Comparing Different Storytelling Approaches for Virtual Guides in Digital Immersive Museums

Marcello Carrozzino[1(✉)], Marianna Colombo[2], Franco Tecchia[1], Chiara Evangelista[1], and Massimo Bergamasco[1]

[1] Scuola Superiore Sant'Anna, Perceptual Robotics Laboratory, Institute of Communication, Information and Perception Technologies, Pisa, Italy
{m.carrozzino, f.tecchia, c.evangelista, m.bergamasco}@santannapisa.it
[2] Department of Filologia, Letteratura e Linguistica, Università di Pisa, Pisa, Italy
marianna.colombo89@gmail.com

Abstract. Virtual museums are becoming increasingly popular, especially thanks to the recent spread of low-cost immersive technologies enabling even small cultural realities to enrich their technology-based offer. The growing availability of computational power and high-quality visual elements enables the use of virtual characters in order to add depth to the virtual visit experience, up to nowadays often limited to the exploration of lifeless environments. This paper presents a pilot study aimed at investigating the positive effects that the use of avatars can provide to a virtual cultural experience, proposing a virtual museum with three different alternatives of storytelling, including one featuring virtual humans, and comparing the results in terms of engagement and understanding of the proposed content.

Keywords: Virtual Reality · Immersive technologies · Cultural heritage
Virtual humans · Avatar · Storytelling

1 Introduction

The meaning of cultural heritage is constantly evolving. Because of the increasing request of a more engaging experience while learning, the traditional way to communicate culture has been progressively enhanced with the use of new technologies.

The most popular technology used for reenacting past cultures and link them with their original environment is Virtual Reality. Creating 3D virtual worlds for cultural heritage applications is being credited as the most affordable, dynamic and interactive option for integrating environment, objects and knowledge associated with a culture [1]. Bringing together culture and new technologies can improve the learning experience of the non-experts, such as children, students and enthusiasts.

However, 3D virtual reconstructions of sites are often only composed of architectural components, objects and other unanimated elements, thus providing a limited experience to the users. Consequently, the current trend is to populate these virtual

L. T. De Paolis and P. Bourdot (Eds.): AVR 2018, LNCS 10851, pp. 292–302, 2018.
https://doi.org/10.1007/978-3-319-95282-6_22

worlds with *avatars* or virtual humans, that have an appearance and behavior very similar (hopefully identical) to the original historical inhabitants of the reconstructed sites. Such virtual actors may interact with the environment, with one another and with the user by human-like behaviour (moving, speaking, using tools and other artefacts, etc.).[1] These new figures provide a more stimulating way to learn historical facts, as users are typically more involved and their concentration level is stronger.

The aim of this paper is to study of the positive effects that the use of avatars can provide to the educational experience, proposing a virtual museum with three different alternatives of storytelling, including one featuring virtual humans, and comparing the results in terms of engagement and understanding of the proposed content.

2 Previous Work

A virtual human is defined as a computer-generated character with the appearance of a human being. This term is used to indicate two virtual figures having different abilities: virtual agents and avatars. A virtual agent is a digital and interactive character, with or without human appearance, presenting several human abilities - like cognition or verbal/non-verbal communication – managed by some form of artificial intelligence. An avatar is a graphical digital representation of a human user, in either a two-dimensional (images, static or moving) or a three-dimensional form (i.e. characters from video games or virtual worlds).

A thorough study of the state of the art about this topic [1] has identified two macro categories of applications: those providing interaction with the users, and the ones that do not. Other possible taxonomies subdivide applications by content, identifying growing levels of immersion and interaction (if present): 2D animated, 3D animated, 3D immersive, Serious games. Users can relate to virtual agents or simply listen to their explanations. Single virtual humans with enhanced interaction capabilities are commonly used [2, 3] as the user's avatar in 3D immersive environments or when the situation requires a single counterpart (like in the case of virtual guides [4]), in other words in circumstances where the VH and the user must interact closely [5]. Multiple avatars [6, 7] (even crowds [8, 9]) are instead opportune to populate virtual reconstructions where the user explores large environments and the VHs are used for increased realism.

Modern cultural heritage application involving VHs nowadays provide a high quality visual interface, at the same time fostering the enhancement of the interaction between the user and the virtual agent, through natural and body language or other means of non-verbal communication. In the case of virtual guides, a high degree of realism and human-like verbal skills and personality traits, are required features in order to maximize their social acceptance and to increase the user engagement. Conclusions from this research point out that the use of virtual actors is able to stimulate a high quality learning experience, especially for young users. Many applications using

[1] Virtual Humans in cultural applications: a review. Octavian M. Machidon, Mihai Duguleana, Marcello Carrozzino, p. 1.

virtual actors have been developed following the *edutainment* concept, where learning is associated with having fun.

3 The Study

The research hereby presented is related to the investigation about the opportunities provided by virtual humans in digital learning environments dealing with cultural elements. The experiment presents users with an interactive visit of a virtual gallery of art objects, subdivided in rooms each related to a specific historical period, proposing three different types of aid assisting the visit:

- traditional expository panels, where explanations about the artworks are shown in text form;
- narrating voice, simulating a typical museum audio guide;
- virtual human, guiding the users around the virtual gallery.

The development of the application has taken into consideration the typology of contents in order to study the different emotional- cognitive reactions of the users. During the preliminary writing of the texts for the expository panels and the narrating voice, particular attention was paid to the definition of complete and concise information. In the same way, the creation of the virtual model involved an accurate selection of elements (3D models and textures) immediately conveying the idea of the historic period of the rooms.

The virtual gallery, modeled using Autodesk 3D Studio Max, is composed by five rooms:

- Entrance
- Egyptian room
- Classical room
- Medieval-Viking room
- Art gallery

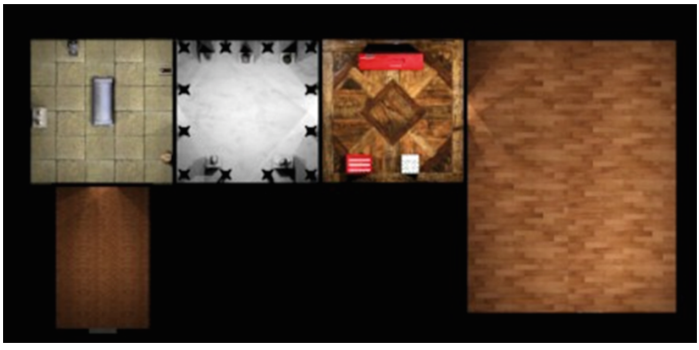

Fig. 1. Top-view of the virtual gallery

The entrance is a basic room providing access to the virtual gallery. The Egyptian room includes artworks from the V dynasty (2500–2350 B.C) and the XXVI dynasty (672–525 B.C). The environment features materials coherent with the historic period: hieroglyphics on the walls and rock tiles on the floor, to avoid a too invasive appearance. The Classical room includes sculptures and busts from the period 130–200 B.C. The space is composed by a pink marble colonnade, and other architectural elements textured with white marble, following the appearance of ancient classical temples. The Medieval-Viking room conserves different typologies of items: weapons, shields, coins and chess pieces. The style is rustic; walls are textured with rock bricks and the floor with dark wood. The last room reunites the most important artworks from Renaissance (XIV–XVI sec.) to contemporary age (1789-present). The pieces are key examples of every artistic movement, to help the user understand the history and evolution of the different artistic styles.

The items in the rooms are 3D models downloaded from *Sketchfab*[2]; in particular the British Museum offers several free 3D models of famous artworks with textures, many of which has been used in the Virtual Gallery. Coins and paintings have been modeled texturing simple 3D models with a jpg image of the original artwork. Every room features photometric lights, whose effects have been embedded in the 3D models

Fig. 2. Views of the virtual gallery rooms

[2] https://sketchfab.com/.

of artworks using lightmaps. The interactive application presenting the three virtual itineraries has been developed with XVR [10], an integrated development environment for the rapid development of Virtual Reality applications.

The structure of the virtual itinerary is composed by a series of viewpoints, which enable the exploration of the virtual gallery in a predefined sequence. Once in a room, the user is free to move around and explore the objects; however, when leaving the room, the next room is always the one predefined in the sequence.

4 The Three Itineraries

4.1 First Itinerary: Panels

In the first itinerary the learning process is regulated by fluctuating expository panels recalling a traditional museum visit. The fluctuating expository panels have been implemented as billboard, constantly pointing toward the observer. The style of explanations is fluid, not educational: only main concepts are presented in order avoid a decline in the attention level of the user. In the art gallery, panels focus on the meaning of the painting and on the idea of the artist.

During the virtual tour the panels are arranged in strategic points: one at the entrance of each room, with information related to the corresponding historic period, and one near every item, presenting the corresponding description. Panels appear when the user is located at a specific position and disappear when the user proceeds towards the next positions.

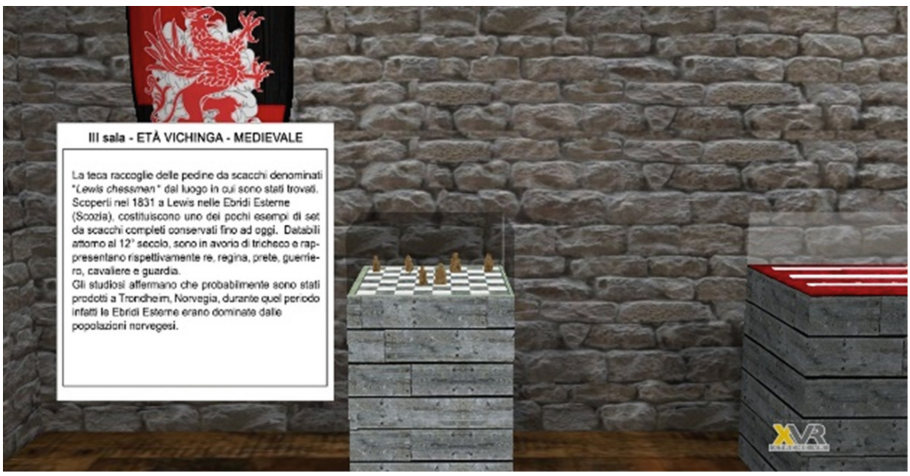

Fig. 3. View of the first itinerary, panels

4.2 Second Itinerary: Audio Guide

In the second itinerary, the user is guided through the virtual gallery by a narrating voice which describes and explains the content of each room. To keep a high attention level, the voice is young and comfortable. Contents are brief, with some sentences devoted to guide the users during the visit through the rooms, welcoming and inviting the user to proceed to the next. Following the same structure of the first itinerary, the audio content start when the user is placed in a specific position. The audio tracks have been implemented as 3D audio objects placed in the same position of the examined artwork.

4.3 Third Itinerary: Virtual Guide

In the third itinerary information is provided by a virtual actor with the same narrative outline of the second itinerary. The integration of the virtual actor in the XVR environment has been made by means of the Speaking Avatar class, an XVR class [11] offering features for the management of character animation and lip-sync with audio files or speech to text, based on the following mid-level tools:

HALCA (Hardware Accelerated Library for Character Animation), a library [12] based on Cal3D allowing to visualize and animate simultaneously several realistic avatars on simple display desktop, HDM's and CAVE-like devices.

SAPI (Speech Application Programming Interface), a Windows API providing a speech recognition feature retrieving text from audio data identifying phonemes and their duration. The Annosoft Speech Recognizer, based on SAPI, creates a text file of the sequences of such phonemes which are furtherly translated into *visemes* (the visual equivalent of phonemes, consisting in basilar postures of the mouth) in order to create a plausible lip sync.

The virtual actor is a young male character whose appearance of the avatar is extremely realistic: he always looks at users and engages with them, with a tone and a

Fig. 4. View of the third itinerary, virtual guide

voice coherent with his appearance and his casual outfit. The agent is positioned aside of every item, according to the behavior of a traditional museum guide.

5 Pilot User Study

The virtual application has been tested in a non-immersive desktop setup by 15 people, using a combination of mouse and keyboard for their interaction. Users have been divided into three groups, each one following a specific itinerary. Before starting the virtual tour, users had to compile a brief questionnaire aimed to collect personal information and previous knowledge with VR and digital cultural environments.

The user sample is composed of 7 male and 8 female subjects aged between 22 and 32. The education level varies from high school (2 users, 1 M and 1 F), bachelor degree (5 users, 2 M and 3 F), master degree (6 users, 3 M and 3F) and Ph.D. (2 users, 1 M and 1 F).

The duration of the test was different depending on the specific itinerary. The visit of the first itinerary lasts approximately 30 min, as it requires the additional effort of reading the panels, the other two last approximately 20 min, a difference that may has consequences on the attention level. After the visit, users completed a test with specific questions related to the items/artwork exposed in the gallery and questions about their emotional feelings, cognitive experience and engagement during the experience.

Questions relative to feelings, cognitive experience and comprehension had answers ranging from 1 (nothing) until 7 (totally).

The first question was: "How much was your senses involved in the virtual visit?" The diagram in Fig. 1 shows a high engagement level for the third itinerary and a strong variability for the second itinerary with a different interest level. This score could be due to a different academic background. Users with lower experience in virtual museums gave a high score in the test, while users with more competence in virtual cultural heritage gave a low score. Previous experiences in VR field have probably influenced the level of engagement during the virtual experience (Fig. 5).

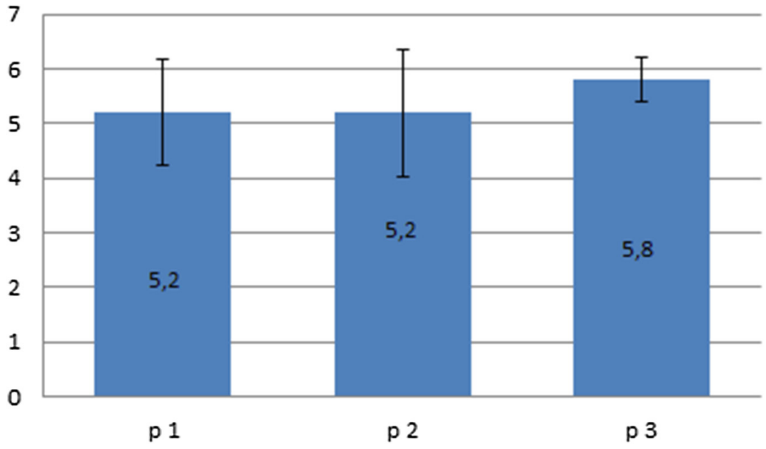

Fig. 5. Engagement of senses

The general appearance of the virtual museum involved mainly users experiencing the third itinerary (Fig. 2). This confirms findings of recent studies proving that the presence of a virtual realistic actor can generally increase engagement and attention of the user [5]. The first itinerary received instead the lowest score (4.8): we can argue that the absence of a vocal narration, leading to rely only on vision, made visitors more focused on the task and therefore less available to feel engaged in the environment (Fig. 6).

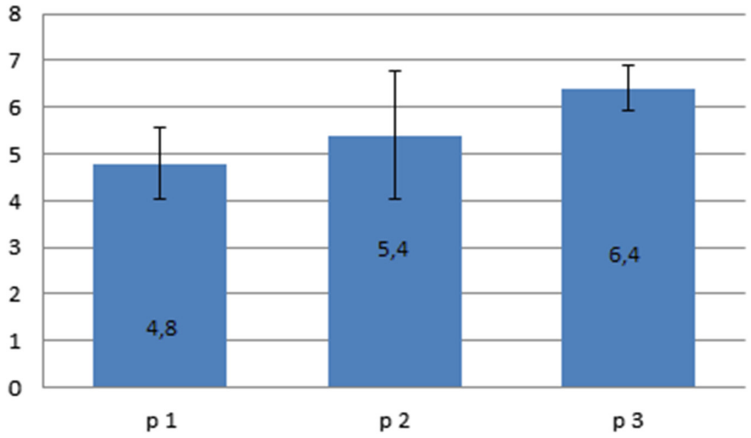

Fig. 6. Engagement of the visual experience

A comprehension question followed which was different depending on the itineraries ("Were the information panels clear and easy to read?" for the first one, and "Was the narrating voice clear and the provided information easy to understand?"). Results of this question demonstrate a general accuracy in the composition of the contents, in the planning of the expository panels and the choice of narrating voice (Fig. 3a). The third itinerary achieved the highest score (7): explanations provided by the virtual actor have been deemed complete and able to maintain a higher attention level.

Fourth question was related to pleasantness ("Was the text of the panels appropriate?" for the first itinerary, "Was the narrating voice pleasant?" for the second and "How annoying was the presence of the virtual actor?" for the third; in the latter case, 1 was a positive mark, i.e.: the presence of the virtual actor was not annoying at all. None of the users believed that the virtual agent could be annoying (all of the users gave a mark of 1) (Fig. 7).

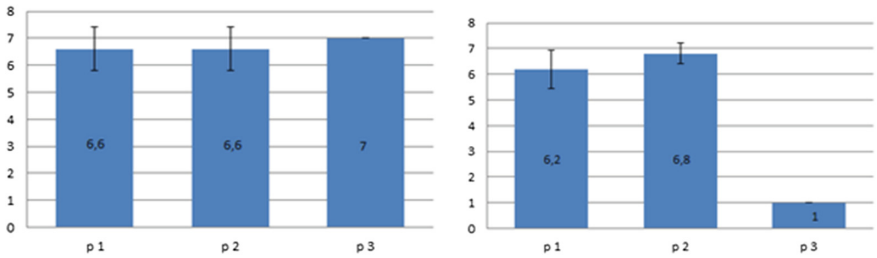

Fig. 7. (a) Third question: comprehension of contents (b) Fourth question: accuracy of contents

The last question concerned how much the experience helped the user to understand the content. The first and third itinerary received a very high score (Fig. 4). The second itinerary presents a greater variability of the results; this could be due to a not uniform level of experience in virtual museum of the second group of users that might have an impact on engagement but also on focusing attention (Fig. 8).

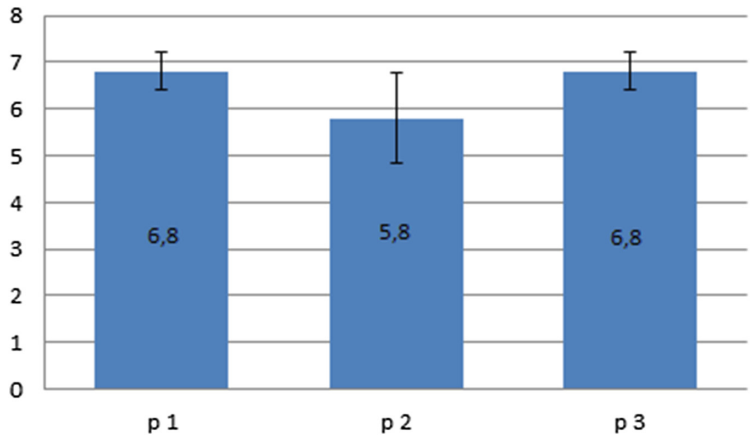

Fig. 8. Fifth question: how much the experience helped to understand the contents

The final results demonstrate that the third itinerary is the most complete one, users who visited gallery through the avatar's explanations found it easier and easily understandable. Users who attempted the first and second itineraries found them less clear and not useful for a complete learning process. The research confirms the important role a virtual agent can perform in a didactic application.

A second questionnaire aimed to estimate the efficacy of the implemented storytelling approaches in terms of learning and memory retention, posing two questions for each visited room related to the artworks thereby contained. No relevant differences were found with the partial exception of the users of the first itinerary who on average made more errors than the others. This could probably suggest a better efficacy of the latest two approaches in delivering content and in helping to keep a deeper attention.

At the end of the questionnaire, an open question asked users about suggested improvements. Besides some improvements requested on the visual quality of the environment, some users suggested to improve the quality of the visit with the insertion of more vocal narration, especially in the transition between rooms, in order to create a more organic and integrated experience and stimulate the visitor during the visit. Interestingly, all the gathered comments come from users of the first two itineraries, suggesting that the presence of a virtual guide lead to a better impression of completeness.

6 Conclusions

This paper proposes a pilot user study aimed to evaluate the efficacy of the presence of a virtual guide in order to implement an avatar-based storytelling approach to assist users in the visit of a virtual museum.

Preliminary findings confirm the hypothesis that an embodied virtual agent is able to stimulate attention and involvement, and contributes to a better content delivery and learning. We plan to extend this pilot study to a larger number of users, refining the comparison protocol, and to use a complete immersive setup based on head mounted displays.

Acknowledgments. This paper is supported by European Union's Horizon 2020 research and innovation programme under grant agreement No. 692103, project eHERITAGE (Expanding the Research and Innovation Capacity in Cultural Heritage Virtual Reality Applications).

References

1. Machidon, O.M., Duguleana, M., Carrozzino, M.: Virtual humans in cultural heritage ICT applications: a review. J. Cult. Herit. (2018, in press)
2. Kennedy, S., Fawcett, R., Miller, A., Dow, L., Sweetman, R., Field, A., Allison, C.: Exploring canons & cathedrals with open virtual worlds: the recreation of St. Andrews Cathedral, St. Andrews day, 1318. In: Digital Heritage International Congress (Digital Heritage 2013), vol. 2, pp. 273–280. IEEE, October 2013
3. Fanini, B., Pagano, A.: Interface design for serious game visual strategies the case study of "Imago Bononiae". In: Digital Heritage 2015, vol. 2, pp. 623–626. IEEE, September 2015
4. Kopp, S., Gesellensetter, L., Krämer, N.C., Wachsmuth, I.: A conversational agent as museum guide – design and evaluation of a real-world application. In: Panayiotopoulos, T., et al. (eds.) IVA 2005. LNCS (LNAI), vol. 3661, pp. 329–343. Springer, Heidelberg (2005). https://doi.org/10.1007/11550617_28
5. Swartout, W., et al.: Ada and Grace: toward realistic and engaging virtual museum guides. In: Allbeck, J., Badler, N., Bickmore, T., Pelachaud, C., Safonova, A. (eds.) IVA 2010. LNCS (LNAI), vol. 6356, pp. 286–300. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-15892-6_30
6. Bogdanovych, A., Rodríguez, J.A., Simoff, S., Cohen, A., Sierra, C.: Developing virtual heritage applications as normative multiagent systems. In: Gleizes, M.-P., Gomez-Sanz, J. J. (eds.) AOSE 2009. LNCS, vol. 6038, pp. 140–154. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-19208-1_10
7. De Paolis, L.T., Aloisio, G., Celentano, M.G., Oliva, L., Vecchio, P.: Experiencing a town of the middle ages: an application for the edutainment in cultural heritage. In: IEEE 3rd International Conference on Communication Software and Networks (ICCSN), pp. 169–174. IEEE, May 2011
8. Maïm, J., Haegler, S., Yersin, B., Mueller, P., Thalmann, D., Van Gool, L.: Populating ancient pompeii with crowds of virtual romans. In: Proceedings of the 8th International Symposium on Virtual Reality, Archeology and Cultural Heritage-VAST (No. VRLAB-CONF-2008-151) (2007)

9. Carrozzino, M., Piacentini, V., Tecchia, F., Bergamasco, M.: Interactive visualization of crowds for the rescue of cultural heritage in emergency situations. SCIRES-IT-SCIentific RESearch Inf. Technol. **2**(1), 133–148 (2012)
10. Carrozzino, M., Tecchia, F., Bacinelli, S., Cappelletti, C., Bergamasco, M.: Lowering the development time of multimodal interactive application: the real-life experience of the XVR project. In: Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology, pp. 270–273. ACM, June 2005
11. Brondi, R., Carrozzino, M., Lorenzini, C., Tecchia, F.: Using mixed reality and natural interaction in cultural heritage applications. Informatica **40**(3), 311 (2016)
12. Normand, J.M., Spanlang, B., Tecchia, F., Carrozzino, M., Swapp, D., Slater, M.: Full body acting rehearsal in a networked virtual environment—a case study. Presence Teleoperators Virtual Environ. **21**(2), 229–243 (2012)