Environment spatial portrait through interactive sound system

Tiange Zhu
Universitat Pompeu Fabra
Music Technology Group
tiange.zhu@upf.edu

Victor Nuñez Tarifa
Universitat Pompeu Fabra
Music Technology Group
victor.nunez02@estudiant.upf.edu

Luis Joglar Ongay
Universitat Pompeu Fabra
Music Technology Group
luis.joglar01@estudiant.upf.edu

ABSTRACT

In this paper, we propose a sound installation implemented with PureData, sonifying the emotional impact one may receive during a confinement time. The sound can be altered in real-time by audiences via their mobile devices, while performers control its spatial distribution and play samples. The paper explains the concept and inspirations of this project in the introduction section. The "design and implementation" section includes technical details of building our project using Pure Data and web programming. In the conclusion section, it discusses evaluation and reflections on our performance. Finally, some potential technical improvements are elaborated in the future work section.

Author Keywords

Computational Music Creativity, Music Installation, Interactive Performance

1. INTRODUCTION

We are living a significant moment in our lives: we are going through the first pandemic of the 21st century. It is affecting many lives globally in different ways and strengths. Curfews and lockdowns are being adopted by many countries in response to the pandemic. The situation that one has to stay at home during most of the day for months, may crucially affect one's mind, leading to overwhelmingness or even claustrophobia. The ways of our communication are changing under these circumstances, as most of us can only communicate and interact with the exterior of their space through internet connection.

Our project aims to transmit the feelings of someone in this state of reclusion. The audience will be immersed in a surrounding sound formed by drones moving in space, while being able to manipulate some characteristics of the sound without knowing for certain what they are manipulating. The audience will also hear samples of certain sounds that may be familiar to the ones heard during the confinement.

Drones are deployed as the basic structural principle in many musical traditions with long histories, such as Indian Rāgas and Middle Eastern bagpipe music. In 1958, La Monte Young introduced it to experimental music, which opened up the possibility of drones in contemporary music composition. We decided to compose drones for our installation, since the constant micro-movements within drone music seemed to be aesthetically suitable for conveying our concept.

A project where the author aims to induce in the audience a similar reaction is PHOBIA [9]. In this artwork, the author intends to induce in the person inside his installation emotions like excitement, discomfort, anxiety, and even paranoia to convey what he calls "digital claustrophobia". He employs a 4ft by 4ft by 8ft room with projections in all 4 walls and a surrounding audio system with repetitive and drone like sounds related to the digital world.

Apart from Maassen's project, a few more NIME papers are related to our work. Tim Shaw et al. in 2015 presented a work where they explored the sound diffusion using the smartphones held by audience members as a diffuse array of loudspeakers [13]. In 2013, Sang Won Lee and Jason Freeman introduced their work "echobo" [8] and instrument for the audience to play simultaneously with their smartphones connected to a network. Jesse Allison et al. created NEXUS [1], a system based on modern web technologies to be able to communicate between user interfaces that smartphones can connect to via website, and sending OSC information to a realtime audio/video processing software.

From the sound perspective, our motivation is to discover how sound alters when modifications are made on different components of the synthesis. We invite audiences to modify the amplitude of each harmonic component during performance. The fundamental frequency of each drone is set to change every five seconds. The shape and length of the envelope is adjustable. We also apply dynamic patching strategy, that harmonics with a new frequency and amplitude can be added in real-time. In addition, we investigate how space interacts with sound. Using pan control to spatially distribute sound with four speakers, we try to introduce the feeling of the sound flowing through the space.

The installation is designed to be open and collaborative. Choosing Performer-System-Audience interaction described by [2], the audience and us interacts with system in real time. While controlling the spatial distribution and play samples during performance, we invite audience engagement, without requiring any previous knowledge on music theory or technology. Even a three-year-old child could perform our installation with sliders using a connected mobile device. The authors themselves no longer obtain the dominant hierarchy during performance, in contrast to the traditional musical performances where the audience usually only watch artists performing on stage.

2. DESIGN AND IMPLEMENTATION

The proposed sound installation is designed using PureData [11], an open source visual programming language for multimedia, together with a website for audience interaction. Implementation comprises three parts: sound synthesis and

Computational Music Creativity'20, June 3-6, 2020, Barcelona, Catalonia, Spain.

mapping, pan control for spatial distribution, and the remote control enabling audience interaction using web sockets.

2.1 Sound Synthesis and Mapping

2.1.1 Sound Synthesis

Additive synthesis is amongst the first audio synthesis techniques widely used in computer music. The term "additive synthesis" refers to sound being formed by adding together sinusoidal components modulated by relatively slowly varying amplitude and frequency envelopes [12].

We utilized this concept in sound synthesis for making drones. In order to enrich the sound generated by additive synthesis, our system comprises three drones. The drone is made of several harmonic partials with different frequency multipliers and adjustable amplitudes. Fundamental frequency of each drone periodically changes in between three randomized digits. The amplitude of each harmonic partial can be adjusted in real-time by audiences with their mobile devices, to achieve the sound effect of a constantly changing drone.

Such a system is implemented with PureData patches. For each harmonic partial, a [harmonics] patch receives a fundamental frequency, a multiplier of the fundamental frequency, and a slider for adjusting amplitude. These numbers are then converted into soundwaves. Each harmonic swells and decays over time due to modulation from an ADSR envelope. A [envelope-wave] patch initializes parameters of the envelope, where its shape and total length are adjustable, in case a different timbre is needed. [envelope-wave] patches are embedded in the [wave] patches. For each drone made of seven harmonics, a [wave] patch contains its volume control and a periodically changing fundamental frequency (Figure 1).

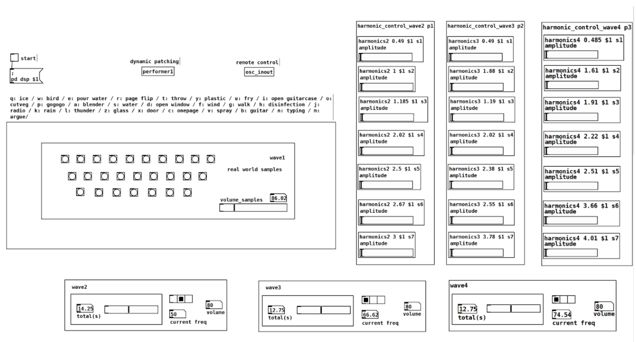

Figure 1: Section of the interface patch for the drones and samples.

We also implemented a patch [performer1] following the concept of dynamic patching. Accompanying every mouth click, a new [harmonics] patch would be generated with randomized frequency and amplitude. Dynamic patching introduces more randomness into the performance, and helps the system become more powerful in terms of sound complexity. The amplitudes sliders of each new [harmonics] patch generated by [performer1] can be played on the PureData interface, albeit due to time limitation, we did not realise its remote control feature. Further possibilities will be elaborated in "future work".

2.1.2 Mapping

Part of the samples we chose are from the online sound database Freesound [7]. The rest are field recordings made by our friend during his quarantine in Wuhan, the city where the covid-19 pandemic originated. From sounds related to human activities such as cutting vegetables or closing a door, to natural environment sounds like a thunderstorm, these samples are curated for the purpose of mimicking the scene of a person's everyday life at home under quarantine.

In the PureData patch, samples are read as arrays. We made a one-to-one mapping from 26 samples to 26 keys of the QWERTY keyboard. We also included [bang] for every sample on the PureData interface. A mouse click or a keystroke would activate an event which plays the corresponding sample for real-time performance. Anyone could play it through the connected keyboard or mouse.

2.2 Spatialization of Sound

A four-speakers configuration is designed for spatialization. The audience is placed in the centre of the room, and the speakers are located in the corners, in order for them to experience the "drones" movement around the space (Figure 2).

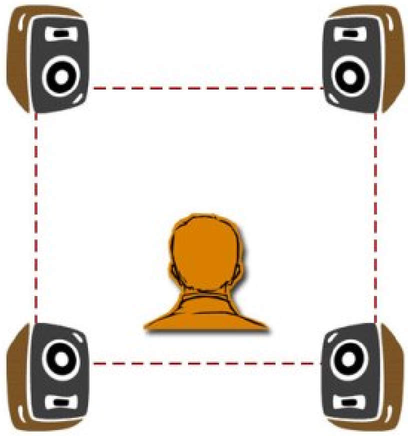

Figure 2: Diagram of the speakers in a room. [3]

The sources are distributed to the different speakers using the else/pan4 object. This object provides an algorithm for quadraphonic panning, that outputs different volume values for the four speakers, according to the desired position in the space. This position is represented in a two dimensional plane with coordinates x and y values, and are defined by us using a midi controller, so we are part of the performance (Figure 3).

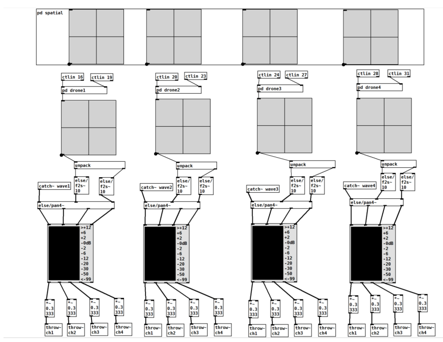

Figure 3: Section of the interface patch for the spatialization.

2.3 Audience Interaction

An important part of this project is the interaction with the audience. The goal of this interaction is twofold: firstly contribute to the surrounding experience like in [13] and [4], and secondly engage the participation of the audience, through some kind of device with WiFi connection, to influence the sound of the performance as [10, 5].

In the Interactive Performance for Ambisonics, Smartphones and Video “Future Perfect”, Garth Paine explores this idea by sending sound to audience smartphones depending on their position in the room and allowing the audience to modify the sound via an XY pad in their screen. Tate Carson, also explores giving the audience the capacity to affect the performance. In this case the audience by how long they listen to a sound or another, what he calls “audience-controlled smartphone speaker array”, influences the main generative algorithm of the performance.

To improve the immersive experience, we took advantage of the loudspeakers of the audience devices. Every device connected to a provided website could reproduce the sound of a drone like the ones built in PD and explained in the section “sound synthesis”, but this time developed using the Web Audio API [14]. The Web Audio API is, as defined by the WC3, “a high level JavaScript API for processing and synthesizing audio in web applications”. Finally the website had a very simple user interface (Figure 4), consisting of one slider. With only this one value all the amplitudes of the harmonics of the drone were controlled by having a mapping table for each harmonic. This mapping table was created by a random function with a softening factor so the consecutive values do not have big jumps.

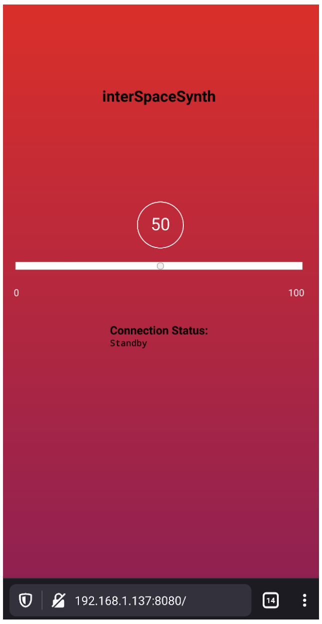

Figure 4: Website user interface.

On the other hand, people could participate in the performance by modifying the amplitude of the harmonics of the drones with their devices. This interaction from the audience was limited to only three participants, each controlling one of the drones in the system. As explained before, the user interface had a slider with a numeric value. This value was sent to the Pure Data patch to control the amplitudes of the harmonics, using the same mapping technique explained above for the drone sounding the phones.

The developed system consists of a simple website using html5 for the layout, javascript for dynamic user interface (the sliders the users can play with) and web audio API for the sound generation. This website is hosted and launched from a node.js server, which will be the center node of the communications. To send information from the node.js to the PureData patch we use the library socket.io on the server side, and the libraries OSC and mrpeach on the PureData side. With these we can create a web socket connection to send the information from the devices up to the PureData patch (Figure 5).

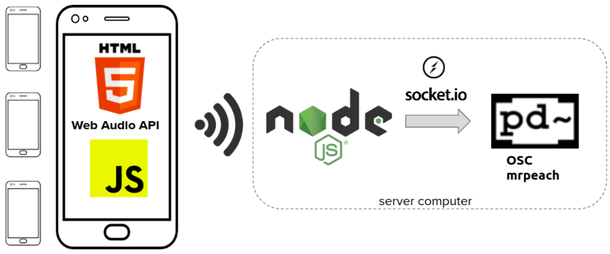

Figure 5: Diagram of the interaction system and the technologies and libraries used.

The Web Audio API construction of the drones consists of the following. First we need to create an instance of AudioContext(), this being the father object of any other Web Audio element. Then to create the drone we create as many oscillators as we need, and assign a frequency to each of them. We also create a gain node for each of the oscillators. As the next step, we build the amplitude mapping tables for each oscillator. Finally we connect the elements following the flow oscillator -> gain -> masterGain -> destination (the destination would be the final step to have sound, like the DAC element in Pure Data). Once everything is created, set and connected we can start the sound.

This example of code shows how to build an oscilloscope:

```
let audioCtx = new AudioContext();
let masterGain = audioCtx.createGain();

let osc1 = audioCtx.createOscillator();
let osc1Gain = audioCtx.createGain();
var table1 = [];

osc1.frequency.value = 100;
while (table1.length < 100) {
    table1.push(softrandom());
}
osc1Gain.gain.value = table1[0];

osc1.connect(osc1Gain);
osc1Gain.connect(masterGain);
masterGain.connect(audioCtx.destination);

osc1.start();
```

Listing 1: Web Audio Code for one oscillator

3. PERFORMANCE

Before the performance, we placed one speaker on each corner of the rectangle space where we performed in. We invited the audience to connect to a private wifi network, and then to a provided url. This information was written in a place where they have access to during the whole performance. When each of them was connected to the website, a standby notification appeared. Once everything was set, the performer sent a ready message to all the connected phones, so the audience could know if they were active.

Thus, the performer and audience could interact throughout the whole performance. The performers controlled the spatialization of the drones and played samples, whilst the audience controlled the drones via their devices.

4. CONCLUSIONS

The performance was an overall success. Several audiences connected their devices to the website we provided, they managed to play with the sliders to modify harmonics amplitudes in real-time. The effect of sound flowing in-between space was obviously perceived, when we controlled spatial distribution via the PureData interface.

From comments we received from a few members of the audience after the performance, they seemed to enjoy the concept and expression of the installment, as well as its sound quality. The performance only lasted a few minutes, and the atmosphere was not fully built up to the extent that audiences could empathize with the feelings aimed to be transmitted. Nonetheless it is still a success for us, as the main goal is just to make them think a bit about the concept we proposed.

However, we believe there are few perspectives that could be improved. For example, we could try to make the experience more interesting and immersive, as well as engage more the audience into interaction of the system. Some ideas are discussed in the following section.

5. FUTURE WORK

There are several aspects that we could improve in the domain of sound synthesis.

First, additive synthesis may lead to a rather thin sound due to its nature: when amplitudes of some harmonics in additive process changes, the complex harmonic structure may be partially destroyed. Other methods such as FM synthesis could be introduced to our system, providing more possibilities on sound.

We directly played samples with keystrokes during performance. In the future, we could explore the samples in the time domain, such as changing the speed when we play it, playing it backwards, or playing a segment of samples. Such methods can twist the emotional expression of recorded samples. These are aesthetically interesting, given their correlation with the current situation: the sense of time we perceived is unusual under confinement.

Finally, enabling the audience to play with more parameters could be a huge improvement for user experience of our installation. For example, we could give audiences the access to samples and spatial distribution. The remote control feature for dynamic patching could also be developed, so that our audience may freely add or reduce harmonics in real-time, to understand sound synthesis in a deeper way.

From the interaction side of things there are also a few aspects that could be improved. The first one is that the loudspeakers of the phones are not very well suited to manage the low frequencies of the implemented drones, so it would be better to change the sound to a range of frequencies that the devices can reproduce properly. Another issue the audience observed is that the interaction was too unnoticeable and that made them disengage from the performance. Maybe a more noticeable interaction would make the audience engage more with the performance, but would that pull them out from the intended experience? That is something to study, and as discussed previously there are a few elements that could be controlled by the audience. Lastly, it would be nice to include more intuitive ways of interacting with the sound other than an interface on the screen, like maybe use the sensors on the phone to capture the movement and use that movement to map the desired parameter. This approach is used by Chien-Wen Cheng in his work [6].

6. REFERENCES

[1] J. T. Allison, Y. Oh, and B. Taylor. Nexus: Collaborative performance for the masses, handling instrument interface distribution through the web. In *NIME*, 2013.

[2] B. Bongers. Physical interfaces in the electronic arts. interaction theory and interfacing techniques for real-time performance. In *Trends in Gestural Control of*, pages 41–70, 2000.

[3] T. Brett. From quadraphonic to good enough sound, 2011.

[4] A. Bundin. Concert for smartphones and orchestra. In J. Monschke, C. Guttandin, N. Schnell, T. Jenkinson, and J. Schaedler, editors, *Proceedings of the International Web Audio Conference*, WAC '18, Berlin, September 2018. TU Berlin.

[5] T. Carson. A more perfect union: Composition with audience-controlled smartphone speaker array and evolutionary computer music. In J. Monschke, C. Guttandin, N. Schnell, T. Jenkinson, and J. Schaedler, editors, *Proceedings of the International Web Audio Conference*, WAC '18, Berlin, September 2018. TU Berlin.

[6] C.-W. Cheng. Interactive audio-visual performance system controlled by smartphone based on arm movements. In *2016 International Conference on Advanced Materials for Science and Engineering (ICAMSE)*, pages 516–519, 2016.

[7] F. Font, G. Roma, and X. Serra. Freesound technical demo. In *ACM International Conference on Multimedia (MM'13)*, pages 411–412, Barcelona, Spain, 21/10/2013 2013. ACM, ACM.

[8] S. W. Lee and J. Freeman. echobo : Audience participation using the mobile music instrument. In *NIME*, 2013.

[9] A. Maassen. Phobia, 2015.

[10] G. Paine. Future perfect, 2018.

[11] M. Puckette. Pure data: another integrated computer music environment. In *in Proceedings, International Computer Music Conference*, pages 37–41, 1996.

[12] J. Smith, S. U. C. for Computer Research in Music, Acoustics, and S. U. D. of Music. *Spectral Audio Signal Processing*. W3K, 2011.

[13] S. T, P. S, and B. J. Fields: An exploration into the use of mobile devices as a medium for sound diffusion. In *Proceedings of the New Interfaces For Musical Expression*, WAC '18, pages 281–284, June 2015.

[14] W3C. Web audio conference w3c candidate recommendation, 2018.