Adding the virtual reality component to audio stories

Ajla Sukrija[*]
Sanda Sljivo[†]
Supervised by: Selma Rizvic[‡]

Faculty of Electrical Engineering
University of Sarajevo
Bosnia and Herzegovina

Abstract

Computer graphics can provide a connection between the literature and visual arts. Humans learn through emotions and personal experience. Virtual Reality is providing them with both, and therefore it nowadays has many applications not only in entertainment and edutainment but also in therapy. We have selected Virtual Reality for a visual representation of audio stories. The goal of this work is to augment the experience the users have while they listen to audio stories. In this paper, we have performed a case study with different Virtual Environments for three audio story styles: children poem, old traditional Bosnian romantic poem and science fiction story. An initial user evaluation has been carried out in order to see the impact that Virtual Environments have on the user experience.

Keywords: virtual reality, user experience, user immersion, audio stories

1 Introduction

The movie industry successfully makes personalization of literature. Movies are still one of the most common forms of entertainment. However, with the breakthrough of new technology, Virtual Reality (VR) and its accessibility in the mainstream with the launch of the mass-market headsets for consumers, movies can easily be surpassed. Humans learn through emotions and personal experience. VR is providing them with both, and therefore it nowadays has many applications not only in entertainment and edutainment but also in therapy. The modernization of our society has brought us many benefits. As one of the negative sides of this, people have lost their passion for reading books. VR can modernize the way that people read and rekindle that lost passion.

Until now audiobooks have been used to teach children to read and to increase reading comprehension [3]. They are also useful for people with visual impairment. However, they are also often used by people who are usually

working on something while listening to an audiobook. Audiobooks can increase a person's interest in the text and allow them to enjoy a book, even if it is above their reading level.

We have selected VR for a visual representation of audiobooks because of the immersion factor it provides. According to Aronson-Rath et al. [1], VR is an immersive media experience with a real or imagined environment that allows users to feel as if they are there. The goal of this work is to show that VR can augment the experience the user has while they listen to audiobooks.

The structure of this paper is as follows: in Introduction and Related work chapters, we give our aim and motivation and a short review of similar projects. In the VR simulation chapter, we describe a case study with different Virtual Environments (VE) for three audio story styles (children poem, old traditional Bosnian romantic poem and science fiction story). In the Evaluation chapter, an initial user evaluation shows the impact that VEs have on the user experience. In the Conclusions chapter we discuss our conclusions and propose future work.

2 Related Work

There is a number of projects which evaluate user immersion in VR applications. In this chapter, we discuss some of the relevant projects to our research.

Jones [4] has studied the impact of VR in immersive journalism narratives. She proposed a framework for further research into immersive technologies and storytelling. She performed user evaluation which measured the user's empathy and immersion in 360° video stories. The results showed excellent user immersion in character-led narratives, but in the reporter-led narratives case, users felt restricted and guided by the reporter.

Another project that studied immersive narratives in journalism is implemented by Longhi [6]. The author presented a review of two VR applications where users could experience the lives of a prisoner and refugees. Similarly to our project, their stories are visualized through VEs, and then the user immersion was measured. The evaluation has

---

[*]asukrija2@etf.unsa.ba
[†]sanda.sljivo@etf.unsa.ba
[‡]srizvic@etf.unsa.ba

shown that the user immersion depends on how strongly the emotions are presented in the narrative used in VR application.

Rupp et al. [8] were measuring user immersion and presence and their impact on learning new information throughout the use of VR in education. The authors wanted to show that the use of VR in education can enhance the learning process, but the evaluation showed surprising results where users were so overwhelmed by the VR experience that they could not focus on the narrator in the 360º video.

The work of Ziegler [10] is slightly different from the already mentioned projects in the sense that it has used real-life objects instead of VEs to augment the user experience of audiobooks. A team of museum curators and artists made an audio drama with a corresponding animated film which was projected onto the exhibits and walls of the exhibition space. In that way, not only the narrative was augmented but the exhibits themselves. Another case study from this paper is a room used as a time machine where a VR content projected onto the walls augments the user experience while the users are listening to a narrative about Chinese history. This method might be improved with the use of VR head-mounted displays (HMDs) instead of real-life environments.

Another method that could benefit with the use of VR is by Molnar and Kostkova [7]. The authors have presented an Interactive Digital Storytelling (IDS) formalization using Final State Machines. They have implemented this logic within their authoring tool for development facilitation of the IDS based games. In the end, they have performed a user evaluation in several schools. The authoring tool produced mostly text-based games, and because the children's reading proficiency was not very good, most of the children would stop playing the game before even finishing it. This tool could be improved with the use of VR technology in future iterations.

The idea to augment user experience while reading books is not recent. In Billinghurst et al. [2] the authors had developed an Augmented Reality (AR) application using for today's standards an outdated technology. The user would hold an AR hand-held display connected to a computer in a network and read a book with AR markers and had an augmented experience. They could also see other users in the network who were reading the same book at that time. Because AR technology has advanced recently, it would be interesting to see the modernized version of this application. The big advantage this approach has unlike ours is the ability for multiple users to use the application at the same time.

3 VR simulation

Our initial idea was to design a VE that will represent a universal story narration environment, where the user can insert any audiobook of their choosing. However, we have decided to create three different environments instead of a universal one, because a single environment could not represent the emotions and atmosphere of different stories as we wanted to stimulate the users' emotions. In order to achieve the full user immersion, several different factors that will work in synergy must be considered. As the perception is one of the principal factors, we have changed it in different contexts, trying to bring the user closer to the environment to experience emotions. The approach selected for narration is different for each audiobook: narration without characters, narration through gestures, and speech animation.

3.1 Application Structure

We developed a pilot VR application for the HTC Vive HMD in the Unity game engine and designed dedicated scenes for each VE. As the first environment where the user starts, we have chosen a library, because a library is a place where one can go and read the books. From there, the user can read the instructions and choose which story they want to experience using the Menu as shown in Figure 1. The menu consists of pickable 3D objects with a collider and rigidbody components added which were attached to HTC Vive controllers. Also, the user can call the menu from a scene where they currently are and switch to some another scene of their choosing. The library scene is the only scene without any background music or narration.

The other three scenes represent the children poem, old

Figure 1: The library scene

traditional Bosnian romantic poem, and a science fiction (SF) story.

3.2 Scenes

The first implemented scene was a children poem VE (Figure 2). This environment puts the user in a fantasy cartoony world while an animated fox near the fire tells the story. The idea was not only to make the adult user feel like a child but to also bring the children closer to the environment through something familiar to them, like cartoons. In this VE, we have focused on character animation because we thought that this approach would be the best for the representation of the children poem. The environment has been implemented in a way that the user could immerse in storytelling regardless of the inserted

Figure 2: The environment of the children poem

Figure 4: The environment of the SF story

Figure 3: The environment of the romantic poem

Figure 5: A user watching the children story

story or the poem. The focus for the current solution was to visualize and present the possibility of such narration. The application has no possibility of uploading audio recordings of different lengths to accommodate the animation. In future work, we will further discuss this possibility.

The second scene was an environment with romantic elements, based on a Bosnian folk poem, with a male and female character. The characters are standing across each other while the song was played in the background (Figure 3). The surroundings were made according to the song. As a narration, we have chosen "Emina" an old romantic poem by Aleksa Šantić and sang by Himzo Polovina. It is a poem about forbidden love. The man is an Orthodox Christian and in love with the Muslim girl Emina. He watches her in the garden of her house and secretly suffers because he cannot be with her. In this case, we have focused on the animation of the characters' gestures moving with the pace of the song because we considered that this approach is the best for the representation of the songs. This scene was animated by hand, and the environment was modeled to match the storyline of the song.

The third environment was the SF scene which was completely different from the previous two scenes in a way that it did not have any animated characters, just the vast sky with the stars, large mountain hills, and the atmosphere of the another planet (Figure 4). In this scene, we have created a particle animation that reacts to the narrator's voice frequency. Since there are no actual characters in the scene, it is possible to add another story with different duration to this environment.

We have used different assets for the implementation

of these environments, all free for commercial use [9]. Particles used were fog, falling stars, and imitation of the polar light. All materials and 3D models are open source assets.

4   User Experience Evaluation

The initial user experience evaluation involved 15 participants. The participants were self-selected, and the demographic balance is presented in Table 1. All of them had normal or corrected to normal vision and no hearing impairments. We had almost equal number of male (8) and female participants (7). Most of our participants were in age group between 21 and 30 years (6) and over 50 years (4). Since one of the stories is dedicated to children, we invited two children as participants.

Before watching the stories (Figure 5), participants filled out the demographic questionnaire including questions whether they had read audiobooks and tried VR applications before [5]. They also rated how much they liked three types of stories: SF, romantic and children story.

The experiment was conducted in the following manner: immediately after watching each story participants filled out the part of the questionnaire related to the story. In this way, we provided for immediate feedback and avoided mixture of experiences. In the end, the participants were asked additional questions considering overall satisfaction with the application.

Question	Responses	Number of participants	Percentage %
Gender	Male	8	53.3
	Female	7	46.7
Age	0-10	1	6.7
	11-20	1	6.7
	21-30	6	40
	31-40	2	13.2
	41-50	1	6.7
	Over 50	4	26.7
Prior use of VR	YES	8	53.3
	NO	7	46.7
Have you ever listened to the audiobooks before?	YES	11	73.3
	NO	4	26.7
How often do you read books?	Yearly	5	33.3
	Monthly	7	46.7
	Weekly	0	0
	Daily	3	20
How often do you watch movies?	Yearly	0	0
	Monthly	4	26.7
	Weekly	8	53.3
	Daily	3	20

Table 1: A demographic questionnaire with additional questions

The questionnaire contained five sections. The introductory part had questions for user profiling (Table 1); the evaluation part was organized separately for each story. Each evaluation part contained Likert items with 7-point scale, grouped into two groups: experience and implementation. Each group had four statements. All statements except the statement I3 were expressed as positive logic. The score for I3 presented in Table 2 was complemented before it is used in the analysis. Table 2 shows mean values for all three VEs. The concluding part contained eight Likert items with 7-point scale and also open questions where participants could express their opinion about the application.

Some of the examples of Likert items used in the questionnaire:

- This approach to story narration has great potential;

- The VE has depicted the poem narration;

- With listening to the song, experience with the VE has become better;

- I did not lose interest during the narration;

- I would like to listen to this and other SF stories again;

The participants have given mostly positive feedback; they liked all three approaches to the visualization of audiobooks. The most positive feedback was for SF VE. Based on the experiment we have organized it is not easy to identify the motivation for this score: if it is linked to narration approach, SF style, the attractiveness of the particular scene, or the participants were influenced by order of watching. To clarify the motivation for this score, we should add the additional questions in our evaluation. The results are presented in Table 2 and visualized in Figures 6 and 7. In both figures and the Table 2 can be seen that the results are mostly positive.

It should be noted that on a weekly or daily basis, only 20% of them read books, but most of them (73.3%) watch movies. Because of that, we could say that they prefer this type of media over other forms. This was interesting because we could see that the users prefer visual over audio presentation. Our application was also appealing to children. All the participants except one would recommend this application to children. In the section with open questions, some of the participants said that the application could be improved if narration and VE are more synchronized.

5 Conclusions

The books were one of the first media for entertainment. Then came the movies and audiobooks. Nowadays virtual reality offers the user what books, audiobooks, and movies lack; a feeling of presence.

Our results showed that the feeling of immersion and presence are important to users. Considering the results of the user evaluation, we can conclude that VEs have a positive impact on user experience with audiobooks. With further development of VR technology, we can only

Statement	SF	Romantic poem	Children poem
Experience			
E1. I like the scene with the narrator voice/song in the background	6.4	6.4	5.5
E2. The scene reflected narration	5.8	6.1	5.3
E3. Listening to the narration enhanced my experience	5.6	5.7	4.9
E4. I would like to experience this VR story or similar again	5.7	5.6	4.9
Implementation			
I1. I carefully listened to everything through the end	6.3	6.8	5.5
I2. While listening, I was exploring Virtual Environment	6.7	6.9	6.7
I3. I lost interest before the end	5.2	5.7	5.7
I4. I have remembered some details from the narration/poem	5.5	6.7	4.9

Table 2: Statements with the mean values of the levels of agreement for all three Virtual Environments
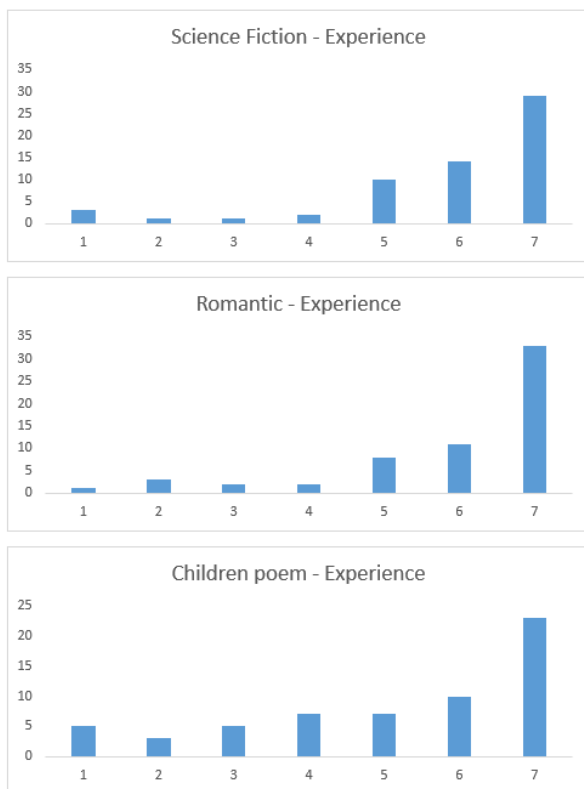

Figure 6: The frequency of responses considering the experience of the user (1 – strongly negative; 7-strongly positive)
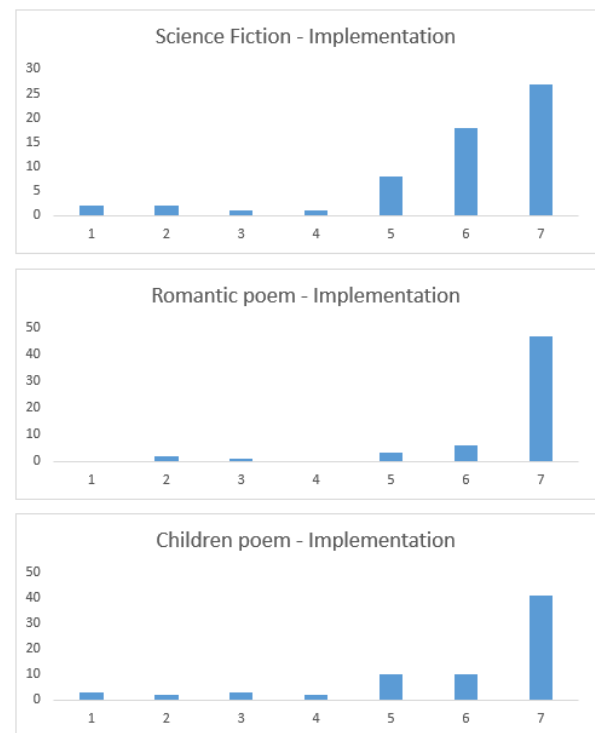

Figure 7: The frequency of responses considering implementation (1 – strongly negative; 7-strongly positive)

dream what the next level of immersion awaits us.
However, our approach is not suitable for the type of users who choose audiobooks because of the possibility of multitasking while listening to the narration.

For future work, we will add more interactivity in VEs and explore how to make VEs with the help of artificial intelligence and intelligent agents.

References

[1] R. Aronson-Rath, J. Milward, T. Owen, and F. Pitt. Virtual reality journalism. `https://towcenter.gitbooks.io/virtual-reality-journalism/content/index.html`, 2015.

[2] M. Billinghurst, H. Kato, and I. Poupyrev. The magicbook: a transitional ar interface. *Computers Graphics*, 25, 2001.

[3] J. Hodgson. American english, listen up! using audio books for english teaching. `https://americanenglish.state.gov/files/ae/resource_files/listen_up_using_audio_books_in_the_classroom.pdf`.

[4] S. Jones. Disrupting the narrative: immersive journalism in virtual reality. *Journal of Media Practice*, 18(2-3):171–185, 2017.

[5] J. Lazar and and Hochheiser H. Feng, J. H. *Research Methods in Human-Computer Interaction*.

[6] R. R. Longhi. Immersive narratives in web journalism. between interfaces and virtual reality. *Estudos em Comunicação*, 1(25), 2017.

[7] A. Molnar and P. Kostkova. Interactive digital storytelling based educational games: Formalise, author, play, educate and enjoy! - the edugames4all project framework. *Transactions on Edutainment*, XII:1–20, 2016.

[8] M. A. Rupp, J. Kozachuk, J. R. Michaelis, K. L. Odette, J. A. Smither, and D. S. McConnell. The effects of immersiveness and future vr expectations on subjec-tive-experiences during an educational 360 video. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 60, pages 2108–2112. SAGE Publications Sage CA: Los Angeles, CA, 2016.

[9] Unity. Unity asset store. `https://assetstore.unity.com/`.

[10] T. Ziegler. Innovative, participatory listening and seeing experiences in museums. In *Proceedings of the Conference on Electronic Visualisation and the Arts*, pages 306–313. BCS Learning & Development Ltd., 2015.