology in its present form takes the shape of metaphorical lenses that make it possible to augment the human visual system. Recent novel approaches toward interactive AR allow tactile feedback for virtual objects and an alignment of sound waves so as to appear perceptually natural. AR as a looking glass into previously unsensable space seems to be the past, present, and future area of principal focus for the research field.

The far-reaching consequences of any research field cannot be predicted at its inception. The development of transistors, CCD, and computer vision algorithms have made AR feasible in everyday life. However, we can only speculate on its possible implications down the road. But we can try to understand it as it exists today. My own research attempts to shed light on the properties of the content we choose to augment. One interesting finding observed during one of our think-aloud sessions perhaps reveals something about the purpose of AR. In that study, participants presented their view of how the technology as a medium can be useful. In *Remediation* (2000), Boulter and Grusin classify AR as a potential new medium, and we take this perspective while looking at the artifact in the study. We exposed the users to different types of AR media content during the THA, and found that the users perceived different types of content as carrying more meaning and usefulness. The types of content we were looking at could be characterized as either flat (i.e. images and video) or 3D. The flat augmented content consisted of perceptually thin overlays of images and video onto a paper marker, whereas the 3D content was something that popped up from the marker and into the world. Most of the participants stated that the idea of flat content on flat media made no sense, did not enhance the experience of AR, and was thus pointless. On the other hand, the 3D content, having more than two dimensions, carrier meaning. The participants chose to interact significantly more with this type of content and wanted to explore it further. This leads us to believe that the potential—at least in handheld augmented reality—is not in the introduction of objects that are easily obtainable in other mediums, like photographs and videos. Rather, AR's potential lies in the realm of bringing virtual 3D objects into our reality. In regard to Haraway's (1991) transition from representations to simulations, this may be partially true. With this technology, it is achievable to simulate representations of anything and naturally embed and represent simulations in our ubiquitous technology environment. We can choose to live in this environment, yet these new simulations lack the intangible element that would make them preferable to the real representation.

What we choose to put into AR will affect its initial purpose. When computers were first introduced into the mass market, their principal purpose was for use as business machines, doing important business. Computers have gradually changed; they are now ubiquitous, as cost and usability have lowered the user threshold. We believe that AR technology is not just an extension of this desktop computer paradigm, but brings something new to the table. The interaction in AR systems, which combines the real and the virtual at the same time, carries with it something new.

The mixing of the real and the virtual in this feigned new reality is the pretext for AR. It brings with it unthinkable possibilities, but it puts users in an interpretation stance at the outset. We are already skeptical of our own senses and acknowledge that we bring with us the baggage of politics and culture when engaging in scientific progression. The vision AR provides allows us to overlay reality and adjust the level of detail. What level of detail and the perspective we choose as our vantage point on the world will be determined by the optimization of the AR technology. How should we choose to optimize our perception of the world? What do we really want to see? AR allows us to remove inconvenience. Or, if we have no control over the optimization, someone else might choose to optimize our viewpoint for us.

Figure 28: A stapler is removed from the field of vision

Some novel AR research exploits the fact that logos are trackable by computer vision; in contrast to AR, the researchers call this field diminished reality (DR) (see Zokai, Esteve, Genc, & Navab, 2003). DR exploits the same progression in technology that AR does. It uses the same tracking algorithms and can be put into any AR-capable device. However, instead of introducing virtual objects, it removes real objects. Figure 28 illustrates how unwanted elements can easily be removed from your context of observation—live and in real time.

As an example, you may have a strong phobia of spiders. So on your fieldwork in the Amazon, you decide not to see them at all. This may obviously lead to the wrong conclusions; perhaps the spiders are an important factor in your fieldwork, and you have removed them from your perceived reality or have removed them entirely from your life from a young age. When this technology is intertwined to such a degree with your comfort of living, it becomes increasingly hard to perceive actual reality. Your senses are optimized to your liking; however, this optimization reduces your ability to see. This trivial example is already achievable, and it is not unthinkable that it may progress even further as the technology becomes increasingly sophisticated. Most advanced Internet users employ some sort of advertisement blocking tools in their web browsers; similarly, we can use AR and DR to remove unwanted or intrusive elements from our head-mounted and handheld augmented reality displays.

In the end, it is hard to predict the impact AR will have on greater scale. The most pertinent ethical ramifications of AR and HMAR are that the augmentation layer provides another way of altering the senses. It is in the hands of the developers to make users aware of this aspect of the technology.

8. Conclusions

MAR as a field has grown from a conceptual idea from the lab in the mid-90s to practical applications in daily use on a growing number of smartphones. The motivation for doing this research is founded in the research community's call for the study of practical applications. Research in the field is and has been largely technical nature, as the field is still young and the technical aspects of AR have not yet been solved satisfactorily. In this study, we shed light on several non-technical aspects of MAR. This thesis has put forward a detailed analysis of what MAR is on a conceptual level, how one can undertake a video study of MAR, an analysis of what content is suitable for HMAR, and design guidelines for (H)MAR systems.

On the conceptual level, a taxonomy for discussing and framing research on HMAR systems has been presented. This taxonomy allows an explanation of HMAR to laymen, professionals, and researchers. Past conceptual models of MAR have addressed categorization of mobile augmented reality on different levels, from a very broad perspective, to narrow categorizations related to functionality. The taxonomy presented in this thesis is aimed at bridging this gap, allowing a wide range of people to understand the fundamental concepts of HMAR specifically.

Furthermore, we have sought to make the reader aware of the qualities of MAR as a new medium. When the content becomes the interface, care needs to be taken when designing for such a platform. As AR's viability as a commercial technology grows further, practitioners and researchers need to familiarize themselves with the potential and drawbacks of the technology as an interface to a new medium. Meaning-making has been used to discuss how users make sense of their experiences with a remediating system, and we have concluded that the qualities of the medium must not be disregarded when designing new media experiences.

For other media, there exists a plethora of guidance with regards to how to produce content (e.g. radio), and the media themselves are well-described and understood.

An approach to evaluating, analyzing, and presenting findings with regard to MAR technology—arguably a very visual medium—has been presented in this thesis, with reproducible steps. An iterative process over several years has led to the approach presented in this paper. Researchers interested in performing video-based qualitative research on a handheld devices (even beyond MAR) can use the steps provided in this thesis. Furthermore, we present a novel approach to presenting findings using a cartoon-strip style. While perhaps a time consuming undertaking, the technique makes the findings arguably more accessible to the general reader, while still satisfying the scrutinizing eye.

Finally, several design guidelines specifically for MAR systems have been presented. Modern systems development requires a wide range of practitioners with various expertise working together. In the case of AR technology used for print media, people with backgrounds in graphic design, 3D modeling, systems development, interaction designer and so forth, are needed for collaboration. Design guidelines that enable a common vocabulary and understanding of the principles of interaction with MAR are thus helpful when designing MAR experiences.

Design science research has been used as a framework to guide this research. DSR acknowledges the contribution of the artifacts themselves. By making these artifacts available to the public, we hope to share not just illustrative descriptions but real usable systems that are available for testing and evaluation by anyone. An overview of the construction process behind the artifact ARad, which shows how it was developed and maintained throughout the thesis work, has been presented, and the application can be downloaded and tried out for free.

This research on interaction with mobile and handheld augmented reality has resulted in concrete digital artifacts as well as theoretical tools for understanding these systems. The ways in which meaningful interaction with HMAR can be designed have been demonstrated to the reader through a description of the design and evaluation of the digital artifacts. In conclusion, we argue that this is a tangible contribution to the field of AR, and MAR specifically.

8.1 Future research

We believe that the tools and methodologies described in this thesis can be used as a blueprint for developing frameworks in similar areas of AR. A taxonomy of head-worn HMD-based AR systems such as Google Glass can be beneficial to illustrate its internal and external relation to components. Furthermore, taxonomies for static display AR, similar to the EyePet and Kinect-based games, can be helpful to their respective fields.

While this study has used an app designed for remediation of different types of content as a case for determining design guidelines, performing similar evaluations of apps designed for other domains will shed light on other aspects of MAR. Of particular interest would be an evaluation of a GPS-based AR system to look at how the user interacts with the world when using such systems.

Further refinement of the video recording setup to have even less user intrusion is becoming possible as streaming of display data has larger resolution and capture equipment is getting lighter. Furthermore, quantitative and automated analysis of video data from static display-based AR looks promising. Some preliminary experiments on this type of analysis have been carried out by the author in cooperation with Aleksander Krzywinski. See the Appendix Section 9.3, “Quantitative analysis of video data.” Furthermore, by looking to the field of video abstraction, there may be potential in automating comic strip generation. This might enable the direct use of comic strips in analysis, as a new port of entry to transcribed video data.

References

Abras, C., Maloney-Krichmar, D., & Preece, J. (2004). User-centered design. *Encyclopedia of Human-Computer Interaction, 37*(4), 445–456.

Anthony, S. (2013, April 13). *US military developing multi-focus augmented reality contact lenses.* Retrieved from Extremetech.com.

Avery, B., Thomas, B. H., & Piekarski, W. (2008). User evaluation of see-through vision for mobile outdoor augmented reality. Proceedings from ISMAR 2008: *7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, 69–72. doi:10.1109/ISMAR.2008.4637327

Azuma, R. (1997). A survey of augmented reality. *Presence-Teleoperators and Virtual Environments, 6*, 355–385.

Bennerstedt, U., & Ivarsson, J. (2010). Knowing the Way. *Managing Epistemic Topologies in Virtual Game Worlds,* 201–230. doi:10.1007/s10606-010-9109-8

Bolter, J. D., & Grusin, R. (2000). *Remediation - Understanding New Media* (Paperback edidtion. pp. 1–290). The MIT Press.

Bolter, J. D., Engberg, M., & Macintyre, B. (2013, January). Media Studies, Mobile Augmented Reality, and Interaction Design. *Interactions*. doi:10.1145/2405716

Borchers, J. O. (2001). A pattern approach to interaction design. *Ai & Society, 15*(4), 359–376. doi:10.1007/BF01206115

Buchenau, M., & Suri, J. F. (2000). *Experience prototyping*. Presented at the Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques, (pp. 424–433). New York: ACM. doi:10.1145/347642.347802

Bødker, S., & Andersen, P. B. (2005). Complex mediation. *Human-Computer Interaction, 20*(4).

Caudell, T. P., & Mizell, D. W. (1992). *Augmented reality: an application of heads-up display technology to manual manufacturing processes*. Presented at the Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences, (Vol. 2, pp. 659–669).

Cheon, J. H., Choi, J. H., Hur, J. H., Jang, J., Shin, H. S., Kyeong Jeong, J., et al. (2006). Active-matrix OLED on bendable metal foil. *IEEE Transactions on Electron Devices, 53*(5), 1273–1276. doi:10.1109/TED.2006.871873

Cross, N. (2001). Designerly Ways of Knowing: Design Discipline Versus Design Science. *Design Issues, 17*(3), 49–55. doi:10.1162/074793601750357196

Damala, A., Cubaud, P., Bationo, A., & Houlier, P. (2008). *Bridging the gap between the digital and the physical.* Presented at the 3rd International Conference on Digital Interactive Media in Entertainment and Arts.

De Sà, M., & Churchill, E. F. (2013). Mobile Augmented Reality: A Design Perspective. In W. Huang, L. Alem, & M. A. Livingston, *Human Factors in Augmented Reality Environments* (pp. 139–164). New York: Springer. doi:10.1007/978-1-4614-4205-9_6

DIS, ISO. (2009). 9241-210: 2010. Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems. International Standardization Organization ISO. Switzerland.

Dix, A., Finlay, J. E., Abowd, G. D., & Beale, R. (2003). *Human-Computer Interaction* (3rd Edition) (3rd ed.). Prentice Hall.

Dubois, E., Gray, P., & Nigay, L. (2002). ASUR++: a design notation for mobile mixed systems. In *Human Computer Interaction with Mobile Devices* (Vol. 2411, pp. 22–24). Springer Berlin Heidelberg.

Duncker, K. (1945). On problem-solving. *Psychological Monographs, 58*(5), i–113. doi:10.1037/h0093599

Dünser, A., & Hornecker, E. (2007). *An Observational Study of Children Interacting with an Augmented Story Book*. Presented at the Edutainment'07 Proceedings of the 2nd international conference on Technologies for e-learning and digital entertainment, (pp. 305–315). Berlin: Springer Berlin Heidelberg. doi:10.1007/978-3-540-73011-8_31

Dünser, A., Grasset, R., & Billinghurst, M. (2008). A survey of evaluation techniques used in augmented reality studies. Human Interface Technology Laboratory New Zealand.

Dünser, A., Grasset, R., Seichter, H., & Billinghurst, M. (2007). Applying HCI principles to AR systems design.

Eisner, W. (1990). *Comics & Sequential Art* (Expanded Edition. pp. 1–166). Poorhouse Press.

Engberg, M. (2012). Writing on the world: augmented reading environments. Sprache Und Literatur, 67–78.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87*(3), 215–251.

Feiner, S. K. (2002, April). Augmented Reality: A New Way of Seeing. *Scientific American Magazine*, 1–8.

Ferdinand, P., Müller, S., & Ritschel, T. (2005). *The Eduventure-A New Approach of Digital Game Based Learning Combining Virtual and Mobile Augmented Reality Game Episodes*. Presented at the Pre-Conference Workshop “Game based Learning” of DeLFI 2005 and GMW 2005 Conference (Vol. 13). Rostock.

FitzGerald, E., Adams, A., Ferguson, R., Gaved, M., Yishay, M., & Rhodri, T. (2012). *Augmented reality and mobile learning: the state of the art*. Presented at the 11th World Conference on Mobile and Contextual Learning (mLearn 2012) (pp. 62–69). Helsinki, Finland.

Foxlin, E., Calloway, T., & Zhang, H. (2014). Improved registration for vehicular AR using auto-harmonization. 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 105–112. doi:10.1109/ISMAR.2014.6948415

Fragoso, V., Gauglitz, S., Zamora, S., Kleban, J., & Turk, M. (2011). *TranslatAR: A mobile augmented reality translator.* Presented at the WACV '11: Proceedings of the 2011 IEEE Workshop on Applications of Computer Vision (WACV, IEEE Computer Society.

Gibson, J. J. (1977). *The theory of affordances.* Hilldale.

Gjøsæter, T. (2008). Bergen Open Research Archive: The development and evaluation of CrimeSceneAR - An augmented reality application for crime scene investigation and reconstruction.

Goodwin, C. (2009). Embodied Hearers and Speakers Constructing Talk and Action in Interaction. *Cognitive Studies, 16*(1), 51–64.

Haraway, D. (1991). *Simians, Cyborgs and Women* (Routledge. pp. 1–309). New York: Routledge.

Harrison, S., Tatar, D., & Sengers, P. (2007). *The three paradigms of HCI*. Presented at the Alt. Chi. Session at the SIGCHI Conference on Human Factors in Computing Systems (pp. 1–18). San Jose, USA.

Hassenzahl, M. (2013). *User Experience and Experience Design. The Encyclopedia of Human-Computer Interaction* (2nd ed.). Aarhus, Denmark: The Interaction Design Foundation. Retrieved from https://www.interaction-design.org/encyclopedia/user_experience_and_experience_design.html

Heath, C., Hindmarsh, J., & Luff, P. (2010). Video in Qualitative Research (Introducing Qualitative Methods series). Sage Publications Ltd.

Henrysson, A., Billinghurst, M., & Ollila, M. (2005). *Face to face collaborative AR on mobile phones*. Presented at the Proceedings of the Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality (pp. 80–89). doi:10.1109/ISMAR.2005.32

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *Mis Quarterly, 28*(1), 75–105.

Hevner, A., & Chatterjee, S. (2010). *Design Research in Information Systems.* Springer Science & Business Media.

Hoonhout, J. (2008). Let the Game Tested do the Talking: Think Aloud and Interviewing to Learn about the Game Experience. In *Game Usability*. Morgan Kaufmann.

Huang, Z., Hui, P., Peylo, C., & Chatzopoulos, D. (2013, September 17). Mobile augmented reality survey: a bottom-up approach. arXiv.org.

Hürst, W., & Wezel, C. (2012). Gesture-based interaction via finger tracking for mobile augmented reality. *Multimedia Tools and Applications*, 1–26. doi:10.1007/s11042-011-0983-y

Iivari, J. (2007). A paradigmatic analysis of information systems as a design science. *Scandinavian Journal of Information Systems, 19*(2).

Ivarsson, J. (2010). Developing the construction sight: Architectural education and technological change. *Visual Communication, 9*(2), 171–191. doi:10.1177/1470357210369883

Jacob, R., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., & Zigelbaum, J. (2008). Reality-based interaction: a framework for post-WIMP interfaces. Presented at the Proceedings of the ….

Jefferson, G. (1984). Transcript notation. Structures of Social Action: Studies in Conversation Analysis, 346–369.

Jørgensen, K. (2013). Gameworld Interfaces. MIT Press.

Kang, S., Lee, J., Jang, H., Lee, H., Lee, Y., Park, S., et al. (2008). *SeeMon: scalable and energy-efficient context monitoring framework for sensor-rich mobile environments.* Presented at the MobiSys '08: Proceeding of the 6th international conference on Mobile systems, applications, and services, ACM Request Permissions. doi:10.1145/1378600.1378630

Kaptelinin, V., & Nardi, B. (2012). *Affordances in HCI: toward a mediated action perspective*. Presented at the CHI '12 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 967–976).

Klein, G., & Murray, D. (2008). *Compositing for small cameras*. Presented at the ISMAR '08: Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, IEEE Computer Society (pp. 57–60).

Klein, H. K., & Myers, M. D. (1999). A set of principles for conducting and evaluating interpretive field studies in information systems. *Mis Quarterly, 23*(1), 67–94.

Kruijff, E., & Veas, E. (2007). *Vesp'R - Transforming Handheld Augmented Reality*. Presented at the ISMAR '07: Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality (pp. 1–2). doi:10.1109/ISMAR.2007.4538865

Lamb, P. (2003, June). *ARToolKit.* Www.Hitl.Washington.Edu. Retrieved January 21, 2013, from http://www.hitl.washington.edu/artoolkit/

Laurier, E. (2014). The Graphic Transcript: poaching comic book grammar for inscribing the visual, spatial and temporal aspects of action. Geography Compass (pp. 1–18).

Liarokapis, F., Macan, L., & Malone, G. (2009a). Multimodal augmented reality tangible gaming. *Vis Comput.*

Liarokapis, F., Macan, L., Malone, G., Rebolledo-Mendez, G., & de Freitas, S. (2009b). *A Pervasive Augmented Reality Serious Game*. Presented at the Games and Virtual Worlds for Serious Applications, 2009 (pp. 148–155). VS-GAMES '09. Conference in. doi:10.1109/VS-GAMES.2009.40

Lund, A. M. (2001). *Measuring Usability with the USE Questionnaire. STC Usability SIG Newsletter. Usability and User Experience Newsletter of the STC Usability SIG.* Retrieved August 20, 2012, from http://www.stcsig.org/usability/newsletter/0110_measuring_with_use.html

Macintyre, B., Bolter, J. D., Moreno, E., & Hannigan, B. (2001). *Augmented reality as a new media experience*. Presented at the 2001 Proceedings. IEEE and ACM International Symposium on Augmented Reality (pp. 197–206). doi:10.1109/ISAR.2001.970538

Marzo, A., Bossavit, B., & Hachet, M. (2014). Evaluating controls for a point and shoot mobile game: Augmented Reality, Touch and Tilt. *2014 IEEE International Symposium on Mixed and Augmented Reality - Media, Art, Social Science, Humanities and Design (IMSAR-MASH'D)*, 59–62. doi:10.1109/ISMAR-AMH.2014.6935439

McCarthy, J., & Wright, P. (2004). *Technology As Experience* (pp. 1–225). London, England: The MIT Press.

McLuhan, M. (1967). *The medium is the message.* New York: Bantam Books.

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1994). Augmented reality: A class of displays on the reality-virtuality continuum. *Proceedings of Telemanipulator and Telepresence Technologies, 2351*(34), 282–292.

Mohring, M., Lessig, C., & Bimber, O. (2004). Video See-Through AR on Consumer Cell-Phones. Proceedings of the 3rd IEEE/ACM … (pp. 252–253). IEEE Computer Society. doi:10.1109/ISMAR.2004.63

Morrison, A., Mulloni, A., Lemmelä, S., & Oulasvirta, A. (2011). Collaborative use of mobile augmented reality with paper maps. *Computers & Graphics, 35*(4), 789–799.

Morrison, A., Oulasvirta, A., Peltonen, P., Lemmela, S., Jacucci, G., Reitmayr, G., et al. (2009). *Like bees around the hive: a comparative study of a mobile augmented reality map.* Presented at the CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM Request Permissions. doi:10.1145/1518701.1518991

Murphy, K. M., Ivarsson, J., & Lymer, G. (2012). Embodied reasoning in architectural critique. *Design Studies, 33*(6), 530–556.

Nielsen, J., Clemmensen, T., & Yssing, C. (2002). Getting access to what goes on in people's heads?: reflections on the think-aloud technique. *Proceedings of the Second Nordic Conference on Human-Computer Interaction,* 101–110.

Nordbø, T. (2011). *ARTree. Designing & evaluating handheld augmented reality art.*

Norman, D. A. (1988). *The design of everyday things.* Basic books.

Olsson, T., & Salo, M. (2012). *Narratives of satisfying and unsatisfying experiences of current mobile augmented reality applications.* Presented at the 2012 ACM annual conference (pp. 2779–2788). New York, New York, USA: ACM. doi:10.1145/2207676.2208677

Otitoju, K., & Harrison, S. (2008). *Interaction as a component of meaning-making.* Presented at the 7th ACM conference (pp. 193–202). New York, New York, USA: ACM. doi:10.1145/1394445.1394466

Pasman, W., & Woodward, C. (2003). *Implementation of an augmented reality system on a PDA*. Presented at the Proceedings of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality (pp. 276–277). doi:10.1109/ISMAR.2003.1240718

Pavlik, J. V., & Bridges, F. (2013). The Emergence of Augmented Reality (AR) as a Storytelling Medium in Journalism, 15(1), 4–59.

Piekarski, W., & Thomas, B. (2002). ARQuake: the outdoor augmented reality gaming system. *Communications of the ACM, 45*(1). doi:10.1145/502269.502291

Piumsomboon, T., Altimira, D., Kim, H., Clark, A., Lee, G., & Billinghurst, M. (2014). Grasp-Shell vs gesture-speech: A comparison of direct and indirect natural interaction techniques in augmented reality. *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 73–82. doi:10.1109/ISMAR.2014.6948411

Plowman, L., & Stephen, C. (2008). The big picture? Video and the representation of interaction. *British Educational Research Journal, 34*(4).

Preece, J., Rogers, Y., & Sharp, H. (2007). *Interaction design*. NY: John Wiley & Son.

Reeves, S., Sherwood, S., & Brown, B. (2010). *Designing for crowds*. Presented at the NordiCHI (pp. 1–10). Reykjavik, Iceland.

Rekimoto, J. (1996). *TransVision: A hand-held augmented reality system for collaborative design.* Presented at the 11th World Conference on Mobile and Contextual Learning (mLearn 2012) (pp. 85–90).

Rogers, Y., Sharp, H., & Preece, J. (2011). *Interaction Design: Beyond Human - Computer Interaction* (3rd ed.). Wiley.

Rohmer, K., Buschel, W., Dachselt, R., & Grosch, T. (2014). Interactive near-field illumination for photorealistic augmented reality on mobile devices., 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 29–38. doi:10.1109/ISMAR.2014.6948406

Rosenblum, L. J., Feiner, S. K., Julier, S. J., & Swan, J. E. (2012). The Development of Mobile Augmented Reality. In Expanding the Frontiers of Visual Analytics and Visualization (pp. 431–448). London: Springer London. doi:10.1007/978-1-4471-2804-5_24

Roto, V., Oulasvirta, A., Haikarainen, T., Kuorelahti, J., Lehmuskallio, H., & Nyyssonen, T. (2004). EXAMINING MOBILE PHONE USE IN THE WILD WITH QUASI-EXPERIMENTATION. Helsinki Institute for Information Technology (HIIT), Technical Report 1 (Vol. 1, pp. 1–19).

Schinke, T., Henze, N., & Boll, S. (2010). Visualization of off-screen objects in mobile augmented reality. … On Human Computer Interaction with Mobile ….

Schmalstieg, D., & Wagner, D. (2007). *Experiences with Handheld Augmented Reality.* Presented at the ISMAR '07: Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality. doi:10.1109/ISMAR.2007.4538819

Schmalstieg, D., Langlotz, T., & Billinghurst, M. (2011). Augmented Reality 2.0. In VIRTUAL REALITIES. Wien: Virtual Realities. doi:10.1007/978-3-211-99178-7

Schmidt, A. (2013). Context-Aware Computing: Context-Awareness, Context-Aware User Interfaces, and Implicit Interaction. *The Encyclopedia of Human-Computer Interaction.* The Interaction Design Foundation.

Sharp, H., Rogers, Y., & Preece, J. (2002). Interaction design: beyond human-computer interaction (pp. 1–551).

Shilkrot, R., Montfort, N., & Maes, P. (2014). nARratives of augmented worlds. … *Augmented Reality-Media*, 35–42. doi:10.1109/ISMAR-AMH.2014.6935436

Simon, H. A. (1996). *The Sciences of the Artificial* (3rd ed., pp. 1–241). London, England: MIT Press.

Steuer, J. (1992). Defining Virtual Reality: Dimensions Determining Telepresence. *Journal of Communication, 24*(4), 73–93.

Sutherland, I. E. (1965). *The Ultimate Display*. Presented at the Proceedings of the IFIP Congress (pp. 506–508).

Swan, J., & Gabbard, J. (2005). *Survey of user-based experimentation in augmented reality.* Presented at the Proceedings of 1st International Conference on Virtual

Reality.

Tullis, T., & Albert, W. (2008). *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics (Interactive Technologies)* (1st ed.). Morgan Kaufmann.

van Dam, A. (1997). Post-WIMP user interfaces. *Communications of the ACM, 40*(2), 63–67. doi:10.1145/253671.253708

Wagner, D., & Schmalstieg, D. (2003). *First steps towards handheld augmented reality*. Presented at the ISWC '03 Proceedings of the 7th IEEE International Symposium on Wearable Computers (pp. 127–127).

Wagner, D., Billinghurst, M., & Schmalstieg, D. (2006). *How real should virtual characters be?* Presented at the ACE '06 Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology, ACM. doi:10.1145/1178823.1178891

Wetzel, R., McCall, R., Braun, A.-K., & Broll, W. (2008). Guidelines for designing augmented reality games. the 2008 Conference (pp. 173–180). New York, New York, USA: ACM. doi:10.1145/1496984.1497013

Winograd, T., Bennet, J., De Young, L., & Hartfield, B. (1996). *Bringing Design to Software.* (T. Winograd). Addison-Wesley. Retrieved from http://hci.stanford.edu/publications/bds/

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. Clinical *Neurophysiology, 113*(6), 767–791.

Wu, H.-K., Lee, S. W.-Y., Chang, H.-Y., & Liang, J.-C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education, 62*. doi:10.1016/j.compedu.2012.10.024

You, Y., Chin, T. J., Lim, J. H., Chevallet, J.-P., Coutrix, C., & Nigay, L. (2008). *Deploying and evaluating a mixed reality mobile treasure hunt: Snap2Play.* Presented at the MobileHCI '08: Proceedings of the 10th international conference on Human computer interaction with mobile devices and services, ACM Request Permissions. doi:10.1145/1409240.1409282

Zheng, F., Schmalstieg, D., & Welch, G. (2014). Pixel-wise closed-loop registration in video-based augmented reality, *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 135–143. doi:10.1109/ISMAR.2014.6948419

Zhou, F., Duh, H. B.-L., & Billinghurst, M. (2008). *Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR*. Presented at the ISMAR '08: Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality (pp. 193–202).

Zokai, S., Esteve, J., Genc, Y., & Navab, N. (2003). *Multiview Paraperspective Projection Model for Diminished Reality*. Presented at the ISMAR '03: Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality, IEEE Computer Society (pp. 217–226). doi:10.1109/ISMAR.2003.1240705

9. Appendix

The appendix contains images and illustrations meant to show the project progression and to document it for posterity.

9.1 Documentation of development

This section of the appendix documents the development process for ARad from its early beginnings in 2010 until release.

9.1.1 Prototype version

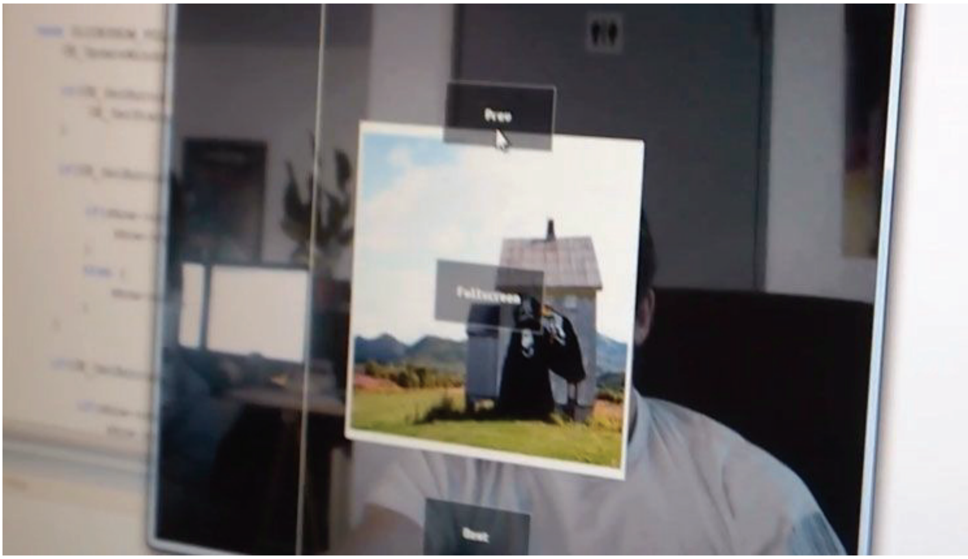

Figure 30: Feb 2010 – Development in the emulator of the ARad engine, showing the first attempt at full-screen mode for images.

9.1.2 Datatypes

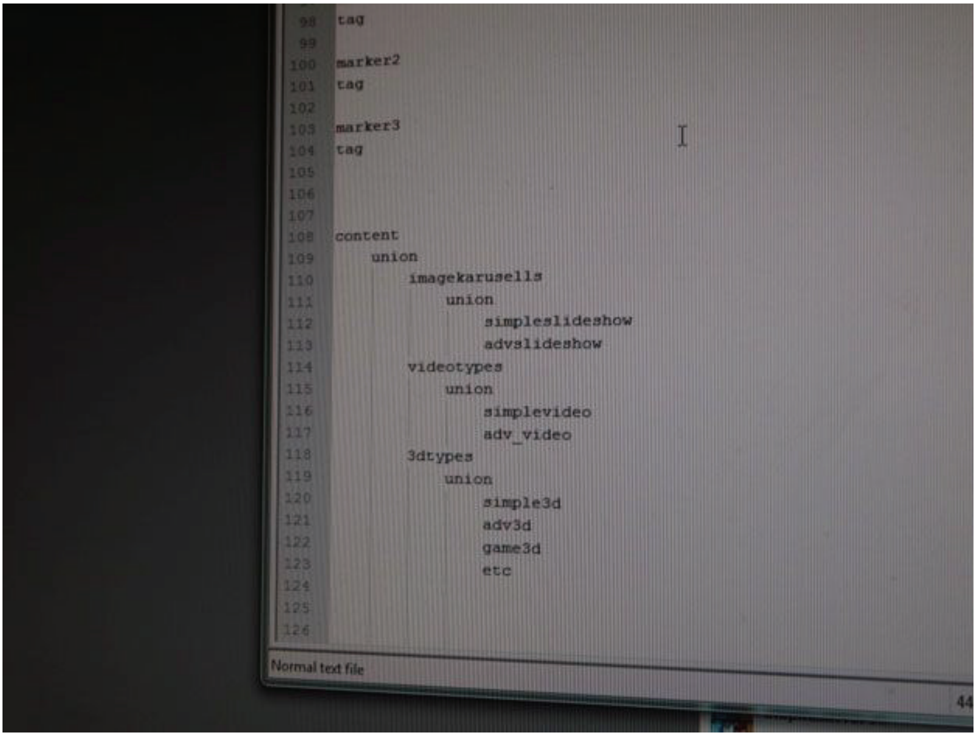

Figure 31: Feb 2010 –First reasoning about data types and whether we should use the union datatype from C to express it. We did.

9.1.3 First splashscreen

Figure 32: The splash screen of the first version of ARad released to the App Store. In this version, all content was offline.

9.1.4 Snoweffect

A great deal of the content was related to the *Trollhunter* movie and the fact that it took place in the snowy mountains. Some effort was spent to create a mixed reality snowflake effect based on the orientation of the device.

Figure 33: Snow appearing when the J marker is in view.

9.2 Posters and prints

This appendix shows the different printed materials associated with ARad.

9.2.1 Full-page advertisement

Figure 34: One of the full-page ads printed in the VG newspaper. The margin surrounding the marker dots was very small. It worked during testing, but in print, people had trouble getting good tracking.

9.3 Quantitative analysis of video data

This section shows some of the early experiments with analyzing quantitative video data gathered from a think-aloud session.

9.3.1 Visualization of video data

Figure 35: An experiment with visualization of video data.

9.3.2 Interacting with video data

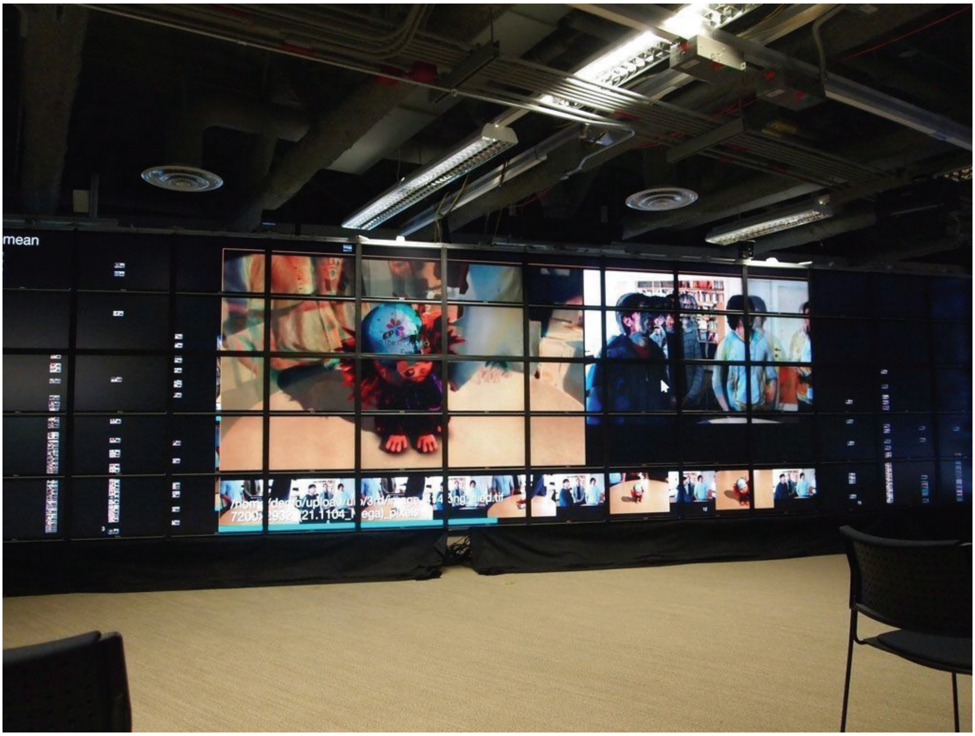

Figure 36: Interacting with categorized video data.

9.4 Evolving graphics

This section documents the evolution of illustrations and models. It is added as a historical document, similarly to Appendix Section 9.1, but also to document the process as being a search process.

9.4.1 Camera vision

Figure 37: An attempt to express the issues with camera vision. It was eventually thrown out, as it was not needed to support the argument in Paper 3.

9.4.2 Video action sketches

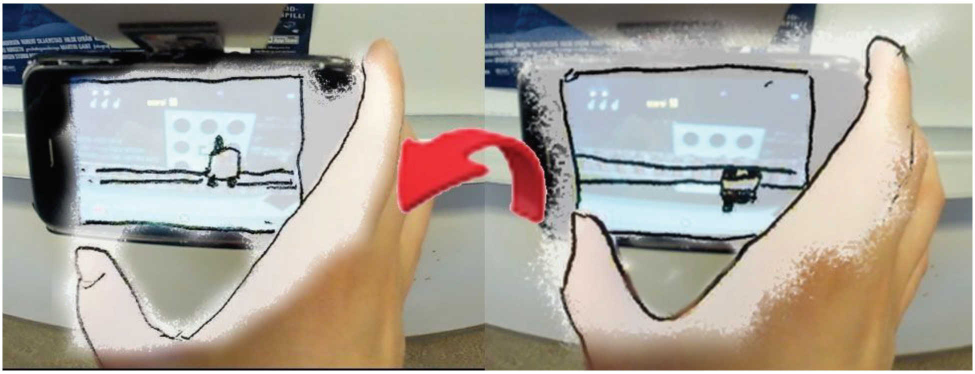

Figure 38: An early sketch of how to visualize action in video data of HMAR applications.

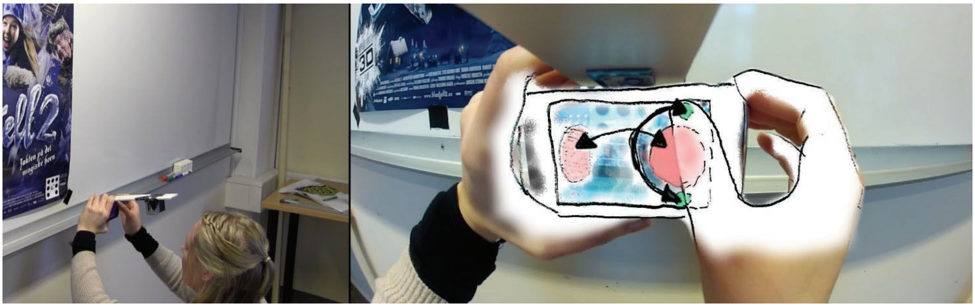

Figure 39: A sketch attempting to illustrate user understanding of a device.

9.4.3 Longer sequence

Figure 40: An attempt to illustrate in a black-and-white representation. A significant amount of information is lost.