
ViReSt - Storytelling with Volumetric Videos

Gerhard Sprung, Andreas Egger, Alexander Nischelwitzer, Robert Strohmaier, Sandra Schadenbauer

Dep. Information Management, FH JOANNEUM, Graz, Austria

Email: [first].[last]@fh-joeanneum.at

Abstract

The aim of the project ViReSt is to explore the extent to which virtual reality, volumetric video and binaural audio can be utilized to create an immersive experience for museum visitors. Museum curators should be able to develop, test and create the story themselves without programming skills. Therefore, a workflow was developed and tested by creating a small prototype. This prototype was then tested with a group of 10 participants to assess immersion, reception of volumetric video and binaural audio.

Keywords — VR; volumetric video; storytelling

I. INTRODUCTION

In the near future Virtual Reality (VR) will be mature enough to be widely used in museums. The project ViReSt (VIRtual REALity StoryTelling) addresses the question of how to create compelling content for such projects. With our prototype, we explored what tools are or would be needed to give museum designers without programming knowledge the ability to create interactive stories that are immersive, emotionally engage visitors and provide a level of interactivity.

The “wow-effect”, which is usually sufficient to attract the visitors’ attention to an exhibit that is presented with VR, will no longer suffice as soon as the visitors will already have experience with VR-devices. In the same way as in the advent of film and again in the early stages of the computer game industry, traditional methods and approaches for creating immersive experiences have to be rethought, reworked, redefined or even created from scratch [1].

II. STATE OF ART

A. Storytelling in Museums

To offer immersive experiences in museums, several technologies are currently to be considered. Audio guides with radio plays and spoken stories are already widely used [2][3] although interactivity is mostly limited to starting and stopping the recording. There are also approaches where the contents are location-dependent [5] or the visitors can select a role from whose perspective the story is told [6]. Films are often used, sometimes presented with methods that increase immersion like the Pepper’s Ghost Effect¹, projection mapping, projections on spheres and dome projection.

Museums try to personalize the visitor’s experience, make it interactive and adapt it to interests and needs. The personalization of the experience during a museum visit is

possible but brings along various technical and content-related problems: in many museums, it not feasible for the exhibit to be viewed by only one person at a time [3]. This raises the problem that everyone sees and hears the same thing. Even if the acoustic information can be perceived independently with the help of headphones or sound domes [7], the obstacle still arises that the playing system must be aware of the identity of the visitor, including their previously chosen role [6], the information already received and the path through the exhibition so far. This data has then to be used to calculate and coordinate the trajectory and distinct information that is presented [8]. However, if the information should also be optically adapted to the role and trajectory of every particular visitor, this can only be done by spatial separation like using separate rooms, booths or head mounted displays.

B. Characters in VR

If a museum uses storytelling to touch visitors emotionally, they have to provide them with believable and convincing protagonists to interact with.

Several researchers found that high behavioral realism and realistic visualization of the characters in a movie increase the believability [9]–[12]. But on the other hand, artists [13] and other researcher stress the problem of the “Uncanny Valley-Effect” [14]. This effect describes the phenomenon that the audience find small discrepancies in the behavior or the realistic visualization of extremely realistic characters very disturbing (often described as “eerie”) [14][15][16] whereas a cartoony stylization is very forgiving and can concentrate on the personality of a character [17][18].

Concerning the creation and digital representation of the virtual actors there are several technical possibilities.

1) Mesh and Bones

The surface of a digital character is stored digitally and, with the help of shaders (instructions how to calculate the color of each pixel) and light simulations, an image is rendered. Game engines are capable of creating very realistic scenes in real-time.

To create such so called “meshes” artists can use Computer Graphics (CG) programs to construct and sculpt the surface. Alternatively, real-life persons can be measured and digitized. These values can be used to reconstruct the person as a mesh. One wide-spread approach to achieve this is by scanning with a laser, infrared (IR) or structured light. Another approach is the

¹ The Pepper’s Ghost effect is an optical illusion that utilizes large pieces of glass to partly reflect a scene which cannot be directly observed by the audience. The viewers see the reflections as “ghosts”.

calculation of the mesh with the help of Photographs and Structure from Motion (SfM) or other photogrammetry techniques.

To animate these meshes usually digital bones are constructed to provide an underlying structure which is utilized to deform the skin. The movement of these bones either can be animated by an animator (this is known as “keyframe animation”) or is acquired from human motion with the help of motion capture techniques.

2) Images and Movies (Billboards)

A second approach is to use 360° movies, often called spherical or 360/180 movies. In this approach, real-life scenes are filmed with a special camera to record a panoramic view. Therefore, creating an interactive application with this technique requires recording all possible versions. If the user makes a decision, the movie has to switch to the corresponding variation.

It is possible to combine spherical movies with 3D-representation. Here the spherical movie is projected onto the background and the characters are placed in 3-D around the virtual camera. Characters can be represented as 3-D-meshes as described above or as 2-dimensional objects showing a movie (this is called a “billboard”). In this case, the shown characters can be life-action footage or stylized representations.

Such films can then be viewed with Head Mounted Displays (HMDs) but, apart from the head rotation, they do not allow much interaction and are accordingly rarely immersive experiences.

3) Volumetric Video

In order to facilitate Free Viewpoint Video (FVV), techniques are developed to store video in a way that the viewer can choose his viewpoint arbitrarily. Volumetric Video allows the recording of people and their movement utilizing the depth information of depth-cameras (“RGBD-camera”) [19][20], photogrammetry or lasers. To store the results, the surface is either represented as a point cloud² [21][22] for each frame of an animation, a dynamic mesh including all movement of the vertices³ or the depth information can be stored separately as an image. In this way it is possible to store animations without using bones and skin deformations [23][24].

C. Story Planning for VR

1) Planning

In the film industry, the description of a film in form of a movie script is highly standardized. To prevent ambiguities and enable effective cooperation scripts are expected to have a special form and structure. Another tool for planning a film, Storyboards, are visualizations of each shot of the movie. Storyboards and the predecessor of the modern movie script (Continuity Scripts) have been used since 1928 [25][26]. With the help of storyboards, one can virtually watch the movie by looking on the images in the right order.

Neither of these methods is applicable for 360° video, VR or interactive storytelling. Since the viewer can choose their

perspective and their position in the scene as well as influence the story new techniques to plan the experience are needed [27][28].

Because of the many different ways a non-linear story can unfold, a system to previsualize an interactive story has to be capable of using variables, conditions and user input. Also it is necessary to deal with the distance to the viewer, their possible positions and what the viewer sees [28].

Interactive story editors like Twine⁴ are designed to create text-based stories. Tools like articy:draft⁵ (Figure 1) are already used to create and organize interactive and non-linear game content. Programs that allow drawing, modeling and animation directly in VR [29][30][31], as well as manipulation of the flow of an interactive application [32] help to position and control visual cues. This prevents situations where the viewers oversee important clues for the understanding of the story.

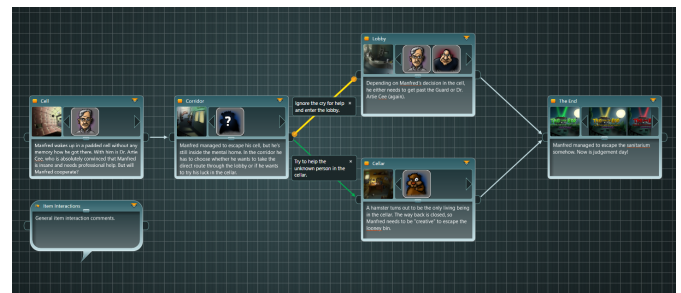


Figure 1: Articy:draft

Although there are tools to support the development of interactive content, real interactive utilities are still missing in order to clearly present and manipulate the complex processes as well as provide visual aids and automatically check for weak points and dead ends.

2) Directing

In traditional movies the director guides the attention of the audience by choosing an appropriate frame and therefore restricts the perceptual boundary. The director can decide upon the position of the camera (Point of View), the focus, movement of camera and actors and so forth. In 360°-movies and VR-applications the viewer himself to a high extent is choosing the point of view. Although there is important research being done concerning the use of focus and depth of field in VR, these possibilities are not yet widely available [33].

In order to increase immersion and to understand the story it is extremely important to make viewers look and move in the right direction at specific times. Therefore the director has to use visual [34] and acoustic [27] cues to help the audience to understand what is happening and to guide them through the story.

If the viewers are able to move freely in the scene, their possibilities should be restricted by obstacles (like walls or fences) in the virtual world. In this case, the visitors have to be provided with methods to cover larger distances. Flying and being moved in the virtual world often leads to nausea because

² A point cloud is a set of points in space

³ A vertex is a point in space used to define a mesh

⁴ <http://twinery.org/2>

⁵ <https://www.nevigo.com/en/articydraft/>

of the discrepancies between real movement and the movement shown in the VR-Application[35][36]. This effect seems to be independent of the amount of immersion [37]. One promising approach to tackle this problem is teleporting. By pointing at the point where one wants to be the viewer instantly changes position [38][39][40]. Sometimes stories need different ways of navigating like a wheelchair[41].

One important help to orient in a virtual world is binaural audio [27]. By manipulating the frequencies of an audio event, the frequency change occurring in the outer ear (Head Related Transfer Function HRTF) can be simulated [42][43]. Therefore the sound seems to be coming not only from the side but also from above, below or behind [44]. The sound sources are placed at the appropriate positions in the virtual space. The sound is manipulated depending on the position of the viewer, the direction of the head and the size and surface of the room and obstacles, and send to the headphones. Thus the user can recognize if a sound event is occurring e.g. behind or above him [45].

With the help of these technologies it is possible to create the basic conditions for the so-called immersion or presence [46].

III. PROOF-OF-CONCEPT

Based on this research we developed a workflow to create interactive stories incorporating volumetric video, spatial sound, interaction triggered by various sensors and events and speech input.

A. Planning

For the first phase –planning– we experimented with different tools like Twine, several script editing Tools like celtX⁶, scrivener⁷ and Trelby⁸. Additionally storyboarding tools like storyboarder⁹ and artcity:draft were evaluated. The highly branched structure[47][48], which is also dependent on conditions and variables has led us to exclude all script-editors and linear storyboarding tools. In future projects we will use Twine, a simple Editor for interactive text stories which is capable of visualizing the structure and the flow of the story (Figure 2). With this software, even variables and conditions can be defined and interactively tested.

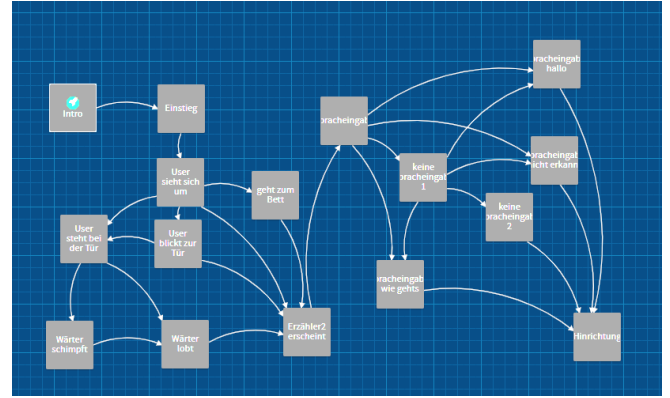


Figure 2: Story structure in Twine

B. Recording Volumetric Video

The method we have chosen to employ for our proof of concept is volumetric video. To allow a simple volumetric representation, we restricted the representation to a 2.5-D representation that spatially visualizes about half of the human body. Our recording software encodes both the depth and color information of a RGBD-Camera as two RGB-images into a movie file. One image containing the color information and one image the depth information are stored side-by-side (Figure 3).

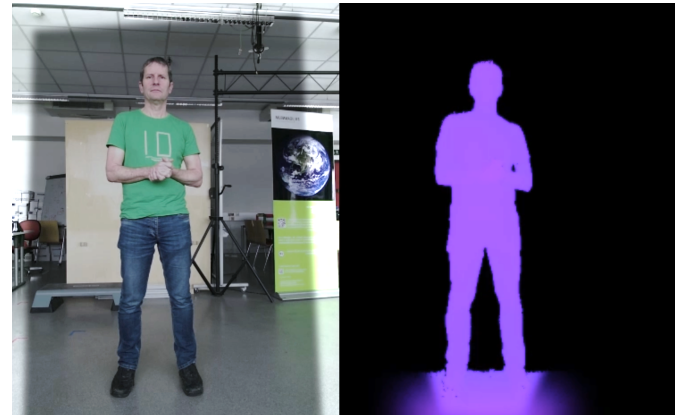


Figure 3: Side by side film

We used the Microsoft Kinect 2, which has 512x424-pixel resolution for the depth information, which was sufficient for our tests. To improve the results, we processed the images with the post-production software Blackmagic Fusion¹⁰. We extracted the background, improved color and contrast, softened the edges and enhanced the depth information.

C. Editor

To add interactivity, we had to develop an editing application that was capable of defining objects, events, sensors and triggers (Figure 4).

⁶ <https://www.celtx.com/index.html>

⁷ <https://www.literatureandlatte.com/scrivener/overview>

⁸ <https://www.trelby.org/>

⁹ <https://wonderunit.com/storyboarder/>

¹⁰ <https://www.blackmagicdesign.com/at/products/fusion/>

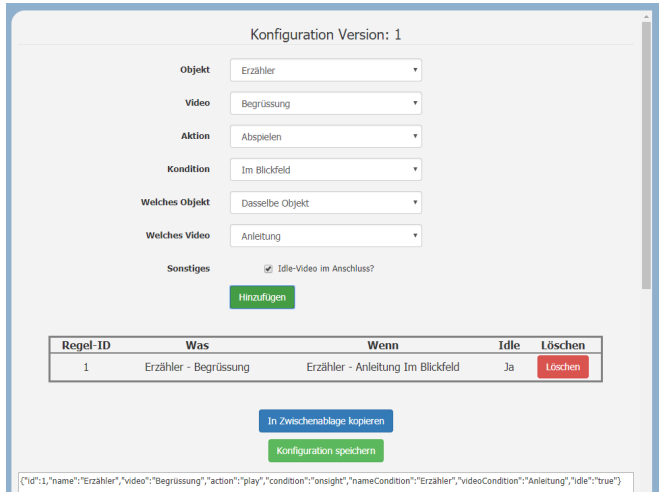


Figure 4: Screenshot of the Editor

Objects can play animations and sound files if they are triggered and all conditions are met. Therefore, features for defining variables, timers and possibilities for user-input had to be provided [49]. With this editor, even persons who are not used to writing code are able to design complex interactive scenes and branched plots. The resulting definitions of scenes can be tested in a generic 3D-environment to detect any gaps, loops or dead ends of the story and to test different paths through the story.

For future applications we will integrate visual editing. For this purpose, screenshots will be created and displayed in the 3D environment. The connections between events will be displayed as lines and the corresponding conditions and variables will be visualized.

We also intend to introduce interactive story design in a VR-environment.

D. Player

Within the game-engine Unity3D¹¹, a mesh is automatically prepared utilizing the previously described video containing the depth information (Figure 5). This is achieved by the help of our vertex-shaders, which refine (“tessellate”) the mesh and displace the vertices according to the brightness of the depth-image stored in the movie. Thus a 3-dimensional object is created and shown in the scene.

These volumetric videos are placed in the appropriate positions according to a file created by the editor. The runtime-module (player) displays, starts and ends the presentation of these videos depending on the variables in the scene description. It also reacts to user inputs to trigger processes and animations. The triggers that are implemented in the application are capable of responding to gaze (is the viewer actually seeing an object), time events (has something already happened), and the end of other animations. One additional experimental trigger was the human voice. The Windows Speech API which is available in Unity3D provides the required phrase recognition capability¹². We implemented the possibility that under certain conditions the user can talk to the character in the scene. The engine searches

for predefined phrases in the spoken text and reacts accordingly, e.g. by starting an animation.

To implement binaural audio in our application we used the Resonance Audio SDK by Google¹³. Resonance Audio allows to apply head-related transfer functions (HRTFs) to the soundfield to enable spatial hearing.

As the space that can be detected by the trackers of the Oculus Rift HMD is limited, the space of the story had to be limited accordingly. This was achieved by creating a scene in a prison cell. The walls of the cell represent the borders of the space, which could be tracked. It would have been possible to put real walls where the users saw walls just as it would have been possible to put real chairs and a bed into the scene. The downside of this approach is that it would require additional effort for calibrating the trackers.



Figure 5: Screenshot of the application

IV. TEST

To assess immersion, reception of volumetric video and of binaural audio we conducted user tests on 10 participants at the ages of 20-26. In a prequestionnaire, the former experiences of the participants concerning VR and interactive applications in museums were surveyed.

In the user test, the participants had to explore an interactive VR-scene. With a HMD (HTC Vive) and headphones including a microphone they were able to experience a scene in a prison cell as a prisoner and interact with other persons in the scene. All actions and reactions of the participants were recorded. With a postquestionnaire, they were questioned about their understanding of the situation and their role in the virtual scene, as well as their emotional involvement and satisfaction.

In the second part of the test, we wanted to test the reception of and reactions to different ways of presenting the volumetric videos. The participants had to rate different versions of visualizations of human characters in the VR-application (Figure 6).

¹¹ <https://unity3d.com/>

¹² <https://developer.microsoft.com/en-us/windows/speech>

¹³ <https://developers.google.com/resonance-audio/>

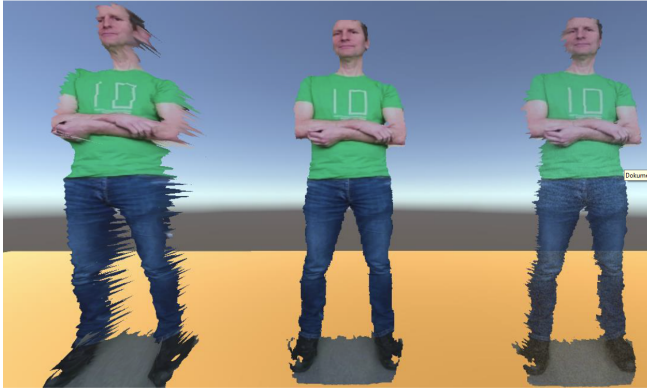


Figure 6: different versions of visualization

V. RESULTS

Ninety percent of the participants were able to identify their role as a prisoner without problems, one participant thought to be an uninvolved observer in the scene. All participants agreed that they were emotionally more involved as they would be in a traditional film. They were able to empathize with the prisoner and understand his feelings. The use of volumetric video was received very positively.

All participants recognized the binaural sound and were positively surprised. They reacted to the acoustic clues and turned exactly into the right direction, often moving toward the perceived source to explore it.

The results of the second part of the test showed that 80% of the participants favored a clear flat representation over more three-dimensional representation. 60% preferred a version where the virtual person automatically faces the viewer ("billboard"). In the last test, the participants were confronted with differently sized characters. All participants reported feeling as intimidated or frightened by the oversized characters. It has to be mentioned that the height of the camera in the scene is dependent on the body height of the user because the sensors of the HMD gauge its exact position and transfer it to the virtual Camera.

One problem, which was not unexpected, was the participants' hesitation to speak loudly and to be seen publicly when wearing a HMD. These problems have to be taken into account when creating such applications for museums. One solution could be to provide small rooms or booths where the visitors can explore the application in private.

We were able to show one possible workflow for creating interactive VR-applications without programming knowledge. In spite of many technical problems that have to be solved to achieve compelling results we were able to realize a small project using volumetric video and an easy to use editor. The first user tests suggest that the approach seems to be very promising concerning the emotional involvement and therefore supporting storytelling in museums.

VI. BIBLIOGRAPHY

- [1] S. Kirchhof, "Storytelling for Volumetric VR," in *Proceeding SIGGRAPH '18 ACM SIGGRAPH 2018 Educator's Forum*.
- [2] L. Fosh, S. Benford, S. Reeves, B. Koleva, and P. Brundell, "See Me, Feel Me, Touch Me, Hear Me," *Proc. SIGCHI Conf. Hum. Factors Comput. Syst. - CHI '13*, p. 149, 2013.
- [3] L. Pujol, M. Roussou, S. Poulou, O. Balet, M. Vayanou, and Y. Ioannidis, "Personalizing interactive digital storytelling in archaeological museums: the CHESS project," *Archaeol. Digit. Era. Pap. from 40th Annu. Conf. Comput. Appl. Quant. Methods Archaeol.*, 2013, unpublished.
- [4] M. Vayanou, A. Katifori, M. Karvounis, V. Kourtis, M. Kyriakidi, M. Roussou, M. Tsangaris, Y. Ioannidis, O. Balet, T. Prados, J. Keil, T. Engelke, and L. Pujol, "Authoring Personalized Interactive Museum Stories," pp. 37–48, 2014.
- [5] J. Alvermann, "Mobile media in the museum space," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct - UbiComp '16*, 2016, pp. 1509–1512.
- [6] S. Benford, G. Giannachi, B. Koleva, and T. Rodden, "From Interaction to Trajectories: Designing Coherent Journeys Through User Experiences," *Chi2009 Proc. 27th Annu. Chi Conf. Hum. Factors Comput. Syst. Vols 1-4*, pp. 709–718, 2009.
- [7] J. Kuutti, J. Leiwo, and R. Sepponen, "Local Control of Audio Environment: A Review of Methods and Applications," *Technologies*, vol. 2, no. 1, pp. 31–53, 2014.
- [8] K. Salo, V. Zinin, M. Bauters, and T. Mikkonen, "Modular Audio Story Platform for Museums," *Proc. 22Nd Int. Conf. Intell. User Interfaces Companion*, pp. 113–116, 2017.
- [9] W. Hai, N. Jain, A. Wydra, N. M. Thalmann, and D. Thalmann, "Increasing the feeling of social presence by incorporating realistic interactions in multi-party VR," *Proc. 31st Int. Conf. Comput. Animat. Soc. Agents - CASA 2018*, pp. 7–10, 2018.
- [10] E. Kokkinara and R. McDonnell, "Animation realism affects perceived character appeal of a self-virtual face," *Proc. 8th ACM SIGGRAPH Conf. Motion Games - SA '15*, pp. 221–226, 2015.
- [11] M. Latoschik, D. Roth, D. Gall, J. Achenbach, T. Waltemate, and M. Botsch, "The Effect of Avatar Realism in Immersive Social Virtual Realities," in *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, 2017.
- [12] L. Freina, R. Bottino, and M. Tavella, "From e-learning to VR-learning: An example of learning in an immersive virtual world," *J. E-Learning Knowl. Soc.*, vol. 12, no. 2, pp. 101–113, 2016.
- [13] R. McDonnell, "Appealing virtual humans," *Motion in Games*, vol. 7660, pp. 102–111, 2012.
- [14] M. Mori, "The Uncanny Valley," 1970. [Online]. Available: <https://spectrum.ieee.org/automaton/robotics/humanoids/the-uncanny-valley>. [Accessed: 08-Nov-2018].
- [15] R. Wages, S. M. Grünvogel, and B. Grützmaier, "How Realistic is Realism? Considerations on the Aesthetics of Computer Games," vol. 3166, pp. 216–225, 2004.
- [16] T. Chaminade, J. Hodgins, and M. Kawato, "Anthropomorphism influences perception of computer-animated characters' actions," *Soc. Cogn. Affect. Neurosci.*, vol. 2, no. 3, pp. 206–216, 2007.
- [17] F. Thomas and O. Johnston, *The Illusion of Life: Disney Animation*. New York, NY, USA: Hyperion, 1995.
- [18] K. F. MacDorman, R. D. Green, C. C. Ho, and C. T. Koch, "Too real for comfort? Uncanny responses to computer generated faces," *Comput. Human Behav.*, vol. 25, no. 3, pp. 695–710, 2009.
- [19] "Microsoft Kinect." [Online]. Available: <https://developer.microsoft.com/en-us/windows/kinect>. [Accessed: 19-Sep-2018].
- [20] "Brekel Pro Pointcloud." [Online]. Available: <https://brekel.com/>. [Accessed: 19-Sep-2018].
- [21] Z. N. Sultani and R. F. Ghani, "Kinect 3D Point Cloud Live Video Streaming," *Procedia Comput. Sci.*, vol. 65, pp. 125–132, Jan. 2015.
- [22] M. Chuang, P. Sweeney, D. Gillett, D. Evseev, H. Hoppe, A. Kirk, S. Sullivan, and D. Calabrese, "High-Quality Streamable Free-Viewpoint Video," *ACM Trans. Graph.*, vol. 34, no. 4, p. 69, 2015.

- [23] “8i.” [Online]. Available: <http://8i.com/>. [Accessed: 19-Sep-2018].
- [24] “Microsoft Mixed reality capture.” [Online]. Available: <https://www.microsoft.com/en-us/mixed-reality/capture-studios>. [Accessed: 19-Aug-2018].
- [25] C. Finch, *The Art of Walt Disney*, 2nd ed. New York, NY, USA: Portland House, 1988.
- [26] J. Hart, *The art of the storyboard: A filmmaker’s introduction*. CRC Press, 2013.
- [27] P. Jayesh, S. I. Azif, and C. Herold, P., “Grammar of VR Storytelling : Visual Cues,” in *VRIC ’17 Proceedings of the Virtual Reality International Conference - Laval Virtual 2017*, 2017, pp. 1–4.
- [28] M. Alger, “Visual Design Methods for Virtual Reality,” 2015. [Online]. Available: http://aperturesciencellc.com/vr/VisualDesignMethodsforVR_MikeAlger.pdf. [Accessed: 08-Nov-2018].
- [29] “gravitysketch.” [Online]. Available: <https://www.gravitysketch.com/>. [Accessed: 20-Sep-2018].
- [30] “AnimVR.” [Online]. Available: <https://nvrmind.io/>. [Accessed: 20-Sep-2018].
- [31] “quill.” [Online]. Available: <https://quill.fb.com/>. [Accessed: 20-Sep-2018].
- [32] “spatial stories.” [Online]. Available: <http://spatialstories.net/>. [Accessed: 20-Sep-2018].
- [33] L. Shi, F.-C. Huang, W. Lopes, W. Matusik, and D. Luebke, “Near-eye light field holographic rendering with spherical waves for wide field of view interactive 3D computer graphics,” *ACM Trans. Graph.*, vol. 36, no. 6, pp. 1–17, 2017.
- [34] L. T. T. Nielsen, M. B. B. Moller, S. D. D. Hartmeyer, T. C. M. C. M. Ljung, N. C. C. Nilsson, R. Nordahl, and S. Serafin, “Missing The Point: An Exploration of How to Guide Users’ Attention During Cinematic Virtual Reality,” *Proc. 22nd ACM Conf. Virtual Real. Softw. Technol.*, pp. 229–232, 2016.
- [35] N. Norouzi, G. Bruder, and G. Welch, “Assessing vignetting as a means to reduce VR sickness during amplified head rotations,” *Proc. 15th ACM Symp. Appl. Percept. - SAP ’18*, pp. 1–8, 2018.
- [36] S. Davis, K. Nesbitt, and E. Nalivaiko, “A Systematic Review of Cybersickness,” in *IE2014 Proceedings of the 2014 Conference on Interactive Entertainment*, 2014.
- [37] S. von Mammen, A. Knote, and S. Edenhofer, “Cyber sick but still having fun,” in *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology - VRST ’16*, 2016, pp. 325–326.
- [38] E. Bozgeyikli, A. Raij, S. Katkoori, and R. Dubey, “Point & Teleport Locomotion Technique for Virtual Reality,” *Proc. 2016 Annu. Symp. Comput. Interact. Play - CHI Play ’16*, pp. 205–216, 2016.
- [39] J. Liu, H. Parekh, M. Al-Zayer, and E. Folmer, “Increasing Walking in VR using Redirected Teleportation,” *Uist 2018*, pp. 521–529, 2018.
- [40] N. Coomer, S. Bullard, W. Clinton, and B. Williams-Sanders, “Evaluating the effects of four VR locomotion methods,” *Proc. 15th ACM Symp. Appl. Percept. - SAP ’18*, pp. 1–8, 2018.
- [41] J. Cho, T.-H. Lee, J. Ogden, A. Stewart, T.-Y. Tsai, J. Chen, and R. Vituccio, “Imago,” *ACM SIGGRAPH 2016 VR Village - SIGGRAPH ’16*, pp. 1–2, 2016.
- [42] Y. Suzuki, D. Brungart, Y. Iwaya, K. Iida, D. Cabrera, and H. Kato, *Principles and applications of spatial hearing*. Singapore: World Scientific Publishing, 2011.
- [43] J. Villegas and M. Cohen, “HRIR~: modulating range in headphone-reproduced spatial audio,” in *Proceedings of the 9th ACM SIGGRAPH Conference on Virtual-Reality Continuum and its Applications in Industry*, 2010, vol. 1, no. 212, pp. 89–94.
- [44] W. Zhang, P. Samarasinghe, H. Chen, and T. Abhayapala, “Surround by Sound: A Review of Spatial Audio Recording and Reproduction,” *Appl. Sci.*, vol. 7, no. 6, p. 532, 2017.
- [45] C. H. Larsen, D. S. Lauritsen, J. J. Larsen, M. Pilgaard, and J. B. Madsen, “Differences in human audio localization performance between a HRTF- and a non-HRTF audio system,” *Proc. 8th Audio Most. Conf. - AM ’13*, pp. 1–8, 2013.
- [46] B. G. Witmer and M. J. Singer, “Measuring Presence in Virtual Environments: A Presence Questionnaire,” *Presence Teleoperators Virtual Environ.*, vol. 7, no. 3, pp. 225–240, Jun. 1998.
- [47] C. Crawford, *On interactive storytelling*, 2nd ed. New Riders, 2012.
- [48] M.-L. Ryan, *Narrative as virtual reality: Immersion and interactivity in literature and electronic media*. Baltimore: The Johns Hopkins University Press, 2003.
- [49] N. Szilas, “Structural Models for Interactive Drama,” *Proc. 2nd Conf. Comput. Semiot. Games New Media (COSIGN 2002)*, pp. 22–27, 2002.