Exploring Dual-Camera-Based Augmented Reality for Cultural Heritage Sites

RIGBY, JACOB,MARK

RIGBY, JACOB,MARK (2013) *Exploring Dual-Camera-Based Augmented Reality for Cultural Heritage Sites*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/8484/

Exploring Dual-Camera-Based Augmented Reality for Cultural Heritage Sites

Thesis for the Degree of Master of Science by Research

Jacob Mark Rigby

School of Engineering and Computing Sciences
Durham University
England

2012

Contents

List of Figures

List of Tables

Abstract

Context: Augmented Reality (AR) provides a novel approach for presenting cultural heritage content. Recent advances in AR research and the uptake of powerful mobile devices means AR is a viable option for heritage institutions, but there are challenges that must be overcome before high-quality AR is commonplace.

Aims: This project details the development of an AR "magic camera" system featuring novel dual-camera marker-based tracking, allowing users to take AR photos at outdoor heritage sites using a tablet computer. The aims of the project were to assess the feasibility of the tracking method, evaluate the usability of the AR system, and explore implications for the heritage sector.

Method: A prototype system was developed. A user study was designed, where participants had to recreate reference images as closely as possible using an iPad and the AR system around the University grounds. Data, such as completion time and error rates, were collected for analysis. The images produced were rated for quality by three experts.

Results: Participants responded positively to the system, and the new tracking method was used successfully. The usability study uncovered a number of issues, most of which are solvable in future software versions. However, some issues, such as difficulty orientating objects, rely on improving hardware and software before they can be fixed, but these problems did not affect the quality of the images produced. Participants completed each task more quickly after initial slowness, and while the system was frustrating for some, most found the experience enjoyable.

Conclusion: The study successfully uncovered usability problems. The dual-camera tracking element was successful, but the marker-based element encountered lighting problems and high false-positive rates. Orientating objects using inertial sensors was not intuitive; more research in this area would be beneficial. The heritage sector must consider development, maintenance and training costs, and site modification issues.

Declaration of authorship

I, Jacob Mark Rigby, declare that this thesis entitled "Exploring Dual-Camera-Based Augmented Reality for Cultural Heritage Sites" and the work presented in it are my own. I confirm that no part of the material provided has previously been submitted by the author for a higher degree in Durham University or any other University. All the work presented here is the sole work of the author. The research has been documented or is related, in part, within the publications listed below:

- Jacob Rigby and Shamus P. Smith. Novel Tracking Techniques for Augmented Reality at Outdoor Heritage Sites. In *Adjunct Proceedings of the 16th International Symposium on Wearable Computers*, Newcastle, UK, 2012.

Acknowledgements

I would like to thank my academic supervisor, Dr. Shamus P. Smith, for all of his help and support during this project. I would also like to thank all of the participants of my user study for giving their time to help me, Alastair Barber for providing image rating results, and my parents for their unconditional support in everything I do. This project was funded by a one-year scholarship from the School of Engineering and Computing Sciences, Durham University.

Chapter 1

Introduction

This thesis intends to investigate the use of augmented reality (AR) technology on mobile devices as a method of presenting cultural heritage content in an outdoor environment.

1.1 Background

Cultural heritage sites and museums are tasked with providing information about the past to members of the public in a clear and easily digestible manner, and in a way that does not require large amounts of the visitors' time. However, there are challenges to overcome to achieve this - walls of text overwhelm visitors, complex and specialist language is used where the readers have only a passing interest, and uninteresting pictures fail to engage. For this reason, the heritage sector often seeks new ways to engage visitors with their sites, and are often keen to harness technology to achieve this. AR interfaces are one such technology.

1.2 Objectives and research questions

This thesis documents the specification, design, implementation, and evaluation of a mobile AR system for use at an outdoor heritage site as a way of visualising 3D content. A novel hybrid marker-based and dual-camera switching tracking approach was employed as a way of testing the limitations of conventional marker-based tracking, whilst also attempting to utilise other hardware functionality in an attempt to find new tracking approaches. A user study was undertaken to evaluate the usability of the entire system, whilst also exploring the best ways to evaluate AR interfaces. Domain specific issues were also explored to uncover the issues that might be encountered by the heritage sector, which is often not technologically-minded.

The research questions can be summarised as follows:

1. *"Does a tracking system utilising two cameras on a mobile device present a feasible method of tracking for AR?"*
2. *"How does the use of a dual-camera paradigm impact usability?"*
3. *"What are the implications for the heritage sector if AR technology is adopted?"*

1.3 Thesis structure

This thesis is structured as follows:

Chapter 2: Literature review. This chapter explores past research relevant to the topics of AR, cultural heritage and mobile devices. The findings of this chapter were used a basis for the rest of the research.

Chapter 3: Design. This chapter details the design of a new AR system intended for outdoor use at a cultural heritage site, including the software engineering principles used and the rationale behind the decisions made. The literature review was used to inform the specification of requirements for the system.

Chapter 4: Implementation. This chapter describes the implementation of the system from the design in the previous chapter. Detailed in this chapter are the tools used to develop the system and the problems that were encountered.

Chapter 5: User study. This chapter details the method that was used to evaluate the implemented system, including the procedure of the usability study and the data that was collected. This chapter also presents the results that were obtained from the user study. Statistical analysis techniques were used to draw conclusions from the data.

Chapter 6: Discussion. This chapter discusses some of the overarching issues that have arisen as part of this research. These issues are discussed critically, and suggestions of how some of them could be solved are presented.

Chapter 7: Conclusion. This chapter presents the conclusions of the research, and sets out the areas where future work could be undertaken.

Chapter 2

Literature Review

2.1 Introduction

This chapter presents some of the existing research in the domains of AR and cultural heritage. The findings of this chapter were used to help inform the rest of the research presented in this thesis.

2.2 Background

Various technologies have been employed previously as a means of presenting cultural heritage content. These have included static computer kiosks to provide information to visitors [26], audio tour guides [34], mobile device-based tour guides [1, 19, 10], and even robot tour guides [15, 85].

AR also offers a novel and interactive way of presenting information to visitors. An AR system is a system which enriches ("augments") the real world with computerised information and objects (an example of this is shown in Figure 2.1). Milgram et al. [49] defined the Reality-Virtuality continuum, which provides a taxonomy for all Virtual Reality (VR) based systems and encompasses everything from

Figure 2.1: A mobile AR application showing a ruined building superimposed with an intact version.

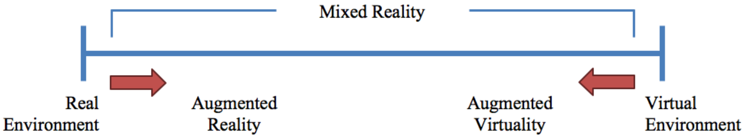


Figure 2.2: The reality-virtuality continuum (adapted from Milgram et al. [49]).

reality (the real world) to complete *virtuality* (a completely immersive, entirely virtual world). According to their classification, AR is a type of *Mixed Reality*, which describes any combination of reality and virtuality (see Figure 2.2). It should also be noted that their definition is not tied to a particular type of hardware, notably Head-Mounted Displays (HMDs), which have been a popular choice for AR systems.

Azuma [3] also notes that AR is not tied only to HMDs. Azuma builds upon the definition, and observes that in addition to being a blend of the real and virtual, AR should be interactive in real time and registered in 3D. Azuma considers that, although AR is most commonly associated with augmenting the users' sense of sight, it can also cover the other senses such as hearing [9, 35].

This chapter focuses on the application of AR to the cultural heritage domain, and covers common approaches taken and issues that arise. This particular section gives a general introduction to AR and cultural heritage. Section 2.3 shows how cultural heritage sites can be broadly divided into indoor and outdoor categories, and discusses the challenges faced from a location perspective. Section 2.4 provides an overview of some common approaches to tracking in AR systems. Section 2.5 presents some examples of the different hardware used in cultural heritage AR systems, and Section 2.6 presents common software toolkits and frameworks used to aid development. An summary of the trends and challenges faced in the future for different stakeholders is given in Section 2.7, and a summary table of the example systems detailed in this section is given in Section 2.8.

2.2.1 Augmented Reality for cultural heritage

The United Nations Educational, Scientific and Cultural Organization (UNESCO) defines a cultural heritage as the following [88]:

Monuments: architectural works, works of monumental sculpture and painting, elements or structures of an archaeological nature, inscriptions, cave dwellings and combinations of features, which are of outstanding universal value from the point of view of history, art or science;

Groups of buildings: groups of separate or connected buildings which, because of their architecture, their homogeneity or their place in the landscape, are of outstanding universal value from the point of view of history, art or science;

Sites: works of man or the combined works of nature and man, and areas including archaeological sites which are of outstanding universal value from the historical, aesthetic, ethnological or anthropological point of view.

This definition has since been broadened to also include artefacts, works of art, etc., such as those commonly found in museum exhibitions [18, p. 2]. In this document, the term “cultural heritage site” refers to any site of cultural significance. Cultural heritage sites can be broadly categorised into two groups - outdoor and indoor. Indoor sites often take the form of museums and visitor centres, where historic artefacts are displayed for the general public. An outdoor site could simply be a field where buildings once stood, or could be an entire city consisting of many (still intact) buildings.

There are an increasing number of applications for using AR for cultural heritage, including tour guides [51], virtual museums [94], serious games [12], and monument reconstruction [90]. One reason why AR is presenting itself as a feasible technology for the heritage sector is that it is now possible to develop applications for consumer-level mobile technology - many people now carry a powerful, lightweight, networked device in the form of “smart phones” and tablet computers. These devices are being purchased by an increasing number of people, and online distribution platforms such as Apple’s *App Store* provide an easy way for developers to deliver their software to consumers. Due to this, expensive specialist hardware need not be purchased by the heritage sites themselves, and large static computer kiosks need not take up often valuable exhibition space. Outdoor sites could particularly benefit, as in the past such systems may not have been possible due to concerns over weatherproofing and vandalism of installations.

2.3 Location issues

Both indoor and outdoor sites present many of the same challenges that must be addressed to successfully implement AR systems, including content acquisition [64], content storage and categorisation [43], tracking and calibration [4], marker placement, usability [31], and ergonomic issues [6]. However, there are different issues that must be overcome that are specific to either indoor or outdoor sites.

2.3.1 AR for indoor heritage sites

Previous applications of AR to indoor cultural heritage sites have often taken the form of "virtual museums", where visitors use AR technology to view objects that may otherwise be inaccessible to them. This may be due the value or fragility of such objects, or simply due to space limitations within the museum or because the physical object is at another museum. A prominent example of this the European Commission-funded ARCO project [94], which was designed to provide museum curators with a complete system to facilitate the creation and exhibition of a virtual museum using both Virtual Reality (VR) and AR techniques, from content acquisition, to content management, to content presentation. Models are acquired using photogrammetry techniques[1], which are then stored in a database where they are categorised and managed. Exhibitions are presented either via a web browser (i.e., VR) or on location (i.e., AR). The AR portion of the system uses pre-fabricated 3D models from the database combined with fiducial markers[2] to allow visitors to manually inspect artefacts. Visitors are able to manipulate the objects via manipulation of the markers themselves, allowing for rotation and adjustment of zoom level of the object which would generally not be possible in a "glass case" exhibition. Multiple objects can be viewed simultaneously, and users can select which objects are shown and which are hidden, which allows for a meaningful comparison of objects. Part of the ARCO system was also modified and presented separately by Liarokapis and White [41]. A similar application of AR was implemented by Caarls et al. [16], who extended the concept with the addition of rapid-prototyped clones of objects which were created with a 3D printer and combined with markers. This allowed visitors to touch and manipulate these prototypes, which were then

---

[1] Photogrammetry is the practice of extracting measurements from photographs [42, p. 1]. In this context, these measurements are used to create 3D models.

[2] Fiducial markers are planar markers which serve as a real-wold placeholder for a virtual object - a marker is detected by the system and the correct virtual object is superimposed over it (see Section 2.4.1).

augmented with 3D renderings. For an in-depth discussion of virtual museums, see the survey by Styliani et al. [83] and the article by Carrozzino and Bergamasco [17].

While fiducial markers can provide adequate tracking in situations where the camera has a good view of the marker and is not subject to constant and dramatic relocation, tracking for larger areas requires larger markers to be used and in greater numbers to maintain accurate poisoning tracking. However, the placement of such markers may not be desired due to aesthetic implications, or may not be allowed if it requires placing markers on features of a protected site.

Even though the “virtual museum” concept, in which 3D models of artefacts are exhibited, is a popular one, it is not the only application of AR for indoor heritage sites and museums. Miyashita et al. [51] also presented an indoor AR system for museums, but it was not for exhibiting digital models. Their system was developed for the Louvre in Paris for a temporary exhibition on Islamic art, and had two parts: an “artwork appreciation” component, which provided 3D information about important parts of the artwork to the user directly in front of the exhibition via a PC station and hand-held component; and a “guidance” component which guides users through the museum using an animated character via an ultra-mobile PC (UMPC). Due to the use of fiducial markers being disallowed, the project utilised a highly accurate (within 1mm) markerless tracking approach that performs well under low-lighting conditions that are common in such exhibitions.

The indoor portion of the Cultural Heritage Layers system presented by Zöllner et al. [98] also did not serve as a method of exhibiting artefacts, but instead allowed users to view a tabletop satellite image of Berlin augmented with a 3D model of the Berlin Wall and urban developments from 1940 - 2008. The same technology was used in conjunction with the Rome Reborn project[3] to present 3D Roman monuments which were augmented over a large floor map. Users would

---

[3] `http://www.romereborn.virginia.edu/` [last access 3rd December 2011]

walk over the map and point a handheld device at points of interest, which would then superimpose the 3D replicas on screen [99].

The main issues affecting the design of AR systems for indoor sites are those of marker placement if using marker based tracking, and ensuring that systems are easy to use for all age groups and levels of computer literacy. It is also important to ensure that hardware used is powerful enough to support AR applications, and that they are structurally robust if being lent to the public.

2.3.2 AR for outdoor heritage sites

Developing AR systems for outdoor applications is arguably more difficult than it is for indoor applications [4]. The environment and resources, such as lighting conditions and electrical power, cannot be as tightly controlled, and hardware cannot usually be left outdoors. This typically means that mobile computer systems must be used, which can be uncomfortably heavy to wear and expensive if it is a wearable system combined with an HMD. The lack of ideal conditions often means that marker-based tracking systems cannot be used, which leads to a reliance on other methods such as those based on Global Positioning System (GPS) and inertial sensors, which can be inaccurate. Nevertheless, numerous systems for outdoor sites have been successfully developed and implemented.

A significant example of an application of AR to outdoor sites is the ARCHEO-GUIDE project [90], which used an HMD and wearable computer combination to guide the user through an Ancient Greek temple site. The system also used a tablet computer to display location-sensitive information to the user, such as pre-rendered 3D reconstructions and images of archaeological finds, which are streamed to the device via a wireless network. The project was successfully deployed in 2001, using off-the-shelf hardware components. However, during testing it was found that the hardware was uncomfortable to wear for long periods [89]. The Augurscope project

[74] took a different approach. There were no wearable components as it combined a wheeled tripod with a computer and camera system, allowing users to wheel the device around a castle site in Nottingham, England. However, despite its mobility and that fact that it was designed to be used while moving, it was found that users were reluctant to move the device (perhaps due to the size and weight of the device and uneven ground surface), and that groups of people of different heights meant constant adjustment of the device was necessary. With these issues in mind, the Augurscope was later refined into the Augurscope II [75], which more people were willing to move it around the site.

Another outdoor system using mobile technology is MARCH [20]. Unlike ARCHEOGUIDE, the then contemporary (year 2009) consumer mobile phones had the power and features to implement an AR system. The MARCH system was developed for the Gargas prehistoric cave site in France, and uses AR to superimpose enhanced images of cave painting over the remains, which can be difficult to interpret. As the system uses a mobile phone, it is completely mobile and does not suffer from the comfort problems encountered by ARCHEOGUIDE.

The Mobile Augmented Reality Tour (MART) system [76] also demonstrates a mobile outdoor AR system. The researchers used popular tourist spots in Gyeongbokgung, South Korea to test the system's tracking technology. Using their system, 3D characters are correctly superimposed in numerous environments. However, even though the technology is targeted at mobile phones, the results presented were obtained from a prototype that ran on a laptop.

The main issues faced in designing AR systems for outdoor sites are those of tracking effectively without the use of markers in an environment that may be devoid of features to use for tracking, and also ensuring any apparatus used is weather-proof and vandal-proof. Also, as with indoor sites, hardware used must be powerful enough to support AR applications. There is also the issue of making software available to visitors - if the user has to download and install software then there needs to be infrastructure in place to allow this.

2.4 Tracking

For an AR system to be effective, the position of the device in the world must be accurately tracked so that virtual objects can be superimposed in the correct location. AR generally requires tracking in 6 degrees of freedom (DoF) - x, y, and z combined with pitch, roll, and yaw - to enable digital data to be superimposed seamlessly into the real world, though there are some exceptions [62]. A number of methods are popular, including marker-based, inertial, optical, and location-based. However, none of these technologies are perfect. For example, GPS can be too inaccurate; inertial sensors are subject to drift (loss of accuracy over time); and optical methods can be computationally expensive. As such, hybrid approaches that mitigate the shortcomings of using a single technology are common [82, 95, 68, 67]. The suitability of each method depends on many factors, primary of which is the domain location (some methods are more suited to indoor than outdoor and vice-versa), but also target hardware, and how sensitive the domain is to modification.

It is important to distinguish tracking from calibration. Calibration refers to the initialisation phase of a tracking system (for example, determining the initial position of a device using GPS), and tracking refers to the continued re-evaluation of the scene and device position so that objects are correctly located. However, much of the literature reviewed for this project to does not discuss calibration in much detail, or it is discussed as part of the tracking system and not explicitly mentioned. Also, sometimes calibration is not actually a separate initial phase of the tracking process, but happens during every step of the tracking process. An example of this is in some marker-based tracking systems, where calibration happens every frame.

2.4.1 Marker-based tracking

Planar marker tracking systems employ the use of a camera to detect markers placed in the real world that are used to describe the position and orientation of virtual objects. The use of fiducial markers as a means of tracking is widely used in the field of AR in general, due to its efficient performance and the ease and inexpense with which markers can be produced and placed. While this is still considered a method of optical tracking, due to it being so common and such a large category it is in this document categorised it as distinct from other optical methods, which are described in Section 2.4.3. There are some disadvantages to using marker-based tracking: it is (i) only suitable under good lighting and visibility conditions; (ii) generally not feasible for outside use; and (iii) the markers may not be aesthetically pleasing or permitted for use. To make markers less intrusive, it is possible that a marker system based on watermarks could be used [39], or even completely invisible markers using infrared ink [61]. Tool support for marker-based tracking solutions are described in Section 2.6.1. Cultural heritage AR systems that have used marker-based tracking include the ARCO project [94], the system presented by Caarls et. al [16], and MARCH [20].

The ARCO project [94] used a fiducial marker based system for its tracking system, which was based on that of ARToolkit (see Section 2.6.1). As this was to exhibit virtual objects, the placement of markers was acceptable, as it did not involve placing them on walls or other features.

The system presented by Caarls et al. [16] also used marker-based tracking system, which used markers very similar to those of ARToolkit.

The MARCH system [20] used a custom marker-based tracking system with unique “colour target” markers [22]. The marker detection system was designed to be run in real time on mobile phones, but the actual MARCH system achieved performance of only 14 frames per second.

2.4.2 Inertial sensor-based tracking

Tracking via inertial sensors in a device (for example the gyroscopes and accelerometers found in many modern mobile phones) provides a method of tracking that is entirely internal to that device, i.e., the sensors are not reliant on any markers or other electronic devices after any initial calibration or set-up stage. However, inertial sensors are subject to drift [96, 53], and some are sensitive to environmental changes such as electromagnetic interference [87, p. 204]. Also, as an entirely inertial system has no visual input it cannot account for object occlusion[4]. Tracking entirely by inertial sensors is rare in AR; it is usually combined with another method to increase accuracy. No example could be found within the cultural heritage domain. Commonly used inertial sensors are gyroscopes, which measure orientation, and accelerometers, which measure acceleration. These are commonly paired to allow accurate position and orientation tracking, both in AR and for a number of other applications such as ship navigation [37, p. 642].

2.4.3 Optical tracking

Optical tracking methods typically achieve tracking by detecting environmental geometry like building corners or picture frame edges, such as the Speeded Up Robust Feature (SURF) [8] and Scale-Invariant Feature Transform (SIFT) [45] methods [30, 7], and other techniques that compare the current scene to reference images [84, 90]. However, this is not always possible if there are no easily-distinguishable features in the environment, such as an archeological site in a field [55]. Cultural heritage AR systems that have used optical tracking include the Interactive Museum Guide [7], the GAMME project [86], Cultural Heritage Layers [98], and the Augmented Reality Presentation System for Remote Cultural Heritage Sites [99].

---

[4] Object occlusion in this context refers to when a real-world object is placed in front of a virtual one. The desired result of this is the virtual object being partially or totally obscured by the real-world object.

The Interactive Museum Guide [7] utilised SURF, where the objects in the current scene are compared to a database of images of objects and matched to their respective “interest points”. The image with the greatest number of matches is selected in an attempt to select the object that the user is looking at. This approach does not require objects in the scene to be compared to photos of each object in isolation, which can increase production speed of the database of images. It was also found to work with a low quality camera, and does not need colour images.

The Cultural Heritage Layers system [98] utilised a similar entirely optical tracking system which used a two-phased approach. An initialisation stage was used for calibration and initialistaion, where input from the camera is analysed to see if it matches a number of pre-specified “spots”, and a tracking phase was used to update the position of objects every frame after a spot has been detected. Knowledge of the spots is supplied to the system in the form of reference images, and for each of these a tree data structure of easily-detectable points is created and stored. This can be done before the system is used. Their tracking system had two modes, one for indoor environments and one for outdoor. The indoor tracking mode was intended for use with objects close to the user, and would accurately track in 6DoF. The outdoor mode assumed that objects would be far away and assumed a panoramic view around the user, and because of this would only track in 3DoF to measure the change in rotation of a stationary user [97]. The system used only 2D overlays as opposed to 3D models in an effort to maximise performance, but still only ran at 15 frames per second.

2.4.4 Location-based tracking

This category of tracking relies on satellites or beacons to calculate the position of the device.

GPS is another commonly used method of tracking for outdoor applications. GPS requires a GPS receiver and un-occluded line-of-sight to 4 or more GPS satellites to calculate the device's position, and is therefore only suitable for outdoor use but often performs poorly in urban environments [32, p. 338]. Another disadvantage of GPS is that most consumer-grade receivers are only accurate to within a few metres, and can often be out by as much as 20 metres [40], and such receivers integrated into mobile devices are just as inaccurate [11]. This lack of accuracy means that GPS is rarely used as the sole tracking technology, but is commonly paired with another technology so that the GPS provides coarse-grain tracking to give a rough idea of location, which is refined by the other tracking method to provide the necessary accuracy. The only AR cultural heritage system found that used only GPS for tracking was the outdoor portion of the Cyberguide system [1].

Infrared beacons can also be used as a means of tracking [69]. A cultural heritage system that used this approach is the indoor portion of the Cyberguide tour guide system [1], where multiple beacons were placed around an indoor setting.

It is also possible to use wireless networks as a means of tracking, where wireless network devices act as beacons to allow for the calculation of location [36, 65]. However, no examples could be found in the cultural heritage domain.

2.4.5 Hybrid tracking

Hybrid tracking systems are common as they ensure the high-accuracy of position tracking needed for AR, and using one tracking method in tandem with one or more others can make up for technological shortcomings in each tracking method respectively. Cultural heritage AR systems that have used hybrid tracking include ARCHEOGUIDE system [90], the Augurscope [74], the Augmented Reality Museum Guide [51], and MART [76].

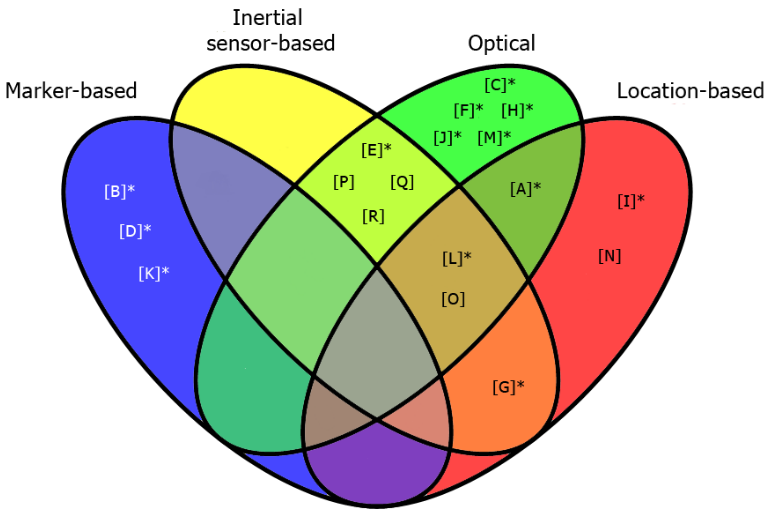
The ARCHEOGUIDE system [90] utilised a custom optical tracking approach combined with GPS position data. The GPS receiver would first give a rough estimation of position and viewing angle, and then the optical tracking algorithm would further refine this by attempting to match the users current view to a series of images stored in a database on a frame-by-frame basis. This approach supported 15 frames per second at a resolution of 320 x 240 pixels.

The Augurscope [74] used a hybrid GPS and inertial sensor system. The results were found to be satisfactory, but there were problems with the accuracy of the GPS which was only accurate to within 2 - 4 metres and had a 2 second update time, meaning that without any further software smoothing the user would experience a lot of jittering.

The MART system [76] combined GPS, inertial sensors and optical tracking. Its tracking model consisted of a *sensor-based tracking flow* which encompasses readings from GPS and inertial sensors, and a *vision-based tracking flow* which matches the current camera scene to a database of reference images. These tracking flows can then be combined to make a *hybrid tracking flow*, which also incorporates other readings, such as those from light sensors.

2.4.6 Summary

Figure 2.3 shows the tracking technologies used by the example systems detailed in this chapter. This shows that cultural heritage AR systems generally favour the use of markerless or optical tracking. It can also be seen that systems that use marker-based tracking do not tend to use any auxiliary tracking methods. This may be because marker-based tracking is only chosen where the lighting conditions can be controlled, and when the user is intended to be up close to the objects in the system, for example in a museum setting - because of this, marker-based methods alone may be perfectly adequate. The diagram also shows that tracking systems that combine three or more tracking technologies are rare.



System	Tracking	Key
ARCHEOGUIDE [90]	Optical, GPS	[A]*
ARCO Project [94]	Marker-based	[B]*
An augmented fine-art exhibit [86]	Optical	[C]*
Augmented Reality for Art, Design and Cultural Heritage [16]	Marker-based	[D]*
Augmented Reality Museum Guide [51]	Optical,inertial	[E]*
Augmented Reality Presentation System for Remote Cultural Heritage Sites [99]	Optical	[F]*
Augurscope [74]	GPS, Inertial	[G]*
Cultural Heritage Layers [98]	Optical	[H]*
Cyberguide [1]	IR (Indoor), GPS (outdoor)	[I]*
Interactive Museum Guide [7]	Optical	[J]*
MARCH [20]	Marker-based	[K]*
MART [76]	Optical, inertial, GPS	[L]*
Mixing virtual and real scenes in the site of ancient Pompeii [59]	Optical	[M]*
[65]	WiFi	[N]
[67]	Optical, inertial, GPS	[O]
[68]	Optical, inertial	[P]
[82]	Optical, inertial	[Q]
[95]	Optical, inertial	[R]

Figure 2.3: Venn diagram and legend showing tracking technologies used by AR systems used in this chapter. Items marked with an asterisk (*) are systems specific to the cultural heritage domain.

2.5 Hardware

At a basic hardware level, an AR system requires four components: a computer system; a camera; a display; and some kind of tracking mechanism (though this may be a software component). There are three main categories of hardware used in AR systems: fixed computer systems; wearable computer systems; and mobile devices.

2.5.1 Fixed computer systems

AR systems can be used with static computer terminals with attached cameras. Some examples in the cultural heritage domain are the museum guide system for the Louvre [51], where a small industrial camera was used in conjunction with a PC for their presentation room, and the MovableScreen [99], which mounts a 24” iMac on a pillar that can be rotated 360 degrees. The advantages of such systems include a potentially stable power supply and network connection if mains power and a wired connection is used, and they are generally easy to develop for as they are often simply standard desktop PCs. They are also able to be used by more than one user at a time so the AR experience can be shared amongst families or groups of visitors. Disadvantages include lack of mobility and the large amount of space they take up.

Although the Augurscope [74] was designed to be moveable, it was essentially a fixed system mounted on a wheeled tripod (the system was not designed to be used when moving the tripod; the wheels served only as a method to move between locations). The system used a laptop computer to minimise weight, and an attached camera.

2.5.2 Wearable computer systems

While not required, many AR systems traditionally used HMDs as the display device [3]. Two types of HMDs are available today: optical see-through, where a semi-transparent surface placed in front of the eyes allows the user to see both the real world and have digital images reflected into their eyes; and video see-through, where one or more cameras provide a video feed of the real world, which is then combined with digital images [71]. Unfortunately HMDs can be very expensive and uncomfortable, and usually have to be paired with some kind of wearable computer. Nowadays, integrated displays such as those featured in smart phones and tablets are preferable because of their favourable size, inexpensiveness, and durability.

Cultural heritage systems that have used wearable computers include ARCHEOGUIDE [90] and the system developed for Pompeii by Papagiannakis et al. [59]. However, for many of the systems that used wearable computers, the then-current mobile hardware such as PDAs, tablets and smart phones were not powerful enough, so to provide enough computational power this often meant a high-end laptop would need to be used which would often be heavy and uncomfortable [5]. However, advances in off-the-shelf consumer hardware now means that contemporary mobile devices can provide enough computing power for such applications.

2.5.3 Mobile devices

Due to ubiquity of powerful mobile devices such as smart phones and tablets, mobile devices are becoming popular platforms for AR systems. Table 2.1 shows the specifications of four popular mobile devices (a number of similar devices are available from competing manufacturers). These devices also feature accelerometers and gyroscopes suitable for position tracking in 6DoF. Their powerful features make all these devices suitable devices for a high-quality AR experience. This means that, as visitors will already own their own devices, cultural heritage institutions that

Table 2.1: Specifications of popular mobile hardware.

Device	**CPU**	**RAM**	**Display**	**Cameras**
Apple iPhone 4S	800MHz dual-core	512MB	3.5”, 640 x 960 px	8 MP back, 0.3 MP front
HTC Rezound phone	1.5GHz dual-core	1GB	4.3”, 1280 x 720 px	8 MP back, 2 MP front
Apple iPad 2	1GHz dual-core	512MB	9.7”, 1024 x 768 px	0.7 MP back, 0.3 MP front
Samsung Galaxy Tab 10.1	1GHz dual-core	1GB	10.1”, 1280 x 800 px	8 MP back, 2 MP front

take of advantage of this will not have to procure expensive specialist equipment to lend to visitors, and greater homogeneity of hardware platforms makes it easier to develop software. However, with multiple software platforms being popular (such as Android, Windows and iOS) it can mean that multiple versions of the same software need to be produced even though the hardware configuration is largely similar.

The presence of both forward- and backward-facing cameras on these devices as opposed to only a traditional backward-facing one is significant: a backward facing camera only allows the user to see an augmented view of the scene in front of them (a *magic window* or *magic lens*), whereas devices with forward-facing cameras also allow the user to see and augmented view of themselves (a *magic mirror*). This would allows visitors to heritage sites and museums to virtually “try on” historical clothing, or even wear suits of armour, and enable them to view themselves wearing such items as well as viewing their friends and family, making for a more collaborative experience which aligns itself well with the family audience that many heritage sites aim to attract [28].

Example AR systems for cultural heritage that have used mobile hardware include the Cultural Heritage Layers system [98] that utilised a Sony UMPC as a mobile device for users, the museum guide system for the Louvre [51] that utilised a Fujitsu UMPC, the MARCH system [20] system that utilised a Nokia N95 mobile phone.

While more powerful hardware is becoming ubiquitous, real-time processing for AR systems (especially tracking and rendering) can still be very computationally expensive [93]. To combat this, it is possible to take advantage of the networking features of some devices to move a lot of the processing to a more powerful server [33, 91]. However, this obviously relies on the presence of a server and a stable network infrastructure.

In addition to the hardware used by the users, it may also be necessary to install other hardware at a site, for example to provide infrastructure for networking or tracking. This obviously has cost implications, and it may not be desirable to leave expensive equipment outside in all weather conditions or at unmanned sites.

For an in-depth discussion of mobile technologies for AR, see the survey by Papagiannakis et al. [60].

2.6 Software

Many past cultural heritage AR applications have used mobile devices, and the designers chose to use a client-server architecture, storing content on a central server and streaming it to the user's device over a network [90, 94]. This makes sense if there is a lot of content as mobile devices are often lacking in storage space, and it means that only content specific to the user's location and interests need be streamed (assuming the system is aware of them) [23].

It is possible to view the structure of an application from a development perspective in layers, as in Figure 2.4. The operating system and APIs used are largely dependent on the choice of target hardware (for example, if developing for the

Apple iPhone the Apple iOS and Apple development tools must be used) but there is more freedom when it comes to toolkits. There are many toolkits available to aid the development of AR applications, mainly to provide robust tracking solutions without the need to code a system from scratch. When developing for AR applications, as with any other kind of development, a number of decisions must be made including those regarding software architecture, target platforms, toolkits, etc. Many AR systems use custom tracking methods, but some developers do not wish to "reinvent the wheel" and so make use of toolkits to save time. A good analogy is that of computer game design, where it is common for game developers to use a licenced game engine, as well as other middleware, which has been developed by another company. This allows the developers to focus more on content creation and reduce development time, but there are the downsides of less customisation and licence fees. Some commonly used toolkits are presented in the following section, which focuses mainly on tools that are available to use for free.

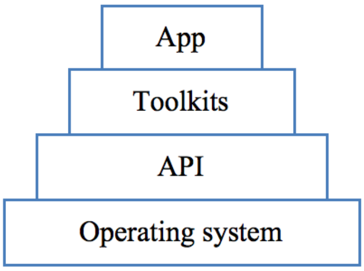


Figure 2.4: The layered structure of an application.

2.6.1 Tool support

A common library used for marker-based tracking is ARToolkit[5], a free, multi-platform C and C++ tracking library which supports both video- and optical-see through AR. It is well supported by documentation and tutorials, and there is an

---

[5] `http://www.hitl.washington.edu/artoolkit/` [last access 2nd November 2011]

active user community who are able to communicate via mailing lists and forums. The library works by converting the video input into a binary image and detecting the large black frame around the edge of the marker (Figure 2.5 shows an ARToolkit marker). ARToolkit then calculates the relative position and orientation of the marker to the camera and matches the symbol on the marker to those stored in pattern files (for storing 12 possible orientations of the marker) to select the correct virtual object. The object is the transformed to align it with the marker and then rendered[6]. Markers are customisable by the developer, so images relevant to the domain can be used. There is also the option to use barcode-style markers. There are, however, some drawbacks to ARToolkit. It often detects markers where there are none (a *false positive* reading), and often confuses markers (*inter-marker confusion*). Despite some shortcomings, ARToolkit is arguably the most popular software toolkit for developing AR applications, and is available for numerous platforms and hardware configurations including popular mobile platforms such as Google's Android and Apple's iOS. It is open source and free to use for non-commercial use, but also commercially licenced for professional use.

Figure 2.5: An ARToolkit marker.

ARToolkit plus, an extended version of ARToolkit, was designed as its successor [92], but both libraries were continually developed alongside each other. ARToolKitPlus has itself since been succeeded by the Studierstube Tracker [73], and Studierstube ES [72] for embedded systems and mobile phones, but these are not

---

[6] `http://www.hitl.washington.edu/artoolkit/documentation/userarwork.htm` [last access 2nd November 2011]

publicly available.

ARTag is another marker-based tracking library, which was created in an attempt to address some of the shortcoming of ARToolkit - namely the high rate of false positives and frequent inter-marker confusion. ARTag differs from ARToolkit in that it does not match the images to pattern files (the default and most common type of marker used in ARToolkit), but instead uses a digital symbol method which encodes a binary code with checksums and error correction redundancy, inspired by the Data Matrix barcode system [29]. Figure 2.6 shows an ARTag marker. The ARTag library is free for non-commercial use.

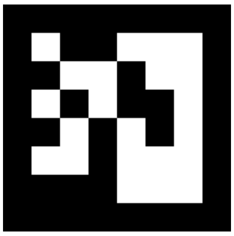


Figure 2.6: An ARTag marker.

Another derivative of ARToolkit is OSGART (ARToolkit for OpenSceneGraph) [44], which combines ARToolkit functionality with that of OpenSceneGraph[7], a 3D graphics API which allows a 3D scene to represented in a graph data structure. OSGART does not only provide tracking functionality, but also advanced visualisation features such as shadows and reflections. It is cross-platform, and free version is available under the GPL licence for non-commercial use.

The Designer's Augmented Reality Toolkit (DART) [47] provides a high-level way of developing rapid prototypes of AR software by using a scripting language. It is aimed more at designers than computer scientists, and attempts to provide an easy way of creating content and making AR applications without the need for in-depth technical knowledge.

---

[7] `http://www.openscenegraph.org/` [last access 17th November 2011]

2.7 Trends and challenges

As research has intensified over recent years and technology has improved, AR now presents itself as a feasible option to the heritage sector. A number of state- and EU-funded projects have seen fruition in the form of ARCHEOGUIDE [90], the ARCO project [94], iTacitus[8], and the GAMME project [86], showing that AR is not just suitable for research environments but also for real museums and heritage sites, and that it is being taken seriously by government funding councils as a method of bringing heritage to the masses. However, there are still many challenges that must be addressed before AR becomes commonplace in the heritage sector.

2.7.1 Technology

The most obvious trend observed is the increase in uptake and development of powerful mobile devices that make mobile AR applications possible, and the increasing number of applications for such devices that apply AR is testament to this. Due to the ever-increasing power of mobile devices and increasing interest in AR, such applications are predicted to become even more pervasive. However, while many potential visitors may own the necessary AR-enabling hardware, this still only represents a small percentage of potential visitors. Many mobile devices are expensive, luxury items, and there are accessibility issues associated with this. Most museums and heritage sites are keen to attract as many people as possible, and if users must purchase expensive hardware to get the full experience then this does not present a high level of accessibility. The issue of cost is also an important one for both the visitor and the heritage site or museum. Most heritage sites and museums try to attract families to visit together, but few families can afford to buy every family member an expensive tablet or smart phone. From the perspective of the heritage site or museum, AR systems can be expensive to develop and deploy

---

[8] `http://www.itacitus.org/` [last access 14th December 2011]

which may not be cost-effective for them if only a small number of visitors will be using it. A solution for this is to provide hardware to loan out to visitors, but the added cost this incurs for the museum and the rate at which hardware must be replaced may not be attractive, and it may not be desirable to allow visitors to borrow expensive and delicate hardware especially if they are to be used by young children.

2.7.2 Location

Many heritage sites are protected from modification of any description, and this can even extend to the placing of fiducial markers or other necessary items to facilitate AR. This narrows the choice of available technologies for some sites, meaning that markerless AR must be used or that other unobtrusive tracking technology must be used. Even if the placing of markers is permitted, it can be undesired for aesthetic reasons.

Many heritage sites and museums disallow the use of cameras and photography, and some disallow the use of mobile phones. If such devices are not permitted, this can preclude the use of mobile AR. A solution to this problem may be to have clearly defined zones in which the devices may be used for AR, but this could be difficult to enforce. Another problem is that visitors using devices for AR while moving can present health and safety issues. If visitors are not paying attention to where they are walking they may trip and fall, or walk into exhibits or other people. Clearly defined, uncluttered AR zones could also go some way to solving this problem.

From a technological perspective, there is still much to be achieved with regards to tracking, specifically for outdoor rural sites such as those situated in fields.

2.7.3 Development and deployment

For developers and programmers, developing for AR applications can be problematic because of the lack of easy-to-use development environments tailored towards AR. The Designer's Augmented Reality Toolkit [47] sought to address this problem by providing an environment to support high-level, rapid prototyping of AR systems without the need for extensive low-level programming. However, the system was designed to run on top of Macromedia (now Adobe) Director which may constrain development somewhat, as well as leading to decreased performance and larger software footprint. No example systems using this toolkit in the cultural heritage domain could be found. While ARToolkit and other toolkits ensure developers to not have to completely invent the wheel, a series of toolkits and frameworks that specifically aid the development of cultural heritage applications (which may be fairly similar across a wide range of sites) would be welcome as a means to reduce time and cost levels.

The type of hardware used can also affect development time and cost. If a standard desktop PC is used, such as in a kiosk system, development is relatively easy compared to a mobile platform as the developers are free to use any software and tools they wish and they will only need to support a single platform. If mobile hardware is to be used then developers are often constrained as to which development tools they can use, and they may be required to support multiple platforms. This in turn brings about licensing issues for software used and the software in developments - for example, developing for Apple hardware and distributing via their *App Store* requires the use of Apple's own development tools, as well as requiring Apple to accept the software on their terms for distribution.

There is still much to be achieved from a cultural heritage perspective. Many of the example systems in this chapter exist only as research projects, often as a vehicle to test new tracking technologies or hardware, instead of providing an entire system for use by heritage institutions. The majority of those working in the heritage sector are not technology experts, and curators and other heritage specialists should not be expected to have to acquire such knowledge in order to use or manage such systems. The ARCO project [94] went some way to address this by providing an entire system the facilitate inventory management and presentation, but the project does not seem to be in active development. The ARTECT tool [38] created a high-level authoring tool for non-technical heritage workers to use, which focused on creating an experience-centred toolkit rather than a software-centred one. Multiple iterations were developed with input from domain experts, and the project was tested and found to be fit for purpose. However, development of the toolkit seems to have ceased, and only a prototype version exists for testing purposes[9]. The CHESS project[10] is a project that is also developing authoring tools for non-technical personnel, in order to allow them to create an interactive storytelling experience for a variety of audiences [66]. The project is still in progress, and the associated software tools have not yet been publicly released. Bruno et al. [14] presented a methodology to facilitate the creation of a virtual museum in a cost-effective manner, from digitising archaeological finds to presenting them with their MNEME software system and portable hardware. However, it still required a team of specialists to digitise the objects over a lengthy three-month period. Also, it used virtual reality as opposed to AR, but a similar methodology could be applied to AR exhibitions.

The majority of AR systems developed to date, including those for cultural heritage, give the user a primarily passive experience with little interaction. Tour guide systems, for example, often simply provide users with context-sensitive information which requires no reaction from the user. However, some systems facilitate basic interaction with objects. The ARCO Project [94] allowed users to move mar-

---

[9] http://www.cs.nott.ac.uk/ ktg/install.html [last access 5th October 2013]

[10] http://www.chessexperience.eu/ [last access 5th October 2013]

kers, and the system presented by Caarls et al. [16] allowed users to touch replicas of objects, but the user is only moving real-world placeholders and not really interacting with the software system. One way to encourage interaction is through games. *Serious games*, games that are used as educational aids rather than purely for recreation, have been widely used in the cultural heritage sector (for further reading on this, see the state-of-the-art review presented by Anderson et al. [2]), but serious games that employ AR have not been widely explored. AR serious games could be an extremely effective way of engaging visitors, especially children. Digital characters could not only inform visitors about exhibitions, but also ask them to complete tasks such as scavenger hunts, or ask them to answer a series of questions about what they have seen. One system that did attempt an AR game in a museum environment was the Mobile Augmented Reality Quest (MARQ): Expedition Schatzsuche[11], which implemented a team-based treasure hunt-style game on top of the Studierstube ES tracking system[72]. However, the project ceased in 2007 and little academic literature was produced from it. Games could also allow visitors to participate in a story (much like modern-day computer games) from the moment they enter the heritage site, which could develop as they progress through the site, performing tasks as they go. Such games can be made all the more believable using today's immersive AR technology. If collaboration (or competition) between users, such as in MARQ, was also possible then it may be especially of benefit to groups of schoolchildren.

2.8 Summary

This chapter has shown that AR systems within the cultural heritage domain are becoming more and more feasible, mainly thanks to the uptake of powerful consumer-

---

[11] http://handheldar.icg.tugraz.at/marq.php lLast access 5th October 2013]

level mobile hardware and technological improvements in tracking systems. The problems that must be overcome for such systems to be successful are also highlighted; these include the limitations of current tracking technologies (especially the lack of accurate tracking systems in outdoor environments), the difficulties of developing AR software without established standards, frameworks, and easy to use toolkits, the cost of hardware, and the cost and duration of system development and content acquisition.

Table 2.2 presents the example cultural heritage systems used in this chapter, comparing the technologies and techniques used by each system.

Table 2.2: Example AR systems for cultural heritage discussed in this chapter.

Name	**Indoor/outdoor**	**Tracking**	**Hardware**
ARCHEOGUIDE [90]	O	Optical, GPS	HMD with wearable laptop, PDA
ARCO Project [94]	I	Marker-based	Static terminal
An augmented fine-art exhibit [86]	I	Optical	UMPC
Augmented Reality for Art, Design and Cultural Heritage [16]	I	Marker-based	HMD with wearable PC, static terminal
Augmented Reality Museum Guide [51]	I	Optical, inertial	UMPC, static terminal
Augmented Reality Presentation System for Remote Cultural Heritage Sites [99]	I	Optical	Rotatable fixed screen, UMPC
Augurscope [74]	O	GPS, inertial	Tripod-mounted PC
Cultural Heritage Layers [98]	I/O	Optical	UMPC
Cyberguide [1]	I/O	IR (Indoor), GPS (outdoor)	UMPC
Interactive Museum Guide [7]	I	Optical	Tablet PC with attached camera
MARCH [20]	I/O	Marker-based	Mobile phone
MART [76]	I/O	Optical, inertial, GPS	Laptop with attached camera, mobile phone
Mixing virtual and real scenes in the site of ancient Pompeii [59]	I/O	Optical	Mobile workstation

Chapter 3

Design

3.1 Introduction

Included in this chapter is the design of the system itself from requirements (including changes that had to be made), and the design of the user study which was used to evaluate the software.

The general scenario envisaged for a proposed system was as follows: the system would be in the form of an application on a tablet computer, which the user would carry around with them when visiting an outdoor heritage site. The user would then use the system to allow them to view structures and features that are no longer present at the site in question, such as buildings or earthworks, in situ. The use of AR technology would allow them to view these from different angles and positions, which is not possible with traditional heritage interpretation material. The design of the actual system decided upon to facilitate this is described here.

3.2 Software process model

Software process models are an abstract way of describing a software process [80, p. 8]. This project utilised an *evolutionary development model*, whereby the system is repeatedly refined over multiple iterations of requirements and software. Sommerville [80, p. 68-69] suggests that evolutionary development is the best approach to take for small and medium-sized systems, and where the system requirements are not well understood and/or likely to change. This made such an approach ideal for this project, which was a small project with a single developer, and where maturing technology was used which made it likely that requirements would change if some elements of the technology were found to be incapable or unsuitable.

3.3 Evolutionary protoyping

Prototyping allows software with an incomplete set of features to be used and interacted with by stakeholders. They can take the form of *throwaway protoypes*, of which multiple are developed and then discarded, or *evolutionary prototypes*, where a single prototype is continually refined and developed into the final system [70, p. 390]. This project used an evolutionary prototyping approach, where the system was continually added to and modified so as to suit the changing requirements. This approach to development works well with the use of an evolutionary development model.

3.4 Requirements

Before designing a system, it is very important that the requirements for the system are well understood so as to avoid serious problems later in the software life cycle [80, p. 75]. Typically, developers will consult their clients and other stakeholders in this process, but as this was a research project (where a greater amount of freedom is afforded), there were no clients as such. However, the needs of the target users (heritage site visitors) and potential clients (heritage sites) could still be considered. Furthermore, many of the requirements for this system were drawn from the results of the literature review (Chapter 2), which highlighted elements of previous systems that worked particularly well or particularly badly, as well as showing areas where new ideas and research could be developed.

3.4.1 Functional requirements

The functional requirements of a software system describe what the system should do and the functionality it should provide [80, p. 120].

The literature review shows that while AR systems are reasonably successful in controlled indoor environments, there is still some way to go before AR systems in outdoor environments reach a similar level of robustness. For this reason, it was decided that a system would be developed to be used in an outdoor heritage site environment.

The literature review also shows how Augmented Reality systems can now take advantage of powerful mobile hardware. For this reason, the decision was made that the system that would be developed would be a mobile system utilising a touch screen tablet device with a large screen, so as to facilitate portability whilst still being as immersive as possible.

The tracking system was an important consideration. As stated in Section 2.7.2, many heritage sites are protected from modification. This means that installing any hardware or other apparatuses often disallowed, but it seems feasible that the placement of fiducial markers would be permitted, even if only temporary. For this reason, the use of a fiducial markers for tracking seemed plausible. Furthermore, there are numerous software toolkits available to aid the development of such a system (see Section 2.6.1), and it also provided an opportunity to evaluate the use of a marker-based tracking system in an outdoor environment.

Many of the systems cited in the literature review are the work of teams of people, often over a long period of time. Due to this, it was decided that third-party toolkits and software would be used as much as possible, in order to minimise development time spent on low-level tasks that had already been successfully implemented by others.

The initial list of functional requirements for the system was as follows:

IFR1: The system should be used in an outdoor setting

IFR2: The system should allow for the visualisation of virtual objects over the camera feed

IFR3: The system should include at least one virtual object representing a historical structure

IFR4: The system should run on a mobile device with a touch screen

IFR5: The system should allow users to take photographs of the scene

IFR6: The system should be able to store virtual objects of multiple structures or features

IFR7: The system should allow the user to walk around the site whilst visualising virtual objects

3.4.2 Non-functional requirements

The non-functional requirements of software system describe requirements that are not concerned with specific functionality provided by the system, but rather impose constraints on the system. Such constraints may be related to reliability, timing and the development process [80, p. 119-120].

The initial list of non-functional requirements for the system was as follows:

Usability:

INFR1: The system should allow the user to visualise different objects at their respective locations without having to restart the application

INFR2: The system should be usable by users of varying heights

Accessibility:

INFR3: The file size of the entire application should be no larger than 100MB to allow users to download it via a mobile internet connection if necessary

Deployment:

INFR4: Any additional equipment or materials needed should be used in a way that requires the least amount of site modification, and should not permanently modify the site

3.5 Design decisions

After having produced a list of requirements for the system, a number of decisions had to be made regarding the hardware and software to be used before development could commence.

3.5.1 Tools and development kits

A number of different tools are available to aid in the development of AR systems (see section 2.6.1). The use of ARToolkit was decided upon, which is very popular in modern AR systems [21]. It provides a robust marker-based tracking suitable for the proposed system, and is well supported by the developers. It is also frequently updated, and supports the latest version of iOS and the iPad and Android devices.

3.5.2 Software platforms

There are two major software platforms for modern mobile devices in use today: Apple's iOS and Android, which is led (but not owned) by Google. One of the main considerations for the choice of software platform was the ARToolkit support. It seemed that the iOS version of ARToolkit was more robust than the Android version, and looking at the support forum[1] showed 237 topics and 973 posts in the ARToolkit for iOS forum versus 29 topics and 98 posts in the Android forum, suggesting a far more active developer community and better user support for the iOS version.

The iOS platform is very versatile and robust, having been through many iterations and currently being at version 5.11 at the time of writing. iOS applications ("apps") are written using the Objective-C programming language, a version on the C programming language extended with extra features. Some extra time was required to become properly versed in these features, and new syntax had to be learned, but an existing working knowledge of the C programming language provided an adequate foundation.

3.5.3 Hardware platforms

Having decided to develop the system on a tablet computer, the exact make and model of the device needed to be finalised before development could commence, as this will impact other design and development choices. Numerous tablet devices are available from many different manufacturers, but ultimately an Apple iPad 2 had to be used as it was the only tablet with two cameras supporting iOS at the

---

[1] `https://www.artoolworks.com/community/forum/` [last access 13th August 2012]

time. A newer version of the iPad has since been released and would probably have been chosen if development had started later. Even though the choice of device was effectively predetermined, its fully-featured hardware meant that it was by no means a bad one. The iPad 2 features a large, high-quality multitouch screen, front- and rear-facing cameras, and a number of internal sensors, and is well supported with development toolkits and frequent updates. In many ways it represents the standard tablet computer that consumers have come to expect, and is a household name, which could help user acceptance levels.

3.6 Initial design

After performing the literature review and looking for areas of the AR and cultural heritage domains that would benefit from further research, the initial design of the system sought to specify a AR system which would: i) feature a mobile device; ii) be usable in outdoor heritage site environments; and iii) make use of existing software toolkits as much possible to avoid "re-inventing the wheel". Due to these decisions and the specialised hardware needs required by the project (e.g. the camera and tablet requirements), it was not feasible to produce a completely abstract design.

As stated in the Section 2.3.2, facilitating AR in outdoor environments is difficult, mainly due to tracking difficulties. After careful thought, a number of novel tracking systems were devised which would satisfy these requirements.

3.6.1 Tracking method

The design of the tracking system is very important in developing an AR system. The tracking system should take into account the environment in which the system will operate, the hardware device, the infrastructure available, and the the target user group. Four potential tracking systems were devised for the AR system, all attempting to make novel use of fiducial markers in outdoor environments.

3.6.1.1 “Over-the-shoulder” fiducial markers

This technique requires a device that features both front- and refacing cameras, such as the Apple iPad 2 or Samsung Galaxy Tab 10.1. Fiducial markers are placed on objects around the site, and situated such that the front-facing camera can read them over the user’s shoulder (this may involve the user having to hold the device in an off-centre fashion). The marker is tracked by the front-facing camera and virtual content augments the view from the rear-facing camera. This is illustrated in Figure 3.1.

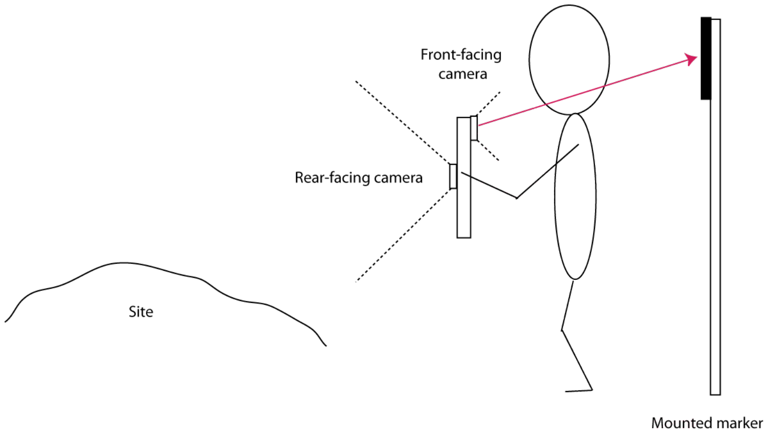


Figure 3.1: “Over-the-shoulder” markers.

3.6.1.2 “Marker hugging” technique

This technique also requires a dual-camera device. A fiducial marker is placed on a pole or mount, facing in the same direction as the user’s view (towards the site). The user then places this marker between themselves and the mobile device, so the front-facing camera can read the marker. The front-facing camera then tracks the marker, which is used to augment virtual objects correctly over the input from the rear-facing camera. This is illustrated in Figure 3.2.

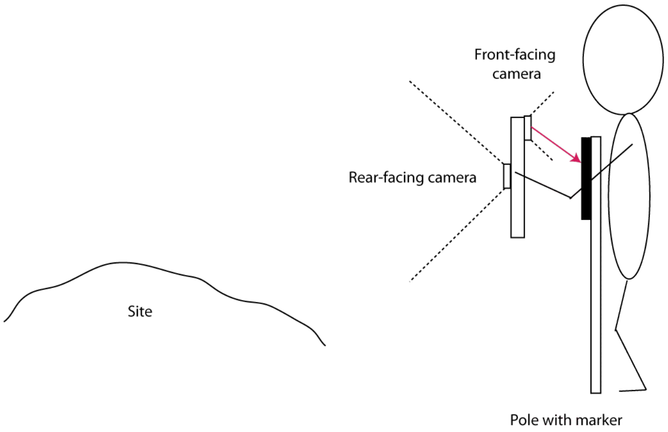


Figure 3.2: "Marker hugging" technique.

3.6.1.3 "Flagstick" device housing

A housing designed to hold the tablet device is placed at an unchanging location which is pre-calibrated (via GPS or direct measurement). By using a fiducial marker on the housing, the system pre-registers and calibrates the location of the device. Another option is for the user to carry a single housing with them which is placed in holes around site, much like a flagstick in golf. Inbuilt inertial sensors can allow the position to be tracked if the housing allows for rotation. This is illustrated in Figure 3.3.

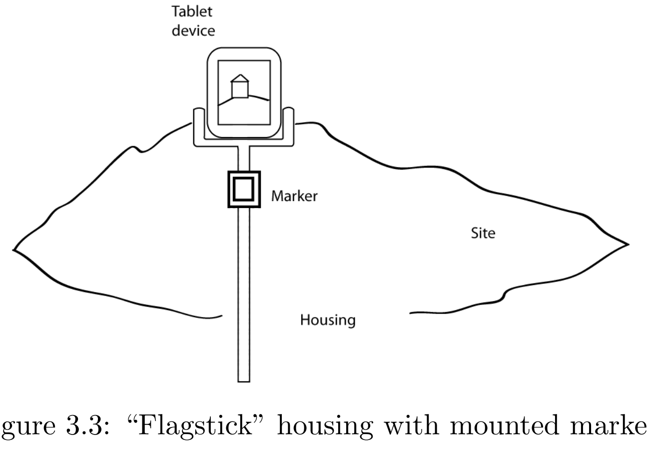


Figure 3.3: "Flagstick" housing with mounted marker.

3.6.1.4 "Doormat" markers

A calibration mat featuring a marker is placed by direct measurement. The user stands at the mat so the device can read the marker to allow the system to register the initial location. Further tracking is achieved in real-time using the device's internal sensors. This is illustrated in Figure 3.4.

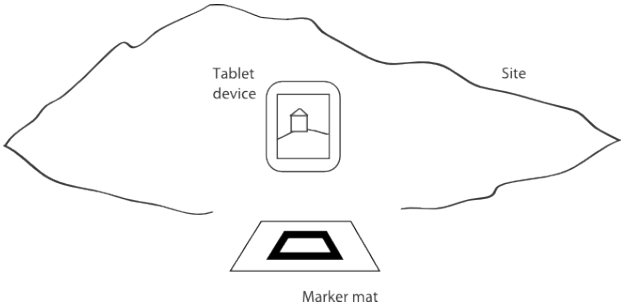


Figure 3.4: "Doormat" markers.

3.6.1.5 Chosen tracking system

The "over-the-shoulder" marker system was ultimately decided on, as it implements a novel dual-camera tracking method that has not been published in any of the literature reviewed for this project. In accordance with INFR4 (see section 3.4.2), it is the method would result in the least site modification, requiring only markers (possibly only temporary and made of paper) to be placed on site walls or other features, and would not necessitate the user to have to carry round any poles or mounts. By using markers placed behind the user, it allows for visualisation of large objects whilst still allowing the user to get close enough to the marker to achieve accurate tracking, and also allows for the visualisation of objects far in the distance, such as in a field.

3.6.2 Issues and refinement

After deciding upon a design for the system, development was started on a prototype iOS app to allow the technology to be quickly implemented and tested. During development of the prototype, it was found that it was not possible to have two camera feeds (one for the the front-facing camera and one from the rear-facing camera) active simultaneously. This was an limitation of the hardware and not the software, meaning that no practical workaround could be developed. This meant that the design had to be altered to accommodate this.

To retain a dual-camera tracking system but without the need for simultaneous camera feeds, a modified design was devised that incorporates obtaining an initial reference location with the from-facing camera and marker, then switching to the rear-facing camera. The user starts the system with the front-facing camera active, and uses this to detect a marker placed behind them over their shoulder. Once the system detects a marker, the user may touch the screen to switch to the front camera. When the screen is touched, the markers detected and their positions are the stored, and when the camera is switched the objects that were detected are drawn at their last known position, but this time over the input of the rear facing camera. After this, tracking is achieved using the inertial sensors in the device (see Figure 3.5).

3.6.2.1 A magic camera system

While attempting to integrate 3D tracking functionality into the prototype, it was found that accurate 3D position tracking using the internal sensors of the device was not feasible due to the inherent inaccuracy of the sensors themselves. A position can be computed but the error present compounds over time, leading to an ever more inaccurate reading. Due to this limitation, the design was changed to specify a *magic camera* system. Similar to how the *magic lens* and *magic mirror* systems in

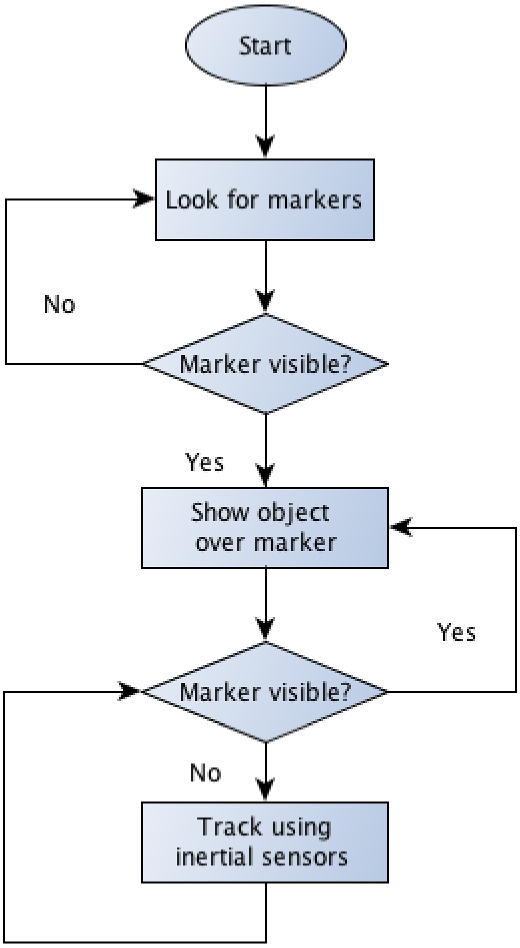


Figure 3.5: Flow chart showing the operation of the tracking loop.

Section 2.5.3 work, the system allows users to have an augmented view of the scene in front of them, but also allows them to take *augmented photos* of the scene with virtual objects in place. This moves the focus of the system onto the photo-taking requirement of the software.

While accurate position tracking is not feasible, accurately measuring only the rotation of the device is possible, so this functionality was preserved in the prototype as a compromise. This allows the user to position the virtual objects with greater accuracy before they take a photo, but does not allow the user to walk around the virtual object and view to from different angles and positions.

The use of a modular, evolutionary prototyping methodology meant that extra functionality could be added to the system easily without the need to completely restart development. If, as above, problems were encountered during the development of new functionality into the prototype, it could simply be removed or deactivated and a workaround or design change could be thought of and implemented in its place.

3.6.3 System architecture

The final design of the system specifies a "magic camera" type AR system designed to be used at outdoor heritage sites. Figure 3.6 shows a block diagram of the system architecture.

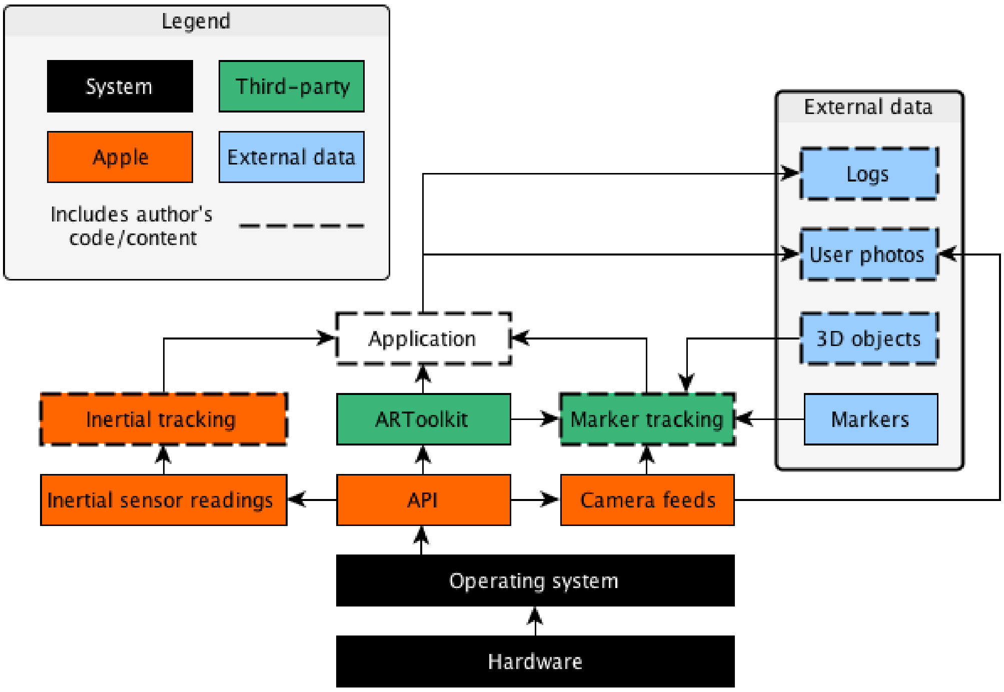


Figure 3.6: Block diagram showing system architecture. Colours represent origins of the code/content.

3.6.4 Final requirements

The final list of functional requirements for the system was as follows:

FR1: The system should be used in an outdoor setting

FR2: The system should allow for the visualisation of virtual objects over the camera feed

FR3: The system should include at least one virtual object representing a historical structure

FR4: The system should run on an Apple iPad 2

FR5: The system should feature a marker-based tracking system facilitated by the use of ARToolkit

FR6: The system should track the orientation of objects using the device's internal sensors when markers have ceased to be visible

FR7: The system should allow users to take photographs of the scene

FR8: The system should be able to store virtual objects for multiple marker sites

FR9: The system should allow the user to view the augmented photos they have taken

FR:10 The software for the system should be in the form of an iOS app

The final list of non-functional requirements for the system was as follows:

Performance:

NFR1: The system should not detect incorrect markers

NFR2: The system should feature only one virtual object for each marker site

NFR3: The system should log events to a file for later analysis

NFR4: Camera switching should happen within five seconds

Usability:

NFR5: The system should allow the user to switch between cameras to start the tracking process again without having to exit the application

NFR6: The system should allow the user to visit different marker sites separately without having to restart the application

NFR7: The system should be usable by users of varying heights

NFR8: The system should be enjoyable to use

Accessability:

NFR9: The file size of the entire application should be no larger than 100MB to allow users to download it via a mobile internet connection if necessary

Deployment:

NFR10: Any additional equipment or materials needed should be used in a way that requires the least amount of site modification, and should not permanently modify the site.

3.6.5 Using the system

Typical use of the system was envisaged as follows:

1. The user loads the software on to the device.
2. The user arrives at the outdoor heritage site.

3. The user walks to the desired marker site and starts the application. The front-facing camera is active.

4. The user positions the tablet so the camera can see the marker over their shoulder.

5. Once the user has lined up the marker and an object has appeared, they touch the screen.

6. The rear-facing camera is activated. The objects are rendered in their last known position before the camera was changed, but this time they are superimposed over the input from the rear-facing camera (what the user sees in front of them). Figure 3.7 illustrates this operation.

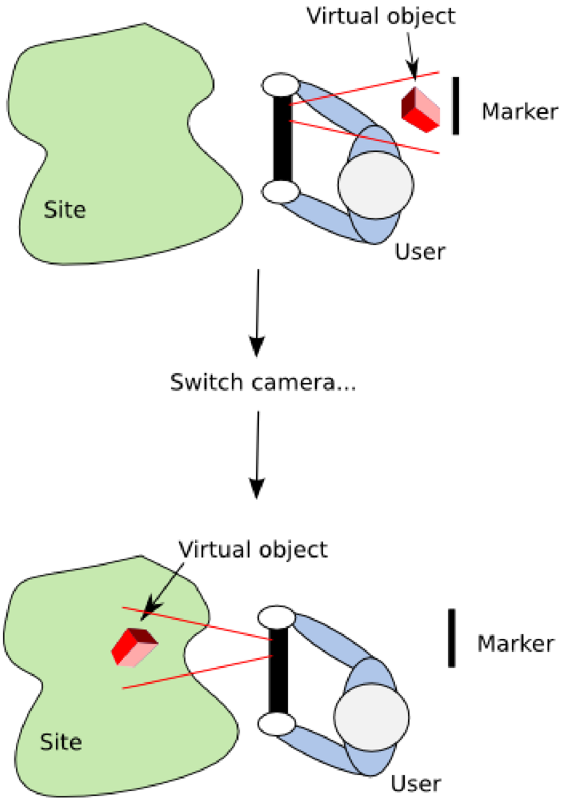


Figure 3.7: Illustration of the camera switching operation.

7. The user is free to move the device around to position the object correctly. The rotation of the object is tracked by the device, so that the object's rotation will act accordingly when the user moves the device.

8. Once the user has positioned the item, the they touch the screen once again to take a photo. The camera is then switched back and the user is free to start again. The user can view the photos they have taken in case they wish to take another.

Figure 3.8 shows a UML activity diagram of the typical system use.

3.6.6 User interface and interaction

The system involves the user having very little interaction with the device. The user is only required to tap the screen, firstly to switch from the front-facing camera to the rear-facing camera, then again to take a photo and and switch back to the front facing camera. As such, no traditional graphical user interface was deemed necessary. The only other interaction is switching between the application and Apple's *Photos* application, which is natively supported by the operating system by double-pressing the home button on the device then selecting the desired application.

3.7 User study

Effective and rigorous evaluations of AR systems are often missing from published literature, possibly due to a lack of suitable methods to achieve this, and because there is no agreement as to which methods are most suited [25]. From the outset of this project it was decided that a thorough user study would be carried out to test the usability of the system.

Nielsen [56, p. 26] defines usability as five different quality attributes:

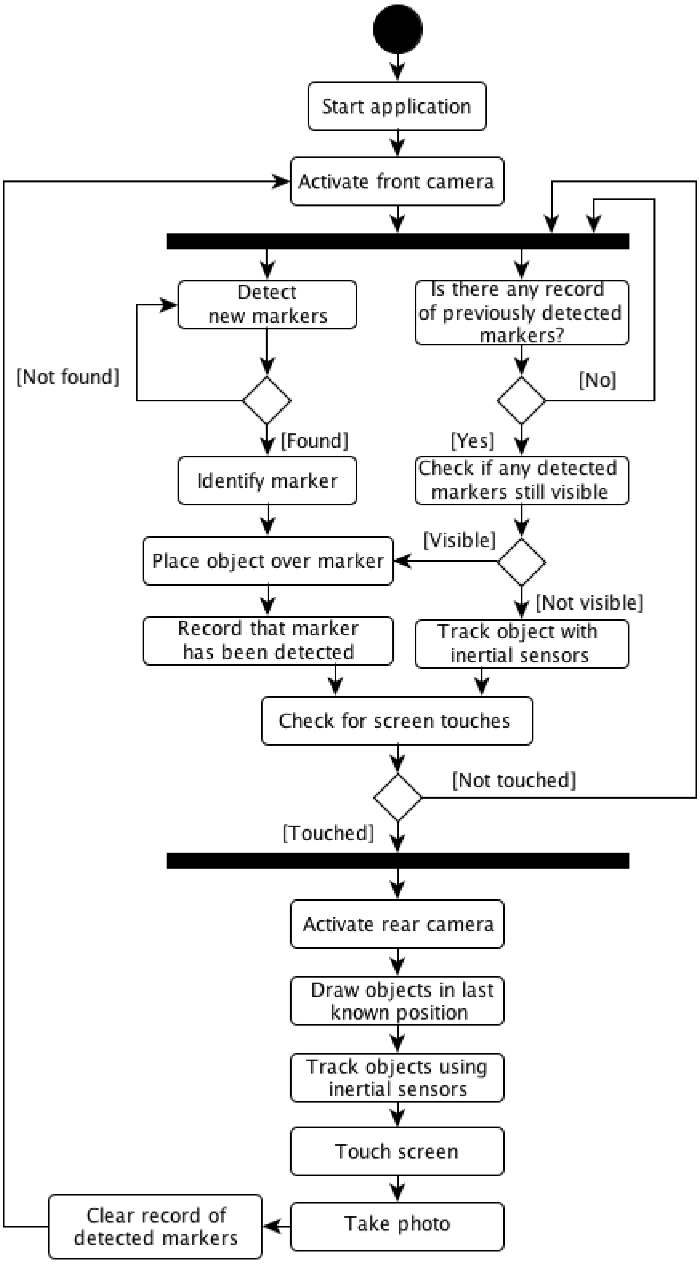


Figure 3.8: UML activity diagram showing system usage.

Learnability: How easy is it for users to accomplish basic tasks the first time they encounter the design?

Efficiency: Once users have learned the design, how quickly can they perform tasks?

Memorability: When users return to the design after a period of not using it, how easily can they reestablish proficiency?

Errors: How many errors do users make, how severe are these errors, and how easily can they recover from the errors?

Satisfaction: How pleasant is it to use the design?

The study was carefully designed to focus on these quality components, with the exception of memorability as it would have required the participants to use the system over two sessions, which would have been impractical.

3.7.1 Study design

The study was designed to assess usability by giving the participants a series of similar tasks to complete while using the system in an outdoor test environment. These tasks involved the participants being shown a number of reference photos, each featuring a single virtual object. The participants were then required to recreate these images as closely as possible using the AR system. The number of tasks was set at 4 so the study would not take a long time for each participant, but the number was hoped to be enough to see if participants got better at using the system after each task. Each task also resulted in an output (a photo that the participant takes) which would allow how well they had completed the task to be assessed using a quality metric.

A pilot study was carried out before any real participants were enlisted so that any major problems would be encountered before the study commenced. It was found that the original amount of tasks to complete (6) was too many, and that some of the locations markers had been placed were impractical. For this reason, the number of tasks was cut to 4. It was also found that the amount of equipment to be carried during the experiment would require extra apparatus, for instance the evaluator making notes and entering data while showing images to the participant was impractical; a music stand was used in the real study to allow images to to be shown to the participants without having to hold them.

3.7.2 Questionnaire design

Participants were required to complete both a pre-session and a post-session questionnaire. The post-session questionnaire (see Appendix A) was intended to collect qualitative usability data that could be compared with the data collected during the experiment. It featured fourteen comments, each of which had a five-level Likert scale with which which the participant had to rate their level of agreement. An odd number of choices was chosen so as to allow the participant to express indifference to a comment.

The comments were carefully chosen to asses different aspects of the usability of the system. Each question was chosen so as to relate to one or more of the quality attributes described in Section 3.7. Each question had a section that allowed the participant to prove any further comments about that question, and there were two open questions at the end asking how they would improve the system and if there was any further feedback. This allowed for the collection of extra qualitative data that could explain why the participant had answered in a particular manner.

3.7.3 Summary

This section describes the design process that was used to specify and test the AR system. After some iterations due to technological setbacks, the final design settled on was a “magic camera” system, which allows the user to superimpose digital objects over a scene and take photos of the result. A user study was also designed to evaluate the usability of the system, and to see if a dual-camera, marker-based tracking system was feasible.

Chapter 4

Implementation

4.1 Introduction

After the design of the system had been specified, work was begun on the implementation stage of the project. This chapter details the programs and tools that were used to implement the design, and the problems that arose.

4.2 Xcode Integrated Development Environment

As the system was being developed with Apple's iOS as the target platform (see FR4, Section 3.6.4), there was no choice but to use Apple's Xcode integrated development environment (IDE). Xcode has many features, including an editor with formatting, code completion and search tools, a compiler, numerous debug tools including the ability to set breakpoints and view the contents of variables.

Xcode provides an easy way to transfer the apps developed to the device, and provided the device is connected to a computer running Xcode debugging can be done on the device. It is also possible to easily transfer data created when the program is used from the device by using the *organizer* tool. The was also used to transfer the log files from the app developed in for project onto the development computer for analysis, which was necessary to fulfil NFR3 (see Section 3.6.4).

Xcode also provides simulators for different Apple devices, allowing developers to test their apps without owning the target device. However, the simulators provide no support for camera feeds so could not be used during this project. This meant that some features, such as the advanced code profiling and debug tools called *instruments*[1] could not be used as they are not supported when running an app on the device.

All iOS developers must be enrolled on the iOS Developer Program. In turn, each developer must have a developer certificate, each device used must be registered and tied to the developer, and each app must be registered and have an ID. This information is then combined into a provisioning profile, which must be valid and is necessary to sign the code to allow apps to be run on a device. Xcode attempts to streamline this, but it is still a complex and long-winded process which was the cause of some problems during the early stages of development.

4.3 Objective C programming language

iOS applications are written in the Objective-C programming language, which is an expanded version of the C programming language. This meant that the advanced features and new syntax had to be learned during the development stage of this project, but developers can also use C and C++ code if they so wish. However, this means that different syntax can be used to accomplish the same task, so it was important that a single style was decided upon for this project, and used throughout the project for the purposes and readability and maintenance. This also means that

---

[1] `https://developer.apple.com/library/mac/#documentation/developertools/conceptual/InstrumentsUserGuide/Introduction/Introduction.html` [last access 15th August 2012]

it is compatible with a great number of non-iOS specific third-party libraries, many of which are written in C.

Objective-C is an object-oriented programming language, which allows for a modular design to be easily implemented and modified as the design changes. Modules can easily be added, removed or modified without the need to make large changes to other parts of the program, and code can easily be reused throughout the program where similar functionality is required.

Automatic memory management is not a feature of the Objective-C language, meaning the developer must closely keep track of allocated and deallocated memory. While this was the cause of some problems during the implementation of this project, it did help to promote the importance of memory management and led to better, more efficient code in the end. Once the correct practices were learnt they could be used as a matter of course.

4.4 iOS Software Development Kit

The iOS Software Development Kit (SDK) is a fully featured collection of over 1,200 Application Programming interfaces (APIs)[2] that provide easy access to many functions and services provided by Apple, as well as exposing the hardware of iOS devices to developers for use in their applications.

Heavy use of the iOS SDK was made during the development of the app for this project. One example is the use of the *Core Motion* framework, which provides access to the device's location services (such as the GPS and magnetometer) and also to the devices internal motion sensors (such as the accelerometer and gyroscope). Frameworks can easily be added to and removed from the project in Xcode.

The SDK includes an `NSNotificationcenter` class, which allows objects to

---

[2] `https://developer.apple.com/technologies/ios5/` [last access 14th August 2012]

send notifications to other objects that have been registered as "observers". The notification centre provides an easy way of communicating between classes that are unrelated without the need for pointers or other inelegant methods. The ARToolkit examples use it natively to accomplish tasks such as sending to notifications to registered objects when markers are detected or disappear, and this was further extended to inform the program when new inertial tracking data was available so the necessary updates could be made to 3D model positions.

4.5 ARToolkit

The ARToolkit marker tracking library was used prominently in the development of the AR system for this project, in accordance with FR5 (see Section 3.6.4). The evolutionary prototype that was developed into the final system was gradually evolved from the ARToolkit Open Scene Graph[3] example project included in the library which simply showed a textured, animated aeroplane object over a marker (see Figure 4.1). Modification of this example included overhauling the tracking system to include a dual-camera switching method and adding inertial sensor tracking (see Section 3.6.2), adding device logging at selected points in the execution, and adding the ability to take screenshots of the scene.

ARToolkit features a selection of marker types that can be used, including pattern markers and barcode-type markers with different error-correction methods included (see Figure 4.2). After some experimentation, it was found that the barcode markers provided more reliable tracking than the pattern markers, which would help fulfil NFR1 (see Section 3.6.4). For the project, 80mm x 80mm markers were printed on a standard laser printer on plain white paper. The markers were then cut to size, ensuring a large white border was left around the black marker.

---

[3] `http://www.openscenegraph.org/projects/osg` [last access 15th August 1012]

Figure 4.1: An ARToolkit sample project, which was developed into the final system.

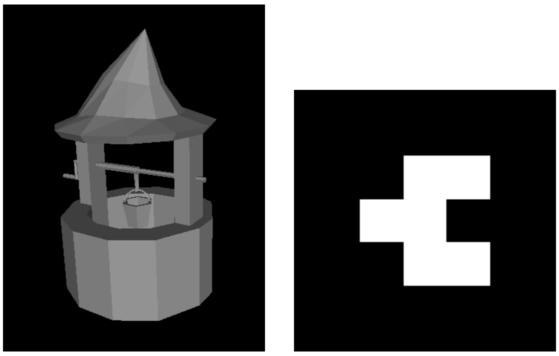
(a) An ARToolkit pattern marker.

(b) An ARToolkit barcode marker.

Figure 4.2: ARToolkit markers.

4.6 3D Studio Max and OSGExp

The 3D models used in the app were sourced from some of the many royalty-free sources on the internet[4,5], and also from a CD of models created for a heritage site in Yeavering, Northumberland. The models used and their corresponding markers are shown in Figure 4.3, Figure 4.4, Figure 4.5, and Figure 4.6. Autodesk 3D Studio Max[6] was used to edit and resize the models where necessary. However, while the objects used were in the common OBJ file format, the prototype being developed utilised the OSG file format, which is not natively supported by many 3D modelling applications. For this reason, OSGExp[7], an OSG conversion plugin for 3D Studio Max was used to export the files to the necessary format.



(a) Well model. (b) Marker.

Figure 4.3: 3D well model and its corresponding marker.

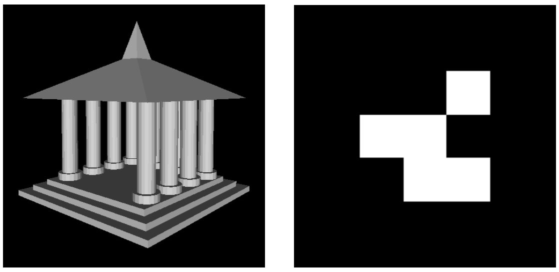
---

[4] `http://thefree3dmodels.com` [last access 12th September 2012]

[5] `http://www.3dmodelfree.com/` [last access 12th September 2012]

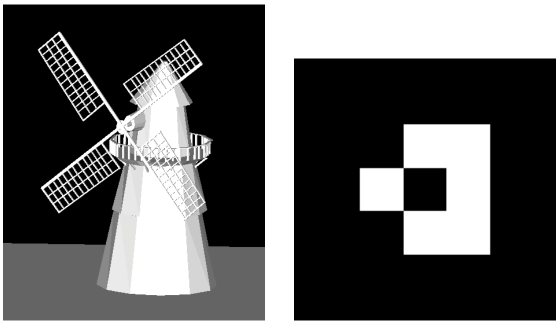
[6] `http://usa.autodesk.com/3ds-max/` [last access 15th August 2012]

[7] `http://sourceforge.net/apps/mediawiki/osgmaxexp/index.php?title=Main_Page` [last access 15th August 2012]



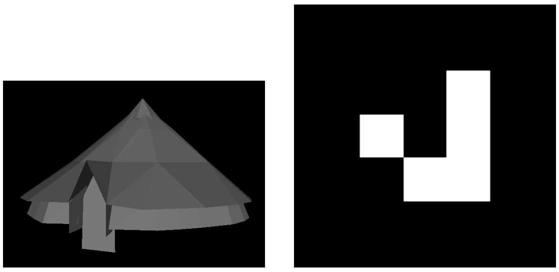
(a) Temple model. (b) Marker.

Figure 4.4: 3D temple model and its corresponding marker.



(a) Windmill model. (b) Marker.

Figure 4.5: 3D windmill model and its corresponding marker.



(a) Roundhouse model. (b) Marker.

Figure 4.6: 3D roundhouse model and its corresponding marker.

4.7 Inertial sensor tracking code

In accordance with FR6 (see Section 3.6.4), the system featured inertial sensor tracking. The orientation estimation algorithm presented by Madgwick et al. [48] was initially used, in the form of a C library that had been uploaded to his website[8]. The algorithm used sensor fusion to combine the readings of the device's accelerometer, gyroscope and magnetometer to generate a representation of the device rotation in 3D space in the form of a quaternion. This quaternion then had to be converted to an OpenGL-compatible rotation matrix. However, when integrated into the application it was found the the drift was too much to be acceptable; the orientation of 3D objects was tracked in relation to device movement, but the objects slowly rotated randomly which suggested that any errors in the sensor readings were being compounded.

As an alternative to Madgwick's algorithm, Apple's CMAttitude class from the Core Motion framework was used, which gives a representation of the devices orientation in relation to a reference frame (for example, where the Z axis is vertical and the X axis is on the horizontal plane). It was necessary to use the class's `multiplyByInverseOfAttitude:` method which gives the change in attitude in comparison with the attitude passed as a parameter, as opposed to the change in relation to a reference frame. Unlike Madgwick's algorithm, this can be represented directly in rotation matrix format which bypassed the need for a conversion step. The CMAttitude class also uses sensor fusion to combine the readings of the internal sensors of the device, and was found to be far more stable - little or no noticeable drift was observed, unlike with Madwick's algorithm.

4.8 Issues

Numerous issues prevented the implementation of the project running smoothly, which meant that the development time exceeded the time that had been planned for at the beginning of the project.

---

[8] `http://www.x-io.co.uk/node/8` [last access 15th August 2012]

4.8.1 ARToolkit issues

While ARToolkit did provide usable, marker-based tracking for the project, a number of problems were encountered which often used up a lot of time trying to resolve them. Firstly, the documentation for the iOS version of ARToolkit was found to be lacking. There were no clear tutorials, and the user manual simply consisted of a description of each function and what it accomplished. A number of example projects are included with the library which demonstrate various different functionalities, but there was no accompanying documentation aside from code comments, which were sometimes found to be insufficient in order to properly understand the code. Some information about the program design was however included in the release notes, so by using this and inspecting the code and objects created while the program was running the structure of the program could be better understood. A full user manual with tutorials that explain the program flow step by step would have been helpful to save time, and architecture diagrams of the program design would have increased overall understanding.

Secondly, working out how to change the co-ordinates of a 3D object manually (irrespective of markers) proved to be very difficult. This was necessary because a rotation matrix calculated from the inertial sensor readings would need to be applied to the object to reflect the change in device rotation, but even simply attempting to move the object on one axis was difficult. There was no documentation regarding this so it had to be asked on the support forum, but it had to be asked multiple times and took some weeks before a satisfactory resolution was found.

Thirdly, there was no lightweight way to switch between front and back cameras without completely destroying the scene. This meant that when the camera was switched, the view controller's `[stop]` method had to be called, which deallocated all the objects in use. These then had to be recreated when the other camera was activated, and all the 3D objects had to be loaded. Between these operations,

(a) Before the camera switch. The virtual object is tracked using the marker over the user's shoulder.

(b) After the camera switch. The virtual object is tracked using the inertial sensors, and shown over the scene in front of the user.

Figure 4.7: The system in use from the user's perspective, showing the camera switching operation.

all of the co-ordinates, translations and visibility status of each object had to be saved and restored. This seemed like a very inefficient way to perform the camera switching operation, and it was not clear exactly which attributes of each object should be saved and restored which meant that a lot of trial and error was necessary. However, this was successfully implemented, and can be seen in Figure 4.7.

Updating to the latest version of ARToolkit also presented problems. When a new version of the library was released, it was added to the project but it would not compile and run. After a lot of time was spent trying to fix this, it was decided to roll back to the previous version which had been running without problems, though this did mean the new features and improvements could not be used.

Upon trying to add photo-taking functionality to the prototype in accordance with FR7 (see Section 3.6.4), the code included in the example projects was found to produce blank white images. After spending time trying to fix this without success, a question was posted regarding this on the support forum and it was established the code itself was at fault. Replacement code was suggested by the ARToolkit developers which was used in the program without problems, but this problem ended up taking up yet more development time unnecessarily.

4.8.2 Updating iOS

Early in the implementation stage of the project, the iPad was updated to the later version of iOS available at the time. Unfortunately, this caused a number of problems. Firstly, it meant that the latest version of the Xcode developer environment had to be used, as older versions were not compatible with the newest version of iOS. However, it was not possible to obtain the newest version of Xcode for the current version of the OSX operating system on the development machine (OSX 10.6 Snow Leopard), which meant that a paid upgrade was necessary to update the operating system to version 10.7 to allow the newest version of Xcode to be installed. This caused unnecessary delays to the development, but also meant that in order to avoid similar problems further upgrades to newer versions of iOS were avoided, even though they may have fixed bugs in the operating system or added new functionality.

4.9 Software Engineering practices

Where possible, a high level of object orientation and modularity was employed in the code. This ensured that the addition and removal of features was easy, and allowed each part of the program to be tested on its own in isolation. This approach also promoted a high level of code reuse to reduce development time and make maintenance easier.

4.10 Summary

For the implementation of the design, Apple's Xcode was used as the development environment. It provides all the tools necessary to develop iOS applications, including access to the iOS SDK which allows the developer to easily implement complex functionality and access low-level hardware features, such as the inertial sensors and cameras. The Objective-C language was used, along with some standard C code. The 3D models for the program were sourced from websites providing royalty-free models, and edited where necessary using 3D Studio Max.

Unfortunately, this stage of the project encountered numerous problems which increased development time greatly. This was mainly due to the use of ARToolkit which, while ultimately provided the necessary functionality, is still not mature enough to provide a smooth development experience and would benefit from more extensive documentation. Aside from these problems, the project was successfully implemented. Figure 4.8 shows the system in use during the user study.

(a) (b)

Figure 4.8: AR system in use during the user study.

Chapter 5

User study

5.1 Introduction

This chapter describes the experimental procedure that was used to carry out the user study for the project, as well as the results that were obtained. The outcome of the study was mainly intended to assess two things: the usability of the system; and to see if a dual-camera tracking system was feasible for mobile devices.

5.2 Evaluation approach

The study was designed to require participants to complete a series of tasks using the implemented AR system (see Section 3.7.1). To do this, the four markers shown in Section 4.6 were placed around the University Science Site at different locations in accordance with the outdoor requirement of FR1. As stipulated by NFR2 (see Section 3.6.4), each marker had a different virtual object associated with it which would allow the participants to create a different scene with a different object at each site. Figure 5.1 shows an example of a marker located around the University.

Figure 5.1: A marker placed on a building at the Science Site.

5.2.1 Reference images

Before the study began, an augmented photo was taken at each marker site, with the correct object for that site as a feature of the photo (see Appendix B). These photos were used as reference images which the participants would then be required to recreate as closely as possible. One of the reference images used in the study is shown in Figure 5.2. These photos were taken to each site and and placed on a music stand to allow the participant to see them easily. Only the photo relevant to the current marker was visible at each respective marker site, and the participant was free to move the music stand to a more convenient location for them if necessary.

5.2.2 Equipment used

The following equipment was used during the study:

- An Apple iPad 2, with the AR app loaded on the device
- A Kodak PlaySport mini-camcorder to record video footage of the participants
- A tripod to mount the camcorder
- Four fiducial markers, one at each of the marker sites around the University
- Four reference images corresponding to the markers, each in a plastic wallet to protect against the rain
- A metal music stand to allow the participants to view the reference images without having to hold them
- Data sheets to allow the evaluator to record results and observations

Figure 5.2: An augmented reference photo featuring a virtual temple.

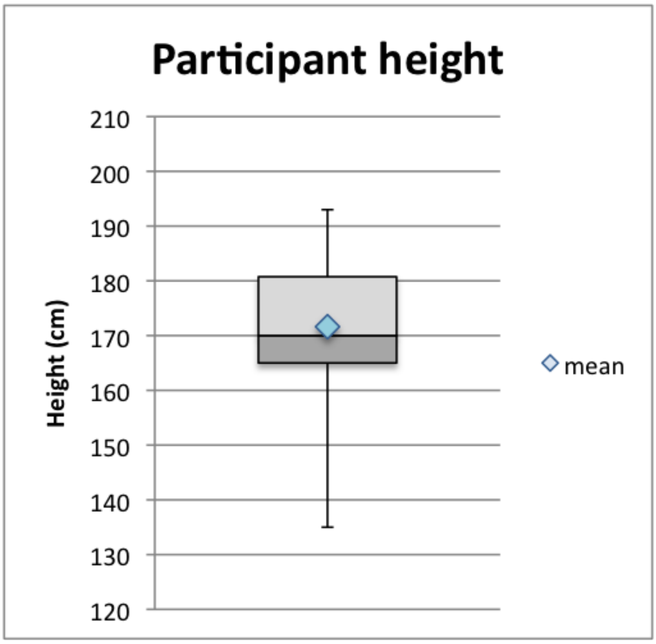


Figure 5.3: Box plot showing the heights of the study participants.

5.2.3 Participants

The participants for the study were recruited via word of mouth, social networking websites, and posters placed around the University. The study was completed by 16 participants; 8 males and 8 females. All of the participants were aged 18 - 25, which was due to the fact that most of the participants were student friends of the evaluator, or friends of friends who were also students. Only one of the participants was left-handed, the other fifteen were right handed. The heights of the participants are shown in Figure 5.3

5.2.4 Study procedure

When the study began for each participant, after reading the participant information sheet and completing a consent form they were required to fill out a short pre-session questionnaire to collect demographic data (see Appendix C). Once completed, they were taken to the first marker site and given a brief training session, where the purpose of the system was stated and they were shown how to use the system by example. This involved the evaluator using the system at the first marker site, while talking the participant through the intended use of the hardware and software and answering any initial questions. The order in which the marker sites were visited was randomised so the learnability of the system could be evaluated, independent of whether some marker tasks were "easier" to complete than others.

After the training session, the participants were given the device and asked to use the system to recreate the reference images to the best of their ability. A co-operative evaluation protocol was in use. This is an extension of the "think aloud" protocol, where the participant is encouraged to about the processes they are undertaking (i.e. verbalise their thought process) and state any problems or difficulties they are having, but in addition to this they can also ask the evaluator for help, and the evaluator can give them prompts or ask them questions [52]. This changes the think aloud protocol from a one-way process to a collaborative two-way process, and can help the evaluator to gain more useful information. It can also make the evaluation process seem more informal to the participant, which may help them relax and use the system in a more natural way.

5.2.5 Data collection

During the study, each participant was filmed for in case later analysis was necessary. A data entry sheet (see Appendix D) was created for the study, which allowed the evaluator to easily record important information. This included when user errors had occurred, such when the app had crashed, the number of verbal hints given to the participant, and the number of photos taken by the participant. The data entry sheet used a tally system. In addition to this, the start and finishing times of the tasks were also recorded, as were the weather conditions, and any important comments that the participants made during the study.

As well as manually-recorded data, the system was made to log important events to the device. This satisfied NFR3 (see Section 3.6.4). When the system was started, a new log file was created using the time and date as a unique filename which could be cross-referenced with the times manually recorded by the evaluator for easy identification. The system recorded when the system was started, when markers were detected, when the camera was switched, when photos were taken, etc. Also, the timing intervals were also recorded as a list of comma-seperated values so the times could easily be input to a spreadsheet program for later analysis.

When the study was complete, each participant was required to fill in a post-session questionnaire to collect qualitative data (see Appendix A).

5.3 Results

The results of the user study are present in the following sections.

5.3.1 Time taken to complete tasks in order

Figure 5.4 shows the completion time of marker tasks in the order they were completed (irrespective of the actual marker) in box plot form.

From this it can be seen that, in general, after completing the first task the participants completed the second task more slowly. They then got faster and faster with the third and fourth tasks. While the times for the second and third task are reasonably widely spread, the upper and lower quartiles of the fourth task are tight which shows that the final task was generally completed noticeably faster than the previous ones. The reason for this could be as follows: the first marker task was completed directly after a short training session (see Section 5.2.4), part of which consisted of the experimenter demonstrating the use of the software. Due to this, the correct way to use system was fresh in the participant's mind, and

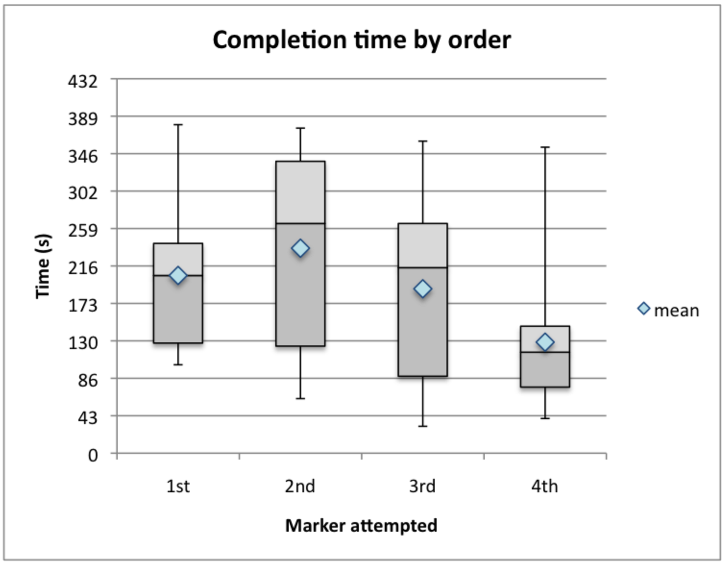


Figure 5.4: Box plot showing the completion time of the tasks in the order they were attempted, irrespective of which marker it was.

they were more willing to ask questions and received more help as it seemed like a continuation of the training session. After this, participants attempted a second marker with less help, which meant it took slightly longer to complete. After they successfully completed this, the next two markers were completed more quickly as they got used to using the system. A Friedman Test test was performed over the first, second, third and fourth markers attempted and indicated there was no statistically significant difference in time taken to complete the tasks over time (χ^2 (3, n = 16) = 7.33, p = 0.0621). However, a Wilcoxon Signed Rank Test was performed on the first and fourth (last) marker competition times and showed there was a statistically significant difference between the two times, z = 2.31, p = 0.021. The median time decreased from 205 seconds for the first marker to 117 seconds for the fourth.

Figure 5.5 shows a box plot of the number of verbal hints given to the participants by the evaluator for each task in the order they were completed. A small downward trend in the number of hints given can be seen.

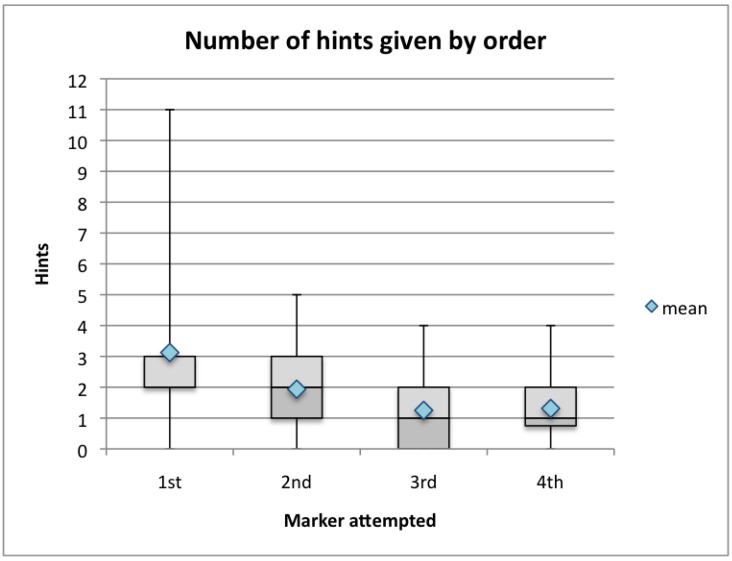


Figure 5.5: Box plot showing the number of verbal hints given by the evaluator for each task in the order they were attempted, irrespective of which marker it was.

5.3.2 Time taken to complete tasks individually

After looking at the time taken to complete each marker task individually (irrespective of the order in which they were attempted), it was found that the task at marker B took noticeably longer to complete than the other markers. This is illustrated in Figure 5.6.

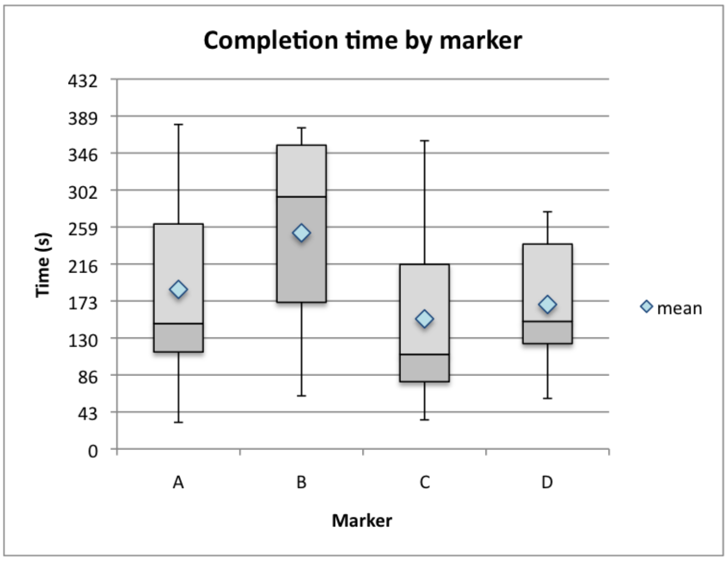


Figure 5.6: Box plot showing the completion time of the tasks by marker, irrespective of the order in which they were attempted.

This can be explained by examining the number of false positive marker detections for each marker task, which are a common problem with ARToolkit [27]. During the study, if an incorrect marker was detected or a marker was detected was there was none, the application had to be closed and restarted, meaning the participant had to restart the task and more time was taken to complete it. An example augmented photo showing a false positive marker detection can be seen in Figure 5.7. The number of false positives for each marker is shown in Figure 5.8, showing that marker B recorded the most. A Friedman Test test was performed and indicated there was a statistically significant difference in time taken to complete the tasks across the individual markers (χ^2 (3, n = 16) = 10.03, p = 0.0183). Inspection of the median values shows that marker B has by far the largest with a median of 294; the next largest is marker A with a median of 186. The presence of numerous false positive marker detections means that NFR1 (see Section 3.6.4) was not fulfilled.

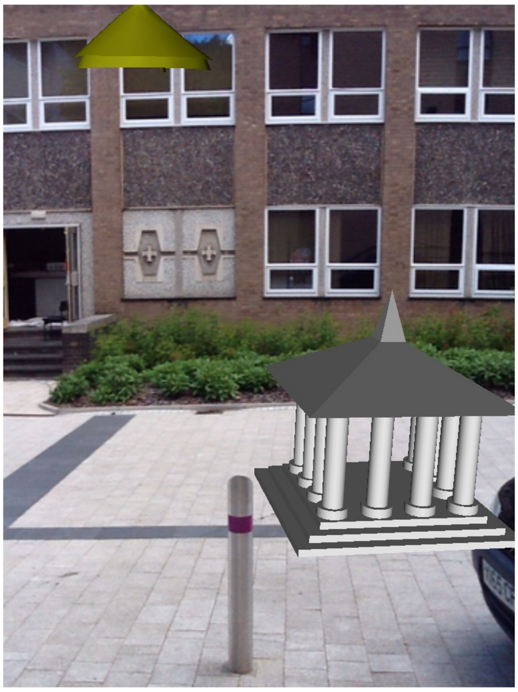


Figure 5.7: An augmented photo showing a false positive marker detection, resulting in an unwanted virtual object.

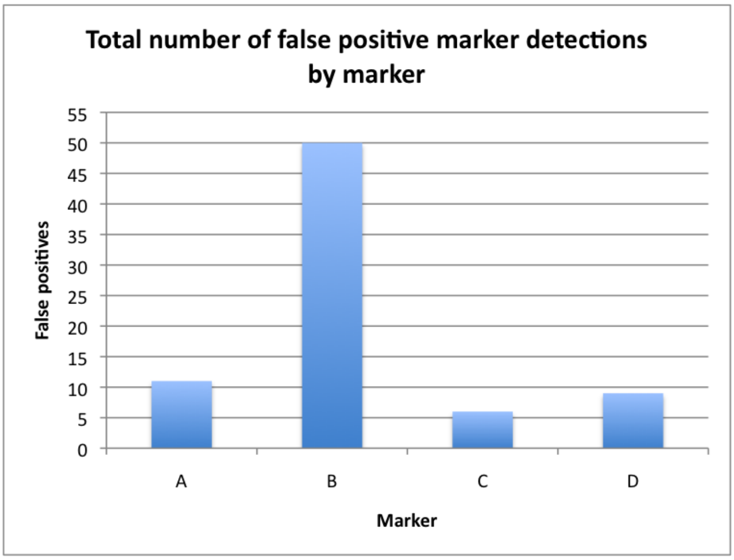


Figure 5.8: Bar chart showing the number of times markers were incorrectly detected for each marker.

5.3.3 Total completion time

The total completion time of the entire experiment by gender was also examined. The results are summarised in Figure 5.9. This shows that males took on average 39% longer to complete all of the tasks. It can also be seen that the times for the female group are generally far less spread; the inter-quartile range for females was 110 seconds, versus 375 for the males. A Mann-Whitney U Test was carried out and revealed that there is a significant difference between the time taken to complete all of the tasks by males (median = 862, n = 8) and females (median = 627, n = 8), U = 53, z = -2.15, p = 0.0316, r = -0.5375.

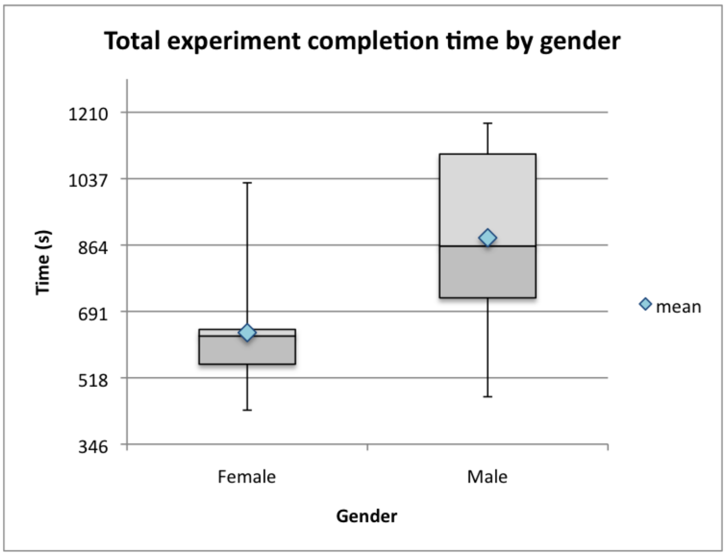


Figure 5.9: Box plot showing total completion times for males and females.

5.3.4 Time taken to take each photo

The amount of time participants took to take photos was also investigated, in order to see if a longer overall time to complete each task meant that more photos were taken. This could show if some people took a long time to get their photos correct, or if they took a large number of photos in a short space of time after making

Table 5.1: Results of analysis of the time taken to take each photo, using Spearman's rank correlation.

Marker	r	n	p
A	0.640	16	0.0076
B	0.205	16	0.4473
C	0.383	16	0.1426
D	0.125	16	0.6456

many small corrections. The average time taken to take a photo at each marker is shown in Figure 5.10. Scatter plots showing the number of photos and time taken for each participant are shown in Figure 5.11, and these relationships were investigated using Spearman's rank correlation coefficient, the results of which are shown in Table 5.1. From these results, it can be seen that only for Marker A was there a strong correlation between time and the number of photos taken with statistical significance (r = 0.064, p = 0.0076), showing that more time spent on this marker task resulted in more photos.

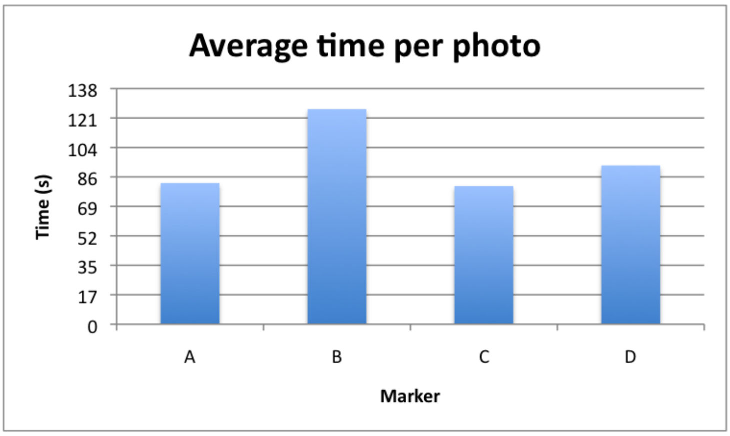


Figure 5.10: Bar chart showing the average time taken for each photo at each marker.

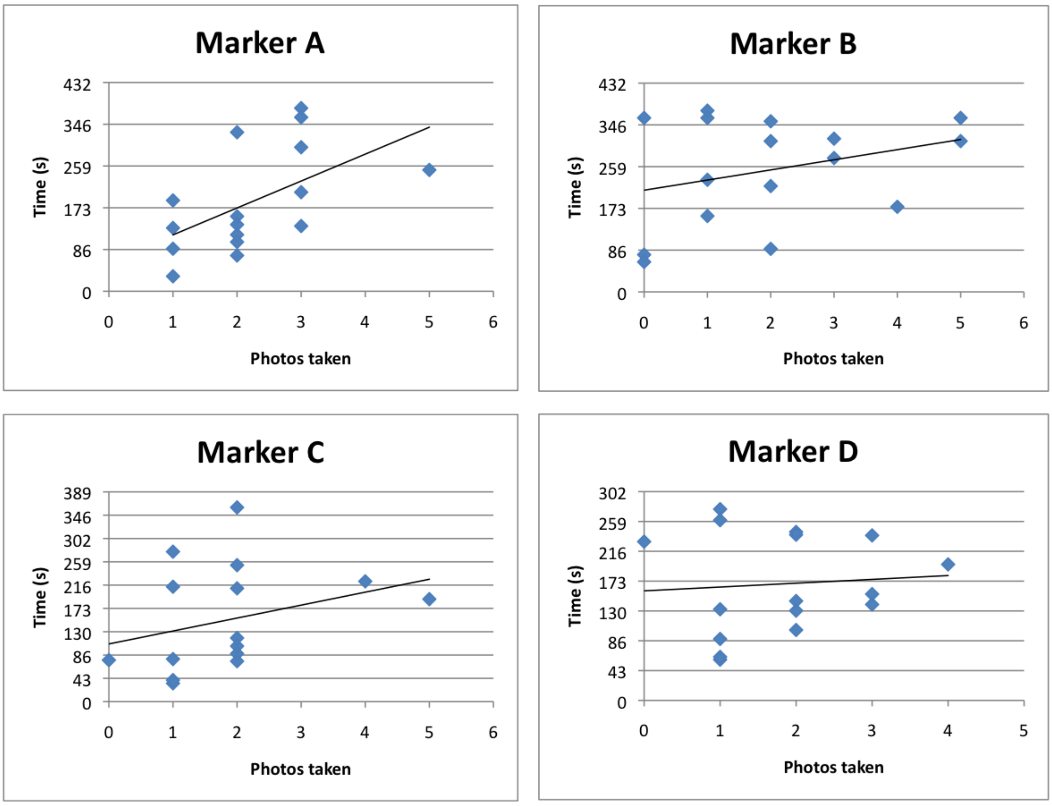


Figure 5.11: Scatter charts showing the number of photos taken against total time by each participant for each marker.

5.3.5 Image quality metric

In order to further evaluate the participants' task performance, a quality metric was used to attach a perceived quality value to each augmented photo taken.

Initially, two experts (the author and primary supervisor) rated four aspects of each image: position in the image frame; size of the virtual object; orientation of the virtual object; and the position on the virtual object in the 3D scene. A 1 - 3 rating was used, with 1 being completely wrong, 2 being a fairly good attempt and 3 being very good or perfect. A mark of 0 could also be given if the expert was unable to successfully rate one of the four aspects, for example if there was no virtual object present none of the aspects could be rated, or if the camera was not looking in the correct direction then the position of the object in the 3D scene could not be rated. This gave a total mark out of 12 for each image. After the ratings were recorded, a Kappa analysis was carried out to analyse the agreement between the experts. As each higher category represented a higher quality score a weighted Kappa analysis was used, and linear weightings were used to reflect the fact that the difference between each consecutive score was the the same (e.g. the difference between 1 - 2 is the same as 2 - 3). Across the individually-rated aspects of the images, Kappa values of 0.4108, 0.5119, 0.6477, and 0.7283 respectively were observed. It was felt that this was not a high enough level of agreement between the two judges, so another method was sought.

A previously used method to subjectively rate images was required, and ultimately the method used by Shelley et al. [77] was adapted for use in this project. In using this method, three experts (the author, primary supervisor, and an image processing PhD student) rated each aspect of each image using the same scale as before, and the results were averaged. To illustrate this, Figure 5.12 shows a reference image, the highest rated image (11.667/12; 97.225%), and the lowest rated image (4.0/12; 33.333%).

The overall average results for each image is shown in Table 5.2. The total scores show that all of the images were, overall, given a very similar rating; the results of a Friedman Test carried out over the total average image rating for all of the participants' four images show that there is not a statistically significant difference in the quality ratings between the images produced for the four tasks (the images at markers A, B, C and D), χ^2 (3, n = 16) = 0.39, p = 0.9423. This shows that even though marker B recorded the most errors, highest completion time, and greatest time taken to produce each photo, the quality of the images produced was not compromised.

(a) Reference image 3.

(b) The augmented photo with the highest average rating.

(c) The augmented photo with the lowest average rating.

Figure 5.12: Example images with the lowest and highest quality ratings.

Table 5.2: Overall average results of subjective image ratings by three experts.

Image	**Pos. in frame**	**Size**	**Orientation**	**Pos. in scene**	**Total /12**	**Total %**
1	2.250	2.292	2.146	2.333	9.021	75.174
2	2.271	2.208	2.104	2.438	9.021	75.174
3	2.167	2.313	2.104	2.292	8.875	73.958
4	2.125	2.229	2.000	2.167	8.521	71.007
Total /12	8.813	9.042	8.354	9.229	-	-
Total %	73.438	75.347	69.618	76.910	-	-

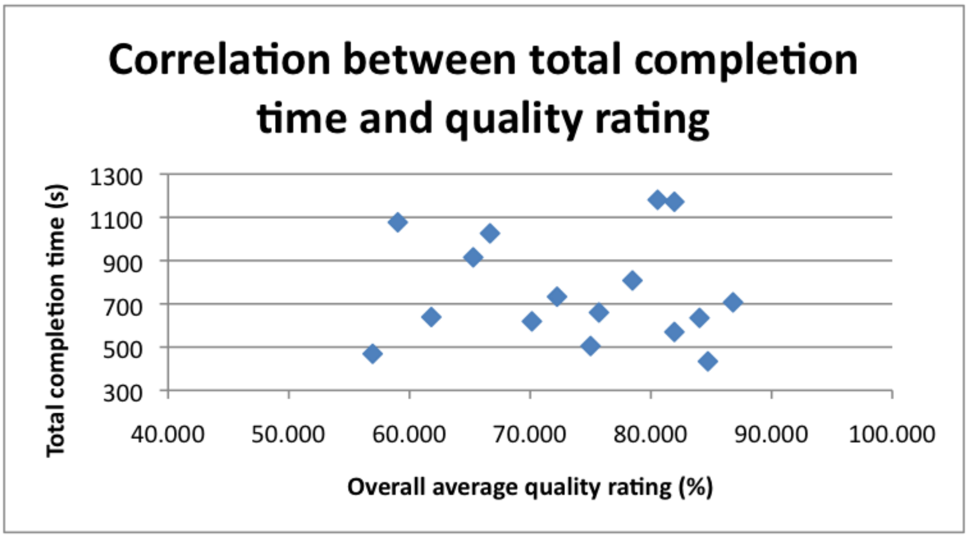


Figure 5.13: Scatter chart showing each participant's total completion time plotted against average quality rating.

Table 5.2 also shows the average rating for each image aspect. This shows that the orientation of the virtual objects was the lowest rated aspect, and the position of the object in the scene was the highest. This agrees with the way many users had difficulty in orientating the 3D objects (see Section 5.3.7).

Each participant's overall average quality rating was also examined. When these ratings are plotted against each participant's total competition time in a scatter chart (Figure 5.13), it can be seen that there is little correlation between the two. This was investigated using Spearman's rank correlation, which supported this (r = -0.071, n = 16, p = 0.795).

5.3.6 Qualitative questionnaire data

Nielsen [56, p. 26] defines usability as consisting of five quality components (see Section 3.7). Each of the 14 Likert scale-measured questions in the questionnaire (see Appendix A) were categorised into these categories. Two questions covered learnability, two questions covered efficiency, one question covered errors, and eight questions covered satisfaction.

The results of the questions covering learnability are shown in Table 5.3, and in the form of a box plot in Figure 5.14. This shows high scores for the learnability aspects of the system, which is confirmed by the downward trend observed in Section 5.3.1. However, it does highlight the necessity of a training session, which is confirmed by some of the comments made in the questionnaire for this question, which say that without any training the system would be very difficult to use.

Table 5.3: Results of the post-session questionnaire questions covering learnability (1 = disagree, 5 = agree).

Question	**Mean**	**Standard dev.**
5. I found each new task easier to complete than the last one.	3.813	1.167
11. The initial training provided before the session was helpful.	4.750	0.577

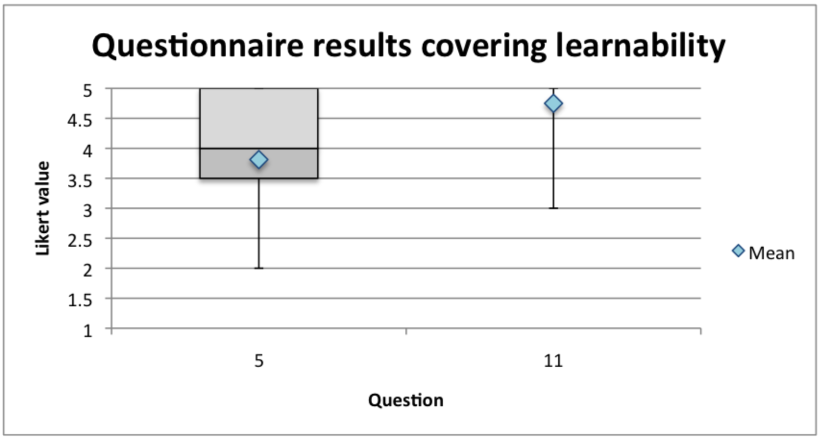


Figure 5.14: Box plot showing the results of questionnaire questions covering learnability (1 = disagree, 5 = agree).

The results of the questions covering efficiency are show in Table 5.4 and Figure 5.15, and shows that there was a varied opinion regarding this aspect of the system. Most participants did not have a strong opinion either way.

Table 5.4: Results of the post-session questionnaire questions covering efficiency (1 = disagree, 5 = agree).

Question	**Mean**	**Standard dev.**
1. I found it easy to recreate the reference images using the tablet.	3.188	1.167
3. I could quickly recreate the reference images using the tablet.	3.000	1.265
12. I had to be reminded how to use the system during the study.	3.063	1.436

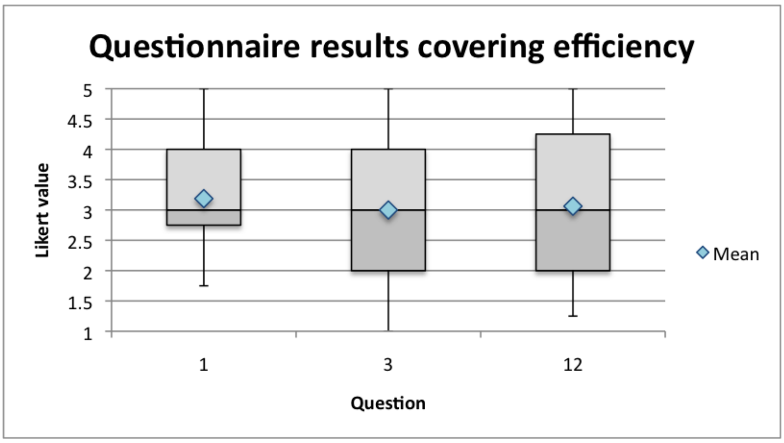


Figure 5.15: Box plot showing the results of questionnaire questions covering efficiency (1 = disagree, 5 = agree).

The results of the questions covering errors are show in Table 5.5 and Figure 5.16. The results were quite spread, but it can be seen that in general people did make mistakes when using the system.

Table 5.5: Results of the post-session questionnaire questions covering errors (1 = disagree, 5 = agree).

Question	**Mean**	**Standard dev.**
14. I made a lot of mistakes when using the system.	3.188	1.276

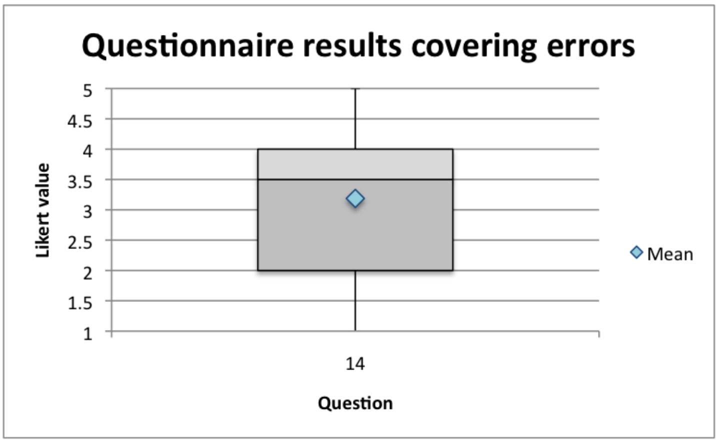


Figure 5.16: Box plot showing the results of questionnaire questions covering errors (1 = disagree, 5 = agree).

The results of the questions covering satisfaction are show in Table 5.6 and Figure 5.17. Although some of the responses are quite spread, generally positive responses can be seen for this usability aspect, and even the low responses are actually positive as they are about negative aspects of the system ("The session was physically tiring" and "Using the system frustrated me"). Of note are the particularly high responses for questions 7, 9, and 13 ("I would use the system again", "I found that the visual quality of the system was high", and "I found the system easy to use"). These results at least partially satisfies requirement NFR8 (see Section 3.6.4).

Table 5.6: Results of the post-session questionnaire question covering satisfaction (1 = disagree, 5 = agree).

Question	Mean	Standard dev.
2. The session was physically tiring.	3.188	1.276
4. The system was enjoyable to use.	3.938	1.063
6. Using the system frustrated me.	2.625	1.147
7. I would use the system again.	4.125	0.957
8. After my site visit, I would share the augmented photos taken.	3.813	1.328
9. I found that the visual quality of the system was high.	3.938	0.929
10. The markers were positioned at the correct height for me.	3.188	1.377
13. I found the system easy to use.	3.750	1.000

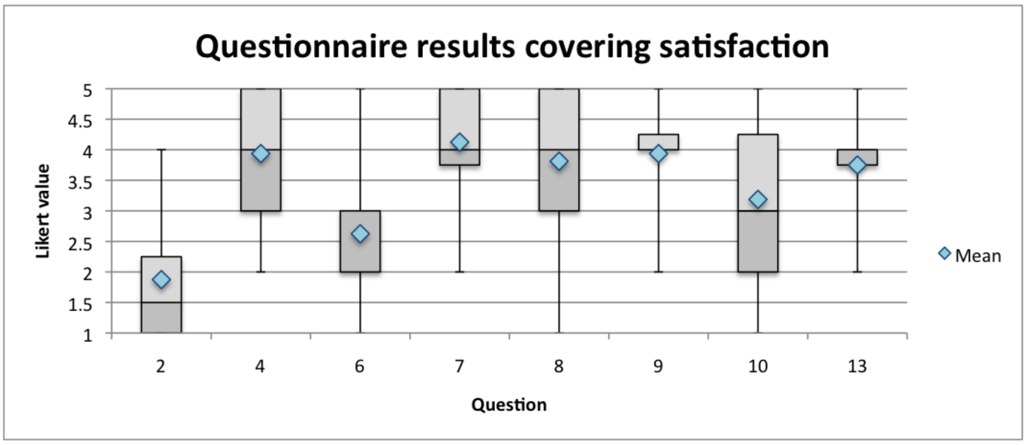


Figure 5.17: Box plot showing the results of questionnaire questions covering satisfaction (1 = disagree, 5 = agree).

5.3.7 Usability problems

During the study, any usability problems encountered were recorded on a data sheet. These included problems that the participants vocalised and problems that were directly observed. All of these problems, along with problems that were commented on in the post-session questionnaire, were listed along the number of participants who encountered each problem and the results were made into a word cloud (see Figure 5.18). This effectively highlights the most common usability problems.

To further analyse the usability problems, the problems were grouped and tabulated as in the co-operative evaluation analysis protocol applied by Smith and

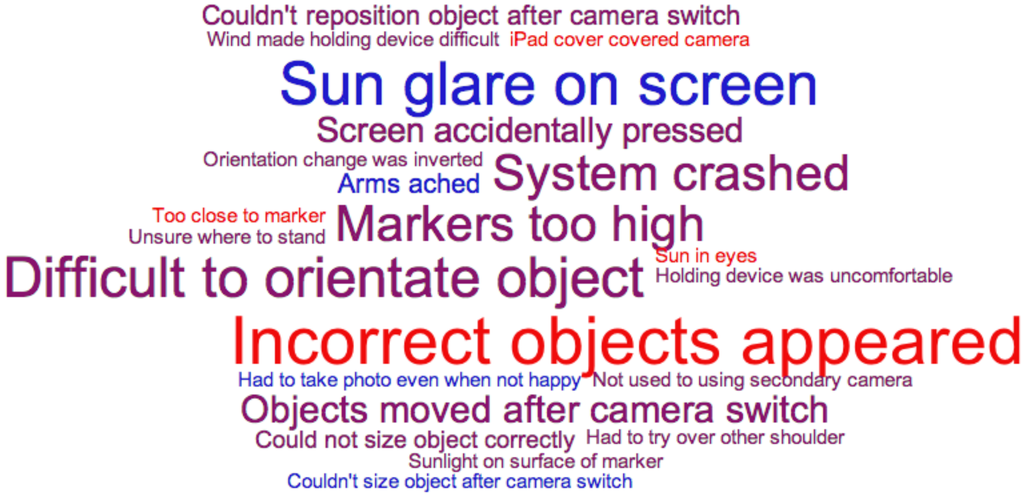


Figure 5.18: Word cloud showing usability problems encountered. Size is relative to the number of participants who encountered the problem.

Todd [78]. A total of 24 usability problems were identified, and these were sorted into 6 groups as follows:

A. Object manipulation issues: It was difficult to orientate the object, could not size object correctly, when changing the orientation by rotating the device the rotation was inverted, could not reposition the object after camera switch, could not size the object after camera switch;

B. Marker interaction issues: The participant found the markers too high, sunlight reflecting against the surface of the marker stopped marker detection, the markers were too low, participant had to try detecting the marker over the other shoulder, participant was unsure where to stand, the participant was too close to the marker, the participant was too far away from the marker;

C. Device issues: Screen accidentally pressed, the iPad cover was covering the camera, it felt strange using the front camera after being used to using rear cameras;

D. Environmental issues: Glare from the sun obscured view of screen, the wind made it difficult to hold the device steady, the sun was in participant's eyes;

E. Ergonomic issues: Participant's arms ached, holding the device was uncomfortable;

F. System interaction issues: Incorrect objects appeared, the system crashed, objects increased in size and changed position after camera switch.

The problems observed were sorted into these groups and ranked either high or low quality by the author, high being a problem that could not be resolved with further training or usage of the system and may be cause for a system redesign, low meaning that the problem is likely to be overcome with more usage or training. The results are shown in Table 5.7. From this it can be seen that by far the most problems fall into the *object manipulation issues* (21 issues) and *system interaction issues* (24 issues) categories, and further analysis of the data collected showed than difficulties orientating objects were very common, as were system crashes and incorrect objects appearing. Participants encountered between 3 and 8 of the issues overall; participant 2 recording the most, and also one of the longest total completion times at 256 seconds.

5.4 Summary

This chapter presented the method that was used to carry out the usability study, whereby a number of participants used the system and performed representative tasks in order to uncover some of the usability issues with the system. The study took place at the University Science Site, and required participants to recreate a number of reference images as closely as possible using the AR system and iPad. A co-operative evaluation protocol was in use to help the participants vocalise their

Table 5.7: Summary of usability issues by user (L = low quality, H = high quality)

Problem group Participant	**A**	**B**	**C**	**D**	**E**	**F**	Total
1	1L	1H	1L	-	-	3H	6
2	1L	2L	-	1L	1L	3H	8
3	2L	-	-	1L	-	3H	6
4	1L	1L	-	1L	-	1H	4
5	2l, 1H	1H	-	1L	-	1H	6
6	2L	-	1L	1L	-	1H	5
7	3L	1H	1L	1L	-	1H	7
8	1L	1H	1L	-	-	1H	4
9	-	1H	-	1L	-	1H	3
10	-	1H	-	-	1L	1H	3
11	1L	1H	-	1L	-	1H	4
12	1L	1H	-	2L	-	2H	6
13	-	1H	-	1L	-	2H	4
14	1H	-	-	1L	-	2H	4
15	1L	-	1L	1L	-	1H	4
16	1l, 2H	1H	1L	1L	-	-	6
Total	21	13	6	14	2	24	**80**

experience and therefore uncover ore usability issues. Data was collected in various forms, both manually and automatically, throughout the study for later analysis.

This chapter also presented the findings of the user study. It was found that after taking longer to complete their second task than the first, participants then completed each new task more quickly. This suggests that the system has a high degree of learnability.

When examining the time taken to complete the tasks at each marker independently the highest average time was recorded at marker B, possibly due to a large number of false positive marker detections which meant the system had to be restarted. Marker B also recorded the highest average time per photo taken. However, the images produced by each participant were subjectively rated by three AR experts, which were then averaged to give an overall rating, and quality of the images produced did not suffer as a result of the high number of false positive readings.

Female participants were found to have noticeably quicker total completion rates than males; the reason for this is not known and beyond the scope of the project, but owing to the small sample group it could be due to chance. However, previous research has shown that males often tend to use technology for it's own sake and spend more time using technology recreationally than females, for instance playing computer games [46], while females have been shown to adopt technology that benefits their everyday practice [50]. Due to this, it is possible that the males saw the AR app as a game and were therefore happy to spend time playing with it, while females may have just seen it as a means to an end to produce the augmented photos.

The qualitative questionnaire data was analysed to determine the degree of usability experienced by participants across 4 usability aspects - learnability, efficiency, errors, and satisfaction. Results were generally positive, but sometimes quite spread.

All of the usability issues recorded by participants were grouped into categories: object manipulation issues; marker interaction issues; device issues; environmental issues; ergonomic issues; and system interaction issues. Each issue was ranked either high or low - low ranked issues are ones that can be overcome with more training, high ranked issues are ones that would require a system redesign to be overcome. Most usability issues observed could be fixed in a future version of the software.

Chapter 6

Discussion

6.1 Introduction

The results of the user study show that an AR system using hybrid dual-camera and marker-based tracking was usable in a cultural heritage context, and allowed outdoor sites to be enriched with digital reconstructions of objects. However, there are still some problems with the system and the technology in general. This chapter explores a number of important issues highlighted by the research here, including AR evaluation studies, environmental conditions, considerations for the heritage sector, and technological issues.

6.2 User study

As AR is still a maturing research topic, there are limited amounts of established methods for evaluating systems. Dünser and Billinghurst [25] describe how AR researchers must find appropriate ways to measure effectiveness and efficiency in their software, and cannot rely on guidelines used for traditional user interfaces. Dünser and Billinghurst [25] conclude that it is questionable as to whether a set of general guidelines could be developed for AR systems, partly because such guidelines may only be applicable to the myriad implementation possibilities at a very high level of design. They also state that it is often necessary for researchers to look beyond the computing discipline for suitable evaluation techniques, for example the image quality analysis technique used here from Shelley et al. [77].

With the above points in mind, the user study was carefully planned, and the use of relevant HCI [24, ch. 9] and VR [13] literature supported the design of the user study. However, there are some possible threats to validity. Firstly, as the author ran the user study, it is possible that he was pre-empting common problems and issues as more participants took part, meaning there may have been an element of evaluator bias which could have influenced the amount of errors recorded and also shortened the overall completion time. The three longest completion times were recorded in the first half of the participants, and the three shortest completion times were recorded in the last half. Evaluator bias can be a downside to using the co-operative evaluation protocol, as experienced by Smith and Todd [78].

Secondly, the order in which the participants completed the tasks may affect the validity. Even though the order in which each participant completed the tasks was randomised, it can be seen in Section 5.3.2 that in many cases the task at marker B took much longer to complete than the other markers. It can also be seen in Section 5.3.1 that on average the second marker task attempted took the longest to complete, and it is possible that this is because marker B was randomly selected as the second marker for a greater number of participants than any other. A balanced assignment would have been better here, though 24 participants would have been necessary for a completely balanced assignment as there are 4! possible orderings.

Thirdly, the number of participants in the study was small. Even though it is often sufficient to test only a small number of users [57], this makes statistical analysis less effective as many statistical tests work best with large data sets. However, the sixteen candidates that were recruited was an acceptable number for the statical tests used (namely the Friedman Test, the Wilcoxon Signed Rank Test, the Mann-Whitney U Test, and Spearman's Rank Correlation) whilst still being manageable for a single evaluator.

Finally, the method used to assess the quality of each image produced was a subjective method, and therefore had the scope to vary widely between the three experts. However, unless a completely objective method is used, any subjective method could suffer from this problem. This also raises the question of how quality is defined, and what is a "good" quality rating; if it is decided that 75% is an acceptable quality level, then the images produced by seven of the sixteen participants (43.75%) were not of an acceptable quality. However, if it is decided that 50% is acceptable then all of the participants produced images of acceptable quality, showing how different interpretations can affect the conclusions drawn from the results. This is illustrated in Figure 6.1, where a reference image is shown along with a 50% quality and 75% quality rated image. Ultimately, such visual quality assessment will always be subjective, and the harshest critics are the users themselves, who were given the opportunity to make another attempt at reproducing the image if they were not happy with the result.

6.2.1 Environmental issues

Environmental issues presented some difficulties during the study, including weather conditions and marker placement. The weather during each iteration of the experiment presented a possible confounding factor - on particularly sunny days, many participants experienced problems with screen glare due to the sun reflecting against the highly polished glass surface of the iPad screen, which may have increased the number of errors or time taken. Similarly, windy days made it difficult for some participants to hold the iPad steady, and on the some occasions the music

(a) Reference image 3.

(b) Augmented photo with a 50% quality rating. (c) Augmented photo with a 75% quality rating.

Figure 6.1: Example images with quality ratings.

stand used to hold the reference images was blown over. The issue of lighting is known to affect marker-based tracking systems [54], and the sun also caused some problems due to reflections against the marker surfaces. The markers used were paper enclosed in a shiny plastic wallet, so reflection issues could be minimised by manufacturing markers out of something with a durable matt surface instead.

During the study, it became clear that the placement of the markers was very important for stable tracking. Marker B recorded the most false positive marker readings, which was most likely due to the building opposite the marker, which featured many square windows and doors which could easily have been detected as markers. A greater amount of false positives were recorded after some blinds were removed from a prominent set of double doors[1], approximately half way through the user study. In order to minimise the number of false positive marker detections markers should be placed so they are not opposite features that can easily be confused by the system. Another way to reduce such confusion would be to simply disable the marker-based tracking element of the system when not needed, i.e., after the camera switch. This would remove the majority of these problems.

6.3 The tracking system

Implementing suitable and accurate tracking can be a big technological hurdle for AR systems, especially those outdoors (see Chapter 2). The issues caused by the tracking system used in this project was also the source of some of the more prominent usability problems experienced by the study participants.

There were some minor issues that were encountered due to the dual-camera element of tracking system. Some participants noticed that objects changed size or position after the camera switch, which was confirmed by the ARToolkit developers

---

[1] These double doors were part of the building opposite the marker site, and the removal of the blinds was outside of the evaluator's control.

as being due to the difference in the fields of view between the cameras[2] and it is feasible that this could solved fairly easily by the ARToolkit developers correcting this in the software. While it was a minor annoyance, it did not significantly affect the usage of the system. The actual act of switching between the two cameras did not appear to cause many problems for the users, as no comments about it were recorded during the evaluation and there were also no comments alluding to any issues in the questionnaires.

Most problems related to the tracking system were due to it being marker-based instead of being directly related to the camera switching, for instance the presence of false positive marker detections. Also, many participants commented that the markers were too high for them to use comfortably, and this is an important issue for consideration especially if such a system is designed to be used by a wide variety of age groups. If one suitable marker height suitable for all could not be found then multiple markers could be used for different heights, each using the same virtual object but calibrated differently to account for the difference in height. Other possible solutions could be markers with adjustable heights, or a simple step to allow shorter users to get closer to the markers.

Tracking objects using the device's inertial sensors proved to be problematic, as most participants in the study felt that orientating and position objects was difficult or even frustrating. One reason for this was that the rotation of the object in relation to the rotation of the device was inverted, so that if the device was rotated anti-clockwise the object would rotate clockwise. This was so the user would be able to rotate around the object as if it were in the real world, but it was only convincing if the user kept the device in one position and only rotated it. This was because position tracking on the x, y and z axes was not possible due to technology limitations both in the sensor hardware and software algorithms, so the

---

[2] `http://www.artoolworks.com/community/forum/viewtopic.php?f=22&t=1776` [last access 30th October 2012]

compromise was to preserve the rotation element. If 3D position tracking cannot be added, then a useful addition to a future version of the software would be to allow the user to choose whether they would like the rotation inverted or not. However, it is hoped that with improving technology that more accurate inertial sensors will be included in future devices, and that better sensor fusion algorithms can make full 6 DoF position tracking possible.

A number of participants remarked that they would like to opportunity to reposition the object with greater precision after the camera switch. Two methods suggested were with arrow buttons on the screen, e.g., arrows to move the object forwards and backwards by a small increment, and via multitouch gestures, e.g., swiping the screen with two fingers could rotate an object and using three fingers could reposition the object. These suggestions would certainly make object positioning much easier, and would be a very useful edition to a future edition of the system for greater usability. Adding such functionality would not be too difficult; the iPad has a multitouch screen and multitouch gestures are well-supported in the iOS SDK. Manipulating objects in 3D space is still an active research area; two methods for achieving this using multitouch surfaces are explored in Smith et al. [79].

6.4 User satisfaction

The purpose of the study was to formatively evaluate the software and uncover as many of the usability problems that could be experienced with the system as possible. Section 5.3.7 shows that all of the participants experienced some kind of usability problem, but this is acceptable in a usability study as usability engineering is an iterative process that happens throughout the software lifecycle [56, p. 71], and if this project were a commercial system, then multiple iterations of the usability study would be recommended.

Norman [58] defines the difference between the two types of user errors: *slips*, where the incorrect action was performed for the desired outcome, but the correct action was intended, and *mistakes*, where the incorrect action for the desired outcome was performed with itention. During the study, only 5 users made slips, and fewer still made mistakes. This suggests that the participants had a good mental model of the system [81].

While all users experienced problems, responses to the questions about satisfaction were generally positive but not without room for improvement. No participants said that they absolutely disagreed with the statement that they found using the system enjoyable, although some did fully agree that they found using the system frustrating. This could have been due mainly to the object orientation issues described in Section 5.3.7, but also due to the false positive marker detections (see Section 6.3) and the frequent crashes that occurred. The crashes indicate that more debugging is necessary before deploying the system, but it did appear that many of the crashes were due to the third-party ARToolkit code and therefore out of the control of the developers using the toolkit. It is fair to assume that ARToolkit stability will be improved with each new release, along with the addition of added functionality which could be used to improve the system.

6.5 Considerations for the heritage sector

A number of issues should be taken into consideration for such a system to be deployed in the heritage sector. The development of the system was time consuming and required specialist knowledge in areas including AR, iOS app development and 3D graphics - knowledge which most heritage sector workers do not posses (see Section 2.7). This means that external developers would need to be contracted to carry out the work, which could be expensive for a project which has taken so long to develop. Due to this it would be beneficial to work towards a general framework or engine that can be applied to multiple heritage sites, for which custom content can be made and then inserted to reduce development costs. Figure 6.2 shows which parts of an AR cultural heritage system could be reused, and which must be considered on a site-by-site basis. Maintenance costs should also be considered.

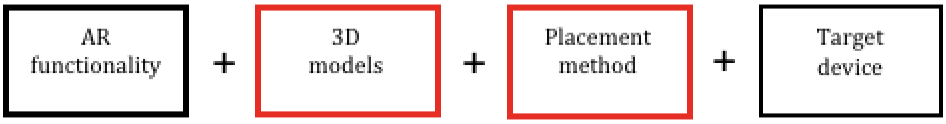


Figure 6.2: A block diagram showing the basic elements required for AR at a heritage site. Black borders are elements that could be completely “plug and play”; red borders are elements that would require special training or expertise for each individual site.

The training needed to use the system should also be considered. If the system is in use at a staffed heritage site, then it may be necessary for staff to have the skills needed to provide a training session to visitors prior to them using the system. If the system is to be used at an unmanned heritage site, adequate documentation needs to be provided to allow visitors to use the system. This could be in the form of posters or displays at the site (if permitted), or videos included as part of the software if they are of a small enough file size.

6.6 Development and technology issues

AR system developers have to overcome numerous problems to develop a working AR system. Many of these are because AR is maturing technology, meaning they they are often exploring uncharted territory with their systems; this project was no different.

One issue that faces developers is the way that computer systems formulated and developed in a lab setting are then moved to an outdoor setting, as this can present a new set of problems to developers and researchers [63]. As mentioned previously, the weather caused some problems during the user study in this project. The system was effectively unusable during rain (even when fairly light) for fear of water damage to the iPad, and also because water droplets obscured the screen and camera and made the device slippery. People are less likely to visit heritage sites in heavy rain, but in light drizzle many people would be happy to venture outdoors if suitably dressed. Possible solutions to this include some kind of covering for the iPad or small shelters at marker sites.

An issue that was not problematic during this project but would be in a real-world setting is the issue of software distribution. The final app size was 40.5MB; smaller than the maximum 100MB that was specified in the non-functional requirements (see NFR9, Section 3.6.4) but still fairly large. It is possible the app could be reduced in size by compressing or some of the 3D models and removing any unused content, or even by optimising the AR code, but even if the size was halved 20MB is still fairly large to download over a potentially slow mobile network. Apple is also aware of this issue, and for this reason imposes a 50MB limit on over-the-air app downloads[3]. It is infeasible that many heritage sites can afford to have any kind of wifi infrastructure to facilitate software downloads, so for now the best solution is for visitors to download such apps before arriving. With speeds and coverage of data networks improving it can be assumed that this issue will be solved in the near future.

One of the major development issues faced during this project concerned the usage of ARToolkit. The toolkit seemed to be fairly robust in its original state, but

---

[3] `http://developer.apple.com/library/ios/documentation/LanguagesUtilities/Conceptual/iTunesConnect_Guide/iTunesConnect_Guide.pdf` [last access 4th December 2012]

trying to modify any of the examples to do anything not intended by the developers was problematic. The documentation was found to be lacking and support was through the community forums, which was often of a poor quality. However, the alternative to using such third-party middleware is to develop one's own tracking library - a project which could have significant development overheads and requires image analysis and 3D graphics expertise. For developing low-cost systems for heritage sites, who often to not have a large budget, using third-party tools seems to be the way forward. However, if such a toolkit were developed specifically for the heritage sector (perhaps funded by a large organisation or a group thereof, such as the EU-funded Arco project [94]), many of the desired domain-specific functionality could be considered at the design stage, which would remove the need for system developers to shoehorn unintended functionality into another organisation's code. Using third-party code also requires the developers using it to rely on the third party to release maintenance updates to support new hardware and software updates; with the rapid changes in mobile technology this is an increasingly big issue.

There were only minor problems encountered with using the iPad and iOS as a development platform. One issue was the somewhat convoluted issue of having to link the developer accounts, application and device together to digitally sign the code, but once resolved it caused no problems. However, this should not be forgotten as part of the maintenance of the system, as the certificates are only valid for a certain amount of time. The actual development process using Xcode and its associated tools was fluid, and the iOS SDK provided a wide variety of very robust functionality with excellent documentation. The only other major issue encountered was the installation of a new version of iOS for the iPad, which was incompatible with the version of Xcode being used. This meant that that a new version of Xcode was needed, along with an upgrade to the latest version of iOS for the development machine to support it. These compulsory upgrades should also be considered in planning for the maintenance of future projects, as they could cause the software to be incompatible with newer devices. Overall, iOS development is generally made very easy, but some process required to deploy the software have a steep learning curve for new developers.

6.7 Summary

This chapter discussed a number of issues that were encountered throughout the course of the project, as well as some of the overarching problems that that face contemporary AR system development. Many of the issues are a result of the relatively short time the field has been in existence - for instance, when compared with traditional graphical user interfaces there are few established design, development and evaluation practices. The toolkits used during the project also suffered from this, as it was felt they were not as robust as they could have been as some of them are still works in progress.

As a prototype of a novel AR mechanism, this project identified a number of usability problems uncovered during the study, many of which could potentially be fixed in future iterations in the software by using the solutions shown in this section. As this is the case, it can be concluded that the usability study was a worthwhile activity. However, through the development of standard evaluation practices future studies could be take a more focused approach based on thorough computer science and software engineering research.

Chapter 7

Conclusions and future work

7.1 Introduction

This chapter summarises the main findings of the project and sets out potential directions for future work.

7.2 Main outcomes

This project has seen the successful design and implementation of an AR “magic camera” system to be used at an outdoor cultural heritage site using a tablet computer. A rigorous usability study was carried out in order to identify usability problems and gather information which could be used to improve the system in a future version of the software.

7.2.1 Revisiting the research questions

The research questions defined in Section 1.2 were:

1. *"Does a tracking system utilising two cameras on a mobile device present a feasible method of tracking for AR?"*

   Although the initial design where both camera feeds were active simultaneously could not be implemented due to technology limitations, the final dual-camera switching mechanism part of the tracking system was a success. The operation worked as intended without problems, and the entire process of switching took around three seconds. In summary, yes - utilising two cameras on a mobile device does present a feasible method of tracking.

2. *"How does the use of a dual-camera paradigm impact usability?"*

   The usability study was very effective at uncovering usability problems. Analysing the study results showed that the AR system demonstrated good learnability and satisfaction. However, all users encountered some kind of usability problem, some of which were deemed high-level problems which may require a system redesign to solve. It was found there there is no "best" way to evaluate an AR system, and that there is a lot of scope for different evaluation methods due to large variations in AR systems. Relevant literature from a variety of disciplines can be used to inform the evaluation design.

3. *"What are the implications for the heritage sector if AR technology is adopted?"*

   The use of third-party development tools was the source of a number of problems in this project, and as such should be carefully considered when planning the development of similar systems. This applies especially with regards to new technologies, such as AR, that are still maturing. However, developing one's own tools for new technology can require a high level of expertise and training in niche subject areas. Due to this, development for such systems can take a long time and cost a lot of money. This will often not agree with the small budgets of many heritage sites, therefore development tools targeted specifically at the heritage sector would be beneficial. The heritage sector would also need to consider additional training needed for heritage site staff who are not technology experts. Deployment and distribution of AR hardware and software also present significant problems for the heritage sector.

7.2.2 Other outcomes

The other relevant outcomes of the research can be summarised as follows:

- Using marker-based tracking outdoors was partially successful. Lighting conditions and false positive marker detection were the main cause of problems, but often it worked without issues.
- 3D rotation tacking only, i.e. 3 DoF tracking of rotation without position tracking, was not intuitive to use and was a source of frustration for some users.
- The weather caused issues when using the system outdoors. Lighting affected the tracking and the users' ability to view the screen, and rainy or windy conditions made it difficult or impossible to use the system as intended.
- User studies are very effective at uncovering usability problems, but they must be carefully planned. Applicable literature from a number of areas can be used to inform the design if none can be found within the AR discipline.

7.3 Limitations of proposed solution

While the system was a success on the whole, there are a number of limitations. Firstly, the user is not able to move around much while using the system, or the effect of the virtual object being part of the real world is lost. This is due to the aforementioned lack of 3D position tracking. Secondly, the marker-based element of the tracking system suffers with the usual issues associated with markers, for example, the poor performance in low light conditions and possibility of false positives. Furthermore, effective use of the system is also subject to weather conditions - high levels of sunlight cause glare, making the screen difficult to see, and rain could cause damage to the tablet.

7.4 Future work

As a result of this project, a number of areas have been identified where further work would be beneficial.

7.4.1 Tracking

There would be a lot of benefit in improving the inertial tracking system. Firstly, full 6 DoF position and orientation tracking using sensor fusion algorithms would greatly improve the user experience. If such tracking functionality was added to the system then users would have a much immersive experience available to them, as they would be able to walk around a site with digital objects remaining in position. Secondly, the marker-based tracking part of the system could also be improved. Different types of markers should be investigated to find ones that suit particular outdoor environments best and result in fewest false positive readings.

7.4.2 Interaction

More interaction with the AR system could provide a more engaging experience to heritage site visitors. This system had a small amount of interaction, but this was added more by necessity than by design (see Section 3.6.2.1). Some participants said they found the positing of objects frustrating, but conversely some said they actually enjoyed the process. To add more interest, interactive features that utilise the multitouch functionality of the iPad could be implemented; for example, the user could touch different objects on screen for additional content, or they could slide their finger over a timeline to show how a particular building has changed over time.

7.4.3 Evaluation

Improvements could also be made to the way the user study was carried out. Firstly, a greater number of participants would be beneficial, as well as more varied participants. Most of the participants recruited were students, meaning that they were all of a similar age and education level and most were technologically able. It would have been useful to recruit some older participants who may not be so comfortable using technology, as well as some children, as these groups also reflect possible users of an AR system at a heritage site.

A future study could also include an alternative method of visualising cultural heritage objects and buildings, so that it could be compared directly to the AR system. This way, the research could suggest whether or not AR systems provide a “better” method of visualisation than another method.

If a future study were to still include the subjective image quality rating element, then more experts could be used to get more consistent results as in Shelley et al. [77], where a panel of ten experts was used. Also, a more finely-grained rating scale could be used, for example 1 - 5 instead of 1 - 3, which may allow the experts to provide more precise ratings. A short questionnaire for each image could be used as an alternative to simply rating four aspects of each image as in Shelley et al. [77], but this may increase the time required for each of the experts to complete their ratings, and could result in the experts becoming fatigued.

7.5 Final conclusion

AR interfaces are often heralded about as being "the next big thing" in technology that will permeate our entire lives. However, technology limitations are the significant barrier to this, and until recently AR technology has often been used as a gimmick for advertising, or used only by computer scientists as part of their throw-away research projects. However, recent projects that have attracted the attention of the media and the general public, such as Google's head-mounted AR display project called *Project Glass*[1], promise to bring AR technology to the masses. In each new generation of technology, new hardware features allow developers to improve the user experience, and this project has shown how this technology can be harnessed and used to help visualise cultural heritage content in an effort to present it in a more engaging way. The final system was not without its problems, but this document has detailed ways in which these problems could be overcome, as well as how how the system could be extended to provide an even more interesting and immersive experience for the heritage site visitor.

---

[1] `http://plus.google.com/+projectglass/about` [last access 10th December 2012]

Appendix A

Post-session questionnaire

Augmented reality user study - Post-session questionnaire

Date: ___________________________________

Candidate ID: ___________________________

Based on your experiences during the study, please state your agreement with the following statements. Please circle your answers and provide any extra information you feel is important.

1. I found it easy to recreate the reference images using the tablet.

   (Disagree) 1 2 3 4 5 (Agree)

   Additional comments:

2. The session was physically tiring.

   (Disagree) 1 2 3 4 5 (Agree)

   Additional comments:

3. I could quickly recreate the reference images using the tablet.

   (Disagree) 1 2 3 4 5 (Agree)

   Additional comments:

4. The system was enjoyable to use.

   (Disagree) 1 2 3 4 5 (Agree)

   Additional comments:

5. I found each new task easier to complete than the last one.

   (Disagree) 1 2 3 4 5 (Agree)

   Additional comments:

6. Using the system frustrated me.

(Disagree) 1 2 3 4 5 (Agree)

Additional comments:

7. I would use the system again.

(Disagree) 1 2 3 4 5 (Agree)

Additional comments:

8. After my visit to a heritage site, I would share the augmented photos I took with friends and family, e.g. on Facebook or another social network.

(Disagree) 1 2 3 4 5 (Agree)

Additional comments:

9. I found that the visual quality of the system was high.

(Disagree) 1 2 3 4 5 (Agree)

Additional comments:

10. The markers were positioned at the correct height for me.

(Disagree) 1 2 3 4 5 (Agree)

Additional comments:

11. The initial training provided before the session was helpful.

(Disagree) 1 2 3 4 5 (Agree)

Additional comments:

12. I had to be reminded about how to use the system during the study.

(Disagree) 1 2 3 4 5 (Agree)

Additional comments:

13. I found the system easy to use.

(Disagree) 1 2 3 4 5 (Agree)

Additional comments:

14. I made a lot of mistakes when using the system.

(Disagree) 1 2 3 4 5 (Agree)

Additional comments:

15. How would you improve the system? Please state below.

16. Please state any further feedback below.

17. If you would like to be informed about the results of this study, please write your email address below.

Appendix B

Reference images

Figure B.1: Reference image 1.

Figure B.2: Reference image 2.

Figure B.3: Reference image 3.

Figure B.4: Reference image 4.

Appendix C

Pre-session questionnaire

Augmented reality user study - Pre-session questionnaire

Date: ___________________________________

Candidate ID: ___________________________

1. Handedness: [ ] Left [ ] Right [ ] Ambidextrous
2. Age: 18-25 26-32 33-39 40-46 47-52 53+
3. Gender: [ ] Male [ ] Female
4. Height:
5. Please indicate your typical computer usage: [ ] Daily [ ] Weekly [ ] Monthly [ ] Never
6. Do you own a smart phone (e.g. an iPhone)?: [ ] Yes [ ] No
7. Do you regularly (at least once a week) use a tablet device with a touch screen?: [ ] Yes [ ] No
8. Before this study, were you aware of Augmented Reality technology?: [ ] Yes [ ] No
9. Before this study, had you used Augmented Reality technology?: [ ] Yes [ ] No
10. Please read and complete the consent form.

Appendix D

Data sheet

Augmented reality user study - data sheet

Date: Participant number: Start time: End time:

	Marker 1 S: F:	Marker 2 S: F:	Marker 3 S: F:	Marker 4 S: F:
Incorrect marker detected				
Screen mistakenly pressed				
User takes photo				
Question asked by participant				
Hint given to participant				
Application crashed				

Notes:

Bibliography

[1] Gregory Abowd, Christopher Atkeson, Jason Hong, Sue Long, Rob Kooper, and Mike Pinkerton. Cyberguide: A mobile context-aware tour guide. *Wireless Networks*, 3:421–433, 1997. ISSN 1022-0038.

[2] Eike Anderson, Leigh McLoughlin, Fotis Liarokapis, Christopher Peters, Panagiotis Petridis, and Sara de Freitas. Developing serious games for cultural heritage: a state-of-the-art review. *Virtual Reality*, 14:255–275, 2010. ISSN 1359-4338.

[3] Ronald Azuma. A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, 6:355–385, 1997.

[4] Ronald Azuma. The challenge of making augmented reality work outdoors. In *Mixed Reality: Merging Real and Virtual*, pages 379–390. Springer-Verlag, 1999.

[5] Ronald Azuma, Yohan Baillot, Reinhold Behringer, Steven Feiner, Simon Julier, and Blair MacIntyre. Recent advances in augmented reality. *IEEE Computer Graphics and Applications*, 21:34–47, November 2001. ISSN 0272-1716.

[6] Chris Baber, James Knight, D. Haniff, and L. Cooper. Ergonomics of wearable computers. *Mobile Networks and Applications*, 4:15–21, 1999. ISSN 1383-469X.

[7] Herbert Bay, Beat Fasel, and Luc Van Gool. Interactive museum guide: Fast and robust recognition of museum objects. In *Proceedings of the International Workshop on Mobile Vision*, 2006.

[8] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 110(3): 346–359, 2008. ISSN 1077-3142. Similarity Matching in Computer Vision and Multimedia.

[9] Benjamin B. Bederson. Audio augmented reality: a prototype automated tour guide. In *Conference companion on Human factors in computing systems*, CHI '95, pages 210–211, New York, NY, USA, 1995. ACM. ISBN 0-89791-755-3.

[10] F. Bellotti, C. Berta, A. de Gloria, and M. Margarone. User testing a hypermedia tour guide. *Pervasive Computing, IEEE*, 1(2):33–41, April-June 2002. ISSN 1536-1268.

[11] Fehmi Ben Abdesslem, Andrew Phillips, and Tristan Henderson. Less is more: energy-efficient mobile sensing with senseless. In *Proceedings of the 1st ACM workshop on Networking, systems, and applications for mobile handhelds*, MobiHeld '09, pages 61–62, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-444-7.

[12] Erkan Bostanci and Adrian F. Clark. Living the past in the future. In *Proceedings of the 2nd International Workshop on Creative Science*, Nottingham, UK, 2011.

[13] Doug A. Bowman, Joseph L. Gabbard, and Deborah Hix. A survey of usability evaluation in virtual environments: classification and comparison of methods. *Presence: Teleoperators and Virtual Environments*, 11(4):404–424, August 2002. ISSN 1054-7460.

[14] Fabio Bruno, Stefano Bruno, Giovanna De Sensi, Maria-Laura Luchi, Stefania Mancuso, and Maurizio Muzzupappa. From 3D reconstruction to virtual reality: A complete methodology for digital archaeological exhibition. *Journal of Cultural Heritage*, 11(1):42–49, 2010. ISSN 1296-2074.

[15] Wolfram Burgard, Armin B. Cremers, Dieter Fox, Dirk Hähnel, Gerhard Lakemeyer, Dirk Schulz, Walter Steiner, and Sebastian Thrun. The interactive museum tour-guide robot. In *Proceedings of the 15th national conference on Artificial intelligence*, AAAI '98, pages 11–18, Menlo Park, CA, USA, 1998. American Association for Artificial Intelligence. ISBN 0-262-51098-7.

[16] Jurjen Caarls, Pieter Jonker, Yolande Kolstee, Joachim Rotteveel, and Wim van Eck. Augmented reality for art, design and cultural heritage: system design and evaluation. *Journal on Image and Video Processing*, 2009:5:2–5:2, February 2009. ISSN 1687-5176.

[17] Marcello Carrozzino and Massimo Bergamasco. Beyond virtual museums: Experiencing immersive virtual reality in real museums. *Journal of Cultural Heritage*, 11(4):452 – 458, 2010. ISSN 1296-2074.

[18] Michael M. Cernea. *Cultural heritage and development: a framework for action in the Middle East and North Africa*. World Bank Publications, 2001.

[19] Keith Cheverst, Nigel Davies, Keith Mitchell, Adrian Friday, and Christos Efstratiou. Developing a context-aware electronic tourist guide: some issues and experiences. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '00, pages 17–24, New York, NY, USA, 2000. ACM. ISBN 1-58113-216-6.

[20] Omar Choudary, Vincent Charvillat, Romulus Grigoras, and Pierre Gurdjos. MARCH: mobile augmented reality for cultural heritage. In *Proceedings of the 17th ACM international conference on Multimedia*, MM '09, pages 1023–1024, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-608-3.

[21] David Claus and Andrew W. Fitzgibbon. Reliable automatic calibration of a marker-based position tracking system. In *Proceedings of the 7th IEEE Workshops on Application of Computer Vision*, volume 1 of *WACV-MOTION '05*, pages 300–305, Washington, DC, USA, 2005. IEEE Computer Society. ISBN 0-7695-2271-8-1.

[22] J. Coughlan, M. Mutsuzaki, and H. Shen. Rapid and robust algorithms for detecting colour targets. In *Proceedings of the 10th Congress of the International Colour Association*, AIC Colour '05, page 959, 2005.

[23] N. Davies, K. Cheverst, K. Mitchell, and A. Efrat. Using and determining location in a context-sensitive tour guide. *Computer*, 34(8):35–41, August 2001. ISSN 0018-9162.

[24] Alan Dix, Janet E. Finlay, Gregory D. Abowd, and Russell Beale. *Human-Computer Interaction (3rd Edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2003. ISBN 0130461091.

[25] Andreas Dünser and Mark Billinghurst. Evaluating Augmented Reality Systems. In Borko Furht, editor, *In Handbook of Augmented Reality, Borko Furht*, chapter 13, pages 289–307. Springer, 2011. ISBN 978-1461400639.

[26] Maria Economou. The evaluation of museum multimedia applications: Lessons from research. *Museum Management and Curatorship*, 17:173–187, 1998.

[27] M. Fiala. Comparing ARTag and ARToolkit plus fiducial marker systems. In *Proceedings of the 2005 IEEE International Workshop on Haptic Audio Visual Environments and their Applications*, page 6, October 2005.

[28] M. Fiala. Magic mirror system with hand-held and wearable augmentations. In *Virtual Reality Conference, 2007. VR '07. IEEE*, pages 251–254, March 2007.

[29] Mark Fiala. ARTag, an improved marker system based on ARToolkit. National Research Council Canada, Publication Number: NRC: 47419, 2004.

[30] Gerald Fritz, Christin Seifert, and Lucas Paletta. A mobile vision system for urban detection with informative local descriptors. In *Proceedings of the 4th IEEE International Conference on Computer Vision Systems*, pages 30–, Washington, DC, USA, 2006. IEEE Computer Society. ISBN 0-7695-2506-7.

[31] J.L. Gabbard and J.E. Swan. Usability engineering for augmented reality: Employing user-based studies to inform design. *Visualization and Computer Graphics, IEEE Transactions on*, 14(3):513–525, May-June 2008. ISSN 1077-2626.

[32] G.F. Gartner and K. Rehrl. *Location based services and telecartography II: from sensor fusion to context models.* Number 2 in Lecture notes in geoinformation and cartography. Springer, 2008. ISBN 9783540873921.

[33] Juergen Gausemeier, Juergen Fruend, Carsten Matysczok, Beat Bruederlin, and David Beier. Development of a real time image based object recognition method for mobile AR-devices. In *Proceedings of the 2nd international conference on Computer graphics, virtual Reality, visualisation and interaction in Africa*, AFRIGRAPH '03, pages 133–139, New York, NY, USA, 2003. ACM. ISBN 1-58113-643-9.

[34] S. Gebbensleben, J. Dittmann, and C. Vielhauer. Multimodal audio guide for museums and exhibitions. In *Proceedings of SPIE Electronic Imaging - Multimedia on Mobile Devices II*, volume 6074, pages 60740S–1 – 60740S–12, 2006. VG-Wort: 32 Standardseiten.

[35] A. Harma, J. Jakka, M. Tikander, M. Karjalainen, T. Lokki, J. Hiipakka, and G. Lorho. Augmented reality audio for mobile and wearable appliances. *Journal of the Audio Engineering Society*, 52(6):618–639, June 2004.

[36] Andrew Howard, Sajid Siddiqi, and Gaurav S. Sukhatme. An experimental study of localization using wireless ethernet. In *Proceedings of the 4th International Conference on Field and Service Robotics*, 2003.

[37] E.D. Kaplan and C.J. Hegarty. *Understanding GPS: principles and applications.* Artech House mobile communications series. Artech House, 2006. ISBN 9781580538947.

[38] Boriana Koleva, Stefan Rennick Egglestone, Holger Schnädelbach, Kevin Glover, Chris Greenhalgh, Tom Rodden, and Martyn Dade-Robertson. Supporting the creation of hybrid museum experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 1973–1982, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-246-7.

[39] Hong Ro Lee, Jung Sub Shin, and Chi Jung Hwang. Invisible marker tracking system using image watermarking for augmented reality. In *Consumer Electronics, 2007. ICCE 2007. Digest of Technical Papers. International Conference on*, pages 1–2, January 2007.

[40] M. Lehtinen, A. Happonen, and J. Ikonen. Accuracy and time to first fix using consumer-grade GPS receivers. In *Proceedings of the 16th International*

Conference on Software, Telecommunications and Computer Networks, SoftCOM '08, pages 334–340, September 2008.

[41] F Liarokapis and M White. Augmented reality techniques for museum environments. *Architecture*, 1(2):95–102, 2005.

[42] W. Linder. *Digital Photogrammetry: A Practical Course*. Springer, 2009. ISBN 9783540927242.

[43] Jyi-shane Liu and Mu-Hsi Tseng. Mediating team work for digital heritage archiving. In *Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries*, JCDL '04, pages 259–268, New York, NY, USA, 2004. ACM. ISBN 1-58113-832-6.

[44] Julian Looser, Raphael Grasset, Hartmut Seichter, and Mark Billinghurst. OSGART - a pragmatic approach to MR. In *Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality*, ISMAR '06, pages 22–25, 2006.

[45] D.G. Lowe. Object recognition from local scale-invariant features. In *Proceedings of the 7th IEEE International Conference on Computer Vision*, volume 2, pages 1150–1157, 1999.

[46] Kristen Lucas and John L. Sherry. Sex differences in video game play: A communication-based explanation. *Communication Research*, 31:499–523, 2004.

[47] Blair MacIntyre, Maribeth Gandy, Steven Dow, and Jay David Bolter. DART: a toolkit for rapid design exploration of augmented reality experiences. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*, UIST '04, pages 197–206, New York, NY, USA, 2004. ACM. ISBN 1-58113-957-8.

[48] S.O.H. Madgwick, A.J.L. Harrison, and R. Vaidyanathan. Estimation of IMU and MARG orientation using a gradient descent algorithm. In *2011 IEEE International Conference on Rehabilitation Robotics*, ICORR '11, pages 1 –7, June–July 2011.

[49] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. Augmented Reality: A Class of Displays on the Reality-Virtuality Continuum. In *Proceedings of the SPIE Conference on Telemanipulator and Telepresence Technologies*, volume 2351 of *SPIE '95*, pages 282–292, Boston, Massachusetts, USA, November 1995.

[50] Ananda Mitra, Jennifer Willyard, Carrie Anne Platt, and Michael Parsons. Exploring web usage and selection criteria among male and female students. *Journal of Computer-Mediated Communication*, 10(3), 2005.

[51] T. Miyashita, P. Meier, T. Tachikawa, S. Orlic, T. Eble, V. Scholz, A. Gapel, O. Gerl, S. Arnaudov, and S. Lieberknecht. An augmented reality museum guide. In *Proceedings of the 7th IEEE and ACM International Symposium on Mixed and Augmented Reality*, ISMAR '08, pages 103–106, Washington, DC, USA, 2008. IEEE Computer Society. ISBN 978-1-4244-2840-3.

[52] A. Monk. *Improving your human-computer interface: a practical technique*. BCS practitioner series. Prentice Hall, 1993. ISBN 9780130100344.

[53] K. Muthukrishnan, S. Dulman, and K. Langendoen. Towards a rapidly deployable positioning system for emergency responders. In *Proceedings of Ubiquitous Positioning Indoor Navigation and Location Based Service*, UPINLBS '10, pages 1–12, October 2010.

[54] Leonid Naimark and Eric Foxlin. Circular data matrix fiducial system and robust image processing for a wearable vision-inertial self-tracker. In *Proceedings of the 1st International Symposium on Mixed and Augmented Reality*, ISMAR '02, pages 27–, Washington, DC, USA, 2002. IEEE Computer Society. ISBN 0-7695-1781-1.

[55] Ulrich Neumann, Suya You, Youngkwan Cho, Jongweon Lee, and Jun Park. Augmented reality tracking in natural environments, 1999.

[56] Jakob Nielsen. *Usability Engineering*. Morgan Kaufmann Series in Interactive Technologies. AP Professional, 1994. ISBN 9780125184069.

[57] Jakob Nielsen and Thomas K. Landauer. A mathematical model of the finding of usability problems. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, CHI '93, pages 206–213, New York, NY, USA, 1993. ACM. ISBN 0-89791-575-5.

[58] Donald A. Norman. Design rules based on analyses of human error. *Communications of the ACM*, 26(4):254–258, April 1983. ISSN 0001-0782.

[59] George Papagiannakis, Sébastien Schertenleib, Brian O'Kennedy, Marlene Arevalo-Poizat, Nadia Magnenat-Thalmann, Andrew Stoddart, and Daniel Thalmann. Mixing virtual and real scenes in the site of ancient pompeii: Research articles. *Computer Animation and Virtual Worlds*, 16:11–24, February 2005. ISSN 1546-4261.

[60] George Papagiannakis, Gurminder Singh, and Nadia Magnenat-Thalmann. A survey of mobile and wireless technologies for augmented reality systems. *Computer Animation and Virtual Worlds*, 19(1):3–22, 2008. ISSN 1546-427X.

[61] Hanhoon Park and Jong-Il Park. Invisible marker tracking for AR. In *Proceedings of the 3rd IEEE and ACM International Symposium on Mixed and Augmented Reality*, ISMAR '04, pages 272–273, Washington, DC, USA, 2004. IEEE Computer Society. ISBN 0-7695-2191-6.

[62] Joonsuk Park and Jun Park. 3DoF tracking accuracy improvement for outdoor augmented reality. In *Proceedings of the 9th IEEE and ACM International Symposium on Mixed and Augmented Reality*, ISMAR '10, pages 263–264, October 2010.

[63] Jason Pascoe, Nick Ryan, and David Morse. Using while moving: HCI issues in fieldwork environments. *ACM Transactions on Computer-Human Interaction*, 7(3):417–437, September 2000. ISSN 1073-0516.

[64] George Pavlidis, Anestis Koutsoudis, Fotis Arnaoutoglou, Vassilios Tsioukas, and Christodoulos Chamzas. Methods for 3D digitization of cultural heritage. *Journal of Cultural Heritage*, 8(1):93–98, 2007. ISSN 1296-2074.

[65] A. Peternier, F. Vexo, and D. Thalmann. Wearable mixed reality system in less than 1 pound. In *Proceedings of the 12th Eurographics Symposium on Virtual Environments*, 2006.

[66] Laia Pujol, Akrivi Katifori, Maria Vayanou, Maria Roussou, Manos Karvounis, Marialena Kyriakidi, Stamatia Eleftheratou, and Yannis E. Ioannidis. From personalization to adaptivity - creating immersive visits through interactive digital storytelling at the acropolis museum. In *Intelligent Environments (Workshops)*, pages 541–554, 2013.

[67] Gerhard Reitmayr and Tom Drummond. Initialisation for visual tracking in urban environments. In *Proceedings of the 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, ISMAR '07, pages 161–160, Nara, Japan, November 13–16 2007.

[68] Gerhard Reitmayr and Tom W. Drummond. Going out: Robust modelbased tracking for outdoor augmented reality. In *Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality*, ISMAR '06, pages 109–118, 2006.

[69] M. Ribo, A. Pinz, and A.L. Fuhrmann. A new optical tracking system for virtual and augmented reality applications. In *Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference*, volume 3 of *IMTC '01*, pages 1932–1936 vol.3, 2001.

[70] Y. Rogers, H. Sharp, and J. Preece. *Interaction Design: Beyond Human-Computer Interaction*. John Wiley & Sons, 2011. ISBN 9780470665763.

[71] Jannick P. Rolland and Henry Fuchs. Optical versus video see-through head-mounted displays in medical visualization. In *Presence: Teleoperators and Virtual Environments*, pages 287–309, 2000.

[72] Dieter Schmalstieg and Daniel Wagner. Mobile phones as a platform for augmented reality. In *Proceedings of the IEEE VR 2008 Workshop on Software*

Engineering and Architectures for Realtime Interactive Systems, pages 43–44, Reno, Nevada, USA, March 2008. Shaker Publishing.

[73] Dieter Schmalstieg, Anton Fuhrmann, Gerd Hesina, Zsolt Szalavári, L. Miguel Encarnação, Michael Gervautz, and Werner Purgathofer. The studierstube augmented reality project. *Presence: Teleoperators and Virtual Environments*, 11:33–54, February 2002. ISSN 1054-7460.

[74] Holger Schnädelbach, Boriana Koleva, Martin Flintham, Mike Fraser, Shahram Izadi, Paul Chandler, Malcolm Foster, Steve Benford, Chris Greenhalgh, and Tom Rodden. The augurscope: a mixed reality interface for outdoors. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves*, CHI '02, pages 9–16, New York, NY, USA, 2002. ACM. ISBN 1-58113-453-3.

[75] Holger Schnädelbach, Boriana Koleva, Michael Twidale, and Steve Benford. The iterative design process of a location-aware device for group use. In Nigel Davies, Elizabeth D. Mynatt, and Itiro Siio, editors, *Ubicomp*, volume 3205 of *Lecture Notes in Computer Science*, pages 329–346. Springer, 2004. ISBN 3-540-22955-8.

[76] Byung-Kuk Seo, Kangsoo Kim, Jungsik Park, and Jong-Il Park. A tracking framework for augmented reality tours on cultural heritage sites. In *Proceedings of the 9th International Conference on Virtual Reality Continuum and its Applications in Industry*, VRCAI '10, pages 169–174, New York, NY, USA, 2010. ACM. ISBN 978-1-4503-0459-7.

[77] Andrew Martyn Shelley, Paul Brunton, and Keith Horner. Subjective image quality assessment of cross sectional imaging methods for the symphyseal region of the mandible prior to dental implant placement. *Journal of Dentistry*, 39(11):764 – 770, 2011. ISSN 0300-5712.

[78] Shamus P. Smith and Susan. Todd. Collaborative evaluation of a haptic-based medical virtual environment. In *Proceedings of the 4th INTUITION International Conference on Virtual Reality and Virtual Environments, 4-5 October 2007, Athens*, pages 102–110. INTUITION, October 2007.

[79] Shamus P. Smith, Elizabeth L. Burd, Linxiao Ma, Iyad Alagha, and Andrew Hatch. Relative and absolute mappings for rotating remote 3d objects on multi-touch tabletops. In *Proceedings of the 25th BCS Conference on Human-Computer Interaction*, BCS-HCI '11, pages 299–308, Swinton, UK, 2011. British Computer Society.

[80] Ian Sommerville. *Software Engineering (7th Edition)*. Pearson Addison Wesley, 2004. ISBN 0321210263.

[81] Nancy Staggers and A. F. Norcio. Mental models: concepts for human-computer interaction research. *International Journal of Man-Machine Studies*, 38(4):587–605, April 1993.

[82] Andrei State, Gentaro Hirota, David T. Chen, William F. Garrett, and Mark A. Livingston. Superior augmented reality registration by integrating landmark tracking and magnetic tracking. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, SIGGRAPH '96, pages 429–438, New York, NY, USA, 1996. ACM. ISBN 0897917464.

[83] Sylaiou Styliani, Liarokapis Fotis, Kotsakis Kostas, and Patias Petros. Virtual museums, a survey and some issues for consideration. *Journal of Cultural Heritage*, 10(4):520 – 528, 2009. ISSN 1296-2074.

[84] Gabriel Takacs, Vijay Chandrasekhar, Natasha Gelfand, Yingen Xiong, Wei-Chao Chen, Thanos Bismpigiannis, Radek Grzeszczuk, Kari Pulli, and Bernd Girod. Outdoors augmented reality on mobile phone using loxel-based visual feature organization. In *Proceeding of the 1st ACM international conference on Multimedia information retrieval*, MIR '08, pages 427–434, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-312-9.

[85] S. Thrun, M. Beetz, M. Bennewitz, W. Burgard, A. B. Cremers, F. Dellaert, D. Fox, D. Hahnel, C. Rosenberg, N. Roy, J. Schulte, and D. Schulz. Probabilistic algorithms and the interactive museum tour-guide robot minerva. *International Journal of Robotics Research*, 19:972–999, 2000.

[86] A.B. Tillon, E. Marchand, J. Laneurit, F. Servant, I. Marchal, and P. Houlier. A day at the museum: An augmented fine-art exhibit. In *Proceedings of the 2010 IEEE International Symposium On Mixed and Augmented Reality - Arts, Media, and Humanities*, ISMAR-AMH '10, pages 69–70, October 2010.

[87] J. Turner. *Automotive Sensors*. Comprehensive Sensors Technology. Momentum Press, 2009. ISBN 9781606500095.

[88] UNESCO. Convention concerning the protection of the world cultural and natural heritage. page 50, Paris, 1972.

[89] Vassilios Vlahakis, John Karigiannis, Manolis Tsotros, Michael Gounaris, Luis Almeida, Didier Stricker, Tim Gleue, Ioannis T. Christou, Renzo Carlucci, and Nikos Ioannidis. Archeoguide: first results of an augmented reality, mobile computing system in cultural heritage sites. In *Proceedings of the 2001 conference on Virtual reality, archeology, and cultural heritage*, VAST '01, pages 131–140, New York, NY, USA, 2001. ACM. ISBN 1-58113-447-9.

[90] Vassilios Vlahakis, Nikolaos Ioannidis, John Karigiannis, Manolis Tsotros, Michael Gounaris, Didier Stricker, Tim Gleue, Patrick Daehne, and Luís Almeida.

Archeoguide: An augmented reality guide for archaeological sites. *IEEE Computer Graphics and Applications*, 22:52–60, September 2002. ISSN 0272-1716.

[91] Daniel Wagner and Dieter Schmalstieg. First steps towards handheld augmented reality. In *Proceedings of the 7th IEEE International Symposium on Wearable Computers*, ISWC '03, pages 127–, Washington, DC, USA, 2003. IEEE Computer Society. ISBN 0-7695-2034-0.

[92] Daniel Wagner and Dieter Schmalstieg. ARToolKitPlus for pose tracking on mobile devices ARToolKit. *Matrix*, pages 139–146, 2007.

[93] Daniel Wagner and Dieter Schmalstieg. Making augmented reality practical on mobile phones, part 2. *IEEE Computer Graphics and Applications*, 29(4): 6–9, July-August 2009. ISSN 0272-1716.

[94] Rafal Wojciechowski, Krzysztof Walczak, Martin White, and Wojciech Cellary. Building virtual and augmented reality museum exhibitions. In *Proceedings of the 9th international conference on 3D Web technology*, Web3D '04, pages 135–144, New York, NY, USA, 2004. ACM. ISBN 1-58113-845-8.

[95] S. You and U. Neumann. Fusion of vision and gyro tracking for robust augmented reality registration. In *Proceedings of the Virtual Reality 2001 Conference*, VR'01, pages 71–78, March 2001.

[96] Zhiqiang Zhang. Ubiquitous human motion capture using wearable microsensors. In *Proceedings of the 2009 IEEE International Conference on Pervasive Computing and Communications*, pages 1–2, March 2009.

[97] Michael Zöllner, Alain Pagani, Yulian Pastarmov, Harald Wüst, and Didier Stricker. Reality filtering: A visual time machine in augmented reality. In *Proceedings of the 9th International conference on Virtual Reality, Archaeology and Cultural Heritage*, VAST '08, pages 71–77. Eurographics Association, 2008. ISBN 978-3-905674-14-9.

[98] Michael Zöllner, Jens Keil, Timm Drevensek, and Harald Wüst. Cultural heritage layers: Integrating historic media in augmented reality. In *Proceedings of the 2009 15th International Conference on Virtual Systems and Multimedia*, VSMM '09, pages 193–196, Washington, DC, USA, 2009. IEEE Computer Society. ISBN 978-0-7695-3790-0.

[99] Michael Zöllner, Jens Keil, Harald Wüst, and Daniël Pletinckx. An augmented reality presentation system for remote cultural heritage sites. In *Proceedings of the 10th International Symposium on Virtual Reality, Archaeology and Cultural Heritage*, VAST '09, pages 112–116, Msida : University of Malta, 2009.