

Augmented Real-Time Virtual Environments

Vanja Jovišić
Faculty of Electrical Engineering
University of Sarajevo
Sarajevo / Bosnia and Herzegovina

Abstract

The focus of our research was to experiment with different approaches for digital storytelling in virtual environments. It consists of including a real avatar into the real time environment as a storyteller. Our avatar is an actor recorded against the green screen background and inserted into the real time virtual environment (VE) using two different techniques. The paper discusses the comparison of these techniques for obtaining the best result – VE augmented using the real avatar.

Keywords: Avatars, Virtual Reality, Virtual Heritage, Digital Storytelling

1 Introduction

Virtual reconstructions often provide highly realistic visual experience of the reconstructed object (environment) whether the real time VE or precompiled ones are used. The main aim of the virtual reconstruction should not be limited only to a photorealistic visual presentation of different examples of cultural heritage, but should provide a meaningful and interactive way to experience different intangible elements of cultural heritage. Digital storytelling [7] has been used to communicate that information (a story) to the user of the virtual environment, because stories have been an important part of all human culture. Thus, virtual environments are often populated with interactive virtual humans (avatars) who are either storytellers or main actors of the story [2]. Considering the complexity of modeling and animation of virtual avatars used for illustrating real life, and the fact that creating a convincing user experience is the most challenging task, we tried to use another approach for inhabiting virtual worlds.

The goal of this work was to create an augmented virtual world where the real-world avatars will be used for populating the virtual environment in order to enhance the user's experience.

This paper describes the process of creation of the real avatars and embedding them into the VE. The paper is organized in the following manner: Section 2 describes the present status in the research in this area, Section 3 offers the insight into the Travnik Fortress virtual reconstruction project, Section 4 explains the methodology used for obtaining the goals set and Section 5 presents the results of the comparison of the presented

techniques. Section 6 offers the conclusion and our future work directions.

2 Background

Virtual Avatars have been extensively used for populating virtual environments in many projects. Recent research presents various frameworks for assembling the scenes and for design and animation of Virtual Humans using speech and sign language for communicating stories to the users. User experience is improved with natural language interactivity within the environment [3]. On the other hand, it raises the question how the users react to the synthetic avatars used for conveying information or the stories. It is especially true in the case of human like faces who focus the attention of the user who has high expectations when it comes to the realism of the protagonists of the real life stories [3]. It is shown that the real life avatars conveyed the information with greater success than virtual avatars in virtual heritage storytelling applications. [1]

In order to classify our work we were exploring the definitions of virtual reality, augmented reality and augmented virtuality.

Trying to define our approach and place it into the Virtuality Continuum (VC) [4], the closest approach used is in the case of Augmented Virtuality where an remotely viewed video is perceived as "real" and computer generated environment as "virtual". This approach is mainly used in the area of tele-immersion (tele-presence). Users at geographically distributed sites can collaborate in the real time in a shared, simulated environment. Applications of this approach can be seen in design and prototyping, archeological reconstructions and other technologies or 3D video/conferencing.

The connection between two approaches is in using chroma-keying to distinct the non important physical environment (in our case, the background of the actors). Still, our example cannot be classified onto any of these categories as we are using the video captured prior emersion into the real time virtual environment.

In our work, the major role of avatars was to tell the story about the reconstructed object through two different scenes. The similar approach to digital storytelling can be found in the case study of Vindolada Roman Fort [6] where the story is describing the different scenes from the life in the fort. Here, the video footage was composited with pre-rendered 3D environment and included into the multimedia application in the archeological museum.

3 Travnik Fortress Project

The Fortress "Old Town" in Travnik represents one of the most beautiful and most preserved facilities of medieval Bosnia, on which the later historical periods left their specific marks. Reconstruction of the photorealistic look of the fortress has the final outcome in form of a user-friendly, multimedia interface (Figure 1) that provides a comprehensive user experience and raises the awareness about cultural and historical importance of the reconstructed object.

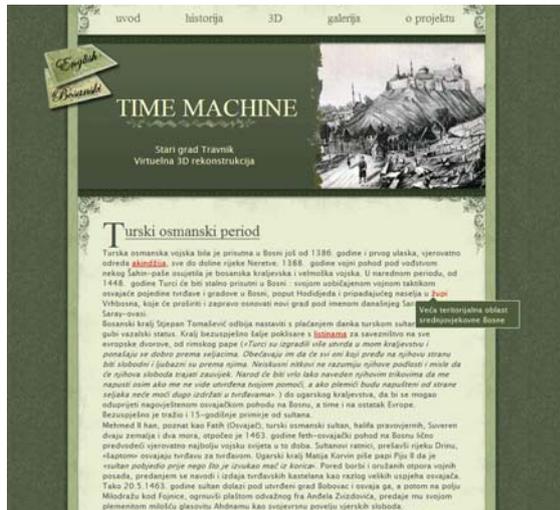


Figure 1: Home page of the project

The main element of the project was a digital story about the fortress Old Town in Travnik in one specific period of its long history, the period of Ottoman Empire, illustrating the life of its inhabitants in that period.

The story consists of the following scenes:

- First mention of the name of the city and the fortress in the history
- Conquering the fortress by the troupes of Sultan II El Fatih
- Military function of the fortress (different scenes showing the everyday life of the soldiers in the fortress)
- Fortress as a prison (casemate)

The user is guided through the story by a narrator and the experience is enhanced with different language versions of the narration and music background which was created using the traditional Ottoman music as an inspiration. Figure 2 shows a screenshot from the movie.



Figure 2: The scene of invading the Fortress

The first step was to create a detailed 3D model of the Fortress using the following procedure:

- Data collection
- Presentation design
- Modeling of the objects in Autodesk 3ds Max software
- Modeling of the terrain using DEM terrain data
- Textures generating and mapping
- Illumination of the model
- Camera animation and rendering

After this, we recorded the actors wearing traditional costumes from the given period against the green screen background and composited them with the pre-rendered model environment using Adobe Premiere Pro 2.0 and After Effects Adobe CS2.

This approach is used instead of making a completely interactive virtual environment because it offered more realistic user experience purposely giving up the interactivity. Users' feedback was very positive to this approach as the user does not need to be technically skilled person to be able to enjoy the story about the Fortress. [5]

The next step in the project development was to experiment with a different approach, making an interactive environment from the 3D model and enhancing it by inserting the real avatars - storytellers.

4 Real Characters in the VE

The goal of our work was to experiment with the new ways of inhabiting the virtual environments. Two different scenes were selected for immersion with the real-time environment. One represents the narrator (in the previous version of the project, narrator was off-sound) who is telling a short story about the history of the Fortress. The second scene is illustrating a short example of everyday life of the soldiers in the fortress, where two actors, costumed into the traditional clothes for that period, are playing chess.

Although we had many scenes recorded from the previous production phase, the video footage used in this project was recorded again due to the technical issues appeared in the previous project, where the following problems were fixed:

- Lighting of the green/blue key background
- Size of green/blue key background
- Planning the location where the live scenes will be positioned in the 3D model

4.1 Filming and Postproduction

Video footage was recorded against the green screen background. The previous experience showed that it was the best to use one fixed camera and capture the whole character, so he can be set on the planned location in the 3D environment.

The video is imported into Adobe After Effects CS3 software and keyed with the Keylight 1.2 plug-in which enabled the better quality of the final video sequence.

The process of keying the material depends on the quality of the green screen set, including the background and light conditions. If the raw material has a good quality, not many modifications of the default settings need to be done. But, there are many techniques of keying and even material with a bad quality can be improved to give the usable results. In the Figure 3 we see an example of the source material.

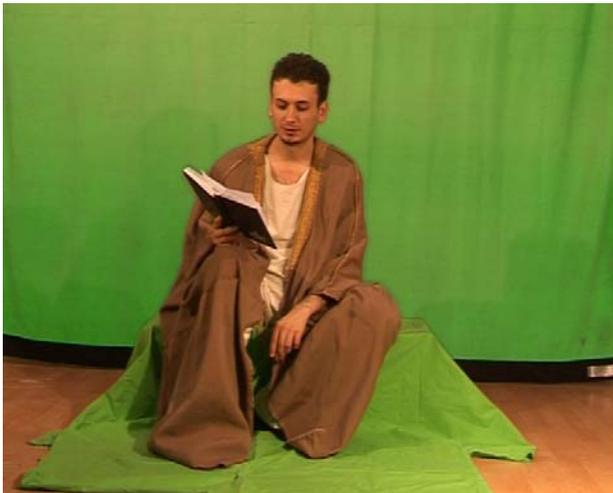


Figure 3: Video footage recorded against the green screen

Figure 4a shows the video material with the garbage mask applied to remove the unnecessary environment and the final result of the keying process (checker background illustrates transparency - Figure 4b).



Figure 4a: Video footage before keying



Figure 4b: Keyed video footage

Videos have the Hue/Saturation values adjusted based on the values of the background environment. This way, the real avatars blended better to the real time virtual environment.

The output format of the video is uncompressed Video for Windows (.avi) with RGB and Alpha channels included and Millions of Colors+ color depth. These settings showed the best results for our purpose.

This way, alpha information (opacity) is linked to each of the pixels and enables superposition of video, minus background.

4.2 Importing into Web 3D

For comparison reasons, two different applications for creating Web3D content were used. First is a commercial, highly sophisticated software for creation of the virtual environments, Quest3D. The experiment was performed on the time limited demo version with license obtained by the software reseller. The second approach was to experiment with open source methods and standard Web3D formats such as VRML, VRML 2.0, X3D.

X3D is the only Web3D standard that supports alpha channel of the video footage. This X3D feature is used to import 3D object mapped with video texture of the certain scene into the virtual environment. The real

avatar is represented by a simple Plane object mapped with the video texture.

For QUEST 3D import we used the following methodology:

1. Preparation of the model for export consisted from the following steps:
 - All unnecessary objects were removed from the scene. This included helpers, camera paths, cameras, lights, all polygons on the objects that are not visible to the user
 - Objects are marked with UVW and UnwrapUVW modifiers, so they can be regularly exported to Quest3D software. The main reason is that UV coordinates specify the application of each pixel of the texture on the surface. When the object is correctly mapped once, due to the relative coordinates used, any size of texture map can be applied to it. This way, specific objects on the scene can be mapped with low/resolution maps if they are not visible or important for the user. This is critical for the performance of the environment. Some other methods are used for optimization such as clipping planes on the camera which defines virtual limits of rendering scene, or by including Fog as an atmospheric element which settles sometimes very harsh cut of the clipping plane.
 - For lighting and realistic shadows, the light mapping support of Quest3D is used. Light maps are baked textures keeping the data about lights and shadows on the object. They are applied as a second layer of the object texture, with diffuse maps containing only color data applied as the first layer.
2. Export of the model using in Autodesk 3ds Max software into the Quest 3D compatible format. There are a couple of exporters on the market used for this process and one of the most popular is Panda Exporter.

The model is imported into the Quest 3D scene as an individual channel. Other channels were the chess player and storyteller planes with applied movie texture. Storyteller has also a sound channel connected, so the effect of 3D sound can be created. Walkthrough camera is used for exploring the object.

The simplified organization of the connections between the nodes is shown in the Figure 5.

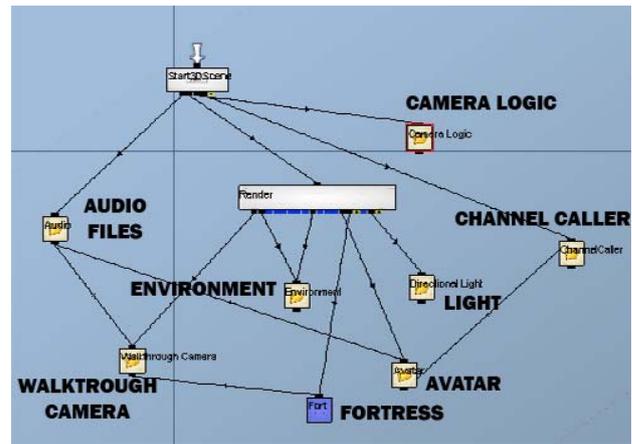


Figure 5: Channel organization in Quest 3D

In order to achieve the transparency of the video footage background, we needed to use specific alpha blending of the applied movie texture.

Creating the interactive virtual environment in the X3D format is a straight forward procedure. The model is exported from the Autodesk 3ds Max software using its embedded exporter. This file is then imported into Vivaty Studio Beta 1.0, Web3D authoring tool which enabled experimenting with different export formats. Using this tool, the movie texture of the stories can be applied to the object and enables that different user navigation possibilities can be added to the environment.

5 Description of experiences - Quest 3D versus X3D

Let us consider the comparison of the input parameters for Quest3D and X3D environment.

Number of polygons in the model:

Scene exported into Quest 3D consists of 159888 polygons.

Scene exported into X3D consists of 78190 polygons (large terrain was excluded).

Size of textures:

In Quest3D, objects are mainly textured with 512x512 textures (taking 1365KB of texture memory each). Eight objects that were considered important for obtaining realism are textured with 1024X1024 textures (each taking 5461KB of texture memory). All other objects are covered textured with light maps of the same size. Overall, 74 texture maps are used.

In X3D the same textures were used, without light maps.

File Size of the captured video:

Narrator scene: 80 seconds of uncompressed video with resolution of 324x240 which is chosen as optimum between quality of appearance and size has 567MB.

Chess Players: 120 seconds of uncompressed video with resolution of 324x240 and size of 1024MB.

Quest3D software offers an extremely efficient way for creating highly realistic VR environments even for the

users with no programming experience. The published outcome (software supports different output formats) is a very fast and realistic standalone environment created using a very friendly user interface. Quest3D cannot perform the complete real-time keying, which can be noticed when the image texture with alpha channel is applied. Movie texture applied to a plane is blended with the environment and is slightly transparent (Figures 6 and 7).



Figure 6: Real avatar embedded into the VE using Quest3D



Figure 7: Real avatars embedded into the VE using Quest3D

X3D demands better understanding of the VRML standards and programming techniques, but offers the expected result. The navigation through the environment is slower even though the same image textures were used in both examples.



Figure 8: Real avatar embedded into the VE using X3D



Figure 9: Real avatars embedded into the VE using X3D

The main drawback of both methods lies in losing the third dimension of the avatar (Z axes). Thus, the user cannot have a side view of the object containing avatar. Output parameters we considered were: the size of the exported VE environment, the realism of the scene achieved, alpha blending result, navigation modes and available export formats.

Comparison results of the two techniques are shown in the following tables:

Input parameters	Quest3D	X3D
Number of polygons in the model:	159888	78190 large terrain excluded
Size of textures:	512x512 1365KB each 8 objects 1024X1024 5461KB each 74 texture maps	same textures without light maps.
Captured video File Size	Narrator scene 567MB Chess players 1024MB	same

Table 1 - input parameters comparison between Quest 3D and x3D

Output parameters	Quest3D	X3D
VE Size	~ 22,237MB with media files included and compressed	~ 2,777MB with external media files referenced and uncompressed
Alpha blending	Movie texture appears partially transparent and shows that not only opacity pixels were considered during keying	Movie texture is keyed properly
Realism	Enhanced with 3D sound, light maps	Advanced options required more work
Navigation	Depends on the chosen navigation format	Standard VR navigation tools
Export Format	.exe, web width	X3D format

Table 2 - output parameters comparison between Quest 3D and x3D

- [4] P. Milgram, F. Kishino, *A Taxonomy of Mixed Reality Visual Displays*. IEICE Transactions on Information Systems, Vol E77-D, No.12 December 1994.
- [5] V. Jovišić, S. Rizvić, *Photorealistic Reconstruction and Multimedia Presentation of the Medieval Fortress in Travnik*. Review Of The National Center For Digitization, 13/2008, Faculty of Mathematics, Belgrade, Serbia, ISSN 1820-0109
- [6] A. E. Carillo, K.Rodriguez-Echavarria and D. Arnold, *Digital Storytelling and Intangible Heritage: Vindolada Roman Fort Case Study*. In Proc. VAST 2008, Short Projects and Papers volume, pp. 85-89.
- [7] A. S. Glassner, *Interactive Storytelling: Techniques for 21st Century Fiction*. AK Peters Ltd, 2004, pp 27
- [8] A. Trnka, *Implementation Of The Raster Image With Alpha Channel In X3d Virtual Environments*. Diploma Thesis, Faculty of Electrical Engineering Sarajevo, July 2008.

6 Conclusion and Future Work

The techniques described in our research show that there is a convenient way of combining real avatars with real time virtual environments. This combination results in much more realistic and immersive user experience. When it comes to the complex stories from the past that need to be presented to the user, real avatars are much better choice with more satisfying results.

Our future work will focus on the further exploration of digital storytelling applied to the cultural heritage virtual environments. We will concentrate on discovering the most efficient way to present the augmented real/time environment with the highest level of immersion of the user. We will also explore the different storytelling techniques for achieving the best results.

References

- [1] A. Sadzak, S. Rizvic, C. Dalton, A. Chalmers, *Information Perception in Virtual Heritage Storytelling Using Animated and Real Avatars*, SCCG 07
- [2] D. Arnold et al, *Tools for Populating Cultural Heritage Environments with Interactive Virtual Humans*. EPOCH Conference on Open Digital Cultural Heritage Systems, 2008
- [3] Janet H. Walker et al, *Using Human Face in an Interface*. Human Factors in Computing Systems, 1994