HoloCities: A Shared Reality application for Collaborative Tourism

Roberto Iacoviello and Davide Zappia
RAI Radiotelevisione Italiana – Research and Technology Innovation Centre
E-mail: roberto.iacoviello@rai.it
E-mail: davide.zappia@rai.it

Abstract. Communication is a key arena for shared-reality application: our mixed reality application is developed on Microsoft Hololens and it has been designed to provide new engaging ways to discover the city using augmented reality. Most augmented reality tourism applications isolate the user, therefore, this application has been made multiplayer and collaborative to encourage shared experiences and socialization. In this project two scenarios are described: In the first scenario there is a real guide that exposes to the group, each equipped with Head Mounted Display, the peculiarities of the monument visited. Each user share the 3D models manipulated by the guide that can be in the same room or it can be a remote guide leveraging on the 5G network low latency. Moreover, the guide could highlight and label to convey relevant information about the objects. While in the second scenario there is no real guide, so the application automatically recognizes the framed monument thanks to a visual search engine running on a 5G-ready infrastructure and shows shared information to the whole group of tourists. Moreover, it uses the huge amount of multimedia data stored in the archives of the Italian Public Broadcaster RAI. Thanks to the fact that the new HMDs are able to track the position and orientation of the user, we have created a menu capable of following the user while exploring the city. Moreover, the shared-interaction with virtual objects is performed by hands or vocal commands. Both scenarios allow users to visit a city in a new and intelligent way, promoting participation and collaboration with others and with the guide, unlike previous approaches that involve only one tourist using the information provided by the app.

1. Introduction
In the Internet of Things Era, people are increasingly surrounded by technology and devices that interact with each other in a huge network which is bigger and smarter day by day. In this environment Augmented Reality is one of the topics in which large companies are investing specially for industrial maintenance and training fields.

Few years ago, people could see the media contents in a passive way and talk about them only with friends and neighbours. Nowadays people can interact with the entertainment contents, lead the storytelling and share it worldwide. This is possible only with devices that are more powerful and wearable every day, and with the internet network that connects them. 5G is the latest technology to bring the internet connection to the next level: with 5G people can share and see much more contents in real time with incredible interactions.

The aim of this work is to propose an augmented support to a tourist guide, that can help a group of users on a city tour and support the real guide in his explanation. This project was created to increase people’s interest in Italian monuments and, at the same time, to create a
bridge with the multimedia contents of our archive: RAI "Teche". There has been much recent activity on Virtual Reality (VR), Augmented Reality (AR) and Mixed Reality (MR), and this has involved all relevant sectors, including device manufacturers, games companies, audio-visual content suppliers and service providers, and has covered application areas both inside and outside the media and entertainment industries.

Important players have introduced in the market many Head Mounted Displays (HDMs): Magic Leap One, Microsoft Hololens2, Meta2; Facebook has also mentioned several times that it’s building AR glasses, partnering with Ray-Ban maker Luxottica [3]. Moreover, many internationals bodies are working on the topic. Virtual Reality Industry Forum (VRIF) is working end to end chain guidelines and lexicon of terms, with the purpose to spread the use of common formats for interoperability [7]. MPEG is working on MPEG-I (Immersive) targeting point cloud, 3D graphics and light field. DVB (Digital Video Broadcasting) has recently completed a study mission report to determine whether VR video is likely to be commercially successful [10]. This article focuses on one particular application sector, public service broadcast and on one particular technology that we call Shared Reality.

The goal of public service broadcasting is to provide a balanced service of entertainment, information, and education. This means providing a mix of programs that are unique, immersive and challenging, following the technology that leads to new opportunities. One of the challenge is to give a second life to the huge amount of multimedia data stored in the broadcaster’s archives. Mixed Reality extends Augmented Reality: physical and artificial objects are blended together in a hybrid space, with some kind of interaction for the user. We combined Mixed Reality and multimedia data with a shared view of the holograms between the participants and the tourist guide in order to give to the user a collaborative experience, following the idea that technology can bring people together, and it allows them to stay connected to each other, which can significantly improve their experience.

In this work, carried under the H2020 European project 5GCity, we used visual tracking algorithm and cloud computing to enrich cultural tourism thanks to the bridge with the RAI archive and to the shared reality experience between the participants and the tourist guide.

2. Background

In recent years experts have been talking about Extended Reality (XR), meaning an umbrella term that includes Virtual Reality, Augmented Reality and Mixed Reality. They share some common goals (although these technologies are different from each other): to improve the user experience by introducing virtual assets and technological requirements (e.g. 3D models). The goal is to use all these technologies in a seamless way to provide users with the best possible experience, according to each user’s hardware, location, network capacity and network latency. Potentials of augmented reality has already been amply demonstrated, showing how it can simplify people’s lives in industrial context and education, or help people in cultural tours [9], because it changes way tourists observe their surroundings. Moreover, it has been noted how the mixed reality increases even more realism because holograms can interact with the external environment, so the user can see a perfect fit of holograms in real life.

In the literature there are many examples of Augmented Reality applications for cultural heritage, but many of these are developed with Mobile Augmented Reality (MAR) [1], and most are created for indoor cultural tours [11].

The problems of this development are that the user has to see the AR contents through a smartphone, therefore the immersion and the realism are very poor; while the development of indoor AR applications is easier because it is more immediate to find the user’s position and recognize the objects in known environment.

Examples of outdoor application are The Archeoguide project [1], which uses a laptop computer to manage computer generated assets and track the scene and it allows to correctly align and
overlap synthetic contents to the real world. The Archeoguide project has been also extended in order to provide users with a powerful mobile device for outdoor AR applications with off-the-shelf components [4]. Authors in [6] used a visual search engine to automatically recognise outdoor monuments with the HMD camera so that the user can freely look at the object for which the augmented contents are available and interact with the virtual assets. The drawback of this project is that the user feels a sense of detachment from friends and the group of people around him. None of them simultaneously achieve the following design goals: immersion thanks to the use of a standalone HMD, computer vision algorithms for the monument recognition to avoid GPS related issues (e.g. GPS denied environment), the cloud computation to save battery life and the sharing of holograms and experiences among participants. In the near future, the most significant advances will be on educational tourism and the real time sharing of the exploratory experience will open doors to a new world for visitors to learn by improving the interaction between Man, Nature and Art.

3. Proposed Solution

3.1. Technologies

The device used for this work is the Microsoft Hololens, because it is a standalone device that allows the user to walk freely in the city, without cable or connection to a laptop. With its capabilities (spatial mapping and gesture recognition) the user can feel surrounded by holograms that interact with the environment. The application is developed with Unity Game Engine and the MixedRealityToolkit SDK are the libraries used for the mixed reality implementation, while Azure Spatial Anchors are used for the multidevice spatial synchronization.

In Figure 1a is depicted the Hardware Architecture for the first scenario: Hololens devices are connected to a Java Web Service through a 5G module. For the second scenario, Figure 1b, on a remote server on cloud, a MPEG (Moving Picture Experts Group) standard CDVS (Compact Descriptor Visual Search) was used to compare the picture taken by the Hololens with the monuments database. In the first scenario the hardware architecture is simpler because the real guide does not need the automatic recognition of the monument.

Figure 1. Hardware Architecture: (a) First scenario, (b) Second scenario
3.2. Software Architecture

In first scenario, the application developed in Unity shows the enriched information and the 3D model of the monument chosen by the real guide and creates a world anchor, which is a spatial landmark that allows connected tourists to watch exactly what the guide is showing them. In this way, leveraging on the 5G low latency capabilities, the virtual assets are synchronized both in time and space.

In the second scenario, Figure 2, the application developed in Unity takes picture from the Hololens front camera at fixed rate (e.g. every 2 seconds) and manages the connection with the remote server in the cloud through a Java Web Service (step 1). After that the picture is sent (step 2) to the visual search engine in the cloud. The MPEG CDVS performs a feature extraction, generates an image descriptor and compare it with those previously inserted in the database (3). If an object depicted in the picture is recognized the best matching label is selected and sent back to the Hololens (step 4). At this point the Unity application shows enriched information and the 3D hologram related to the detected object. There are many level of interactions available with user gestures and voice commands: Details Button shows enriched information about the object; Media Content Button shows videos taken from RAI archive; Manipulate Button offers to the user the possibility to manipulate the holograms (zoom, translation, rotation); Reset Button is a refresh of the viewing parameters.

3.3. Graphic user interface and interaction model

In the first scenario (with the real guide), a study has also been done on the user interface, in order to make the commands and interactions as intuitive and productive as possible.
Augmented Reality HMDs are becoming very common and applications are becoming more and more useful and aimed at the business world. For this reason, the type of interaction between man and virtual objects is becoming a very important topic. In Gutwein [5] there are different types of augmented reality menus, but they can be divided into three main categories:

- **2D menus in 3D environments**
- **Enhanced 2D menus**
- **Augmented reality specific menus**

It has been noticed that the 2D menus are immediately usable by users because they already know that style of interaction, but in extended usage these menus are less intuitive than the others just because they are not optimized. The enhanced 2D menus want to solve these problems, being already familiar to users, but with an interaction optimized for AR.

Finally, in Davis [2] and Mitchell [8], menus are designed specifically for augmented reality. For this reason, these menus are often very intuitive, but with interaction patterns unknown to users, so they need a small training.

In the first version of the application we used a Hololens default style menu, which is an Enhanced 2D menu. It is a 3D representation of a conventional 2D menu; the interaction has been adapted for specific use in augmented reality. In this way the user already knows the interaction model and feels more skilled using the menu from the start of the application.

We’ve encountered several drawbacks, e.g. occlusions, and in the second version we used a different approach. The goal was to create a menu that would always follow the user in all positions, but never unintentionally occupy his field of view. This is because the user can move and manipulate the 3D models so that the menu may block the view in certain positions (Figure 3).

Figure 3. New user interface.
3.4. Application Use Case
At the start of the application, the user can choose between two modes, first mode is with a human tourist guide, second is without the guide.

3.4.1. First Scenario
In this scenario (Figure 1a) the real guide is the master of our multiplayer session. The guide provides users with some Hololens devices and then starts the application in guide mode. The guide hosts the session and the users join his session. With this mode the CDVS for visual search is disabled, because the guide knows the surrounding environment, so he only chooses the monument he wants to explain. The guide can manipulate the holograms and show users the most important features of the monument, this is possible thanks to the connection between the Hololens devices. Each user can move around the 3D model and see every interaction of the guide in real time. The guide can also annotate the environment by placing virtual lines or text on monument and buildings in his surrounding, enhancing the storytelling and the engagement of the tourist group.

3.4.2. Second Scenario
This scenario (Figure 1b) was created for a group of users who are visiting a city without a tourist guide. They have little knowledge of the environment but have a strong interest in navigating through their surroundings to visit different places in the city. The first user who starts the application and select "no-guide mode" becomes the host of the multiplayer session in order to share the tour with others. Then the host’s Hololens use 5G network connection to send the picture to the remote cloud for the monument recognition with the CDVS visual search. Once a monument is recognized, each user can see the holograms and the information about that monument. In this way the group of users can stay close to the holograms, interact with them thanks to the gesture recognition and discuss together, encouraging cultural tours and socialization.

4. Test and Results
This project aims to overcome two different challenges: helping a tour guide to make his work more intuitive and interesting for clients, but also helping groups of tourists to easily discover the artistic heritage of a city. Therefore, in order to evaluate the real contribution of the application, we organized objective and subjective tests. The tests were created in order to have subjective and objective evaluations regarding the user interface and the ease of use of the application. After the test a questionnaire was submitted to the testers, the evaluations for each question ranged from 1(strongly disagree) to 5(strongly agree). The testers were chosen at random, distributed in men and women aged 20 to 50. They had a different knowledge of augmented reality, and some of them had never tried such application.

We tested both scenarios but with a particular attention to the one with the real guide. In the test one of our collaborators had the role of a tour guide, while the testers listened to a description of the Guinigi Tower of Lucca (one of the city of the EU project 5GCity). The testers could move around in the environment, watch all the tourist guide’s annotations and interact with the holograms. The questionnaire was structured in two parts in order to assess the application from different points of view. The first part of the questionnaire was related to the ease of use and intuitiveness of the application: most users think that the application is very intuitive and easy to use. The second part is related to the user interface: the users think that the interaction is always intuitive, comprehensible and easy to use. Finally, it is very relevant to point out that almost all testers believe that the collaboration within the Shared Reality adds another dimension to the possibilities of such systems by supporting the tourist guide and enhancing the cultural tour of the city.
5. Conclusion
Our mixed reality application is designed to offer an engaging and collaborative Share Reality experience capable of providing information in outdoor while visiting a city. The application automatically enhances historical monuments with multimedia information and 3D models, exploiting all the multimedia data stored in the archives of the Italian Public Broadcaster RAI (Figure 4). Moreover, multiple users can collaborate, or a tourist guide might share the user’s reality and the holograms. The 3D manipulation enables participants to zoom into the details, annotating content and capturing display views. The collaboration feature can be used to create unique AR experiences. For example, a group of tourist can share 3D virtual objects and interact with them, or a tourist guide can annotate the monument in real time. The overall goal is to augment the face-to-face collaborative experience leveraging on the emotional immersion given by the sharing of the reality among users. In our future work we would like to add the option to send all the contents to the user’s mobile phone in the scenario where the HDM is provided from the city’s tourist office. Moreover, we would add a remote guide mode, and give to the user the emotional involvement of a human guide into the scene, thanks to a tourist guide avatar and his real voice.
References

[1] P. Dahne and J. Karigiannis. “Archeoguide: system architecture of a mobile outdoor augmented reality system”. In: Feb. 2002, pp. 263–264. ISBN: 0-7695-1781-1. DOI: 10.1109/ISMAR.2002.1115103.

[2] Matthew M. Davis et al. “Depth-based 3D gesture multi-level radial menu for virtual object manipulation”. In: Proceedings - IEEE Virtual Reality 2016-July (2016), pp. 169–170. DOI: 10.1109/VR.2016.7504707.

[3] “Facebook says it will build AR glasses and map the world”. In: http://bit.ly/2HiDDC1 (2019).

[4] Tim Gleue and Patrick Dähne. “Design and Implementation of a Mobile Device for Outdoor Augmented Reality in the Archeoguide Project”. In: VAST ’01 (2001), pp. 161–168. DOI: 10.1145/584993.585018. URL: https://doi.org/10.1145/584993.585018.

[5] Sonja Gutwein. “Computer support for collaborative creativity”. In: Beyond the Desktop (2013), pp. 15–22.

[6] Zappia D. Iacoviello R. et al. “Enhancing cultural tourism by a mixed reality application for outdoor navigation and information browsing using immersive devices”. In: IOP Conference Series: Materials Science and Engineering 364 (2018), p. 012048. DOI: 10.1088/1757-899x/364/1/012048. URL: https://doi.org/10.1088%2F1757-899x%2F364%2F1%2F012048.

[7] “Lexicon-VRIF”. In: https://www.vr-if.org/wp-content/uploads/Lexicon-VRIF.pdf (2017).

[8] Peter Mitchell and Brett Wilkinson. “Periphery triggered menus for head mounted menu interface interactions”. In: (2017), pp. 30–33. DOI: 10.1145/3010915.3010964.

[9] G. Reitmayr and D. Schmalstieg. “Collaborative Augmented Reality for Outdoor Navigation and Information Browsing”. In: Proceedings of the 2nd Symposium on Location Based Services and TeleCartography., 2003, pp. 53–62.

[10] Christian Timmerer. “Immersive media delivery: Overview of ongoing standardization activities”. In: IEEE Communications Standards Magazine 1.4 (2017), pp. 71–74. ISSN: 24712825. DOI: 10.1109/MCOMSTD.2017.1700038.

[11] Z. Yovcheva, D. Buhalis, and C. Gatzidis. “Overview of Smartphone Augmented Reality Applications for Tourism”. In: eRTR 10 (Jan. 2012).