Mobile Egypt

Investigating Mobile Augmented Reality in Museums

Francesca Madeddu (803623)

Department of Computer Science
University of Swansea

Project Dissertation submitted to Swansea University in Partial Fulfilment for the Degree of Master of Science

Singleton Park September 2015

Declaration

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

This dissertation is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by giving explicit references. A bibliography is appended.

I hereby give consent for my dissertation, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Francesca Madeddu (803623)
September 2015

Acknowledgements

I wish to thank my dissertation supervisor Daniel Archambauld, not only for supporting me to produce this work, but also for having helped me to achieve the most from my year of study.

I also wish to thank Rita Borgo, without whom this project would not have been possible.

Finally I wish to thank Harri Mansikkamaki and Tom Owen (Leadin UK Ltd) for their insightful comments. Their support throughout has been instrumental in the success of this project.

This work is part-funded by the European Social Fund (ESF) through the European Union's Convergence programme administered by the Welsh Government.

ESF

Llywodraeth Cymru
Welsh Government

Ewrop & Chymru: Buddsoddi yn eich dyfodol
Cronfa Gymdeithasol Ewrop
Europe & Wales: Investing in your future
European Social Fund

Abstract

Augmented reality is a technology that is becoming more widely available and on a variety of devices, including smartphones. The technology is now being applied to many new applications. When considering museums, augmented reality seems to provide a solution. Often artefacts in a museum cannot be touched or manipulated and augmented reality can offer this more direct form of manipulation. In this thesis, it is presented *Mobile Egypt*, an augmented reality game for smartphones that engages visitors with artefacts in a museum along with two user studies. First, it is performed a guessability study aimed at discovering intuitive gestures with augmented reality on smartphones. Next, the elicited gestures are integrated into the game. Finally, it is performed a qualitative study on the game in a museum setting with volunteers interested in ancient Egypt. The two studies provide a gesture taxonomy for interaction with virtual objects and qualitative validation of the developed application.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

The current project aims at developing and evaluating an augmented reality (AR) game for smartphones. The game consists of a treasure hunt: visitors, using the camera view finder of their mobile devices, need to explore the museum looking for special cardboard boxes called *target objects*. When a target object is scanned, a 3D virtual model of an Egyptian artefact appears on the screen. Visitors can capture the artefact by dragging it into their treasure hunt bag, where all the collected arteafacts are stored. Every time a new artefact is collected, the application displays text about the artefact and a 3D model that can be further manipulated and explored. Reading and understanding the text is needed in order to find clues to locate the next artefact to collect. In fact, as each artefact is tightly connected to the story game, artefacts can only be collected following a predefined order. In this way, as visitors need to understand the story, they implicitly learn about Egyptian history. Additionally, visitors can explore the content of the bag, where the collected objects are stored, anytime they want. The game ends when all the artefacts are captured.

Fig. 1.1 The digitalisation process of an artefact, from the real Anubis to the virtual object.

1.1 An AR Game to Overcome Museums Limitations

People visit museums for a variety of different reasons. They see museums as interesting places to meets friends, or as alternative refuges for when the weather is miserable. They could also be interested in a specific museum because it is highly recommended by a friend or they can have a deep interest towards a specific subject. Different motivations mean that each visitor comes to a museum with different expectations and needs. However, despite all the new possibilities that the modern technology now have available, museums tried to engage the visitors' curiosity with supporting tools such as books, audio-guides, plaques with written explanations or, in the best case, tour-guides. Therefore, the purpose of the thesis is to propose a method, supported by the modern technologies, to break this typical approach: visitors will be offered with an engaging, pleasant and satisfying user experience through an AR game for smartphones.

The following paragraphs present the potential benefits of associating AR with the development of a smartphone game application in the context of a museum.

Overcoming artefact fragility constraints Museums are typically places characterised by a low interactivity, where visitors are afraid to touch anything. One of the main reasons is that, by definition, museums usually store unique, precious and delicate material. Some of the artefacts are so delicate that becomes impossible for the museum to let the visitors see them, and they must be stored in a warehouse where no one can benefit from them. Through the use of AR, one can overcome these constraints. In fact, an artefact which is not accessible or touchable, can be acquired as 3D model and shown to the visitor by the game. Through the interaction with the augmented artefact, visitors have, for the first time, a chance to fully interact with the environment around them.

Overcoming time and space constraints Space can be a problem for a museum under several perspectives. There can be space constraints, as in certain time of the year or on specific days, museums can become overcrowded, not allowing people to see all the artefacts they want to see. Sometimes, there is not enough space to show all the pieces owned by the museum. Finally, museums have physical boundaries, meaning that once the visitors leave the museum, they cannot longer access all the information. The same effect which is caused by the fact the museums clearly also have opening times. Developing a game, which allows visitors to take virtual pieces of the museum outside its physical and temporal borders, helps overcoming all the cited constraints. In fact, a functionality of the game is to allow visitors to retrieve and explore the collected artefacts anytime they want. Potentially, extra information could be also made available to visitors even once the have left the museum, by using a internet connection.

Supporting learning As mentioned, in order to find the next target object to scan, visitors need to read some information on the previously collected artefacts, as clues are embedded in the text. Therefore, if visitors want to win the game, they need to learn about the story of the game, which is a story about the Egypt culture. Hence, visitors learn in a subtle and enjoyable way. Moreover, as challenge is a important factor for maximising learning, the game appears as a perfect vehicle to deliver this knowledge.

1.2 Understanding User Interaction with AR

With the increasing power and decreasing cost of mobile devices such smartphones and tablets, technologies once confined to research laboratories are becoming available to the mass market. This trend not only increases the accessibility of these technologies in diverse settings, but it also makes it easier from programmers to develop new applications. Today, even casual programmers can develop virtual reality applications simply using a smart phone mounted on a *Google cardboard*[1], a low cost headset equipped with two lenses to create the impression of stereoscopic 3D. Many organisations are investing in AR products such as the *Google Glasses*[2], and the Microsoft *HoloLens*. All of these mixed reality products reside in the virtuality *continuum* [48], which extends from reality to a completely synthetic virtual environment. Within this continuum, AR is the technology that overlays virtual information over the real world. AR enhances the user's perception and interaction with the real world by providing them with information that cannot be directly detected with their senses, helping them perform real-world tasks [6].

AR has traditionally focused on the medical [67, 7, 15, 88, 75], military [31, 46, 95, 73], manufacturing [18, 65, 81], visualisation [76, 63, 24, 35, 3, 30, 23], and robotics [34, 80, 59, 27, 78, 93, 25, 62] domains, but more recently have involved applications for urban

[1]https://www.google.com/get/cardboard/
[2]https://developers.google.com/glass/

planning and civil engineering [1, 87, 10, 42, 39, 11], navigation and path planning [56, 44, 74], marketing [77, 17, 22], entertainment and game [19, 5, 20], education [13, 96, 82, 40, 21], and tourism [29, 85] [32, 38, 28]. With the increased accessibility of this technology, AR is now reaching non-traditional application areas such as the museums [55] and archaeological cites [85]. However, these applications traditionally support a one way interaction, where users are presented with this information and cannot interact with it directly. These applications do not allow the user to fully exploit potential of AR technology. For example, museums often store valuable and delicate items that usually are not available for physical manipulation. Similarly, certain views of large objects are inaccessible due to their size (such as the rear of a heavy object that cannot be moved). AR not only provides an an alternative way to access these objects but it also makes interaction with them virtually possible.

Before introducing a new technology such as AR into these application areas, we need to understand how people interact with them. If we do not, we could develop a fully functional application, that cannot be used by visitors to the museum. Therefore, we need to first answer the following questions: How are people going to try approach the new technology? How are they going to interact with it? How does the previous experience with technology affect their interaction? Is there a way to design the interaction so that it supports the natural capabilities of the user? In trying to find answers to these questions, it is essential to involve users in the process from the first step.

1.3 Aims and Objectives

Aim of the project is to develop an educational AR game, for visitors of the Egypt centre, to support learning and offer an engaging user experience. Therefore, project objectives are:

- Developing *Mobile Egypt*, a game using AR for the Egypt centre museum;
- Creating a story game and identifying key aspects to deliver knowledge and support learning;
- Defining a taxonomy for gestures and understanding user's mental models when interacting with AR in laboratory settings;
- Investigating user interaction with AR to understand how to support and develop strategies for encouraging interaction;
- Evaluating end-user elicited gestures and understand user's mental models in AR in the field;

- Understanding how *Mobile Egypt* behaves in the real museum environment by testing it with real users;

- Defining a possible application for AR in museums.

1.4 Target Users

Visitors of the Egypt centre are mainly children accompanied by schools or families. However, because involving children in the design and evaluation process would be challenging, the application is developed without a specific target range age in mind. A positive aspect is that, as mentioned in *the Tate Kids Guide to Creating Games for Galleries*, by Sharna Jackson [9]:

> *"Children today are very media literate and savvier* [...] *in practice this means avoiding any slang, limiting the use of cool, refraining from using emoticons, being wary of over punctuating!!!!! (it looks desperate), definitely no comic sans (*shudders*), sparing the use of primary colours and absolutely no swapping s's for z's. Don't try to be 'down' with them".*

Therefore, it is not the case that developing a game which satisfies several adults public needs, will not meet children's requirements too.

1.5 Content Organization

The content of this thesis is divided as following:

Chapter 1 introduces the current thesis presenting motivation, aims and objectives.

Chapter 2 presents a brief history overview of AR, introducing involved enabling technologies, technical limitations, fields of applications, and research areas. Then, related work are presented, divided according to the categories this thesis aims to contribute, that are *interaction with AR*, *mobile interaction* and *application of AR in museums*.

Chapter 3 introduces different possible approaches for designing interaction and it justifies the choice of opting for a particular participatory design method, known as guessability study. Therefore the chapter presents results of a guessability study performed to gather an understanding on user interaction with AR and to define a gesture taxonomy for AR.

Chapter 4 presents *Mobile Egypt*, the fully functional application developed following the design implications defined through the guessability study. In particular application workflow, user interface layout and usability aspects are presented.

Chapter 5 presents a second study, a qualitative study performed in a real environment, the Egypt centre of Swansea. The study aimed to verify that results from the guessability study hold, to further investigate user interaction with AR and to investigate possible applications of AR for museums.

Chapter 6 presents conclusion.

Chapter 7 presents possible future directions of the current work.

Further details on story game, application development, application prototyping and implementation of the qualitative study are reported in Appendix A, B, C and D

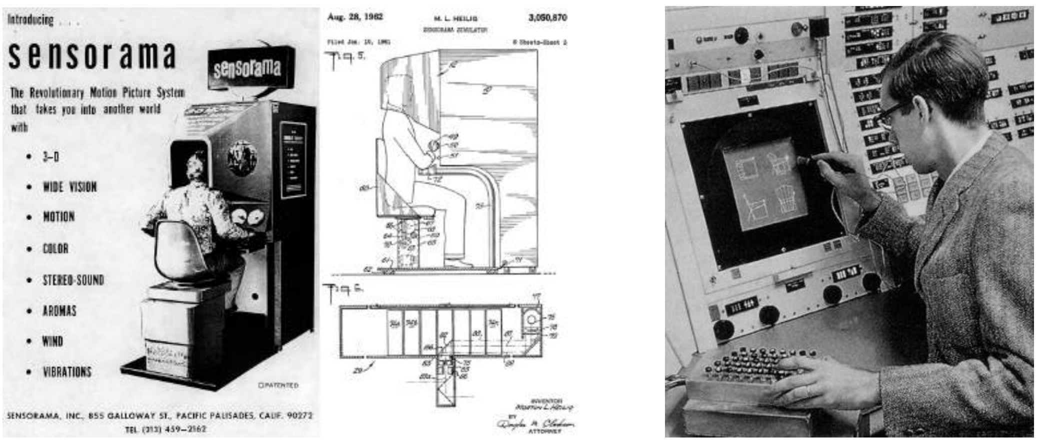
Chapter 2

Related Work

2.1 A Brief History of Augmented Reality

The idea of augmenting reality with additional virtual information is not new. *Sensorama*, a machine developed by the cinematographer Morton Heilig [52] in the 1962, can be already considered an example of immersive, multi-modal technology implementation (figure 2.1a). Only one year later, Ivan Shutherland [79] developed *Sketchpad*, a system which is now considered the ancestor of modern CAD programs and which using a light-pen was able to create and manipulate objects in engineering drawings (Fig. 2.1b).

However we have to wait until the 1990 to see the *augmented reality* term conied by Caudell [18], a Boing researcher who worked on the development of the first heads-up, see-



(a) Sensorama. (b) Sketchpad.

Fig. 2.1 First examples of AR systems.

through, head-mounted display (HMD), and other 8 years to assist to the first AR conference, the international Workshop of Augmented Reality '98.

Azuma et al. [4, 6], give one of the widest accepted definition, describing an AR system to have the following properties: 1) combines real and virtual objects in a real environment, 2) runs interactively, and in real time; and 3) register (aligns) real and virtual objects with each other. This definition underlines three different aspects: AR is not restricted to any particular display technology such as the HMD, as it used to happen at the beginning; AR can involve all of our senses and not only sight; AR can either add virtual objects to a real environment or remove real objects, task commonly defined as *diminished reality*. Milgram [48] defines a *continuum* where AR is the "middle ground" between *virtual reality* (VR) and telepresence (figure 2.2). While in VR the user is completely immersed in a synthetic environment and he cannot see the world around him, in AR virtual objects are integrate inside a real environment. Therefore, AR supplements reality, rather than completely replacing it [6].

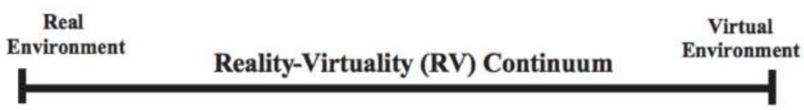


Fig. 2.2 Reality-Virtuality Milgram's Continuum [48].

Besides the technical progresses of last decades, elementary components of AR technology remains the same and they include displays, trackers, graphic computers and software. Display technologies are presented in the following section, while trackers and graphics in subsection 2.1.2.

2.1.1 Displays

Considering only visual displays (aural and haptic are possible and object of research, while olfactory and gustatory are less developed), the augmented layer can be presented using three different technologies: video see-through, optical see-through and projective.

Video see-through display In a video see-through display the video is acquired by an integrated camera and it functions as background for the AR layer. See-through displays are usually characterised by a relatively low resolution, low contrast, low brightness and small field of view. Moreover, they suffer of parallax error, due to the fact the camera cannot be mounted in the true eye location. However, they are lightweight, inexpensive and easy to implement. Since the reality is digitalised, they also allow to easily solve the occlusion problem, by mediating or removing objects.

Optical display In a optical display the AR is overlaid a transparent display, such as in the *Goole Glasses*. Optical displays are in real world resolution, they are inexpensive and they do not suffer of the parallax error. They are also safer as not subject to power failures, making them more suitable for military and medical domains. However, they require the integration with other devices, such as cameras, for registration, and they are characterised by low brightness and contrast, making them not suitable for outdoor environments [84]. Other approaches directly involve the human eye, either by drawing the AR on the retina using low-power lasers [33], such as the virtual retina display (VRD) or through the use of bionic contact lenses [45].

Projection display In projection displays the AR layer is directly projected on the physical surface to be augmented. In the simplest case the AR layer is coplanar with the surface. In this case there is the advantage of covering wide surfaces for a wide field of view. In most sophisticated solutions the use of a head mounted projective display (HMPD) is essential to project the AR layer along the user's line of sight. As in the case of traditional optical displays, projection displays are characterised by low contrast and brightness, therefore they are only suited of indoor environments.

An additional classification, based on the position between the user and the real environments allow to classify displays as head-worn, hand-handled and spatial displays.

Head-worn display Head worn displays (HWD) are mounted over the user's head, with the display positioned in front of their eyes and they include the already mentioned VRDs and HMPDs. Limitation of HWD is that, because they have to be connected to graphic computers, they affect the user mobility.

Hand-held display Hand-held displays are discussed in subsection 2.2.1.

Spatial display Spatial display are statically installed within the environments and they are suited for large presentations and exhibition where a very little interaction is required. These include screen based video, see-through displays, spatial optical see-through displays and projective displays. A problem with spatial display is that they could appear wrongly aligned from users prospective depending on their position.

Table 2.1 summarise characteristics, advantages and disadvantages of each approach. Most of these solutions consist in specific purpose systems, developed ad hoc to display the AR layer. This represents a challenge when trying to make AR accessible for people. Not every organisation could have the possibility to invest huge budget for the purchase of

Positioning	Head-worn				Hand-held	Spatial		
Technology	**Retinal**	**Optical**	**Video**	**Projective**	**All**	**Video**	**Optical**	**Projective**
Mobile	+	+	+	+	+	−	−	−
Outdoor use	+	±	±	+	±	−	−	−
Interaction	+	+	+	+	+	Remote	−	−
Multi-user	+	+	+	+	+	+	Limited	Limited
Brightness	+	−	+	+	Limited	+	Limited	Limited
Contrast	+	−	+	+	Limited	+	Limited	Limited
Resolution	Growing	Growing	Growing	Growing	Limited	Limited	+	+
Field-of-view	Growing	Limited	Limited	Growing	Limited	Limited	+	+
Full-colour	+	+	+	+	+	+	+	+
Stereoscopic	+	+	+	+	−	−	+	+
Dynamic refocus (eye strain)	+	−	−	+	−	−	+	+
Occlusion	±	±	+	Limited	±	+	Limited	Limited
Power economy	+	−	−	−	−	−	−	−
Opportunities	Future dominance	Current dominance			Realistic, mass-market	Cheap, off-the-shelf	Tuning, ergonomics	
Drawbacks		Tuning, tracking	Delays	Retro-reflective material	Processor, Memory limits	No see-through metaphor	Clipping	Clipping, shadows

Table 2.1 Characteristics of display technologies for AR [84].

enabling technologies. Therefore, when developing a game for museums, the most obvious choice focused on general hand-held devices such as smartphones. In this way, the museum can be freed by the bundle of investing and maintaining expensive technological systems, as visitors bring their own devices.

2.1.2 Tracking

From a technological point of view, the *registration* problem is one of the most basic problem limiting AR. The registration problem comes from the fact that real and virtual objects must be properly aligned with each other, in order to perpetuate the illusion that the two worlds coherently coexists together. This problem becomes even more serious in application such as the medical one, where there cannot be a margin for error. For prepared indoor environments, visual tracking techniques offerer nowadays good results and they usually rely on distribute fiducial markers along the environment in well known locations. On the other hand, for outdoor unprepared environments, because of the impossibility of using markers, tracking and calibration are usually not so accurate. Other technological limitations include the hardware portability, as devices should be small and light, and the processing power, as it should be fast enough to display graphics. Low latency and system delays, for example, are often the largest source of registration error.

Because the game had to be developed for a museum, therefore in a indoor environment, fiducial markers where chosen to manage tracking. For the purpose of the game, a particular set of fiducial markers, called target objects, was developed. Target objects are presented in Chapter 4

Table 2.2 Examples of developed target objects created adding fiducial markers on the top of cardboard boxes.

2.2 Research Areas

Zhou et al. [97] identify eleven AR research areas. Table 2.3 shows percentages of papers published at ISMAR for each area from 1998 to 2007.

Topics	**% Papers**	**% Citations**
Tracking	20.1	32.1
Interaction	14.7	12.5
Calibration	14.1	12.5
AR App.	14.4	12.5
Display	11.8	5.4
Evaluations	5.8	1.8
Mobile AR	6.1	7.1
Authoring	3.8	8.9
Visualization	4.8	5.4
Multimodal	2.6	0
Rendering	1.9	1.8

Table 2.3 Trends in AR at ISMAR from 1998 to 2007

According to this classification, this thesis aims to contribute to three different areas. It contributes to interaction by defining a set of gestures to interact within AR on smartphones by running a guessability study and subsequently evaluating it in the field through a fully functional AR game. Secondly, it contributes to applications, through the development of an AR game for museums to deliver knowledge and enhancing user experience. Finally, it contributes to mobile, by using hand-held devices in our studies, in particular smartphones, to deploy the application. Thus, related work are divided along these lines.

2.2.1 Interaction in AR

Azuma et al. [4] recognise two trends in supporting AR interaction: 1) using heterogeneous devices to leverage the advantages of different displays and 2) integrating with the physical world through tangible interfaces. An example of basic interaction is offered by the Raskal et al. [64] spatially-augmented reality technique which uses multiple ceiling-mounted light projectors inward to graphically augment table-top scaled physical models of buildings. Projectors can be also embedded into wearable devices, allowing to align the AR layer to the viewer's line of sight. In doing so, head-worn projectors remain the most commonly used approach [49, 37, 70]. Other approaches, such as wrist-worn projection displays [14, 12], have also been investigated. Displays for AR are divided between optical and video see-through (see Capter 1 for details on displays). In all these approaches, interaction is unilateral and AR is used for annotation or visualization purposes. More sophisticated types of interaction include tangible interfaces [83] where markers are applied to tangible real objects which act as portals to digital information. This interaction metaphor is known as tangible augmented reality and since its introduction, it is one of the most commonly used [41]. An example of this bidirectional approach is given by the haptic user interfaces [84], such as gloves. FingARtips, for example, supports a fingertip-based interaction using image processing software and gloves with fiducial markers to track gestures [16]. Even though gloves are commonly used in virtual reality, they struggle to succeed in AR contexts, as they limit the use of hands for normal tasks and are considered socially awkward [84].

Hand-held Interaction in AR

According to the Rekimoto [66] magnifying approach, an hand-held device can assume the role of a portable window into the virtual world. Following this metaphor, Rhos [69] developed three game prototypes using ubiquitous product packaging and other passive media as backgrounds for hand-held augmentation. Ahn and Han [2] developed an indoor AR evacuation system for a smartphone using personalised pedometry. In a collaborative context, AR implementations have used PDAs for multi-user games [86] and smartphones to interact with public displays [8]. Lv [47] proposed a novel smartphone wearable hybrid interaction framework which supports users interaction with hand/foot gesture.

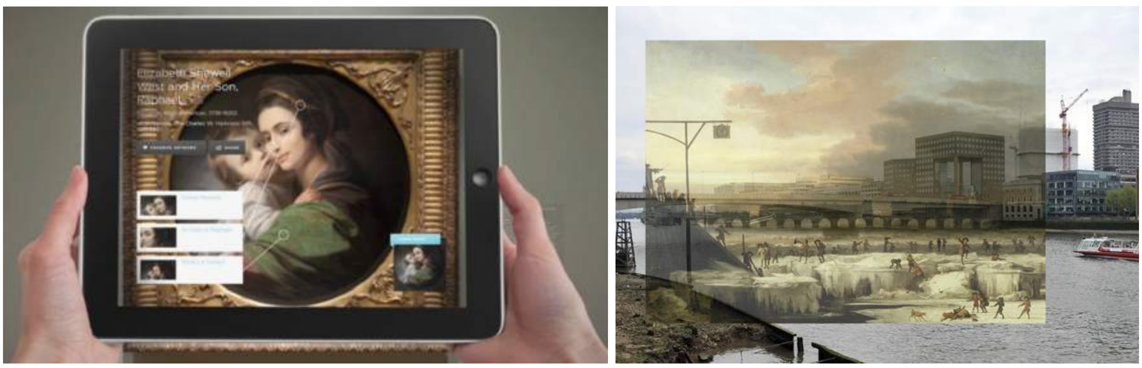
Aim of this thesis is to take the Rekimoto approach further. While traditionally hand-held devices are seen as portable windows into a virtual world, where the user is only allowed to look at it through the device's screen, the current work shows how a support for a more engaging interaction can be developed. Instead of passively observing, users should be able to actively engage with the virtual objects. Smartphones, thanks to their pervasiveness, portable size and familiarity, are well suited to support this interaction.

2.2.2 Mobile Interaction

Smartphones, can potentially support a wide variety of interaction styles. Motion gestures can easily take advantages of all smartphone's built-in sensors [71] while other modalities exist that investigate non-conventional channels [68, 43]. Non traditional standard surface interaction techniques, such as single point of pressure [50] and touch key design for one-handed thumb interaction [60] have been also proposed. Despite all the attempts to introduce novel ways of interaction, surface gestures over a smartphone touchscreen are still considered as standard.

Therefore, this thesis, aims to understand if and how the introduction of a new technology such as AR will affect the user interaction with well known devices such smartphones. Because AR is a metaphor for reality, it leads to ask if it will push users to ask for a more concrete way of interaction, or if the previous experience and learnt idioms will have a greater effect.

2.2.3 Applications for AR in Museums



(a) ArtLens (b) Street Museum

Fig. 2.3 Examples of AR mobile applications for museums.

Between AR domains there is often overlapping. In the context of museum exhibitions, contributions could come from many different areas. From visualisation, by annotating museums spaces or collections with virtual location-based information. From touring and data navigation, by investigating how to support the visitor's journey through exhibits [74]. Good examples of AR applications created and currently used by museums are ArtLens [54] and StreetMuseum [55]. ArtLens, a tablet application developed by the Cleveland museum, can recognise artworks in the collection seamlessly, offering access to additional interpretive content [53]. StreetMuseum is a smartphone application which uses either maps or GPS to determine the user location and to overlay historical photographs over the real world. The coexistence of virtual objects and real environment, allows learners to understand complex spatial relationships and abstract concepts [3], making AR a key technology for educational purposes.

In museum contexts AR is traditionally used to present 2D virtual layers of information that can be accessed but not manipulated. This thesis aims to take this approach further. Therefore, it is here explored the possibility of using an AR game for smartphones to enable access and interaction with virtual artefacts of a museum collection. This work approach aims tonovercome several museum limitations such as physical spatial constraints and low interaction.

Chapter 3

Guessability Study for AR with Smartphones

The purpose of the guessability study is to investigate user interaction with smartphones in an AR context. In particular the study was designed in order to: 1) define a gesture taxonomy with a specific focus on selection, rotation and zoom tasks, 2) uncover insights on the participants' mental model, 3) discover implications for the development of AR applications.

Results from the guessability study were accepted and presented as an extended abstract at EGUK Computer Graphics & Visual Computing (CGVC 2015) conference[1]. Results were also included in a paper submitted for the CHI 2016 conference[2].

3.1 Methodologies for Interaction Design

Two common approaches exist for defining a set of gestures: the designer invents them or they are elicited from the end user. Within the first category, Hurst et al. [36] defined a set of gestures for interacting with AR for basic tasks as translation, rotation, and scaling of virtual objects. They measured the performance (time and accuracy) and engagement (subjective user feedback) of each interaction. Recognising the limitation of surface gestures, they created finger tracking supported gestures. Finger tracking can be performed through fiducial markers [16] but some researches avoid the use of markers or gloves as they can be cumbersome [57]. Rather than simply focusing on the definition of a gesture set, Wu et al. [94] developed and evaluated a novel set of design principles to support multi-hand

[1]http://cgvc15.cs.ucl.ac.uk/.
[2]http://chi2016.acm.org/wp/.

gestural interaction on direct-touch surfaces. White et al. [89] developed visual hints for tangible gestures in AR.

A downside of letting an HCI experts explicitly design a gesture set is that they are often complex and far from natural for the end user. In fact, Morris et. al [51] performed a study to compare two gesture sets for interactive surfaces: one created by an end-user elicitation and one created by a group. Results showed that participants preferred the first set of gesture as those defined by HCI experts were usually perceived as physically and conceptually complex.

Therefore, a better methodology would involve the user in this process from the beginning. For this purpose, Wobbrock [92] developed the guessability study methodology. In this methodology, gestures are elicited by end users: participants are given an effect and are asked to perform the gesture they think will cause that effect. Capturing the interaction of the participants allows researchers to identify a set of gestures using guessability and agreement statistics. Several researchers within and outside AR have used the Wobbrock methodology to define gesture sets for novel technologies such as surface computing [92], interactive data graphics [90], mobile [72] and motion interaction with children [26].

Although Piumsomboon et al. [61] defined a set of hand gestures for AR using HMDs, interaction with hand-held devices such as smartphones remains unexplored. As such, for the first study of the current work, a guessability study methodology to elicit a set of gestures in this context is used.

3.2 Study Design

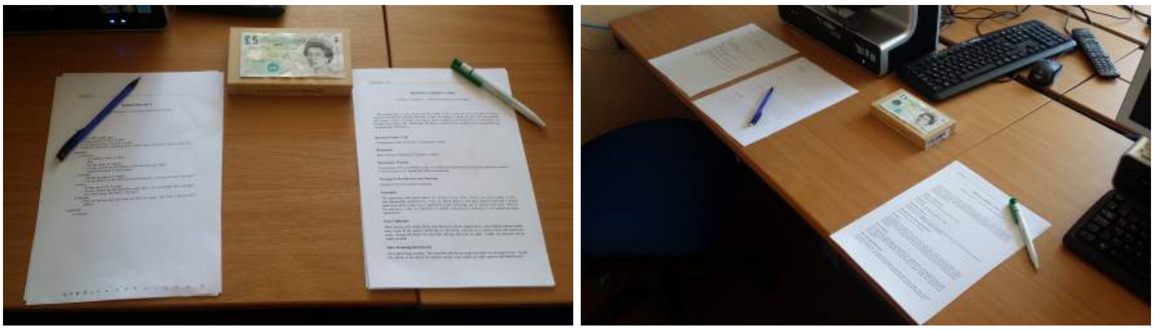


Fig. 3.1 Guessability study settings.

3.2.1 Users

The study involved 15 participants (60% male 40% female) with an age range of 19 to 40 years old. Most participants were students (93%), of whom 73% were from the Computer

Science department. The majority (93%) of participants declared to be between normal to expert mobile users and all of them owned at least one mobile device. Finally, even though 40% knew what AR was, only 26% had previous experience with AR applications.

3.2.2 Apparatus

For the purpose of this study, a custom application to display the virtual objects and to record users interaction with the smartphone screen was developed. The smartphone used to deploy the custom application was a Samsung Galaxy Note 3, measuring 151x79x8mm. For each surface gesture, the application captured the coordinates of the user's fingers and saved them as images, where the colour of each line represents a different finger. Each time the participant touched the screen a screenshot was also saved to contextualise the captured gesture. The application was developed to give no feedback.

The target object was a rectangular box with an image on top. This image, once captured by the camera, triggered the AR layer (figure 3.2). The choice of using a box, rather than a sheet of paper, aimed to improve the affordance for manipulation. In this way, the bias of surface and motion gestures over direct manipulation was mitigated.

3.2.3 Data Gathering

A structured questionnaire was used to gather demographic information and technological knowledge of mobile technologies and AR. During the experiment the researcher used a data sheet to take notes on participant's interaction with the smartphone. Interaction was also partially recorded (only surface gestures) by the preinstalled logging application. Finally, a semi-structured interview was conducted to discuss with participants the elicited gestures. All stages of the experiment were audio-recorded.

3.2.4 Tasks

Participants were given a smartphone which, when pointed at the target object, showed two virtual objects: a 3D model of an Anubis artefact and a bag. They were then asked to complete four tasks: moving Anubis inside the bag (selection), looking at Anubis close up (zoom) and looking at other sides of Anubis (rotation); finally, they were asked to freely explore the scene. To avoid bias in the explanation, the words chosen for the task description were focused on the effect to achieve rather than on the action needed to obtain it.

Fig. 3.2 Example of images captured by the logging application for a rotation task.

3.2.5 Procedure

Firstly, participants were welcomed and informed about their rights. Secondly, they were given an experiment description paper to introduce them to the AR concept and the study procedure. Then, they were given a questionnaire to complete. As the questionnaire was completed, participants were given a smartphone and asked to perform the four tasks following a talk-aloud protocol. As target recognition was outside the scope of the experiment, the researcher demonstrated to participants how to make the AR appear the first time. Finally, the researcher conducted the final interview. The experiment lasted 15 to 30 minutes and the participant sat next to the researcher in front of a rectangular desk. The target object was positioned in front of them at a reachable distance.

3.3 Results

Table 3.1, 3.2 and 3.3 respectively list the 18 gestures elicited for selection, zoom and rotation tasks. Three types of gestures can be identified: *surface gestures* are performed on the touch screen of the device; *motion gestures* are performed by moving the device and are captured by the device sensors (e.g. gyroscope, accelerometer); *direct manipulation* are performed manipulating the target object. The *mixed* category indicates those cases where participants performed a combination of two different gestures (such as moving the device and Anubis closer to each other at the same time).

During the experiment an additional gesture classification emerged. When participants were asked to guess the gesture to cause a specific effect, they performed their *first attempt gesture*, e.g. the first gesture which came to their mind. However, it frequently happened that participants did not stop trying, gradually discovering new gestures which, in some cases, were defined as more satisfactory, hence defined as *preferred gestures.* Finally, the *standard gesture* was the one that, according to participants' opinion, most of people would try when asked to guess a gesture to achieve a specific effect. For example, for the zoom task, several participants tried to move the smartphone close the Anubis as first attempt gesture, and they recognised it as their favourite; however, when they were asked about their opinion on which kind of gesture other people would have done, some of them chose the pinch to zoom surface gesture.

3.3.1 Selection

Eight gestures were elicited for the selection task. The strongest consensus was over a drag and drop gesture from Anubis to the bag, in fact 93% of participants tried the gesture as the first attempt and they also recognised it as standard. Similarly, two participants tried the same gesture but in the opposite direction, that is from the bag to Anubis; however, this gesture was never tried as first attempt. All participants agreed on the expected feedback, that was to see the Anubis statue (or the bag) moving following their finger, as if it was anchored to it, according to a traditional drag and drop behaviour. Other elicited surface gestures consisted in single or double tap on different areas of the screen. For example, two participants respectively tapped on the screen corners and on Anubis in the attempt of triggering the appearance of hidden menus or contextual information. Another participant also tapped on the bag in the attempt to open it, to then drag Anubis into it. Only two participants tried a motion gesture, a rigid motion from Anubis to the bag or shaking the phone. When performing the gesture they verbally explained that they expected a direct cause-effect relationship between the smartphone and Anubis, so that sliding the phone, for example, should have caused Anubis to move towards the same sliding direction. Finally, none of the participant tried a direct manipulation by touching the target object.

		1st	2nd	3rd	Std
Surface	One finger drag'n'drop from Anubis to the bag	14		1	14
	One finger drag'n'drop from the bag to Anubis	1	2	1	
	Single tap on Anubis		2		
	Double tap on Anubis		1		
	Single tap on the bag			2	
	Single tap on background		1		
Motion	Rigid phone motion from Anubis to the bag	1			
	Shake the phone		2		1

Table 3.1 Elicited gestures for selection.

3.3.2 Zoom

Five gestures were elicited for the zoom task. Few participants pointed out that the most natural gesture could have depended on their distance from the target object. In particular they expressed a preference on moving the device for closer objects and on the pinch to zoom for farther objects. However, this contradicts results. In fact, most of participants tried surface (60%) rather than motion (26.6%) gestures even if in the experiment settings the target object was close to participants. Within surface gestures, 53.3% of participants tried the pinch to zoom as first attempt, but the percentage rises up to 73.3% when it comes to identifying a standard gesture for the task. The rest of the participants tried mixed gestures: moving the device (motion) and the target object (direct manipulation) closer to each other at the same time or moving the device closer to Anubis (motion) and pinching to zoom (surface) at the same time. In terms of feedback, 86.6% of participants expected all the scene (both virtual objects and background) to be enlarged as consequence of surface gestures. However one third of participants mentioned that they would have preferred if only Anubis was enlarged, as it was the model, and not the background, the object of their interest.

		1st	2nd	3rd	Std
Surface	Pinch to zoom	8	2		11
	Double tap	1	1	1	
Motion	Move the phone closer to the Anubis's target object	4	1		2
Mixed	Move the phone and the Anubis's target object closer to each other	1	2		1
	Move the phone closer to the Anubis's target object + pinch to zoom	1			1

Table 3.2 Elicited gestures for zoom.

3.3.3 Rotation

Five gestures were elicited for the rotation task. Overall we saw a trend change and elicited gestures were divided in 40% for direct manipulation, 33% for motion and 22% for surface gestures. For the first time, the direct manipulation rotation gesture was not only most tried as first attempt, but also the one recognised as standard. Surface gestures consisted mainly in a transposition of the rotation concept on screen: one or two fingers were used to draw a semi-circle shape. Clockwise and anti-clockwise gestures were grouped in the same one, under the assumption, mostly confirmed verbally by participants, that the gesture direction had a direct mapping on the rotation direction. Finally, the only motion gesture consisted in moving the device closer to Anubis.

It has to be noted that, while a bias mitigation was provided for the direct manipulation, since the target image was stitched to a rectangular box to improve its graspable affordance, no mitigation was implemented for motion gestures. That is, being the participant sat on the side of a square table, it was inconvenient for him to physically move the smartphone around the target object. For any further study it is suggested to use a round table, and to invite the participant to stand up.

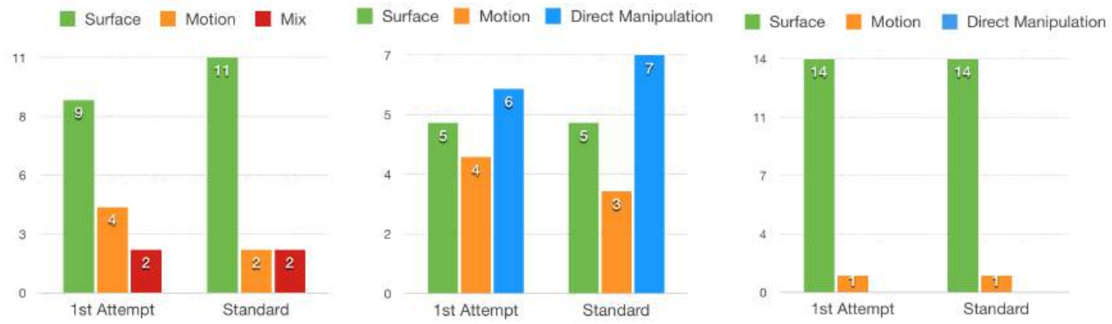
		1st	2nd	3rd	Std
Surface	One finger circle rotation	2			1
	Two fingers circle rotation	2	1	2	2
	One finger straight line	1			1
Motion	Move the phone around the Anubis's target object	4	1	1	
Direct Manipulation	Rotate the Anubis's target object	6	2		7

Table 3.3 Elicited gestures for rotation.

3.3.4 First attempt, Standard and Preferred Gestures

During the experiment an additional gesture classification emerged. When participants were given with the smartphone and asked to guess the right action to cause a specific effect, they performed his *first attempt gesture*, so the first gesture which came to their mind. However it frequently happened that participants did not stop trying, gradually discovering new gestures which, in some cases, were defined as more satisfactory than the first one, hence defined as *preferred gestures*. Finally, the *standard gesture* was the one that, according to participants opinion, most of people would have tried when asked to achieve a specific task.

For example, for the zoom task, several participants tried to move the smartphone close the Anubis as first attempt gesture, and they recognised it as their favourite; however, when they were asked about their opinion on which kind of gesture other people would have done, some of them chose the pinch to zoom surface gesture (table 3.2). Similarly, for the rotation task, it was noticed a difference between first attempt and standard gestures (figure 3.3b). Finally no difference between standard and first attempt gesture was declared for the selection task, which underlines the strong consensus toward the elicited gesture (figure 3.3c).



(a) *Zoom* task gestures. (b) *Rotation* task gestures. (c) *Selection* task gestures.

Fig. 3.3 Elicited gestures by category for each task.

3.3.5 Guessability and Agreement

Wobbrock defines the *Guessability* G as:

$$G = \frac{\sum_{s \varepsilon S} |P_s|}{|P|} * 100\%$$

where P is the set of proposed symbols for all referents, and P_s is the set of proposed symbols using symbol s, which is a member of the resultant symbol set S [91].

Moreover Wobbrock defines the *Agreement* A as:

$$A = \frac{\sum_{r \varepsilon R} \sum_{P_i \subseteq P_r} \left(\frac{|P_i|}{|P_r|}\right)}{|R|} * 100\%$$

where r is a referent in the set of all referents R, P_r is the set of proposals for referent r, and P_i is a subset of identical symbols from P_r [91].

So for example the A value for the selection task would be:

$$A = \left(\frac{8}{15}\right)^2 + \left(\frac{4}{15}\right)^2 + \left(\frac{1}{15}\right)^2 + \left(\frac{1}{15}\right)^2 + \left(\frac{1}{15}\right)^2$$

In figure 3.4 are showed *guessability* and *agreement* values for the most performed gesture of each task [91]. Guessability and agreement give as a score to define the consensus over the identified gestures. The selection gesture, a drag and drop one finger surface action from Anubis to the bag, is characterised by high guessability and agreement. The zoom gesture, the surface pinch to zoom, is characterised by a lower value of guessability, but it still is slightly over the 50% threshold; the agreement over the gesture, however, is only the 36.8%. Finally, the rotation gesture, a direct rotation manipulation on the target object, is characterised by a even less certain guessability and agreement, respectively 40% and 27%.

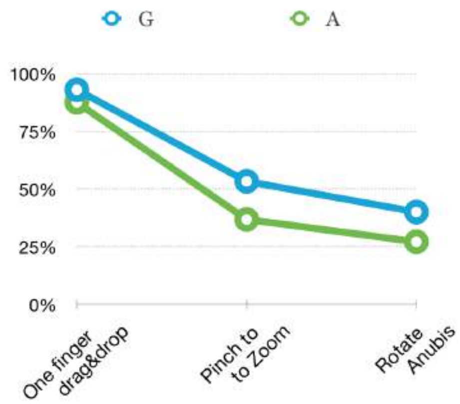


Fig. 3.4 Guessability and Agreement values for the elicited gestures.

3.4 Discussion

In the following sections quantitative results are discussed in the context of the qualitative findings gathered thorough the talk-aloud protocol and the final interview.

3.4.1 Previous Experience vs Tangible Interaction

For a group of participants, as already observed by Wobbrock [92], the interaction seemed to be strongly influenced by their previous experience. Therefore, because touch is the primary smartphone modality, they tried to apply common surface gestures, such as the drag and drop and the pinch to zoom, to the AR context. For example, on the selection task one participant stated:

> cit1. *"Yes, it* [thinking of a gesture to perform] *was easy but I always actually think of the previous experience* [...] *it's like simulating what we are doing when we use a computer when you put a file into a folder and you drag one into each other."*

In some cases participants expressed concerns towards direct manipulation approach speculating that the target object could be fixed somewhere or too heavy to be moved. It has to be noted that these ideas were not supported by the laboratory settings, in fact in the experiment the target object was light, easy manipulable and close to participants. However, these observations provide an additional explanation for the general preference of surface gestures over the others. Some people also stated they preferred surface gestures because, being the smartphone already in their hands, they were more immediate (assumption confirmed by Fitt's law).

Other participants, even if they agreed on identifying surface gestures as standard interaction, preferred motion and direct manipulation gestures. They justified the preference, explaining that, being AR a metaphor for reality, motion and direct manipulation approach seemed better suited to simulate a real interaction. One participant, for example, stated:

> cit2. *"I like the idea of augmented reality and I think that the interaction should be augmented based* [...] *interaction shouldn't be difficult because it is the same of the real world* [...] *I think this is the point of augmented reality."*

The contrast between the two approaches emerged in particular with the rotation task, where a standard gesture could not be identified. As a participant stated:

> cit3. *"The zoom and the selection, it's something we are used to every day, the zoom with the camera for example, but the rotation is not so common so maybe other people will try different gestures."*

3.4.2 Need of consistency and coherence

As soon as AR reality was accepted as equal to reality, participants asked for the same consistency and coherence they would expected in the real world. For example one participant tried to move Anubis around the table to see if the lighting over the object was consistent with the lighting of the room. The same participant also moved a pen first between the camera and the target object, then behind the target object to test the occlusion coherence. Another participant stated that he would have expected Anubis to fall down from the target object as if it was responding to the gravity force.

When interacting through standard surface gestures participants seem to suspend their critical judgment, but they look for coherence and consistency when they had to think of a custom gesture. For example, participants were not disturbed by the fact that pinch to zoom and drag and drop triggered a 3D effect even if they were 2D gestures. However, when they had to think of a surface gesture for the rotation task, most of participants underlined the incongruence.

3.4.3 Abstract and Concrete gestures

A further classification on gestures can be made according to their abstract or concrete nature. For example, shaking the phone to achieve selection is an abstract gesture as the relationship cause-effect is not clear whereas moving the phone around the object or rotating the target object itself, are a concrete gestures, in the sense they emulates the action we would perform in a real world.

3.4.4 Ambiguous Gesture and UI Interaction

Some elicited gestures caused a conflict. For example, the double tap was used both to achieve zoom but also selection. In general, it suggests that, when a gesture does not have a high agreement, it is better to provide users extra information, to solve the ambiguity.

During the free exploration task several participants pointed out that they would have expected other UI elements to appear to help them to interact with the scene. For example, for the zoom task often it was hypothesised a *plus* or *minus* button for the zoom-in and zoom-out while for the rotation task it was often suggested the visualisation of the dimensional axis for the manipulation of the object.

3.5 Implications and Conclusion

In this chapter findings from a guessability study, which aimed to elicit end-user gestures for selection, zoom and rotation tasks in AR environments with smartphones, were presented. No restriction was imposed in terms of category so that participants were free to perform surface, motion or direct manipulation gestures.

Results showed that, once a category was identified, the choice of a specific gesture was straightforward. In particular, for the selection task almost the totality of participants (93%) chose a surface drag and drop gesture; for the zoom task, even if the agreement was lower, more than a half (53%) chose the pinch to zoom surface gesture. Finally, for the rotation task, the general trend looked inverted and for the first time the highest agreement was towards a direct manipulation gesture, that is participants directly rotate the target object. However, the agreement for the gesture was the lowest for the three tasks, in fact only 40% participants tried the gesture in their first attempt.

According to Wobbrock's results [92] participants behaviour seemed in general highly influenced by their previous experience, especially where a common gesture already existed for the task. In those cases, the fact that they were interacting with an AR environment did not affect their capacity to think of a gesture for the task, which was always considered as natural and intuitive. However, when a standard gesture did non exist, as for the case of the rotation task, their behaviour was affected by their critical judgment. In these cases it emerged a need for coherence and consistency that was previously ignored.

There are several implications for these findings. Firstly, as the biggest indecision is on which gesture category the user will choose, rather than the specific gesture within a certain category, when developing an application we should put an effort in two different directions. On one hand attention should be paid on making the underrated categories more visible: that is, if users are not tempted by the direct manipulation because they are afraid to touch objects, then the target object affordance for manipulation should be improved. On the other hand, because for some tasks, such as the rotation one, agreement is divided between several categories, then the application should probably implement a gesture for each category. Finally, when the agreement is too low, additional UI elements should be provided to help users interacting with the application. In fact, as Wu et al. state, a system where the learning only refers to gestures, is a system hard to learn [94]. Participants' suggested ideas such as contextual menus, buttons for zooming and axis slider for rotation, are good example of how to achieve the goal.

The current research also confirmed that participants prefer a simple interaction, apart from one occurrence that happened during the pilot study, when a participant tried a surface grabbing gesture involving all five fingers, participants always interact with the screen using one or two fingers.

During the experiment also emerged that, when no common gesture exists for the required task, participants tend to use their judgment to figure out a new gesture. Implications are that we should always prefer simple commonly used gestures over complex new ones.

Finally, participants showed a need for consistency and coherence. Implication are that we should pay attention to the relationship between the AR and the real world rules. This problem is not new in HCI, and it usually appears when dealing with metaphors. For example, in a desktop environment, the basket is positioned above instead of under the desktop: breaking the metaphor is needed or the basket would not be invisible. However, the more the metaphor is close to reality, the more a breakage can confuse the user's mental model. As AR is a metaphor for reality, we should always be careful when breaking rules. Example of design choices that could break the metaphor are: implementing the zoom function in a way which enlarges only the models and not the background or implementing the rotation through surface gestures as we would allow a 2D gesture to cause a 3D effect. As in the desktop case, breaking the metaphor does not have to be considered wrong, but it has to be done consciously and the choice need to be justified.

Chapter 4

Mobile Egypt: Application Design

Having defined a set of standard gestures and having gathered some insights into the users' mental model, the application design can be defined. In the current chapter, application workflow and layout, usability aspects and target objects are presented. Details on the story game are reported in Appendix A.

4.1 Application Core Concept

The application starts introducing players into the story. That is, players have to look for five target objects corresponding to five virtual artefacts. When the correct target object is found, the virtual object to collect appears on top of it (figure 4.1a) and it is available for collection. Once the virtual object is into the bag, a window with a 3D model and information of the artefact appears (figure 4.9). When the target object found corresponds to an artefact that cannot be collected yet, 3D virtual text appears stating "Collect other artefacts first" instead of the virtual artefact (figure 4.1b).

(a) (b)

Fig. 4.1 Examples of AR content triggered by target objects.

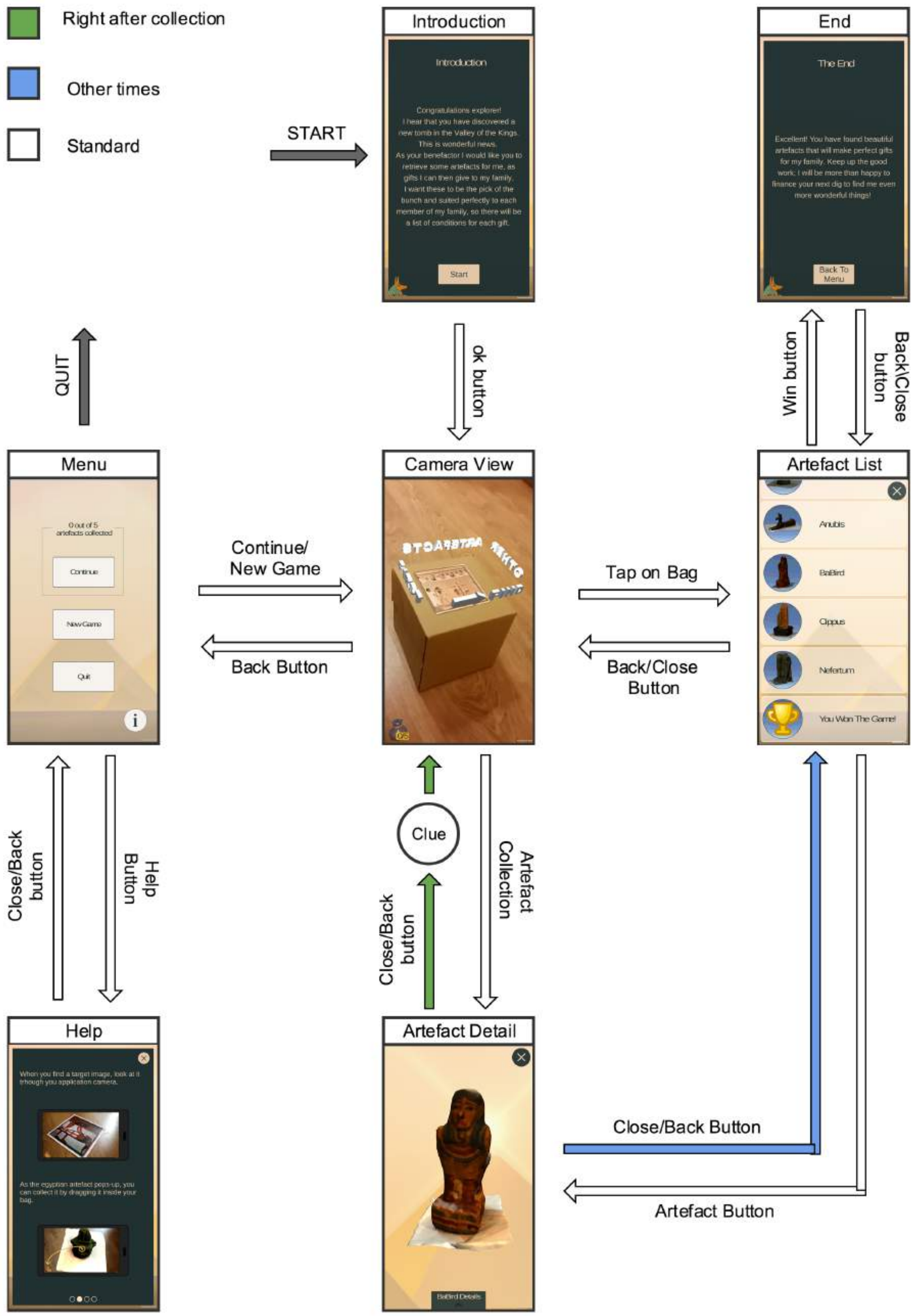


Fig. 4.2 Simplified workflow of the application.

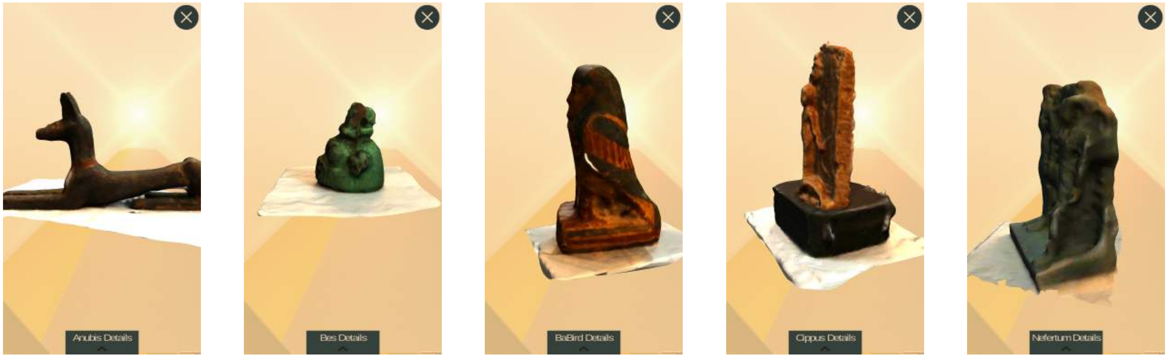


Fig. 4.3 The five game artefacts.

4.2 Main Functionalities

Main functionalities of the application include: scan of a set of target objects to show the virtual artefacts; collection of the virtual artefacts through the surface drag and drop gesture; manipulation of 3D artefact models through zoom, rotation and translation over the Y axis; implementation of the story game in the application.

4.3 Study Implication for Interaction

Three different gestures, determined from results of the guessability study, are supported for interacting with virtual artefacts: selection, zoom, and rotation. Selection is implemented using a surface drag and drop gesture as it emerged from the guessability study as a clear result with high agreement. The zoom gesture did not have such a clear result with surface gestures being the most common with motion gestures in second. When the AR object has been captured and the 3D model displayed, the pinch to zoom is supported. However, due to technological limitations, the gesture could not be implemented when capturing virtual objects. More specifically, when part of the target object leaves the field of view, the AR registration cannot function properly. The second most popular gesture, motion gestures, were fully supported. Finally, for rotation, direct manipulation was supported as suggested by the guessability study. Surface gestures were also supported in the 3D model view.

During the guessability study, it was found that two main factors were influencing the users interaction with AR: previous experience and environment. As users previous knowledge cannot be influenced, the current work focuses on the environment manipulation. In particular, three different strategies have been identified to encourage an active user interaction. First of all, the target object was developed to improve its affordance for manipulation by choosing a specific size and shape (details on target objects are reported in subsection 4.6). Secondly, different types of AR content were created to encourage exploration. This was realised through instructions, the 3D virtual text, 360 degrees around and just above the target object (figure 4.1b) which should encourage visitors to walk around it (motion gesture) or move it (direct manipulation gesture). Finally, tips were embedded tips into the UI. As the guessability study suggested that participants might expect widgets to support interaction, we used UI tips to suggest how to interact with a specific scene when displayed for the first time (figure 4.13). Tips are further described in subsection 4.4.7.

4.4 Application Workflow and User Interface

The final application is composed by five main sections: *menu*, *help*, *camera view finder*, *artefact list* and *artefact details*. *tips* and *clues* are considered as transition states. In figure 4.2 it is presented how these sections are related with each other by a simplified version of the application workflow. During the early stages of the development, paper prototyping was used to define workflow and user interface of the application. In particular Popapp[1], an android application, was used to create low fidelity dynamic prototypes (details on paper prototyping are reported in Appendix C). The development went through a series of iterations at the end of which the application was evaluated to adjust requirements. In the following sections it is presented how the application changed over time to adhere to the new requirements.

4.4.1 Layout

The layout of the application was developed using as baseline colours the ones of the Egypt centre logo, in order to improve the brand identity.

Fig. 4.4 Egypt Centre logo and application colours.

4.4.2 Menu

From the *menu* it is possible to access the *help* section, the *camera view finder* by starting a new game or continuing the current one, or to quit the game. An indicator showing the state

[1]https://popapp.in/

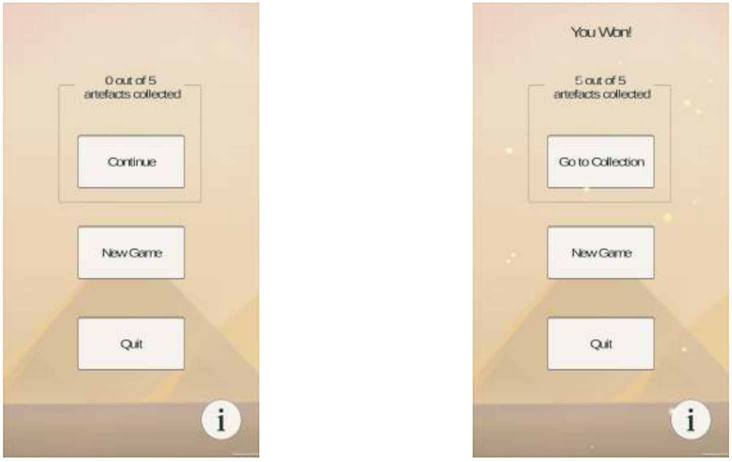


Fig. 4.5 From left to right: the menu at the beginning and at the end of the game.

of the game, that is the number of artefact already collected, and a win message in case of victory, is shown on the top. When the game is won, a particle effect animates the screen. Moreover, buttons content change dynamically according to the game state. For example, the button 'Continue' changes into 'Go To Collection' when the game is completed.

4.4.3 Help Section

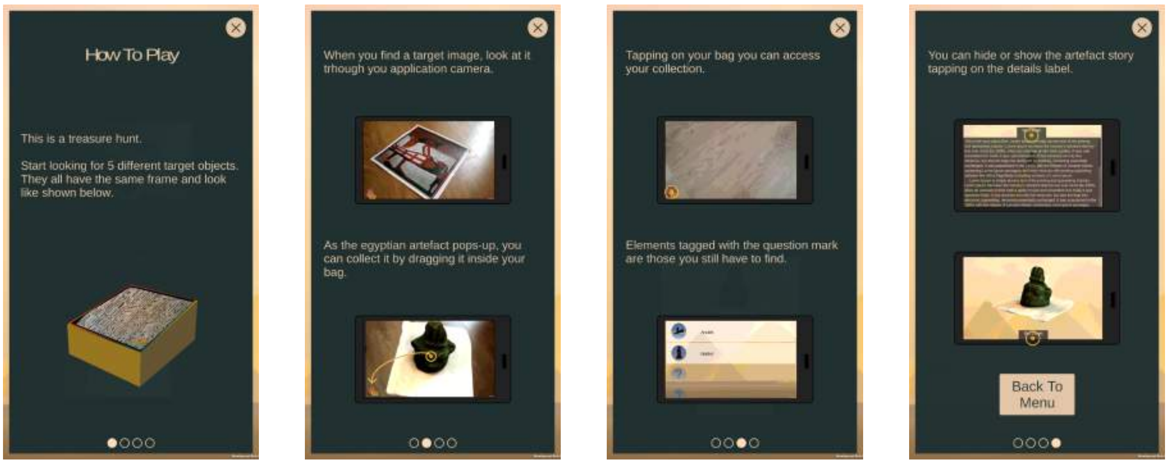


Fig. 4.6 The four cards of the Help section.

The application includes an *help* section to walk the user through the first use of the game. In the early versions of the application, this section also functioned as welcome screen, therefore it was the first section visited by the user. However, after performing usability testing, it appeared clear that the content of the introduction was too long, even when reduced to four screens. Users tended to skim reading the text, without really paying attention to it, so that the content delivery was not really effective. Therefore, *tips* were introduced to support users during the first use of the application (details on *tips* are reported in subsection 4.4.7). Being the *help* section more exhaustive, it was kept in the final version of the application, where it is accessible by the main menu through an info button (figure 4.5). The *help* section is implemented as a series of four cards, browsable with swipe gestures. An indicator on the bottom shows the number of the current card, and from anywhere it is possible to close the section by tapping the close button on the top corner. Each card is composed by screenshots of the application integrated with explicative text.

4.4.4 Camera View Finder

Fig. 4.7 Camera view finder.

The *camera view finder* is a crucial section of the application, allowing the interaction with virtual objects. From this section it is possible to accede the *artefact list* by tapping on the bag, or the *artefact detail* when a new artefact is collected. The *camera view finder* is the only section which is not equipped with a close button, however it is possible to go back to the *menu* using the integrated back button of the smartphone. Because the *camera view finder* works as window to interact with AR, it was considered best not to include unessential elements, in order to leave the whole screen space available for interaction.

Interaction with Virtual Objects

Once a virtual object is located, it can be explored through motion gestures, by moving the phone closer and around the target object, or through direct manipulation gestures, by moving the target object. These functionalities are implemented in the *Vuforia* library by default (see Appendix B for details). The surface selection gesture was instead implemented from scratch. The gesture lets the user to select the virtual object only to drag it inside the bag. If the object is not collected, it goes back to its original position. This strategy allowed to overcome the technological limitations introduced in section 4.3. Implementation details are reported in Appendix B.

The Treasure Hunt Bag

The treasure hunt bag provides access to the collected artefacts. It also shows the current state of the game by displaying the number of artefacts already collected through the associated dynamic label. The bag also provides interactive feedback. In particular, when it is touched by a single tap, it responds with an elastic stretching movement and the background is highlighted. Moreover, when an uncollected artefact is dragged inside the bag, it responds by spinning around and launching a particle effect, to notify that the artefact has been collected. Implementation details are reported in Appendix B.

4.4.5 Artefact List

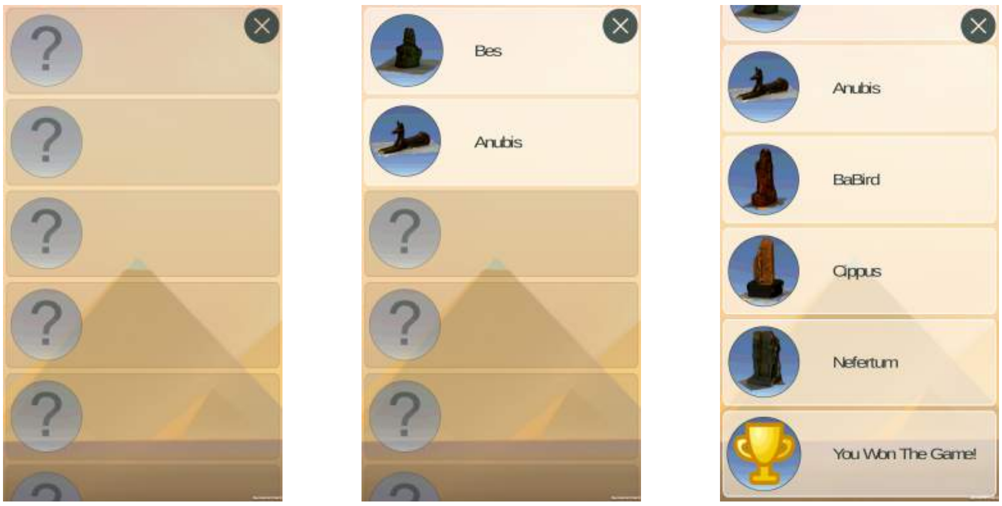


Fig. 4.8 From left to right: the artefact list at the beginning of the game, during the game and when the game is completed.

From the *artefact list* it is possible to go back to the *camera view finder*, by tapping on the close button, or to see the detail of a collected artefact, by tapping on the corresponding artefact button. A total of six elements are presented in the list, where the first five correspond to the five artefacts to collect, while the last one is a win button. The win button appears when the game is completed and takes the player to the final screen. The *artefact list* changes dynamically according to the game state. At the beginning of the game, all elements are characterised by a semi transparent colour and a question mark is displayed for each element. As the game proceeds, buttons corresponding to collected artefacts are filled with artefacts' images and names.

4.4.6 Artefact Detail

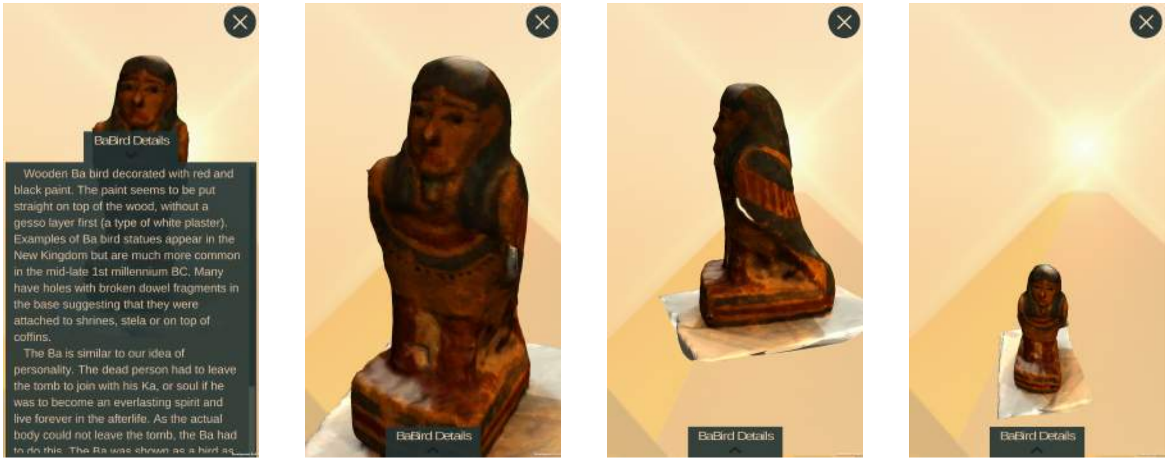


Fig. 4.9 Artefact detail.

The *artefact detail* section is and end point. From this section it is only possible to go back to the *artefact list* or to the *camera view finder*, respectively if an artefact has just been collected or not.

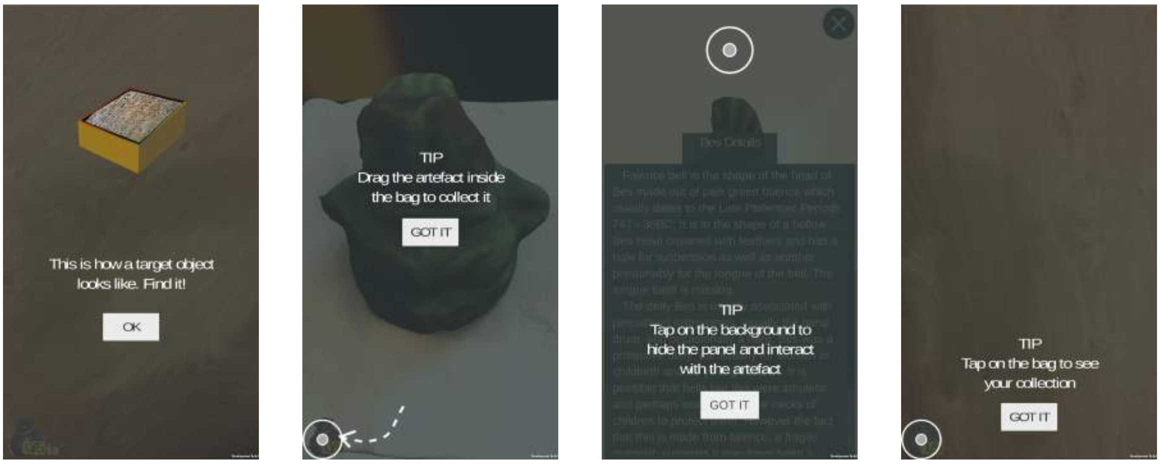
Artefact Informative Text Panel

During the first stage of the development this section was designed as divided in two areas, one containing the 3D artefact model, and the other one the informative text (see Appendix C for details). However, given the small screen dimension, the design was later changed to give the user a bigger interaction area. In the final design the informative text is wrapped inside a panel that can be hidden. When the artefact is collected and this section is visited for the first time, the panel appears open. On the contrary, when the section is accessed later on from the *artefact list*, the panel is hidden, showing the 3D model. During usability testing emerged that, when the screen was accessed and the panel was open, very few users realised that the panel could be closed to interact with the 3D model (which was interpreted as a background image). Therefore, two different strategies were adopted to improve the discoverability of the 3D model. Firstly, a tip indicating how to hide the panel is showed when the section is acceded the first time (figure 4.10c). Secondly, while in the first version of the application to hide the panel it was necessary to tap on the panel label, in the final version it is sufficient to tap on the screen background, making the active surface wider.

Artefact 3D Model

Interacting with the 3D model is possible through surface gestures. In particular the model can be rotated, by horizontal swapping, translated over the Y axis, by vertical swapping, and finally be enlarged by pinching to zoom. Details on the implementation of surface gestures are reported in Appendix B.

4.4.7 Tips



(a) ‘This is how a target object looks like. Find it!’ (b) ‘Drag the artefact inside the bag to collect it.’ (c) ‘Tap on the background to hide the panel.’ (d) ‘Tap on the bag to see your collection.’

Fig. 4.10 Tips.

Tips were introduced to provide a more flexible and contextual solution to support users interaction and to substitute the *help* section. Tips were developed as short messages on the top of a semitransparent layer and were displayed only the first time that users acceded a specific scene. They aimed to inform users of application features that would have been otherwise hardly discoverable. For example, during usability testing emerged that when reading the artefact informative text, users rarely tried to close the panel to interact with the 3D model. Therefore, a tip to inform users about this functionality was added. Other tips were used to show a target object looked like (figure 4.10a), how to access the collected artefacts (figure 4.10d) and how to collected an artefact (figure 4.10b) Once a tip was dismissed, it never appeared again even if a new game was played again, under the assumption that the user already learnt how to interact with the application.

4.4.8 Pop-Up messages

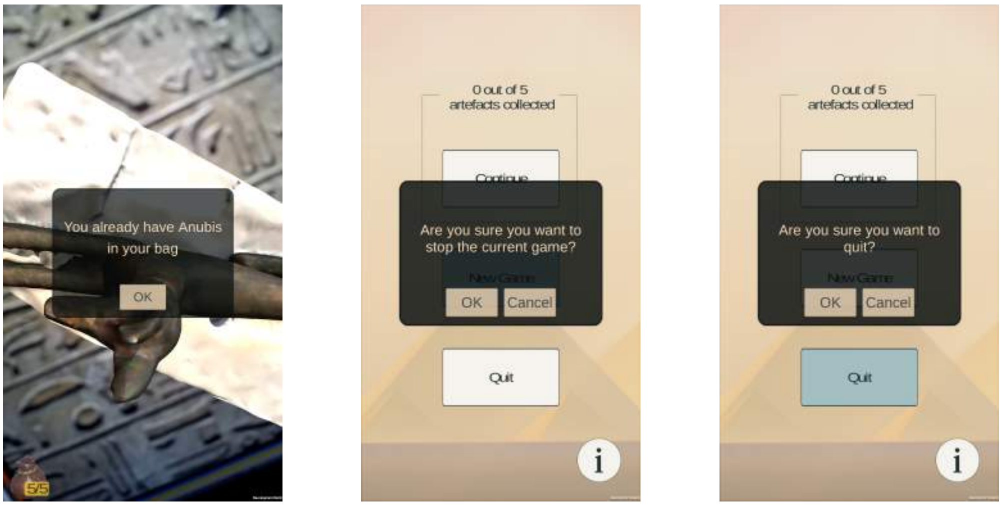


Fig. 4.11 Pop-up messages.

Pop-up messages were used for confirmation purposes or to inform users of particular circumstances. In the first case, the aim of the pop-up was to prevent error and to make an action reversible. For example, because when quitting or starting a new game, all the progress were lost, a confirmation message was needed. A pop-up message was also shown when the player tried to collect again an artefact that was already collected, making clear to users why the action was not allowed.

4.4.9 Clues

Clues are cards showed only one time between artefacts collection. Introduction and end screen are considered special cases of clues. Clues contain informative text useful to find the next object to collected and once dismissed they cannot be longer accessed. Clues take the navigation workflow to the *camera view finder*.

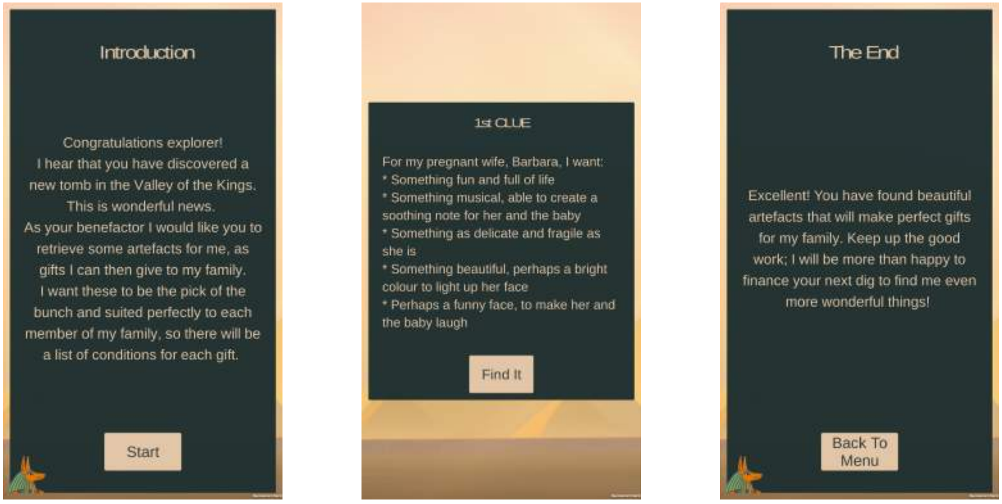


Fig. 4.12 Clues.

4.5 Usability Aspects

The game was developed with respect of usability guidelines. In particular, in the following sections, it is reported how the application implemented all the ten usability heuristics defined by Nielsen [53].

Visibility of system status

> *"The system should always keep users informed about what is going on, through appropriate feedback within reasonable time."*

Indicators are used in the application to inform users on the current state of the game. Examples are the label on the treasure hunt bag and the meta information showed in the menu, indicating how many artefacts have been already collected. Other examples of indicators are the circles on the bottom of the each card of the *help* section, which shows the user how many cards are present in total, and which one is the current one, or the arrow in the informative text panel of *artefact detail* section, which indicates if the panel is hidden or not. Moreover, the different layout used to represent elements of the *artefact list* helps users to identify which artefacts have been already collected and which are still to collect. Finally, feedback is always properly implemented to help users seeing the results of their interaction. Example of feedback are the change in aspect of buttons when pressed and the particles and elastic effects of the bag when pressed or selected after a collection.

A critical aspect emerged during the development of the application due to performance limitations. While, overall, the game always responded well and feedback usually appeared instantly, the access to the *camera view finder* after each collection was always slow. Therefore, a loading screen before accessing the section was added, to avoid overrun errors [58]. That is, to prevent users to press buttons more than once because of the lag in the section loading.

Match between system and the real world

> *"The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order."*

The application layout follows real world conventions. For example, showing question marks in the *artefact list* to identify which artefacts still have to be collected, makes use of well known and familiar concepts to inform users on the state of the game. The elastic feedback on the treasure hunt bag, to suggest that the bag has been filled with a new object, is another example of how real world knowledge is used to help the user understanding. When not familiar concepts are introduced, such as for the case of the target object, tips are always provided to help the user understanding. Words in tips and informative text are always carefully chosen to be as much familiar as possible. For example, the term virtual object is never mentioned and the word artefact is used instead.

User control and freedom

> *"Users often choose system functions by mistake and will need a clearly marked 'emergency exit' to leave the unwanted state without having to go through an extended dialogue. Support undo and redo."*

Users are usually allowed to revert their last action and to go back to a previous screen by tapping the close button, which is present in every screen (except for for the *camera view finder* for the reasons explain in subsection 4.7) or by pressing the back button of the smartphone.

Consistency and standards

> *"Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions."*

The application layout has been implemented in order to facilitate the system learning. Example of implementations which follow standards are the back button to go back to the previous screen, the close button to close the current screen, a scrollable list with clickable elements to select artefacts, a *menu* to manage the game and an info button to reach the *help* section. Also surface gestures such as the swipe to navigate between cards of the *help* section, or to interact with the 3D model, follow standard conventions.

When the application breaks consistency, the choice is always aware and justified. For example, the close button on the *artefact detail* section causes different consequences according to the different state of the system. If the user has just collected an artefact, the navigation goes back to the *camera view finder*, otherwise, it goes back to the *artefact list*. The choice is justified by the fact that, with the workflow implemented in this way, the number 'click' for task is optimised. In fact, it is reasonable to think that after a collection the user wants to go back to the *camera view finder* for the next artefact to collect rather than on the *artefact list*.

Error prevention

> *"Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action."*

A custom screen, to notify users that the application is loading, was implemented to prevent the user to tap multiple times on the screen causing unwanted effect (overrun errors are already discussed in paragraph 4.5). Pop-up messages for confirmation are also used to prevent users to perform unwanted actions such as losing game progresses because of unintentionally quitting or resetting the game. Moreover, buttons in confirmation pop-up are placed to minimise the risk of pressing the wrong button, that is confirmation button on the left side, and cancellation button on the right side.

Recognition rather than recall

> *"Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate."*

The *artefact list* is implemented as a list where for each element it is given name and image of the artefact. The use of an image helps users to identify the correct artefact, without having to rely only on the name. The application workflow and layout is made essential and important elements, such as the close button, are always visible.

Flexibility and efficiency of use

> *"Accelerators - unseen by the novice user - may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions."*

The go back to a previous section, users can choose to press the smartphone back button or the application UI close button. Moreover, both landscape and portrait modalities are supported so that users are allowed to use the the modality which suits the best for them or for a specific task. Finally, allowing to close the informative text panel of the *artefact details* section by tapping on the panel label or on the screen background, is another example of how flexibility and customisation are achieved.

Aesthetic and minimalist design

> *"Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility."*

The application layout is essential and only key elements are embedded in the UI.

Help users recognise, diagnose, and recover from errors

> *"Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution."*

Informative messages are shown to users when not allowed actions are tried. For example when trying to collect an artefact already collected, an informative pop-up explaining why the action is not possible appears.

Help and documentation

> *"Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large."*

Help and documentation are implemented through the *help* section and the use of tips.

4.6 Target Objects

A set of five different target objects was developed for the game. Target objects were crated using cardboard boxes, measuring 14x14x7cm, and applying on the top target images. That are fiducial markers that once captured by the smartphone camera trigger the AR layer. Target images were created by framing images related to the Egypt history. The frame was developed to recall the Egypt centre logo, therefore to support the brand identity. Using the same frame for each image also provided to target objects an unique characteristic that improved their recognisability during the treasure hunt. Details of how target images are recognised by the application are reported in Appendix B.

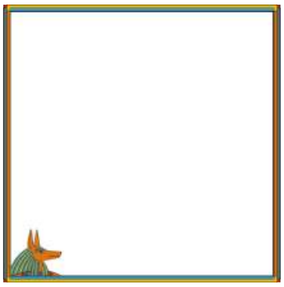


Fig. 4.13 From left to right: the empty frame with the Egypt centre logo, a target image created by adding an Egypt image and the final target object.

Chapter 5

Evaluation in the field: a qualitative study

The purpose of the qualitative study was to investigate user interaction with AR using smartphones in a real museum environment. The study was designed to: 1) verify that results elicited from the guessability study hold in a real environments 2) further understand the participants' mental model; and 3) investigate the possible applications of AR in museums.

Results from the qualitative study were included in a paper submitted for the CHI 2016 conference[1].

5.1 Study Design

5.1.1 Participants

The study involved 15 volunteers employed at an Egyptian museum, 46% of the participants were female and 54% were male. The volunteers age ranged from 18 to 65. Most volunteers (81%) self reported as normal to expert smartphone users. Only two volunteers had experience with AR prior to the study. One had used a Google glass application with voice commands and the other had used a Nintendo portable console.

5.1.2 Apparatus

Mobile Egypt, the application described in Chapter 4, was used for the study. The application was deployed on the same device used for the guessability study. Additionally five target objects were created, by applying five different target images to the top of boxes measuring

[1] http://chi2016.acm.org/wp/.

14x14x7cm. Each target object displayed a unique virtual artefact that formed part of the story.

5.1.3 Procedure

Volunteers were recruited concurrently from the same room where the study took place or from other areas of the museum. After signing consent forms and being explained the purpose of the study, participants completed a questionnaire to gauge computer, smartphone, and AR experience. Participants were then given the device with the application already running in the portrait mode and were invited to play the game. No specific task was assigned to participants and they were asked to use the application to try to complete the game. While participants were using the application, the researcher followed them around the room and acted as an external observer. The researcher only intervened when participants explicitly asked for help. After the game was completed, a debriefing interview took place. Participants took 25 to 45 minutes to complete the study. As participants used the supplied smartphone, all screen interaction and verbal feedback were recorded using a preinstalled application. A camera partially filmed the session and also captured participants' interaction with the environment.

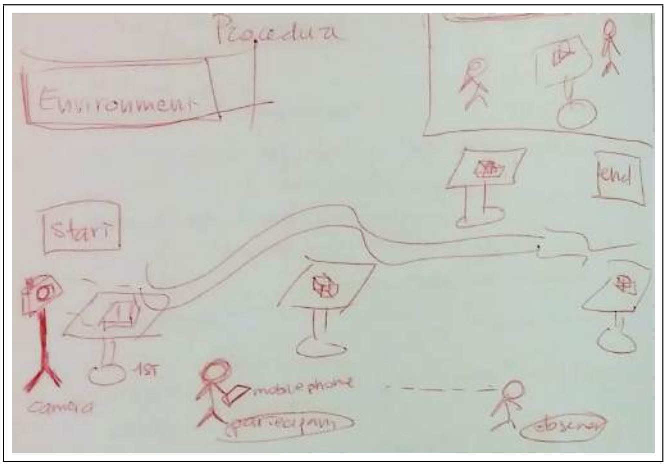


Fig. 5.1 A sketch of the study procedure.

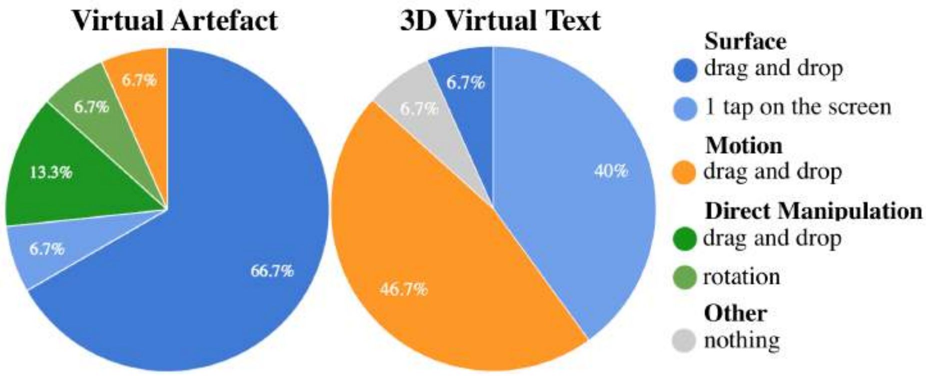


Fig. 5.2 First approach to the AR layer.

5.2 Results

5.2.1 Interaction Over Time

On average, participants took 6.45 minutes to complete the game. With few exceptions, most of the time was spent wandering around the museum and looking for the right target object to scan, rather than interacting with virtual artefacts, 3D artefacts models or reading text.

5.2.2 Elicited Interaction

During the use of the application only a new gesture was observed. A participant tried a two finger pinch gesture (a thumb-first finger grabbing gesture) to collect the virtual artefact, a gesture similar to one already observed in the guessability pilot study (but performed with five fingers). Therefore, elicited gestures can be still classified as surface, motion, or direct manipulation. The first approach with the AR layer depended on the virtual object they were interacting with (figure 5.2). When the 3D text was found first, the reaction was equally divided between motion (walking around the text trying to read it) and surface gestures (mainly a single tap on the screen looking for hotspots). Only one participant observed the scene without trying to interact with the target object in any way. When the virtual artefact was found first, the most common form of interaction was through surface gestures (73.3%), while a minor number of participants tried direct manipulation gestures (20%). Only one participant used a motion gesture. Moreover, while for the virtual artefact a message indicated how to collect it (figure 4.13), the 3D text was not accompanied with any such tip (figure 4.1b).

Over time, the approach for interaction changed. For example, participants who first tried surface gestures for a specific task, later discovered they could interact though motion gestures. Surface and motion gestures were the most common approach and only two participants tried direct manipulation. During the debriefing interview the researcher demonstrated how the virtual artefact could have been manipulated, by moving the target object, to participants who did not try direct manipulations. Reactions to this new way of interaction was extremely positive. Most participants laughed and started playing with the game again as it was a completely new experience. Example of participants statements were:

> cit1. *"Ooh I see* [laughing] *my gosh that's clever* [laughing] *how do you do that, that's really cool!"*

> cit2. *"I didn't realise the boxes would be out where you can do that* [...] *that's a really cool I am having fun just playing again now, yeah that's fantastic!"*

> cit3. *"I see! I see! I assumed because I was in a museum that I wasn't allow to touch* [...] *That's very good actually!"*

From the second and third statements also emerged an aspect often underlined by participants. Because in museums interaction is usually forbidden, participants were afraid to touch the target objects ad they thought they were not allowed.

On the effect of previous experience with AR, one participant stated:

> cit4. *"From experience I've had with augmented reality in the past everything stays in the place where it is meant to be and the actual movement is limited just to the screen, to the digital side rather than to the physical side."*

Finally, as the game was a treasure hunt, some participants stated that their attention was focused on finding and collecting virtual objects rather than on exploring them.

Overall, the participants reactions to direct manipulation were consistent with results of the guessability study. That is, there was a difference between what participants tried as a first gesture, what they recognised as standard, and what they considered as the best form of interaction. In particular, all participants recognised surface gestures as the standard way to interact with smartphones but preferred direct manipulation and motion gestures for virtual objects.

5.2.3 Difficulties

While interaction with virtual artefacts was generally straightforward, a few difficulties emerged during the use of the application.

One of the first problems involved target objects. Given the smartphone with the application running, participants usually neither understood what they had to look for nor how to

Fig. 5.3 One of the volunteer playing the game at the Egypt centre.

trigger the AR layer. In these cases, participants always asked for the researcher help. This issue only appeared the first time participants were required to interact with the target object. Once the first virtual artefact was acquired, the remainder of the game proceeded fluently.

Other problems were related to the AR content. Sometimes happened that, depending on the distance from which the target object was scanned, the virtual artefact appeared so large that participants had to step back to see it properly on the screen. In regards of the virtual 3D text, even if the 360 degree of text encouraged the discovery of motion gestures (46% of participants walked around the text in order to read it), participants found it annoying. In most cases, they found the text hard to read and they did not understand the purpose of walking around it. To compound matters, as a few of the participants were not able to read the text properly and did not understand the message. For example, some thought that the 3D text was the object that they had to find, while others that the arrow that indicated the reading direction was pointing in the direction of the next collectable artefact.

5.2.4 Mobile Interaction

Participants always used the application by holding the smartphone in portrait mode even if the application was developed to support the landscape too. In most cases, the drag and drop gesture was performed with the right hand thumb, holding the smartphone with only one hand. Even if the smartphone was not small, participants never complained or seemed to struggle. The screen seems also big enough to explore the scene. Only in one case a participant mentioned the possibility of exporting the acquired information to another device such as tablets or PCs, to explore the 3D model at a higher resolution.

5.2.5 Technical Limitations

Due to low light in the room and reflections, on a few occasions the camera was not able to recognise the target object to trigger the AR layer or the virtual object was not properly aligned. However, participants always responded well, immediately understanding the nature of the problem and walking around the target object to scan it from a different angle. When mentioning this in the debriefing interview, participants never stated to be annoyed.

5.3 Analysis

5.3.1 Accessibility

Five different ways in which the developed AR game improved accessibility were identified: interaction, observation, information depth and information retrieval.

Interaction. As virtual objects are outside cases, visitors can finally manipulate them. In the developed application a two ways interaction is implemented as surface drag and drop gesture, motion of the smartphone and direct manipulation of the target object.

Observation. Motion and direct manipulation gestures allowed participants to observe virtual objects from any angle. This feature is particularly important in museums where artefacts are often located in cases or next to walls and cannot be completely explored.

Information depth. AR potentially allows layers of multimodal information to be displayed. One participant suggested two new functionalities: seeing the original state of the artefact or its original setting. Visitors could explore this information at the desired depth. Additional layers, such as audio, can be also easily integrated.

Information retrieval. By scanning target images, objects, or QR codes, contextual information can be displayed to visitors. This simplifies the task of finding information for themselves in a flexible and customisable way.

5.3.2 Conveying Knowledge and Support Learning

Two different approaches to convey knowledge by the game were used: game clues and artefact details. Clues are bits of text information that help players to locate the right target object. Because the artefact collection has to follow the story game order, clues are important to players and they are likely to read them. Once a virtual object is collected, players can interact with the 3D model of the artefact or reading the story of the artefact. However, this information, is not crucial for the story development. In line with our expectations, participants read carefully clues, but only few of them spent a long time reading the artefact story. Few participants mentioned that they were mainly paying attention to find and collect virtual artefacts rather than learning about them. Therefore, it is difficult to infer if the game actually improved participants learning experience. However, in general, participant agreed that using the application actually taught them something about the Egypt culture. The personal involvement in the story, the fact that they had to think of virtual artefact as gift to offer to their family, helped them to memorise at least names and basic information of each artefact.

5.3.3 A Treasure Hunt to Support Navigation

Because participants had to walk around the museum to find the target objects, several ways in which the application can be used to support navigation tasks were identified. By locating target objects in specific places, many virtual journeys to explore the museum can be created. In this sense, the application can help visitors in exploring and orienting themselves around the museum. The application could also lead people to explore different areas of the museum. Along these lines, a participant stated:

> cit5. *"Normally I'd avoid all the pottery section but if there was a box there I would obviously go there and then look at other objects as well."*

Therefore, the application has the potential to change the way people visit and experience museums. Finally, a few participants found it useful to select and view important pieces of a collection, especially when they did not have enough time to visit the entire museum.

5.3.4 Independency and Freedom

Compared to the other approaches traditionally adopted by museums, participants appreciated the higher independence and freedom offered by the smartphone application. For example, staff are occasionally unavailable and it is difficult to receive assistance. On the other hand, even when staff are available, there are many reasons for more independence and autonomy. Visitors can be shy and may not want to interact with other people. They also might want to visit a museum at their own pace. As one participant stated:

> cit6. *"I like the fact that once you have the app you can just you don't need to be led or supervised* [...] *I think for children it is good because they can go and explore on their own, they don't need adults with them, to tell them what to do."*

Some participants stated that they also liked the fact that having the application on their smartphone, they could just stop it and resume it any time they liked.

Enhancing the Visitor User Experience

Because smartphones are familiar to people, they were appreciated more when compared to other supporting tools such as books or information written on labels. The game dimension also allowed to engage visitors more by involving them in an immersive game experience. A few participants suggested to add a narrative voice to walk the player through the game, creating an active conversation. Moreover, enhancing the game dimension with AR, increased personal involvement. As a participant stated:

> cit7. *"I think* [visiting the museum using the application] *would be a lot more enjoyable especially for kids as they no longer have to go to the cases, they can walk and find things and it may feel like it is their own things that they are finding."*

Using a technology that was unfamiliar to most participants proved to be cognitively challenging and therefore stimulating. Several participants tried to manipulate the scene in an attempt to understand how the technology worked, and they were generally very fascinated by it. Surprise and curiosity were the main feelings expressed when the virtual object first appeared and also when they were shown how to perform manipulation gestures. Typical participants reaction were:

> cit8. *"It was fantastic the way it pops up is actually quite cool to be able to look around the way you would be if you had a real object in front of you, it makes you really wanna be more in the game."*

> cit9. *"It was quite good, even if I was expecting something like that I got surprised when the virtual object actually popped out."*

> cit10. *"I thought, how on the earth are they doing that? How is that done? I was wondering if there was another devices somewhere near the box. It's very clever."*

After the game completion, participants were asked to give a definition for AR. Most of participants stated that the technology augmented their senses in some way. Example of definition were:

> cit11. *"Adding things to your perception on the environment that otherwise would not be there through technology."*

> cit12. *"The app made things appear that were not actually there, so you could say some sort of magic."*

5.4 Discussion and Conclusion

This chapter presented findings from a qualitative study run using *Mobile Egype*, a custom smartphone AR application, with volunteers at the Egypt centre. The study aimed to verify that the insights elicited from the guessability study hold in a real environment, and to further investigate user interaction in AR as well as possible applications for AR in museums.

Results showed that the gestures support for AR implemented in the game (surface gestures for selection, motion and direct manipulation gestures for rotation and zoom) was successful. All participants were able to finish the game and, with rare exceptions, no one struggled or complained when trying to interact with the AR layer. Surface and motion were the most tried gestures. However, once direct manipulation gestures were discovered, they were the most appreciated ones.

During the study, three main aspects affecting users interaction, emerged: environment settings, previous experience and affordance. Understanding these aspects is crucial to the development of strategies for encouraging more concrete ways of interaction.

The environment setting affected the interaction limiting the kind of gestures that participants were able to perform. This was particularly true for motion and direct manipulation. For example, reading the 3D text by walking around a target object was sometimes made difficult due to physical constraints. Moreover, because in museums interaction is usually forbidden, participants were not sure if they were allowed to touch the target objects.

The previous experience affected users interaction in two different ways. On one hand, because the standard method to interact with smartphones is through touch screens, it gives a reason to explain the participants preference towards surface gestures. On the other hand, because AR is traditionally used for displaying virtual objects rather than interacting with them, it explains why participants with previous experience with AR did not try direct manipulation gestures.

Other two important aspects affecting user interaction are AR content and the target object affordance. The first time that the AR layer was triggered, reactions changed depending on the type of the displayed virtual object. For virtual artefacts, participants mainly tried to interact touching the device screen (surface gestures). For the 3D virtual text, they mostly walked around the target object trying to read the text (motion gesture). Similarly, when a virtual object appeared too big to be displayed in the device screen, it caused participants to take a step back from the target object, discovering motion gestures. Therefore, the idea of using AR content to encourage motion and direct manipulation gestures, can be considered a partial success. In fact, even if the strategy worked, in both cases participants found actions annoying. For example they did not understand why the reading task was made so difficult by displaying the text over a 360 degree surface. The relationship between AR content and users interaction is worth investigating. Strategies to develop AR content to suggest a concrete interaction, without annoying users, are possible. A solution could be, for example, to embed information in a specific spot of the virtual object, to encourage exploration. That could be, in the case of the application game, to ask players to find the precious stone hidden behind the head of a virtual artefact.

Affordance affected user interaction also in regards to the target object. Using a squared box to improve affordance for manipulation partially worked. In fact, two participants grabbed the target object while trying to achieve the selection task on their first approach with AR. Therefore, it is worth investigating other strategies such as trying different shapes, sizes and materials. A possibility could be to include written instructions on the target object itself or in the game, to help participants discovering direct manipulation. Moreover, instructions could also be applied in the real environment, creating an additional connection between the real and the virtual world.

In general, the use of UI tips proved to be essential and efficient in supporting user interaction. As Wu et al. [94] state, a system where the learning only refers to gestures, is a system hard to learn. However, while participants went through the use of the game, they started dismissing tips without even reading them. Therefore it is needed to find the right balance between providing the right amount of information without overwhelming the user.

Overall the application was positively appreciated by participants, giving evidence that it can change how people approach museums by offering and supporting an engaging user experience.

Chapter 6

Conclusion

AR is a promising technology that allows the display of virtual objects and it can be successfully applied for different applications. The interest for developing an AR application for museums is motivated by the belief that AR is well suited to overcome many museum limitations such as low interactivity (museums store delicate and valuable artefact that traditionally cannot be touched) and spatial and temporal constraints (museums do not often have enough space to show all pieces of a collection, certain pieces of a collections because of their size cannot be observed from different angles or in details, due to their size need to suggest visitors a way to be explored). Therefore, an AR application for smartphones was identified as a better suited alternative to traditional tools such as books, audioguide or staff support.

However, before thinking about how AR can be successfully applied, as it is a new technology for people, we needed to understand how people are going to interact with it. Therefore, it was run a guessability study to elicit end user gestures for basic interaction tasks such as selection, zoom and rotation and to gather insights on user mental models in AR. In a guessability study, participants are shown or described an effect and they are asked to perform the gesture they think is the right one to cause the effect. Main results from the study were that participants' interaction was mainly affected by their previous experience and by the environment settings. Participants used standard gestures when a standard gesture for the task existed. That is, they usually interacted by performing surface gestures as touch screens are still the most common way to interact with smartphones. However, when a standard gesture from their previous experience could be not identified, such as for rotation, new approaches were taken: for the rotation task most of participants tried in fact a direct manipulation gesture. The other aspect which highly affected the user interaction was the environment. For example participants declared that sometimes they did not try motion gestures because they were constraints by the physical limitations of the environment. That is, sitting in front a table with the target object in front of them, it was impossible to move the smartphone around the target object and being still able to look at the device screen.

Having defined a set of standard gestures and having gathered insights on users' mental models in AR, *Mobile Egypt* was developed, a fully functional AR game application for smartphones. The application, structured as a treasure hunt game, required players to wander around a museum to find five target objects corresponding to five virtual artefacts from the ancient Egypt. The collection order was tight to the story game, so that participants needed to actually read the text (trying to ensure learning) in order to find the next artefact to collect. Insights from the guessability study were used to define and implement a set of design guidelines to support user interaction. In particular, three main intervention areas to improve a chance for active interaction were identified. Firstly, tips embedded in the UI were used to suggest users how to interact with certain scenes when first presented (figure 4.13). Secondly, target objects were designed to improve affordance for manipulation, by creating them as rectangular boxes. Finally, the AR content was designed to help the discover of new ways of interaction. In particular when a target object not corresponding to the next object to collect was found, a 3D virtual text "Find other artefacts first" appeared. Being the 3D text disposed on a 360 degrees surface, users had to walk around the target object to read the text, pushing them to find motion gestures.

A qualitative study was then performed, using *Mobile Egypt* in a real environment, the Egypt centre, involving 15 volunteers. Aims of the study were verifying that the findings from the first study hold in a real environment and investigating possible applications of AR in museums. This second study consisted in passively observing participants while using the application, and in a debriefing interview to discuss the experience had while using the application. Overall, results from this study were consistent with what discovered during the guessability study: once again it was found previous experience and environment to be the two major causes affecting user interaction. No new gesture emerged during the game and a preference over surface gestures was found as well. Only two participants tried direct manipulation gestures. However, during the final interview, when participants were shown the possibility of manipulating the target object, they responded very positively engaging the new type interaction and starting to play with the game as it was the first time. While the interaction with the AR layer was mainly natural and intuitive, most of participants struggled to understand the first part of the game. That is, looking for target objects and scanning them for the first time. In most cases, they asked for more clear instructions embedded in the game to understand what they had to do.

Results from the two studies bring us to some conclusions. In regards to the previous experience, we have to accept that the user interaction will be highly affected by the participants previous experience with the devices, that is surface gestures in case of smartphones. Therefore, support for standard gestures must be provided. However, during the study, it was found that there was no correspondence between what was considered standard and what was identified as preferred and more engaging. This opens a challenge, how can we help users to find new possible ways of interaction?

The developed guidelines proved to be only partially successfully. Firstly, participants asked for more clear tips embedded in the UI to achieve difficult tasks. However, it was found that, while some tips were read and understood (such as the tip suggesting how to drag and drop the artefact inside the bag), others that occurred later in the game were dismissed before even been read. So the challenge of how support the user through difficult tasks providing the right amount of information remains open. Secondly, even if designing AR content to help discover new ways of interaction worked (most of participants tried a motion gesture when the 3D virtual text was showed the first time), it was judged as annoying. Participants did not understand why they had to walk around the target object to read the text and they found the task uselessly challenging. Therefore, we need to find alternative solutions to help users discovering new ways of interaction avoiding to annoy them. A possible solution could be making the task more meaningful: for example a clue could suggest how look at a specific side of an artefact to discover something about it. Finally, even if an effort was put to try to improve affordance for manipulation for the target object, the effort proved not to be effective enough. An explanation for the lack of direct manipulation attempts can be given by previous experience and environment. On one hand, participants who had already used AR declared AR as something that can be only observed but not manipulated. On the other hand, because museums are place characterised by a low interaction, where touching objects is usually forbidden, participants declared to be afraid of touching, being unsure they were allowed to. Possible solutions go from further acting on the target object and trying different shapes or materials to embed instructions on the UI, on the target object itself and in the real museum environment.

Chapter 7

Future Work

Findings from the the guessability study (chapter 3) and the field study (chapter 5) led to identify key aspects of user interaction with AR, opening many questions that worth investigating.

A first aspect involves the relationship between AR content and user interaction. Different type of virtual objects proved to cause different reactions in participants, pushing them to discover a strategy for interaction rather than another. It is the case of the 3D virtual text that, during the field study, caused participants to walk around to read it (discovering motion gestures), versus the virtual artefact that did not cause the same reaction. In the field study it also emerged that active interaction, therefore direct manipulation gestures, is the preferred but also the most under-sighted way of interaction. Therefore, investigating on how to develop AR content to encourage manipulation should be further investigated. A feasible attempt could be encouraging exploration through the story game. For example the treasure hunt could lead participants to find specific details of a certain virtual object rather than the entire object itself. That is, visitors could be asked to find the precious stone hidden on the back of the Anubis statue. At the current state the main challenge is to develop strategies that can achieve results without annoying users with not meaningful, hard to accomplish tasks.

Another aspects involves the investigation of how to develop target objects to improve the affordance for manipulation. Because of the implicit consequences that the museum context brings with it, that is visitors are traditionally discouraged to touch, it is needed a strategy to enable interaction, but also to make visitors aware that the interaction is actually enabled. A possible approach could consist in testing on the filed different implementation of target objects, characterised by different shapes, sizes or materials, to verify which one suggests a better affordance of manipulation.

Beside intervening on hidden environment characteristics such as target objects or virtual objects affordance, more explicit approaches could be investigated as well. For example, in the current version of the application tips were embedded in the UI to support user interaction in specific contexts. Tips proved to be only partially successful. In fact as the game proceeded, participants often dismissed them before even reading them. However, the tips role is crucial, as participants highly asked for clear instructions when the game was not clear and easy to understand. Therefore, it is needed to investigate better strategies to provide tips to users. A possible solution could be to disseminate tips not only in the application UI but also in the real environment, creating another invisible link between the real and the virtual world. On these lines, for example, White et al.[89] presented visual hints, which are graphical representations in AR of potential actions and their consequences in the augmented physical world used to enable discovery, learning, and completion of gestures and manipulation in tangible AR.

Finally, a last aspect regarded the use of smartphones and hand-held devices to deploy the application. Smartphones proved to be well suited for the job, bringing a familiar dimension to the user interaction with a novel technology. However, other hand-held devices, such as tables, could provide interesting features as a bigger display screen for a more detailed AR exploration. On the other hand, precisely because of the bigger dimension, it leads to ask if it would discourage the use of direct manipulation gestures. In general, how other familiar hand-held devices can support user interaction with AR still worth investigating.

References

[1] Ahlers, K. H., Kramer, A., Breen, D. E., Chevalier, P.-Y., Crampton, C., Rose, E., Tuceryan, M., Whitaker, R. T., and Greer, D. (1995). Distributed augmented reality for collaborative design applications. In *Computer Graphics Forum*, volume 14, pages 3–14. Wiley Online Library.

[2] Ahn, J. and Han, R. (2012). An indoor augmented-reality evacuation system for the smartphone using personalized pedometry. *Human-Centric Computing and Information Sciences*, 2(1):1–23.

[3] Arvanitis, T. N., Petrou, A., Knight, J. F., Savas, S., Sotiriou, S., Gargalakos, M., and Gialouri, E. (2009). Human factors and qualitative pedagogical evaluation of a mobile augmented reality system for science education used by learners with physical disabilities. *Personal and ubiquitous computing*, 13(3):243–250.

[4] Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., and MacIntyre, B. (2001). Recent advances in augmented reality. *Computer Graphics and Applications, IEEE*, 21(6):34–47.

[5] Azuma, R., Neely, H., Daily, M., and Leonard, J. (2006). Performance analysis of an outdoor augmented reality tracking system that relies upon a few mobile beacons. In *Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 101–104. IEEE Computer Society.

[6] Azuma, R. T. et al. (1997). A survey of augmented reality. *Presence*, 6(4):355–385.

[7] Bajura, M., Fuchs, H., and Ohbuchi, R. (1992). Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. In *ACM SIGGRAPH Computer Graphics*, volume 26, pages 203–210. ACM.

[8] Baldauf, M. and Fröhlich, P. (2013). The augmented video wall: multi-user ar interaction with public displays. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pages 3015–3018. ACM.

[9] Beale, K. (2011). *Museums at play: games, interaction and learning*. Museums.

[10] Behzadan, A. H. and Kamat, V. R. (2005). Visualization of construction graphics in outdoor augmented reality. In *Proceedings of the 37th conference on Winter simulation* , pages 1914–1920. Winter Simulation Conference.

[11] Ben-Joseph, E., Ishii, H., Underkoffler, J., Piper, B., and Yeung, L. (2001). Urban simulation and the luminous planning table bridging the gap between the digital and the tangible. *Journal of Planning Education and Research*, 21(2):196–203.

[12] Betsworth, L., Bowen, H., Robinson, S., and Jones, M. (2014). Performative technologies for heritage site regeneration. *Personal and Ubiquitous Computing*, 18(7):1631–1650.

[13] Billinghurst, M. and Duenser, A. (2012). Augmented reality in the classroom. *Computer*, (7):56–63.

[14] Blasko, G., Coriand, F., and Feiner, S. (2005). Exploring interaction with a simulated wrist-worn projection display. In *Wearable Computers, 2005. Proceedings. Ninth IEEE International Symposium on*, pages 2–9. IEEE.

[15] Blum, T., Stauder, R., Euler, E., and Navab, N. (2012). Superman-like x-ray vision: towards brain-computer interfaces for medical augmented reality. In *Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on*, pages 271–272. IEEE.

[16] Buchmann, V., Violich, S., Billinghurst, M., and Cockburn, A. (2004). Fingartips: gesture based direct manipulation in augmented reality. In *Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, pages 212–221. ACM.

[17] Carmigniani, J., Furht, B., Anisetti, M., Ceravolo, P., Damiani, E., and Ivkovic, M. (2011). Augmented reality technologies, systems and applications. *Multimedia Tools and Applications*, 51(1):341–377.

[18] Caudell, T. P. and Mizell, D. W. (1992). Augmented reality: An application of heads-up display technology to manual manufacturing processes. In *System Sciences, 1992. Proceedings of the Twenty-Fifth Hawaii International Conference on*, volume 2, pages 659–669. IEEE.

[19] Cavallaro, R. (1997). The foxtrax hockey puck tracking system. *Computer Graphics and Applications, IEEE*, 17(2):6–12.

[20] Cavallaro, R., Hybinette, M., White, M., and Balch, T. (2011). Augmenting live broadcast sports with 3d tracking information. *IEEE MultiMedia*, 18(4):38–47.

[21] Chang, H.-Y., Wu, H.-K., and Hsu, Y.-S. (2013). Integrating a mobile augmented reality activity to contextualize student learning of a socioscientific issue. *British Journal of Educational Technology*, 44(3):E95–E99.

[22] Chen, J.-H., Chen, W.-Y., and Chen, C.-H. (2014). Identification recovery scheme using quick response (qr) code and watermarking technique. *Appl. Math*, 8(2):585–596.

[23] Clark, A. and Dunser, A. (2012). An interactive augmented reality coloring book. In *3D User Interfaces (3DUI), 2012 IEEE Symposium on*, pages 7–10. IEEE.

[24] Claus Brenner, V. P., Haunert, J., and Ripperda, N. (2006). The geoscope-a mixed-reality system for planning and public participation. In *25th Urban data management symposium*.

[25] Collett, T. and MacDonald, B. A. (2006). Developer oriented visualisation of a robot program. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pages 49–56. ACM.

[26] Connell, S., Kuo, P.-Y., Liu, L., and Piper, A. M. (2013). A wizard-of-oz elicitation study examining child-defined gestures with a whole-body interface. In *Proceedings of the 12th International Conference on Interaction Design and Children*, pages 277–280. ACM.

[27] Daily, M., Cho, Y., Martin, K., and Payton, D. (2003). World embedded interfaces for human-robot interaction. In *System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on*, pages 6–pp. IEEE.

[28] Díez-Díaz, F., González-Rodríguez, M., and Vidau, A. (2007). An accesible and collaborative tourist guide based on augmented reality and mobile devices. In *Universal Access in Human-Computer Interaction. Ambient Interaction*, pages 353–362. Springer.

[29] Feiner, S., MacIntyre, B., Höllerer, T., and Webster, A. (1997). A touring machine: Prototyping 3d mobile augmented reality systems for exploring the urban environment. *Personal Technologies*, 1(4):208–217.

[30] Fjeld, M. and Voegtli, B. M. (2002). Augmented chemistry: An interactive educational workbench. In *Mixed and Augmented Reality, 2002. ISMAR 2002. Proceedings. International Symposium on*, pages 259–321. IEEE.

[31] Foxlin, E., Altshuler, Y., Naimark, L., and Harrington, M. (2004). Flighttracker: a novel optical/inertial tracker for cockpit enhanced vision. In *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality* , pages 212–221. IEEE Computer Society.

[32] Fritz, F., Susperregui, A., and Linaza, M. T. (2005). Enhancing cultural tourism experiences with augmented reality technologies. 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST).

[33] Furness III, T. A. and Kollin, J. S. (1995). Virtual retinal display. US Patent 5,467,104.

[34] Green, S. A., Billinghurst, M., Chen, X., and Chase, G. (2008). Human-robot collaboration: A literature review and augmented reality approach in design.

[35] Hugues, O., Fuchs, P., and Nannipieri, O. (2011). New augmented reality taxonomy: Technologies and features of augmented environment. In *Handbook of augmented reality*, pages 47–63. Springer.

[36] Hürst, W. and Van Wezel, C. (2013). Gesture-based interaction via finger tracking for mobile augmented reality. *Multimedia Tools and Applications*, 62(1):233–258.

[37] Inami, M., Kawakami, N., Sekiguchi, D., Yanagida, Y., Maeda, T., and Tachi, S. (2000). Visuo-haptic display using head-mounted projector. In *Virtual Reality, 2000. Proceedings. IEEE*, pages 233–240. IEEE.

[38] Ingram, D. (2003). Trust-based filtering for augmented reality. In *Trust Management*, pages 108–122. Springer.

[39] Ishii, H., Ben-Joseph, E., Underkoffler, J., Yeung, L., Chak, D., Kanji, Z., and Piper, B. (2002). Augmented urban planning workbench: Overlaying drawings, physical models and digital simulation. In *Proceedings of the 1st International Symposium on Mixed and Augmented Reality*, page 203. IEEE Computer Society.

[40] Johnson, L. F., Levine, A., Smith, R. S., and Haywood, K. (2010). Key emerging technologies for elementary and secondary education. *Tech Directions*, 70(3):33.

[41] Kato, H., Billinghurst, M., Poupyrev, I., Imamoto, K., and Tachibana, K. (2000). Virtual object manipulation on a table-top ar environment. In *Augmented Reality, 2000.(ISAR 2000). Proceedings. IEEE and ACM International Symposium on*, pages 111–119. Ieee.

[42] Kensek, K., Noble, D., Schiler, M., and Tripathi, A. (2000). Augmented reality: An application for architecture. In *Proc. 8th International Conference on Computing in Civil and Building Engineering, ASCE, Stanford, CA*, pages 294–301.

[43] Kim, J., Huo, X., and Ghovanloo, M. (2010). Wireless control of smartphones with tongue motion using tongue drive assistive technology. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 5250–5253. IEEE.

[44] Kiyokawa, K., Billinghurst, M., Campbell, B., and Woods, E. (2003). An occlusion-capable optical see-through head mount display for supporting co-located collaboration. In *Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, page 133. IEEE Computer Society.

[45] Lingley, A. R., Ali, M., Liao, Y., Mirjalili, R., Klonner, M., Sopanen, M., Suihkonen, S., Shen, T., Otis, B., Lipsanen, H., et al. (2011). A single-pixel wireless contact lens display. *Journal of Micromechanics and Microengineering*, 21(12):125014.

[46] Livingston, M. A., Rosenblum, L. J., Julier, S. J., Brown, D., Baillot, Y., Swan, I., Gabbard, J. L., Hix, D., et al. (2002). An augmented reality system for military operations in urban terrain. Technical report, DTIC Document.

[47] Lv, Z. (2013). Wearable smartphone: Wearable hybrid framework for hand and foot gesture interaction on smartphone. In *Computer Vision Workshops (ICCVW), 2013 IEEE International Conference on*, pages 436–443. IEEE.

[48] Milgram, P. and Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329.

[49] Mistry, P., Maes, P., and Chang, L. (2009). Wuw-wear ur world: a wearable gestural interface. In *CHI'09 extended abstracts on Human factors in computing systems*, pages 4111–4116. ACM.

[50] Miyaki, T. and Rekimoto, J. (2009). Graspzoom: zooming and scrolling control model for single-handed mobile interaction. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*, page 11. ACM.

[51] Morris, M. R., Wobbrock, J. O., and Wilson, A. D. (2010). Understanding users' preferences for surface gestures. In *Proceedings of graphics interface 2010*, pages 261–268. Canadian Information Processing Society.

[52] Morton, H. (1962). Sensorama simulator. US Patent 3,050,870.

[53] Nielsen, J. (1995). 10 usability heuristics for user interface design. *Fremont: Nielsen Norman Group.[Consult. 20 maio 2014]. Disponível na Internet.*

[54] of Art, T. C. M. (2015). Artlens.

[55] of London, M. (2015). Streetmuseum.

[56] Ogi, T., Yamada, T., Yamamoto, K., and Hirose, M. (2001). *Invisible interface for the immersive virtual world*. Springer.

[57] Oka, K., Sato, Y., and Koike, H. (2002). Real-time fingertip tracking and gesture recognition. *Computer Graphics and Applications, IEEE*, 22(6):64–71.

[58] Oladimeji, P. (2008). *Devices, Errors and Improving Interaction Design-A case study using an Infusion Pump*. PhD thesis, Swansea University.

[59] Oyama, E., Tsunemoto, N., Tachi, S., and Inoue, Y. (1993). Experimental study on remote manipulation using virtual reality. *Presence: Teleoperators and Virtual Environments*, 2(2):112–124.

[60] Park, Y. S. and Han, S. H. (2010). Touch key design for one-handed thumb interaction with a mobile phone: Effects of touch key size and touch key location. *International journal of industrial ergonomics*, 40(1):68–76.

[61] Piumsomboon, T., Clark, A., Billinghurst, M., and Cockburn, A. (2013). User-defined gestures for augmented reality. In *Human-Computer Interaction–INTERACT 2013*, pages 282–299. Springer.

[62] Portilla, H. and Basañez, L. (2007). Augmented reality tools for enhanced robotics teleoperation systems. In *3DTV Conference, 2007*, pages 1–4. IEEE.

[63] Qi, W. (2005). A vision-based augmented reality system for visualization interaction. In *Information Visualisation, 2005. Proceedings. Ninth International Conference on*, pages 404–409. IEEE.

[64] Raskar, R., Welch, G., and Chen, W.-C. (1999). Table-top spatially-augmented realty: bringing physical models to life with projected imagery. In *Augmented Reality, 1999.(IWAR'99) Proceedings. 2nd IEEE and ACM International Workshop on*, pages 64–71. IEEE.

[65] Reinhart, G. and Patron, C. (2003). Integrating augmented reality in the assembly domain-fundamentals, benefits and applications. *CIRP Annals-Manufacturing Technology*, 52(1):5–8.

[66] Rekimoto, J. (1997). A magnifying glass approach to augmented reality systems. *Presence*, 6(4):399–412.

[67] Roberts, D. W., Strohbehn, J. W., Hatch, J. F., Murray, W., and Kettenberger, H. (1986). A frameless stereotaxic integration of computerized tomographic imaging and the operating microscope. *Journal of neurosurgery*, 65(4):545–549.

[68] Robinson, S., Rajput, N., Jones, M., Jain, A., Sahay, S., and Nanavati, A. (2011). Tapback: towards richer mobile interfaces in impoverished contexts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2733–2736. ACM.

[69] Rohs, M. et al. (2007). Marker-based embodied interaction for handheld augmented reality games. *Journal of Virtual Reality and Broadcasting*, 4(5):1860–2037.

[70] Rolland, J. P., Biocca, F., Hamza-Lup, F., Ha, Y., and Martins, R. (2005). Development of head-mounted projection displays for distributed, collaborative, augmented reality applications. *Presence: Teleoperators and Virtual Environments*, 14(5):528–549.

[71] Ruiz, J. and Li, Y. (2011). Doubleflip: a motion gesture delimiter for mobile interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2717–2720. ACM.

[72] Ruiz, J., Li, Y., and Lank, E. (2011). User-defined motion gestures for mobile interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 197–206. ACM.

[73] Sanders, B., Crowe, R., and Garcia, E. (2004). Defense advanced research projects agency–smart materials and structures demonstration program overview. *Journal of Intelligent Material Systems and Structures*, 15(4):227–233.

[74] Seungjun, K., Hojung, K., Seongeun, E., Mahalik, N. P., and Byungha, A. (2006). A reliable new 2-stage distributed interactive tgs system based on gis database and augmented reality. *IEICE transactions on information and systems*, 89(1):98–105.

[75] Sielhorst, T., Feuerstein, M., and Navab, N. (2008). Advanced medical displays: A literature review of augmented reality. *Display Technology, Journal of*, 4(4):451–467.

[76] Silva, R. L., Rodrigues, P. S., Oliveira, J. C., and Giraldi, G. (2004). Augmented reality for scientific visualization: Bringing data sets inside the real world. *LNCC–National Laboratory for Scientific Computing, Petropolis, RJ, Brazil*.

[77] Spies, R., Ablaßmeier, M., Bubb, H., and Hamberger, W. (2009). Augmented interaction and visualization in the automotive domain. In *Human-Computer Interaction. Ambient, Ubiquitous and Intelligent Interaction*, pages 211–220. Springer.

[78] Stilman, M., Michel, P., Chestnutt, J., Nishiwaki, K., Kagami, S., and Kuffner, J. (2005). Augmented reality for robot development and experimentation. *Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-05-55*, 2(3).

[79] Sutherland, I. E. (1964). Sketch pad a man-machine graphical communication system. In *Proceedings of the SHARE design automation workshop*, pages 6–329. ACM.

[80] Suzuki, N., Hattori, A., and Hashizume, M. (2008). Benefits of augmented reality function for laparoscopic and endoscopic surgical robot systems. *navigation*, 1:6.

[81] Tang, A., Owen, C., Biocca, F., and Mou, W. (2003). Comparative effectiveness of augmented reality in object assembly. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 73–80. ACM.

[82] Thomas, R. G., William John, N., and Delieu, J. M. (2010). Augmented reality for anatomical education. *Journal of visual communication in medicine*, 33(1):6–15.

[83] Ullmer, B. and Ishii, H. (1997). The metadesk: models and prototypes for tangible user interfaces. In *Proceedings of the 10th annual ACM symposium on User interface software and technology*, pages 223–232. ACM.

[84] Van Krevelen, D. and Poelman, R. (2010). A survey of augmented reality technologies, applications and limitations. *International Journal of Virtual Reality*, 9(2):1.

[85] Vlahakis, V., Ioannidis, N., Karigiannis, J., Tsotros, M., Gounaris, M., Stricker, D., Gleue, T., Daehne, P., and Almeida, L. (2002). Archeoguide: an augmented reality guide for archaeological sites. *IEEE Computer Graphics and Applications*, (5):52–60.

[86] Wagner, D. and Schmalstieg, D. (2006). Handheld augmented reality displays. In *Virtual Reality Conference, 2006*, pages 321–321. IEEE.

[87] Webster, A., Feiner, S., MacIntyre, B., Massie, W., and Krueger, T. (1996). Augmented reality in architectural construction, inspection and renovation. In *Proc. ASCE Third Congress on Computing in Civil Engineering*, pages 913–919.

[88] Wen, R., Tay, W.-L., Nguyen, B. P., Chng, C.-B., and Chui, C.-K. (2014). Hand gesture guided robot-assisted surgery based on a direct augmented reality interface. *Computer methods and programs in biomedicine*, 116(2):68–80.

[89] White, S., Lister, L., and Feiner, S. (2007). Visual hints for tangible gestures in augmented reality. In *Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 1–4. IEEE Computer Society.

[90] Willett, W., Lan, Q., and Isenberg, P. (2014). Eliciting multi-touch selection gestures for interactive data graphics. In *Short-Paper Proceedings of the European Conference on Visualization (EuroVis)*. Eurographics.

[91] Wobbrock, J. O., Aung, H. H., Rothrock, B., and Myers, B. A. (2005). Maximizing the guessability of symbolic input. In *CHI'05 extended abstracts on Human Factors in Computing Systems*, pages 1869–1872. ACM.

[92] Wobbrock, J. O., Morris, M. R., and Wilson, A. D. (2009). User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1083–1092. ACM.

[93] Wörn, H., Aschke, M., and Kahrs, L. (2005). New augmented reality and robotic based methods for head-surgery. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 1(3):49–56.

[94] Wu, M., Shen, C., Ryall, K., Forlines, C., and Balakrishnan, R. (2006). Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces. In *Horizontal Interactive Human-Computer Systems, 2006. TableTop 2006. First IEEE International Workshop on*, pages 8–pp. IEEE.

[95] Yu, D., Jin, J. S., Luo, S., Lai, W., and Huang, Q. (2010). A useful visualization technique: a literature review for augmented reality and its application, limitation & future direction. In *Visual Information Communication*, pages 311–337. Springer.

[96] Yuen, S., Yaoyuneyong, G., and Johnson, E. (2011). Augmented reality: An overview and five directions for ar in education. *Journal of Educational Technology Development and Exchange*, 4(1):119–140.

[97] Zhou, F., Duh, H. B.-L., and Billinghurst, M. (2008). Trends in augmented reality tracking, interaction and display: A review of ten years of ismar. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality* , pages 193–202. IEEE Computer Society.

Appendix A

Game Storyline

The story game comprises an *Introduction*, to walk the player through the game, and *End*, displayed when the game is completed, a series *Clues*, to help the player identifying the next target object to find, and finally a *Artefact Detail* with information and 3D model of the artefact. The story game was provided by the Egypt centre of Swansea.

A.1 Introduction

"Congratulations explorer! I hear that you have discovered a new tomb in the Valley of the Kings. This is wonderful news. As your benefactor I would like you to retrieve some artefacts for me, as gifts I can then give to my family. I want these to be the pick of the bunch and suited perfectly to each member of my family, so here is a list of conditions for each gift."

A.2 Bes

A.2.1 Bes Clues

"For my pregnant wife, Barbara, I want:

- *Something fun and full of life;*
- *Something musical, able to create a soothing note for her and the baby;*
- *Something as delicate and fragile as she is;*
- *Something beautiful, perhaps a bright colour to light up her face;*
- *Perhaps a funny face, to make her and the baby laugh."*

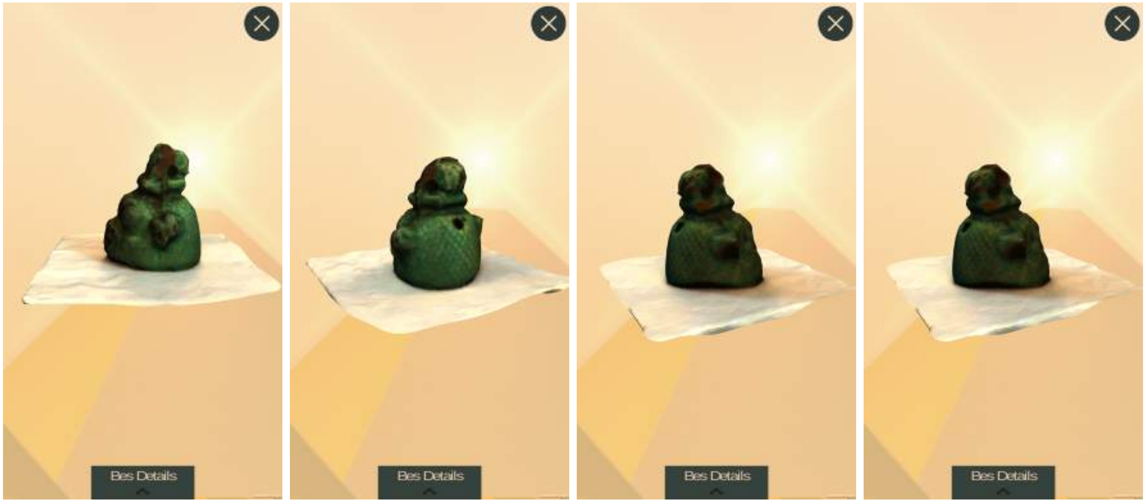


Fig. A.1 Bes 3D Model.

A.2.2 Bes 3D Informative text

"Faience bell in the shape of the head of Bes made out of pale green faience which usually dates to the Late Ptolemaic Periods 747 - 30BC. It is in the shape of a hollow Bes head crowned with feathers and has a hole for suspension as well as another, presumably for the tongue of the bell. The tongue itself is missing. The deity Bes is usually associated with percussion instruments, usually the hand drum, and occasionally a harp. Bes was a protective deity, particularly for women in childbirth and for young children. It is possible that bells like this were amuletic and perhaps worn around the necks of children to protect them. However the fact that this is made from faience, a fragile material, suggests it may have been a votive offering and not intended to be used as a bell at all."

A.3 Anubis

A.3.1 Anubis Clues

"For my son, Alan, I want::

- *Something deep from the depths of the tomb;*
- *Something linked to mummification, that's sure to fascinate him;*
- *Nothing too scary, make it seem friendly as well;*
- *He loves animals, especially dogs."*

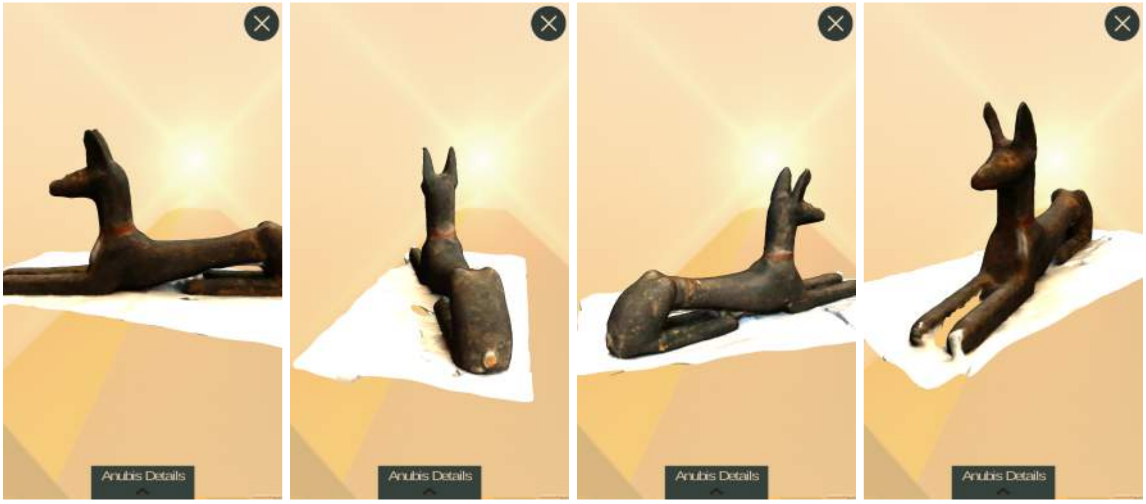


Fig. A.2 Anubis 3D Model.

A.3.2 Anubis Informative Text

"Wood statue of a crouching jackal, representing the god Anubis. When depicted in the form of a jackal like this he is often shown wearing a red collar or a tie. Objects such as this were often placed on top of coffins, and date to the Late Dynastic Period 747-332 BC. Anubis was the god of the dead and was particularly associated with mummification. He was believed to have performed the very first mummification on Osiris, god of the afterlife. Anubis was shown in the form of a jackal because the ancient Egyptians would see these creatures lurking around tombs at night and connected this with the god who would lead the dead from their tombs into the afterlife."

A.4 Ba Bird

A.4.1 Ba Bird Clues

"For my daughter, Amelia, I want: :

- *Something unusual, but with a friendly face;*
- *She loves to visit the market, so something connected with gathering provisions;*
- *Something with lots of colour, to brighten up her day;*
- *She has a fondness for birds, she loves to listen to their beautiful song."*

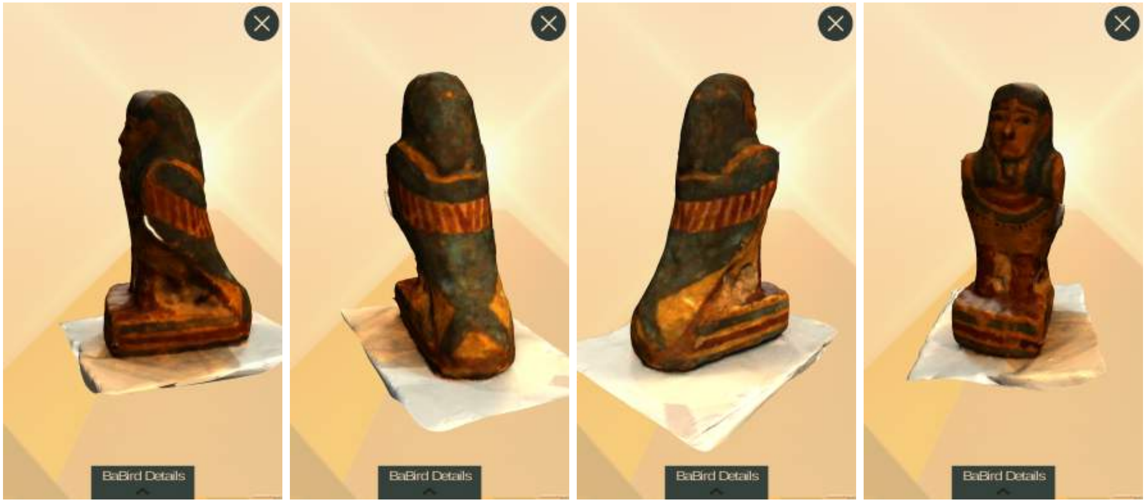


Fig. A.3 Ba Bird 3D Model.

A.4.2 Ba Bird Informative Text

"Wooden Ba bird decorated with red and black paint. The paint seems to be put straight on top of the wood, without a gesso layer first (a type of white plaster). Examples of Ba bird statues appear in the New Kingdom but are much more common in the mid-late 1st millennium BC. Many have holes with broken dowel fragments in the base suggesting that they were attached to shrines, stela or on top of coffins. The Ba is similar to our idea of personality. The dead person had to leave the tomb to join with his Ka, or soul if he was to become an everlasting spirit and live forever in the afterlife. As the actual body could not leave the tomb, the Ba had to do this. The Ba was shown as a bird as it could fly away and leave the tomb. It would have a human head representing the deceased."

A.5 Cippus

A.5.1 Cippus Clues

"For my brother, Howard, I want::

- *He's interested in Egyptian religion, something to do with that;*
- *Something more homely, a bit more domestic;*
- *Something made of stone;*
- *Something to remind him of his son."*

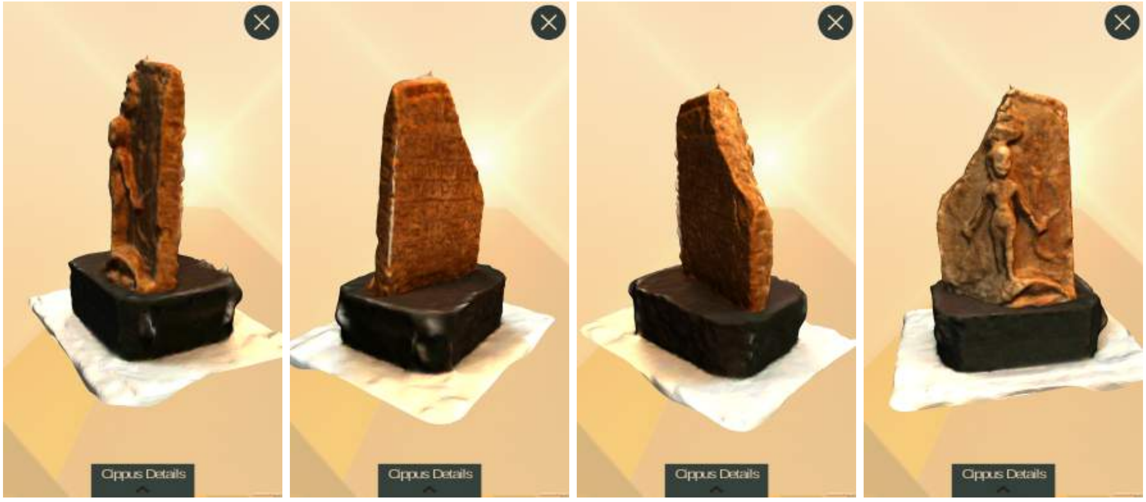


Fig. A.4 Cippus 3D Model.

A.5.2 Cippus Informative Text

"This type of stela is called a 'cippus'. This object is made from a stone called steatite and is likely to have come from Abydos. Items such as these were found in the home of both rich and poor and usually date from the New Kingdom to Graeco Roman times. It shows the god Horus in the form of a child standing on a crocodile and holding snakes in his hands. In this way, he displays victory over dangerous animals. Above would have depicted the head of the protective deity Bes, although this is now badly damaged. The stela also has magical spells written on the reverse and around the edge. It is usually believed that such items had magical qualities against illness; if you poured water over the Cippus, the figure of Horus the Child together with the magic spell on the back and the image of Bes would impart healing powers to the water."

A.6 Sekhmet and Nefertum

A.6.1 Sekhmet and Nefertum Clues

"For my mother, I want::

- *Something divine, depicting Egyptian gods;*
- *Something showing a mother and son to remind her that I'm thinking of her;*
- *Something with a back-support, she's getting older;*
- *Something not too bright, subtle colours are best;*

Fig. A.5 Sekhmet and Nefertum 3D Model.

- *Something brave and fierce, she?s a feisty woman!.'*

A.6.2 Sekhmet and Nefertum Informative Text

"This statue dates to the Late Dynastic Period 747-332 BC, and is made of faience which is a type of early ceramic made in a similar way to glass. It depicts the god Nefertum and his mother the goddess Sekhmet. Nefertum is closely connected to the lotus flower, its perfume and its association with creation. As such, he would have been a strong symbol of life. His mother, the lioness headed Sekhmet is goddess of war, pestilence and toothache. Although she would have been believed to be responsible for causing sicknesses, the ancient Egyptians would have turned to her in the hopes she would make them better and remove their illnesses. The back support of this faience statue has prayers to Sekhmet, her husband Ptah and their son Nefertum for 'all life, health and graciousness of heart'."

A.7 End

"Excellent! You have found beautiful artefacts that will make perfect gifts for my family. Keep up the good work; I will be more than happy to finance your next dig to find me even more wonderful things!"

Appendix B

Application Development

B.1 Development Supporting Tools

In the following section software used for the application development, debugging, maintenance, and versioning control is presented.

Unity3D[1] is a cross-platform game engine developed by Unity Technologies and used to develop video games for PC, consoles, mobile devices and websites. Despite its cross-platform nature, the game was developed uniquely for Android devices. Unity3D was used as main platform to develop the application game.

MonoDevelop[2] is the integrated development environment supplied with Unity. It was used to write and debug code.

Vuforia[3] is an augmented reality software development kit for mobile devices that enables the creation of augmented reality applications. The library was integrated in the Unity3D environment to support the application game development.

Target Manager[4] is a web-based tool that enables to create and manage target databases on-line. Images need to have predefined characteristics in order to work well with recognition algorithms. Therefore, when images are uploaded, they are ranked from zero to five to indicate the goodness of the image in terms of recognising characteristics. Once that a set of

[1] https://unity3d.com/
[2] http://docs.unity3d.com/Manual/MonoDevelop.html
[3] https://www.qualcomm.com/products/vuforia
[4] https://developer.vuforia.com/target-manager

images is ready, it can be downloaded and exported as asset to be imported in a Unity3D project.

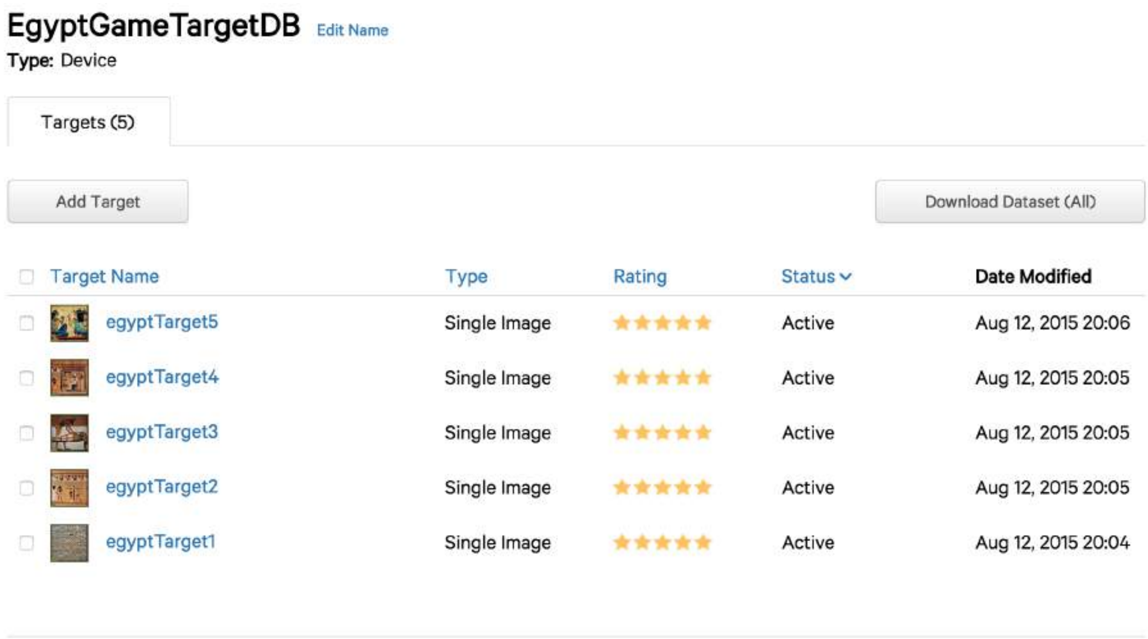


Fig. B.1 Target Image Vuforia Website.

Tinkercad[5] is a free, on-line application that can be used to create and print 3D models. It was used to generate 3D content for the application.

Unity Assets Store[6] is the Unity3D on-line store where assets can be bought and sold. It was used to buy the 3D model of the treasure hunt bag.

Logcat[7] is the Android logging system which provides a mechanism for collecting and viewing system debug output. It was used used for the debugging purposes.

Unity3D-LogCat-extension[8] is an extension for Unity3D which allows to use the Logcat feature directly inside the Unity3D environment.

[5] https://www.tinkercad.com/
[6] https://www.assetstore.unity3d.com/
[7] http://developer.android.com/tools/help/logcat.html
[8] https://github.com/dzonatan/Unity3D-LogCat-extension

Crashlytics[9] is a light weight crash reporting solution that can be integrated in an Android project to collect and analyse crashes. Because Crashlytics can be easily integrated in Android Studio but not in Unity3D, few passages were needed for integrating of the tool in the final application. That is, the game application had to be first exported from Unity3D as an Android project and then integrated with the Crashlytics plugin using Android Studio.

Android Studio[10] is the official integrated development environment for Android applications. It was mainly used to integrate the Crashlytics plugin into the final application and to use the Logcat feature.

GitHub[11] is a web-based Git repository hosting service. It was used locally and remotely for versioning control the Unity3D project. The Unity3D project can be retrieved at https://github.com/moodymood/EgyptGameAR-Demo.

B.2 *Mobile Egypt* Unity3D Project

In the following sections are presented the main components of the Unity3D project.

B.2.1 Scenes

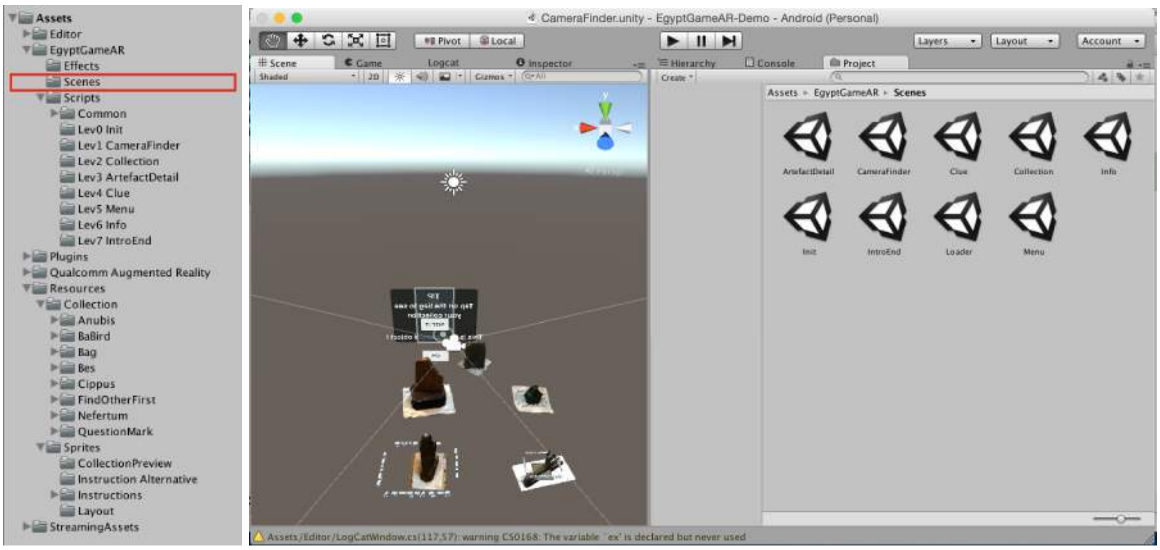


Fig. B.2 Unity3D Project Scenes.

[9]https://fabric.io/kits/android/crashlytics
[10]http://developer.android.com/tools/studio/index.html
[11]https://github.com/

Scenes contain the objects of the game, and they are usually used to create main menu or individual levels. The game was composed by eight different scenes. Generally, scenes correspond to the different sections introduced in Chapter 4. From figure B.2:

Menu corresponds to the *menu* section;

CameraFinder corresponds to the *camera view finder* section;

ArtefactDetails corresponds to the *artefact details* section;

Collection corresponds to the *artefact list* section;

Clue, Intro/End correspond to the *clue* transition;

Loader correspond to the *loader* transition;

Init is an empty scene used to initialise the game parameters when the application is first loaded;

B.2.2 3D Models

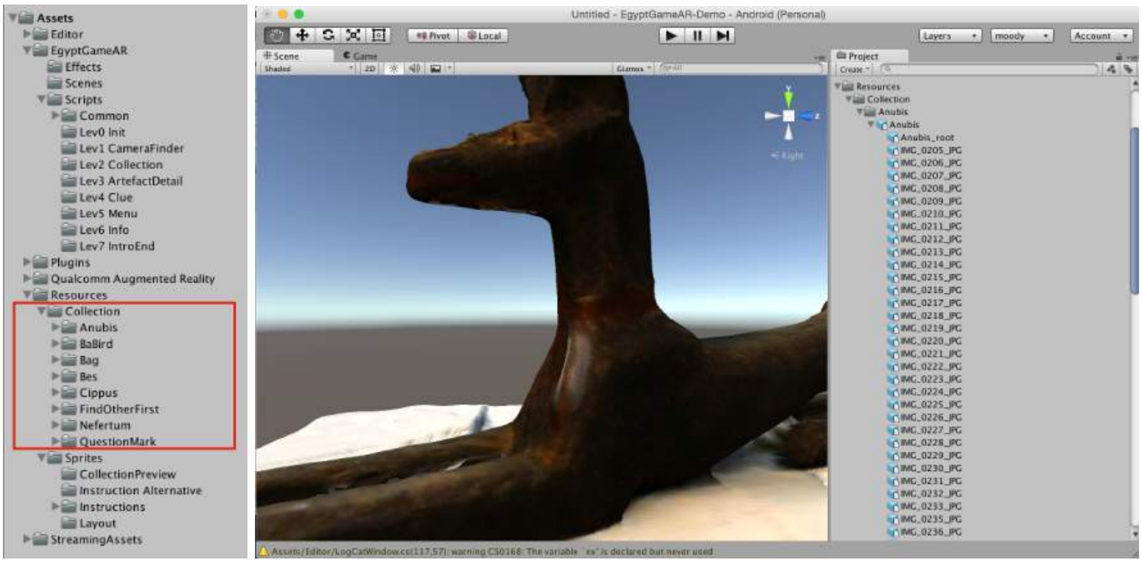


Fig. B.3 Anubis 3D Model Asset.

In the application were used a total of five 3D models of artefacts plus a 3D text. The 3D text was generated using Tinkercad, while 3D artefacts model were imported as external assets on concession of Lewis Hancock, who personally acquired the models from the Egypt centre for his BSc dissertation project *Mobile Egypt: Educational Game Development Using Object Recognition*.

B.2.3 Scripts

Project scripts are the core of the game, allowing to dynamically change the application behaviour. Scripts belongs to two main categories. When they inherit from `MonoBehaviour`, they are strictly related to the game logic. That is, they manage the game workflow, they control game objects, or they are used to programmatically construct the user interface. It is also possible to develop generic classes to create a backbone for the game.

Game Management

Different classes were developed to support the game. For example, `Collection` is a class to manage list of `Artefact`. Given an XML file properly formatted, it generates a structure containing a list of `Artefact` where for each element is stored: id, name, clue and information text. In this way the logic of the game is kept separate from the content. New artefacts can be easily added or changed without interfere with the core structure of the game. The complete code of the project can be retrieved at https://github.com/moodymood/EgyptGameAR-Demo.

```
// Collection.cs
public class Collection : IEnumerable<Artefact>{

  private List<Artefact> collection;

  // Empty Constructor
  public Collection(){
    collection = new List<Artefact> ();
  }

  // Constructor reading from XML file
  public Collection (String XMLContent){
    collection = new List<Artefact> ();
    XDocument xdoc = XDocument.Parse(XMLContent);
    var queryResult = from artefact in xdoc.Descendants("artefact")
    select new {
      id = artefact.Attribute("id"),
      artefactNames = artefact.Descendants("name"),
      artefactStories = artefact.Descendants("story"),
      artefactClues = artefact.Descendants("clue")
    };
```

```
    foreach (var artefact in queryResult){
      addArtefact(new Artefact(
        Int32.Parse(artefact.id.Value),
        artefact.artefactNames.FirstOrDefault().Value,
        artefact.artefactStories.FirstOrDefault().Value,
        artefact.artefactClues.FirstOrDefault().Value));
    }
  }

  // Basic methods to add or get Artefacts from the Collection

  public void addArtefact(Artefact artefact){
    collection.Add (artefact);
  }

  public void addArtefact(int id, string name, string story, string clue){
    collection.Add (new Artefact (id, name, story, clue));
  }

  public Artefact getArtefactByName(string artefactName){
    return collection.Find((x => x.getName() == artefactName));
  }

  public Artefact getArtefactById(int artefactId){
    return collection.Find((x => x.getId() == artefactId));
  }

  public int getTotal(){
    return collection.Count;
  }

  public IEnumerator<Artefact> GetEnumerator()
  {
    for (int index = 0; index < getTotal(); index++)
      yield return collection[index];
  }

  IEnumerator IEnumerable.GetEnumerator()
  {
```

```
      return GetEnumerator();
   }
}


public class Artefact{

   private int id;
   private string name;
   private string story;
   private string clue;

   public Artefact(){
      this.id = -1;
      this.name = null;
      this.story = null;
      this.clue = null;
   }

   public Artefact(int id, string name, string story, string clue){
      this.id = id;
      this.name = name;
      this.story = story;
      this.clue = clue;
   }

   public int getId(){
      return id;
   }

   public string getName(){
      return name;
   }

   public string getStory(){
      return story;
   }

   public string getClue(){
```

```
    return clue;
  }

  // Definition of the Equality operator for Artefact objects
  public override bool Equals(System.Object obj)
  {
    if (obj == null)
      return false;

    Artefact artefact = obj as Artefact;
    if ((System.Object)artefact == null)
      return false;
    return (getName() == artefact.getName()) && (getId() ==
        artefact.getId());
  }

  public bool Equals(Artefact artefact)
  {
    if ((object)artefact == null)
      return false;
    return (getName() == artefact.getName()) && (getId() ==
        artefact.getId());
  }

  public override int GetHashCode()
  {
    return getId ().GetHashCode () + getName ().GetHashCode();
  }

}
```

3D Model Interaction

Interaction with the 3D model is possible through zoom, rotation and translation over the Y axis. In the following snippet it is reported how the 3D model interaction is achieved.

```
//ArtefactGestures.cs
// If only one finger is touching the screen, rotation or translation over
    the Y axis
```

```
if (Input.touchCount == 1) {
   var touch = Input.GetTouch (0);
   if(touch.phase == TouchPhase.Moved){
         artTransform.Translate (0, touch.deltaPosition.y * 0.3f, 0);
         artTransform.Rotate (0, -touch.deltaPosition.x * 0.3f, 0);
         if (!isVector3DInsideBottomWindow (artTransform.position)) {
            artTransform.Translate (0, -touch.deltaPosition.y * 0.3f, 0);
            artTransform.Rotate (0, +touch.deltaPosition.x * 0.3f, 0);
         }
   }
// If two finger are touching the screen, pinch to zoom
} else if (Input.touchCount == 2) {
   Touch touchZero = Input.GetTouch (0);
   Touch touchOne = Input.GetTouch (1);

   Vector2 touchZeroPrevPos = touchZero.position - touchZero.deltaPosition;
   Vector2 touchOnePrevPos = touchOne.position - touchOne.deltaPosition;

   float prevTouchDeltaMag = (touchZeroPrevPos - touchOnePrevPos).magnitude;
   float touchDeltaMag = (touchZero.position - touchOne.position).magnitude;

   float deltaMagnitudeDiff = prevTouchDeltaMag - touchDeltaMag;

   GetComponent<Camera> ().fieldOfView += deltaMagnitudeDiff *
       perspectiveZoomSpeed;
   GetComponent<Camera> ().fieldOfView = Mathf.Clamp (GetComponent<Camera>
       ().fieldOfView, 20.1f, 120.9f);
}
```

To be noted that the artefact must be confined inside the view port, otherwise it will disappear from the screen and users will not longer be able to access it. Therefore a specific function has been implemented, to constraints the 3D model interaction within the view port space. The complete code of the project can be retrieved at https://github.com/moodymood/EgyptGameAR-Demo.

```
// Given a Vector3 the function return true if the point belongs to the view
    port, false otherwise
public bool isVector3DInsideBottomWindow(Vector3 artefactPos){
   Vector3 pos = Camera.main.WorldToViewportPoint (artefactPos);
```

```
    pos.x = Mathf.Clamp01(pos.x);
    pos.y = Mathf.Clamp01(pos.y);
    if (pos.x > 0 && pos.x < 1 && pos.y > 0 && pos.y < 1)
        return true;
    else
        return false;
}
```

AR Interaction

Interaction with virtual objects is made possible through motion and direct manipulation gestures for zoom and rotation tasks. Those functionalities are implemented by default by the Vuforia library. The selection task was implemented with a custom class. The complete code of the project can be retrieved at https://github.com/moodymood/EgyptGameAR-Demo.

```
// DragArtefact.cs
if (Input.touchCount == 1) {
    Vector3 touchedPoint = Input.GetTouch (0).position;
    Ray ray = Camera.main.ScreenPointToRay (touchedPoint);
    RaycastHit hit;

    bool isArtefactHit;

    // If an object has been hit
    if (Physics.Raycast (ray, out hit)) {
        if(hit.transform.name == "BagModel"){
            bagTransform = hit.transform;
            isArtefactHit = false;
        }else {
            artefactTransform = hit.transform;
            game.setCurrentArtefact(game.getCollection().
            getArtefactByName(artefactTransform.name));
            isArtefactHit = true;
        }

        switch (Input.GetTouch (0).phase) {

            case TouchPhase.Began:
```

```
        // The hit object exists in the collection
        if(isArtefactHit){
            draggingArtefact = true;
            startPoint = artefactTransform.position;
        }
    break;

    case TouchPhase.Moved:
    // The artefact is anchored to the user finger
    if(isArtefactHit && draggingArtefact){
        hit.transform.position = Camera.main.ScreenToWorldPoint (new
            Vector3 (touchedPoint.x, touchedPoint.y, startPoint.z));
        // The artefact is hitting the bag
    }else if(draggingArtefact){
        // The artefact hasn't been collected yet
        if (!game.getCollectionStatus ().isCollected
            (game.getCurrentArtefact())) {
            game.getCollectionStatus().setAsCollected(game.getCurrentArtefact());
            game.setArtefactJustCollected(true);
            isFirstTime = true;

            StartCoroutine(callParticlesEffect(artefactTransform.gameObject));
            StartCoroutine(resize());
            StartCoroutine(rotate());

            StartCoroutine(loadLevel());
            // The artefact has been collected
        }else{
            // Artefact collected, so show error message
            if(!isFirstTime)
                GameObject.Find("CameraFinderSceneManager").
                    GetComponent<PopUpManager>().setShowPopUp(true);
                GameObject.Find("CameraFinderSceneManager").
                    GetComponent<PopUpManager>().setPopUpMessage("You already
                        collected " + game.getCurrentArtefact().getName();
        }
    }
    break;
```

```
        case TouchPhase.Ended:
           //The hit object exists in the collection
           if(draggingArtefact){
              draggingArtefact = false;
              artefactTransform.position = startPoint;
           }
        break;
     }
  }
}
```

Appendix C

Prototyping

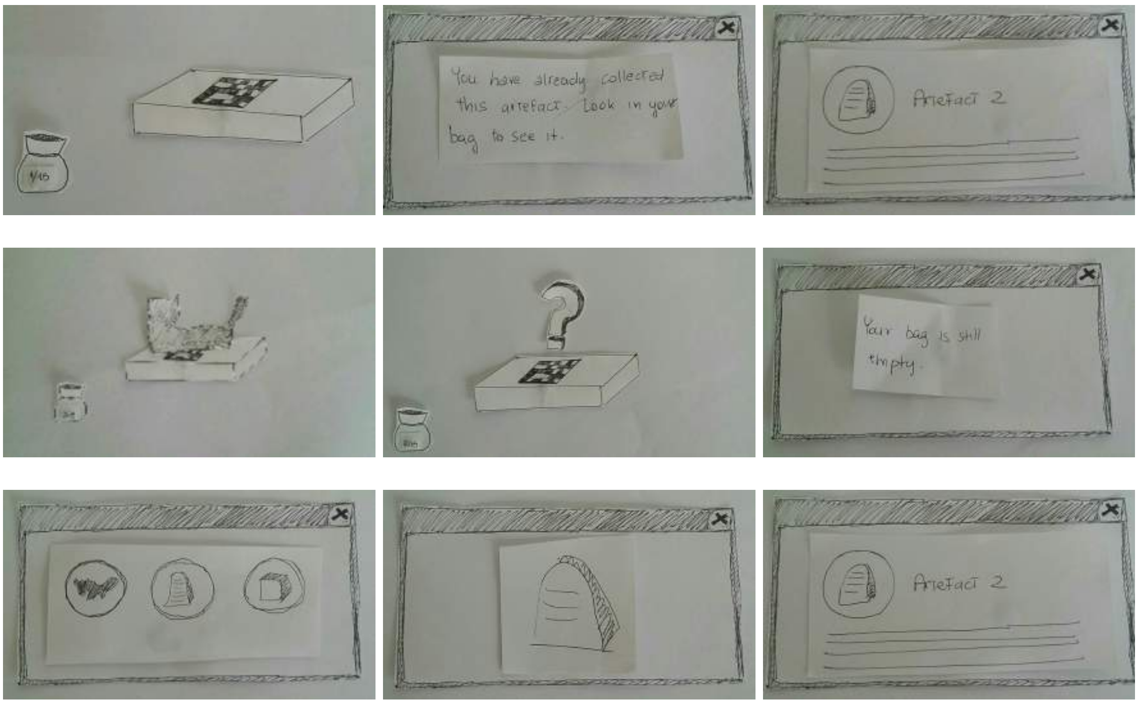


Fig. C.1 Paper prototyping with PopApp.

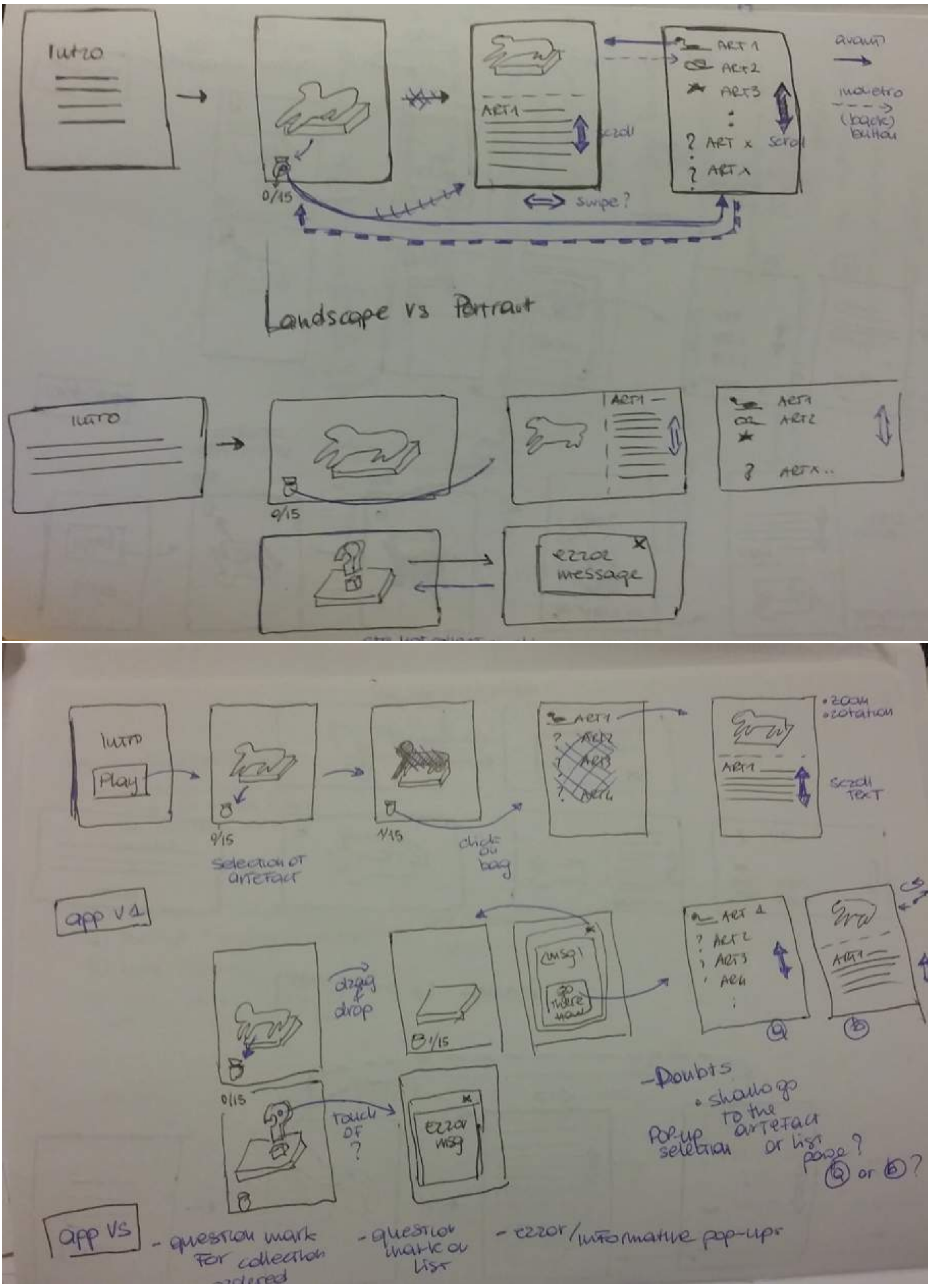


Fig. C.2 Workflow and layout paper prototyping.

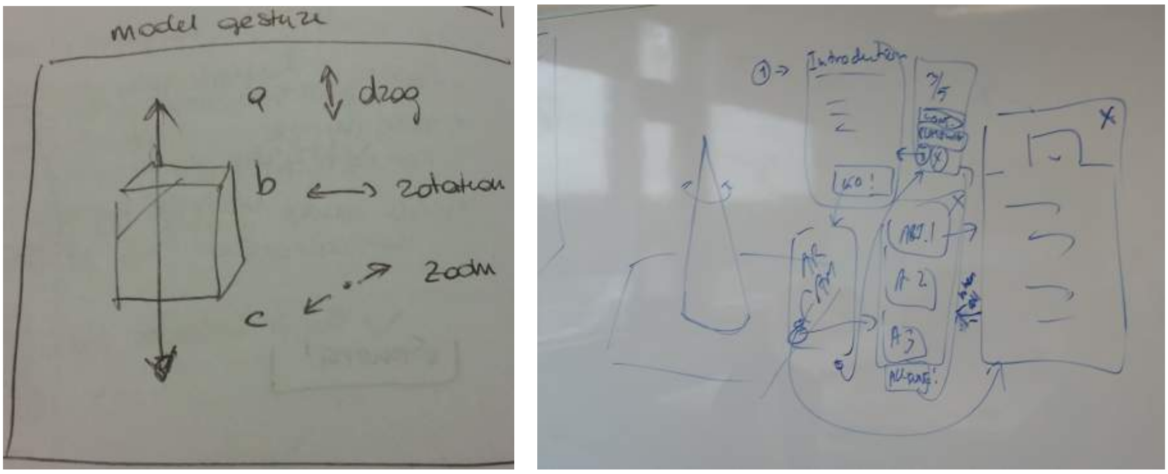


Fig. C.3 3D model interaction prototyping.

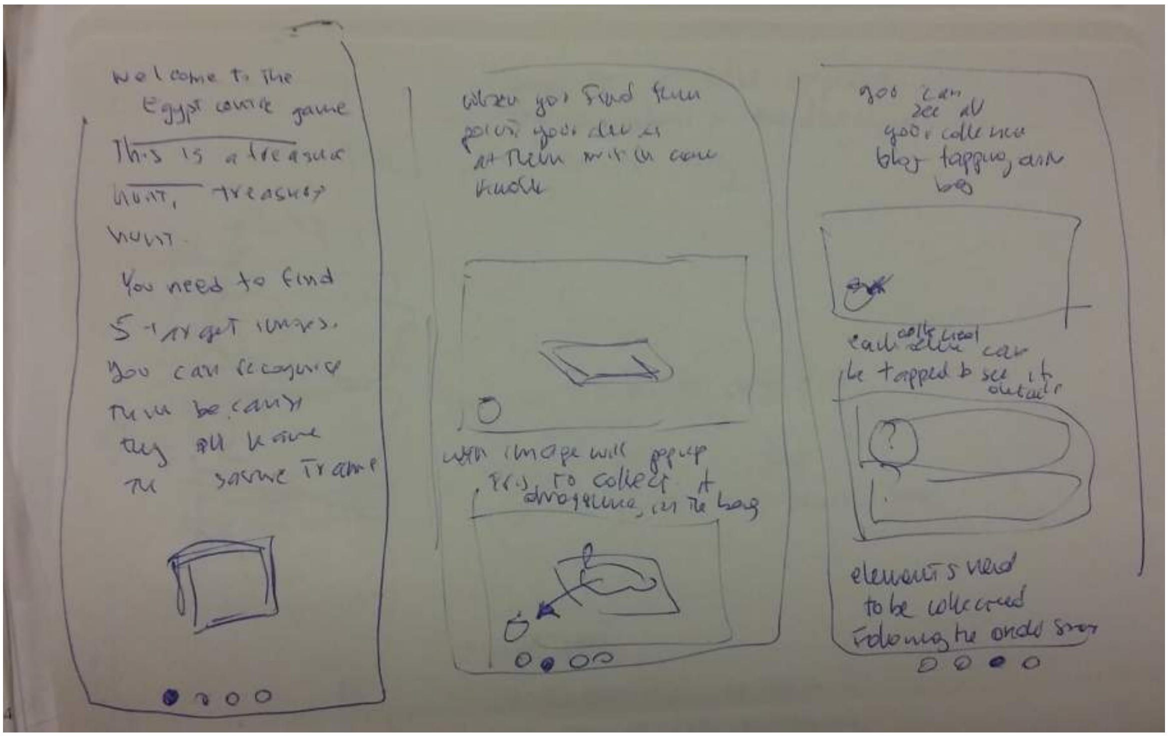


Fig. C.4 Help section paper prototyping.

Appendix D

Study in the Field

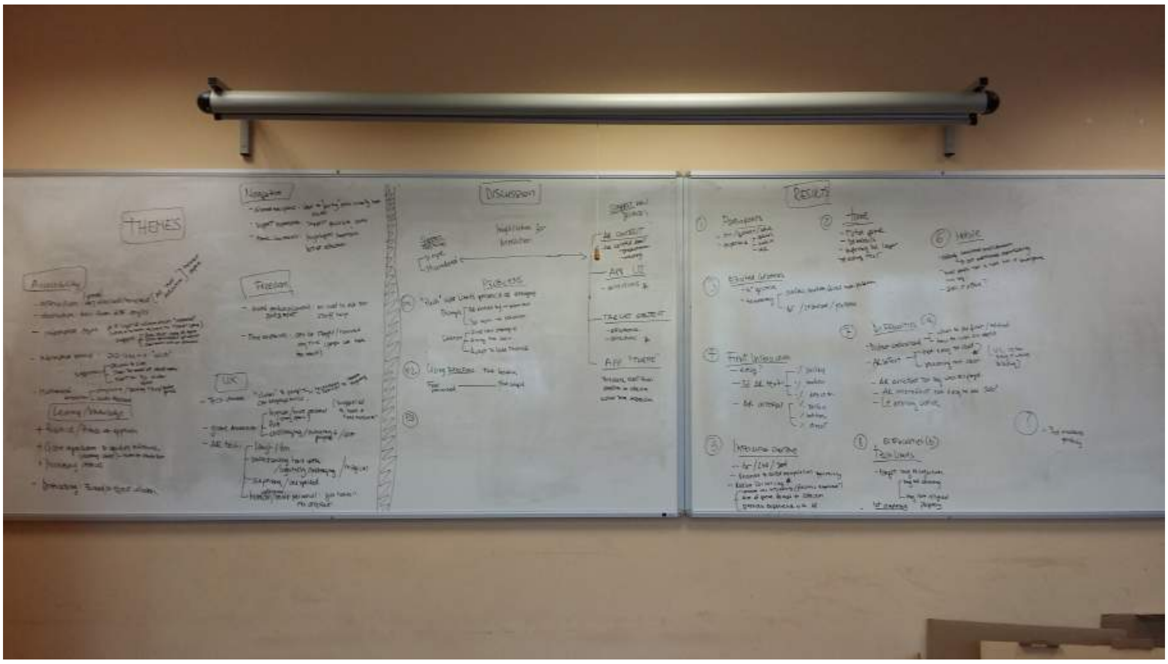


Fig. D.1 Whiteboard for the thematic analysis of the field study.

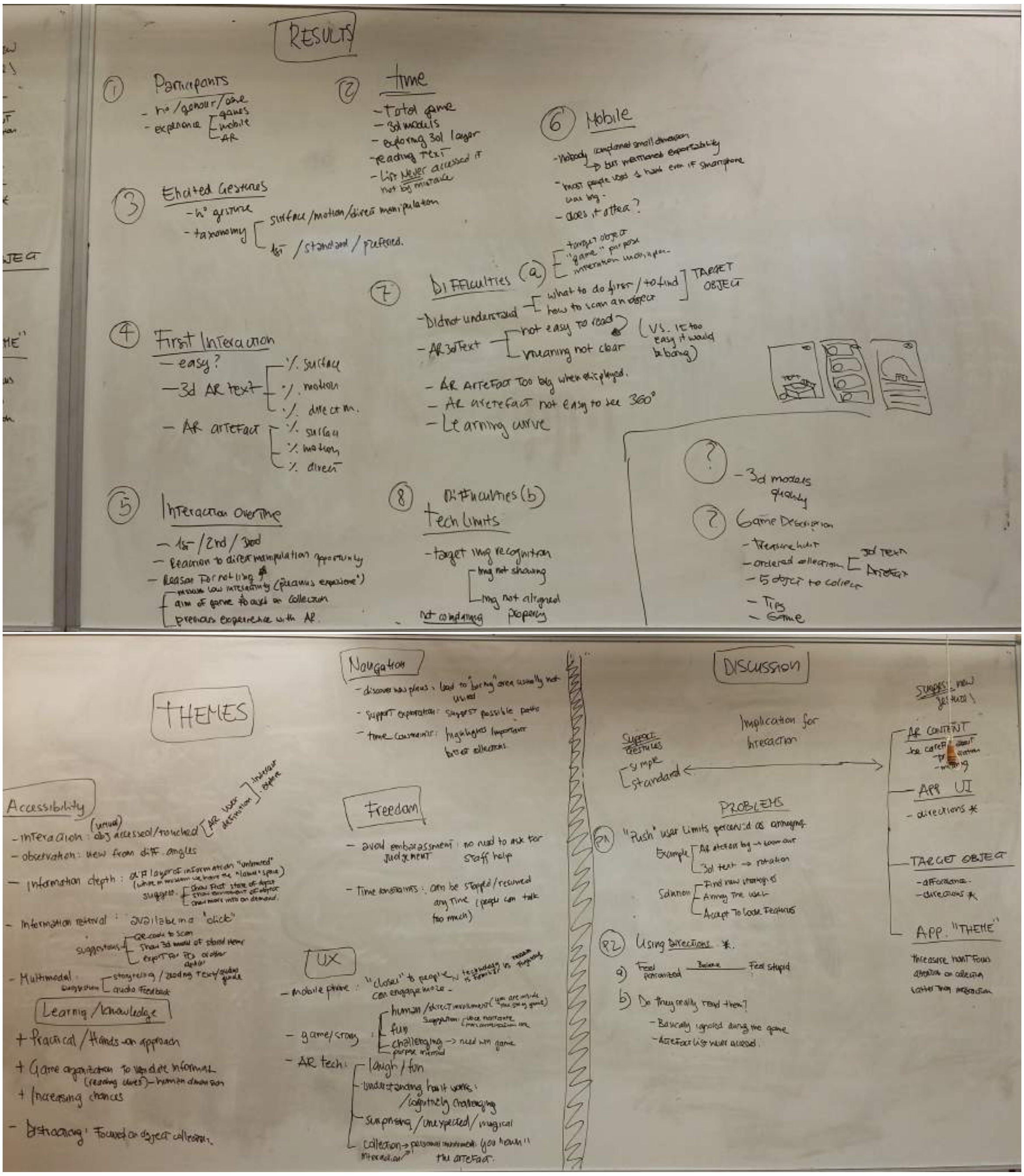
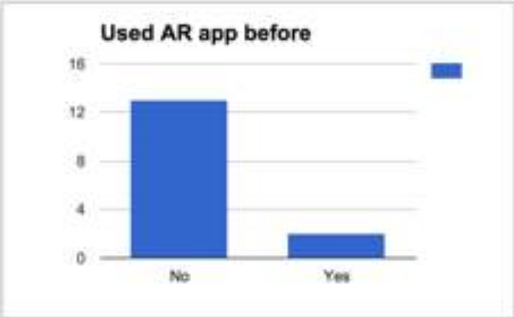
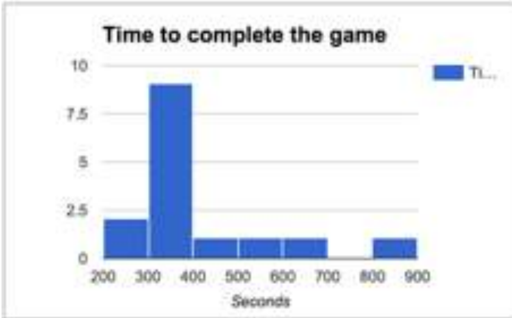


Fig. D.2 Thematic analysis for the field study: close up.





AR definition

- "A system of interacting with physical objects without physical contact"
- "It's like having 3D images which they might not be there but you feel like they are there"
- "Basically you have the phone with your camera on it and if you point the pictures on the boxes it should 3D objects on your phone that are not actually there"
- "Walking through the physical world to interact with objects in a digital plane"
- "Pointing a smartphone at an object or pictures and it is going to show you the artefact which is associated to that image"
- "Through a digital device to interact with something which is actually not there"
- "It's like a virtual.. it's like 3d pictures it's cool"
- "Adding things to you your perception on the environment that otherwise would not be there through technology"
- "You look through the screen and you see something you couldn't see otherwise"
- "there were things, the app make things appear the were not there, so you could sort of", "magical?"
- "Something between some sort of hologram in 3D where you can explore the object"

First reaction to AR

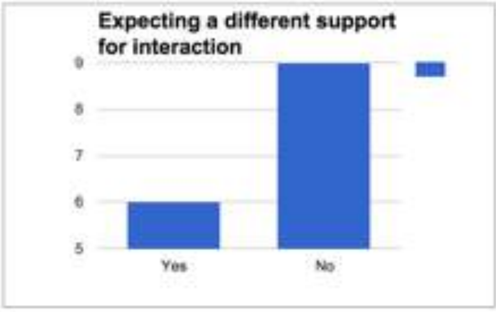
- "It was quite impressed with it"
- "It was quite unique, I thought it was fun"
- "It pop-up just too big, but I thought it was quite impressive I thought it was very cool that you could see all the details"
- "I liked the way it pop-up up and the fact the you can actually interact with it dragging it inside the bag"
- "I thought where it came from (laughing) I was pointing at a box not at a statue coming out of it"
- "It was quite attractive for my attention, it would attract people attention, it's a good way" (not a native speaker)
- "It was fantastic the way it pops up is actually quite cool to be able to look around the way you would be if you had a real object in front of you, it makes you really wanna be more in the game"
- "It was quite good, it was what I was expecting but I got surprised when the 3D model popped out"
- "That's amazing, that's brilliant"
- "It was good"
- "It was cool"
- "Oooh, well done! Nicely done, it looks real it looks nice, fun!"

Expecting a different support for interaction

Talking about interaction with the AR MODEL expressed the desire for surface gestures as for other application. "I was a bit surprised because when the object first appeared it looked like something I could actually turn around by myself but then I discovered I that to make the object move on the screen I had to actually move myself ". "There is nothing wrong with that but it was not what I was expecting from my experience".

- Preference over motion gesture

- He was very focused on the game so he said he din't think of interacting too much with the AR as he was thinking about the aim of the game (collection) but that he would have rather enjoy to explore those possibility at the end of the game. Speaking of motion gesture "to see the artefact from behind would be difficult". On surface and motion gesture "I think both are interesting, at the beginning is funny to walk around and see objects but if you want to examine it close without turing around it could be quite convenient to have surface gestures"



- "If the target object are not accessible because in cages then it would be nice to be able to just drag it to turn it around". Even if she thought having surface gestures was good and common, she still preferred motion and direct manipulation gestures.
- "I would like it more (direct manipulation)" because then I could move it freely without accidentally putting it in the bag" "It was ok (the surface gesture) but I go confused from the first message, I though I had to drag the real object"
- "I'd like a surface gesture so that you can spin and rotate but make sure that it stays in one spot and it is not going to travel"
- "I'd try surface gesture but that's just because of the way I am use to use the phone"

When showed the direct manipulation gesture reaction
- "I see! I see! I assumed because I was in a museum that I wasn't allow to touch (...) That's very good actually!"
- "Aaahn okay, that's quite fun (laughing)!" "I just didn't think of it, that's actually makes sense a lot", "maybe a round base would be suggest more that you can turn"
- "Oh right that's very cool, I like that, I like that a lot" "It's more and hands on approach". On interaction, when asked why she didn't think about touching the object she said "I wasn't sure I was supposed to move the object where if I did it the image on the phone could sort o disappear"
- "From experience I've had with augmented reality in the past everything says in the place where it is meant to be and the actual movement is limited just to the screen to the digital side rather than to the physical side". "It's something which would need to be sign posted and highlighted so that people actually think to do it".
- "I didn't realize the boxes would be out where you can do that, that's good I would say as long as it is accessible to be turn, that really cool I am having fun just paying in that, no yeah that's fantastic"
- "No I didn't (laughing), oooh I see, I my Gosh that's clever (laughing), how do you do that, that's really cool"
- "That's crazy, spacial awareness I didn't know I could move the box, so I just kept moving around the box to try to read it"
- "Oh no! I haven't though about that really, I though everything was on the screen, I wouldn't move that one" "That's cool"

Difficulty on interaction
- I didn't realize it was a treasure hunt;
- I didn't understand how to scan the object
- I didn't understand what I had to find
- I didn't understand how to drag the object
- Reading the text was not easy
    - Reading the text was not easy, BUT if it was super-easy it would be boring
    - The meaning of the 3D text was confusing (only read/understood 'you found it'
- Sometimes the object went upside down

- Sometimes the image was now recognized or on focus
- Struggling to understand what he had to scan and also the difference between real and virtual objects
- The object was too big to be seen entirely, so I had to move away
    - It was confusing at the beginning because I was very very close to the object and it wasn't until the second one that realize that I had to actually put out the phone a little bit so that it gives the space to the screen space to fit it the object, but after that it was good"

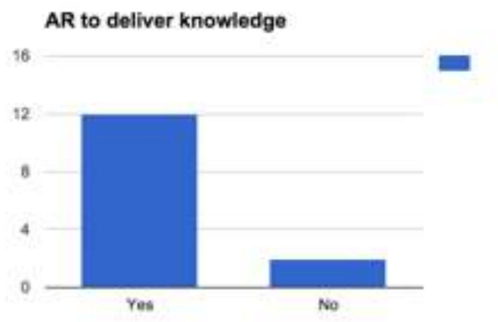
Improvements
- Better quality for the 3D models (Talking about the quality of the image "it is easy recognisable but it lacks in detail");
    - vs "The quality was good"
- More objects
- clear instructions when the purpose of the game is not clear
    - "I wanted instruction telling me what to do instead of walking around and be challenged by the application"
    - I didn't like the fact that I didn't know what to do, I felt a bit silly
- Auditorial feedback
- Reading the text feature as an audio-guide
- Make it more "personal" to give it a human aspect like a conversation more like "just to show you are not alienating people"
- Making it exportable for PC so that it can be consulted and explored later
- "Another thing could be interesting would be like a square code next to the object that you can scan and the information just comes up because at the moment you have to go to the book you have to flick through but I could just scan the box next to the object and have the full description or even better a 3D things of it that would be the possible outcome really"
- "Could be useful to recreate the first state of an object which is deteriorated or the area around it such as buildings"
- "if we could to it for every object instead of having the information on the card I guess.. but that would probably be expensive"

AR museums applications
- "It's often very frustrating while visiting museums because you cannot puck up the object, or look at it throughout different angles. If they are in a case you can never look what it is behind".
- like the fact that once you have the app you can just you don't need to be lead or supervised (...) I think for children is good because they can go and explore on their own, they don't need adults with them, to tell them what to do (...) and of course it's a game, kids will automatically respond to that"
- As an adult in a much larger places with more items (referring to the Egypt museum as comparison) it would be very useful, especially if I was in a hurry, and I wanted to see only the interesting stuff, and I had a guide like this it would be very useful"
- "I think (visiting the museum using these kind of apps) would be a lost more enjoyable especially for kids as they no longer have to go to the case they can walk and find things and it may feel like it is their own things that they are finding?"
- "I think it would be definitely useful in museums, it would engage people a lot more, especially children I can see"
- "I think it could probably take you to areas of the museum that you wouldn't probably go normally.. so normally I'd avoid all the pottery section but if there was a box there I would obviously go there and than look at other objects al well"
- "It was a lot more interactive am I think it's reading is fine but that's a lot more fun to use and sort of presents informations in a much more different way than books and all the other things in a museum do"

- "I think there is going to be a lot of attractions with visitors I think it's gonna be sort of come and try I see a lot of people come in"
- "Because it brings out the information about the objects, because people they have to go and find them they might be actually more interested in interested in reading what they are"
- "It can be more education as well"
- "It will appeal to young audience showing them that we are getting them involved in technology as well, I see the novelty value so people who have been here before who wants to come back because it's new and it's shiny"
- "I think it's going to be more enjoyable and interactive, of course I am slightly biased about museums but I have friend and family who don't particularly like them because all the see are a loads of objects in cases, there is not much to actually DO whereas this is much more interactive"
- "It think it can appears more free to people compared to other activities you can try and if you want to do something else you can close the application and I think it can involve more because they are not feeling too much observed"
- "It would make museums so much more accessible I would enjoy so much more as a kid."
- "It makes you feel like you're really interacting with the museum rather than just looking at things in cases (...) and it gives you more of the history for me for example I was shy and I didn't want to go and ask people about the story of an object so to be able to find it and then read about it is actually quite nice"
- "I think it will improve their (visitors) experience, especially with younger children because we bring younger children in and they don't always fully grasp but being able to at it and see the object come out is going to be helpful for them"
- "I feel the fact it's more hands on it will make accessible the museum to a wider range of people"
- "Doing activities because obviously you go to a museum to see the actual objects in cases but in term of activities it is interesting it's cool to be able to see the object on the screen as well"
- "It's definitely a 10 out of ten, and everybody own a phone nowadays so they are definitely going to use them"
- "I think it would be great for children because with cases you can't see up close but with a picture like that you can, it's a really good idea"
- "Make it better because you have more interaction with objects instead of having to look at the through cases it means you have more fun""
- I can imagine people pumping in each other while walking, it will walk hilarious"
- "Highlight bits of the collection because at the moment they are quite limited in how much description they can have on a object and you have to consult a folder"
- "Sometimes if you are shy and you don't want to go and ask a person"
- "you haven't got enough time and you don't want to have them talking for ages"
- "or also other museums they just don't have so many people around as here to ask so it's good that it is something that you can do by yourself"
- "and kids are on the phone all the time so it would be good to give them something related to a museum"
- "it would bring the museum alive, lost of museum don't offer activities""
- Mentioned that it could be really useful to show artefact which are stored in the storage and cannot be accessed nor consulted"
- "You are more focused on the boxes and walk around and explore more""
- "For kids though I think adults they just want to read"
- "Museums sometimes you just got exhibition and pieces of papers just sticked on the wall telling you a little bit a about it whereas an app like that gives you a sort of, you'll be more excited about it, you know not everyone just one to stand there and read it"
- "I am not tetchy either but I think it would be a very good idea"".
- "Direct them an define things, following what the apps does, you can learn about the object instead of waling around with no direction and then missing something"

- “It's a more interactive way to learn about the objects, more playful than just go around and look at the objects “

AR to deliver knowledge



- "Maybe most people will read it, there are going to be people that are just going to skip over it but at least they will know the object and the object name and it's something”
- "I think it works quite well, because it is a sort of hands on experience. Not everybody is suited to just sit down and reading from a textbook”
- "Yes because you have to think about what you are doing fist and then what you are looking for”
- "It depends on the organisation of the game”
- "It's good to deliver the knowledge and convey the history of the artefact”
- "Is very interactive, and you definitely can learn more, obviously you need to read to find the clues (...) you have to get the five objects and you need to read to get the next one”
- "If you get someone who's actually pay attention to it yea, if you get someone like me who is just like ok right I got the object let's find the other object they are not going to read the information but then I guess the kind of people who usually go to museums they want to read the information so.. if you read it you would like it but then it applies to the other labels in museums as well”
- "I am very easily distracted but if implemented in the right way I think it has got potential (…)""
- Sometimes some people come to museums and they just see shelves and things but if you have to go and look around for things and stuff than you may can get more interested in the moment you find the object”
- “Kids they are probably not going to read to much text, but they wouldn't do it anyway (using normal tools offered by a museum) so you're just increasing the chance they'll to it or at least they remember something from the game”
- "I think you are focused on the next object to find so you wouldn't focus on learning. The information was good, but you'd just be distracted”
- "You're thinking about mum and and then you're thinking about the artefact you were looking for her, that relates to people, there is more of a human element as well”