SpectraScore VR: Networkable virtual reality software tools for real-time composition and performance

Benedict Eris Carey
Hamburg University of Music and Theatre
Harvestehuder Weg, Hamburg, Germany
benedict.carey@hfmt-hamburg.de

ABSTRACT

Keywords: Virtual Reality, Spectralism, Network Music, Real-time Composition

This paper describes a package of modular tools developed for use with virtual reality peripherals to allow for music composition, performance and viewing in 'real-time' across networks within a spectralist paradigm. The central tool is SpectraScore~, a Max/MSP abstraction for analysing audio signals and ranking the resultant partials according to their harmonic pitch class profiles. This data triggers the generation of objects in a virtual world based on the 'topography' of the source sound, which is experienced by network clients via Google Cardboard headsets. They use their movements to trigger audio in various microtonal tunings and incidentally generate scores. These scores are transmitted to performers who improvise music from this notation using Leap Motion Theremins, also in VR space. Finally, the performance is broadcast via a web audio stream, which can be heard by the composer-audience in the initial virtual world. The 'real-time composers' and performers are not required to have any prior knowledge of complex computer systems and interact either using head position tracking, or with a Oculus Rift DK2 and a Leap Motion Camera.

1. DESIGN

Instruments that are designed for music composition through a score interface such as the proposed VRCPS platform [1] and the Quintet.net package [2] are accessible to those who read and write music, yet also incorporate control elements that do not require such skills. These tools, as was the case with Percy Grainger's Free Music Machine [3] and the ANS synthesiser [4] before them, are not just enhanced traditional scores, but more like score/instrument hybrids. This practice, whereby scores are generated in real-time is a type of 'real-time composition' [5]. SpectraScore VR is designed with this method in mind, and optimally so, in terms of Georg Hajdu's 'real-time composition' scenarios as it encompasses unidirectional, bidirectional and multidirectional interaction [5]. It also includes novel interface features that separate the 'real-time composer' from the concept of notated music entirely, replacing this paradigm with a 3D virtual reality environment influenced by gesture, which is used to generate music for live performers controlling Theremin like synthesisers, thereby enabling the audience to become 'real-time composers', in a similar fashion to that used in Kevin Baird's 'No Clergy' [6]. SpectraScore VR is comprised of 3 networkable modules, which forward data to each other over networks via a central server, which synchronises performance and combines this data into a musical composition and broadcasts it back at the audience/composers, and anyone else who wishes to tune in. WebSocket connections are used for larger byte arrays associated with score creation, and the rest is handled via TCP/IP sockets and a Photon Server account (Fig. 1.). These interfaces were designed in Unity 5, Java, Max/MSP and JavaScript but also make use of the MaxScore Package [2], Zsa.descriptors library of externals [7] and the Max/Unity Interoperability Toolkit [8].

Inspired by an early VR experiment, which did not strain one network stretching between the continental United States and Japan, simply 'because only the participants motions needed to be sent, not the entire surface of each person' [9], this system also extends the practice of real-time composition and performance into VR over large distances. As was the case with Tele-immersion [9], or when transferring high-resolution audio and video streams as is done in VOIP, network overheads must be reduced as much as possible for the user experience to be of an acceptable quality. Quintet.net made synchronous performance possible by reducing the overall bandwidth required by only transferring control signals between participants in the performance as OSC messages over TCP/IP connections, much like the early VR experiments mentioned by Lanier.

While a component of this tool is in the same vein as the tradition of gesture based virtual reality instruments such as those developed by Mäki-Patola et al. [10], the point of this system is to synthesise elements of networking, real-time composition, score-creation and virtual reality interaction to reflect on the notions of 'real-time composition', and re-synthesis in order to question how far they can be extended, over distance and time before becoming something else entirely.

2. SPECTRAL ANALYSIS

According to the users specification, an FFT Analysis can be performed on any incoming audio signal, or buffered sound file using the zsa.descriptors objects zsa.freqpeak~ and zsa.energy~. This analysis is routed through a histogram, which reduces the final collection of partials for resynthesis to those that are: A) above a nominated amplitude threshold and B) occur the highest numbers of time in the analysis window relative to all other partials collected over the analysis time (known as HPCP). This collection of partials is thus ranked according to each partials prominence in the analysed spectrum over the user specified time window. The user has access to a number of controls as to how the spectrum is analysed including: the maximum number parts to render, input frequency thresholds (low and high cut-offs), MIDI velocity scaling (min, max values), the FFT window and step size, an input for the large range of MaxScore commands, when to generate and when to output parts, and the dimensions of the master score and parts. Once these parameters are set, the audience can begin their interactive composition session.

3. NETWORK AUDIENCE/COMPOSERS

SpectraScore VR relies on a Photon Server to connect audience members together in a virtual space and synchronise their actions, as it is done in a First Person Shooter video game, while they use smartphones running a Google Cardboard compatible app to experience the performance. A separate server handles score data, and another handles audio streams so as to avoid bottlenecking and offer some redundancy. Each connected audience member's smartphone utilises its local system resources to generate graphics and render incoming audio/data streams as well as transmit control data to the Max/MSP Server (See Fig. 1.).

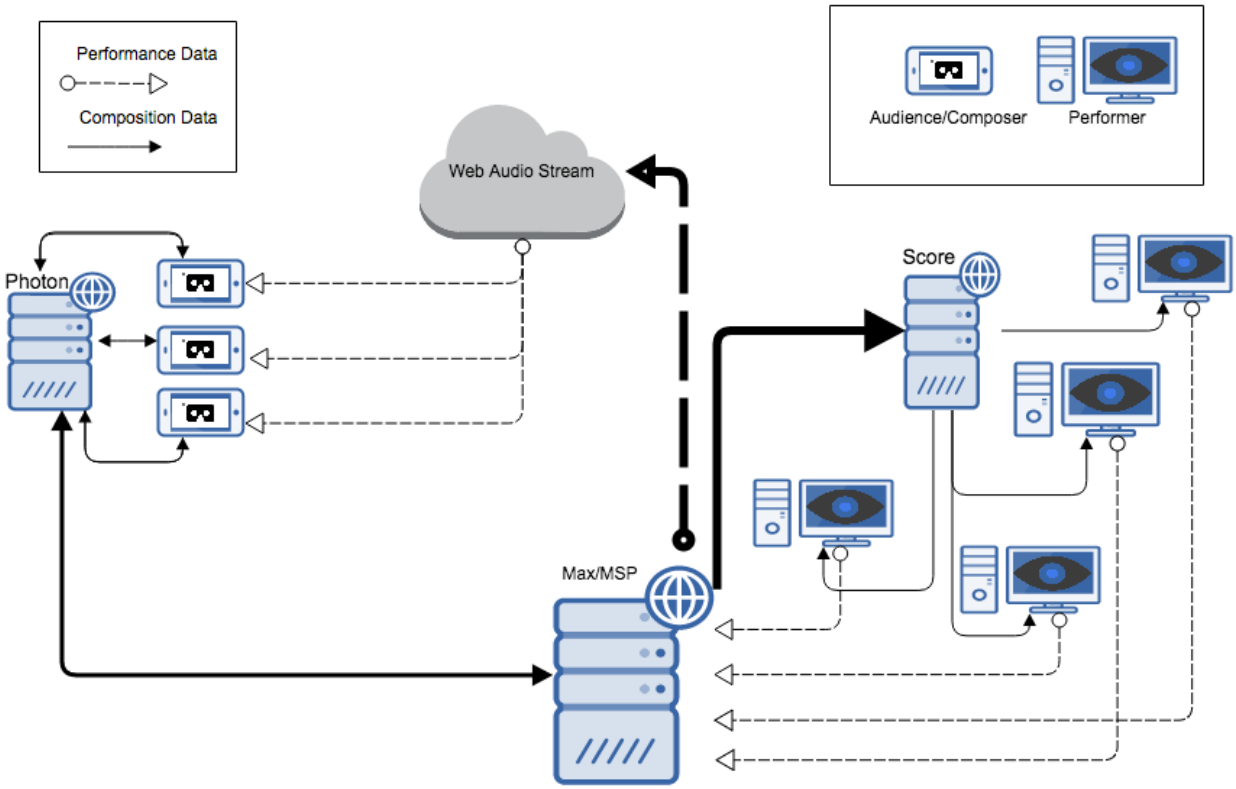

Fig 1. Illustrating the dataflow of a performance with SpectraScore VR. 4 Separate servers synchronise the performance.

The job of producing the final rendered audio is left to the central server running a Max/MSP patch that directs control-data streams and render the composition for broadcast. A web radio cast handles the playback of the composition as it is created. This also allows for users to connect and listen if they do not wish to be visually immersed in the composer/audience space. Communication between Max/MSP and the connected Google Cardboard and DK2 clients is made possible with an edited version of the Max/Unity interoperability toolkit. At present a script is used to instantiate sphere objects in the virtual world once they are detected in the FFT analysis. This technique could also potentially be used to procedurally generate spaces for the participants to perform in.

4. NETWORK PERFORMANCE

A VR interface created using Unity 5 allows the real-time performers to manipulate Leap Motion Theremins, which trigger OSC values over the network. The performers control avatars that are locked to a music stand in the performance space facing each other with their generated scores. Their movements are mapped onto their avatars, collected with a Leap Motion controller. The instrument that they control is a simple chordal Theremin in design with an axis for pitch bend and amplitude. Each finger detected triggers a note from within the spectral analysis collection. This interface is available for OS X and Windows as a VR environment designed for the Oculus Rift DK2.

REFERENCES

[1] Egan, M. P. (2010). Virtual reality composer platform system, Google Patents.

[2] Hajdu, G. (2008). Real-time composition and notation in network music environments. International Computer Music Conference (ICMC).

[3] Grainger, P. A. (1996). "Free music." Leonardo Music Journal: 109-109.

[4] Kreichi, S. (1995). "The ANS synthesizer: Composing on a photoelectronic instrument." Leonardo: 59-62.

[5] Hajdu, G. (2005). "Quintet.net: An environment for composing and performing music on the Internet." Leonardo 38(1): 23-30.

[6] J. Freeman and M. Godfrey. (2010) "Creative Collaboration Between Audiences and Musicians in Flock." Digital Creativity, Vol. 21, No. 2.

[7] Malt, M. and E. Jourdan (2008). "Zsa. Descriptors: a library for real-time descriptors analysis." Sound and Music Computing, Berlin, Germany.

[8] Papastamatiou, N., et al. (2010). Virtual reality applications with user interface for dynamic content development. Proceedings of the 4th WSEAS international Conference on Computer Engineering and Applications.

[9] Lanier, J. (2001). "Virtually there." Scientific American 284(4): 66-75.

[10] M, T., et al. (2005). Experiments with virtual reality instruments. Proceedings of the 2005 conference on New interfaces for musical expression. Vancouver, Canada, National University of Singapore: 11-16.

NIME'16, July 11-15, 2016, Griffith University, Brisbane, Australia.