Virtual reality and audiovisual experience in the *AudioVirtualizer*

Adinda van 't Klooster

Freelance artist

adinda_18@hotmail.com

Nick Collins

Durham University

nick.collins@durham.ac.uk

ABSTRACT

The *AudioVirtualizer* is an interactive and generative Virtual Reality environment created as an immersive art installation utilizing the HTC Vive Pro and Oculus Rift headsets. This interactive audiovisual system can be controlled live by the voice and by a selection of musical compositions by the authors, though in theory any audio file can be fed to the system. A Unity plug-in was developed to implement realtime machine listening (released open source to accompany this paper) and VR deployment utilized SteamVR. We describe and briefly evaluate the work with respect to the nascent VR literature.

INTRODUCTION

Virtual Reality has great potential for new artworks (Grau 2003; Wilson 2002), including new musical instruments (Ciciliani 2019; Hamilton 2019; Serafin et al. 2016; Serafin et al. 2018; Tatar et al. 2019), enabling new ways to explore the diverse possible mappings between sound and image (Alexander and Collins 2017; Velardo 2019; Weinel 2019). We created an interactive live system that generates graphics in VR in response to sound input, driven either by microphone or from a selection of audio files.

Existing audiovisualisers in VR include *plane9* with 250 different scenes the user can mix, *FantaSynth*, *Chromestasia*, *Raybeam VR*, *CyberDreamVR* inspired by Rave music (Weinel 2019) and *Vision,* STEAM's own music visualizer. There are also musical rhythm games founded in pre-annotated music tracks or beat tracking technology such as *Beat Saber* and *Beat Blocks VR*. None however are responsive to live sound input, with our audio engine tracking low- as well as mid-level features in the sound, allowing more complicated visual returns than previous FFT filterbank based audiovisualisers.

A UNITY PLUGIN FOR MUSICAL MACHINE LISTENING

It is perfectly possible to build inter-connected systems combining Unity and standard computer music environments such as Pd or SuperCollider through Open Sound Control messages, the extOSC Unity asset, or OSC-XR (Johnson 2019). However, such an approach makes standalone building more difficult or impossible, with an application bottleneck in network communication, and possible licensing issues. We preferred to create our own Unity plugin for musical machine listening which could be compiled into an application for distribution, and substantially eased the issue of prototyping onto Oculus Quest in particular, and compiling for Steam or the Oculus Store. Feature extraction code was adapted from SuperCollider machine listening facilities originally written by the second author, alongside a few additional custom C++ implementations. The signal analysis runs with 512 sample hopsize, corresponding (for 44.1KHz sample rate) to around 86Hz and near a typical VR 90Hz visual frame rate.

The list of features comprises spectral centroid, power, spectral irregularity, spectral entropy, sensory dissonance, key clarity, pitch detection via constant Q pitch detector, density of onsets (onsets detected per second), mean Inter Onset Interval (IOI), standard deviation of IOIs, four statistics over the beat histogram (a measure of the strength of different metric periodicities in the music), alongside flags for a detected onset and predicted beat location, and a 'continuous held pitch' derived feature which is larger the longer a consistently held frequency is detected by the pitch tracker. All feature values were normalized to the range [0,1] according to max-min normalization values derived from a large corpus of electronic music. Such normalization assisted being able to easily compare the effect of the different features and map quickly to graphical objects in scenes.

THE AUDIOVIRTUALIZER

The user can explore three different Op Art inspired scenes revealing different mappings between musical input and visual output. In each scene the user can select from a range of different soundtracks, or live microphone input, to investigate how the visuals respond to the sound. The graphics are based on drawings on Skyboxes by artist Adinda van ’t Klooster; these are combined with contrasting live generated 3D graphics that appear in the scene and respond to particular features in the sound. The user can fully navigate each of the 3D scenes to gain different perspectives; clicking on certain ‘gateway sculptures’ allows the user to reach the next level.

Each scene explores a different strategy of responsiveness to the sound input. In the first scene changes in the audio power push spheres with diverse shaders around in a loop; the intensity of this effect changes over the spheres and over time. Detected percussive onsets swap the visible texture on the outside of the sphere and the surface shader is perturbed by power, sensory dissonance (roughness of sound) and irregularity in the spectrum. Black dust particles are also emitted when the system detects an onset. The size of these particles increase with power and the paths of the particles are determined by power and spectral irregularity.

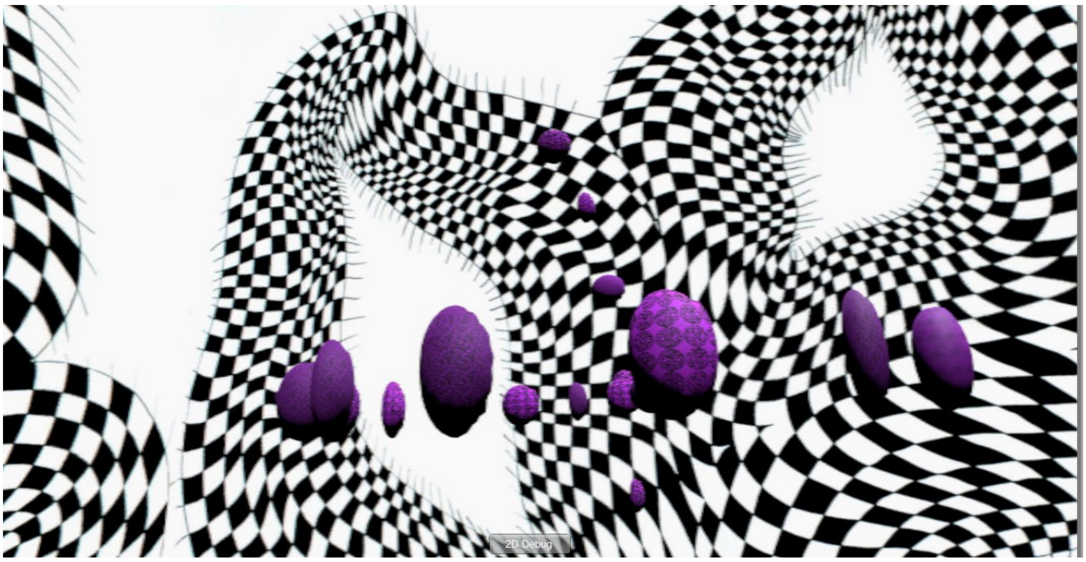

Figure 1. AudioVirtualizer scene 1

In the second scene ambiguous nut/flower-like shapes are generated based on beat detection. They move in a mountainous landscape depending on power and brightness of the sound. They display a flock like behavior and move to a central point if power is sufficient, with a wobble applied to their path depending on brightness of the sound and the density of onsets. This scene works particularly well with microphone input which can be selected by pressing the menu button of the VIVE pro controller.

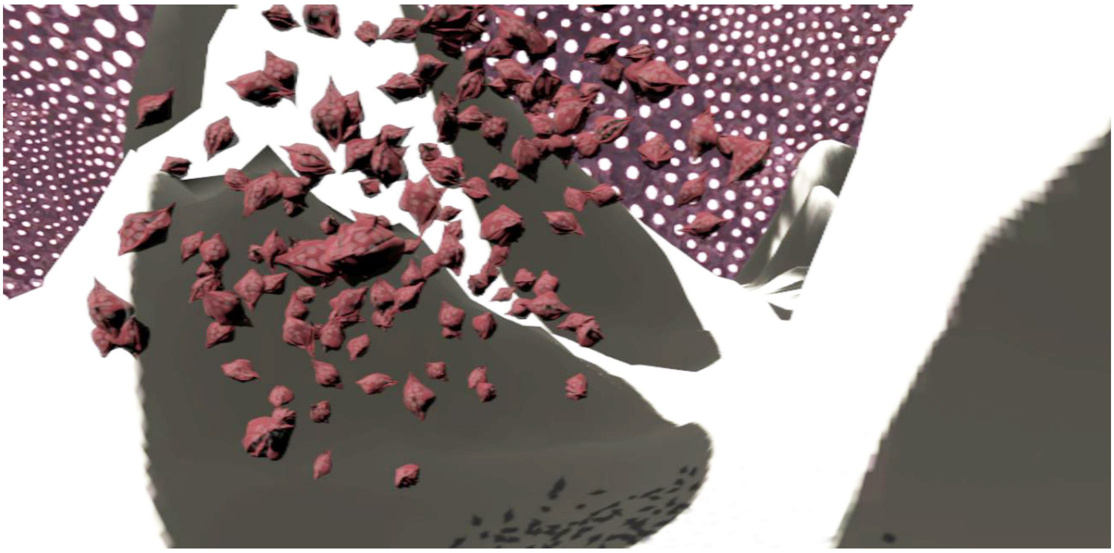

Figure 2. AudioVirtualizer scene 2

The third scene is built around a third Op Art like artwork. White smoke particles are emitted when the system detects sustained pitch, these particles increase with size as pitch is maintained and sensory dissonance increases, sometimes causing the whole scene to go white.

Figure 3. AudioVirtualizer scene 3a

Furthermore, when an onset is detected, ten sculptural shapes are generated starting as a circle and gradually falling apart depending on the brightness of the sound, sustained pitch and spectral irregularity (movement in X is controlled by brightness of the sound, movement in Y by sustained pitch and movement in Z by irregularity).

Figure 3. AudioVirtualizer scene 3b

DISCUSSION

The art installation had an initial public testing at the new tech hub PROTO in October 2019, using versions for HTC Vive Pro and Oculus Rift. The primary finding was that those with previous gaming and VR experience were active navigators, used to and expectant of interactive control, also appreciating the painterly nature of the skybox backdrops as an artistic deviation from video gaming tropes. Those new to VR preferred a more hands-off immersion where they could dwell on looking around them, though this could be due to the learning curve experienced by VR novices in learning to use the controls. Such expectations are therefore likely to change as VR becomes a more available platform for home usage. Most users were happy to spend at least ten minutes using the system and none reported dizziness. As there were people lining up for the experience, most individuals would probably have spent longer if they had engaged with it on their own.
Comparing the headsets, we found the HTC Vive Pro superior as it accommodates existing glasses more comfortably and suffers least from internal sweat being produced after prolonged usage. The Oculus Rift could also run the application, with only some minor coding adjustments. The Unity application is being released as a free art experience together with this paper..

In the future we would like to be able to add more higher level audio features. We would also like to make the system more generative, including creating related skybox graphics on the fly rather than through prepared drawings.

REFERENCES

Alexander, A. and Collins, N. (2017) Live audiovisuals. In Collins, N. and d'Escrivan, J. (eds.) *The Cambridge Companion to Electronic Music* (2nd edition), Cambridge University Press, Cambridge

Ciciliani, M. (2019) ‘Virtual 3D environments as composition and performance spaces’, *Journal of New Music Research* 49,1, 2019, pp. 104-113.

Grau, O. (2003) *Virtual Art: From Illusion to Immersion*, The MIT Press, Cambridge, Massachusetts, p. 196

Hamilton, R. (2019) ‘Collaborative and Competitive Futures for Virtual Reality Music and Sound’, in *IEEE Conference on Virtual Reality and 3D User Interfaces* (VR), pp. 1510-1512.

Johnson, D. (2019) ‘OSC-XR: A Toolkit for Extended Reality Immersive Music Interfaces’ in *Proceedings of Sound and Music Computing*, Malaga 2019

Serafin, S., Erkut, C., Kojs, J., Nilsson, N.C. and Nordahl R. (2016) ‘Virtual Reality Musical Instruments: State of the Art, Design Principles, and Future Directions’ in *Computer Music Journal* 40, 3, pp. 22–40.

Serafin, S., Geronazzo, M., Erkut, C., Nilsson, N.C., and Nordahl, R. (2018) ‘Sonic interactions in virtual reality: state of the art, current challenges, and future directions’ in *IEEE computer graphics and applications* 38, 2, pp. 31-43.

Tatar, K., Prpa, M. and Pasquier, P. (2019) ‘Respire: Virtual Reality Art with Musical Agent Guided by Respiratory Interaction’in *Leonardo Music Journal* 29, pp. 19–24.

Velardo, V. (2019) ‘Making Kantor: Human-machine Collaboration to Build Artistic VR’, Published online Oct 22, 2019, available at: <https://medium.com/the-sound-of-ai/making-kantor-human-machine-collaboration-to-build-artistic-vr-387ceabe59db>

Weinel, J. (2019) ‘Cyberdream VR: Visualizing Rave Music and Vaporwave in Virtual Reality’ In *Proceedings of the 14th International Audio Mostly Conference: A Journey in Sound* (Nottingham, September 2019), ACM, pp. 277-281.

Wilson, S. (2002) *Information Arts: Intersections of Science, Art and Technology*. MIT Press, Cambridge, MA