

Using Mixed Reality and Natural Interaction in Cultural Heritage Applications

Raffaello Brondi, Marcello Carrozzino, Cristian Lorenzini and Franco Tecchia
 Laboratorio Percro of Scuola Superiore Sant'Anna
 Via Alamanni 13, San Giuliano Terme (PI) - Italy
 E-mail: r.brondi, m.carrozzino, c.lorenzini, f.tecchia@sssup.it

Keywords: natural interaction, mixed reality, virtual reality, cultural heritage, museum application

Received: June 27, 2016

In this paper, we present a general architecture for Mixed Reality applications. The proposed solution has been developed in order to provide a useful instrument to develop Cultural Heritage applications. During the design of the system, particular attention was given to intangible knowledge, such as manual activities, performing arts, lost civilizations habits, representing a particular heritage poorly addressed by previous studies. The system aims at providing an easy and engaging infrastructure to develop immersive application to be used for communication / dissemination and education purposes. The proposed architecture exploits Natural User Interfaces solutions as interaction metaphor between the Virtual Environment and the user. Natural interaction in fact provides high sense of presence and immersion to the user, improving the user engagement and fostering the learning process.

The paper presents also two case studies, where two different applications aimed at teaching and disseminating crafts knowledge, in particular printmaking and weaving, have been developed on top of the presented architecture.

Povzetek: Predstavljena je mešana resničnost za namene elektronske kulturne dediščine.

1 Introduction

Virtual Reality (VR) applications create interactive environments in which the observer feels totally immersed: users can move or interact in a completely synthetic world [1]. Differently, in Augmented Reality (AR) applications the digital content is integrated into the real environment [2]. VR and AR technologies are becoming extremely popular and are used nowadays to implement many kind of applications in several different fields: military, medicine, education, visualization, entertainment, etc.... These two technologies represent two different expressions of a common family of technologies and applications falling under the definition of Mixed Reality (MR). Milgram and Kishino [3] theorized the concept of “virtuality continuum” in order to create a classification able to describe this concept. They placed on one extreme of the continuum the real world and on the other side a completely virtual world. The space of virtuality continuum lying between the two extremes, all the technologies and applications, represents different flavors of Mixed Reality (MR), including AR (real environments augmented with virtual contents) and Augmented Virtuality (virtual environments augmented with real contents). Several researches have focused on Mixed — rather Virtual or Augmented — Environments, how their adoption can affect several aspects of the user experience and how at the same time they are affected by the new arising technologies.

VR and MR applications in the context of Cultural Heritage (CH) are nowadays gaining an increasing consent, for a variety of purposes including digital conservation (reconstructing artwork damaged or destroyed) [4][5], for validation of scientific hypotheses in archeological reconstructions [6] and for education [7]. At the same time the recent spreading of depth sensors together with sensorised controllers is nowadays shaping the way we interact with the Virtual Environments (VE). Natural User Interfaces¹ (NUIs) are becoming more and more popular, and new richer interaction metaphors can be designed in order to improve the engagement and sense of presence of the users providing a completely new experience [8]. NUIs enabling visitors to be physically and emotionally involved during a virtual experience are becoming popular also in the Cultural Heritage context [9].

In this paper we present a general architecture for Mixed Reality systems that can be used to provide an immersive experience to Cultural Heritage visitors. The

¹ Natural User Interface is a term used to identify human-computer interactions based on typical inter-human communication. These interfaces allow computers to understand the innate human means of interaction (e.g. voice and gestures) and do not require humans to “learn” the language of computers (e.g., keyboard and mouse).

proposed solution can be used both for dissemination purposes, as it enhances the engagement of the user, and for training/teaching activities. In particular such systems results extremely effective when trying to transmit intangible knowledge, as like as craftsmanship, since it allows the visitors to physically emulate the proposed actions. Moreover the system can be used both with static prerecorded material and with real-time capture of the real context.

2 State of the art

As above mentioned, differently from VR, MR mixes “synthetic” and “real” information making them coexisting in the same environment. Most of the applications developed in the context of CH are based on Augmented Reality. Among the first AR applications developed in the CH field is ARCHEOGUIDE [10]. Using a Head Mounted Display (HMD) visitors of archaeological sites can see virtual reconstructions of the temples and other monuments directly superimposed on the real ruins. Nowadays, thanks to the ubiquitous networking availability and the technological progresses in mobile computing, the ARCHEOGUIDE concept has been further developed. New research is focusing on mobile devices as a gateway to provide augmented cultural content everywhere [11].

Other AR applications developed in the same field aim at providing new ways of interaction between visitors and artworks inside museums. This “augmentation” of the real-world environment can lead to an intuitive access to museum information and enhances the impact of the museum exhibition on virtual visitors [12]. Wojciechowski et al. [13] developed an AR-system composed by an authoring tool and an AR-browser. Using the former instrument, museum superintendents can design Virtual and Augmented Reality exhibitions. Through the AR-browser, installed for example in a kiosk, visitors can see the representations of cultural objects overlaid on the video captured by a camera. Similarly Chen et al. [14] proposed a new AR guidance system for museums based on markers. ARCube [15] exploits a 3D marker to enable 360° interaction with fully reconstructed three-dimensional archaeological artefacts in real-world contexts. Debenham et al. [16] developed an AR-system used inside the Natural History Museum in London which provides visitors with augmented contents through hand-held displays in order to enable an exciting new way to present the evolutionary history of our planet. In [17] Augmented Reality has been used instead to improve the work of the restorers and promote communication and cooperation between them.

The great success of AR in the Cultural Heritage context is mainly related to the fact that they provide an easy, engaging and friendly way to access information related to a particular asset, commonly by keeping the cultural asset in the foreground and enriching its images with digital content. When dealing with intangible assets, such as performing arts, manual activities, lost civilizations habits, there is not a real concrete object to

augment. This kind of evanescent knowledge requires a deeper usage of virtual components [18] because the real part is not physically present or is not always available.

All MR solutions usually needs to merge 3D (live or recorded) information coming from the real environment with a 3D synthetic environment as smoothly as possible. By using immersive displays like HMDs, the user experience can be further enhanced. Tecchia et al. in their work [19] proposed a HMD visualization system including the real time stream of 3D images of user's hands recorded with a depth-camera. The system makes use of two colored markers placed on top of the user hands to enable a basic interaction with the VE. The depth sensor put on top of the HMD was in charge of recording the peri-personal space and in particular the user hands. The acquired stream was used to recreate a representation of the user inside the VE. Moreover, using the combination of RGB and depth information coming from the camera, the system recognizes the fingers movements in the environment allowing the user interaction with the environment (e.g. virtual object pinching).

The metaphors of interaction enabled inside MR applications, from completely Virtual to Augmented ones, represent another extremely important design aspect. Specific solutions impact differently on the sense of Presence, Immersion and Engagement of the user. In the context of CH applications, improving these factors can enhance the impact of a dissemination application. It becomes even more important when the aim of the application is learning or training. Safeguarding and passing over skills and intangible cultural heritage features is the subject of several experiments, as well as large research projects [20][21]. Carrozzino et al. [9] argued that Immersive VEs combined with natural interaction would provide a powerful solution to develop a system to transfer practical skills.

In the last years several researchers and technological industry leaders have focused on the development of different solutions enabling a smooth and simple natural interaction of the user inside the VE. Most of the efforts so far are focusing on hand tracking solutions [22][23][24]. The Leap Motion Controller represents one of the latest technological products created to enable user natural interaction inside the VE based on hand tracking/gestures. It is gaining a lot of popularity due to the ease of use and tracking performance achieved with the latest updates. This device can be easily integrated in any VR/MR application in order to allow users to see and interact with the VE with their own hands. Coupling the Leap Motion controller with an HMD (e.g. the Oculus Rift or HTC Vive) provides developers with an extremely powerful and relatively cheap VR/MR solution that can be used in many Cultural Heritage contexts to provide extremely engaging experiences to the users.

Given these premises, the presented architecture aims at providing an easy way to develop MR applications, exploiting the capabilities of HMDs and immersive displays in general, coupled with devices, like the Leap Motion, able to track user hands in order to

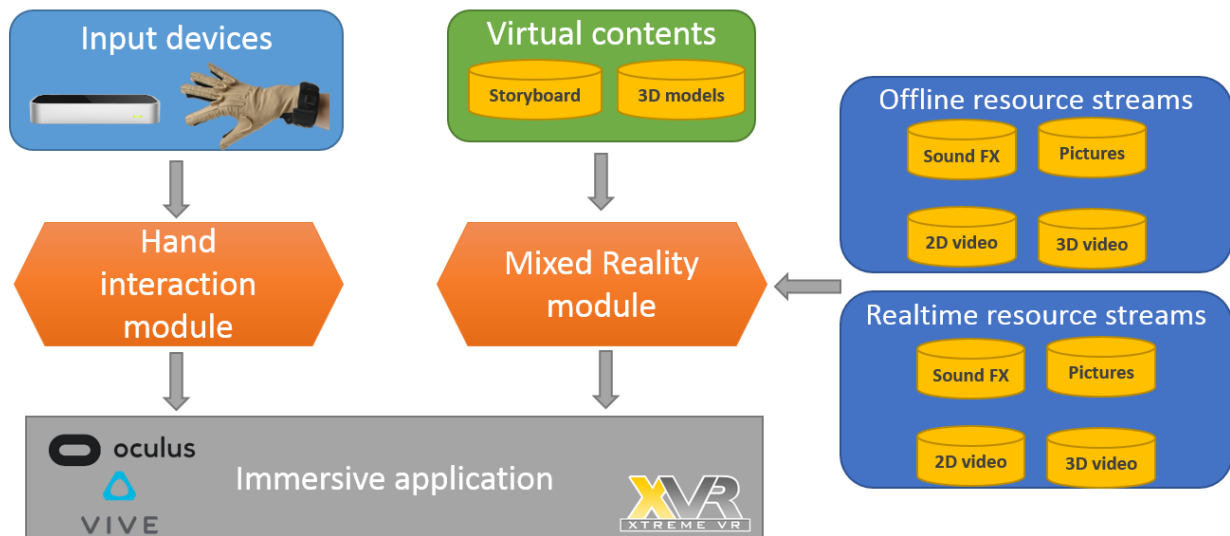


Figure 1: Architecture overview.

create engaging interactive applications in the context of CH.

3 The architecture

The proposed system has been designed and realized on top of the XVR technology [25], an in-house made VR-oriented framework offering a graphics engine for the real-time visualization of complex three-dimensional models and the support to a wide range of VR devices (such as trackers, motion capture devices, stereo projection systems and HMDs). XVR applications are developed using a dedicated scripting language whose constructs and commands are targeted to VR, including the support to 3D animation, 3D sound effects, audio and video streaming and advanced user interaction. This choice has allowed a good flexibility in terms of support of hardware devices and ease of developing dedicated software add-ons able to expand the capabilities of the framework. Figure 1 gives an overview of the developed system which can be divided in three main parts. A low level infrastructure based on XVR and responsible for the direct interaction with the visualization system. This component is in charge of managing not only the stereoscopic rendering on the immersive displays but also the VE update according to the tracking technologies of the visualization system used. This system element allows applications developed on the proposed system to run on the latest HMDs, the Oculus

Rift and HTC Vive, and also on projection based systems like CAVEs displays.

On top of this low level component the system is composed by a Mixed Reality module which is in charge of merging the real content coming from various streams (audio, images or video either 2D or 3D) with the virtual contents. The resources coming from the real world can be either pre-recorded or real-time captures of the world. The system takes care of handling the different streams, registering them inside the VE and rendering them to the user.

Dedicated tools have been developed in order to handle 3D video (RGBD streams) acquired with depth cameras like the Microsoft Kinect, since existing tools lack of the features needed to opportunely post-process this kind of data. In particular the suite of developed tools allow to trim the stream (in order to select specific portions of the stored data) and to clean the video data in order to make it easier to seamlessly mix it with the virtual environment (see Figure 2). The use of these tools allow to separate, with a good precision, the desired content of the stream from the unwanted background/noise.

The virtual content consists of 3D models and Virtual Storyboards (VS). The system interprets the VS, a sequence of instructions defining the application storyboard, in real time. The VS, defined in text files in order to be easily authored by any text tool, provides the possibility of defining custom key-points that can alter

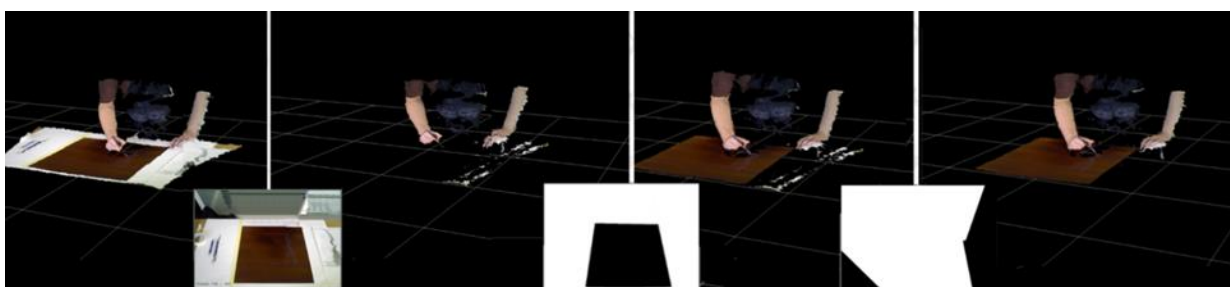


Figure 2: Cleaning 3D video streams acquired with depth cameras using the developed tools.

the flow of the application.

The VS allows also to specify how the real world resources and other elements (e.g. camera animations, interactive elements, movements and dialogues) are displayed in the VE and how the user can interact with them. All the resources and the relationship between the resources, the action of the users, the timing and the environment are defined in these configuration files. This allows to easily replicate the same functionalities in different contexts, loading custom resources and developing different applications.

3.1 Interaction in the VE

The developed architecture contains a hand interaction module dedicated to the management of user interaction with the VE. The module is in charge of tracking the user hands and detecting gestures. Using the tracking information, a virtual representation of the user hands is provided in the VE. The system animates the virtual hands interpreting the information coming from the sensors used to capture the user. For each hand, first it uses the position of the palm in order to evaluate the position of the user virtual representation. Then for each finger, it evaluates the angles to be applied to each phalange.

The detected gestures are used in order to enable actions to undertake in the Virtual Environment. Currently pointing, tapping and pinching gestures have been implemented in the system. These actions can be used to develop the user interaction with the environment (e.g. selection/pressing of GUI elements like a button, object selection and or transformation). The module offers an abstraction to the above infrastructure allowing the use of different input systems. The architecture have been tested with the Leap Motion Controller and with the CyberGlove II. Other hand tracking systems will be included in the future.

3.2 Case Studies

Using our architecture, two different case study applications have been developed, for two different kinds of intangible Cultural Heritage dealing with craftsmanship.

The subject of the first application has been the work of printmakers. The VE replicates a structured print house featuring different locations where artisans can show their job, retracing all the steps involved in the process of making a stamp. An approach similar to "I'm in VR" has been used to stream pre-recorded depth-movies inside the 3D VE [21], in order to reproduce real artisan movements. Depth movies allow to observe human motion with a high level of detail. More complex systems using graphics animations would require the use of very expensive motion tracking systems that can also hinder the artisan work (see Figure 3, right).



Figure 3: The pedagogical agent inside the VE (left) and the artisan pre-recorded depth-movies (right).

Artisans are visualized as pre-recorded depth-movies tessellated in real-time, rendered as polygonal meshes and merged inside the VE. Users can also explore the environment using natural movements or teleporting themselves in different places; they can also interact with the environment by selecting 3D objects and GUI buttons using the provided NUI.



Figure 4: Watching own hands and artisan's hands at the same time.

A virtual character, a pedagogical agent, guides the user through the environment in order to explain the actions of the artisans and to provide information about their work (see Figure 3, left). Visitors, beyond observing the artisans from a classic "third person view" (see Figure 3 right), can observe the manual activities from a "first person point of view" (see Figure 4) as they were seeing his/her own hands. Furthermore, when in first person view, users can try to emulate the movements of the artisan as they see both their own hands, captured by the hand tracking device, and the artisan's ones.

The second application developed with the proposed architecture is related to the work of weavers. Weaving is a repetitive manual job, made on a loom. The developed application recreates an artisan workshop where different

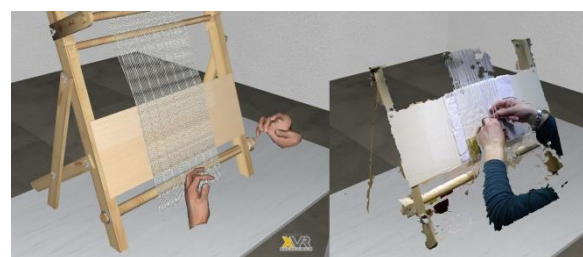


Figure 5: Hands motion recorded with Cyber Gloves vs depth-stream recording reproductions.

weavers are working. Users can freely explore the space around the artisan and, like the previously described application, see their own hands and overlap them to the "ghost" hands of the artisan in order to learn how to perform some of the actions needed during the work of weavers. The artisans' hands can be visualized both as pre-recorded depth-movies and as computer graphics animated "avatars". The 3D avatar animations have been recorded using the Cyber Globe II. Using the data gloves to capture the artisans' hands movements has been possible in this case because weavers commonly wear gloves in their work and therefore this did not hinder the artisan's activities. This allowed also to compare depth videos against avatar animations (see Figure 5) in terms of information delivery and perceived quality.

4 Acknowledgments

The design and implementation of the proposed architecture has been carried out in the context of the AMICA project, funded by Fondazione TIM under the "Beni Invisibili" financing program.

The study of the related work, the setup of the test methodology and the design of future expansions of the described methodology have been carried out in the context of the EU 2020-TWINN-2015 eHERITAGE project (grant number 692103).

5 References

- [1] Riva, Giuseppe. "Virtual reality as communication tool: A sociocognitive analysis." *Presence: Teleoperators and Virtual Environments* 8.4 (1999): 462-468.
- [2] Azuma, Ronald T. "A survey of augmented reality." *Presence: Teleoperators and virtual environments* 6.4 (1997): 355-385.
- [3] Milgram, Paul, and Fumio Kishino. "A taxonomy of mixed reality visual displays." *IEICE TRANSACTIONS on Information and Systems* 77.12 (1994): 1321-1329.
- [4] Carrozzino, Marcello, et al. "The virtual museum of sculpture." *Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts*. ACM, 2008.
- [5] Brondi, Raffaello, and Marcello Carrozzino. "ARTworks: An Augmented Reality Interface as an Aid for Restoration Professionals." *International Conference on Augmented and Virtual Reality*. Springer International Publishing, 2015.
- [6] Barceló, J. A., Forte, M., & Sanders, D. H. (Eds.). (2000). *Virtual reality in archaeology*. Oxford, UK: ArchaeoPress.
- [7] Economou, Maria, and L. Pujol. "Educational tool or expensive toy? Evaluating VR evaluation and its relevance for virtual heritage." *en New Heritage. New media and cultural heritage*, Oxon, Routledge (2006).
- [8] Brondi, R., Alem, L., Avveduto, G., Faita, C., Carrozzino, M., Tecchia, F., & Bergamasco, M. (2015, September). Evaluating the impact of highly immersive technologies and natural interaction on player engagement and flow experience in games. In *International Conference on Entertainment Computing* (pp. 169-181). Springer International Publishing. <http://doi.acm.org/10.1145/161468.16147>
- [9] Carrozzino, M., Lorenzini, C., Duguleana, M., Evangelista, C., Brondi, R., Tecchia, F., & Bergamasco, M. (2016, June). An Immersive VR Experience to Learn the Craft of Printmaking. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics* (pp. 378-389). Springer International Publishing.
- [10] Vlahakis, V., Karigiannis, J., Tsiotos, M., Gounaris, M., Almeida, L., Stricker, D., ... & Ioannidis, N. (2001, November). Archeoguide: first results of an augmented reality, mobile computing system in cultural heritage sites. In *Virtual Reality, Archeology, and Cultural Heritage* (pp. 131-140).
- [11] Brondi, R., Carrozzino, M., Tecchia, F., & Bergamasco, M. (2012, May). Mobile augmented reality for cultural dissemination. In *Proceedings of 1st International Conference on Information Technologies for Performing Arts, Media Access and Entertainment*, Firenze, Italy (pp. 113-117).
- [12] Sylaiou Styliani, Liarakapis Fotis, Kotsakis Kostas, and Patias Petros. Virtual museums, a survey and some issues for consideration. *Journal of cultural Heritage*, 10(4):520–528, 2009, Elsevier.
- [13] Rafal Wojciechowski, Krzysztof Walczak, Martin White, and Wojciech Cellary. Building virtual and augmented reality museum exhibitions. In *Proceedings of the ninth international conference on 3D Web technology*, pages 135–144. ACM, 2004.
- [14] Chia-Yen Chen, Bao Rong Chang, and Po-Sen Huang. Multimedia augmented reality information system for museum guidance. *Personal and ubiquitous computing*, 18 (2):315–322, 2014, Springer.
- [15] B Jiménez Fernández-Palacios, F Nex, A Rizzi, and F Remondino. Arcube – the augmented reality cube for archaeology. *Archaeometry*, 2014, Wiley Online Library
- [16] Paul Debenham, Graham Thomas, and Jonathan Trout. Evolutionary augmented reality at the natural history museum. In *Mixed and Augmented Reality (ISMAR)*, 2011 10th IEEE International Symposium on, pages 249–250. IEEE, 2011.
- [17] Brondi, Raffaello, and Marcello Carrozzino. "Fostering collaboration among restoration professionals using augmented reality." *2014 IEEE 23rd International WETICE Conference*. IEEE, 2014.
- [18] Papagiannakis, G., Ponder, M., Molet, T., Kshirsagar, S., Cordier, F., Magnenat-Thalmann, M., & Thalmann, D. (2002). LIFEPLUS: revival of life in ancient Pompeii, virtual systems and multimedia. In *Proceedings of VSMM 2002* (No. VRLAB-CONF-2007-038).

- [19] Tecchia, Franco, et al. "I'm in VR!: using your own hands in a fully immersive MR system." Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology. ACM, 2014.
- [20] Skills project, CORDIS EU, available at: http://cordis.europa.eu/project/rcn/103956_en.html
- [21] Carrozzino, M., et al. (2015). AMICA-Virtual Reality as a Tool for Learning and Communicating the Craftsmanship of Engraving. Proceedings of 2015 Digital Heritage International Congress, (pp. 187-188)
- [22] Piumsomboon, T., Altimira, D., Kim, H., Clark, A., Lee, G., & Billingham, M. (2014, September). Grasp-Shell vs gesture-speech: A comparison of direct and indirect natural interaction techniques in augmented reality. In Mixed and Augmented Reality (ISMAR), 2014 IEEE International Symposium on (pp. 73-82). IEEE.
- [23] Rautaray, S. S., & Agrawal, A. (2015). Vision based hand gesture recognition for human computer interaction: a survey. Artificial Intelligence Review, 43(1), 1-54.
- [24] Ding, W. and Marchionini, G. 1997. A Study on Video Browsing Strategies. Technical Report. University of Maryland at College Park.
- [25] Tecchia, Franco et al.. "A Flexible Framework for Wide-Spectrum VR Development." Presence 19.4 (2010): 302-312