A Virtual Reality Volumetric Music Video: Featuring New Pagans

Gareth W. Young
YoungGa@tcd.ie
V-SENSE
Trinity College Dublin, Dublin, Ireland

Néill O'Dwyer
ODwyerNC@tcd.ie
V-SENSE
Trinity College Dublin, Dublin, Ireland

Aljosa Smolic
SmolicA@tcd.ie
V-SENSE
Trinity College Dublin, Ireland

Figure 1: A volumetric reconstruction of a bass player (texture and mesh)

ABSTRACT

Music videos are short films that integrate songs and imagery produced for artistic and promotional purposes. Modern music videos apply various media capture techniques and creative post-production technologies to provide a myriad of stimulating and artistic approaches to audience entertainment and engagement for viewing across multiple devices. Within this domain, volumetric video capture technologies (Figure 2) have become an emerging means of recording and reproducing musical performances for new audiences to access via traditional 2D screens and emergent extended reality platforms, such as augmented and virtual reality. These 3D digital reproductions of musical performances are captured live and are enhanced to deliver cutting-edge audiovisual entertainment (Figure 1). However, the precise impact of volumetric video in music video entertainment is still in a state of flux.

CCS CONCEPTS

• **Applied computing** → **Media arts**.

KEYWORDS

Virtual Reality, Volumetric Video, Music Video

ACM Reference Format:
Gareth W. Young, Néill O'Dwyer, and Aljosa Smolic. 2018. A Virtual Reality Volumetric Music Video: Featuring New Pagans. In *Proceedings of MMSys '22*. ACM, New York, NY, USA, 3 pages. https://doi.org/https://doi.org/10.1145/3524273.3532907

MMSys '22, 14 - 17 June, 2022, Athlone, Ireland

ACM ISBN 978-1-4503-9283-9/22/06…$15.00
https://doi.org/https://doi.org/10.1145/3524273.3532907

Figure 2: A volumetric video capture (feat. New Pagans)

1 SUMMARY

Building upon our previous creative works with volumetric video (VV) [1, 3, 4, 9], the proposed demonstration will help establish how users respond to VV representations of music performance via virtual reality (VR) technology. As a sophisticated, interactive music video that can potentially be accessed and presented via multiple augmented and virtual reality (AR/VR) platforms, we will demonstrate new workflows on how volumetric music videos may be captured, edited, and accessed for virtual live performance. This approach to contemporary music interactions will demonstrate to the conference delegates how audiences are likely to react to music videos in an VR context and offer insights into how future music video research may be further developed.

2 ARTIST STATEMENT

Finding new ways to visualize and communicate musical performance in VR is driven by artistic creativity, a desire to innovate

Figure 3: Concept art for the presented works

technologically, and a need to capture new and existing audience attention. Moreover, the potential of human-computer interaction (HCI) evaluation methods for researching digital musical interfaces are diverse [7]. It has long been accepted that "Artificial reality is the authentic postmodern condition, and virtual reality its definitive technological expression" [6, p.169]; therefore, it stands to reason that postmodernist art representations can be expressed within VR. Immersive volumetric music videos are being studied at V-SENSE as an emergent art form in and of themselves as many emergent AR/VR technologies are being applied in this endeavor, such as stereoscopic and 360° audiovisual spatial recording technology. These capture technologies have expanded the traditional viewing medium to include further dimensions of immersion, interaction, and imagination for the audience and were closely tied to advancements in home PC GPU/CPU speeds, head-mounted display (HMD) optics, software data processing capabilities, and AI.

The study of extended-reality (XR) music videos has been used to inform V-SENSE's user-centered design of a custom-made VV VR music video experience, featuring the New Pagans' track Lily Yeats (Figure 3 & 4). The project's pilot study initially highlighted the specific qualities that audiences seek during the consumption of such materials [8]. Iterations of this novel application area are expected to focus on differences between traditional media and new XR experiences and expose and build upon existing HCI studies that focus on music and technology in use, specifically those concerning how users experience music videos presented via six degrees of freedom (6DoF) XR technologies.

3 TECHNICAL DESCRIPTION

The authors propose demonstrating the outputs from our ongoing research in the area of *Creative Experiments* (https://v-sense.scss.tcd.ie/creative-experiments/). Thus, a VR installation will be presented that uses (dynamic) VV and (static) computer-generated 3D world-building techniques combined and displayed using the Unity game engine. Figure 5 highlights the specific requirements for viewing.

The volumetric video (VV) capture follows the underlying principles of photogrammetry to record, measure, and interpret images and create a 3D volumetric, textured model for each video frame.

Figure 4: Scenes from within the proposed VR demonstration.

The musicians (New Pagans: Lyndsey McDougall (Vocals) and Cahir O'Doherty (Guitar/Vocals)) were recorded individually performing in the V-SENSE capture studio (Figure 2). The session was captured simultaneously across twelve video cameras. The cameras were placed in each corner of the performance space and recorded against a green-screen background. The individual camera videos were synchronized, comprising the same number of frames, so the gestures captured on one camera matched the same frame number on any other given camera. The audio was recorded separately in a professional music studio and was played back during the VV capture performance, matching the video files' length and playback speed. This capture method simplified the postproduction stage's chroma-keying processes.

After the individual camera clips were trimmed and aligned, the central performance figure was then segmented to separate the dynamic foreground from the static background, and a thresholded silhouette was generated for each frame of the video [5]. The footage was then exported as a series of images and silhouettes that correspond to each frame [2]. This data is then input to a postproduction process that generates dynamic 3D models (i.e. volumetric video) using 3D reconstruction algorithms. The pipeline for generating the VV content is based on a combination of two-volume estimation techniques, multi-source shape-from-silhouette (MS-SfS) and multi-view stereo (MVS) [5]. The resulting data are merged using a bespoke combinational method that produces a series of complete and accurate 3D models. The VV assets were then imported into the Unity game engine using a custom-built SDK that dynamically loads, displays, and executes the textured models (that constitute the VV) at runtime. At the playback stage, the software iterates through these models at the standard video frame

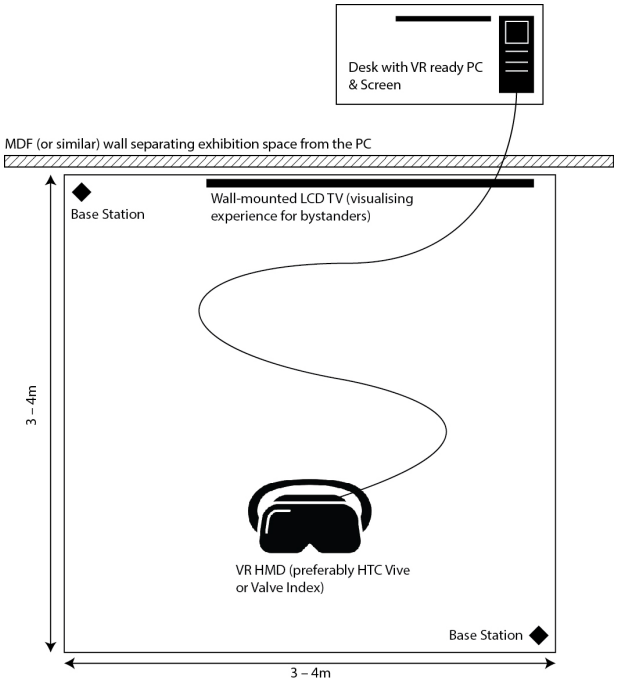

Figure 5: Floor plan (HMD and PC provided by V-SENSE)

rate of 30fps, creating the illusion of movement in the same way as traditional film. Unity is also where the surrounding scene was built, and VR interaction functionality could also be implemented.

The presented VV is closer to a video than computer-generated animation in its current form. Although this technology allows us to capture and display the musicians with high fidelity, it has limitations. Nevertheless, the VR platform demonstrates new forms of music video content that will enable the audience to follow through with a creative narrative that is interactive, immersive, and imaginary. In its current form, the VR experience runs via OpenXR. OpenXR is an open, royalty-free standard for access to multiple VR platforms and devices.

ACKNOWLEDGMENTS

This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under Grant Number 15/RP/2776. The volumetric video content for the music video was generated by Volograms (www.volograms.com). More information about New Pagans and their music can be found at www.newpagans.com.

REFERENCES

[1] Lubna Arielle and Aljosa Smolic. 2020. Bridging the Blue. In *The Expression of Emotion in Humans and Technology*, Ryan Brown and Brian Salisbury (Eds.). ETC Press, Pittsburgh, Chapter 1, 15–27.

[2] Néill O'Dwyer, Nicholas Johnson, Enda Bates, Rafael Pagés, Jan Ondřej, Konstantinos Amplianitis, David Monaghan, and Aljosa Smolic. 2017. Virtual play in free-viewpoint video: Reinterpreting Samuel Beckett for virtual reality. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*. IEEE, Nantes, 262–267.

[3] Néill O'Dwyer, Gareth W. Young, and Aljosa Smolic. 2022. XR Ulysses: addressing the disappointment of cancelled site-specific re-enactments of Joycean literary cultural heritage on Bloomsday. *International Journal of Performance Arts and Digital Media* 0, 0 (2022), 1–19. https://doi.org/10.1080/14794713.2022.2031801 arXiv:https://doi.org/10.1080/14794713.2022.2031801

[4] Néill O'dwyer, Emin Zerman, Gareth W Young, Aljosa Smolic, Siobhán Dunne, and Helen Shenton. 2021. Volumetric Video in Augmented Reality Applications for Museological Narratives: A user study for the Long Room in the Library of Trinity College Dublin. *Journal on Computing and Cultural Heritage (JOCCH)* 14, 2 (2021), 1–20.

[5] Rafael Pagés, Konstantinos Amplianitis, David Monaghan, Jan Ondřej, and Aljosa Smolic. 2018. Affordable content creation for free-viewpoint video and VR/AR applications. *Journal of Visual Communication and Image Representation* 53 (2018), 192–201. https://doi.org/10.1016/j.jvcir.2018.03.012

[6] B. Woolley. 1993. *Virtual worlds: A journey in hype and hyperreality*. Penguin, London.

[7] Gareth W Young and Dave Murphy. 2015. HCI Models for Digital Musical Instruments: Methodologies for Rigorous Testing of Digital Musical Instruments. In *Computer Music Multidisciplinary Research conference*. Springer, Plymouth, 534 – 544.

[8] Gareth W. Young, Néill O'Dwyer, and Aljosa Smolic. 2022. Audience Experiences of a Volumetric Virtual Reality Music Video. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces*. IEEE, Christchurch, 775 – 781.

[9] Gareth W Young, Néill O'Dwyer, Nicholas Johnson, Emin Zerman, and Aljosa Smolic. 2020. Mixed Reality and Volumetric Video in Cultural Heritage: Expert Opinions on Augmented and Virtual Reality. In *International conference on human-computer interaction*. Springer, Copenhagen, 195–214.