A PROOF OF CONCEPT MIXED REALITY APPLICATION FOR AUGMENTED CITY TOURISM

[1]Federico Debandi, [1]Roberto Iacoviello, [1]Alberto Messina, [1]Maurizio Montagnuolo, [2]Davide Zappia

[1]RAI, Radiotelevisione Italiana, Italy; [2]Dipartimento di Automatica e Informatica, Politecnico di Torino, Italy

ABSTRACT

Augmented Reality is gaining attention primarily driven by the availability of consumer devices such as Head Mounted Displays (HDM) on the market. This paper focuses on a different flavour of Augmented Reality, called Mixed Reality and it describes the work carried under the H2020 European project 5GCity. We realized an application running on Microsoft Hololens and designed to provide enhanced experience information on a city scale. The application provides information about historical buildings, thus supporting cultural outdoor tourism. The user experience is enriched with content coming from the archives of the Italian public broadcaster RAI. A cloud application (conceived and designed to run on a 5G-ready infrastructure) based on a visual search engine receives an image flow captured by the HMD by the user and identifies known objects. The user can freely watch at the object for which augmented contents have to be displayed and interact with these contents through a set of pre-defined gestures. Moreover, if the object of interest is detected and tracked by the mixed reality application, also 3D contents can be overlapped and aligned with the real world's one. Subjective evaluations confirmed that the application was fluent and the initial recognition was stable and fast.

INTRODUCTION

In recent years there has been a lot of activity around Augmented Reality (AR): many important players have shown support for AR by introducing HDMs, such as Microsoft Hololens, Samsung Odissey, Meta2, Vuzix Blade AR.

AR and Mixed Reality (MR) are two well-known concepts: the difference is that augmented reality "blends" together real and virtual objects while in mixed reality, objects are positioned and aligned in order to appear as a part of the real world with respect to the user view.

Moreover, with respect to Virtual Reality, the shared feature of AR and MR application is that users have contact with the real world and this has an advantage both from a technical point of view (part of the space exists and a computer-generated model is not necessary) and from a physical point of view (a detachment from the real world can lead to physical and mental discomforts) (1).

Virtual contents are called assets: a set of computer-generated objects overlapped to the real world. They can be: text, audio, 3D models, video.

This paper addresses the problem of aligning assets (such as 3D models) with real-world objects.

Several AR solutions have been used in the context of cultural tourism because of their potential to improve the tourist experience and help the tourist to access relevant information, improving their knowledge of the touristic destination and increasing the user's entertainment.

Yovcheva et al. (2) define augmented tourism as a "*complex construct which involves the emotions, feelings, knowledge and skills resulting from the perception, processing and interaction with virtual information that is merged with the real physical world surrounding the tourist*", explaining that it has not yet been fully exploited.

Moreover, authors conclude that the value added to the overall tourist experience is determined by the fit between context and content, referring to the spatial, temporal, personal, and technical context where the AR system is being used.

Many solutions have been presented for indoor (3), where is quite easy to track the user position and light conditions can be controlled. Several museums provide users AR applications that recognize position and orientation of the user by framing well defined images (markers) 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 (4). 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 (e.g. high buildings block the signal from the satellite in smaller side streets and close to buildings). The second issue is related to the impossibility to alter the environment by markers or target images. Finally, lighting conditions cannot be controlled.

This paper presents a MR application for outdoor environments enhancing the user experience during cultural tours, by anchoring synthetic contents to some positions in the real space and letting the user interact with them. In order to allow the user a full mobility, the Microsoft Hololens has been chosen as a HMD; it is connected by a high speed and high-throughput network (e.g., 4G/5G cellular networks) to a cloud infrastructure.

First, the application captures images of what the user is watching at a constant rate; then in a transparent way, images are sent to the cloud architecture where a visual search engine (based on MPEG Compact Descriptors for Visual Search - CDVS) is running. In this way, objects of interest framed by the user can be identified; finally, a notification about a recognized object is sent back to the user, that can receive augmented contents about the object.

The application can display both textual information and movies about the framed monument/building/artwork and a semitransparent silhouette of the recognized item can be used to align the user with respect the target, thus enhancing the tracking robustness. When the AR application tracks the object, 3D virtual assets can be overlapped to the target.

The paper is organized as follows: basic concepts behind AR and MR are presented in Section 2 as well as some examples of outdoor augmented reality for tourism, Section 3 details the architecture of the proposed application and Section 4 shows tests and results gathered in the Turin's Archaeological Park.

BACKGROUND

Due to the potential of enhancing the immediate surrounding, Augmented Reality has been considered to be of high potential for the tourism industry (5).

Several attempts of using AR and MR to support outdoor cultural tourism are known in the literature.

In (6) the authors demonstrate the use of augmented reality for collaborative navigation and information browsing tasks in an urban environment, having a direct impact on tourism domain. The information is presented to users via a head-mounted display, overlaying the real-world with a combination of text, graphical objects, 3D objects and images. Both navigation and information browsing functions support collaboration. Their solution uses a GPS receiver to determine the position of the system in outdoor environment.

The Archeoguide project (7) uses a laptop computer to manage computer generated assets and track the scene. In particular, the tracking strategy is based on a set of reference images due to the impossibility to put any special markers onto the historical sites. For each viewpoint there is a set of reference images that are compared with the framed one. When a match is found, the transformation to fit the reference image to the video image is computed.

This transformation allows to correctly align and overlap synthetic contents to the real world. A similar approach is also used in (8). The project Archeoguide has been also extended in order to provide users a powerful mobile device for outdoor AR applications with standard off-the-shelf components (9). At the date of (9) it was impossible to reach simultaneously the following design goals: intuitiveness, robustness, powerful and lightness.

Drawbacks and advantages of Mobile Augmented Reality (MAR) are presented in (10), whereas MAR using smartphones is considered in (11). Many other examples of application of the MAR paradigm for tourism are known in the literature such as the LIFEPLUS system (12) (which aims to provide a guide for the archaeological site of Pompeii in Italy) and the work presented in (13), which improves the tracking phase on the basis of a combination between computer vision algorithms and device sensors.

Several works have been devoted also to investigate issue related to the expected user experience. For instance, in (14) it comes to light how users are interested in applications that are easy to use and highly customizable, whereas also the emotional point of view is analysed in (15): MAR services are expected to offer stimulating and pleasant experiences, such as playfulness, inspiration, liveliness, collectively and surprise.

Moreover, many international bodies are working on AR/VR: DVB study mission report try to determine whether VR video is likely to be commercially successful; commercial requirements definition on smartphone–based VR devices is ongoing within the DVB CM-AVC group. W3C is working on WebVR an open standard that makes it possible to experience VR in the browser. VRIF (Virtual Reality Industry Forum) is working on the VR end-to-end chain guidelines, encouraging use of common formats for interoperability. MPEG is developing the Omni Directional Media Applications Format (OMAF) standard as well as the Media Orchestration Interface (MORE) for video stitching and encoding and MPEG – I (Immersive) targeting point cloud, 3D Graphics and light field.

PROPOSED MR APPLICATION

The device chosen for this project belongs to the Head Mounted Display family (HMD), as this kind of device provides users both an immersive experience and a full mobility. The

Microsoft Hololens fulfils the above mentioned characteristics as it is a standalone device that does not require an external computer for computation.

It has semi-transparent lenses and it uses a projection system to render a full-color 3D models at low latency to allow virtual objects to be drawn in the real world. Gestures and voice commands can be created and set to communicate with applications, while navigation can be achieved through gaze.

The detection step is the most important phase for any AR-based application. After the recognition of a monument/building/art work of interest, information gathered in the RAI archives can be sent to the user. The users have to be able to select the wished kind of information (text, audio, video). Only when the recognized item is correctly tracked, 3D contents can be aligned and overlapped to the real object.

From the hardware point of view, the proposed architecture is very simple (Figure 1). The user wearing the Hololens is connected to the Internet through a high speed network (e.g., 4G/5G). The MR application is running on the glasses and it exchanges data in a bidirectional way with a visual search application deployed in cloud on a virtual

Figure 1 - Hardware Architecture

machine. For this purpose, the MPEG CDVS was integrated as a proof-of-concept however any other similar tool can be seamlessly plugged into the system.
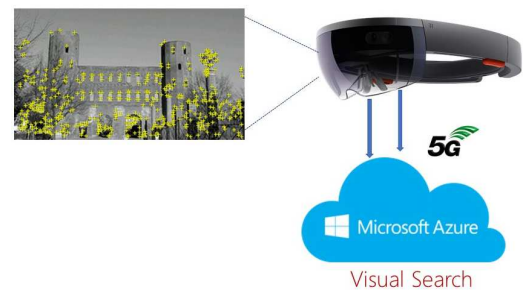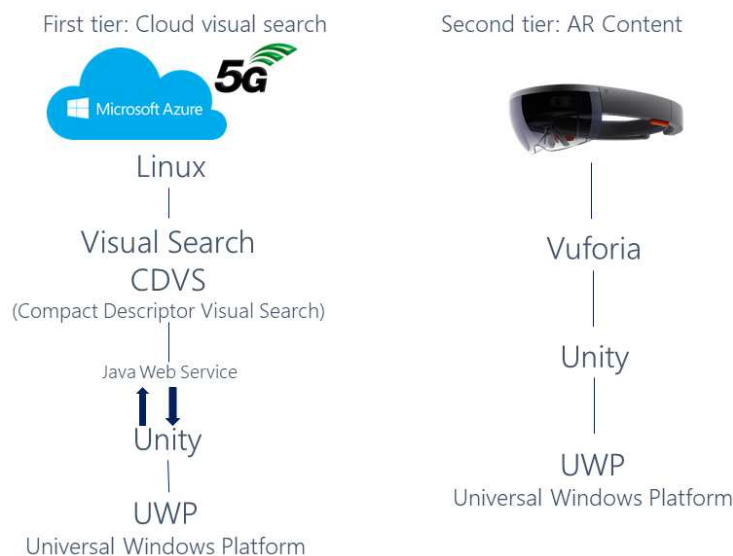

Figure 2 – Software architecture

Two-tier architecture

The architectural concept (Figure 2) is based on a 2-tier model. The first tier manages the connection with the virtual machine and the visual search engine running on it, whereas the second tier is in charge to track objects and display correctly aligned 3D models. The first tier is composed by 5 different layers: Unity (a game engine) and UWP (Universal

Windows Platform - a platform to develop applications for Windows 10, Windows Mobile 10, Xbox One and Hololens) for the application running on the Hololens; Java Web service, the Image recognizer software and Linux for the cloud side. The second tier is composed by three layers: Vuforia, a framework for developing AR application, Unity and UWP; all these three layers refer to the application running locally in the HDM.

Data flow exchange

Figure 3 shows the data flow exchange which is performed transparently from the user point of view:

1. the application takes a picture at fixed rate (e.g. one every five seconds) and manages the connection with the virtual machine in the cloud;
2. the picture is sent to the visual search engine in the cloud;
3. if an object depicted in the picture is recognized only the best matching label for the depicted object (e.g. Palatin Towers, Turin) is selected;

Figure 3 – Data flow exchange

4. at this point the Unity application starts providing enriched information related to the detected object taken from wikis. Figure 4 illustrate an example. An Info Panel shows the enriching information. There are also two buttons: Menu button and Exit button. Selecting the Menu button allows to browse four further buttons: Details (it shows the details of the object), Media Content (it shows a video about the object), AR Content (it shows the virtual reconstruction of the object) and Exit (it exits the menu and search for another object);
5. when the user selects the AR content button the Unity application instantiates a "silhouette": it is a semi-transparent image of the recognized object that helps the user to align and to see the monument from the right perspective. This facilitate the Vuforia's tracking. After Vuforia recognizes the framed object it instantiates a 3D model of the recognized object perfectly aligned with the real one (Figure 5).
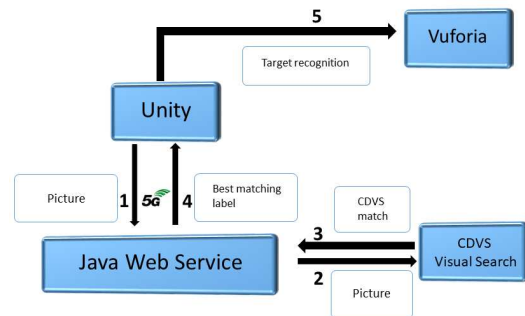

Figure 4 – Info Panel after monument recognition

Figure 5 - 3D model overlay of the monument

Question	1	2	3	4	5
Are commands simple and intuitive?	0	0	0	75%	25%
Instructions have been clearly exposed?	0	0	0	37,5%	62,5%
Can you easily come back to the menu?	0	0	0	87,5%	12,5%
Are text and symbology intuitive and clear?	0	0	25%	37,5%	37,5%
Is the passage through contents fluid?	0	0	12,5%	62,5%	25%
Is starting recognition fast?	0	0	37,5%	12,5%	50%
Is 3D model overlay with Vuforia's tracking well executed?	0	12,5%	25%	25%	37,5%
Is the application well realized?	0	0	0	50%	50%
Is text content clear and visible?	0	0	12,5%	37,5%	50%
Are video clips fluid and useful?	0	0	12,5%	50%	37,5%
Do you think that read textual information is useful?	0	0	12,5%	37,5%	50%
Do you think that see video clips is useful?	0	0	12,5%	50%	37,5%
Do you think this application could helps cultural heritage?	0	12,5%	0	25%	62,5%

Legend: **1**-Strongly Disagree **2**-Partially Disagree **3**-Neutral **4**-Partially Agree **5**-Strongly Agree

Figure 6 – Evaluation questionnaire

Application use case

This project aims to enhance tourists' experience during a city tour. Therefore, in order to evaluate a real usage of this application, some tests have been organized in Turin, Italy. The location chosen for the tests is the Archaeological Park in Turin, and the target monument is the Palatine Towers. This place is impressive, spacious and very popular for tourists. Palatine Towers was the Roman Age city gate, which provided access to the ancient city Julia Augusta Taurinorum.

Tests were focused on various type of parameters, and they have been designed in order to gather both subjective and objective feedbacks. In particular, the usability and the simplicity of usage have been the most interesting target of this tests. A questionnaire has been submitted after the test phase and users could rate a set of claims from 1 (totally disagree) to 5 (totally agree).

Testers have been chosen randomly in order to get a representative set of samples. The group of testers was equally composed by women and men, the age ranges being between 20 and 50, and testers had different knowledge of AR technologies. About one third of participants never tried an augmented reality application (Figure 6) and some of them have not any familiarity with computers. Testers were all volunteers and their participation was not remunerated.

The users had to simulate a normal walking through the Archaeological Park, until the Palatine towers. When they arrived at the monument the application should recognize it and then users could navigate all information as already shown in Figure 4.

After a starting tutorial about the application, the users were left alone and they had to accomplish the following actions: reading initial information about Palatin Towers; tapping on the menu button; tapping on the Details button and reading further information; tapping on the Media Content button to start a video about the monument; tapping on the Augmented Reality button to visualize a 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: ease to point, visibility, fluency, stability and speed of content switching (Figure 6).

Before starting the tests, users were introduced to the pointing system of the Hololens and to the application interface. Furthermore, they were informed about the possibility to move

freely in the park by using only their hands to interact with the system. At the end of the tests, the users filled a questionnaire, thus providing their comments about the application.

User trials analysis and feedback

Before the results analysis, some technical and software limits have to be introduced. One of the most important issue is the limited field of view of Hololens, which is only 30x17 degrees; therefore, the users cannot see large objects when they are too close the framed items. Moreover, the lenses brightness is limited, this makes the contents not really visible when the application is used in very sunny days. Another issue is the instable tracking of the virtual reconstruction of the monument (see Figure 5). Nowadays, it does not exist a free and stable building recognizer to perform augmentation; furthermore, Vuforia is the only SDK supported by Microsoft to develop applications running on the Hololens. Vuforia recognizes an object by comparing the image framed by the camera with a set of pictures pre-loaded in a database; unfortunately, every light change or object movement with respect the images loaded in the database can compromise both the recognition and the tracking.

The users were requested to fill a questionnaire consisting of 16 questions (Figure 6), regarding technician aspects of the applications as well as usability and interface appealing. All participants thought that the commands are simple and intuitive (question 1). Most users found menus and buttons intuitive or very intuitive (question 4). The majority of the testers judged the application fluent and the initial recognition is stable and fast (questions 5 to 7). Some drawbacks relate to the Vuforia's tracking stability (question 7). This depends on the instability of Vuforia in recognizing images. However, a noticeable feedback received from users is that most of them appreciated the application, and they think AR/MR could really help and support the development of the national cultural heritage.

Conclusions

The application is designed to offer an immersive mixed reality experience capable of engaging the final user during a city sightseeing.

Moreover, it shows the potential of new technologies (immersive devices, 5G networks) combined with cultural heritage oriented contest.

The application enhances real monuments with textual, multimedia information and 3D models perfectly merged with the physical one, giving the opportunity to the user to improve their knowledge in a seamless way, increasing their entertainment.

Subjective test evaluations showed that users appreciate the application both from the usability point of view and from the information content point of view. The access to the broadcaster's archive is an added value.

On the other hand, there are some hardware limitations that can affect the user experience (e.g. limited field of view and low visibility in some lighting conditions), and another problem is that rapid changes in outdoor lighting conditions produces tracking and recognition instability. One possible solution could be the use of algorithms based on deep learning detectors. As a future work we would like to add a navigation mode that gives the possibility to the user to select a specific target address or a desired target location of a certain type such as a museum or a church. The system then will compute the shortest path and will drive the user with the help of appropriate graphics.

References

1. Sanna A and Manuri F 2016 A survey on applications of augmented reality Advances in Computer Science: an International Journal, 5 18-27.

2. Yovcheva, Z.; Buhalis, D. and Gatzidis, C.: Engineering augmented tourism experiences. In Information and communication technologies in tourism. Springer Berlin Heidelberg, 2013, http://dx.doi.org/10.1007/978-3-642-36309-2_3

3. Huey L C, Sebastian P and Drieberg M 2011 Augmented reality based indoor positioning navigation tool In IEEE Conf. on Open Systems (ICOS) 256-260.

4. Azuma R T 1999 The challenge of making augmented reality work outdoors Mixed reality: Merging real and virtual worlds 379-390.

5. Fritz, F., Susperregui, A., and Linaza, M.T. (2005). Enhancing Cultural Tourism experiences with Augmented Reality Technologies. The Eurographics Association.

6. Reitmayr, G. and Schmalstieg, D.: Collaborative augmented reality for outdoor navigation and information browsing Proceedings of the Second Symposium on Location Based Services and TeleCartography. 2004

7. Dahne P and Karigiannis J N 2002 Archeoguide: System architecture of a mobile outdoor augmented reality system Proc. of the Int. Symposium on Mixed and Augmented Reality - ISMAR 263-264

8. Stricker D 2001 Tracking with reference images: a realtime and markerless tracking solution for out-door augmented reality applications In Proc. of the conference on Virtual reality, archeology, and cultural heritage - VAST 77-82.

9. Gleue T and Dahne P 2001 Design and implementation of a mobile device for outdoor augmented reality in the archeoguide project. In Proc. of the conference on Virtual reality, archeology, and cultural heritage - VAST 161-168.

10. Hollerer T and Feiner S 2004 Mobile augmented reality Telegeoinformatics: Location-Based Computing and Services Taylor and Francis Books Ltd. London, UK, 21.

11. Yovcheva Z, Buhalis D, and Gatzidis C 2012 Smartphone augmented reality applications for tourism E-review of tourism research (ertr) 10 63-66.

12. Vlahakis V, Pliakas T, Demiris A M and Ioannidis N 2003 Design and Application of an Augmented Reality System for continuous, context-sensitive guided tours of indoor and outdoor cultural sites and museums In Proc. of the conference on Virtual reality, archeology, and cultural heritage - VAST 155-164.

13. El Choubassi M, Nestares O, Wu Y, Kozintsev I, and Haussecker H W 2010 An Augmented Reality Tourist Guide on Your Mobile Devices In Proc. of the Int. Conf. on Multiscale Materials Modeling - MMM 588-602.

14. Han D I, Jung T and Gibson A 2013 Dublin AR: implementing augmented reality in tourism. In Information and communication technologies in tourism 511-523

15. Olsson T, Lagerstam E, Karkkainen T, and Vaananen-Vainio-Mattila, K 2013 Expected user experience of mobile augmented reality services: a user study in the context of shopping centres Personal and ubiquitous computing 17, 287-304.