CART 253 - Project 2 Proposal Document
Alex Henri - 40108348
November 11th, 2021

Artistic Vision: Generative Audio Visualizer

For my final project, I wish to create an audio visualizer which uses generative systems to create a dynamic and visually interesting interpretation of sound and/or music.

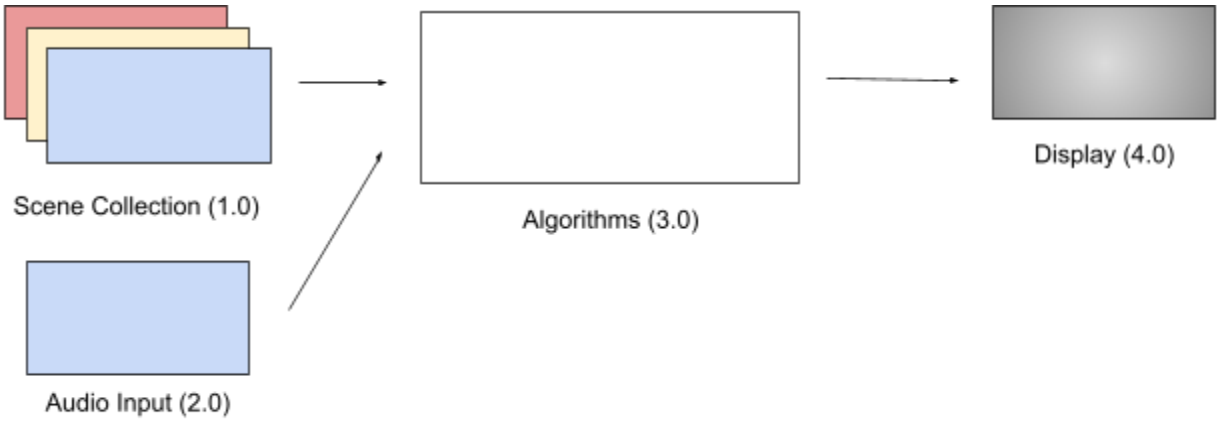
Many years ago, I discovered a digital creative, Raven Kwok. I saw his work in a lyric video for a song called “Greatness”. In the video, there were these crazy shapes and flashing colors and I remember feeling so entranced by it. His work has always felt representative of the music and this is the kind of feeling I want to achieve in my project. More than anything though, I want to challenge myself to make something generative and adaptable to the music.

The goals for my project are:

- Create a program which will use microphone activity (and possibly voice recognition)
- The program will display visual elements that change and “react” to audio activity
- The program’s visuals will be procedurally generated
- The program’s visuals will follow patterns but retain some element of randomness

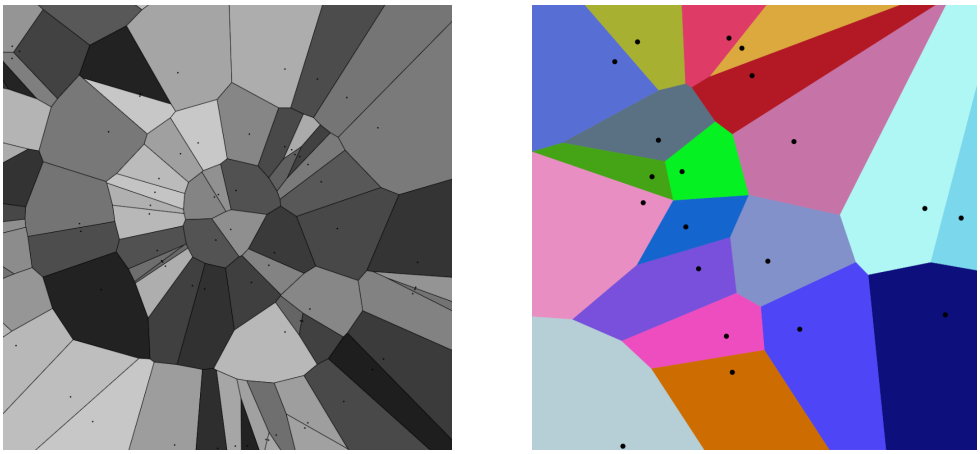
Elements I want in my project are:

- Voronoi Diagrams
- Circle Packing
- Flow fields (time permitting)
- Speech recognition text (time permitting)



I can break my project down into 3 main sectors which all combine into the final product, the display. The first sector is the scene collection (1.0). The scene collection will be composed of “scenes” which are simply mini simulations. These can be swapped between as needed.

Each scene will have a main visual interest, this could be voronoi graphs, circle packing, flow fields, etc. The visuals of the scene will generate automatically and follow a predetermined pattern. It's worth noting that for any visual element, I may decide to apply different patterns, essentially creating its own unique scene.



These two scenes both use voronoi diagrams as the source but they follow different patterns.

The second sector of my project pertains to sound. I wish for the project to be flexible and inclusive in terms of sound input. I will divide sector 2.0 into 2.1: Computer audio (music) and 2.2: Mic activity. I want to be able to use either, both or none of these audio inputs and still be able to display something visually interesting to the screen. This is why my scenes should be self-sufficient in terms of generation and evolution.

Sector 2.1 will use the audio of the user's computer to influence the attributes of the visuals from sector 1.X. I intend to use beat detection, frequency filters and volume to do this. The exact implementations of these will depend on the scene.

Sector 2.2 will use the user's microphone (if they consent to it of course). I intend on using the mic level (volume) to influence the attributes of the visuals from sector 1.X. Perhaps by adding a more disruptive force such as wobbling or distortion. Additionally, if time permits, I might explore using p5.speech to do some basic voice recognition and display what the user says on the screen. Again, if time permits, integrating it into the scene could be a possibility.

The third sector of my project is about the connective tissue. It will cover *how* audio interacts with my scenes. This will likely be the most difficult part of the project yet, the least evidenced (at least in a visual way). Sector 3.0 will have to take the audio inputs from sector 2.0 and apply them to 1.0 to create visual interest. These algorithms will need to be modular because not all audio inputs may be used at any given time and it is also possible that a given scene doesn't use the audio data the same way another scene does.

Technical Challenges

P5.voronoi allows users to draw voronoi diagrams. These diagrams feature a collection of points called “sites”. Each site will be the “seed” from which a circle grows, until it collides with another site’s circle, forming a border. Eventually, the diagram will be filled in, each site being contained within its own region called a cell. The various cells make up the diagram.

The P5.voronoi library allows me to create a diagram of the size that I desire, with the amount of sites that I desire, in the color and stroke that I desire. The challenges I will face are animating the sites’ positions to form visually interesting patterns.

The P5.sound library will serve as a major component in this project. I’ve begun researching how to use the computer’s audio and microphone audio in p5.js. I will also make use of the frequency analyzer functions to get frequency-based feedback from the sound input. This will allow my visual elements to react differently to different frequencies of the audio. Bass, for example, might manipulate the elements in a different way than the higher pitched tones.

The last technical challenge I will face pertains to the patterns and visual elements themselves. I will have to design algorithms which can interact with and alter the parameters of my visual elements. To date, I’ve explored animating the voronoi diagrams by changing the locations of the sites. You can see an example which can be manipulated in my prototype.

If time permits, the p5.speech library will allow me to do basic speech recognition. Being an audio visualizer, I figured it would be fun to make it karaoke-compatible and allow the user to say something and have it represented on screen somehow. I have yet to look into how this works in conjunction with p5.sound.