Spatially Accurate Generative Music with AR Drawing

Kyungjin Yoo
University of Maryland
kyoo@umiacs.umd.edu

Eli Schwelling
University of Maryland
eschwell@umd.edu

ABSTRACT

Recent experiments in semi-automatically generating ambient music have yielded emotionally affecting results, leading scientists and musicians alike to develop and experiment with computational systems for creating audible art with varying degrees of success. Most of these systems are based either in analogue technology such as classic tape-reel recording systems or digital systems like virtual synthesizers triggered by a combination of developer-defined values and random number generation. In this paper, I outline the conceptual reasoning behind and development of one such generative music system which uses a simple but versatile virtual synthesizer to generate sound and sequences of repeating randomly generated notes drawn by the user in augmented reality to formulate the patterns and spatial origin of each sound contributing to the entire generative piece.

CCS CONCEPTS

• **Graphics and Visualization → Reconstruction; Mixed / Augmented Reality**.

KEYWORDS

museum, music, education, augmented reality, art, interactive, visualization

ACM Reference Format:
Kyungjin Yoo and Eli Schwelling. 2019. Spatially Accurate Generative Music with AR Drawing. In *25th ACM Symposium on Virtual Reality Software and Technology (VRST '19), November 12–15, 2019, Parramatta, NSW, Australia.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3359996.3365048

1 INTRODUCTION

While my project initially concerned artwork more directly and audio as sort of an example application of the 3D modelling system I designed (i.e. visualizing existing audio using artwork), I soon allocated equal importance to audible art within my work. After discussing the importance of exploring the intersection between art and emotion within the art community with my advisor, I realized I was entering a highly subjective area, which is difficult to navigate; while science is often objective, emotion is as subjective as subjects can be. People often experience similar emotions in day-to-day life, but their emotional responses to artwork (both visual and audible) have the potential to differ completely. This consideration led me slightly away from the completely automated approach of my previous work, which did not give the user input beyond augmenting the depth scale of the model; I realized that to create a tool and system which explored the artistic emotional intersection user input was key. I chose generative music as the perfect medium for examining this relationship. Generative music systems take cues (which can be an infinite number of things, analogue or digital; I soon explain what the cues in my system are) from the user or developer of the system and randomly produces ambient results given those cues. Thus, by allowing the user to make simple choices and using them to influence a mostly random musical piece, I can allow them to create emotional pieces of ambient music which are generated primarily by my computational system but still retains personal traits of their subjective input. The slight amount of subjectivity of their input should make the resultant pieces especially personally affecting.

2 DESIGN AND IMPLEMENTATION

In my generative system, the cues provided by the user are lines they draw in augmented reality. When the user draws a line in the 3D real-world view, the line becomes a note sequencer which contains as many steps as it takes to span the length of the line. This sequencer then determines a random note each step should play at a random velocity, and loops infinitely at a certain user- defined BPM, playing one step at a time. The sequencer's output is filtered through a 3D panning system, which creates the auditory illusion that the spatial origin of its sound is the position of the scanner in the world space view relative to the user's position (i.e. if a sequencer is to the left of the user, the audio it produces will come mostly through the user's left ear). In addition, the sequencer's overall volume is controlled based on the user's position (i.e. as the user gets farther away from the sequencer, its output will get quieter). This spatial audio system is unique to generative music systems created previously; the idea of using the user's real world position to influence the volume and audible origin of note sequencers, to my knowledge, is entirely novel, and I hypothesize that this intuitive nature of deciding where audio comes from relative to the user will make the environment and sonic landscape generated by the application especially immersive. The user also has subjective input regarding the virtual instrument which each sequencer triggers. Depending on the colour a user selects when they draw a line, a different timbre will be produced by the sequencer. The user can access a GUI through which they can tweak parameters on the virtual synthesizer associated with each colour (each colour corresponds with a different synthesizer setting), thus determining what kind of sound each colour produces. By deciding what kind of timbres should correspond to which colours, the user is implicitly creating a map of their personal association between specific colour and sound; and, as previously discussed and supported later in literature review, this relationship is almost certainly emotional in

VRST '19, November 12–15, 2019, Parramatta, NSW, Australia

ACM ISBN 978-1-4503-7001-1/19/11.
https://doi.org/10.1145/3359996.3365048

nature. Therefore, the system can be used to study and compare between individuals how people's associate sound with colour, and how this connection is influenced by emotion. Perhaps the most difficult aspect of building the application was figuring out an effective and intuitive way to create audio sequencers from user drawings and manage the 3D panning and assigned sound generator of these sequencers at runtime. However, augmented reality world space ended up lending itself quite naturally to this purpose. The 3D positioning of each sequencer's audio could just be set to a point corresponding to the location of the sequencer itself in the world space, and the number of notes in each sequence could be determined by the length of the user's drawing in world space. After determining this, development went rather smoothly.

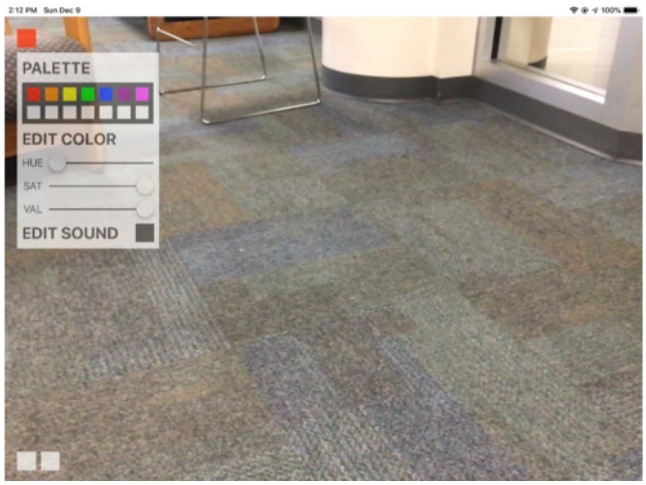

Figure 1: The AR drawing interface with the palette open.

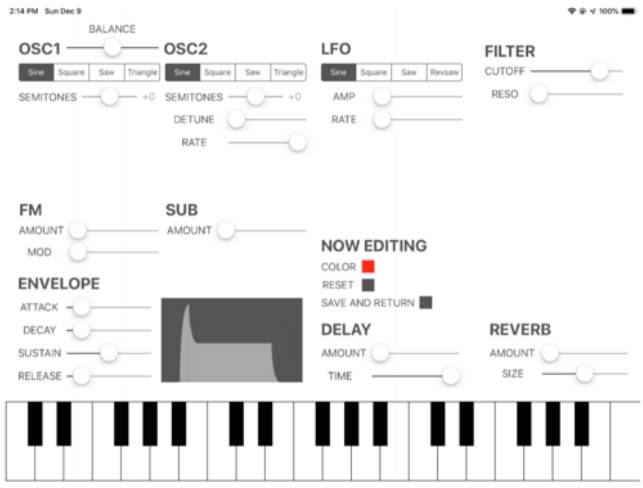

Figure 2: The synthesizer view for editing sounds.

The palette interface serves multiple purposes. Firstly, it stores the 14 available palette cells, which appear in a grid directly under the "Palette" label. The palette cells are instances of a small custom class that extends Swift's native UIButton (see Apple's documentation for further details) but have been modified to store an instance of a synthesizer. As mentioned before, the red button above the palette is used to toggle the palette interface, and changes colour when the user selects a new palette cell to indicate the currently selected cell. The currently selected cell is stored within the view controller for the AR scene as a global variable, so that the selected cell's colour and synthesizer can be accessed and modified from both the augmented reality view and the synthesizer view. In every frame that the scene's renderer is called, the position of the user's device is calculated by extracting the location and orientation components from the point of view's transform and adding them.

When the user holds down the button in the far bottom left and moves their device, if the device is moved beyond a certain distance threshold from its last position, we add a new spherical node to the real-world scene at the device's current position and save the current position as the last position. The distance between the current position and the last position required for a new node to be created is determined by multiplying node size by 1.5. This results in trails of nodes where each node is distinctly visible, as in Figure 2. The colour of the node, as well as the sound the node produces, correspond to the selected palette cell.

As nodes are added, a list keeps track of each one. The first node in a sequence determines its sonic origin point, so when it is drawn, its x, y, and z position in the world space are also stored. When the user taps the button next to the one in the far bottom left, a new instance of a sequence is created with the x, y, and z positions as well as current list of nodes passed into its constructor. Then, the sequencer is added to the global audio manager (which is another class to be explained later), and the drawn sequence is started.

3 RESULTS

The resultant synthesizer produces a very wide variety of sounds and timbres, so I believe most users will be able to formulate sounds they find personally appealing and affecting. In addition, the palette interface allows the user to add custom colors and edit existing colors, further personalizing the cues they send to the generative system. The resultant spatial audio system is relatively convincing and immersive. Again, as explained in the above section, there has been little to no time for conducting user studies, so results are based solely on my personal experience throughout testing. However, I hope that future studies using the application as a tool will prove illuminating.

ACKNOWLEDGMENTS

The authors would like to acknowledge the University of Maryland's FIRE program for the opportunity to pursue this research.