Dissertation Type: Enterprise

DEPARTMENT OF COMPUTER SCIENCE

MuseumAR

Augmented Reality Guide for Children in Museums

Orsolya Lukacs-Kisbandi

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Master of Engineering in the Faculty of Engineering.

Friday 11th May, 2018

Declaration

This dissertation is submitted to the University of Bristol in accordance with the requirements of the degree of MEng in the Faculty of Engineering. It has not been submitted for any other degree or diploma of any examining body. Except where specifically acknowledged, it is all the work of the Author.

Orsolya Lukacs-Kisbandi, Friday 11^{th} May, 2018

Contents

Executive Summary

A Short Description of MuseumAR

Art is an important part of our culture. We are exposed to it from an early age, but often people find visiting galleries uninteresting and do not understand certain pieces due to lack of education or interest. In order to really appreciate art, children should be taught what to look out for when visiting galleries and museums. This is a real challenge for both museums and parents as it is difficult to capture children's attention and present information in a captivating way.

The aim of this project is to build an application for children to use when visiting museums and exhibitions, turning their visit into a treasure-hunt, drawing their attention to interesting details that would otherwise be missed, using Augmented Reality technologies. This technology is suitable to address the problems mentioned before since it can be used to augment the artwork, bringing paintings to life and providing a means to interact with them. To display the information, a hand-held device (smart phone or tablet) is used. The user looks at the displayed artwork through the camera of the device. Pieces are selected beforehand to create a chain of events called a 'story'. The paintings are recognized, triggering audio recordings, animations and augmentations to be displayed. Animated characters give clues and tasks to children by pointing out interesting details of paintings, guiding them through the museum/art gallery in an interactive and entertaining way. The AR characters are guides for a treasure-hunt where children collect different items from paintings. Audio recordings are used to give instructions and information on which children can be tested as part of the teaching process. The children interact with their surroundings by tapping on the screen and moving the device around.

The application also provides a story creation interface for museum employees; creators are able to add their own story content, triggering custom events like personalized audio recordings using intuitive 'drag-and-drop' components. With this system, museums can create entire new stories based on their own paintings by simply uploading a photo and selecting areas of interest.

Building MuseumAR was an experience that taught me many things about project planning, execution and the importance of user testing.

During this project I learnt:

- How to plan and build a mobile application.
- The importance of agile software development style.
- How to collect and interpret qualitative and quantitative user feedback.
- The importance of an intuitive User Interfaces.
- Potential and limitations of Augmented Reality and Object Tracking.
- The Swift programming language and supporting tools.

The specific goals were to research and understand the needs and requirements of the users —in our case museum employees and children; implement the main features of the application and conduct an evaluation of the app through semi-structured informal interviews.

Supporting Technologies

- ARKit [1] — Augmented Reality library developed by Apple
- Xcode — Integrated development environment for macOS
- Swift Storyboard was used to create User Interface mock-ups
- Blender [1] was used for creating and modifying 3D objects and characters

[1] www.blender.org

Notation and Acronyms

3D	:	Three-Dimensional
AR	:	Augmented Reality
BRISK	:	Binary Robust Invariant Scalable Keypoints
ORB	:	Oriented FAST and Rotated BRIEF
PDA	:	Personal Digital Assistant
SFM	:	Structure From Motion
SIFT	:	Scale Invariant Feature Transform
SLAM	:	Simultaneous localisation And Mapping
SURF	:	Speed-Up Robust Features
SSQS	:	Semi Structured Qualitative Study
TEL	:	Technology enhanced learning
TML	:	Technology Mediated Learning
UI	:	User Interface
VIO	:	Visual-Inertial Odometry
VO	:	Visual Odometry
VR	:	Virtual Reality

Acknowledgements

I would like to thank my supervisor, Andrew Calway for his useful comments and enthusiasm, Julia Carver, the curator of Bristol Museum for her input and for allowing me to use pictures of the displayed artwork for MuseumAR and Carina Burghard and Sue Celik for their valuable advice on how to engage and create content for children.

I would also like to thank my parents, my brother, Adam, and my best friend, Csilla for their endless support and encouragement.

Last but not least I would like to thank Gavin, John, Ross and Tuana for making sure there are no grammar mistakes in my thesis.

Chapter 1

Contextual Background

1.1 Main Objectives

The high-level objective of this project is to provide a means to educate children in an interactive and entertaining way and provide museums with an easy to use interface to deliver an enhanced *educational* experience.

Key Objectives of MuseumAR:

1. Provide museums with a useful tool to enhance children's experience.
2. Enable museums to make use of the latest technologies without any technical knowledge.
3. Enhance enjoyment for both children and parents while visiting a museum.
4. Provide an interactive platform for children to explore exhibitions.
5. Connect technology and fine art.

The unique selling points of this project are that it provides a way for children to directly interact with paintings, offers a fluid user experience by not requiring users to scan any markers and turns the museum visit into a treasure-hunt. In addition, it focuses on the needs not only the visitors but of museum workers, providing them a platform where changes and additional information for the visitors are easily implemented. By combining these aspects —to the best of our knowledge— our implementation is unique.

Main Technical Aims:

- Create an intuitive UI for museum workers to create and modify the content visitors see.
- Implement a framework that allows easy addition of new components to storylines.
- Make use of the latest technologies such as Augmented Reality, animations, object recognition and tracking.

1.2 Motivation

With the development of technology, sources of entertainment such as museums and art galleries have become less popular, especially with young people. The number of mobile phones and tablets has significantly increased in the recent years, and today 95% (1.1) of households in the United Kingdom own a smart phone or tablet. We can assume that the majority of the population is comfortable with using these devices and applications. It is likely that, especially children, will easily learn how to interact with new technologies such as augmented or virtual reality.

Studies conducted on media literacy of children show that they are exposed to and learn how to use these devices from an early age [2]. The number of children using media devices is increasing every year, with 94% of 8-11 year olds owning their own device [3].

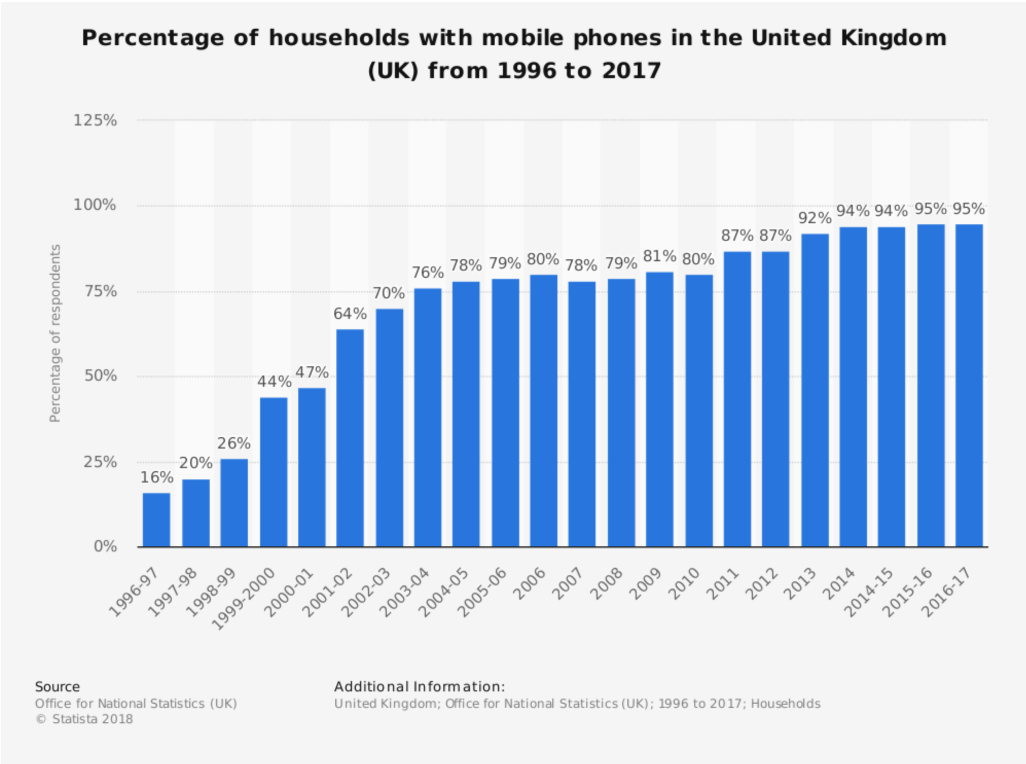

Figure 1.1: [4]

Despite the efforts made by museums to attract young people by involving school groups in creating content about local artifacts and organizing performances, activities and talks[5], the number of children visiting museums is declining. When asked, two-third of parents agree that local museums and art galleries make efforts to engage with children, schools and families [6], but their efforts seem to be unsuccessful.

According to statistics published by the *DCMS*[1], in 2015/16 [7] there was an overall 14.4% decrease in *museum* visits by children (below the age of 16) compared to 2014/15 and a further decrease of 6% in 2016/17 [8]. Only 16% of the overall visitors were children. This is true in most cases, although some museums/galleries such as the *Tate Gallery Group* and *National Gallery* have increased the number of visitors by offering new experiences to them such as interactive tours and family workshops. Figure 1.2 shows the statistics on the number of under-16 visitors at the museums participating in the survey.

Augmented Reality's first big success with the public was Pokemon Go [2]. It successfully convinced children and young adults to go outdoors [9] by allowing them to interact with their surroundings while playing a game. We believe we can harness the interest of the general public in technology to help museums and teachers with getting children interested in arts.

1.2.1 Education

There is a significant amount of prior research on how children learn, what factors enhance their learning, and what are the best methods to motivate them. In this section we will discuss studies conducted on different important aspects of education and how these were motivating factors for building MuseumAR.

1.2.2 Museums as a Learning Environment

Museums are an informal learning environment where children get to explore and voluntarily participate in the learning process. In the past, the role of museums was mainly educational, but in recent years the entertainment aspect has become more important as well. Di Serio et al. [10] explore the effects

[1]Department for Digital, Culture, Media & Sport
[2]www.pokemongo.com

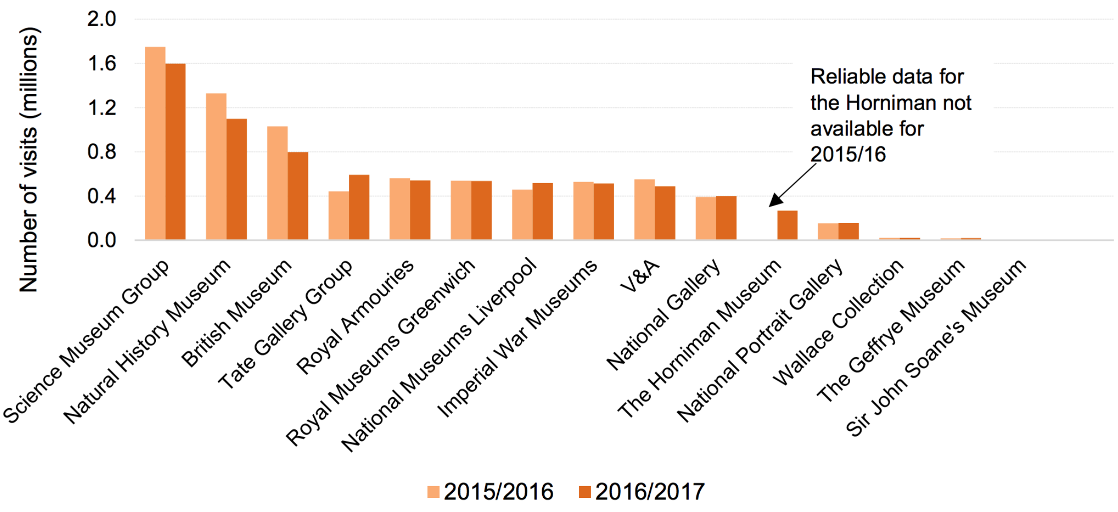

Figure 1.2: Number of child visits 2015/16 - 2016/17 (*image from* [8])

of Augmented Reality on the motivation for learning in middle-school students and conclude that using AR, despite limitations owing to the immaturity of the technology, has a positive impact. Augmented Reality offers different degrees of immersion and interaction that can help students engage in learning activities. This assumption was tested with students who took part in a compulsory visual art course. First, information was presented to them using a slide show. Later, two similar classic artworks were presented using AR to communicate information in a multi modal way using text, audio, video files and 3D models. Their conclusions were that students' motivation for learning increased when using AR. Out of the four factors that describe motivation statistically (Attention, Relevance, Confidence and Satisfaction [11, 12]), significant changes were shown for attention and satisfaction.

1.2.3 Children's Behavior in Museums

Anderson et al. [13] examined the way children behave, engage and interact in museums in order to understand the ways in which they learn to appreciate museum experiences. This process has multiple aspects: cognitive, aesthetic, socio-cultural and collaborative, each playing an important role in learning. They concluded that children were more likely to remember information when they interacted with the displayed objects or could trigger some event by performing a task [14, 15]. Furthermore they reported that children frequently recalled and accurately described museum experiences that were presented as a story and concluded that tactile experiences as well as storytelling provided context with children's own culture and were powerful mediators for learning.

1.2.4 Importance of Interaction with Surroundings

The results presented in Pallud et al. [16] show that use of interactive technologies results in higher levels of cognitive engagement. These results are supported by studies suggesting that psychological factors such as enjoyment [17, 18, 19] and authenticity [20] have a considerable effect on children's learning. They also address the question of how technology-mediated learning (TML) affects learning outcomes. They take into account the ease of use and interactability of the software, which is especially important for systems that rely on communicating information [21].

Studies have shown that interactivity enhances cognitive engagement. Furthermore, the more interactive and information-rich the environment seems to people, the more likely they are to participate and be involved [16, 22]. Interactivity is one of the major factors that influences the efficiency of learning practical skills and advanced knowledge, hence AR/VR technologies would improve children's learning [23].

In another study presented in Charitonos et al. [24] the role of new technologies — such as social media — in meaning making and learning process is discussed. 'Meaning making' stands for the process of understanding, retaining and revising knowledge developing more nuanced and complex understanding.

The focus lays on the importance of engagement with the environment and each other. They state that through interaction with the displayed artwork individuals make connections with their own experiences and prior knowledge that they discuss with their peers in person or through social media platforms.

1.2.5 Gamifying the Learning Experience

Lin et al. [25] conducted a study on game-like teaching methods on museum websites and results showed that this approach resulted in more enjoyment and better learning outcomes. They concluded that game-based software can help users learn complex information while enjoying the learning process more than text based learning. In addition, they stated that 'game-based online learning features attract more attention, provide more positive experiences, fulfill more desires, and create more online enjoyment experiences than traditional text-based web pages'. When participants found the interaction with technology intuitive and interactive they experienced higher levels of cognitive engagement hence becoming more involved with the exhibitions. Their final conclusion is that in order to enhance the learning experience, the technology used has to be intuitive in order to provide an interactive experience resulting in emotional responses such as excitement, curiosity and enjoyment.

It is important to use these technologies appropriately since, as much as they can enhance learning, they can also distract from the information. Dror et al. [26] explores the effects of technology on learning as well as its limitations and potential risks — such as cognitive load — introduced by it. They state that technology can help us shape the information and reduce the cognitive load making learning more efficient.

> It should be clear that technologies offer a great and powerful tool which can enhance learning. This good aspect needs to be considered and understood in light of the limitations of Technology Enhanced Learning (TEL), and the potential bad outcomes that are possible by incorrectly using this tool. [26]

1.3 A Possible Solution

MuseumAR is a new approach for solving the aforementioned issues as it could fill the gap between arts and technology. By providing means for children to interact with paintings we can ensure that they stay focused and see museums as a source of entertainment, learn to appreciate classic art and further develop their knowledge.

Informal interviews with staff at museums suggested that apps and other means of providing information using phones and tablets were difficult to set up. In all cases, they relied on external input from other devices such as sensors (e.g. iBeacons, a proximity sensor used to determine the location of a connected device using low energy Bluetooth) and QR codes. These introduce additional purchase and maintenance costs and provide a less fluid experience for users by requiring them to perform the same task over and over again (e.g. reading in QR codes multiple times). Technology writer *Om Malik* stated that 'QR codes have always been a kind of half-measure, a useful but inelegant transitional technology; the ultimate goal is Augmented Reality.' [27]

Augmented Reality has the potential to be the solution to these problems by making use of the latest object recognition, machine learning and image tracking technologies providing a fluid and intuitive experience.

MuseumAR relies only on the built-in technologies of an iPad/iPhone, it does not need any external input other than the video stream from the camera capturing the environment and can provide real-time performance.

1.4 Museums and Technology

It is debatable whether technology should be introduced in museums. Museums are often viewed as a safe-haven from technology and a place for connecting with our past and history. Some believe it is unnecessary or even inappropriate to introduce technology in these environments as it would distract visitors from appreciating the art and divert their attention from the displayed pieces to the app itself. This would be the case especially if the app offers multiple features or has a storyline.

On the other hand, technology would be a good way to teach people about arts and might attract more visitors as it provides means for them to interact with the artwork (Figure 1.2.4) and, as mentioned in the previous section, would improve the learning experience in the museums (Figure 1.2.2). Given the latest advancements in mobile technologies, museums which ignore the opportunities are missing out on the aforementioned benefits.

Research conducted on interactive VR for education purposes—presented by Walczak et al. [23]—shows that people have a positive reaction to new technologies presented in a museum-like environment. There are also several studies on how AR is a suitable way to communicate information on the displayed artwork [28, 29].

In many situations the presence of interactive technologies can be a reason for social isolation and lack of inter-group interaction. This concept is explored by Heath et al. [30], who analyse video-based field studies conducted in museums and art galleries. They conclude that interactive devices in museums discourage inter-group interactions.

Another motivating factor to build this application was that museums find it difficult to include new technologies in their curriculum, the main question being not whether or not these technologies should be used but how they should be used. Many museums are still skeptical when it comes to technologies—some believe it takes away from the experience instead of adding to it, although a shift of opinions regarding this can be observed in the last few years [31]. The role of museums is transforming from being a means to present information (as it was in the past) into being a potential site for new experiences.

The way new interactive resources are used in museums is discussed by Lehn et al. [32] based on two studies both conducted in art museums/galleries. One introduced Personal Digital Assistants and the other an information kiosk. This article concludes that new technology can be interesting and provide valuable information. Visitors also appreciated the novel approach to presenting this information. An issue with Personal Digital Assistants was the fact that they displaced the original object and observing the displayed artwork was not necessary anymore. This problem is addressed by MuseumAR since the users are looking straight at the artwork through the camera of the device.

Tillon et al. [33] address several questions about the use of Augmented Reality in museums. They explore the impact of virtual guides on the visitor experience, the way it transforms the activity of visitors and the way they interact with technology. They conclude that the displayed augmentations should be clues that guide the observer and encourage them to have a closer look at the paintings instead of displaying any other content. In MuseumAR we are doing exactly that, since we display only small augmentations/hints on the paintings themselves, and use audio to communicate the information. Augmented Reality is a useful tool to redirect people's attention to the significant details in an unintrusive way, potentially changing visitor's perception. This is especially beneficial when it comes to children, since they need special guidance because of their lack of knowledge.

In recent years museums have been able to access increasing funding for introducing digital technologies as part of their exhibitions. The National Gallery's 2017 annual review [34] shows that there is £3.2 million available funds to support projects that inspire learning and engagement such as our app Figure 1.3.

Based on the aforementioned arguments, we can conclude that the use of an AR guide would improve the visitor experience. MuseumAR not only provides a natural experience for the users, but also a completely new experience by turning a museum visit into a treasure hunt where children can explore the hidden corners of exhibitions and learn new things in a playful setting at the same time providing the means to easily input information for museum employees.

1.5 Previous work

Antoniou et al. [35] collaborate with the UCL Grant Museum of Zoology, using micro-augmentations to improve the visitor experience. These micro-augmentations are minimal stimuli visible to the visitor and barely audible sounds, aiming to stimulate visitors' curiosity rather than providing information or teaching. This method would be unsuitable in our case as we would like to enhance learning and engage with the display, not only observe.

The British Museum employed Augmented Reality in multiple projects, targeting different groups of visitors as well as different parts of the exhibition [36]. One project relied on markers and revealed content on them; another used location-based AR for navigation and another one displayed virtual art in the gallery. All these projects were built with the purpose of providing new ways to educate children about arts and all had different difficulties. There were problems reported with wireless signal reception/3G reception. This won't be a problem for us since the app works offline, doesn't require internet connection to operate and the features do not rely on markers other than the paintings themselves. Other museums, such as the *British Museum* experimented with marker-trigger AR interactions as well [37]. These proved

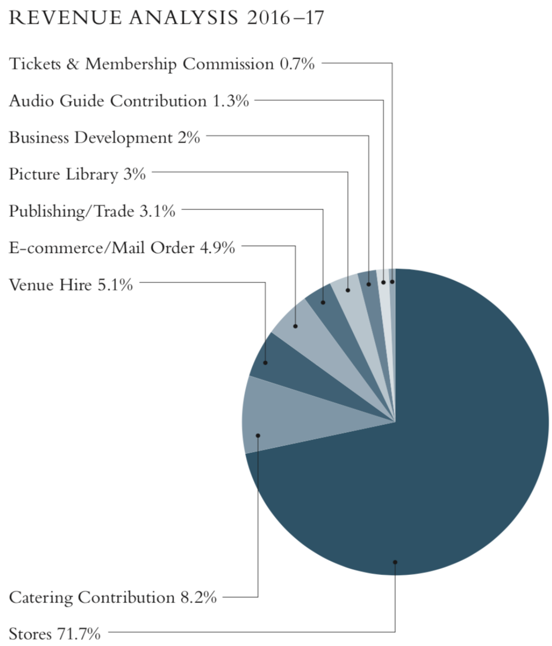

Figure 1.3: National Gallery Revenue 2017 [34]

to be impressive at the time, but since the technology has evolved since 2010 and it is now much easier to implement tracking relying only on the user's surroundings.

In their review Ding et al. [38] discuss the role of AR in museums and the benefits of introducing digital technologies into a traditional museum environment. The Chicago00 [39] project uses Augmented Reality to animate the Chicago Riverwalk between Clark and LaSalle Streets showing events that occurred during The Eastland Disaster. This uses location-based Augmented Reality and it provides a unique experience for people visiting the city.

ArtLens is an award winning app developed at The Museum of Cleveland that allows users to scan paintings that get automatically recognized using an internet connection. Information about the paintings is acquired from a server run by the museum and it is displayed as text on the screen. Visitors can choose artworks and plan a path through the museum making sure none of the items of interest will be missed. The app uses iBeacon technology, that has clear limitations, such as not knowing what exactly the user is looking at and using a location-based event trigger instead of an image recognition based one. It also introduces additional costs of buying and maintaining the beacons.

1.6 Target Audience

The app will be suitable to use for children from primary school on. MuseumAR does not limit the age range, whether or not the app is suitable for a certain age-group depends on the content introduced. Although it was designed with children in mind, it could be used to communicate information to adults as well.

Chapter 2

Technical Background

2.1 Augmented Reality

Augmented Reality is a term used since the early 90's referring to systems that allow to present in real-time both virtual objects and real elements from the physical environment that users can interact with[40]. The main purpose of Augmented Reality is to create new and better forms of interaction between the real and the virtual world, providing the user with additional knowledge. This technology is highly suitable for the task at hand as it allows any media content, such as images, videos, animated virtual objects and augmentations to be layered upon real environments, this way enabling museums to bring collections to life. AR makes use of the physical environment of the user, blurring the line between what is real and virtual. By doing so, it enhances the user experience and provides a new Natural User Interaction (NUS). This is comparable to touch, gesture or eye tracking, creating an interface that can be interacted with in an intuitive manner. Visitors can look straight at the artwork through the camera of a phone, tablet or a specialized headset while listening to audio recordings. They can be shown different augmentations being displayed without intrusion, allowing them to pay close attention to details in the exhibit at the same time.

In the past using Augmented Reality was challenging both because of hardware [41] and software limitations, but thanks to the fast development of technology, lately it has become much more accessible for everyone. Advancements in this area made software-only solutions possible for AR that work even on mobile platforms enabling access to new environments in situ such as museums and art galleries.

This is not the only use of Augmented Reality, this technology is used in other fields such as health care [42, 43] or data visualization [44].

The flow diagram of an interactive Augmented Reality based system has six main components as seen on the figure below (Figure 2.1) [45]. The images captured by the camera of the device are processed using computer vision algorithms and features from the images are extracted to be used for tracking. Rendering is the most commonly used method for augmenting reality, but audio and haptic feedback can also be used to provide a unique user experience. The equipment used for displaying AR content can vary from head-mounted displays to mobile phones or other devices. These devices are either light enough that users can carry them in their hands or they are part of spatial displays, such as desktop monitors or projected displays.

> Augmented Reality (AR) describes user experiences that add 2D or 3D elements to the live view from a device's camera in a way that makes those elements appear to inhabit the real world. ARKit combines device motion tracking, camera scene capture, advanced scene processing, and display conveniences to simplify the task of building an AR experience. [1]

2.2 Tracking & Movement

Motion tracking is a key requirement for Augmented Reality applications, as it enables developers to create a convincing composite, where virtual objects behave the same way as real ones would. Tracking plays an important role in obtaining the illusion of virtual elements being part of the real scene [46]. To align virtual objects with real ones, it is essential to know the relative position of the device to the scene. By accurate motion-tracking we can ensure that objects remain at the same location they were attached to.

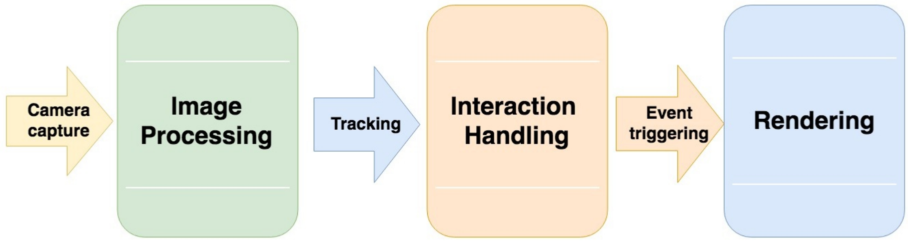

Figure 2.1: Flow Diagram of a vision-based AR System

There are different tracking techniques used in AR that are suitable in different situations. These include sensor-based tracking (e.g. magnetic, acoustic), vision-based tracking or a combination of these two. Sensor-based tracking has many advantages such as speed and robustness against changes in the environment but it is lacking the accuracy of vision-based tracking [47] that is able to detect small details of the environment.

In this section we are focusing on vision-based tracking. Vision-based tracking refers to dynamic sensing and measuring of the spacial properties of the environment [45]. There are two types of vision-based tracking: marker-based and marker-less. Marker-based tracking uses unique patterns printed and displayed in the environment to localize. This method is more reliable than marker-less tracking and it has high accuracy even with a low resolution camera. On the other hand it requires changing the environment, interfering with the user experience and making the experience seem less realistic. Marker-tracking is a special use-case rather than a general-purpose solution when it comes to Augmented Reality.

2.2.1 Marker-less tracking

Marker-less tracking relies on the natural features from the environment to determine the camera location and position. These features, called interest points, are represented by descriptors. These descriptors are highly distinctive, scale and rotation invariant and are robust against change in 3D viewpoint, affine transform, noise and changes in illumination. The extraction of appropriate features for the descriptors is crucial[48]. These features must be well localized in both spacial and frequency domain, reducing the likeliness of disruptions caused by occlusions, clutter or noise [49]. To extract the features an approach called SIFT is used which stands for Scale Invariant Feature Transform [50]. In every image a large number of features are extracted using a cascade filtering approach. This method ensures that more expensive operations are applied only to regions that pass an initial test [51].

2.2.2 Visual Odometry

Visual Odometry (VO) is the process of estimating the devices motion based on a stream of images of the environment [52]. VO is a specific case of Structure From Motion (SFM) tracking that addresses the problem of 3D reconstruction of both the motion of the camera and the structure of the environment from a stream of images by examining changes introduced by motion in consecutive frames [53]. This estimation is time consuming and requires a lot of computational power, because of this, it is often performed offline. This is not an option when it comes to Augmented Reality, instead VO focuses on estimating the 3D motion of the camera sequentially as a new frame is captured in real time.

Nister et al.[54] describe a real-time method for deriving vehicle motion from monocular and stereo video sequences. Their method for localisation is

- tracking features over a number of frames
- triangulating the tracked features into 3D points

- repeat iteratively relative to the first identified 3D point

To avoid error propagation, they decide on a window within they reference to previous 3D points, after starting a new frame, the location gets re-estimated similarly to the very first estimation.

2.2.3 Visual-Inertial Odometry

ARKit[1], Apple's framework for Augmented Reality uses Visual-Inertial Odometry (VIO)[55] to create a correspondence between real and virtual spaces. VIO combines camera sensor data — Visual Odometry — with Core Motion data —Inertial Odometry. Data form the camera is used to identify unique feature points and detect the devices movement relative to these landmarks. Inertial odometry provides additional information to improve the accuracy of the devices understanding of how it is moving in space. These two inputs allow the device to sense it's own movement within a defined spatial location with high accuracy and without requiring any additional calibration. There has been significant amount of research about VO as it proved to be faster than SLAM and has promising results in accuracy. Some methods combine it with depth estimation [56], others with other sensors such as a gyroscope.

For our project, we used Apple's built in monocular VIO as it proved to be reliable for localisation, velocity and rotation estimation [57]. VIO doesn't create a 3D model of the space, only tracks the movement of the device by observing the movement of reference points in consecutive frames estimating the relations between these points through 3-dimensional projective geometry. ARKit world tracking uses six degrees of freedom(6DOF) (Figure 2.2) to track the environment because it allows the placement of realistic virtual objects in the real world and interaction with these objects [58]. The camera's position estimation must be highly accurate since even small errors will result in inconsistency in the world frame that the user sees.

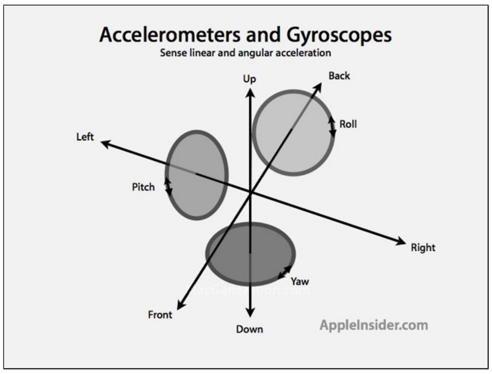

Figure 2.2: Inertial odometry - Following the devices movement in 6DOF

2.3 Image Recognition

Image recognition is one of the most important technologies for this project. It is used to recognize paintings that are the base of MuseumAR. In this chapter we will discuss different techniques that are used to perform image recognition and hypothesize what is happening in the background when using ARKit, since this information is undisclosed.

Bay at al. [59] describe SURF(Speed Up Robust Features) algorithm and its application for the recognition of artwork in museums. SURF is a local feature detector and descriptor that operates in real time. They develop an application that can be used as an interactive guide in museums, it offers robust recognition of displayed artifacts under varying environmental conditions and operates on low-cost hardware.

Dalens et al. present a method for object retrieval to perform painting recognition matching paintings in the user's environment to paintings in a database [60]. Their aim is to perform this operation in real-time with no noticeable latency in order to make it suitable for Google Glass. They compute image descriptors of the images in the database, store that information offline and match features to the images captured by the camera on the user's command on Google Glass. After experimenting with several feature extraction methods (SIFT[50], SURF[61], ORB[62] and BRISK[62]) they conclude that using a hierarchical bag of ORB features has the highest accuracy. This method works well, but it requires the user to signal when to start the image recognition, interrupting the user experience.

Fockler et al. [63] uses a single-layer perceptron neural network to achieve light-weight object recognition that can be performed real-time on phones with a recognition rate of 90%. Ruf et al.[64] performs experiments to compare SIFT and SURF with the aim of building a server-based mobile application that can operate real-time recognizing images from the user's camera and operating as a guide. They conclude that to achieve the expected speed the best solution is to use low-resolution images, for which SIFT seems to perform better. ARKit has a feature that allows the detection of 2D images in the user's environment using ARReferenceImages. The coordinates of the recognized images can be retrieved and used as an anchor for AR objects, the recognition of images can also be used as a trigger for events. These features are all demonstrated in MuseumAR.

For images to be reliably recognized and their position accurately estimated it is important that the physical dimensions of the reference images are provided by either the program or the user. Without this input the scale of the AR content will not match the expected dimensions. The image detection is sensitive to changes in lighting conditions — glare and darkness have a significant impact on the image detection. Images with more content are much easier detected, while single-color images or images with little content might impose difficulties for the recognition algorithm. The recognition and tracking works significantly better when the background of the images is not uniform, since meaningful reference points can be extracted.

When experimenting in real-life conditions, the image recognition performed well when the images had some distortions or were missing some detail, but it was sensitive to extreme lighting conditions [65].

Since the recognition works in real-time we can hypothesize that features from the reference images are extracted when the application is loaded/compiled and these features are used for detection making the recognition less computationally expensive.

2.4 Rendering

Image rendering is an important aspect of Augmented Reality. It is essential that the video feed from the camera is processed and virtual objects are rendered real-time. The rendering of objects relies on information from both the camera feed and other sensors of the device. It's accuracy is crucial as the virtual objects must perfectly align with the real world giving the illusion that they are part of the environment.

When tracking the scene, ARKit performs light estimation resulting in photorealistic rendering of virtual objects.

Chapter 3

Project Execution

3.1 Description

In this section we describe how the app works, its main functionalities, who are we targeting with it, and what the most important aspects are. The app has two main parts one user interface that is intended for the museum curator/staff and another one for the visitors.

Currently the app can have one storyline that is input by the museum staff. The main idea is that the museum curator/workers add paintings one at a time to the story. These paintings are recognized, triggering the appearance of an AR character that when tapped gives a task using audio recordings; on completion, the user receives a clue to find the next painting.

Figure 3.1: Starting screen

3.1.1 UI: Museum staff

Start screen

On the Start screen Figure 3.1, the button on the top left corner takes the user to the museum staff UI where the content can be edited.

Gallery

On the *Painting menu* screen Figure 3.2 paintings can be added edited and deleted. If the user taps on any of the paintings, they are taken to the painting information screen (see Figure 3.3).

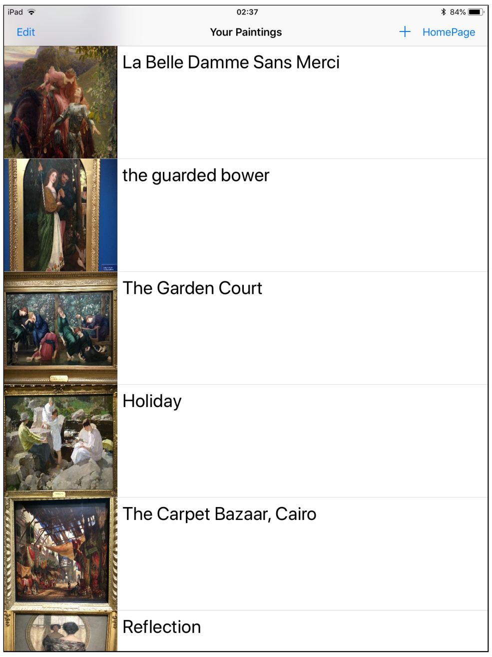

Figure 3.2: Painting menu (Paintings [66, 67, 68, 69, 70, 71])

Painting Information

On the Painting Information screen (Figure 3.3), all the information about a particular painting is displayed, including the elements attached to the storyline (audio recordings, the physical size of the painting, objects of interest on the paintings and any additional information that can be seen by the museum curator). Existing paintings can be easily changed, a new picture can be selected by simply tapping on the image and selecting the new input from the devices' gallery. Audio recordings can be played back or replaced with a new recording. Areas of interest on the paintings can be selected and made part of the story (Figure 3.5).

New Painting

By tapping on the + symbol on the *Gallery* screen Figure 3.2, the new painting screen appears in Figure 3.4. The painting title, the picture of the painting and the physical size are required fields, without them the *Save* button is disabled. The reason for this is that these are the minimum requirements for a painting to be detected. The physical width of the painting is necessary because the image recognition relies on this information to calculate the relative size of the attached AR objects to the physical world and to support accurate anchoring of these objects.

Personal notes and descriptions can be added to the painting, for example the role of this particular painting in the story, a transcript of the audio recording or other useful information for the curator. This content won't be visible to the visitors.

Tapping on the button *Add objects to be detected* the object selection screen is triggered (Figure 3.5). Here the user can draw a bounding box around the areas of interest. These boxes will be used to create touch-sensitive objects overlayed on the paintings that will trigger events. This is a key part of the app, because this allows objects to be to be 'collected' from paintings. These objects can be changed later from the *Painting Information* screen.

The *Record Task* caption is suggesting to the user that an audio recording should be made which tells the visitor the task they need to complete to get the next clue. These recordings should also be used to provide the user with interesting information about the paintings.

The *Record Next* caption is suggesting that an instruction about the visitors' next stop should be recorded, giving a hint about which painting they should look for next.

Figure 3.3: Painting Information (Painting: [68])

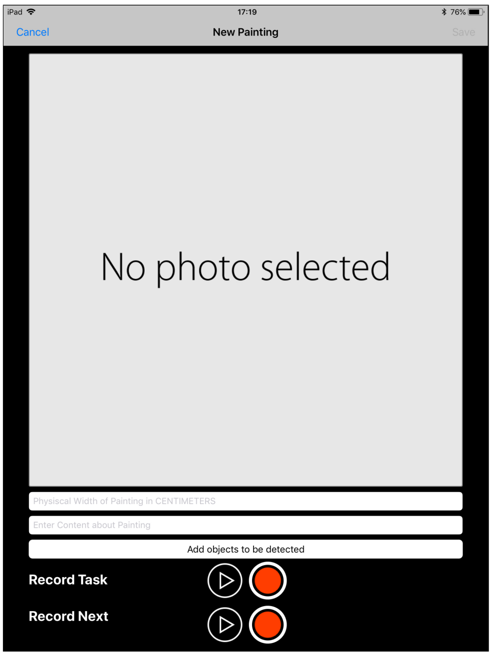

Figure 3.4: New Painting

Figure 3.5: Select objects (Painting: [68])

3.1.2 UI: Visitors

Visitors see the same starting screen as the museum staff (Figure 3.1). After tapping on the *MuseumAR* logo the session is initialized, performing localisation, instantly allowing the user to start exploring the exhibit.

Visitors are asked to walk around looking at paintings. At every painting an AR character appears giving information about the painting to the visitor and instructions on what objects they should find an tap on to get a hint on where to go next.

3.2 Design

The design of the user interface and functionality of MuseumAR was an iterative process that proved to be challenging because there were multiple requirements that had to be met.

Unofficial interviews were conducted with teachers, museum staff, other students and children. These interviews were used to ensure that all users' needs and expectations are met by the application.

The main points that guided the design were:

- Easy-to-use interface
- Intuitive interactions
- The Museum Curator's needs
- Educational value
- Content suitable for children
- Overall positive user experience for both children and museum employees

Studies conducted in museums about the use of AR have shown that the timing of information delivery is crucial when it comes to learning [72], since users seemed to pay more attention when they were experiencing higher levels of enjoyment. We used the findings of this study when planning the timing and type of feedback children receive during the visit.

3.2.1 Use of Audio Recordings

The suggested use of audio recordings is described in the following section.

When a painting is detected an AR character appears. When tapping on the character an audio recording is played providing information about the artwork then explaining the task at the given painting. When the task is completed, the user is given more information about the painting they are looking at before getting a clue about where to go next. By communicating information even after the task is completed it can be ensured that the user is paying attention carefully since they are expecting that after every successful task the next clue will be provided.

Hamilton et al. [73] mention that by constructing a story around the visit, a more engaging experience is created, that is more likely to transfer knowledge than only communicating bare facts.

Other Alternatives

Displaying text was an alternative for providing information, but audio recordings better suited the task at hand. The main reasons for making this decision were that text would have taken up a significant part of the screen, in addition would have made it difficult to target young children since they can not read.

3.2.2 Early Field Studies

Informal interviews were conducted with teachers, museum visitors as well as museum workers at the design stage of the app in order to explore what people's expectations are from an interactive AR guide.

Introducing technology in museums is a delicate matter given the controversy around it and the subjectivity of the displayed artworks. Because of this, before designing the app multiple museums were visited and employees were interviewed to find out more about their expectations from an app designed to use in their exhibitions by children.

Interviews with Museum Staff

Many museums already had an app or other interactive technology related displays that offered some level of guidance for visitors. The interviewed staff thought that their museum would benefit from an application that would specifically target children giving them a new way to communicate information and engage them in exploring the displays. They seemed enthusiastic about the choice of targeting paintings, especially fine art, since these artifacts are the hardest for children to relate to and engage with. They were unaware of the capabilities of the current state of the art technology and thought that the capabilities of our demo (recognising paintings, displaying their title and showing an AR object) was surprisingly fast and easy to use.

Some interviewed museum staff were skeptical about the value the app would add to their exhibitions. Surprisingly, the most critical people were the ones who worked with children, although when they were shown a working prototype they said that the app could be beneficial for the museum. Their main concerns were not about the technology or the distractive element of gamifying the visit but about the content of the app.

This was not an issue for us, as our design allows museum employees to create and curate their own content. The use of the app would require input from the museum curator and an educational consultant to build the experience. We intended to provide the tools for them to create an interactive experience, but not the content.

Interviews With Other Participants

Multiple informal interviews were conducted with teachers, parents and children in order to find the balance between educational content and entertainment. The final conclusion was to create a treasure-hunt style game where children will receive instructions from a character that accompanies them through their visit. At every selected painting they get a task, on completion they receive the next instruction.

Every interviewed person had a different idea of what the app should provide, some children wanted to compete against eachother, others preferred to explore the museum in their own pace. Some teachers suggested that competition would motivate them to pay more attention; others said that it would result in disorder if they were on a school trip. We decided that at this stage of app development, we can't ensure that children's paths don't cross causing clusters around some displayed pieces, hence it would be best to avoid making the treasure-hunt an open competition.

After telling teachers about this idea, they suggested making the app available not only for museums but also for classrooms and school trips. This is a possibility we could further explore in the future. Since the application's design is intentionally simple, it could be easily adapted to accommodate the requirements of using it for general educational purposes.

Issues and Limitations

We are addressing the issue of social isolation in the app by introducing tasks that include instructions for discussing certain aspects of a painting with other visitors. We were also planning to introduce a scoring system that would tell children how well they did during the treasure-hunt, encouraging further discussion about the tasks even after leaving the museum in an attempt to find out what others did and how did they get a different score. This would also motivate children to revisit exhibitions, next time achieving a better score.

Content generation for the app is also a significant challenge. Even museum curators struggle with this because art is extremely subjective and interpretation bias should not be introduced or should be avoided as much as possible. For this reason, the content of the app shown to children should not be viewed as the final version of an education treasure-hunt game but a means to create a story that can be changed based on feedback from professionals and users.

Another limitation is the quality of object tracking. During our study we obtained feedback on object tracking accuracy. Some participants stated that they believe the tracking requires more improvement as sometimes the content was jumping around due to lack of meaningful features in the camera view.

3.3 Implementation

3.3.1 Comparing AR Frameworks

There are two main frameworks released for creating Augmented Reality applications for mobile devices: *ARKit* and *ARCore* [74]. Both of these have limitations when it comes to hardware, as both are only compatible with devices with specific hardware requirements. Both use sparse maps to create world maps and understand space. ARCore uses a process called *COM* —Concurrent Odometry and Mapping— to understand and map the world while ARKit uses VIO —Visual Inertial Odometry— to do the same thing. When it comes to mapping, ARCore performs better, because it is tracking more features points than ARKit [75]. On the other hand, ARKit is more accurate when it comes to distinguishing floor and wall. Both these SDKs are expected to remember the places AR objects are attached to even if the user moves around or even leaves the scene and comes back later. From this perspective ARKit performs better [76]. Both these frameworks are in early stages and there are no significant differences between them. In conclusion, ARKit performs better when it comes to tracking and more devices are supported, ARCore is better at mapping and recovery of the session.

In either case, museums will have to acquire devices with compatible hardware if they would like to use MuseumAR in their exhibitions.

3.4 ARKit

ARKit is Apple's Augmented Reality session-based framework that has different configuration options. The session runs at 60 frames per second and uses Visual-Inertial Odometry, features extracted from the video frame and built-in motion sensors for marker-less motion-tracking. ARKit allows developers to build unique AR experiences and deploying them on iPhones or iPads.

Some of the key classes and methods used in the project from ARKit documentation:

- ARSession: object that manages the main elements for creating the AR experience such as reading data from the motion sensors and camera and preforming analysis on the images.
- ARWorldTrackingConfiguration: class that tracks the device's movement with six degrees of freedom (6DOF): three rotation axes (roll, pitch, and yaw), and three translation axes (x, y, z). Thanks to this type of tracking, the user has the illusion that the object is occupying a 3D space and can be observed from multiple angles (Figure 3.6).
- ARAnchor: a real-world position and orientation anchor that can be used to place virtual objects at the selected point.
- ARReferenceImage: to accurately detect the position and orientation of a 2D image in the real world, ARKit requires preprocessed image data and knowledge of the image's real-world dimensions. These images can be added to the project resources or dynamically as one of the apps' functionalities.
- ARLightEstimate: estimates the light of the scene based on the video feed and applies the same lighting to the virtual objects.

Because ARKit uses one camera vision, the size of the paintings had to be given as a parameter. This was used for depth estimation (accurate tracing). 'The disadvantage is that motion can only be recovered up to a scale factor. The absolute scale can then be determined from direct measurements (e.g., measuring the size of an element in the scene).' [53]

Figure 3.6: Demonstrating tracking with ARKit [77]

3.4.1 World Tracking

As mentioned in Chapter 2.2, ARKit uses VIO for world tracking, a method that combines information from the motion sensors and the extracted features from the scene captured by the camera. In addition, world tracking also analyses and understands the contents of a scene; it detects vertical and horizontal surfaces and allows different interactions with these surfaces such as changing their texture or placing virtual objects on them. Word tracking works best if the device is moving since localisation relies on motion. We included in our implementation a pop-up header that tells the user if tracking is appropriately configured, letting them know when the device is ready to be used.

Developers can configure how the tracking should work by changing the AR configuration. The framework offers three options: a configuration for world tracking, one for only orientation tracking and one for face tracking. These three have different levels of accuracy, speed and functionality. Face tracking focuses on finding faces and recognising expressions, orientation tracking only tracks the device's orientation not its location in the environment. World tracking on the other hand, aims to localize the device and find its orientation.

3.4.2 Limitations

There are limitations when it comes to world tracking, since the features rely on the physical environment, which is not always consistent and is difficult to track in real time with high precision.

ARKit analyzes the images captured by the devices camera and extracts feature-points that are tracked and differences in their position between frames is used to calculate the devices motion. This method is error-prone for several reasons. If there is not sufficient light in the scene, there will be no reliable feature points. In case the environment has a repeating pattern (such as a chessboard) or it has only one color (white wall) the quality of tracking is much lower.

If the lighting conditions are poor the tracking quality is reduced since there are no meaningful features in the images this is the case when the background has uniform color as well [78].

While using ARKit we observed that in case the device loses tracking, it updates all feature points resulting in a jump of the frame. Another important limitation is that the framework is only supported on the latest iOS devices with an A9 or later processor.

3.5 3D Virtual Objects

The objects displayed during the AR experience were created using Blender [79] or downloaded [80] royalty free.

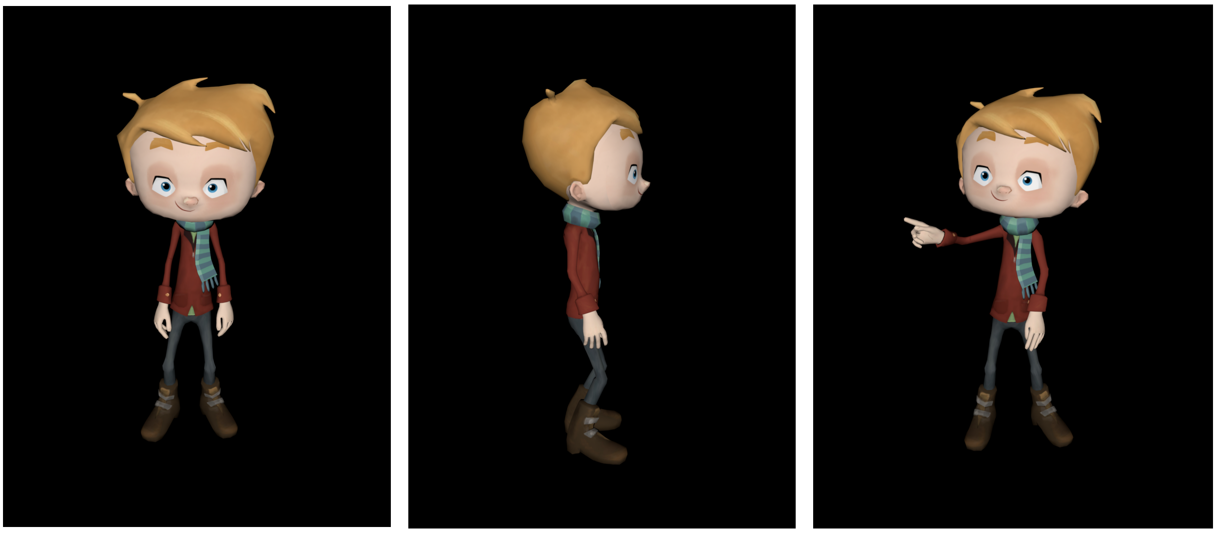

Figure 3.7: Animated Main character [81]

The main character is animated, it is capable of simple gestures such as pointing, waving, looking around and others.

Other objects were also used in the stories created to enrich the experience or to illustrate objects on certain paintings. For example, in case a story has multiple scenes, the 'Medieval house' (Figure 3.8) could be used to indicate to the user where they are in the imagined world. The rose (Figure 3.9) and the knights (Figure 3.10) could be used as clues for the next painting.

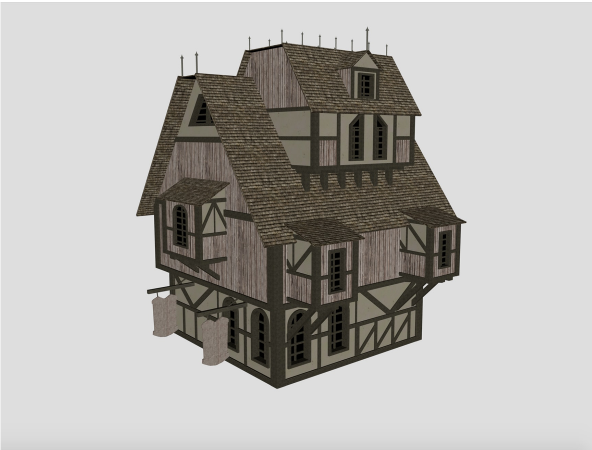


Figure 3.8: Medieval house [82]

Figure 3.9: Rose [83]

Figure 3.10: Knights [84]

Figure 3.11: Other 3D models

Chapter 4

Critical Evaluation

Because the main purpose of MuseumAR is to enhance the experience of both museum staff and children, the most suitable evaluation method was through user testing.

Feedback was obtained in the following ways:

- Interviews with primary school teachers
- Interview with *Julia Carver* — Curator of Visual Art at Bristol Museum and Art Gallery
- Semi-structured interviews with museum employees
- Informal interviews with 10-14 year old children
- Semi-structured interviews followed by a Likert scale questionnaire with other visitors

We used Semi-structured interviews to gather verbal information from participants on their perception of the app and the experience as a whole. In the case of visitors in the museum, after a short discussion participants were also asked to fill in a short questionnaire about their insights and thoughts, gathering not only qualitative but quantitative data. They tried both UIs and gave feedback on the whole experience.

4.1 User Testing

4.1.1 Feedback from Museum Employees

We conducted a small study (six people) involving museum staff.

After the first round of user testing in the museum we realized that the app could be used not only for paintings but for 3D objects as well as long as these objects had a description displayed next to them, the only specific requirement would be that this text would be used as a trigger for the virtual character and there were no augmentations possible on the objects themselves, but information still could be communicated through the character. In this case the task in the treasure-hunt would be to find the given object and look at the description, after listening to the information about this object, the text could be overlayed with a clue for the next artifact.

At the final stage of the application, we conducted an interview with *Julia Carver*, the curator of *Bristol Museum and Art Gallery*. She gave feedback on the UI and shared her thoughts of MuseumAR saying that this is a challenging problem. She thought the interface was intuitive and would meet the expectations of someone adding content to the app. She stated that she would be interested to see further user studies conducted with the app. She agreed, that in fact children are hard to engage during museum visits and she could imagine people using MuseumAR at the Bristol Museum. She suggested that this app could be used to engage adults with modern art as well, since they often struggle to understand more abstract pieces. She particularly liked the idea of creating stories involving the paintings and the option to interact with them.

4.1.2 Feedback from Children

During informal interviews conducted with children aged 10-14 they shared their opinion about MuseumAR and the idea of an interactive guide in museums.

The setting for the interviews was informal, children used the app to look at three paintings receiving instructions from the character on what their task is. Hints on the paintings were displayed giving a subtle indication on where to tap.

The overall feedback was positive even from older children, although the demo storyline was not designed for them but for their younger peers. Most of them stated that it was easy to understand how to use the app, they remembered information mentioned during the session, found it easy to follow the instructions and had enjoyed using the app.

Their favorite parts were:

- Being able to interact with paintings.
- Looking for objects and details on the paintings.
- Looking for the smaller objects that were more hidden.
- Learning new information about paintings.
- Sense of accomplishment when completing all tasks.
- Looking at the paintings through the camera and receiving information about them.
- Interaction with the AR character.
- The ease of use.

They suggested changes such as giving more information about the paintings, using a smaller device to display content on (we conducted the interviews using an iPad, but the app works on iPhones as well), some tasks were too easy, they would have preferred not to have hints on where to tap and they would have liked to see more character animation.

When asked what would they add, they said that a quiz at the end of the visit would be fun and could be used to give them scores. They also suggested that a score based on task completion time could be used. We did think of this when designing the project but decided not to implement it because of concerns that children will race around trying to complete tasks as fast as possible, while not paying much attention. Other suggestions included more paintings, more animated characters and harder tasks.

Formal studies were not conducted with children to test the final version of the app due to the limited time-frame as our Ethics Application did not get approved before the submission deadline. To mitigate this, we conducted a study in the *Bristol Museum* with visitors who were over the age of 18 as well as with students from *University of Bristol* in a mock-up gallery in the atrium of *Merchant Venturers Building*.

4.2 Feedback From other Participants

All of our participants thought children would enjoy using this app.

> *Good idea, it would be definitely more engaging than just going to the museum and reading things about paintings.*

We first conducted a Semi-structured interview, thank asked participants to fill in a questionnaire containing nine questions using *Likert* [85] scale and two open ended questions to get additional feedback. Likert scale is widely used to measure attitudes and opinions with a more nuance than a simple 'yes' or 'no' question. This method provides the means to get more granular feedback. This is one of the most reliable ways to measure opinions and perception of users.

We wanted to avoid asking users too broad questions, but at the same time wanted to provide the opportunity to give creative answers since they might have other additional input about the design or the fundamental idea. To achieve this we asked them which parts they thought could use more work and whether they have further comments. Their answers were particularly useful for future work as they made useful suggestions. One of the participants suggested to reduce noise on the recorded audio by preprocessing it, another said they would prefer having more instructions on how to use the app — as a museum employee — such as a short tutorial at the beginning or suggestions during use.

We evaluated the results of the Likert scale questions by looking at the *median* of the results.

Figure 4.1: Users testing MuseumAR (*Picture taken with the participants' consent*)

Figure 4.2: The museum seen by a visitor exploring the gallery

4.2.1 Analyzing Results

- The unclarity of the tasks (Figure A.1) and the difficulty of using the app (Figure A.4) both had a median score of 2/5. This means that users mostly *disagreed* with these statements.

- Participants *agreed* that museums should fund the further development of this project (Figure A.9), they thought the UI was intuitive (Figure A.7), easy to use (Figure A.5) and instructions were easy to follow (Figure A.2) with a median of 4/5.

- They *strongly agreed* when asked whether they think this app would attract more visitors (Figure A.3), if it is a novel idea (Figure A.6) and children would enjoy using MuseumAR (Figure A.8) with a score of 5/5.

- Numerical results were missing from Question 10 as users preferred to specify their answers instead of giving a score.

By conducting semi-structured interviews and observing the users performing the tasks we could draw several conclusions.

When testing the UI intended for visitors we observed that at the first time users saw the AR character they were unsure how to interact with it. Because the character was waving at them all of them eventually realized that they are supposed to tap on it. We concluded that in the future this should be made more explicit. Most users spent more time looking at the first painting, trying to understand how to perform the tasks they were asked to do. Some of them seemed to struggle with tapping on the objects on the screen because of the size of the device. This could be mitigated by using iPhones instead of iPads, although that would have drawbacks because of the reduced screen size.

When testing the interface for inputting data, users seemed to struggle finding the button to delete a painting. This could be made more explicit in the future.

There are three non-optional fields when adding a new painting to the gallery — title, physical width and an image of the painting. Users struggled to understand why they couldn't save paintings before completing these fields (The *Save* button was disabled while these field were missing, see Figure 3.4). They required clarification and further instructions. This could be solved by having a short notification at the top of the screen explaining that these are the minimum requirements for a painting to be recognized.

In the 'Further comments' section many participants mentioned that they liked the app and suggested to continue developing it.

> *This could definitely work in a real-world setting with with further development.*

The entire questionnaire and the answers to it can be seen in Appendix A.

The answers to questions 10 and 11 can be seen at A.3.1.

4.2.2 Interaction Feedback

When paintings are recognized, an almost transparent layer is overlayed to give visual feedback to the users that the painting was detected. This translucent plane contains hints of where the user is expected to tap in the form of a red box over the area of interest. Some users thought that this made their task too easy, although others found these hints useful or even thought these should be more obvious, helping them find the objects they were looking for. Eight out of the 18 people thought that a visual feedback when they tapped the right objects should be added to the app.

4.2.3 Tracking

Most participants suggested that the tracking is still not accurate enough. Not surprisingly all these participants were recruited in MVB and all of them were Computer Science students. Other participants didn't mention this and seemed impressed with the capabilities of the technology.

4.2.4 Content

Feedback on the content was mostly positive, the participants thought that the information they got was educational and the interactions with the paintings encouraged them to remember details of the paintings (Transcript of the audio used in MVB in A.1 and in Bristol Museum in A.2). Some of them suggested to add more educational content, such as details about the paintings, artistic style and the story behind it. We could indicate this to the museum workers who will provide this content in the future.

The feedback from users was very helpful, in case we would have had more time at our disposal, we would have implemented most of them.

4.3 Technical Issues

4.3.1 Size of Devices

Based on our user testing using an iPad and a study conducted at the British Museum [36] we found that small children have difficulties handling iPads at first because their hands are too small. Their biggest challenge was to hold the iPad with one hand and tap with the other. Surprisingly after a few attempts and by observing older children around them using the iPad, most of them managed to master the skill and experienced completing the task as a much bigger accomplishment resulting in higher levels of satisfaction. These challenges can be beneficial for children as they are more aware of their surroundings and more attentive to their bodies improving their coordination, an effect of kinaesthetic learning. Kinaesthetic learning, or tactile learning is a learning style in which learning takes place by the students carrying out physical activities, rather than listening to a lecture or watching demonstrations.

4.3.2 Audio Recordings

During testing the app, an issue with audio recordings became evident. When the audio was played through the speakers pf the iPad it disturbed other visitors. In some cases users said that they could not hear the instructions clearly because the speakers were too quiet. Another issue in the future would rise when more children would look at the same painting, causing audio playing at the same time.

A possible solution to this is the use of headphones. This would solve the aforementioned issues, but would raise the problem of social isolation. There has to be a compromise made when it comes to this issue.

Chapter 5

Conclusion

5.1 Key Accomplishments

MuseumAR is a prototype app for a potentially much bigger project motivated by the need of technical moderators in the museum environment to engage children and provide them with a better learning experience. The app is not mature enough to be deployed, but it does have all the main functionalities. The reason for this is the limited amount of time at our disposal. Nevertheless we were able to demonstrate the key aspects that made our approach unique and valuable for museums and art galleries.

- Built an easy to use User Interface for museum staff.
- Provided a means for museums to make use of Augmented Reality technology.
- Provided an interactive platform to children to learn about arts.
- Created a demo story to show the potential of the app.
- Attempted to bring closer technology and fine art.

5.2 Ideal User Experience

During the planning and development of MuseumAR I imagined an ideal final outcome for the project. I would like to describe this idea as I believe it would give a better perspective of our goals regarding the user experience.

When creating the chain of events triggered by paintings, it is important to have a well defined story line that guides the whole experience.

As an example we could take a hypothetical museum where one of the main exhibits is about pirates. The main character in the story would be an animated pirate that accompanies children during the visit.

The children would receive a printed sign that would trigger the character to appear and repeat the latest instruction that was heard. When the card is first scanned, the character would start with introducing themselves and telling children how to interact with the app, what their tasks are and give health and safety instructions.

When the adventure begins this character disappears, but it can be evoked by looking at the handed picture any time. Children would be told that the pirate needs their help to go on a quest. They are hired to collect objects necessary for the journey and memorize information that will be useful for the pirate. Children would walk around the exhibit collecting objects from paintings and carefully listening to the educational information about what different objects symbolize on paintings, how were these objects used, differences between different painting styles and other.

At the end of the story, they would get a quiz that would evaluate their knowledge and would give them a score in addition to the time required to complete the tasks giving them a ranking.

Some paintings or signs displayed in the museum could be used as anchors to trigger different scenes. For example, the first room of the exhibition could be the pirates home, the second the tavern and the third the ship. These scenes could be illustrated by displayed 3D objects, such as the one shown at Figure 3.8. There could be also side stories that would require special attention to be discovered and potentially would lead children to a different exhibition.

5.3 Future Work

There is a lot of potential for future work with MuseumAR, since the app currently is a first prototype and would need more iterations to be deployable in a museum. The implementation is not completely developed yet and it is meant to be a proof of concept rather than a market-ready product.

In this section we will discuss potential future work and improvements for the app.

5.3.1 Object Catalog

The initial plan of the app included an object catalog for the museum staff to pick 3D objects from. This was not implemented as it was not a crucial part of the design. The main reason for not implementing this feature was the limited amount of time. Creating 3D objects is time-consuming and as a proof of concept we decided that one character will be enough. The character can perform different actions (implemented animations) that could be chosen by the museum staff. In addition to objects accessible offline we could also link our app to Google Poly [86], an online platform containing open source 3D objects available for everyone.

5.3.2 Scoring System

As mentioned earlier in the *early field studies section*(Section 3.2.2) some participants suggested to implement a scoring system that would also allow visitors to select a name for themselves. At every task they would receive a score, that could be based on time required to complete the tasks, accuracy of completion and potential results from a quiz at the end of the visit. This would have also required building a server where the results could be uploaded.

5.3.3 Multiple Stories

An extension to this app could be built that would allow the creation of more stories and building a server where these stories could be stored and downloaded again in case it was required. If this was the case, we should allow visitors to decide which storyline they would like to explore.

5.3.4 Extending to Adults

Many adult participants from our study suggested that they would like to see a similar guide designed for adults. There are only few changes that would have to be made to accommodate this requirement.

5.3.5 Augmenting Paintings in Detail

When a painting is recognized, it could be overlayed with a version of that painting but an object missing. The user then would have to find what the differences are between the two paintings.

5.3.6 Videos

We could incorporate displaying educational videos about certain paintings in the app. For example a video of how the painting was created could be played, overlaying the original painting.

Bibliography

[1] ARKit. https://developer.apple.com/documentation/arkit#overview.

[2] H. K. Kabali, M. M. Irigoyen, R. Nunez-Davis, J. G. Budacki, S. H. Mohanty, K. P. Leister, and R. L. Bonner, "Exposure and use of mobile media devices by young children," *Pediatrics*, vol. 136, no. 6, pp. 1044–1050, 2015.

[3] Ofcom, "Children and parents: Media use and attitudes report," November 2017.

[4] "Percentage of households with mobile phones in the united kingdom (uk) from 1996 to 2017." www.statista.com/statistics/289167/mobile-phone-penetration-in-the-uk/.

[5] N. Mendoza, "The Mendoza review: an independent review of museums in england," independent report, Department of Digital, Culture, Media & Sport, November 2017.

[6] N. Mendoza, "The mendoza review: an independent review of museums in england," independent report, Department of Digital, Culture, Media & Sport, November 2017.

[7] Department of Digital, Culture, Media & Sport, "Official statistics - sponsored museums performance indicators 2015/16," independent report, Department of Digital, Culture, Media & Sport, 2016. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/595440/Sponsored_Museums_Performance_Indicators_revised.pdf.

[8] Department of Digital, Culture, Media & Sport, "Official statistics - sponsored museums performance indicators 2015/16," independent report, Department of Digital, Culture, Media & Sport, 2017. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/657464/Sponsored_Museums_Performance_Indicators_2016_17.pdf.

[9] F. Y. Wong, "Influence of pokémon go on physical activity levels of university players: a cross-sectional study," *International Journal of Health Geographics*, vol. 16, p. 8, Feb 2017.

[10] Á. Di Serio, M. B. Ibáñez, and C. D. Kloos, "Impact of an augmented reality system on students' motivation for a visual art course," *Computers & Education*, vol. 68, pp. 586–596, 2013.

[11] J. M. Keller, "Development and use of the arcs model of instructional design," *Journal of instructional development*, vol. 10, no. 3, p. 2, 1987.

[12] J. M. Keller, "Strategies for stimulating the motivation to learn," *Performance Improvement*, vol. 26, no. 8, pp. 1–7, 1987.

[13] D. Anderson, B. Piscitelli, K. Weier, M. Everett, and C. Tayler, "Children's museum experiences: Identifying powerful mediators of learning," *Curator: The Museum Journal*, vol. 45, no. 3, pp. 213–231, 2011.

[14] D. Anderson, B. Piscitelli, K. Weier, M. Everett, and C. Tayler, "Children's museum experiences: Identifying powerful mediators of learning," *Curator: The Museum Journal*, vol. 45, no. 3, pp. 221–222, 2011.

[15] F. Rizzo and F. Garzotto, ""the fire and the mountain": Tangible and social interaction in a museum exhibition for children," in *Proceedings of the 6th International Conference on Interaction Design and Children*, IDC '07, (New York, NY, USA), pp. 105–108, ACM, 2007.

[16] J. Pallud, "Impact of interactive technologies on stimulating learning experiences in a museum," *Information & Management*, vol. 54, no. 4, pp. 465–478, 2016.

[17] A. C. Lin, W. D. Fernandez, and S. Gregor, "Understanding web enjoyment experiences and informal learning: A study in a museum context," *Decision Support Systems*, vol. 53, no. 4, pp. 846–858, 2012.

[18] R. Pekrun, T. Goetz, A. C. Frenzel, P. Barchfeld, and R. P. Perry, "Measuring emotions in students learning and performance: The achievement emotions questionnaire (aeq)," *Contemporary Educational Psychology*, vol. 36, no. 1, pp. 36 – 48, 2011. Students Emotions and Academic Engagement.

[19] D. Finch, M. Peacock, D. Lazdowski, and M. Hwang, "Managing emotions: A case study exploring the relationship between experiential learning, emotions, and student performance," *The International Journal of Management Education*, vol. 13, no. 1, pp. 23 – 36, 2015.

[20] L. Bannon, S. Benford, J. Bowers, and C. Heath, "Hybrid design creates innovative museum experiences," *Communications of the ACM*, vol. 48, no. 3, p. 62, 2005.

[21] S. Gupta and R. Bostrom, "Technology-mediated learning: A comprehensive theoretical model," *Journal of the Association for Information Systems*, vol. 10, no. 9, pp. 694–696, 2009.

[22] M. Csikszentmihalyi, "Toward a psychology of optimal experience," in *Flow and the foundations of positive psychology*, pp. 209–226, Springer, 2014.

[23] K. Walczak, R. Wojciechowski, and W. Cellary, "Dynamic interactive vr network services for education," in *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, VRST '06, (New York, NY, USA), pp. 277–286, ACM, 2006.

[24] K. Charitonos, C. Blake, E. Scanlon, and A. Jones, "Museum learning via social and mobile technologies: (how) can online interactions enhance the visitor experience?," *British Journal of Educational Technology*, vol. 43, no. 5, pp. 802–819, 2012.

[25] A. C. H. Lin, S. D. Gregor, and J. C. Huang, "The hedonic experience of enjoyment and its relationship to informal learning: A study of museum websites," *ICIS 2008 Proceedings. 208*, p. 208, 2008.

[26] I. E. Dror, "Technology enhanced learning: The good, the bad, and the ugly," *Pragmatics & Cognition*, vol. 16, no. 2, pp. 215–223, 2008.

[27] O. Malik

[28] T. Miyashita, P. Meier, T. Tachikawa, S. Orlic, T. Eble, V. Scholz, A. Gapel, O. Gerl, S. Arnaudov, and S. Lieberknecht, "An augmented reality museum guide," in *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 103–106, IEEE Computer Society, 2008.

[29] D. Schmalstieg and D. Wagner, "Experiences with handheld augmented reality," in *Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on*, pp. 3–18, IEEE, 2007.

[30] C. Heath and D. vom Lehn, "Misconstruing interaction," *Proceedings of Interactive Learning in Museums of Art and Design*, 2003.

[31] M. B. Stogner, "The media-enhanced museum experience: Debating the use of media technology in cultural exhibitions," *Curator: The Museum Journal*, vol. 52, no. 4, pp. 385–397, 2009.

[32] D. vom Lehn & Christian Heath, "Accounting for new technology in museum exhibitions," *International Journal of Arts Management*, vol. 7, no. 3, pp. 11–21, 2005.

[33] A. B. Tillon, I. Marchal, and P. Houlier, "Mobile augmented reality in the museum: Can a lace-like technology take you closer to works of art?," in *Proceedings of the 2011 IEEE International Symposium on Mixed and Augmented Reality–Arts, Media, and Humanities*, ISMAR-AMH '11, (Washington, DC, USA), pp. 41–47, IEEE Computer Society, 2011.

[34] "The national gallery review of the year 2016-2017,"

[35] A. Antoniou, J. O'Brien, T. Bardon, A. Barnes, and D. Virk, "Micro-augmentations: Situated calibration of a novel non-tactile, peripheral museum technology," in *Proceedings of the 19th Panhellenic Conference on Informatics*, PCI '15, (New York, NY, USA), pp. 229–234, ACM, 2015.

[36] S. Mannion, "Beyond cool: Making mobile augmented reality work for museum education," in *Museums and the Web*, pp. 11–14, 2012.

[37] B. Museum. http://www.britishmuseum.org/learning/samsung_centre.aspx.

[38] M. Ding, "Augmented reality in museums," 2017.

[39] G. A. R. Chicago History Museum, "Chicago00 experiences: making chicago history with new media." http://chicago00.org/experiences.html, May 2018.

[40] R. T. Azuma, "A survey of augmented reality," *Presence: Teleoperators and Virtual Environments*, vol. 6, no. 4, pp. 355–385, 1997.

[41] D. Wagner, T. Pintaric, F. Ledermann, and D. Schmalstieg, "Towards massively multi-user augmented reality on handheld devices," in *Pervasive Computing* (H. W. Gellersen, R. Want, and A. Schmidt, eds.), (Berlin, Heidelberg), pp. 208–219, Springer Berlin Heidelberg, 2005.

[42] K. J. Carlson and D. J. Gagnon, "Augmented reality integrated simulation education in health care," *Clinical simulation in nursing*, vol. 12, no. 4, pp. 123–127, 2016.

[43] I. C. da Silva, G. Klein, and D. M. Brandão, "Segmented and detailed visualization of anatomical structures based on augmented reality for health education and knowledge discovery," *Advances in Science, Technology and Engineering Systems Journal (ASTESJ)*, vol. 2, no. 3, pp. 469–478, 2017.

[44] E. Olshannikova, A. Ometov, Y. Koucheryavy, and T. Olsson, "Visualizing big data with augmented and virtual reality: challenges and research agenda," *Journal of Big Data*, vol. 2, p. 22, Oct 2015.

[45] W. Li, A. Nee, and S. Ong, "A state-of-the-art review of augmented reality in engineering analysis and simulation," *Multimodal Technologies and Interaction*, vol. 1, no. 3, p. 17, 2017.

[46] R. Azuma, "Tracking requirements for augmented reality," *Commun. ACM*, vol. 36, pp. 50–51, July 1993.

[47] P. Daponte, L. De Vito, F. Picariello, and M. Riccio, "State of the art and future developments of the augmented reality for measurement applications," *Measurement*, vol. 57, pp. 53–70, 2014.

[48] C. Tomasi and T. Kanade, "Detection and tracking of point features," 1991.

[49] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[50] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proceedings of the Seventh IEEE International Conference on Computer Vision*, vol. 2, pp. 1150–1157 vol.2, 1999.

[51] F. Pla and J. A. Marchant, "Matching feature points in image sequences through a region-based method," *Computer vision and image understanding*, vol. 66, no. 3, pp. 271–285, 1997.

[52] K. Yousif, A. Bab-Hadiashar, and R. Hoseinnezhad, "An overview to visual odometry and visual slam: Applications to mobile robotics," *Intelligent Industrial Systems*, vol. 1, pp. 289–311, Dec 2015.

[53] F. Fraundorfer and D. Scaramuzza, "Visual odometry: Part i: The first 30 years and fundamentals," *IEEE Robotics and Automation Magazine*, vol. 18, no. 4, pp. 80–92, 2011.

[54] D. Nistér, O. Naroditsky, and J. Bergen, "Visual odometry," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 1, pp. I–I, Ieee, 2004.

[55] D. E. Dilger, "Inside apple's arkit and visual inertial odometry, new in ios 11." https://appleinsider.com/articles/17/10/12/inside-apples-arkit-and-visual-inertial-odometry-new-in-ios-11, 2017.

[56] J. Engel, J. Sturm, and D. Cremers, "Semi-dense visual odometry for a monocular camera," in *Computer Vision (ICCV), 2013 IEEE International Conference on*, pp. 1449–1456, IEEE, 2013.

[57] J. Campbell, R. Sukthankar, I. Nourbakhsh, and A. Pahwa, "A robust visual odometry and precipice detection system using consumer-grade monocular vision," in *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, pp. 3421–3427, IEEE, 2005.

[58] E. Marchand, H. Uchiyama, and F. Spindler, "Pose estimation for augmented reality: a hands-on survey," *IEEE transactions on visualization and computer graphics*, vol. 22, no. 12, pp. 2633–2651, 2016.

[59] H. Bay, B. Fasel, and L. Van Gool, "Interactive museum guide: Fast and robust recognition of museum objects," 2006.

[60] T. Dalens, J. Sivic, I. Laptev, and M. Campedel, "Painting recognition from wearable cameras," contract, Sept. 2014.

[61] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (surf)," *Comput. Vis. Image Underst.*, vol. 110, pp. 346–359, June 2008.

[62] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "Orb: An efficient alternative to sift or surf," in *Computer Vision (ICCV), 2011 IEEE international conference on*, pp. 2564–2571, IEEE, 2011.

[63] P. Föckler, T. Zeidler, B. Brombach, E. Bruns, and O. Bimber, "Phoneguide: Museum guidance supported by on-device object recognition on mobile phones," in *Proceedings of the 4th International Conference on Mobile and Ubiquitous Multimedia*, MUM '05, (New York, NY, USA), pp. 3–10, ACM, 2005.

[64] B. Ruf, E. Kokiopoulou, and M. Detyniecki, "Mobile museum guide based on fast sift recognition," in *International Workshop on Adaptive Multimedia Retrieval*, pp. 170–183, Springer, 2008.

[65] C. Hasbrouck, "Difference-between-google-arcore-and-apple-arkit." https://blog.truthlabs.com/image-recognition-with-apples-arkit-showcasing-it-s-capabilities-3ba460b5898f.

[66] F. Dicksee, *La belle damme sans mercy.* Bristol Museum and Art Gallery.

[67] A. Hughes, *The Guarded Bower.* Bristol Museum and Art Gallery.

[68] E. C. Burn-Jones, *The Garden Court.* Bristol Museum and Art Gallery.

[69] H. Wattson, *Holiday.* Bristol Museum and Art Gallery.

[70] W. J. Muller, *The Carpet Bazaar, Cairo.* Bristol Museum and Art Gallery.

[71] A. Talmage, *The Mackerel Shawl.* Bristol Museum and Art Gallery.

[72] I. Pedersen, N. Gale, P. Mirza-Babaei, and S. Reid, "More than meets the eye: The benefits of augmented reality and holographic displays for digital cultural heritage," *Journal on Computing and Cultural Heritage (JOCCH)*, vol. 10, no. 2, p. 11, 2017.

[73] M. Hamilton and M. Weiss, "The power of storytelling in the classroom," *Excerpt from students tell stories: teaching and using storytelling in the classroom available at http://www. rcowen. com/PDF-s/CTS% 20Ch*, vol. 201, 2005.

[74] ARCore. https://www.blog.google/products/google-vr/arcore-augmented-reality-android-scale/.

[75] M. Usman, "Does googleâĂŹs arcore surpass appleâĂŹs arkit." http://appyspot.com/difference-between-google-arcore-and-apple-arkit/.

[76] C. Hasbrouck, "Image recognition with apple's arkit: Showcasing it's capabilities.." http://appyspot.com/difference-between-google-arcore-and-apple-arkit/.

[77] R. Campbell, "Augmented solar system." https://github.com/ThumbWorks/AugmentedSolarSystem.

[78] ARKit-Limitations. https://developer.apple.com/documentation/arkit/about_augmented_reality_and_arkit.

[79] Blender, "3d creation suite." https://belnder.org.

[80] Maxiamo, "Animated 3d characters." https://www.mixamo.com/#/.

[81] Maxiamo, "Characters." https://www.mixamo.com/#/?page=1&query=&type=Character.

[82] turbosquid, "Medieval house." https://www.turbosquid.com/FullPreview/Index.cfm/ID/606468.

[83] free3d, "Rose." https://free3d.com/3d-model/rose-31675.html.

[84] free3d, "Knights." https://www.turbosquid.com/FullPreview/Index.cfm/ID/606468.

[85] I. E. Allen and C. A. Seaman, "Likert scales and data analyses," *Quality progress*, vol. 40, no. 7, p. 64, 2007.

[86] Google, "Google poly." https://poly.google.com.

Appendix A

User Study Questionnaire

MuseumAR

Questionnaire

On a scale from one to five **one** being **strongly disagree**, **five strongly agree** —
Circle your answer

1. I had issues understanding the tasks. **1 2 3 4 5**

2. I found it easy to follow the instructions in the adventure. **1 2 3 4 5**

3. I believe this app would attract more visitors. **1 2 3 4 5**

4. I found the app difficult to use. **1 2 3 4 5**

5. On a scale of 1 to 5 how easy was the app to use? **1 2 3 4 5**

6. This is a novel idea. **1 2 3 4 5**

7. How intuitive was the user interface? **1 2 3 4 5**

8. I think children would enjoy using this app. **1 2 3 4 5**

9. I think the museum should fund the further development of this project.

   1 2 3 4 5

10. I thought some parts could use more work. Please specify which:

    1 2 3 4 5

11. Further comments and suggestions:

A.1 User testing story and Audio Transcript — MVB Atrium

In addition to the paintings, there was a mark on the wall that triggered an AR character to appear and shortly describe the app and tell visitors how to interact with their surroundings.

The first bullet point at every painting represents the audio played when users tap on the AR characters, the second one is triggered by the users completing all the tasks.

The transcript of the audio recordings attached to the individual paintings for the user testing are as follows:

Introduction

- Welcome to MuesumAR!

  To find out what you have to do next tap on the character appearing next to the painting. Your first task is to find the impressionist painting with complimentary colours on it.

Sunrise — Claud Monet

- This is an impressionist painting by Claud Monet. The complimentary colours are very important. To find out where to go next find the boats and the Sun and tap on them.
- The next painting is another impressionist painting. There the main colours are green and red. Can you find it?

Poppy filed — Claud Monet

- This another painting by Monet. Your task is to find the people on the painting and the house.
- The next painting is a very colourful one. If you look at the brush strokes they are very different from the previous one. Which one did you like more?

Freshness of Cold — Leonid Afremov

- On this painting you can see all the colours of the rainbow. Tap on the ones that make you feel warm, also find the people.
- Congratulations, you completed all the tasks!

A.2 User testing story and Audio Transcript — Bristol Museum

In addition to the paintings, the sign at the entrance was used to trigger an AR character to appear and shortly describe the app and tell visitors how to interact with their surroundings.

As at the previous study, the first bullet point at every painting represents the audio played when users tap on the AR characters, the second one is triggered by the users completing all the tasks.

The transcript of the audio recordings attached to the individual paintings for the user testing are as follows:

Introduction

- Welcome to MuesumAR!

  To find out what you have to do next tap on the character appearing next to the painting. In this exhibition we will explore ‘Places of Desire’ and compare it to everyday life. Your first task is to find the painting with the knight and a lady riding a horse.

La Belle Damme Sans Merci — Frank Dicksee

- This painting was inspired by Kieat’s poem about a knights’ encounter with a beautiful lady capturing the moment they meet. What is the lady wearing in her hair? This object is in opposition with the knights sword. Find these two objects and tap on them.
- On the next painting there are six women sleeping, what do you think happened to them?

The Garden Court - Briar Rose series — Edward Coley Burne-Jones

- This another painting depicting an enchanted scene. It is inspired by the famous fairy tale, *The sleeping beauty* and it shows a part of the castle where six women fell asleep while fetching water. There are is an interesting object on the painting, it represents time. Can you find it?
- The next painting is similar to the first one we saw. There is a man and a woman, and the man is holding a sword.

The Guarded Bower — Arthur Hughes

- On this painting you can see two white birds. What type of birds do you think they are? What do they symbolize. There is another bird hidden in the background. Find all three birds and tap on them.
- The next painting is showing people on holiday.

Holiday — Harry Wattson

- We are back to reality! This picture shows a care-free day at the riverside. How do you think these people feel? To get the next clue, find the umbrella and the hat.
- The next painting is a very colorful one. It shows a warm day at the carpet bazaar. Do you know what a bazaar is?

The Carpet Bazaar, Cairo — William James Muller

- This painting was painted by a Bristol artist years after he returned from his trip to Cairo. The colors are in harmony, just by looking at it you can imagine the heat. There are three lams hidden on the painting, please find them.
- Congratulations! You completed all the tasks!

A.3 Study Results

A.3.1 Answers to Question 10 and 11

- ‘Good idea, it would be definitely more engaging than just going to the museum and reading things.’
- ‘Remove the hint, it makes it too easy.’
- ‘More obvious hint.’
- ‘Add feedback when I tap the correct object.’ x 8
- ‘Improve tracking.’ x 2
- ‘Interactivity of the paintings, doesnâĂŹt feel like you are touching the painting at the moment.’
- ‘Could definitely work in a real-world setting with with further development.’
- ‘Add more character animations.’ x2
- ‘Fit all the information for staff on the screen at the same time.’
- ‘Add more content about the meaning and depiction of the paintings.’
- ‘Keep development going! ItâĂŹs really cool.’ x 2
- ‘Automated audio cleaning.’ x 2
- ‘Put more instructions on how to use the app.’
- ‘Good that we can interact with paintings.’
- ‘Audio should be lauder.’
- ‘I like the Character! It is fun.’

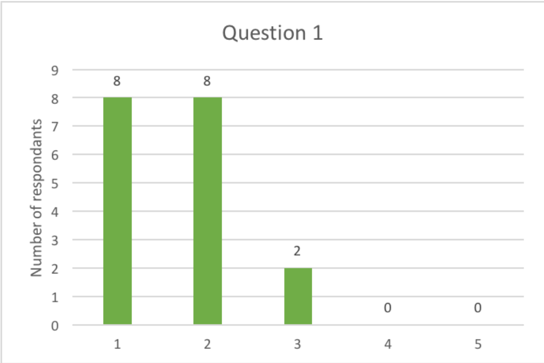

Figure A.1: I had issues understanding the tasks.

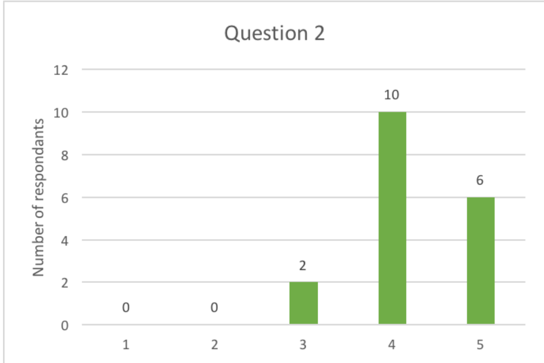

Figure A.2: I found it easy to follow the instructions in the adventure.

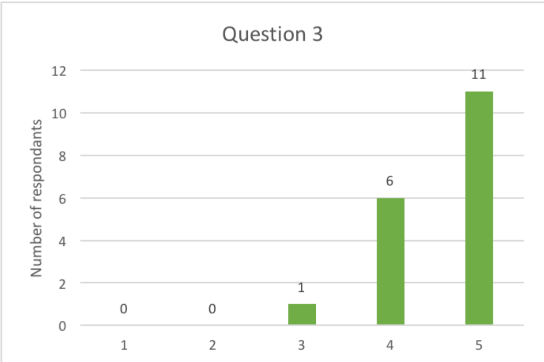

Figure A.3: I believe this app would attract more visitors.

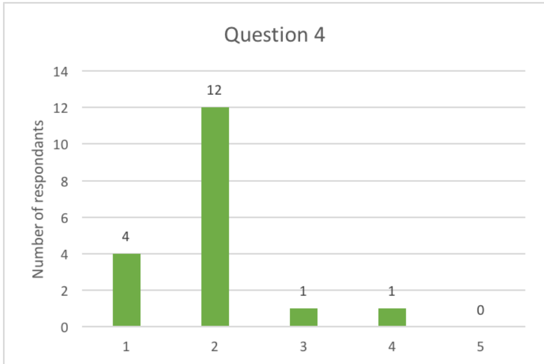

Figure A.4: I found the app difficult to use.

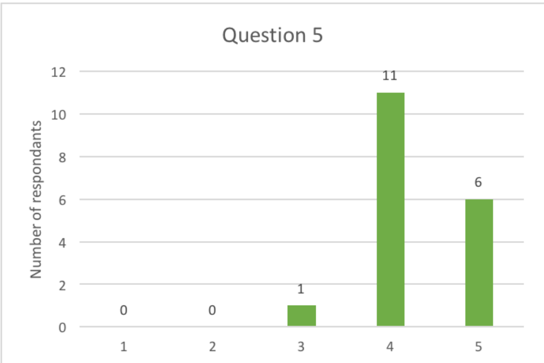

Figure A.5: On a scale of 1 to 5 how easy was the app to use?

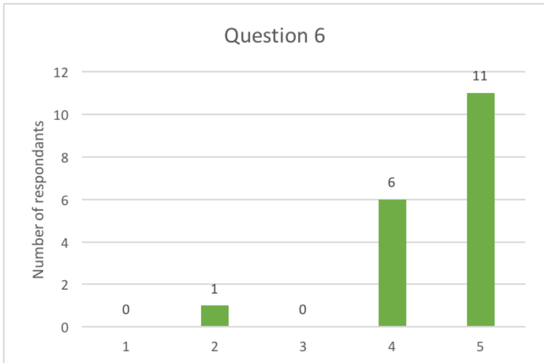

Figure A.6: This is a novel idea.

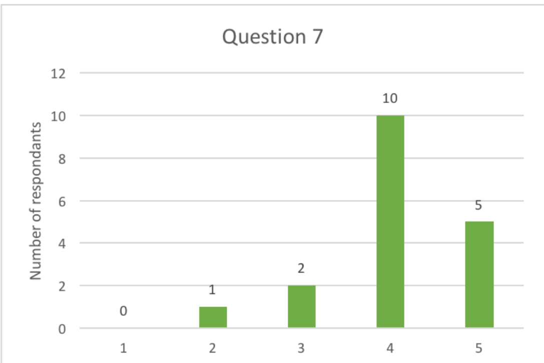

Figure A.7: How intuitive was the user interface?

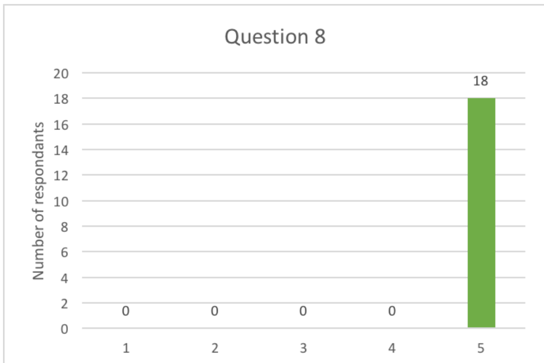

Figure A.8: I think children would enjoy using this app.

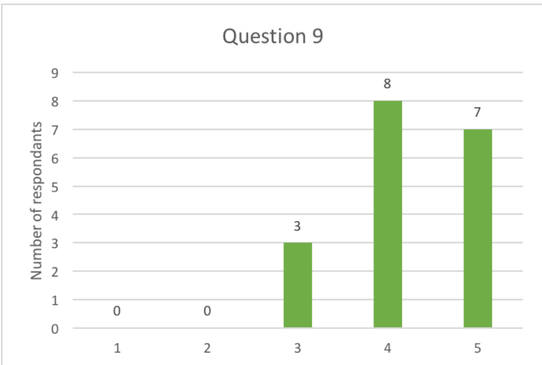

Figure A.9: I think the museum should fund the further development of this project.