



POLITECNICO DI TORINO

Corso di Laurea in Ingegneria Informatica

Tesi di Laurea Magistrale

An Augmented Reality Application to Perform a Virtual Tourist Guide

Relatore

prof. Andrea Sanna

Candidato

Davide ZAPPIA

Supervisore aziendale

RAI CRIT

ing. Roberto Iacoviello

ing. Federico Debandi

ANNO ACCADEMICO 2016-2017

An Augmented Reality Application to Perform a Virtual Tourist Guide

Davide Zappia

Supervised by:

Prof. Andrea Sanna

DAUIN

Ing. Roberto Iacoviello

RAI CRIT

Ing. Federico Debandi

RAI CRIT

Abstract

The idea of this thesis was born in a collaboration between RAI-CRIT (Centro Ricerche e Innovazione Tecnologica) and Politecnico di Torino: in particular, it has been developed in the context of “5GCity”, a European Project which has the purpose to implement and valorize 5G technology, creating innovative application. This thesis work has the aim of studying how mixed and augmented reality could be used in the real people’s life. In particular it wants to make user feel a new experience of city tour, with interactive information and virtual reconstruction of ancient monuments. Tourism is always an important topic in the world, especially in Italy in which are present best monuments and sculptures worldwide. Until nowadays the tourists can have many details about a monument or a cultural site by a physical guide or from any web sites. Unfortunately, a real prepared guide can be very expensive, while web sites can report wrong information or can be difficult to find.

For these reasons a Virtual Touristic Guide could be very useful for a lot of people. This virtual guide should show to the tourist some information about every interesting site, then it could go in deeply illustrating media content, text details and doing a virtual reconstruction in order to create an incredible immersive experience.

Many solutions have been presented for indoor where is quite easy to track

the user position and light conditions can be controlled. Several museums provide users Augmented Reality (AR) applications that framing well defined images (markers), recognize position and orientation of the user and overlap computer generated assets to the artwork the user is interested in. On the other hand, several challenges have to be tackled for outdoor applications. First of all an accurate tracking of the user is not easy; often GPS-based solutions are provided, but they might be not able to accurately provide both the position and orientation. Moreover, GPS-denied environments should be also considered. The second issue is related to the impossibility to alter the environment by markers or target images. Finally, lighting conditions cannot be controlled.

In this project is presented an Augmented/Mixed Reality outdoor solution, that wants to solve this request. Augmented Reality (AR) provides an efficient and intuitive way to visualize computer generated information overlaid and aligned with objects in the real environment, while with Mixed Reality (MR) the virtual object also interact itself with the real object, so the user can try an immersive experience which interact with his daily world. This can be possible thanks to the see-through lenses of the headset. The aim of the application is to enhance the user experience during cultural tours. With Mixed Reality, synthetic contents are anchored and the user can interact with them. The user wears a headset (e.g, the Microsoft Hololens) that is connected by a high-speed and high-throughput network (e.g, 4G/5G) to a cloud infrastructure. The project requirements have been defined in discussing with RAI-CRIT tutors. The application developed in this thesis is a Virtual Tourist Guide, based on Microsoft Hololens and allows to:

- Recognize the monument
- Display general information
- Show some text details
- Reproduce a video content
- Perform a virtual reconstruction of the monument

Based on this work's pre-study, the best way to recognize monument and perform a good usage of 5G technology is to deploy a web service on a cloud node in which a program of visual content analysis and search based on the MPEG Standard CDVS (Compact Descriptors for Visual Search) is running. The guide application is completely installed on the Hololens. The tools used are:

- Unity3D – Game engine and development environment for the Hololens application
- Eclipse as IDE to develop the web service on the cloud
- Blender for 3D model
- Vuforia SDK to implement AR features

The first part of the application, the monument recognition, is invisible to the user. So, he can visit the city while the application search for any monument. When the application finds a monument, a hologram with the general information appears and the user can choose to know deeply about it or search for other monuments. If the user is interested, he can select text details, media content or AR reconstruction through the monument menu.

The project has been tested at Archaeological Park in Turin, with Palatine Towers as target. A group of testers walked through the park until the Towers. Once there, they tried to use all part of the application, from the text details to the 3D model reconstruction. Moreover, the users had to pay attention to some other factors in order to evaluate the usability and visibility of the system. At the end, users had to fill a questionnaire with some questions and an answer scale from 1 (totally disagree) to 5 (totally agree). Tests are focused on various type of parameters, they are being created in order to analyse both subjective and objective impressions.

From the results it is visible that users generally appreciate the whole application, thinking that it can be a useful system to enhance the cultural heritage. Users also liked the pointing system and the natural hands interaction paradigm, that allow them to learn fast how to use the application and make them feel completely immerse in the holograms.

*ai miei nonni,
guida nel passato e nel presente*

Acknowledgements

Dopo l'arduo sviluppo di questo progetto, vorrei ringraziare il mio relatore, prof. Andrea Sanna, per avermi dato la possibilità di svolgere la tesi in una grande realtà come la RAI e per la sua disponibilità e guida costante.

Questo progetto di tesi è stato realizzato in collaborazione con RAI-CRIT (Centro Ricerche e Innovazione Tecnologica), per cui ringrazio tutti coloro che durante lo svolgimento del lavoro in azienda mi hanno dato consigli preziosi.

Un ringraziamento particolare va sicuramente all'Ing.Roberto Iacoviello e all'Ing.Federico Debandi, per avermi sostenuto e indirizzato per tutta la durata del progetto.

Mi sento inoltre in dovere di ringraziare l'Ing.Alberto Messina e l'Ing.Maurizio Montagnuolo per aver messo a disposizione la loro competenza.

Infine, devo ringraziare i miei genitori, i miei cari, mia sorella e Lucia per il costante supporto.

Contents

List of Figures	12
1 Introduction, Motivations and Goals	17
1.1 Introduction	17
1.2 Research Objectives	17
1.3 Motivation	19
1.4 Context	20
1.4.1 5G Networks	21
2 State of the Art	23
2.1 Virtual, Augmented and Mixed Reality	23
2.2 Augmented Reality Applications	24
2.2.1 Gaming and Entertainment	25
2.2.2 Teaching	25
2.2.3 Medicine and Surgery	25
2.2.4 Augmented Reality for Television and Entertainment . .	26
2.2.5 Augmented Reality for Cultural Heritage	27
2.3 Augmented Reality Devices	28
2.4 Augmented Reality SDKs	30
3 Technologies	33
3.1 Microsoft Hololens	33
3.1.1 Hardware	34
3.2 Unity 3D	35
3.3 Vuforia	36
3.3.1 Tracking	36
3.3.2 Image Recognition	37
3.4 Microsoft Azure	38
3.5 CDVS Compact Descriptors for Visual Search	38

3.5.1	Key Point	39
4	Augmented Reality Tourist Guide	41
4.1	Project Requirements	41
4.2	System Architecture	42
4.2.1	Hardware Architecture	42
4.2.2	Software Architecture	43
4.3	The Application	46
4.3.1	First Scene	46
4.3.2	Second Scene	49
4.3.3	Third Scene	54
4.3.4	Web Service	55
4.3.5	User Interface	56
5	Test and Results	59
5.1	Test	59
5.2	Testers	59
5.3	Location	60
5.4	Results	63
5.4.1	Analysis Part	64
6	Conclusion	67
	Bibliography	69

List of Figures

1.1	Hype cycle for emerging technologies (Source: Gartner.com, 2017)	18
1.2	Most relevant trends appear in the Gartner Hype Cycle (Source: Gartner.com, 2017)	19
1.3	5G capabilities expectation (Image source: ETRI graphic, from ITU-R IMT 2020 requirements)	21
1.4	Expected performance of 5G networks (Source: Nokia 2017)	22
2.1	Virtuality Continuum	24
2.2	Pokemon Go	25
2.3	Chemistry augmented reality application for educational field	26
2.4	“Skins and Bones” app in Smithsonian Museum (Source: The Smithsonian National Museum of Natural History)	27
2.5	Microsoft Hololens and Epson Moverio (Source: Michael Leyva)	29
3.1	Technologies used in this project	33
3.2	Microsoft’s newest mixed reality device, Hololens	34
3.3	Hololens’ technical specifications (Source: Microsoft)	35
3.4	Hololens’ display system, Waveguide Lens	35
3.5	Natural feature tracking	37
3.6	Natural feature tracking on Porta Palatina in Turin	37
3.7	CDVS tracking (Source: sistemi-integrati.net)	38
3.8	Key points descriptor system (Source: sistemi-integrati.net)	39
4.1	System’s hardware structure	43
4.2	System’s software structure	44
4.3	GameObjects and their scripts in the first scene	47
4.4	AnnotationHandlingScript Flow	48
4.5	GameObjects and their scripts in the second scene	50

4.6	ManagerScript flow	51
4.7	Web service flow	55
4.8	Air tap gesture	56
4.9	Gaze targeting technique	57
4.10	InfoPlane Example	57
4.11	MenuPlane Example	58
5.1	Initial information after monument recognition	61
5.2	Some text details about the monument	61
5.3	Monument's 3D model overlay	62
5.4	Hololens' field of view	63
5.5	Answers to question 1 "Are commands simple and intuitive?"	65
5.6	Answers about "Application's fluency"	65
5.7	Answers about "Speed of initial monument recognition"	66
5.8	Answers about "Quality of Vuforia Tracking Stability"	66

List of Contributions

- *Ing.Roberto Iacoviello* RAI CRIT;
- *Ing.Federico Debandi* RAI CRIT;
- *Ing.Alberto Messina* RAI CRIT;
- *Ing.Maurizio Montagnuolo* RAI CRIT.

Chapter 1

Introduction, Motivations and Goals

Give me six hours to chop down a tree and I will spend the first four sharpening the axe. *Abraham Lincoln*

1.1 Introduction

In the digital era humans have always dreamed about a smarter life merging reality with virtual object and information. Technology development grows up day by day in an incredible way, making the dream much more possible. Until few years ago this improvement was shown only by books, videogames and cinema, like the futuristic interface of *Minority Report* (2002). The current success of some Augmented Reality applications (Pokémon Go) increases the media and public interest for this topic and enhance the expectation for future development. The meaning of this terms and some other important AR applications are explained deeply in chapter 2. This master's thesis discusses later in the introduction and more deeply in chapter 3 about the technologies used in this project. Then, next chapters will present the work done for this project, the objectives and the results obtained.

1.2 Research Objectives

This thesis work aims to propose a Virtual Touristic Guide as a Hololens application for visualize and learn more about the Italian's monuments. It is been created in order to add value to our country and to let people experience

a unique city tour. This application was developed in the context of “5GCity” European Project. It’s really exiting to have digital objects that fit with the real environment and allow the user to interact with them through the help of the fastest internet connection. According to (Tim Cook, 2017), “I could not be more excited about AR and what we’re seeing with ARKit. It’s all over the place, from entertainment to gaming. I think AR is big and profound, this is one of those huge things that we’ll look back at and marvel at the start of it. I think customers are going to see it in a variety of ways and it feels great to get [AR] going at a level that can get all of the developers behind it.”. Augmented reality can perform unique advantages, like the overlays of virtual object on reals, that never people seen. This new technology can make people life very easier with digital information in the daily routine. Some of this AR application allow the users to be smarter and take advantage in their real world with the combination between senses and digital computing.

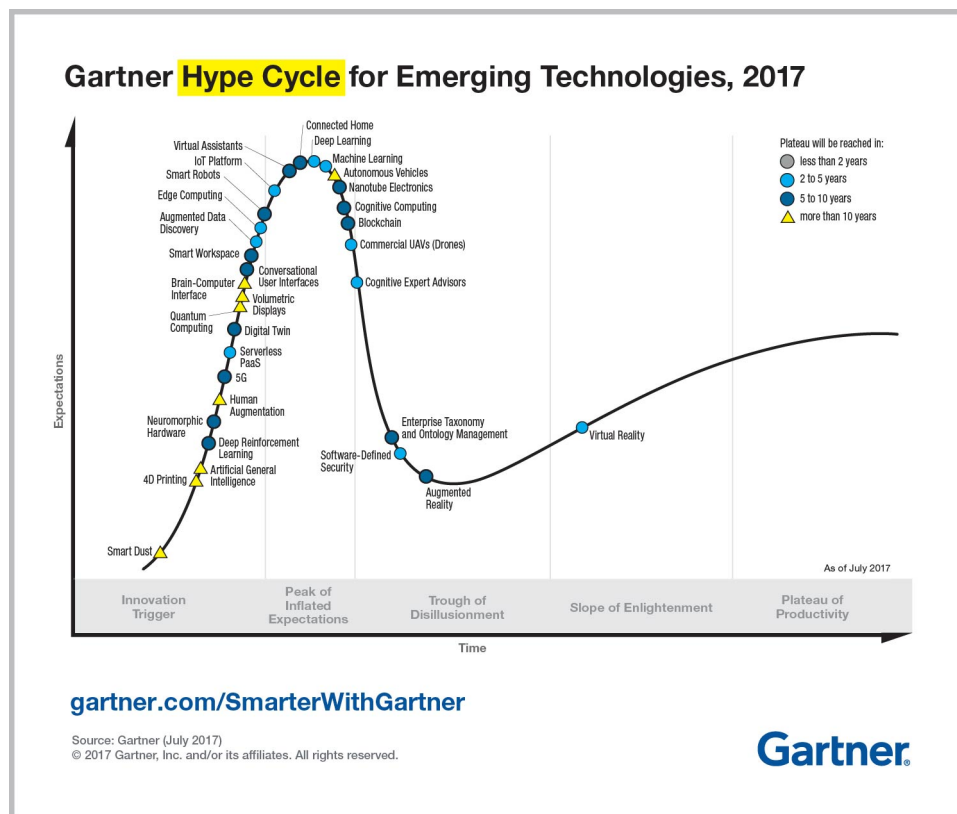


Figure 1.1. Hype cycle for emerging technologies (Source: Gartner.com, 2017)

1.3 Motivation

The technological research is an important topic of this century, with the large use of 3D in the entertainment (cinema, gaming) also the people's expectation is very increment.

This 1.1 Hype Cycle focuses on some technologies that promise a high degree of advantage. Mike J. Walker, research director at Gartner Inc, says: “Our 2017 Hype Cycle reveals three distinct technology trends that profoundly create new experiences with unrivaled intelligence, and offer platforms that propel organizations to connect with new business ecosystems in order to become competitive over the next five to 10 years.”

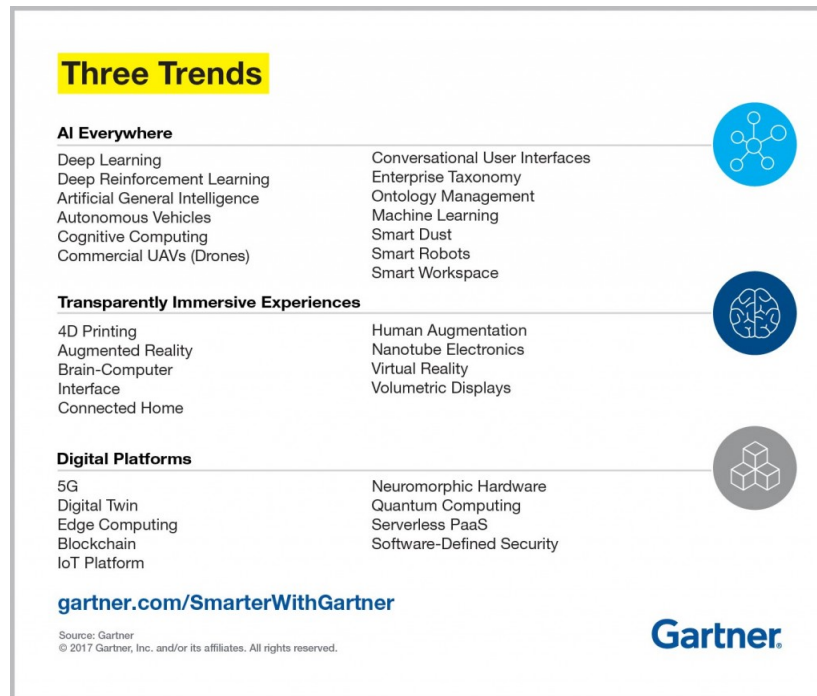


Figure 1.2. Most relevant trends appear in the Gartner Hype Cycle (Source: Gartner.com, 2017)

As we can see in the image 1.2 , Augmented reality is part of one of the most important technology trends. In this work, we decided to follow these trends and to do a step in the future. In the application it's been used the advantages of human interaction of the newest AR device. Additionally, we consider to develop an educational but futuristic app that is perfect for every age.

Nowadays, often AR is used for indoor, where is very simple to know the user position and to track some elements, moreover, the light condition can be controlled. This type of application is being presented from several museums in last years, it allows user to frame images, recognize his position and orientation and overlap virtual objects to the artwork the user interested in. Obviously, an outdoor system presents much more difficulties and challenges. Most important known problems are:

- Accurate user tracking
- Impossibility to alter the environment
- Light conditions

One of the usual solution for the first problem is a GPS-based solution, but usually it might be not able to provide an accurate position and orientation of the user. Furthermore, must be also considered GPS-denied environment. For these reason, in this thesis is presented a Mixed Reality application for outdoor, that make user's cultural tour very immersive and interesting. This innovative application enhances the user experience with synthetic contents that interact with the real environment, thanks to the cloud computing.

1.4 Context

Over the last years media and entertainment are really changed. Some years ago, people can only passive consume the media and entertainment contents, but now they can interact, share and talk with that contents while walking or running etc. Starting from this revolution, devices and media contents are more portable every day, trying to develop entertainment system very smart, simple and useful.

This evolution is sustained also by the technological development, that allow the companies to use larger storage, more powerful processor and better connectivity. Connectivity is one of the most important part of the media applications and in this last years it is very grows. Thanks to this, the applications can offer much more real-time services and useful interaction to the users using the new network capabilities.

5G should be the perfect technology that answer to all the features requested by the developers like in Figure 1.3. This thesis project is perfectly inserted in this context, because of its smart and useful cultural service that needs a performant connection to complete its tasks.

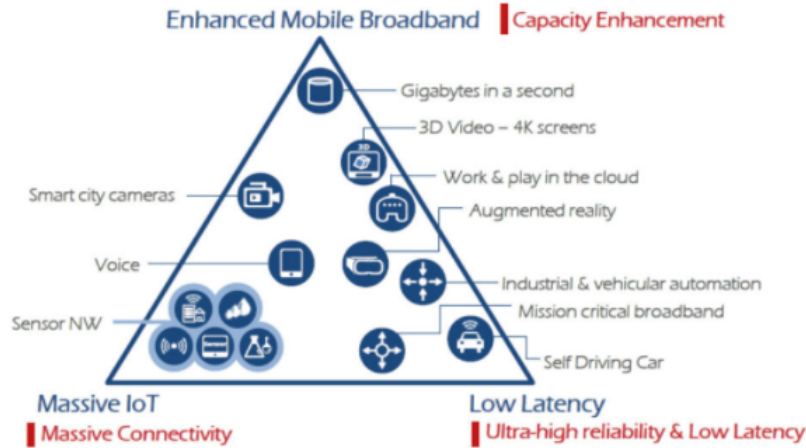


Figure 1.3. 5G capabilities expectation (Image source: ETRI graphic, from ITU-R IMT 2020 requirements)

1.4.1 5G Networks

5G is the name of the fifth generation of cellular networking. In the next future, smartphones and tablet will use this technology for transfer data, and people will see 4G networks very slow and irrelevant. So 5G is an improved mobile network that will much more flexible and performant than the previous systems.

The aim of this powerful technology is to provide and enable a fully mobile connected society, with specific services based on individual customer needs. It is important to know that the development of the 5G has been ongoing for some years, but the standardization is started recently.

Major mobile companies are working together in order to define the technical specification that 5G should have 1.4. However, the essential features and functionalities will be:

- Very low latency
- Ultra fast speed data transfer
- High number of connected devices
- Energy efficiency

With these features 5G is a revolution for Media and Entertainment in the world, because it will be capable to accomplish every company's demand.

And it will be very scalable and adaptable to every needs or user cases.

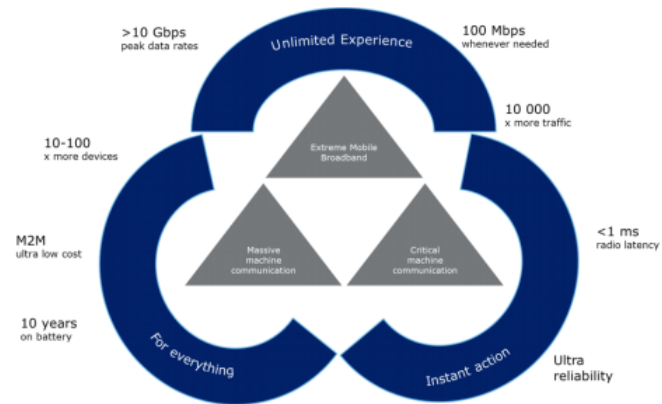


Figure 1.4. Expected performance of 5G networks (Source: Nokia 2017)

Chapter 2

State of the Art

In this chapter will be introduced the differences between the major techniques that allow the user to bring virtual content in the real life, namely virtual reality(VR), augmented reality(AR) and the newest term mixed reality(MR). This last term was created by Microsoft and specify a particular case of augmented reality. Later on, different uses of the augmented reality are presented shortly.

2.1 Virtual, Augmented and Mixed Reality

In the past virtual reality was seen very far from the real life, like two opposite entities. Differently, the fast developing of technology in many sphere of people's life is getting this stuff much more similar, allowing users to live and to experience virtual and physical reality at the same time. A person can walk on the street, while talking with a friend using a new virtual device.

The term *Virtual Reality* (VR) was used for the first time in the 19th century and it was related to the theater and its illusions, but during the years this term has changed to a digital representation. Nowadays these topics are earning a lot of media attention and are promising tremendous growth, but what is the difference between virtual reality, augmented reality and mixed reality?

Real environment is physical life without any virtual representations of the world.

Augmented reality (AR) is an idea where the digital domain fit with the physical world. Information not only is passive seen by person, but when the user looks to an object he can see relevant information, amplifying his intelligence. The view of reality is enhanced by a computer and get richer

one's current perception of reality. Usually the augmentation is done in real time in order to take the right information about the environment and perform a good overlays. AR technologies can transform simple information in interactive and digitally manipulable.

Augmented virtuality (AV) is inverse of Augmented reality. This is an enhanced virtuality, so its aim is to merge real objects in a virtual environment. In this method there is a virtual space, and physical object or person are integrated and can interact in realtime with the virtual environment. It can be possible only with the usage of different technique, physical peripheral that is related to a virtual object or real object seen by any cameras. Most of the elements in the Virtual environment or virtual reality are computer generated, and often they are connected to an input user peripheral in order to interface with the user. Generally, many computer games and applications are part of this category.

Mixed reality (MR), as we can see in 2.1, is a evolution form of augmented reality, that feels to the user a very immersive experience. Mixed reality devices usually have some sensors that allo them to perform an interaction beetwen virtual object and real world. This term was created by Microsoft before Hololens presentation.

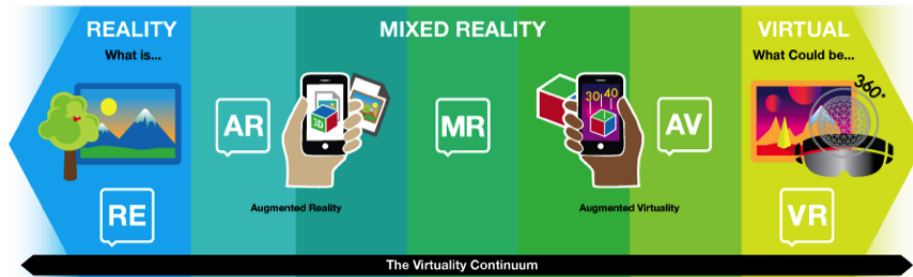


Figure 2.1. Virtuality Continuum

The function of sensor networks and virtual worlds allow the real and virtual worlds to work together as a new interactive medium.

2.2 Augmented Reality Applications

Augmented reality applications have been developed to support many fields, like teaching, medicine, media and entertainment.

2.2.1 Gaming and Entertainment

With AR, the game can interact with the real world, and people can have a direct, personal, and emotional interaction as in traditional table or card games. In augmented reality, the player can influence a game with some input and feedback devices like in 2.2. Furthermore, the user can move around virtual object moving in the real world, moreover he can also interact with other participants.

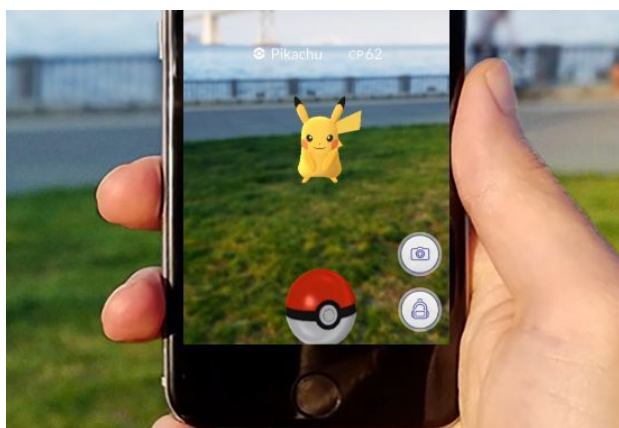


Figure 2.2. Pokémon GO

2.2.2 Teaching

Usage of augmented reality is becoming great also in the educational field, in various sectors including schools and universities.

Swiss researchers have developed an AR virtual chemistry laboratory. Students can view, learn and study the shape, size and methods of these atoms, interacting only with a simple drag and drop. With this application 2.3 the students can learn more about chemistry and can construct their complex molecules following the subatomic rules. This feature is a real advantage over traditional methods of learning models.

2.2.3 Medicine and Surgery

Many medical AR systems were developed and tested in the late 90's like the semi-transparent display for orthopaedic surgery.

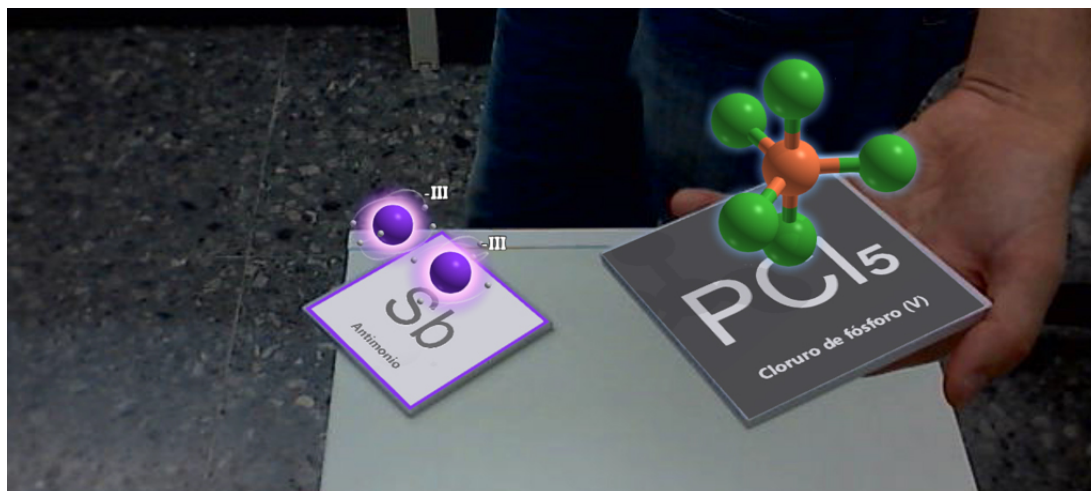


Figure 2.3. Chemistry augmented reality application for educational field

Nowadays, augmented reality is used for surgery and birth simulations, virtual laparoscopy, tracking surgery tools and visualization of anatomical joints in motion.

2.2.4 Augmented Reality for Television and Entertainment

Another important field that is changed by the introduction of Augmented reality is the television.

Augmented Reality can be used in the entertainment context in order to create games, but also make some aspect of important sports more visible. In these events where a huge number of spectator is reached, AR can also implement into advertisers to show virtual ads.

There are many sports(e.g. Football, race, basket) in which video see-through augmentation is used everyday from some years. One of the most important application of this is Fox-Trax system, using to make visible in every moment the hockey puck, but AR is also used in cars racing, swimming, etc. In this predictable fields the AR work very well, using the chroma-keying technique.

2.2.5 Augmented Reality for Cultural Heritage

In recent years the two global trends in museums are digitally mediated personalization and personalized learning. The most important museums in the world are using new mobile-only technology. Through mobile apps, museums can provide supplemental information about an exhibit or the museum itself; or as a personalized mobile guide through the museum collection or gallery spaces.

Now some museums are experimenting with AR apps inside the gallery spaces. For museum visitors, AR apps on their mobile devices will be very easy to use.

ArtLens 2.0

One of the most well-known AR apps that has been designed and developed by art museums is the ArtLens 2.0 by Cleveland Museum of Art.

ArtLens 2.0 is an AR app that uses image-recognition software to recognize a selection of artwork, it gives to the museum's visitors the chance to get information about the artwork they interested in before approach it. It was also designed to model and spark conversations among visitors as they respond to works of art in gallery spaces. Besides AR, the app contained mapping and beacon technology, enabling visitors to discover and create new pathways through the museum's collections. The implementation of different technologies has made the museum visit experience more entertaining.



Figure 2.4. “Skins and Bones” app in Smithsonian Museum (Source: The Smithsonian National Museum of Natural History)

Smithsonian Museum of Natural History

Smithsonian Museum offers to their visitors an AR application (“Skins and Bones”) that is a great example of using AR in a museum, and it is also a perfect educational tool. This application was developed for an old museum exhibition (Bone Hall) that had not upgrades since 1960. With this application a user can only scans some targets to see the skeleton of the target related animal.

ArcheoGuide

ArcheoGuide is a project aiming at providing a personalized virtual guide and tour assistant to archaeological site visitors. The system provides an Augmented Reality reconstruction of ancient ruins, based on user’s position and orientation in the cultural site, and Realtime image rendering. It includes a big database of cultural material for on-line access, virtual visits, and restoration information.

HoloMuse

HoloMuse is an AR application that use HoloLens device to allow users to interact with holograms in a simpler way. The models of archaeological artefacts are from the Anonymous Museum’s collection.

In the last prototype, users can choose a hologram from a holo-gallery, by targeting their gaze on a hologram and selecting it using a voice command or a tap gesture. Users can manipulate the hologram using tap gestures, and then place it within their field of view with a tap.

The user can learn more about an individual artefact and study it carefully, as well as to consider the relations between artifacts.

2.3 Augmented Reality Devices

Nowadays the virtual and augmented reality devices are very popular and there is a big expectation about their capabilities and features. Until now, people have always had to clash with the reality, divided between the futuristic vision of the cinematography and the low performance of the old augmented devices. But now CPUs are getting smaller and the performance of the new technologies are growing day by day, and devices’ features are much more similar to the film’s fantasy.

It is important to know that augmented and mixed reality are still in their infancy, indeed developers and companies are working hard on it and there are updates every week.

In this section will be introduced some of the actual available augmented reality devices.

Microsoft Hololens

Hololens is the device that is been used in the develop of this project, so it will be deeply examined in the next chapter, but it is one of the best all-in-one mixed reality device. It is fully standalone, it has some depth sensor to interact with the real environment and good see-through lenses [2.5](#).

Meta

Meta 2 and Meta 1 also are two augmented reality devices developed by Meta company. These two devices have see-through lenses and a 3D camera that allow them to recognize easily hand gestures or holograms in the real world.

Meta 2 seem to be one of the best AR device, because of his big field of view and his resolution, but it is tethered, so it needs a pc to work correctly.



Figure 2.5. Microsoft Hololens and Epson Moverio (Source: Michael Leyva)

Epson Moverio

Moverio is a very versatile device, it can reproduce VR and AR contents. It uses GPS, camera and gyroscope to understand the user position, because it has not a depth camera. It has two screen for the projection of the contents, but they are very small with a little field of view of 23 degrees [2.5](#).

Vrvana Totem

This is a new type of Head mounted display, because it is very similar to the conventional VR devices, but it can be also an AR device thanks to its front cameras.

With this system it can offer the big field of view available in the AR field (120 degrees), so the user doesn't see the real environment directly, but a reproduction of it through the cameras.

2.4 Augmented Reality SDKs

In this part will be introduced best and popular AR SDKs. They are of very different type, from simple to quite hard, or from cheap to expensive systems.

Vuforia

This is the sdk used for this project, so it will described deeply in the next chapter. Furthermore Vuforia is one of the most popular system that help developers in the Augmented reality applications. It is supported by all big platform and software, it offers many services for free (image target, object scanner, qr code scanner etc.).

EasyAR

EasyAR is another free and easy to use AR libray, it is an alternative of Vuforia. It is also supported by all major companies, but in the actual release offers image recognition only. A lot of people are waiting for 2.0 release that should implement object recognition environment perception etc.

Wikitude

In the latest version, this system offer a powerful SLAM (Simultaneous Localization and Mapping) solution for AR apps. Moreover it can perform image recognition and tracking, 3D tracking and Geo data. It is available only for iOS and Android, it's very expensive.

ARtoolKit

ARtoolKit is an open source AR libray, it has a good support and implements a lot of features like camera tracking, image target, square target etc. This system has a lot of possibilities but it has still a low user usage.

Kurdan

According to user's feedback and review, Kurdan seem to be the main rival of Vuforia, making the augmented reality very easy to use. It uses a SLAM technology, and can perform images and 3d tracking, but it is supported only by iOS and Android. Its free version is only for application testing, while its licence is very expensive.

Chapter 3

Technologies

In this chapter will be introduced the technologies used in this thesis work Figure 3.1, including hardware (HW) components and software (SW) tools.

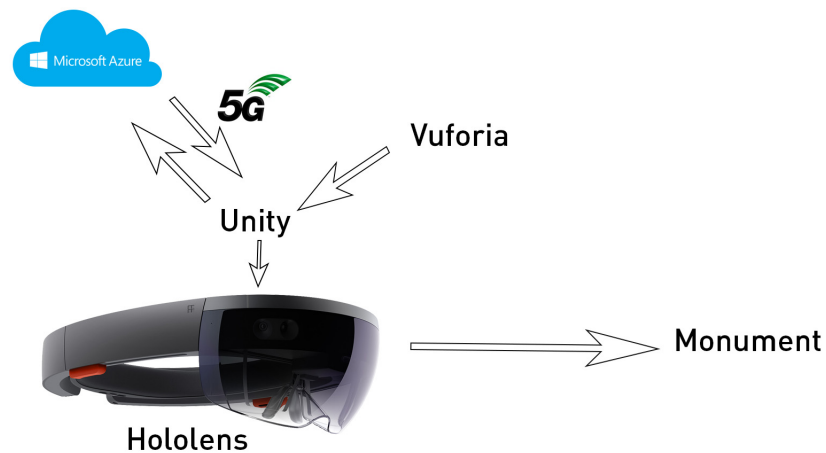


Figure 3.1. Technologies used in this project

3.1 Microsoft Hololens

The aim of this work is to understand features and the usability of this new AR device. Hololens 3.2 is one of the latest Augmented reality products. It overlays virtual objects over the physical environment and it can do this only

thanks to the multitude of sensor that recognize the user's gestures and the surrounding environment.



Figure 3.2. Microsoft's newest mixed reality device, Hololens

3.1.1 Hardware

This device is an untethered Holographic computer: it has own Central Processing Unit, memory and wireless network interface that allow it to work completely standalone. Furthermore, it has the Universal Windows Operating System that is also developed for mobile and PC. This make the user feel in a familiar environment, with the ability to navigate setting configurations and selection of files and applications [3.3](#).

The Hololens diverges from the typical PC interface with the input methods the user generates to navigate the features and with the display method that provides information and feedback to the user. To do that, it has a custom microprocessor centred around augmented reality, called Holographic Processing Unit. The HPU is used for managing sensors input, for recognition of the environment and for determining the graphical displays of the device to the user. One of the most important component of the Hololens is the See-through holographic lenses (waveguides) [3.4](#), that allows the viewer to see both the light of the environment passing through and the light generated by the holograms' presence. This type of lenses allow the device to render only the focused models. This also reduces the battery usage, as lighting methods are usually the biggest power draw of mobile devices, and

Specs

A small computer controlled by Windows 10	WiFi 802.11ac and Bluetooth 4.1
Holographic Processing Unit(HPU)	weight 579 grams and battery for 3/4 hours
4 logical 64 bit CPU	MicroUSB (not C) connector
various sensors - ambient light and others like "Mapping" sensors	Resolution 1268 x 720 per eye
HD camera of 2 MP	3.5 mm audio jack
4 microphones for voice recognition	Spatial sound via 2 HRTF
2GB RAM and 64GB of flash storage	Speakersets

Figure 3.3. Hololens' technical specifications (Source: Microsoft)

it ensure 2-hour battery without any further attachments, avoiding to show constant background rendering in the application.



Figure 3.4. Hololens' display system, Waveguide Lens

3.2 Unity 3D

Unity is a feature rich, fully integrated development engine that provides out-of-the-box functionality for the creation of interactive 3D content, and it is fully compatible with VR and AR technologies. Unity was born as a game engine, so it has some functional and performance limitations for normal applications. One of the most important features of this software is that it provides live changes; any change in the code can be easily shown on the

graphical interface. Unity's script can be developed in C sharp or Javascript, in this project has been used C sharp for simplicity and for the compatibility with UWP (Universal Window Platform). Furthermore, the best advantage of Unity is that allow you to develop cross-platform applications, so you can write one application and then deploy it for many devices or systems (Android, iOS, Windows). This gives the possibility to have more people use and interact with the app.

3.3 Vuforia

Vuforia is one of the best AR framework for mobile programming, and it has natural native support for the Hololens. "Vuforia uses stable, and efficient computer vision-based image recognition technique and provides several features, enabling capability of mobile apps and frees developers from technical limitations." *Amin and Govilkar, 2015*.

Vuforia represents a suitable tool for AR development, it supports iOS, Android and UWP. Combined with Unity it represents one of the best solution to do cross-platform. According to their official website, the Vuforia platform enables a variety of types of recognition. Through Vuforia the developers can bring synthetic objects to life thanks to the recognition and tracking of a broader set of objects. Vuforia Target Manager perform also an image improvement which in meantime leads to the optimization of the application performance. It also recognizes and tracks user defined images that make possible the interaction within AR environment.

3.3.1 Tracking

It is not stated explicitly by Vuforia how their tracking algorithms work, but looking at their documentation it is evident that they use some variation of natural feature tracking.

Natural features [3.5](#) in images are usually sharp details such as high contrast corners or the tips of a triangle. These corner or spot will cause large changes when shifted in any direction. By considering these changes in image intensity, the tracking algorithm can determine the features of the image, and then it creates a pattern of features. Vuforia uses the patterns to define a target.

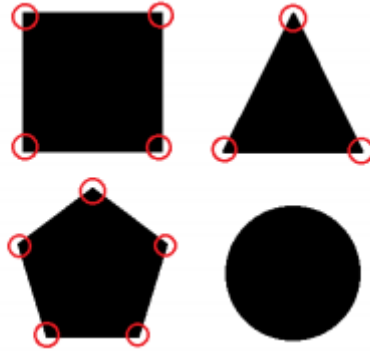


Figure 3.5. Natural feature tracking

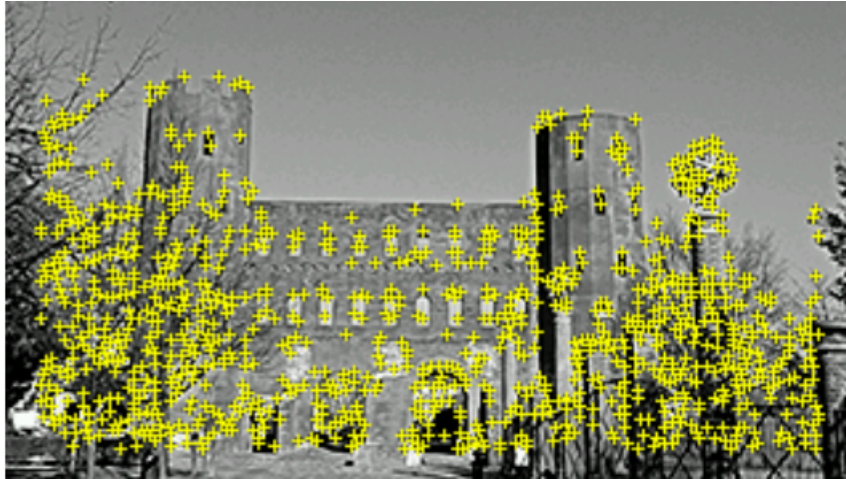


Figure 3.6. Natural feature tracking on Porta Palatina in Turin

3.3.2 Image Recognition

Using the Vuforia's Target Manager the developer can choose an image and use it as an image target. An image is a good target if it has a great number of features, if it has, Vuforia can detect the image easily.

In particular, good targets contain a higher number of unique features [3.6](#). A 2D grid picture has many features but the pattern is very repetitive, this makes impossible for the software to determine what is up, down, left or right.

3.4 Microsoft Azure

Azure is a cloud computing platform developed by Microsoft. Azure offer many solutions for building, deploying and managing in a cloud environment. In particular, in this project, it's been used a Linux virtual machine on a Azure node.

The virtual machine has been configured through a web based portal that is used also in deploying and configuring different components of the environment. Azure has been growing a lot in the past few years and is under constant development with several preview features constantly available for customer feedback and testing.

3.5 CDVS Compact Descriptors for Visual Search

In todays age, the visual research is one of the most important trend in technology world. Until now, an internet research, for the people's mentality is a text research, but with these algorithms the whole idea of search can change, in an image or visual search [3.7](#).



Figure 3.7. CDVS tracking (Source: sistemi-integrati.net)

CDVS (Compact Descriptors for Visual Search) is a new visual search MPEG (Motion Picture Expert Group) standard, that consist in a robust

technology for a performant and fast visual search in image databases. SIFT features are the base of this method, in fact the global and local features extractors are the CDVS core.

The fundamental elements to perform a good visual search is to find some features inside the images. This algorithm scan an image in order to recognize key points, that are descriptors of the picture. Every key point must be unique, light invariant and invariant to the geometric transformations.

This is an important feature in order to recognize the image by night or in sunny or cloudy day. All key points create a image code, that allow us to recognize it with descriptors. This technology can do a fast image matching, it takes the image that receive in input and compare it with all the images in the database. The algorithm uses a lot of local descriptors and a global descriptor that describe the whole image. If some key points are equals, the system recognize a matching and return the result. This method can seem slow and heavy, but it is accomplished in a fast and efficient way.

3.5.1 Key Point

The fundamental point of this technology is the creation of a detailed descriptor of the image, it is possible only with a right determination of the key points.



Figure 3.8. Key points descriptor system (Source: sistemi-integrati.net)

A good key point should represent a characteristic feature of the picture, something that makes it almost exclusive. So it can't be created in a flat

part of the image, because in this way it couldn't add any information to the descriptor.

The best way to chose a key point is in a point that shows significant changing in every direction, like an edge or a cross.

Chapter 4

Augmented Reality Tourist Guide

Enhancing Cultural Tourism by a Mixed Reality Application For Outdoor Navigation and Information Browsing Using Immersive Devices

The future belongs to those who believe in the beauty of their dreams. *Eleanor Roosevelt*

4.1 Project Requirements

The project described in this thesis has been proposed by RAI Radiotelevisione Italiana. The aim of this project is to develop an application that use the augmented reality, the motion awareness and computer vision algorithm to do a smart guide. In particular, the system will be used by the tourists to have some information and additional content about monuments and sculptures. The best device for this project seem to be an HDM, because insures immersive experience that allow the user to learn history in an interesting way.

The device that has been used in this project is Microsoft Hololens. The requirements and features of the project are in the following:

- Recognize the monuments
- Show to the user some information and history
- Possibility to see additional content (pictures, movies)
- Virtual reconstruction of the monument

First point is the most important, because it needs object and image recognition in order to recognize the monument and then ask information to the server about it. Obviously if the monument is not recognized no one of the other point can be reached. After that, the user can see the initial information that allow him to decide if he wants more details and content, or close it and search for other monuments. If the user is interested to know more about the monument, he can choose between some text details, image or video content. The last requirement is also the main theme of the thesis: overlaying a 3D model on the real monument that shows its ancient glory.

4.2 System Architecture

In this section is discussed the system's architecture from every point of view, including hardware and software needed to create this project.

4.2.1 Hardware Architecture

In order to accomplish the requirements and for the device available, Hololens has been used, that is a see-through AR device. This is a very important feature because the user should see the monument that interests him. Moreover every user's input are managed by

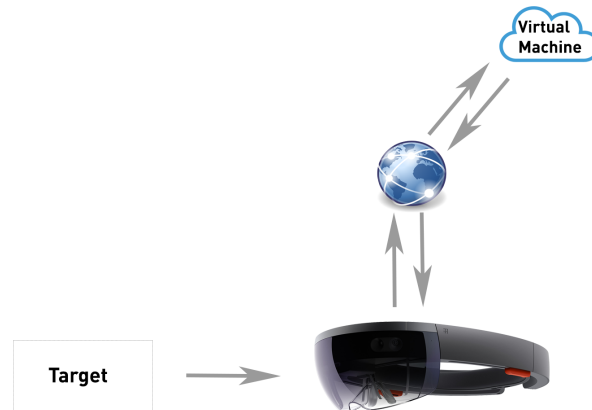


Figure 4.1. System's hardware structure

Hololens itself, thanks to its sensors, allowing an easier and intuitive interaction.

On the Virtual Search server is present a visual search software that compare the picture taken by Hololens with its database to find matches like in Figure 4.1. The virtual machine is installed on an Azure node in the cloud. Hololens can interact with the virtual machine through an internet connection (WiFi or 5G module).

The target of the user should be a monument or sculpture, but it could be also an image of a monument or something like this. It is a very important component firstly because without it the application is useless and also because it is a position reference in order to superimpose the 3D model.

4.2.2 Software Architecture

The application is completely developed in Unity, but it is logically divided in two parts: First is the connection with the virtual machine and the visual search, second is details and 3D interaction about the monument. The first part is composed by 2 different layers: Unity and UWP (Universal Windows Platform). While in the cloud part

there are 3 layers: Java Web service, the Image recognizer software and Linux 4.2. Differently the second part seems like the first but has one more layer: Vuforia, Unity and UWP.

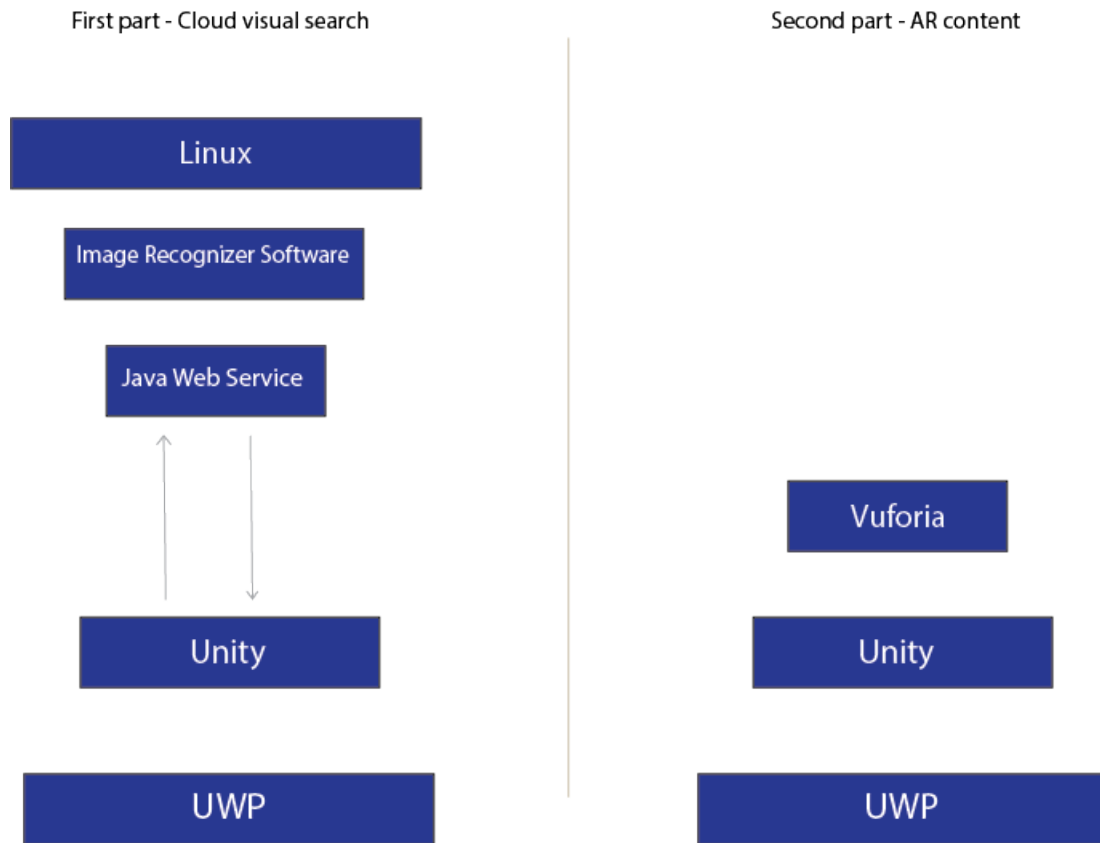


Figure 4.2. System's software structure

Vuforia is a Software Development Kit (SDK), discussed previously, that allows the developer to use its functionalities to create an AR application. Unity is the Game engine used to develop the Hololens' software and it is linked to Vuforia in order to integrate the AR feature.

The web service is developed in Java with Eclipse IDE and uses RESTful API. RESTful API is an application program interface (API) that uses HTTP requests to GET data. Usually developers use REST technology because of its low consumption of bandwidth.

Functionalities

The main functions of the system are the following:

- Recognize a monument
- Show text details about monument
- Show image or video content about monument
- Show virtual reconstruction of the monument
- Interact with the 3D model of the monument

Now will be discussed these points, their implementation will be explained later.

Recognizing

Recognize the monument is the first operation that the application does, it is completely automatic and invisible to the user. When the application is running, it gathers several times per minute images of the user's view to the cloud, until a monument is recognized. After that the user can choose if he wants more details or search for other monuments.

Text Details

When the user is interested to a monument the first function available is text details. This option shows him an information table about the monument he chose. After that he can simply tap on the table and turn back at the monument's menu.

Media Content

The second option presented to the user is media content. In this part he can see video and image about the history of the monument, for example streamed from Teche RAI. The user can stop the video in every moment and return at the monument's menu through the menu button.

AR Content

This function is the most important of the system. If the user chooses the AR content, the application should find the real position of the monument and overlays the 3D model of the ancient monument to it. So, the user can see the reconstructed monument in real size from every point of view.

4.3 The Application

Every element in Unity is a `GameObject`. Every part of the scene like models, lights and camera are `GameObject`. A `GameObject` can be also a container for a script that implement the real behaviour of the component. It's possible to associate a lot of different components to a `GameObject`, because it depends from the role that the `GameObject` has. Unity provides many different components ready to use, like `AudioSource`, `VideoPlayer` and `BoxColliders`, but if need to have a custom component, like an internet connection, this can be done only with the creation of a `Script`. Every particular behaviour or action of a `GameObject` must be implemented by a `Script`.

Moreover in Unity is possible to divide the application in some scenes, in particular in this project there are three scenes, first for the cloud visual search, second for the monument details and third for the augmented contents.

4.3.1 First Scene

The recognizing part is invisible to the user, so there aren't visible `GameObject` in this part of the scene, but there are some important invisible object as shown in Figure 4.3.

All this `GameObject` are empty `GameObject`, except for the Main Camera. Empties `GameObjects` are composed only by the `Transform` Component for specify their position and orientation. They are invisible to the user because they not have any visible elements.

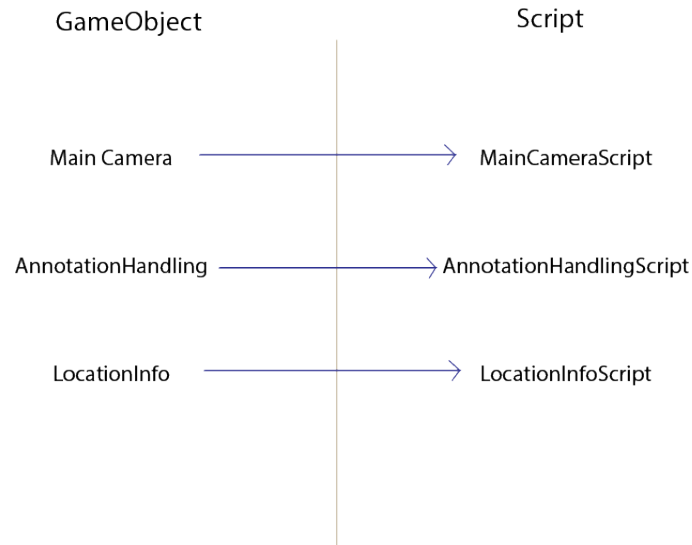


Figure 4.3. GameObjects and their scripts in the first scene

In fact their function is not to show something but to manage some custom behaviour.

The Main Camera is a GameObject imported from Microsoft Hololens SDK, which has some important settings that allow the best usability and comfort of the Hololens. It is responsible for the part of the user's eyes in the virtual world, so the user sees what is visible from the Main Camera. The AnnotationHandling empty GameObject is responsible to take a picture every 10 seconds and manage the connection with the Virtual Machine in the Cloud. Finally, when the monument has been recognized, the information about it are stored in the LocationInfo empty GameObject. This GameObject pass through the scenes because the GameObjects in the second scene need it to know the information about the monument.

AnnotationHandling GameObject

This GameObject is the most important of the first scene, it uses the Microsoft Driver to use the camera to take a picture and it has to manage the connection with the Web Service. When a picture is taken, the script manages the data, it transforms the picture in byte

and then put them into a Unity Web Request, as a Multipart/form and send it. After that, the script waits for a response from the Web Service, and parses the answer [4.4](#).

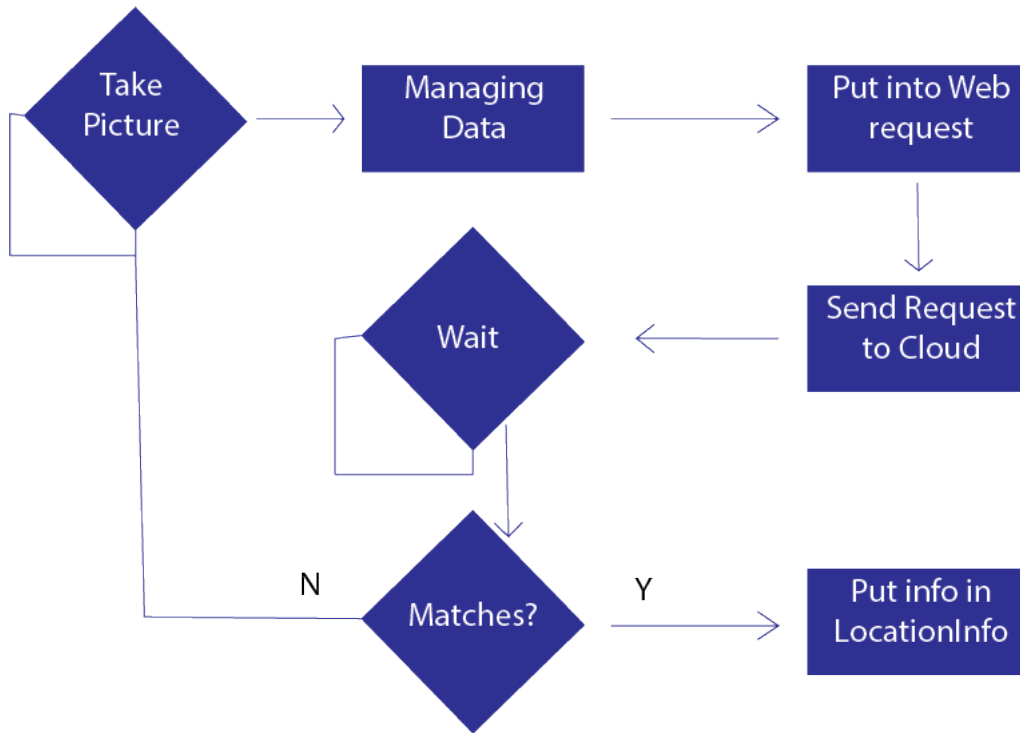


Figure 4.4. AnnotationHandlingScript Flow

AnnotationHandling Script First operation of this script is to Set up the Hololens camera to prepare it to take a picture, the interval between the pictures is editable in the configuration file. In the Start() method a delegate for taking asynchronous picture is created, every camera parameters are initialized and it's called the method for taking picture with the InvokeRepeating() command in order to call it every some seconds. So every interval ExecutePictureProcess() starts its execution, where the picture is taken and every bytes data of the image are put in a buffer that will passed to the UploadFileCo() method. This method firstly creates a web request setting the url at which send the image. Then, it puts the image

buffer in the request, send it and wait for the response. When the answer is captured by the script, it has to parse it.

Now the answer is converted to a string and the script analyses it:

- “No matches” nothing happens, and the script return to the main execution
- “Empty string”, probably an error occurred in Web Service, so the script return to the main execution
- “Match – NameOfMonument”, Image recognizer found a match, so the script call ParseResult() method that writes in the right way the result string.

At the end of the script, the information about the monument are sent to LocationInfo GameObject, and it’s called the second scene for the monument’s details.

LocationInfoScript is only a way to store every information found about the monument through the scenes. It is initialized by AnnotationHandlingScript while ManagerScript in the second scene take the informations from there.

4.3.2 Second Scene

This scene, differently from the first, has many visible object and in general more GameObject, Figure 4.5.

ARCamera Game Object has been imported from Vuforia SDK. It is responsible to add Vuforia AR features to the Hololens Main Camera. The AR virtual reconstruction will be discussed later on, but it’s important to know that ARCamera is a most important point of that feature. The Cursor is an important GameObject from the usability point of view. It works like a viewfinder, every interaction between user and software are done with the gaze of the user and targeting could be impossible without this component. The Cursor is a prefab of the Microsoft Hololens SDK, but its script manages the

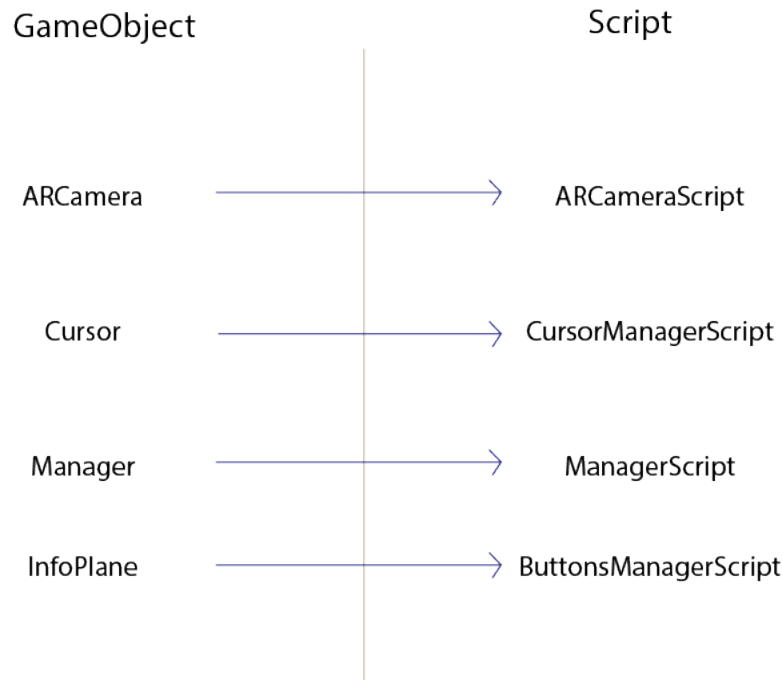


Figure 4.5. GameObjects and their scripts in the second scene

targeting objects in the scene and this information is taken by the manager when the user perform a tap. A prefab in Unity is a custom GameObject, created by developers that can contain 3D objects, script and other component. The prefab is very important in Unity because it allow developers to store a complete GameObject that they can use everytime they need. Moreover every modifies made to a prefab are instantly reflected in each instances of it. Manager is the most important Game Object in this scene because it manages all the visible objects and the interaction between user and models.

The first operation of the ManagerScript, as shown in Figure 4.6, is to verify if a monument has been found, if it is, the script take every information from LocationInfo GameObject and call his first method, DisplayInfo(). DisplayInfo create a new instance of the InfoPlane prefab taken from Resource folder. Then it prepares the material related to the monument and attach it to the InfoPlane.

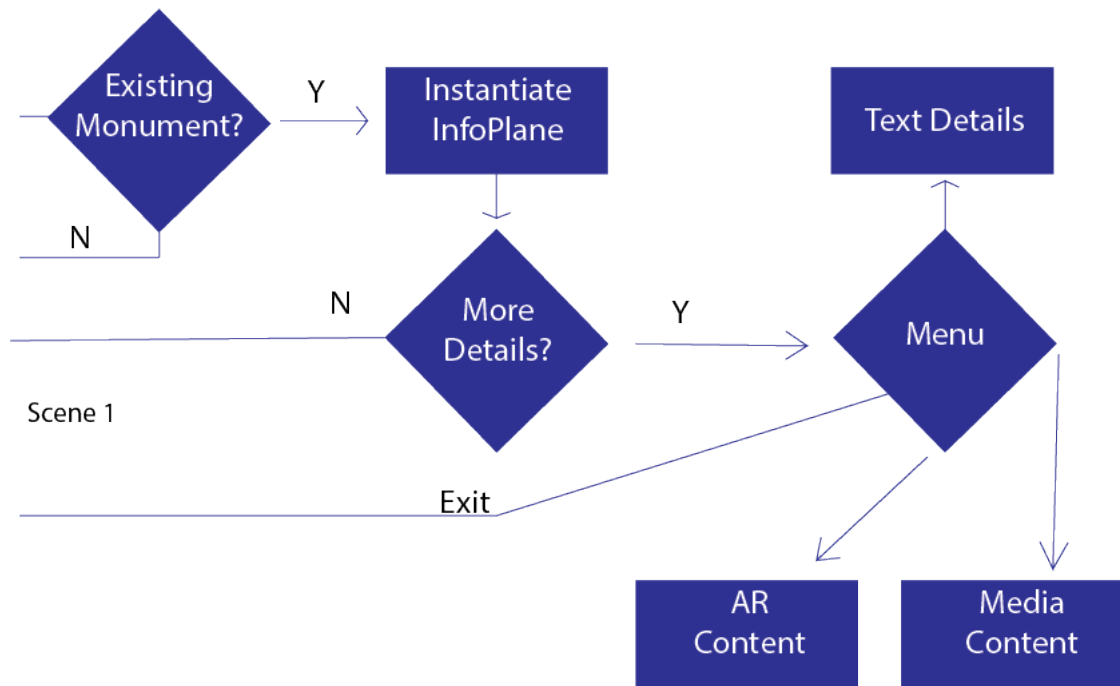


Figure 4.6. ManagerScript flow

Finally, the method attaches Billboard component and SimpleTagAlong component to the instantiated object. Billboard is a component that modify the orientation of the InfoPlane in order to keep it always facing the user. While a Game Object with SimpleTagAlong never fully leaves the user’s view. This component allows the holograms to follow the user around the space.

When a method finishes its execution, Manager wait for a new user’s input.

The others methods of the ManagerScript are:

- DisplayMenu()
 - Creates a new instance of MenuPlane
 - Places it in front of the user
 - Adds SimpleTagAlong component to the menu
- DisplayDetails()

- Creates a new instance of the DetailsPlane
- Applies the material related to the monument
- Places it in front of the user
- Adds SimpleTagAlong and Billboard component to the object
- DisplayMedia()
 - Creates a new instance of the MediaPlane
 - Loads the video content related to the monument
 - Plays the video
- DisplayAR()
 - Creates new instances of the Vuforia image target related to the monument
 - Displays the virtual reconstruction
- ExitScene()
 - Loads the first scene when the user taps on the Exit button

Every method of the ManagerScript is called by a user's input, in each Panel that user views, he can do a choice. The execution of a method is the consequence of user's selection.

InfoPlane GameObject is a panel that shows some information about the monument. So it has a text part that describes the monument and two buttons:

- Menu
 - If the user is interested to the monument, he can know more by clicking here
- Exit

- If the user wants to search for other monuments, he'll click here

MenuPlane GameObject is the main menu of the recognized monument. From here the user sees 4 buttons:

- Details
 - Shows DetailsPlane of the monument
- Media Content
 - Shows a video about the monument
- AR Content
 - Shows the virtual reconstruction of the monument
- Exit
 - Exit the menu and search for other monuments

DetailsPlane GameObject is a panel that shows some specific information about the monument. When the user has read the text, he can only click on the Menu button to return to the menu.

MediaPlane GameObject shows a videoclip about the history of the monument. When the user wants to return to the menu he can click on the menu button, that is the only visible.

While, if the user selects *ARContent*, manager store every useful information and load the next scene, Augmented reality scene.

Each panel have not a proper script, but their buttons have the ButtonsManagerScript. This script sends the name of the button clicked by the user to the Manager GameObject. The most important method of this script is OnSelect(), that starts when a button is clicked and send the right data to the Manager.

4.3.3 Third Scene

At the start of this scene there are very few elements, only the AR camera and ARManager, but a lot of gameobject are instantiate by the manager once it has the location value. When the scene is loaded first two created elements are the camera and the manager. In the Start() method of the ARManager it has to check if the LocationInfo object exist and if its location variable is valid. When ARManager get a valid location, he starts to instantiate every gameobject that needs.

First prefab instantiated is the silhouette, based on the location variable. Silhouette is a semi-transparent image of the monument recognized that helps the user to align and to see the monument from the right perspective in order to allow Vuforia to work. In particular, ARManager create a hologram plane and attach it the monument image, then put it as a child of the ARCamera to be sure of the position, always in front of the user.

At this point ARManager has to instantiate the image target related to the location found. Image target is a component of Vuforia SDK and is composed by the image that Vuforia use to recognize the monument and show the 3D model. In this case the manager creates 5 different image targets with monument's images took in different light condition and different angles in order to increment the recognizing rate. Each image target has a SilhouetteManager that recognize if a target finds the monument.

After that, ARManager instantiate the 3D model of the monument recognized, put it in the scene and disable its visible until any image target tracks the monument. When a target raises a tracking event, silhouetteManager makes invisible the silhouette plane, makes visible the monument 3D model and sets it as a child of the target that found the monument. From now the target has all the control of the 3D model position and put it on the real monument to perform

a virtual reconstruction. If the target loose monument tracking, silhouetteManager sets the 3D model invisible again and makes visible the monument's silhouette to restart the tracking procedure. If the user wants to see a smaller 3D model, the model is shown in front of him and he can examine it from all perspective and views.

4.3.4 Web Service

Web service is an application completely loaded on an Azure Node on the Cloud. It is developed in java in Eclipse IDE, with RESTfulAPI.

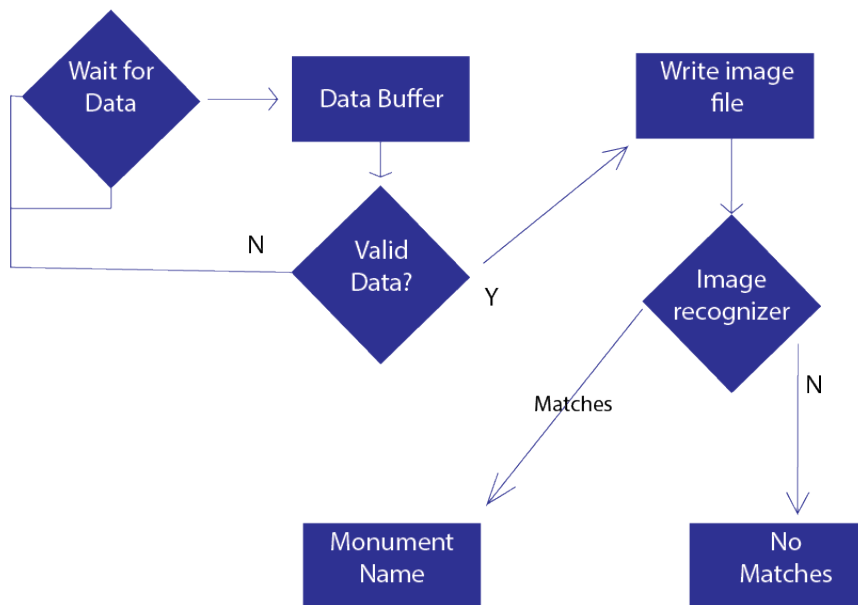


Figure 4.7. Web service flow

Web service is always waiting for data from the Unity application. When the data is arriving, the web service puts the bytes in a buffer and at the end it checks if they are valid. If the bytes are invalid, they are discarded and the web service return to waiting. While, if the bytes are valid, the web service creates an image file and send it to the CDVS image recognizer. Next the web service waits for the

CDVS result, and then it sends the result to the Unity application, Figure 4.7.

4.3.5 User Interface

The experience of human interface with digital application is started with simple peripherals like keyboard or mouse. Differently in the last years the people learned to interact with touch device like smart-phones or tablets using simple gestures. The aim of this project and also of the Microsoft Hololens is to simplify the experience and change it in a natural interaction using user's hands to interface with the virtual object.

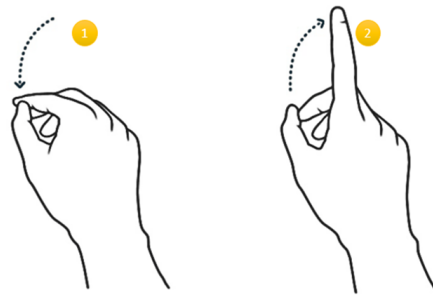


Figure 4.8. Air tap gesture

In this application the user can perform one gesture in order to keep it simpler as possible. The selection or interaction with object is done through the user's gaze and the “air-tap”, so the user aim to something with the eyes and select it performing a tap.

In the first scene the image recognizer is invisible to the user, so there isn't any user interface. When monument is recognized a type of interaction appears. InfoPlane is the first panel that the user can sees. This is a very simple panel, as shown in Figure 4.10, with some general information and two big buttons to make simpler user's choice.

The MenuPlane 4.11 is a simple panel with four buttons. It is the



Figure 4.9. Gaze targeting technique

main monument menu and from this point the user can see every available content. The buttons are always big to make usability one of the most important feature. The others two panel are very similar to this, with one or two buttons and the same style.

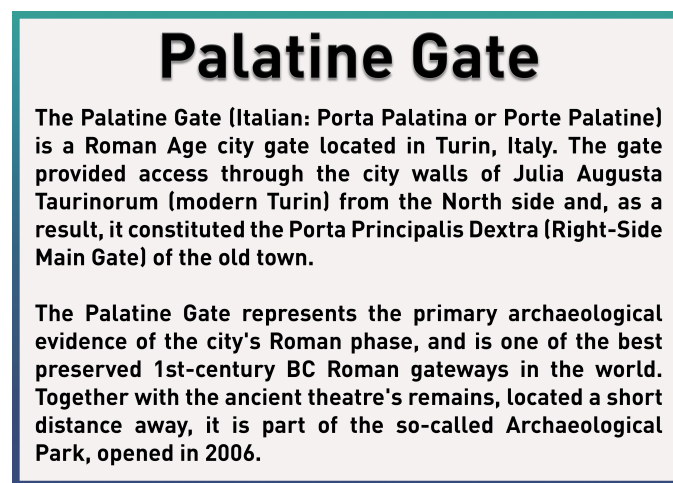


Figure 4.10. InfoPlane Example

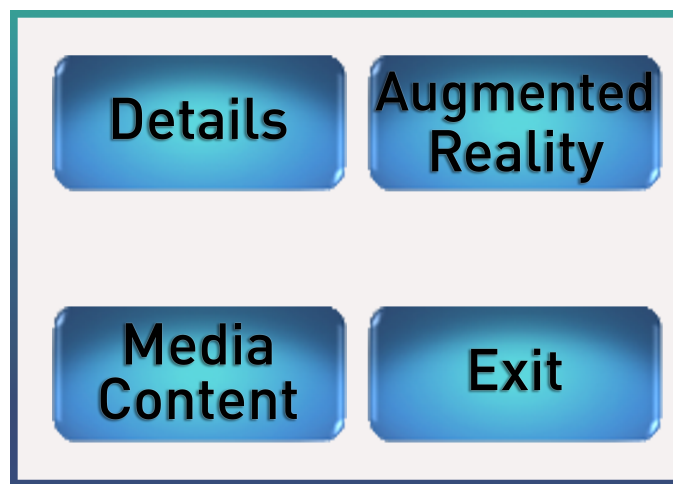


Figure 4.11. MenuPlane Example

Chapter 5

Test and Results

5.1 Test

This project was born with the purpose of inform tourist about the monument and the point of interest that they are seeing in their city tours. So, to understand if this system can really help people to visit and learn more about the cities, some tests have been organized at any monument in Turin.

Tests are focused on various type of parameters, they are being created in order to analyse both subjective and objective impressions. In particular, the usability and the simplicity to use all part of the application are the most interesting target of this survey.

To do this, a questionnaire is being created with some questions and an answer scale from 1 (totally disagree) to 5 (totally agree). The aim of this research is not only to try if the application work correctly, but also to understand if its behaviour is stable when the users perform strange actions or walk freely in the environment.

5.2 Testers

An important part of this application is that it could be used by all type of user, from expertise to simple tourist, so the testers are being choose as random as possible in order to try every aspect of

it. The group is composed half by women and men, the age range is between 20 and 50 years old and they have different knowledge of AR technologies. A half on participant never tried an augmented reality application, and some of them have not any familiarity with computers. All testers were volunteer and their test and participation were not compensated.

5.3 Location

The location choose for the test is the Archaeological Park in Turin, and the target monument are the Palatine Towers. This place is impressive, spacious and very popular for the tourist that come in Turin. Palatine Towers was the Roman Age city gate, that provided access to the ancient city Julia Augusta Taurinorum, so they are a perfect example to try this smart guide application. The users had to simulate a normal walking through the Archaeological park, until the Palatine towers. When they arrived at the monument the application recognized it and then users can navigate all information like in Figure 5.1. After the starting tutorial, the users were alone in the use of the application, and they had to accomplish all actions:

- Read initial information and tap on menu button
- Tap on Details button and read some additional information (Figure 5.2)
- Tap on Media Content button to see the video content about the monument
- Tap on Augmented Reality and try to examine the 3D model of the monument

Moreover, the users had to pay attention to some other factors in order to evaluate the usability and visibility of the system:



Figure 5.1. Initial information after monument recognition



Figure 5.2. Some text details about the monument

- Easy to point
- Visibility
- Fluency
- Stability
- Speed between contents

Before starting the test, users were introduced to the pointing system of the Hololens and to the application's interface. Furthermore, they were informed about the free in movement and that they had to use only their hands to interact with the system.

After completing the visit of the system, they can choose to visit another monument or to exit. At the end the users fill the survey and their comments about the application.



Figure 5.3. Monument's 3D model overlay

5.4 Results

The questionnaire was composed by 16 questions and was divided in two categories:

- About usability and stability
- About the interface

Before the analysis of the results, some technical and software limit must be introduced, to better understand the results.

One of the most important issue, is the limited field of view like in Figure 5.4. Hololens' field of view is very small, only 30×17 degrees, because of the technology limit of the hardware, so the user can't see whole big object, but only part of its.

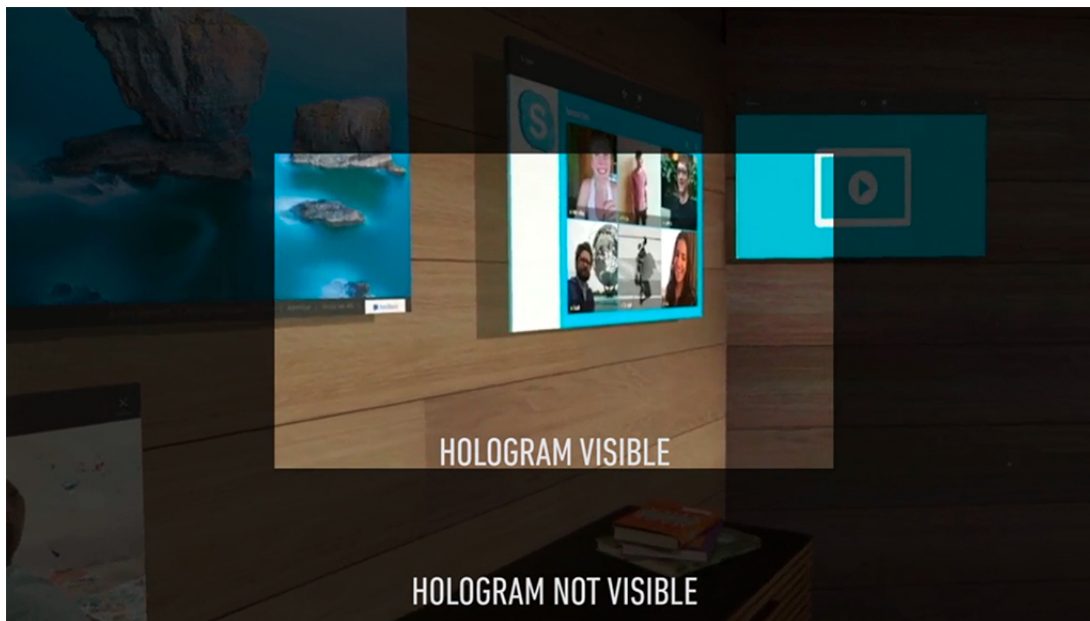


Figure 5.4. Hololens' field of view

This is the big problem of all AR devices and it can be solved only with new systems and research in the see-through technology. Moreover, the lenses' brightness is limited, so it makes the contents

not really visible when a user uses the application in a very sunny day.

Another issue that is been recognized by the users is the instable tracking of the virtual reconstruction of the monument. Nowadays not exist a free and stable building recognizer to perform augmentation, furthermore Vuforia is the only supported by Microsoft to develop application on Hololens. In this project is been tried to improve the Vuforia image recognition, but it remains not very stable yet.

If Vuforia try to compare real images with its image database, is obviously that every light change, or objects movement can compromise the result. However, Vuforia is in a very evolution époque, it has weekly updates and it is getting better every day, so when big object recognition will be implemented this issue will be removed.

5.4.1 Analysis Part

The questions are divided in two parts, one for the technician part and one for the interface and usability. First question is about the usability, in particular it asks about the simplicity and the intuitiveness of the application's commands.

From the first question [5.5](#) is simple to appreciate that all participants think that the commands and the general usability of the application are simple and intuitive. This is a most important result in this project, because the user type is very various and all of them should be able to use the application in an easy way. About this last thing, it is also important to know if the users find intuitive the symbol and the text of the menu and button interface.

Most users find menus and buttons intuitive or very intuitive, so it is also an important achievement in order to create a useful and simple application. Next questions are about the fluency [5.6](#) and the speed in the passage between the content shown, and the speed and stability in the initial monument recognizing [5.7](#).

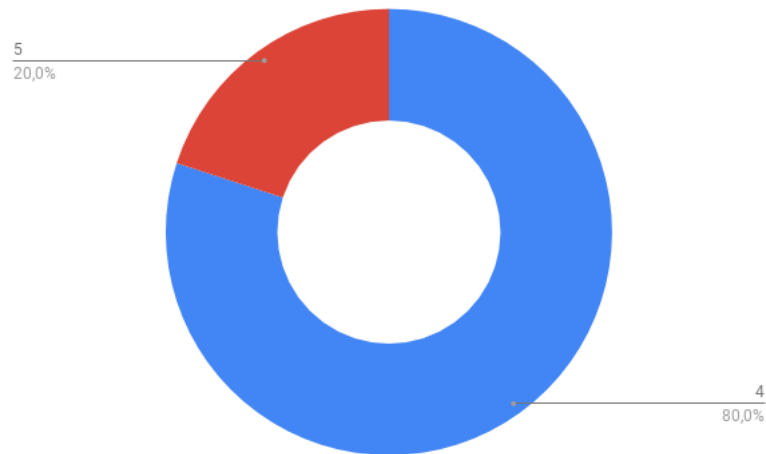


Figure 5.5. Answers to question 1 "Are commands simple and intuitive?"

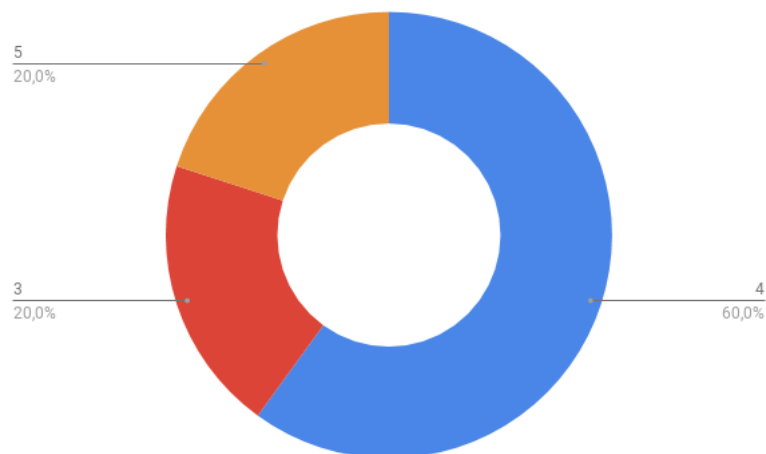


Figure 5.6. Answers about "Application's fluency"

In each case is very relevant to see that 80 percent of the testers say that the application is quite fluent and the initial recognize is stable and fast. The only inconstant results are about the Vuforia using and its tracking. As it is visible in Figure 5.8, there are very different result, based on the user's experience.

This variety of answers depends from the instability of Vuforia

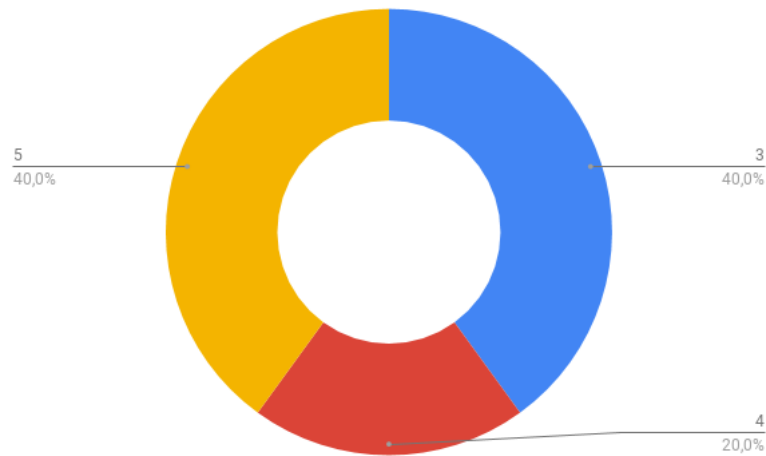


Figure 5.7. Answers about "Speed of initial monument recognition"

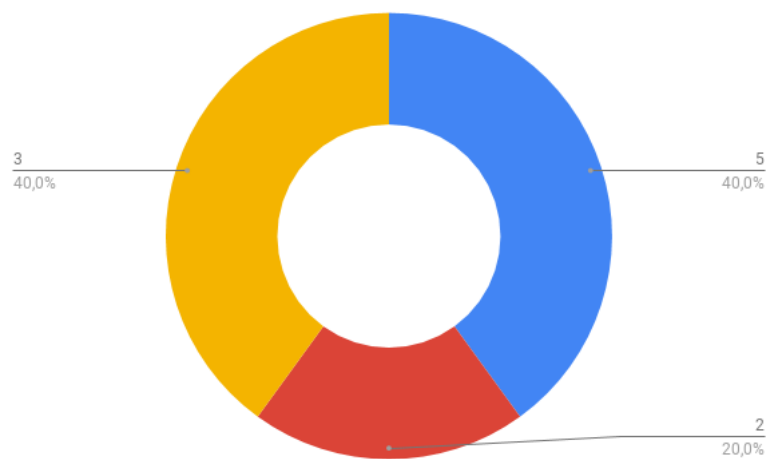


Figure 5.8. Answers about "Quality of Vuforia Tracking Stability"

image recognizing, that it can works for someone and not for others. However, a notable feedback that can be received from users is that most of them appreciated the application, and think that it could really help the development of the national cultural heritage.

Chapter 6

Conclusion

The develop of the project introduced in this thesis is been focalized on the creation of a smart guide that enhance the real environment and monument that tourists can see with text or video holograms, replacing the old paradigm of knowing monuments through paper guides or by Internet.

In the creation of this application has been used knowledge in different fields, like Augmented and Mixed reality, software development and 3D modelling. A study of web developing is being done in order to create the web service, moreover a research work is being accomplished to perform a good knowledge of the available tools and possibilities with these new technologies. The deep understanding of Vuforia and Microsoft Holographic library is being the hardest work.

The aim of this project is to perform a smart guide that help the tourist when he visits a city: the application should enhance the real monument with textual information and curiosity, with media content and with 3D model augmentation. The requirements are being accomplished, with some difficulties. Big problems were in the monument recognize and with the interaction between user – monument – holograms. Furthermore, an important difficulty there was in the Vuforia tracking, that is not very stable yet.

In the application is being decided to use only users' hands to interact with the augmented content to simplify the intuitiveness

and make all the system as natural as possible.

From the results of the users' tests it is visible that they generally appreciate the whole application, thinking that it can be a useful system to enhance the cultural heritage. Users also liked the pointing system and the natural hands interaction paradigm, that allow them to learn fast how to use the application and make them feel very immersed in the holograms.

Obviously, there are some problems, first of all the virtual reconstruction of the monument: Vuforia tracking, as it was written before, is not yet very stable, so users can't fully experience this part. Other issues are all related to hardware limits, such as the holograms' brightness in a sunny day, or little field of view. Surely, future technology developments in the augmented reality field will solve these problems. As said, users generally have been appreciated the whole application, but there are some improvements that can make better the user's experience.

Most important future improvement could be the use of CDVS standard to get the position of key points and use it to replace Vuforia in the virtual monument reconstruction. In this way the monument tracking should be very stable, light robust and position invariable. This modify can make the user's experience much more real and immersive, and allow it to examine 3D model from all perspective and point of view in order to better understand the monument's changing during the ages.

This project was born with the idea of creating a virtual tourist guide that helps people to become interested in culture and make their life smarter; this is just a small drop in the ocean of the future, but I am honoured to be part of it.

Bibliography

- [1] Vassilios Vlahakis, John Karigiannis, Manolis Tsotros, Michael Gounaris, Luis Almeida, Didier Stricker, Tim Gleue, Ioannis T. Christou, Renzo Carlucci, Nikos Ioannidis, *ARCHEOGUIDE: First results of an Augmented Reality, Mobile Computing System in Cultural Heritage*
- [2] Ding, Mandy, *Augmented Reality in Museums*
- [3] Kjeldskov, Jesper(2000), *Human-Computer Interaction Design for Emerging Technologies*
- [4] D.W.F. van Krevelen and R. Poelman (2010), *A Survey of Augmented Reality Technologies, Applications and Limitations*
- [5] Grahn, Ivar(2017), *The Vuforia SDK and Unity3D Game Engine*
- [6] Gottmer, Menno , *Merging Reality and Virtuality with Microsoft HoloLens*
- [7] Ding, Mandy, *Augmented Reality in Museums*
- [8] Xue, Hui (2017), *Augmented Reality Application for Training in Maritime Operations*
- [9] Microsoft, *HoloLens hardware details*, www.developer.microsoft.com/en-us/windows/mixed-reality/hololens_hardware_details
- [10] Emanuele Caimotti, Maurizio Montagnuolo and Alberto Messina (2017), *An Efficient Visual Search Engine for Cultural Broadcast Archives*, Proceedings of the 11th Italian workshop on Artificial Intelligence for Cultural Heritage (AI*CH 2017)
- [11] Christina Pollalis, Whitney Fahnbulleh, Jordan Tynes, Orit

- Shaer (2017), *HoloMuse: Enhancing engagement with Archaeological Artifacts through Gesture-Based Interaction with Holograms*
- [12] Sisvel Technology, Sistemi Integrati CDVS: *il visual search diventa interoperabile*, www.sistemi-integrati.net/articoli/cdvs-il-visual-search-diventa-interoperabile_ID_0_IDa_836.htm
- [13] Alexander Geurtz, Uwe Herzog (2016) *5G and Media and Entertainment*
- [14] 5GCity, *5GCity Project*, www.5gcity.eu/about-5gcity/description
- [15] Henrysson, Anders (2007), *Bringing Augmented Reality to Mobile Phones*
- [16] BTF-VR WORKING GROUP (2017), *Opportunities and challenges for public service media in VR, AR and MR*
- [17] O. Schreer, W. Waizenegger, W. Fernando, H. Kodikara Arachchi, A. Oehme, A. Smolic, B. Yargicoglu, A. Akman, U. Curjel *Mixed reality technologies for immersive interactive broadcast*