Inhabiting the Instrument

Mark Durham
Ravensbourne University London
6 Penrose Way
North Greenwich, London
SE10 0EW
mark.durham@rave.ac.uk

ABSTRACT

This study presents an ecosystemic approach to music interaction, through the practice-based development of a mixed reality installation artwork. It fuses a generative, immersive audio composition with augmented reality visualisation, within an architectural space as part of a blended experience. Participants are encouraged to explore and interact with this combination of elements through physical engagement, to then develop an understanding of how the blending of real and virtual space occurs as the installation unfolds. The sonic layer forms a link between the two, as a three-dimensional sound composition. Connections in the system allow for multiple streams of data to run between the layers, which are used for the real-time modulation of parameters. These feedback mechanisms form a complete loop between the participant in real space, soundscape and mixed reality visualisation, providing a participant mediated experience that exists somewhere between creator and observer.

Author Keywords

Installation, Mixed Reality, Spatial, Mapping, Ecosystemic, Ambisonic, Algorithmic Composition, Simulation, Visualization, Interactive.

1. INTRODUCTION

1.1 Acoustic Ecology and Ecosystemic Design

An ecosystem comprises of a set of organisms that exist in a relationship with their environment. These relationships are often complex and causal in nature, with changes to one element affecting the others within the ecosystem. In our case, specific habitats on Earth are our natural environments. Through evolutionary processes we are tuned to operating in these, and our senses and perceptual system are optimised for gathering the most essential information for this purpose. We also apply that same system to any experience that utilises these senses, for example listening to music, watching a film or engaging with a gallery exhibit.

Understanding the impact and role of sound to us as organisms is a critical component of the field of acoustic ecology. This not only forms the basis for techniques of analysis, but also informs a compositional paradigm that can be applied to many areas of artistic practice, including creating media or sound art works [1, 2, 3]. In parallel, this understanding has also been applied to an ecologically driven approach to music analysis and interpretation [4, 5] that focuses on understanding music as a phenomenon, through experience.

In relation to sound composition, the term ecosystemic was initially coined by Agostino Di Scipio in *'Sound is the interface': from interactive to ecosystemic signal processing* [6]. Di Scipio presents the Audible Eco-Systemic Interface (AESI) project, an installation system that is tuned to the site of installation, through audio analysis of the acoustic environment. Site-specific music is then generated to exist within that space, specific to the acoustics of the site. Musick [7] also explored cyber-physical systems that integrate computational algorithms and physical systems (such as motorised speakers), referred to as *sonic spaces*. In terms of process, much of this work is rooted in generative compositional techniques.

NIME'20, July 21-25, 2020, Royal Birmingham Conservatoire, Birmingham City University, Birmingham, United Kingdom.

1.2 Previous Work and Motivation

The immersive approach and sonic aesthetic of this project are heavily influenced by previous work by carried out by the author [8]. In particular, in a project titled - *Multi-channel Sound Design: Instruments for 360 Degree Composition* the author presented a new motion-controlled granular synthesis instrument, with the aim of enabling intuitive control of multiple simultaneous grain streams with gestural control. The design was heavily influenced by concepts central to many NIME discussions, around *control, mapping* and the resulting effect on *expressiveness* of the instrument.

During the development and testing phases, a range of empirical observations were documented by both users of the instrument, and participants in the space. A notable feature of many of these observations was the depth and perceptual complexity of the sound field generated by the process. A proposed explanation for this phenomenon is that the fluctuation patterns in sound output of the instrument are (in terms of timbral change and dispersion) similar to those exhibited by certain natural phenomena, such as rainfall in natural environments, due to the fluctuating timbres and pitches of the impacts [9]. This was also supported by examining ambisonic recordings of such phenomena using the Harpex-X plugin visualizer, where the dispersion pattern was similar across the sound field. An aim of the current study is to explore this dispersion as an applied immersive technique to more traditionally musical sounds, that are part of a combined audio-visual *spatial* system. In this way audio-visual synchronization is surrounding the participant, who is in-turn connected to, or with a further leap *inhabiting* the instrument, and are a causal factor in the output of the instrument, that is itself inherently environmental in how it encompasses the participant.

In relation to the immersive potential, it was also observed that, whilst controlling the instrument, users found benefit in physically moving around the central point between the speakers. A potential explanation is that we locate sound through orienting ourselves towards the sound source [4] while translating the interaural time differences (ITDs) [9]. In this context, the behavior is unsurprising and supports the theory that complex sound fields encourage exploration as a means of gathering environmental information, to further environmental understanding. Furthermore, there is evidence that embodiment itself is a crucial part of perception. For instance, Hall [10], suggests that "perception is formulated through intermingling with the larger world" and that "the body must be an active participant in this world in order for perception to arise". Therefore, real-time interaction and environment (or system) immersion became key drivers in developing a structure and form for this installation, and central to developing a methodology for the study.

This paper presents a system that brings together a set of technologies to explore immersive approaches to generative composition, in an original way. Section 2 will describe the components of the system, how they are connected, and the decisions that were made in designing and developing the work from a compositional standpoint. Sections 3-4 discuss what was achieved through developing the installation, against the research aims.

2. IMPLEMENTATION

2.1 System Design

The installation combines a generative music system, mixed reality visualisation and an ambisonic speaker array. To run and display the mixed reality visualisation to the user, the Microsoft HoloLens head mounted display is used. The participant enters the space of the installation, at which point the visualisation and music system are activated. The physical space inside the circular speaker array is then mixed with the virtual space where the experience unfolds. The visual component of the experience is a real-time evolving particle simulation, as shown in figure 1. Additional footage of the installation is available to view with a stereo sound mixdown [16].

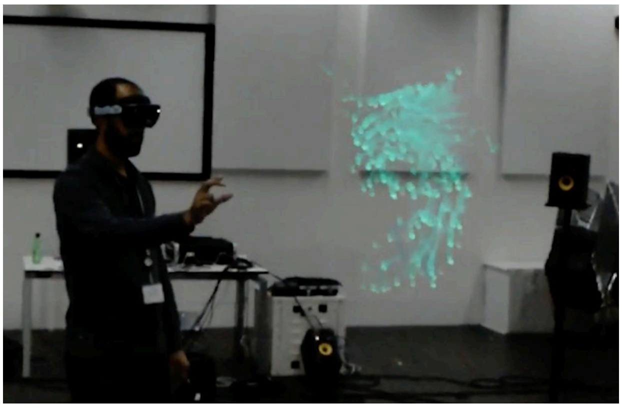

Figure 1. Mixed reality visualization, taken from HoloLens footage [16].

The different elements are designed to form a complete system that, once activated by the participant, is closed but in a state of continuous feedback. It is exactly this approach that forms a system which is ecosystemic in behaviour. Figure 2 displays a map of the system components and connections.

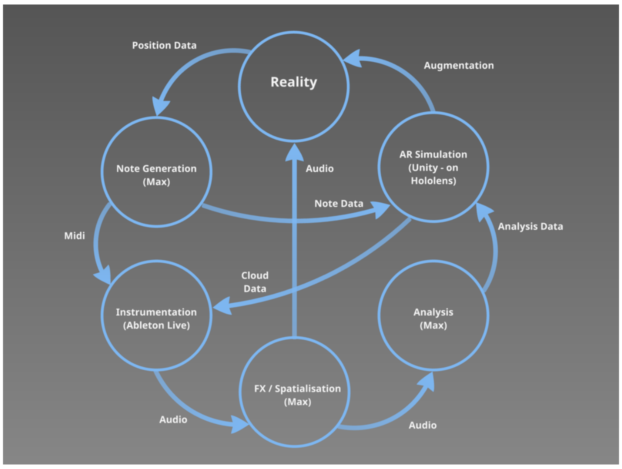

Figure 2. Overview of system and connections.

2.2 Music System

The generative music system comprises of note generation, effects and spatialization components that are created in the software environment Max, and a pair of instances of Collision, a physical modelling synthesizer running in Ableton Live.

Note generation is largely a stochastic process, where notes are being continually generated through a set of weighted probability structures. To link the note generation section to the rest of the system user position is used as a modulator. This takes the form of X,Y data that is generated by the HoloLens based on its position in the space. This positional data affects a probabilistic choice of tempo (60, 90, 120 or 150 bpm), with the likelihood of a faster tempo if the user is towards the center of the simulation, and of a slower tempo when the user position is towards the edges. If the user is outside of the bounds of the simulation (defined by the array of speakers), no notes are triggered.

The system generates note and velocity data for two instruments in Ableton Live. The connection between Max to Ableton is via MIDI. The arrangement shifts between a layered strucure that can produce chords, to an alternating one where only one note per instrument is triggered at a time, resulting in faster melodic sections. Tonally, the instruments most closely resemble struck wooden bars that are close to the sound of a Marimba. Collision data from the particle simulation (average number of collisions per second) is used to modulate the *Material* parameter of the instruments, producing subtle tonal changes within the music that are linked to the level of activity in the particle simulation.

From Ableton Live, the two stereo audio streams of the instruments are sent to the effects and spatialization section of the Max patch. Two effects are used in parallel before the audio is spatialized. The first effect is the *Mutated Texturizer,* a Max for Live device by Timo Rozendahl, that is itself a port of the *Mutable Instruments Clouds* firmware by Émilie Gillet. Sonically, the effect produces an additional granulated layer and probability triggered granular freeze effect. The trigger for this is synchronised to tempo at quarter beat intervals, with a weighted probability trigger. A bespoke buffer-based effect runs in parallel at faster divisions, providing a retrigger or stutter effect, that is also synchronised to the clock in various fixed multiplications. As with many areas of the patch, this effect is controlled by a random probability trigger activated on note divisions. The two interfaces for these effects are shown in Figure 3.

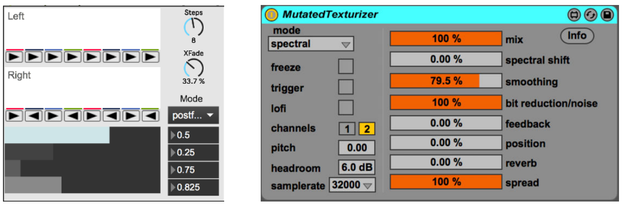

Figure 3. The Mutated Texturizer (right) and bespoke buffer effect (left).

The installation approaches spatialization from a perspective inspired by the idea of sonic world-building, as an additive process. The approach draws from the concept of *Sonic Virtuality* [11], where sounds are placed within the environment by the perceiver, rather than found within the environment by the organism's perceptive system. In this context, the generative system distributes sound from the instruments at both macro and microsound levels of organisation in the composition [12], building the sound field or sonic terrain around the user through spatial placement of sonic material. Through this process, sound fragments are spatially off-loaded to create the environment, but over time they can also be brought back to the participant, in a systemic process that mirrors environmental feedback.

The implementation comprises of an ambisonic panner for each instrument, coupled with a buffer-based sample re-slicer per physical speaker. The ambisonic panning uses the HOALibrary for Max, specifically the object *hoa.2d.map~*, which was selected for the ability to encode position smoothly at signal-rate within the sound field. In control of this are two sine wave oscillators that fluctuate in rate between 0.05hz and 4hz. The oscillators modulate the circular position of each panner as it moves between the speaker outputs, again controlled through a stochastic weighted process. This approach generates some long sections where panning is slowly shifting or close to static, alongside more rapid sections where the rotation is too fast to be tracked by the participant. At these speeds, individual notes become divided between individual speakers forming rapid fluctuations that shimmer like a chorus effect. The effect at speaker level that re-introduces fragments of sound is similar to the buffer-based effect as shown in Figure 3, but working at macro scale from multiple beats up to several bars. These sound slices are stochastically re-ordered and can also be reversed by this process as another means of introducing sonic fluctuations into the system.

2.3 Mixed Reality Visualization

The visual element of the installation was created in collaboration with visual artist Marius Matesan [13] using the Microsoft HoloLens to achieve the mixed reality component of the installation. This choice was governed in-part by the available technology at the time of development, alongside the technical requirements of system. From a broader perspective the HoloLens provided some useful attributes that would enable easy integration into the environment, such as a see-through approach to overlaying graphics into the space, development within the Unity game engine, untethered operation and wireless networking to connect with Max.

Functionally, the visualization is a real-time particle simulation in a three-dimensional space that is overlaid in the room where the installation is located. The centre of the simulation is tethered to the central point between the circle of speakers. As the simulation runs, it can be influenced by forces that repel the particles from the location of the speakers in the space, through the implementation of physics modifiers in the virtual environment. It is also possible for the user to interact with the particle system by using the HoloLens *grab* gesture to introduce a temporary force into the space that attracts the particles.

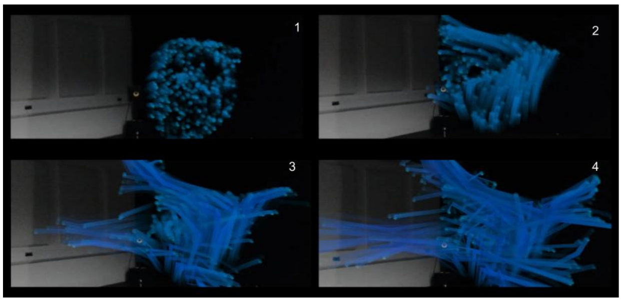

Figure 5. Example 1 of particle development.

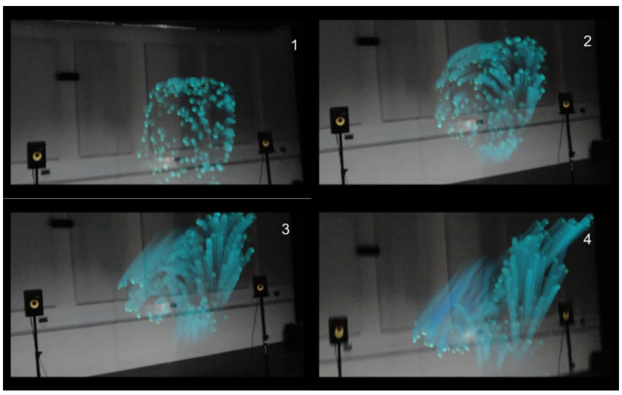

Figure 6. Example 2 of particle development.

2.4 Eco-systemic Modulation

The HoloLens is able to communicate with the computer running Max through a closed Wi-Fi network. Data is exchanged as OSC packets using a bespoke Unity script OSC Flow, seen in Figure 7.

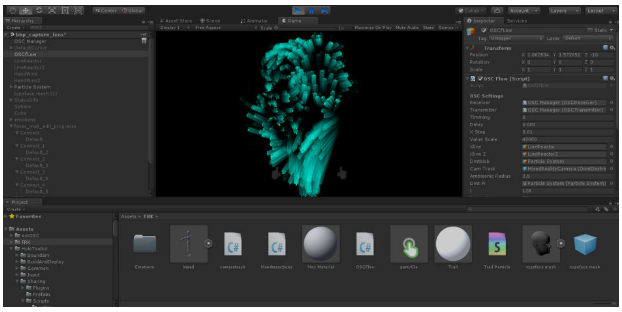

Figure 7. Unity implementation

The bi-directional communication between the computer generating the audio and the HoloLens providing the visuals, is crucial to implementing the experience as a closed system with feedback mechanisms at each level. These mechanisms exist as data streams, or modulators, that run between the components of the system. Expanding on the information presented in Figure 2, the following connections were implemented:

- Position data of the HoloLens in the space is used in determining the tempo of the music system. X,Y data affects a probabilistic choice of tempo (60, 90, 120 or 150 bpm).
- Note data from the note generation part of the Max patch is used to affect the forces in the Unity simulation that attract the particles to the centre of the simulation. As notes are triggered, additional force of attraction is present in this area.
- Collision data from the particle simulation (average number of collisions per second) modulates the *Material* parameter of the Collision instrument in Ableton Live.
- Audio analysis of the direct speaker outputs affects the strength of the forces in the simulation that repel the particles. This part of the system uses the objects zsa.rolloff~ and zsa.slope~ from the ZSA Descriptors Max library [14] to produce the analysis data for Unity. A combination of the data streams from both objects govern whether force is being applied to the particles, and if so, how strong that force is. This is location specific, so the participant will experience forces repelling the particles from the approximate point that sound is emitted at the speakers.

3. ANALYSIS AND FUTURE WORK

3.1 Ecosystemic Design

One of the primary aims of this study was to develop a new immersive composition, that explores soundscape as part of an ecosystemic design. The implementation approach was to link the components of the installation as a feedback mechanism, and to then present music in a way that generates sound as an immersive environment. The resulting composition engages with organisational strategies that exhibit similarities to those found in naturally occurring soundscapes, through the implementation of stochastic processes and a degree of fluctuation throughout many parameters. This is something that Roads would refer to as fuzzy timing [12]. This is inherent in the core macro scale timing system, and continues through to a microsound level in the granular and buffer-based processes that vary the spatial distribution of sound.

3.2 Instrument Interaction and Mapping

Throughout this paper the terms installation and instrument have been used in an interchangeable way as a deliberate means of questioning which is most relevant to this artwork. As an interface, the work is arguably ineffective in providing sufficient feedback to the user on which elements they are in control of, and how these are controlled. However, this is a specific choice, and one that is intended to create an extended period where the user is attempting to make sense of their control of, and position in, the (eco)system. With a more direct 1-1 mapping system, this extended period of questioning would be shorter, along with the perceived complexity of the system. This way, the *learnability* [15] of the system is directly compromised in an attempt to stimulate prolonged engagement from the participant, while they are in the state of trying to understand the interface and potential mappings present.

3.3 Limitations

During the development of the particle simulation we reached the computational power of the onboard computer in the HoloLens. Whilst this was not unexpected, it did limit the potential complexity of the visualization, in terms of particle number and density, and simultaneous connections with external devices. When questioning how immersive the installation is, or how closely it does resemble a natural system, the issue of scale and complexity must be a determining factor. Lastly, with the HoloLens there is a finite number of participants who can simultaneously view the work, currently set at two. This is a definite obstacle to any actual gallery exhibit.

3.4 Future Work

This study is part of an ongoing process that is centered around exploring sound through world-building and ecologically informed design practice. As such, it represents a first step to developing a more detailed system, where the processes of feedback between system components are further refined. As this iteration is now complete, further work in the form of qualitative analysis through exhibition and discussion is to take place, a process is ongoing.

As the current system reached the computational limit of the HoloLens, any further work must either operate within these limits or look for another platform, likely to be another HMD such as HoloLens 2. It is likely that future headsets will be more computationally performant, and will be able to run more complex simulations of this sort, furthering the creative and immersive potential. Similarly, the number of simultaneous users will likely expand, providing more potential for a shared experience between more participants.

4. CONLUSION

A unique mixed-reality artwork that exists between installation and instrument was created. The installation engages the user or participant in an interactive way that encourages exploration and a sense of embodiment within a reactive ecosystemic experience. When viewed as a generative instrument, the system is capable of producing a uniquely spatial soundscape that is informed by principles of soundscape design, and serves as another example of media art engaging with this field. When viewed as an interactive installation, the work explores some new approaches to participant engagement that could be further developed as the technology for augmented and mixed reality experiences matures.

5. ACKNOWLEDGEMENTS

The author wishes to thank Marius Matesan for his collaborative work and support in co-developing this installation, and Ravensbourne University London for their ongoing support with this research.

1. REFERENCES

[1] R. M. Schafer. *The Tuning of the World*. Destiny Books, Rochester, Vermont, (1977).

[2] B. Truax. *Acoustic Communication*. Greenwood Publishing Group, Westport, Connecticut, (2001).

[3] H. Westerkamp. Linking Soundscape Composition and Acoustic Ecology. *Organized Sound. Vol. 7. No. 1.* (2002).

[4] E. Clarke. *Ways of Listening.* Oxford University Press*,* Oxford, (2005).

[5] S. Zargorski-Thomas. *The Musicology of record Production.* Cambridge University Press*,* Cambridge (2014).

[6] A. Di Scipio. 'Sound is the interface': From interactive to ecosystemic signal processing. *Organised Sound, 8*(3). (2003). 269-277. doi:10.1017/S1355771803000244.

[7] M. Musick. Practice-led Research / Research-led Practice Identifying the Theory and Technique of Sonic space Ecosystems. *Ph.D. Thesis*, New York University. 2016.

[8] M. Durham. Multi-Channel Sound Design : Instruments for 360-Degree Composition. *Proceedings of the 12th Art of Record Production Conference : Mono: Stereo: Multi* (pp. 71–88). (Stockholm: Royal College of Music (KMH) & Art of Record Production, 2019).

[9] E. B. Goldstein. *The Encyclopedia of Perception*. SAGE Publications, California, (2010).

[10] J. Hall. Interactive Art and the Action of Behavioural Aesthetics in Embodied Philosophy. *Ph.D. Thesis.* Academic Research and Dissertations. Book 4. (2014)

[11] M. Grimshaw & T. Garner. *Sonic Virtuality*. Oxford University Press, Oxford, (2015).

[12] C. Roads. *Composing Electronic Music, A New Aesthetic*. Oxford University Press, Oxford, (2015).

[13] M. Matesan. Freeklabs. https://freeklabs.net. Accessed 2020-20-01

[14] M. Malt & E. Jordan. ZSA Descriptors: A Library for Real-time Descriptors Analysis. (SMC, Berlin, Germany, 2008).

[15] N. Orio, N. Schnell and M. M. Wanderley. Evaluation of Input Devices for Musical Expression: Borrowing Tools From HCI. (*Computer Music Journal*. 26(3): 62–76, 2002).

[16] M. Durham. Inhabiting the Instrument – Example Footage. *YouTube Video*. (2019) https://youtu.be/Yqb8HL4YDrA. Accessed 2020-20-01