Guidelines for User Interactions in Mobile Augmented Reality

Erik Ortman, Kenneth Swedlund
Department of Applied Physics and Electronics
Umeå University, Sweden
c05eon@cs.umu.se, c06ksd@cs.umu.se
Abstract. In the last couple of years the field of Augmented Reality has transformed from something mainly seen in academic research into several examples of big commercially successful products, and the widespread use of highly capable mobile devices has greatly helped accelerate this trend. The powerful sensors in modern handsets enables designers to bring Augmented Reality implementations to the hands of the users.

This thesis examines how Augmented Reality can be implemented on mobile platforms, mainly the iPhone 4, and surveys existing implementations and solutions for developers. It presents a number of design guidelines for user interactions in AR on mobile devices that can be used for designers as a reference when designing user-centered mobile AR applications.
Table of Contents

1 Introduction ................................................... 4
2 Background ................................................... 6
  2.1 Definition ................................................ 6
  2.2 History .................................................. 6
3 Augmented Reality on Mobile Phones ............................ 7
  3.1 Pattern or Marker based AR................................. 7
  3.2 Markerless AR ............................................ 8
  3.3 Geo-location based AR ..................................... 9
  3.4 Existing Software Development Kits for Mobile Augmented
     Reality .................................................. 10
  3.5 Methods and limitations for designing AR markers ............ 13
4 iPhone 4 Technical Performance ................................. 14
  4.1 Technical Performance .................................... 14
  4.2 Positioning System ....................................... 14
  4.3 Camera .................................................. 17
  4.4 Acceleration and Rotation Sensors ........................... 17
5 Related Work ................................................. 19
6 Methods ...................................................... 21
  6.1 Process .................................................. 21
  6.2 Resources ................................................ 21
  6.3 Generating the Guidelines and Examples ..................... 22
  6.4 Brainstorming ............................................ 22
7 Mobile Augmented Reality Design Guidelines...................... 23
  7.1 Interactions using the Device ............................... 23
  7.2 Interactions using GUI .................................... 27
  7.3 Interactions with the Augmentations ........................ 30
  7.4 Interactions using Markers .................................. 34
8 Result ........................................................ 37
  8.1 Ortman & Swedlund’s heuristics for Mobile Augmented Reality . 37
9 Discussion and Future Work ................................... 39
10 Conclusions ................................................... 40
11 Acknowledgements ............................................. 41
1 Introduction

In the last 5 years the field of mobile development has changed radically. With the introduction of the iPhone and Android handsets to the market, smartphones have become a ubiquitous part of many peoples everyday life, and today we take for granted the possibility of always staying connected wherever we go. This evolution of technology has brought with it another interesting aspect. People are carrying around powerful sensors in their pockets that can gather all kinds of data about the world around them. Through cameras, microphones, GPS systems and others, we now have mobile devices that have the capability of working in very context-aware computing environments.

Augmented Reality (AR) presents itself as a very powerful User Interface (UI) in these environments by mixing the real world with overlays of digital information so that users can get a enhanced perception of their surroundings. The field of AR has been developing quickly in the last couple of years, being widely adopted in large marketing campaigns for the automotive industry and in Hollywood. It has been identified by Gartner Inc. as one of the top 10 strategic IT technologies of our time [1], and Juniper Research forecasts a US1.5 billion revenue stream in the market by 2015.

It has however also been struggling with a large hype and a failure to deliver truly useful adaptations beyond novelty implementations that often require cumbersome markers to be printed by the users. Even AR implementations that avoid the use of markers through for example geo-location features have been struggling with the lacking fidelity of the sensors in mobile handsets, leading to inaccurate implementations that fail to impress a wide audience given the large hype surrounding the technology.

Hype aside, the revenue numbers speak for themselves. There is little doubt that, given the improvements in technology, the field of AR will continue to grow and evolve into something that will play a part in our daily interactions with technology. As our world starts becoming more and more connected, and as the consumption of information increases rapidly, it is becoming increasingly important to help society in consuming information in meaningful and simple ways, instead of actively searching and finding said information on a display. With the help of AR, information can be made available in the "real world", and give the right technology (such as glasses or lenses with built in displays) AR can provide a way of moving the eyes of society away from a square display and instead back to the real world. This way of providing information also brings with it bigger question that will have a real impact on society in the future. For example, who has the right to impose information on certain physical objects? Who has the right to impose information over a person? These are not questions we aim to answer with this thesis, but it will certainly be interesting to see how the debate will take shape in the future.

Mobile AR systems are largely dependent on the computing power and sensor fidelity of handsets to deliver truly immersive and accurate experiences, and the recent introduction of dual-core processors, gyroscopes and more accurate sensors in general hold a lot of promise for the future of mobile AR. In this study
we will explore the AR capabilities on the iPhone 4 handset, and determine what possibilities it has of delivering AR implementations that go beyond novelty and explore the possibility of deeper interaction by the user with AR systems by developing a number of guidelines for user interactions in Mobile Augmented Reality.

The thesis has been written equally by both Erik Ortman and Kenneth Swedlund, but because this thesis is part of the examination for the Master program in Interaction Technology and Design, each performance need to be presented individually. In the theoretical background, Erik Ortman has focused on the history of AR, its definition and related work. Kenneth Swedlund has focused on AR for the Mobile phone and the technical performance for the iPhone 4. For the examples we developed as well as the guidelines, Erik Ortman has focused on interaction for GUI and AR Objects while Kenneth Swedlund has focused on interaction for the device and the marker. Both have worked with all the parts, but with different focus in order to divide the work equally.
2 Background

In this section we will briefly cover the definition of Augmented Reality and talk about its early history of development.

2.1 Definition

There are two separate definitions that are used widely when defining AR. The Virtuality Continuum [2] introduced by Paul Milgram and Fumio Kishino in 1994 defines AR as one part of a Mixed Reality that encompasses both Augmented Virtuality, wherein a real object is placed into a completely virtual world, and Augmented Reality, where a real environment is augmented somehow by computer-generated images.

Another definition has been described by Ronald Azuma [3]. In order to avoid limiting AR to specific technologies, he defines AR as systems that follow three distinct characteristics. For one, the system should combine real and computer-generated images in a real environment. Secondly, it should be interactive in real time. And finally, it should be registered in 3D, aligning virtual objects with physical ones. An important aspect of this definition is that it is not limited only to Head-Mounted Displays, but is adaptable to a wide range of technologies. When discussing Mobile AR in this study, we will consider both these definitions as they both bring interesting aspects to the field.

Fig. 1. The Reality-Virtuality Continuum as defined by Milgram and Kishino

2.2 History

Although AR has in its essence long been prominent in popular culture through famous movies such as The Terminator (1984) and Robocop (1987), it was not until 1992 that AR was officially coined as a term when Tom Caudell at the Boeing aircraft manufacturer applied it to a Head-Mounted Display (HUD) that assisted in the assembly process of electrical wiring on aircraft [4]. It has been argued however [5] that the birth of AR took place back in 1968, when Ivan Sutherland and his students at Harvard University and the University of Utah
developed the first functioning HUD using half-silvered mirrors that allowed users to see both virtual images projected on the screens and objects in the room, simultaneously.

![Early pop-cultural examples of AR in Terminator](image)

**Fig. 2.** Early pop-cultural examples of AR in Terminator

As technology progressed and became more wearable over time, computing and tracking devices also became more powerful and adapt at supporting AR implementations [6]. Towards the late 1990s AR had become a distinct field of research with several conferences being held on the topic.

### 3 Augmented Reality on Mobile Phones

When discussing mobile AR in this study there are a number of different technologies that are used. Here we will introduce and describe those in greater detail, and show how they are being used in mobile phones today.

#### 3.1 Pattern or Marker based AR

When talking about mobile AR today, and AR implementations in general, the most widely adapted technology is marker tracking. This technology works by having the camera detect and track physical markers in the environment, and displaying virtual objects on the screen aligned to the markers. It has become very wide-spread due to how easy it is to calculate the angle of the camera relative to the marker and align the virtual object accordingly [7]. The virtual objects could be 3D models, information, audio or video stream, or basically anything else that can be created digitally. The early popularity of marker based AR came in large thanks to Hirokazu Kato who developed the ARToolkit [8], a now open
source framework for AR and marker tracking that has seen several hundred thousand downloads and has evolved into several spin-offs that are optimized for mobile versions.

![Image](image1.png)

**Fig. 3.** An example of marker based AR using the Qualcomm SDK

In recent years several large players have entered the field of marker based AR, releasing commercial Software Development Kits aimed at mobile platform development. Examples of these include but are not limited to the Sting AR SDK [9], which focuses on creative high-performance 3D model AR on the iPhone, and the Qualcomm AR SDK [10], which as of the writing of this paper gives developers free access to their SDK with royalty-free development and distribution.

### 3.2 Markerless AR

While marker based AR has seen widespread adaptation on mobile platforms due to the simplicity of its implementation, it still struggles with bad user experience due to the fact that users usually have to print cumbersome markers to enjoy the AR experience. As the hardware in mobile phones have improved substantially in recent years, it has opened up for a new kind of AR: Markerless AR. This technology builds on the principle of natural feature tracking, which means that the device creates a visual understanding of an environment it had no previous knowledge of. This allows for new forms of AR applications where digital images are placed and aligned straight into the real environment without the need for physical markers.

Natural feature tracking is a very complex problem that requires high computational power [11] which has limited its functionality on mobile platforms
the past. With the introduction of dual-core processors in mobile phones however, the capability of implementing natural tracking technologies has become more available, and it has also been shown that natural feature tracking can be achieved on the iPhone 3G, a device lacking the dual-core processing power of newer devices, using SLAM (Simultaneous Location And Mapping) technology [12].

![Image](image_url)

**Fig. 4.** Ball Invasion AR app by 13th Lab. Uses SLAM for natural feature tracking

SLAM has also been successfully implemented in the Apple iPad2 app Ball Invasion by developer 13th Lab. The developer has stated that it aims to release a development kit for its technology, making it the first commercial SDK for natural feature tracking on iOS [13].

### 3.3 Geo-location based AR

While both marker based and markerless AR technologies focus around the camera as the most important sensor, geo-location based AR focuses on the positioning sensors of the device such as the GPS, accelerometer, compass and gyroscope. By locating the position of the user relative to the real world, digital information can be overlaid to provide information on the users current location and surroundings. One of the biggest commercial implementations of geo-location based AR has been done by the Dutch company Layar with their Layar Reality Browser, a mobile application that has seen over 10 million installs [14]. Layar
uses the camera, GPS, compass and accelerometer to locate the device's location and displays different digital layers on top of the real world through the display of the device.

![Fig. 5. Layar Reality Browser displaying geo-based digital data](image)

While geo-location based AR is easier to implement due to not requiring the same heavy processing resources that vision based AR requires [15], it also suffers from limitations due to the fact that location sensors in today's mobile devices suffer from a lack of location information precision [16] which leads to inaccurate digital overlays.

### 3.4 Existing Software Development Kits for Mobile Augmented Reality

A few companies have created Software Development Kits (SDK) that focuses on Mobile AR. These SDKs provide both individuals and companies the possibility of developing Mobile AR solutions. Some of these kits are free to use, while others charge licensing fees for commercial use.

**Qualcomm**: Qualcomm has developed a Mobile AR SDK called Vuforia that supports iOS, Android and Unity 3D. Vuforia uses computer vision technology to
recognize 2D and 3D image targets. An image target or "Trackable" is an image that the Vuforia SDK can detect and track. These images can be in the format of JPG or PNG and is uploaded and processed in the online Target Management System, and later the system will recognize these images by comparing their natural features against a known target resource database. Unlike a traditional marker/QR-code these images don’t need special black and white regions or codes to be recognized because Vuforia uses algorithms to detect and track the features that are naturally found in the image itself. By composing several image targets in a fixed spatial relationship, Vuforia provides functionality for multi target tracking. If one image target is detected the system also knows the position and orientation to the other image targets and as long as one of these targets are visible the multi target can be tracked.

Fig. 6. The image shows a multi image target in form of a package of cereals. As long any of the sides of the box are detected a virtual bowl will be added and circulated around the package

Another type of image targets is a special kind of predefined markers called Frame Markers. Vuforia provides 512 predefined Frame Markers where each marker has a unique code of binary pattern around the border of the marker image. Decoding a Frame Marker takes relatively little processing power and this allows for all 512 frame marker to run in one application and around five of them can be detected and tracked simultaneously. [10]

Metaio: Metaio has developed a mobile SDK for AR that also supports iOS, Android and Unity 3D. Metaio’s SDK supports computer vision functionality
for recognizing markerless 2D and 3D object tracking. Their SDK also supports a robust ID marker tracking in 2D as well as QR code and barcode scanning features. To help create a natural AR experience Metaio has also developed a patented method for gravity aware AR that allows the digital augmentations to be aware of the gravitation depending on how they are placed in the physical world. [17]

Layar: Layar is one of the leading companies in geo-location based AR. They have developed a browser application on iOS called Layar Reality Browser. Using this browser, users can download or create different layers of digital content that is placed on top of the real world depending on location. This could for example display twitter updates in a certain area. Other layers lets users play games within their environment, or browse for clothes in a 360-degree virtual shop, or even view artwork that has been digitally placed into the real world like a virtual art gallery.

Layar has also developed its "Layar Vision" technology, which uses computer vision to perform image detection and tracking similar to Vuforia and Metaio. Layar wants the users to use visual triggers in their surroundings such as magazines, posters or newspapers. It has also developed the "Layar Player", a SDK for developers that wants to develop their own applications instead of just creating layers inside the Layar Reality Browser application. [14]
3.5 Methods and limitations for designing AR markers

Systems that use computer vision for recognizing image targets also need to have a system for detecting key features in the image that can later be tracked. If predefined markers are used this key feature detection has already been used and saved. However, if the system allows the developers to use their own images, these images must be analyzed in order to find these key features. All pictures can in theory be used for this kind of system, but the number of key features that are found in an image can vary greatly depending on its quality and layout and that affects how usable it is as an AR marker. [10]

![Fig. 8](image)

**Fig. 8.** In image A there are no sharp edges and no high local contrast. In Image B there is a high local contrast but the key features are too similar. Both images are examples of images that are hard or impossible to detect and track.

In an image it is important to have strong key features, such as sharp edges and high local contrast so that the camera can detect and track the image easily. Strong key features and a high number of key features makes for a good image target. It is important that the key features are as unique as possible, because if a pattern is repeating itself or is too similar in appearance it will provide poor or no detection. To get a high rate of detection, the key features should have a good feature distribution across the whole image and not leave many blank spaces.
4 iPhone 4 Technical Performance

The system for development in this study has been iOS with iPhone 4 as the platform. When developing for mobile devices the technical performance can be a limitation. This chapter will cover the hardware performance for the iPhone 4 and how sensors can be used in supporting AR implementations. This chapter is added in order to help designers understand the technical challenges of the current situation in mobile AR. There is no information in this chapter that is needed for the thesis results, but it can come in handy to help as a technical overview for designers that are interested in working with mobile AR. This chapter is also designed to help those that later wants to continue this research in understanding what has changed in the technology since this thesis was written.

4.1 Technical Performance

The technical performance of today’s handheld devices is constantly increasing, which is necessary for today’s AR technology. However, it is still not close to the processing power and memory bandwidth of a modern workstation CPU. The iPhone 4 uses an Apple A4 chip as its processor. The A4 chip is also used in the iPad where it is clocked at its rated speed of 1 GHz. Many examples of AR applications have been made with even lower performance than that. Therefore, the technical limitations of the smartphones today determines how good the quality of the AR application can be. A high processor speed is needed in order to faster process the data that the camera has to capture in an AR application. If the processor is not strong enough for the graphics that the application is utilizing, it can affect the frame rate. This can then make the graphics irregular and shaky and that will negatively effect the overall feeling of the application. To support the device in processing graphics, the iPhone 4 has a PowerVR, SGX535 GPU chip clocked at 200 MHz. PowerVR’s Series 5 SGX has pixel, vertex and geometry shaders hardware, supporting DirectX 10.1 and OpenGL ES 2.0. Other performance issues on the iPhone 4 that are relevant but not as important as the processor speed is the memory. The more memory a device has supports increased performance and multi-tasking. The memory of a iPhone 4 is 512 MB of eDRAM.

4.2 Positioning System

Reliable location information has become an important part of many applications and is a cornerstone for AR applications. Starting with the iPhone 3G, Apple has been providing three positioning technologies in their handsets: Assisted GPS (A-GPS), WIFI positioning and Cellular Networking positioning. These three positioning technologies all have different coverage, accuracy and reliability that this study will investigate to determine the effectiveness of Location Based Services (LBS) using the iPhone 4.
Global Positioning System (GPS) is the leading technology for LBS and improvements in GPS technology have resulted in a reliable and affordable technology where it provides an accurate location anywhere in the world with a good time effort. Most newer models of cell phones have GPS, but for the processing they use a server side component and is then referred to as A-GPS.

The university of Alaska Anchorage compared the GPS capabilities for the iPhone 4 and other GPS receivers [23]. As can be seen in Figure 9, their results show that the iPhone 4 has an accuracy range between 5 to 50 meters, with a positional accuracy of 10 meters up to 94.38% of the time and a precision of 5 meters 50.66% of the time.

To overcome the limitations of GPS systems, development in different indoor positioning systems has been made. These systems uses everything between cellular network signals, WiFi signals, Bluetooth, infrared, ultrasound or other radio frequencies. Many of these systems have a high accuracy in indoor, highly controlled environments and that results in these systems mostly being used in single buildings. Within this controlled environment many systems are able to continuously track devices or people, and this allows some applications to complement the GPS, resulting in an indoor-outdoor positioning system. For these implementations, WiFi and cellular positioning are today the most commonly used systems in mobile devices.

Wi-Fi Positioning Systems (WPS) takes advantage of wireless Access Points (AP) to find the current location of a device. The global amounts of AP have grown rapidly over the last years [24] and can now been found by individuals, homeowners, businesses, academic institutions, retail store and public buildings. The technology is based around measuring the intensity of the received signals and the method of identifying the device. WPS does not require a connection to an AP, it only needs to record the Unique Device ID (UDID), signal strength and the particular location of the device. The accuracy of the system depends

![Fig. 9. Confidence radii of the iPhone 4 GPS positioning](image)
on the number of positions that has been entered into the database. That means that the coverage of WPS is usually at its best in urban areas and especially indoors.

During a study of iPhone A-GPS, the university of New Mexico also investigated the other technologies for LBS on the iPhone [24]. Their results show that the performance of WiFi positioning on the iPhone at indoor locations was less than the performance of A-GPS outdoors. They could only achieve a 87.7% availability for acquiring a location in urban areas. They compared their results to earlier studies that had been conducted one year earlier, which showed a 100% availability and also claimed that WPS systems are able to achieve a horizontal accuracy of 20 to 30 m. Their newer results however, show a median error of 74 m.

Cellular Positioning System (CPS) uses the cellular towers to get an area of the network. Each area is divided into cells and each of these cells have a cellular tower associated to it. When a user connects to the network, the device is allocated to the cellular tower which shows the strongest signal. There are different techniques to get the devices location. The most basic method is to determine which cellular tower has the strongest signal, and then set the devices location to be the same as that tower. If the device is within range of multiple base stations, more complex methods can be used.

Studies show that the availability of cellular positioning on the iPhone at indoor locations is 98.5% which is much higher than the WiFi positioning for the same location, but the accuracy was much lower than the WiFi where the
median horizontal error was 599m [24]. Still, cellular positioning is of interest because it may be available when other signals are not, such as A-GPS or WiFi.

### 4.3 Camera

Many AR applications need a good camera for either using markers or to be able to use markerless tracking. The iPhone 4 uses a 5 mega-pixel camera with high definition video capabilities that records video at 1280x720 pixels with a speed of up to 30 frames per second [22]. This performance is much higher than earlier mobile phones that usually had video recording restricted to 320x240 pixels [25]. Compared to HMD, that have around 1600x1200 pixels[26], the iPhone 4 is an equal alternative for running AR applications.

One AR solution for smartphones is to run environment mapping on the fly and in the same time use the maps that are generated to track the camera’s orientation. This process requires at least 15ms per frame and that allows the smartphone applications to run at an interactive frame rate (at 30 frames per second). It is important for a visualization system to maintain an interactive frame rate. If the frame rates are too low or arrhythmic the feeling for the system is greatly reduced. [20]

<table>
<thead>
<tr>
<th>Task</th>
<th>PC</th>
<th>Phone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down sample</td>
<td>&lt;0.1ms</td>
<td>0.5ms</td>
</tr>
<tr>
<td>Track low-res</td>
<td>0.5ms</td>
<td>1.9ms</td>
</tr>
<tr>
<td>Track mid-res</td>
<td>0.2ms</td>
<td>1.8ms</td>
</tr>
<tr>
<td>Track high-res</td>
<td>1.0ms</td>
<td>0.4ms</td>
</tr>
<tr>
<td>Map color (b/w)</td>
<td>0.4ms</td>
<td>4.5ms</td>
</tr>
<tr>
<td>Overall (b/w)</td>
<td>2.2ms</td>
<td>15.2ms</td>
</tr>
</tbody>
</table>

![Fig. 11. The timing for a PC(2.5GHz quad-core notebook) and smartphone(Asus P565, 800MHz) when tracking the camera and updating the map for camera image.](image)

The table above shows the average duration for each task on both a PC and a smartphone while it is creating the map. The values depends on the number of key points it needs to track from the environment and the user’s behavior while running the application. [27].

### 4.4 Acceleration and Rotation Sensors

The iPhone 4 comes with 3 separate sensors for registering the acceleration of the device: A digital compass, an accelerometer and a digital 3-axis gyroscope. These three sensors all help in tracking the devices location and orientation in 3-dimensional space.

The digital compass of the iPhone 4 provides a measure of magnetic field strength along x, y, and z directions. It works like a normal magnetic needle
compass and captures the temporary changes in the magnetic field when then magnet is moved [28]. This also means that the accuracy of the digital compass can be affected by magnetic fields or other environmental interferences, so there may be need for the compass to be calibrated from time to time. The compass should therefore only be used as a basic navigation assistant. The accelerometer of the iPhone 4 measures the device’s proper acceleration in 3 different axis, allowing the system to recognize how the device is being moved around in space. When implemented in AR systems this technology allows for a more accurate representation of the augmented digital layer, since the device can better track where it should position the camera in the digital world based on its position in real space. The downside of the accelerometer is that it uses dead reckoning and is prone to drift over time [29], which leads to an inaccurate representation of the device location relative to the real world. This prevents the device from using the accelerometer to create AR implementations that utilize truly accurate tracking representations in six degrees of freedom. The 3-axis gyroscope in combination with the existing accelerometer gives the device what is called six-axis motion sensing (x, y, z for both the accelerometer and gyroscope), which helps the iPhone to know its rotation and orientation. The integration of the gyroscope has allowed for more accurate recognition of movement within a 3D space than before [30].
5 Related Work

While we in our research has focused on creating guidelines specifically for user interactions in Mobile Augmented Reality environments, there has also been research conducted on how to properly design the overall AR user experience, using techniques such as Embodied Interactions. They represent the types of interactions that rely on our pre-existing embodied knowledge of the physical world (naive physics, body awareness and skills, environmental awareness and skills, social awareness and skills) [31]. According to Jacobs, naive physics refers to peoples common sense knowledge about the physical world. Body awareness and skills means that people have an awareness of their own physical bodies, and possess skills for controlling and coordinating their bodies. Environment awareness and skills refers to people having a sense of their surroundings and possess skills for negotiating these environments. Finally, social awareness and skills means that people are generally aware of others in their environment and have skills for interacting with them. Users will refer to this knowledge when they come upon systems that implement similarities to the real world, and it is important to consider this when designing mobile AR systems since they often enter into the physical realm of interactions.

![Fig. 12. The four Reality Based Interaction themes according to Jacobs](Image)

Xu et al. has suggested a number of design pre-patterns that support embodied interactions in Handheld Augmented Reality (HAR) games [32]. Specifically, these pre-patterns are: device metaphor, control mapping, design for the seams, world consistency, landmarks, personal presence, living creatures, hiding or revealing information, and body constraint. These pre-patterns offer a foundation of things to consider when designing for HAR games, but their findings carry over to other AR systems as well even if they carry different affordances and constraints. In their report they cover both aspects of user interaction and also user experience. For example, one of the points they raise in their pre-patterns is World Consistency. While not directly related to interactions made by the user, it is highly relevant to the general feel and experience of the application. World consistency means that if the system consists of a hybrid world of both digital and physical content, the user will expect the digital content to adhere to the same rules and principles of the real world such as physics and gravitation. Xu
et al. states that a designer can the choose wether to adhere to these rules or to defy user expectation, but the point still has to be considered.

<table>
<thead>
<tr>
<th>Title</th>
<th>Meaning</th>
<th>Embodied Skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>Device Metaphors</td>
<td>Using metaphor to suggest available player actions</td>
<td>Body A&amp;S Naïve physics</td>
</tr>
<tr>
<td>Control Mapping</td>
<td>Intuitive mapping between physical and digital objects</td>
<td>Body A&amp;S Naïve physics</td>
</tr>
<tr>
<td>Seamful Design</td>
<td>Making sense of and integrating the technological seams through game design</td>
<td>Body A&amp;S</td>
</tr>
<tr>
<td>World Consistency</td>
<td>Whether the laws and rules in physical world hold in digital world</td>
<td>Naïve physics Environmental A&amp;S</td>
</tr>
<tr>
<td>Landmarks</td>
<td>Reinforcing the connection between digital-physical space through landmarks</td>
<td>Environmental A&amp;S</td>
</tr>
<tr>
<td>Personal Presence</td>
<td>The way that a player is represented in the game decides how much they feel like living in the digital game world</td>
<td>Environmental A&amp;S Naïve physics</td>
</tr>
<tr>
<td>Living Creatures</td>
<td>Game characters that are responsive to physical, social events that mimic behaviours of living beings</td>
<td>Social A&amp;S Body A&amp;S</td>
</tr>
<tr>
<td>Body constraints</td>
<td>Movement of one’s body position constrains another player’s action</td>
<td>Body A&amp;S Social A&amp;S</td>
</tr>
<tr>
<td>Hidden information</td>
<td>The information that can be hidden and partially revealed can foster emergent social play</td>
<td>Social A&amp;S Body A&amp;S</td>
</tr>
</tbody>
</table>

*(A&S is short for Awareness & Skills)*

Fig. 13. The nine design pre-patterns presented by Xu et. al.

They also point out a number of general do’s and dont’s based on the current state of HAR technology. These included supporting one-hand interaction, considering the natural viewing angle, making sure at least one tracking surface is always in view, not tiring players out physically and not encouraging fast actions since this can lead to loss of tracking.
6 Methods

In this section we describe the process of our work, which resources we have used, and how we have worked in the process of generating our guidelines and examples.

6.1 Process

This master thesis has been part of the examination for the Master program in Interaction Technology and Design and as a research project for The Mobile Life. It has consisted of five month full time work divided into different phases. The first month was focused on literature studies and a contemporary social and environmental analysis was also made. These studies have continued over the whole length of the project, but with less focus. After the theoretical studies were concluded, a month of trying different existing technologies was conducted wherein we familiarized ourselves in developing with different SDK’s. The third month was spent generating different AR examples and the fourth month was used to evaluate these examples and generating the guidelines. The last month was dedicated to documenting the work and writing the thesis. The time frame for the project has been long enough for a clear result, but the result could have been better with more focus on different examples and with a clear model earlier in the process. The process was not entirely clear in the beginning and the work could have been better documented, especially during the brainstorming sessions. Even to have more organized brainstorming sessions could have given a higher academic value for our project. The whole project has circled around new ways to use technology and brainstorming has been an important part of our work. To involve more people outside the process and to have a system of analyzing the outcome of our sessions could have helped in improving the overall quality of the project.

6.2 Resources

Different technical resources have been used for this thesis. We used our own iPhone 4 for developing, because we both had this device at hand and were part of the Apple developer program. Qualcomm’s AR SDK, Vuforia, has been used to create the connection between the device and the augmented world. To be able to easily place our augmented 3D content in the physical space, Unity 3D has been used because it works well with the Vuforia SDK. The Mobile Life supported us with iPhone licenses for Unity 3D. For the theoretical part we used the academic program Mendeley to easily organize all the papers from other researchers. For brainstorming and organizing our ideas we used the application MindNode Pro to create mind maps. To share documentation and files, Dropbox and Google Documents has been used. LaTeX was used to write and design the thesis.
6.3 Generating the Guidelines and Examples

The focus for our examples has been mobile AR using markers, because it is technology that can be readily used on the iPhone 4 and it is supported by the Qualcomm Vuforia SDK. The examples have been generated as an iterative process, where we in the beginning used brainstorming scenarios to find good examples and later iterated new examples as we discovered new needs. The examples are different from each other in order to better analyze a larger area in how to interact with mobile AR.

When writing an academic thesis it is always important to limit the area of research. After creating the examples, a process of what should be presented in the thesis started. Early in this stage it was decided to create a model where the different interaction methods are compared for each media that are used and for each of these analyze different ways of interaction. What these medias were and what the interaction methods were has change during the process of the project. In the end the interaction methods were set to the physical world and the virtual world. During one part of the development another "mixed world" interaction method was part of the model, but this was taken out in order to create a more consistent model. The different forms of media have been in the model for the whole process of the project, but with different names and with different responsibilities. In its simplest form the media was separated into simply on-screen and off-screen, but to add more fidelity to the model these two were separated into Device, GUI, AR Objects and Markers. The guidelines have been generated through experiences learned while running and evaluating our examples and reading about other researches experiences. Creating scenarios and brainstorming around possible problems have helped in situations when examples could not be created.

6.4 Brainstorming

During the process different methods for generating ideas have been used. Sometimes these methods have been in the form of spontaneous discussion, and sometimes they have been organized brainstorming sessions. The purpose have mainly been to generate different concepts for interaction with Mobile Augmented Reality, but also to evaluate of our ideas later in the process.

Methods that have often been used more spontaneously are "Reversed Thinking" and "Challenger". Reversed Thinking is about thinking what everyone will typically do in a certain situation and then think about what will happen if they do the exact opposite. In the Challenger method the assumptions in the situation are listed down and then each assumption is challenged. During the brainstorming sessions methods like "Gap Filling", "Brain Writing" and "Meditation" were used. Gap Filling is based on finding a current location and the end goal. Then you try to find what the gap is between these two points and what is needed to fill it. In the Brain Writing method the ideas are written down individually on a paper and after a certain time the paper changes owner, to give new input to the idea. Finally, in the Meditation method a main problem
is stated and then everyone thinks only about how to solve the problem. During the process a solution is written down immediately when it comes to mind.

For the evaluation process a SWOT analysis has been used to find out what kind of Strengths, Weaknesses, Opportunities and Threats the current solution had. During the whole process a mind mapping method was used in documenting and categorizing our work. This has also been used in organizing the whole project and to solve small problems. Both mind mapping program such as MindNode Pro and more traditional methods using pen and paper or whiteboard were used. [33]

7 Mobile Augmented Reality Design Guidelines

In this section we present the different areas wherein we have set our guidelines and show how different examples on how interactions take place in each media. For each section we define its meaning in the context of this thesis, present a number of examples on how the interactions take place, and finally present our guidelines.

The categorizations were listed into four different kind of media. Device, GUI, AR Object and Markers. These are the four different kind of media in which the user can interact with Mobile Augmented Reality. For each of these formats the user gets different perception of how to interact. This largely depends on how the user relates to the media as a digital or physical experience. Therefore each of the formats have been divided into a digital and physical subcategory. A model was created as a 4x2 matrix, wherein each type of media was compared to its digital and physical means of interaction. The model is presented in section 8, Result.

7.1 Interactions using the Device

Definition: The device is defined as the object through which the augmented world is perceived, it could be either a HMD or a mobile device. The interaction methods for the device is the relation between the physical device itself and its space. When interacting between the device and the marker or the augmented object a relation in space is created. Different experiences can be created depending on the distance and angle given between the device and the object. Interactions using the device is to encourage the user to discover the augmented space with their device and to design the virtual space for guiding the user into the right actions with the device.

Physical Interactions: The way the user interacts in the physical space with the device depends on what intention the user has. Physical interaction is how the user can feel they have control over its surroundings in the physical or virtual space. It is about designing a relation between the device and the augmented world and to make the user to feel they are part of it. It is also about helping the user to understand how to use the device through good design.
Examples: In one of our prototypes, we created a AR helicopter which could change scale through pressing a button on the screen and dragging the device closer or further away from the virtual helicopter. This functionality gives the user the option to observe a more detailed object when needed, and also to observe the helicopters surrounding environment. In the same example different text layers was shown depending where in the 3D space the device existed at the moment. In the mobile AR game “Bug Juice” [32], the user use the device as a metaphor for a magnifying glass to burn ants if the device is being hold at the right distance. A more famous example of using devices as metaphors is the Nintendo Wii[34], where the user can use the device as everything between a sports racket , steering wheel or as a mighty sword. The game “Nerdherder” [32], uses the position and orientation of the device itself in order to let the user transform their bodily awareness into action in the physical space. The nerds in the game get scared of the device and run away when the player physically moves the device closer to them.

Fig. 14. Metaphors are used in the game Bug Juice for creating a intuitive interaction.

Guidelines: When designing for interactions with device through physical media, it is important to design for intuitive user control. A common solution in creating intuitive interactions is to use metaphors. Metaphors help the user understand the digital worlds actions by relating it to the actions of the physical world, and this has been used during many years in making digital design more understandable. In the game Bug Juice they use the metaphor of a magnifying glass to give the user a fast intuitive understanding in how to interact with the game. When using metaphors, the designers must be aware that they don’t always transfer existing knowledge perfectly to the new medium. If the metaphor is dead or not strong enough it can create interaction problems. Evaluate the design regularly with user studies to find a good intuitive design, especially for AR applications when then abstraction level can be very high.
Mapping the users physical actions into the digital AR world can have an effect on the total experience. In our prototype with the helicopter, the user has the control to change the scale of the helicopter by moving the device closer or further away. This creates an interaction that makes the user feel in control over the world through the device. The game Nerdherder uses this theory for enhance the feeling of personal presence, and by doing that, improving the game experience. These physical actions create stronger bonds between the physical world and the digital world, which can help the users adapt to the augmented experience.

**Digital Interactions:** The way the virtual space guides the user in the virtual and physical space depends on the design. Digital interaction is how the user is encouraged by feedback from the augmented world in how to interact in the augmented space. It involves setting up limitations or guidance to help the user focus on what the design is intended to do, or finding out where to go next. It also involved giving the user clues on how to navigate in the space and showing the user when it reaches its limitations. It is an important part of the design as the AR environment often is very limited in its space. AR using markers is limited to always having the markers in focus in order to know the position of the device in the space and markerless AR is limited to the predefined designed space. Geo-location based AR doesn’t suffer from the same limitations in space, but when the user is given unlimited freedom, guidance is important so that content can be provided.

**Examples:** In the mobile AR game Joe Warpin [32], the user assumes the role of a sniper in a helicopter. The snipers reticule is rendered on the screen so that the user can aim and shoot. In the game, the users starts to take damage when they reach the limitation of the tracking boundaries. Resulting in an annoying noise and a constantly blinking screen, this causes the player to instinctively move to a new position and restore the tracking. This helps the user avoid going too far away from the marker and always keeping at least one complete marker in the camera view. The mobile AR game "Buffer Busters" [35], lets the user hunt monsters in the AR world. Buffer Busters uses geo-location based AR in placing out monsters around the world. To help the user find these monsters, Buffer Busters displays a radar that indicates where these monsters exist. Another example is Wikitude GPS [36], which is a GPS map app that displays arrows and lines in the AR world to indicate the way a user should take.

**Guidelines:** AR has a lot challenges in the technical limitations that need to be considered during the design. When designing for interactions with the device through digital media, the problem is mainly tracking and registration. The problem for geo-location based AR is accuracy and constant connectivity. These limitations are hard to design around, but they need to considered. Analyzing scenarios when and where the the application will be used can give the designer an idea of possible problems, such as weak signals because of indoor activity.
When considering the user experience, the user can lose focus or get lost in the augmented space if an unlimited freedom is provided. Therefore a design that helps the user in finding what they are looking for or that limits the users space should be provided. In the example Buffer Busters, the user is given a clue where the monsters are located because of the radar. That helps the user choose which is the next monster to find. In the example of Wikitude GPS, the user can choose to get their path painted out into the augmented world. This gives the user a clear way to always follow. When using markers or markerless AR the limitation is to stay within the tracking boundaries.

To have a clear indication of when the application is tracking or when it have lost the tracking is an important part of the design. It is important for the user to know if everything working as intended. When we lose the tracking in our prototypes we removed the augmented object and replaced it with a label telling the user that they have lost the tracking. The design should also encourage the user to stay within tracking boundaries. In the example Joe Warpin the relation to the marker is a part of the game design. When the user is too far away from the marker, they will be punished for that behavior.
7.2 Interactions using GUI

**Definition:** The GUI (Graphical User Interface) is defined as the part of the interface that is made up of traditional interface elements such as buttons, windows, icons and other image based elements that the user can interact with. When talking about Mobile Augmented Reality these elements can be placed either on the screen of the device, or placed as augmentations in the physical space. When placed in the physical space these elements can either be placed as a GUI on top of a physical object, or as GUI on another virtual object that has been placed in the world through augmentation.

**Physical Interactions:** Physical interactions with GUI elements in Augmented Reality systems means that the user is performing the interactions in the physical world rather than through the display of the device. This could be in the shape of touching real physical GUI elements or printed graphical GUI elements, and having the system translate those touches into actions. The GUI elements might also exist purely in digital form on the screen, but the user interacts with them through the physical space, such as tapping a specific point of an objects that has a virtual button displayed on the screen at that point.

This can be implemented in an AR system by tracking for hand gestures or interactions in specific areas of the screen. In the Qualcomm Augmented Reality SDK [10] used to develop the prototype implementations in this report, this has been solved using "Virtual Buttons", which allows the developer to assign specific rectangular areas of the augmentation that acts as interaction hotspots that when touched or occluded in the camera view trigger events in the application.

**Examples:** In our prototype implementations we designed a solution with a business card that was augmented with different GUI elements to enhance the information available when looking at the card. We designed two solutions, one with GUI elements printed onto the card itself in which we augmented the printed GUI with additional digital content, and another with purely digital GUI.
In the version with both physical and digital elements we found through our tests that the user has an easier time to figure out where to perform different interactions on the card, since the GUI elements are clearly visible even without looking through the device. However, we also found that users would sometimes fail to properly point the iPhone towards the card because they were too focused on the printed GUI elements which lead to a cognitive conflict making some users believe that the card itself would change based on their interactions. This could sometimes lead to loss of tracking making the whole interaction have no effect on the system. Another issue exists in that when digital content is displayed on the screen over the GUI elements, and the user interacts with them, the digital content will occlude the users hand, which leads to bad usability. Solutions to the occlusion problem is discussed in the next chapter about interactions with the Augmentations. Furthermore, the digital content, especially when it also takes
the shape of GUI elements, has a very strong affordance. This sometimes lead to the user preferring to perform the interactions on the screen rather than on the physical surface.

**Guidelines:** When designing physical interactions with GUI elements in AR systems it is important to consider the affordance factor of those elements that are real and exist in the physical world compared to digital elements that are being used as augmentations. When working with solutions such as the Virtual Button technology it is also important to consider the placement of interactive elements on the physical surface. Because these button areas trigger interactions by detecting occlusions, they can be triggered by mistake if they are placed in such a fashion that a user might cover one area to reach another.

**Digital Interactions:** We define the digital GUI interactions in mobile AR as the type of interactions with GUI elements that the user perform through the display. This could take the shape of interacting with traditional GUI elements such as any other mobile application. It could however also be in the shape of interacting with GUI elements that has been placed as augmentations in the physical world, but interacting with them through the screen rather than through touching them physically as in the previous section.

**Examples:** Returning to the example used in the previous section with the business card, we also implemented a solution wherein the user interacted with the GUI elements that were augmented onto the card by tapping on the screen. As was mentioned in the previous section this kind of interactions was heavily favored by users, some even tried to perform this kind of interaction even if it was not implemented in the system. This kind of interaction has the advantage that the user can perform all the needed interactions with just one hand, holding the device and tapping on the screen at the same time, while the other kind of implementation forced the user to employ both hands in order to interact with the business card. This leads to greater usability and comfort for the user.

**Guidelines:** Designing regular GUI elements in an AR system follows the same basic principles as designing and regular GUI on a mobile device. One of the key points to consider is screen real estate. When designing any mobile application there is always the issue of having a limited amount of screen available to display content to a user. In Augmented Reality applications, you often want to give as much as possible of the screen area to the camera window, so that the user has as big a viewpoint as possible to the real world. If possible, placing GUI elements on the screen can be completely avoided in favor of placing these elements as augmentations in the world instead, and allowing the user to interact as usual through interactions with the screen. When implementing solutions like these it is also important to consider the same guidelines that we present in the next section for digital interactions with the augmentations.
7.3 Interactions with the Augmentations

Definition: The augmentations are defined as the digital content that is being placed into the physical world, either through projections or by displaying them on a screen which is the case on the iPhone. This does not include regular digital content that is being displayed on the screen such as traditional GUI, but only such content that exists in the physical space.

Physical Interactions: When interacting with augmentations created through AR systems, it might seem strange to consider a physical means of interaction with a purely digital content that you can not touch. However, technology such as the Microsoft Kinect platform makes it possible to detect and track a users movement in front of a camera and connecting these movements as interactions with digital content that can be augmented on top of the physical world. While this technology is still not available on mobile platforms, advances are being made to make it available and simple implementations resembling this technology can be created using the "Virtual Buttons" available in the Qualcomm AR SDK [10] however this is limited to certain areas of the screen as has been mentioned earlier. A problem when integrating real and virtual objects is that the depth information is generally not available, which can lead to incorrect occlusions [37]. That means if a real object is placed in front of the virtual augmented object, the virtual object will still appear as being placed closest to the viewport. Because this prevents the user from telling different depths, usability suffers. To solve this problem Yokoya [38] has proposed that depth information can be acquired by using stereo cameras. Another, commercially ready solution is the Kinect platform which also enables processing of depth information in a compact hardware.

Examples: Included in the Qualcomm SDK is an example wherein the user can place virtual domino bricks by dragging a finger on the display, and then topple
the bricks by moving a hand in front of the camera and flick the bricks as if they were physically placed on the table. While this example has the possibility of creating believable physical interactions, it suffers from the same technical limitations in the Virtual Button technology that has been previously discussed. If the user moves a hand over areas not indicated as Virtual Button areas, the domino bricks will not be toppled.

**Fig. 20.** The domino example included in the Qualcomm AR SDK

*Guidelines:* When designing for physical interactions with an object that does not physically exist in the real space, consider solutions wherein the user does not need to hold, grab or otherwise perform interactions that require a lot of the hand to interact with the object. This helps avoiding the occlusion problems. Interactions such as poking, pushing or otherwise interacting in ways where the user’s hand only interacts with the edge of the virtual object also helps avoid this problem. Another solution is to restrict the user’s interaction with the virtual object through using physical objects with markers that the AR system can detect and track, as proposed by Katoa [37]. Because the form of those physical tools could be known to the system, virtual representations can be overlaid and as such the occlusion problem is avoided. This does have a negative affect on usability as that the user has to have this tool available while using the system, and when using a handheld system it means that both hands will be occupied. But it does however also help avoid the negative effect of incorrect occlusion.

Katoa also suggests that since virtual objects do not occur naturally in the real world, the user will not find it unnatural if virtual objects have a certain transparent property. This way natural objects will not be completely covered by the virtual ones even if they lie on top of one another, which leads to less visual discrepancy for the user. However, in some cases the purpose is to augment the environment with virtual objects that resemble real ones as much as possible,
so the purpose of the augmentation needs to be considered before implementing this type of transparent solutions.

**Digital Interactions:** Digital interactions with the virtual augmentations are defined as the kind of interactions where the user is interacting directly with the objects through a screen, without using GUI elements such as buttons. This includes touch gestures such as tapping, swiping, pinching and others. It also includes interactions where the user makes use of virtual tools to interact with the objects. The main differences from the physical interactions are that the user has a more detached interaction with the object in that the interactions take place on a screen while the objects exist virtually in space, but also that the user can perform more precision interactions without having to watch his or her hand through a screen and determine where it is in relation to the object, which can create disconnected user experiences.

*Examples:* The mobile AR game "Inch High Stunt Guy" [39] allows the player to drag digital objects around on the playing field to create courses and help solve different challenges. By allowing the player to edit and move several digital objects simply by using the screen, the game has the benefit of being playable using only a single marker, instead of using multiple markers to represent the different objects.

In our prototype implementations we also designed a small example where we connect a digital object to the bounds of the screen, in our case a spoon, and then use this as a tool for interacting with the augmentations in the world. This works similarly to the examples mentioned in the previous section where a physical tool is being used, only this time the tool is completely virtual. In our tests this worked well when precision control was necessary, however it led to problems in that this type of implementation might require the user to get very close to the marker, making the system lose track of any trackable objects.
Guidelines: When designing for interactions with the augmentations through digital means, it is important to consider the fact that the augmentations are often anchored in physical space, meaning that while the user is interacting with them on the screen, they are also moving around as the device is being moved. This could lead to breaking interactions if the user is for example grabbing a certain object, then moving the device in such a manner that the object is no longer placed under the users fingers. When designing such implementations the designer can therefore consider solutions where the user can "grab" the digital objects by touching them on the screen, removing their world-anchor during the duration of the touch. Designers can also consider using quick taps as interactions and avoid long pressing or dragging movements of the augmented objects, thus avoiding the problem of having the objects moving too much under the fingers of the user.
7.4 Interactions using Markers

**Definition:** The marker is defined as the object reference point between the augmented space and the physical space. The marker can be a predefined physical object that is detected by the device when running the application. It can also be key points in the physical space that are generated by the application while running. The interaction methods for the marker is the relation between the user and the marker or the device and the marker. It also includes how the user perceives and interact with the marker and how the application attempts to keep the user inside the markers limitation.

**Physical Interactions:** Physical interaction is how the user interacts with a marker as a physical object and how the user perceives the marker in the augmented world. A marker can be a steady reference point in the world which is not meant to move, or used as a tool to interact with. The marker doesn’t always have to be in the center of the space, and sometimes it might become occluded in the virtual space. Multiple markers can be used to enable interaction over larger areas and multiple makers can also be used in creating physical objects that can inspire new kinds of interaction.

*Examples:* The marker is used as a reference point for the real world and the augmented world. Because the marker is a very strong symbol, it could both enhance and reduce the experience depending how it is used. A common problem is that the augmented world is too abstract for the user and that the design of both the marker and the AR objects feels too artificial. This is especially true when using QR-tags as markers. When using multiple markers as reference points, larger worlds or more complex objects can be created. In Qualcomm’s Multimarker example [10], a cereal box is created with markers covering each side. The user can then move around the box and turn it without losing the tracking. With this solution the user can follow augmented objects that moves behind markers and observe it in more angles. In this example a plate of cereals is moving around the box. A marker can also be used as an interaction object. In the game AR marble game, [40] the user interacts with the marker as a classical marble labyrinth game. The marble balls are augmented objects and have properties such as gravitation, so that the user needs to balance the balls on the marker to guide them through the labyrinth. Multiple markers can also be used as interaction objects. In the AR game PuppyPlus [32], the user has one central marker that represent the world and several other smaller markers that acts like tools or objects in the world.

*Guidelines:* When using a marker as a tool of interaction like in the AR marble game, it is important to remember that the user can accidentally cover the marker and therefore lose the tracking. Designing a marker with areas dedicated for holding and interacting can be a solution that helps in not misleading the user. While designing for Mobile AR, the designer must keep in mind that one hand is often occupied with holding the device, and that also creates a limited
Fig. 24. The marker is used as an object for interaction in the AR marble game

relation in space between the marker, user and the device in being able to keep a line of sight through the screen. In the game PuppyPlus the user interacts with multiple markers, and each of them have different functionality. In this case it is important to design each marker in unique ways, both physically and digitally, in order to avoid misleading the user. When interacting with a single marker, the user could accidentally cover the maker, ether with their own body or with other markers. Design the physical markers bigger than the are used for recognition by the system, so that interaction can be performed without loosing the tracking.

Its not always the purpose that the user should interact with the marker. The marker can also be used as a clear reference point between the augmented world and the real world. The user knows where to focus the device and can get a feeling of the limitations of the augmented space. In some designs these limitations are not desirable, and the marker can be used more discreet in both the physical and digital world. Pixel Placement [41] a guide for better realism in the AR experience gives advice on how to use the marker on new ways. If the marker should not be in the center of the experience it can be set aside and instead project the augmented object onto other surfaces. But the benefits to have the augmented world centered around the marker is to help the user in staying centered to it as well. Another solution could be to cover the marker with a layer in the augmented world. The marker could for example transform into a landscape or become a virtual hole in a surface. This can create a stronger feeling of realism in the augmented world.

When multiple marker are used as a reference point, it allows the user to have greater freedom in the augmented world. Ether many markers can be placed in a room, that could allow the user to move around and still be able to continue the tracking. Or it could be multiple markers placed around a physical object to create a experience where the augmented object can become occluded by the physical object or vice versa. If markers are placed around a room, the room becomes a very controlled environment. It is important that the device can always track all the markers and more markers than what is needed should be use for the sake of redundancy. If the markers are place onto one or several objects, like in the Qualcomms cereal example, all the sides and angles of the
Fig. 25. The marker does not always have to be in the center of the design

object that are covered should be trackable to the user, so they don’t lose tracking while the device moving around the object.

**Digital Interactions:** A digital marker consists of key points generated when the application has analyzed its surroundings with natural feature tracking. A digital marker can be generated before using the AR application and it can also be updated on the run. What makes it different from the physical marker is that it doesn’t need to be an object that creates the augmented area. It is instead the physical space itself that creates it. The user will not find any relation to the marker, because, the marker is often not a physical object. This can create cognitive problems in understanding the boundaries of the augmented world.

**Examples:** In the AR game Ball Invasion [13], the user shoots balls from the device onto the AR objects generated in the augmented space. Before the game starts an analyzing phase is run in order to read the natural features of the surrounding area. After the process is completed the user can move around in the physical world and shoot at the targets from different angles. The connection is established as long as a certain number of key points can be tracked. How many key points is needed and for how large area they cover depends on how many unique natural key features that could be generated during the analyzing process.

**Guidelines:** Physical markers have a strong affordance of interaction that connects the physical world and the digital world between the device and the marker. If the marker is removed, this natural connection is also removed and the user doesn’t know the limitations of the augmented world.

Showing the user the key points that has been generated as a digital marker can help the user understand the limitation of the environment. Indicating when the number of key points is too low also creates a feeling of the worlds limitations. These kind of events encourages the user to stay within the limited space.

By creating AR objects in the middle of the space and rarely have objects leaving the center will effect the user to do the same. In markerless AR, the
analyzing process is an important part of the user experience. The user experience can be improved with good game design and storytelling. These guidelines mostly cover today's technical limitations, because this area is both new and technically demanding, even compared to other AR solutions. In a possible future these processes will work faster and be more effective, and a need of hiding it or considering it in the design process will become less necessary.

8 Result

As a result of our research we have generated a number of design guidelines for interactions in Mobile AR applications through analysis of existing applications and systems, discussing and analyzing the different solutions and their relevance to user interaction, and developing a number of small prototype implementations that explored different types of user interaction elements. We extracted different key interaction features that we found created the best user experience. The result was categorized into four different methods of interaction, namely Device, GUI, AR Objects and Markers. This provided us with a clear distinction between the different kinds of media the user is interacting with. We also separated the interactions into two different dimensions; the Physical and Digital.

These categorizations were compiled into the model presented below that shows how each category is separated and what kind of interactions it contains. The specific guidelines are those presented in section 7, Mobile Augmented Reality Design Guidelines.

<table>
<thead>
<tr>
<th>Physical</th>
<th>Digital</th>
</tr>
</thead>
<tbody>
<tr>
<td>Device</td>
<td>Interaction through design elements that encourage the user to explore and understand the digital and physical space</td>
</tr>
<tr>
<td>GUI</td>
<td>Performing the GUI interactions in the physical world rather than through the display of the device</td>
</tr>
<tr>
<td>AR Object</td>
<td>Interactions with virtual augmentations through actions in the physical world, such as touch the objects</td>
</tr>
<tr>
<td>Markers</td>
<td>Relation between user and physical marker and the users perception of the marker and its impact on the augmented experience</td>
</tr>
</tbody>
</table>

Fig. 26. Mobile Augmented Reality Design Guideline Model

8.1 Ortman & Swedlund's heuristics for Mobile Augmented Reality

While analyzing the guidelines a number of key features appeared. These are 5 general principles for designing interactions in Mobile Augmented Reality. We
call them "heuristics" because they are more in the nature of rules of thumb than specific usability guidelines. These heuristics are specific for MAR and other guidelines for usability should also be used for good design.

**Physical interaction** : Use metaphors to create intuitive interactions and map physical actions into the digital world to create a feeling of control.

**Control the level of reality** : Control the level of reality by considering the level of transparency, details and shadows for the virtual objects, depending on the level of realism that aims to be achieved. Also indicate for the user when they are connected to the augmented world or not.

**Embodied Interaction** : Consider the placement of interactive areas or volumes so they don’t disturb other interactions or functions. When performing physical interactions with non-physical objects, consider what physical action the user need in order to avoid the occlusion problem. Keep in mind that one hand is often occupied while using mobile devices.

**Navigation** : Help the user to navigate in the physical and virtual world. Keep markers and objects centered in design to not mislead the user and don’t cause the user to lose connection to the world. Design limitations in the interactions while moving around, because it is hard to simultaneously predict the relation of the user, the device and the virtual objects.

**Design Interaction** : Design markers and objects so that it is easy to understand their function and how they should be used. Keep a clear and easy GUI and use affordance for intuitive interaction.
9 Discussion and Future Work

The guidelines presented here are for Mobile Augmented Reality and therefore just a limited part of the AR technology field has been covered. Other parts of AR technologies have other technical limitations and possibilities, affordances and constraints and other guidelines are therefore needed. But our hopes are that these guidelines can be used for other AR technologies as well, and in inspiring future work. Our focus has been on mobile AR, but during the time we have learned many lessons that can be used in both projection based AR and head mounted AR as well. For example physical interactions through the device discusses in general the relation in space, and that can be applied even if a mobile device is not used. Similarly, in the chapter about interaction with the physical markers, we discuss the relationship between the physical and virtual world, which are key features that need to be considered even if no marker is being used.

When we designed our model and later also our guidelines, we based it on an iterative process with a lot of brainstorming. Where we were testing our own and others applications we found many possibilities and limitations. From these experiences and the experiences of other researchers we created guidelines. This mean that our guidelines are not completely comprehensive and can be further evaluated and redesigned. But its our belief that these guidelines are a good base for all kind of work with mobile AR. Our strength is our model where we have limited our self to think about these four possible kinds of media and for these we covered how to interact both physically and digitally.

The whole field of AR is a new field and constant improvements, both technical and theoretical, are being made. Some of the guidelines have therefore only been created to avoid technical limitations and we hope some of our advice can be redundant in the future. During our six month of work a lot of change has been made in the AR community. The mobile phones have gotten better cameras and and technical performance, which allows more advanced AR and makes existing AR solutions run even better and with better graphic. The big AR companies are also constantly developing their SDK’s, which provides more advanced features in the augmented world and new ways to interact with it. One of the technologies we believe can have a great future within mobile AR is the field of Markerless AR, such as Simultaneous Location and Mapping, since it makes the printed marker redundant and saves the user a lot of work.

Our hope is that this thesis can be used for discussion in designing mobile applications with AR content. Not only by academia, but also by mobile developing companies and indie developers alike. The Mobile Life has found Augmented Reality as the next natural step in the improvements of mobile society. As a leading mobile developer, they want to be one of the first to lay down the ground work for this new technology and they hope that our thesis can provide them with important insights into the future. In the near future, we hope we can use our guidelines in our own design work and improve open them, and even create new ones that did not make it into this paper.
10 Conclusions

Augmented Reality is a technology that has existed in concept for almost two decades. The last few years has seen the field change radically as the technology has become more available to consumers, and companies has started to put more resources in to developing commercial AR products. Even giants such as Microsoft and Sony has entered into the field with their Kinect and PS Vita platforms, and mobile handset producers such as Apple has started gathering lots of patents in the field. Nonetheless, AR is still a very young field and there are no defined standards of development processes since the technology is rapidly changing the possibilities and further removing limitations of the technology. Because of this, we set out to define interaction design guidelines based on development for the iPhone 4 platform, but with the idea of making these guidelines wide enough so they can be considered even when technology has progressed further than what is readily available today.

Interaction design is all about how the user interacts with the system, and also how the system interacts back to the user with feedback. In a sense, in AR context, user interactions also encompasses the sense of immersion since it blurs the edge between physical and digital and sometimes attempts to create a very tactile experience for users. We have classified our design guidelines based on this approach and attempted to include all the different ways a user interacts with such a system. Although our guidelines are in no sense exhaustive in terms of all the possible solutions available in an AR implementation, we feel that it presents a good sense of basic design principles and guidelines for a designer to follow when implementing AR on mobile platforms. One important conclusion we have reached when developing our design guidelines is that there is no one golden rule to follow when designing a AR system, but it is very important to consider the context in which the user is using the system, and the purpose the user is hoping to achieve, and act accordingly to choose the guidelines and design solutions that fit best into that scenario.

Augmented Reality is at its core a very human-centered technology where it is critically important to consider how a user perceives the world in which he or she is interacting, because the purpose of an AR system is to enhance the reality where the users exists. We feel strongly that it is more important than ever to put the user first when designing these kind of AR systems, and we hope that our presented guidelines might serve as a good starting point for anyone attempting to design such a system.
11 Acknowledgements

Working on this master thesis project has been incredibly interesting and has opened our eyes to a whole new field of technology and inspired us about the possibilities of this exciting new technology. We would like to take this opportunity to thank The Mobile Life, in particular Kenneth Andersson and Christopher Magnani for providing us with this subject and letting us work in their Singapore office which has been an incredible experience which we are truly grateful for. We would also like to thank our supervisor at the Department of Applied Physics and Electronics at Umeå University, Ulrik Söderström, for his support and feedback in the process of creating this report.
References


