Talking with Fish: an OpenCV Musical Installation

Gabriel Zalles Ballivian[1] *

UC San Diego
gzalles@ucsd.edu

Abstract. *Talking with Fish* is an interactive multi-media installation in which the movement of fish is translated to musical material using OpenCV. Specifically, the installation makes use of the centroids algorithm to track the position of multiple blobs simultaneously, and the data generated by these blobs is applied to synthesis parameters that generate sound. The X coordinate of each blob provides us with a frequency and position for the voice, the Y coordinate modulates the volume, and the area of each blob modifies the spread of the voice over the loudspeakers. The algorithm is perpetually modulating the musical key, moving clockwise along the circle of fifths, creating a constant harmonic movement for added interest.

1 Introduction

The idea of using fish and a sensing system to create art is not a novel concept. Walker, Kim, and Pends [1] from the Georgia Institute of Technology wrote about this idea in 2007. In this paper, the authors actually noted that sonifying elements of exhibits is not simply an artistic endeavor but actually has deeper implications. The concept the authors were attempting to convey is that visually impaired guests do not have the same experience in museums, or aquariums, as those with 20/20 vision. Therefore, as an approach to increasing inclusion in these spaces, sonifying data becomes imperative for all patrons to derive a meaningful experience. Part of our future goal, therefore, is to explore not just how this tank can be meaningfully sonified, but rather how can we add sound to the entire aquarium, to make a richer experience for visually impaired people[1]

FuXi [2] is another project which featured fish in a real-time music performance system. In contrast to our project, the authors of that work decided to incorporate a MIDI controller into the design, allowing the musician to collaborate with the fish in the music-making process. The authors note how the use of live animals adds indeterminacy to the composition and natural gestures. Rather than using a MIDI controller, our system tracks and reports the musical key while it generates the music. Our vision was that,

* Thanks to Birch Aquarium at Scripps.

[1] This might include a guided audio tour or audio descriptors at our various tanks.

habitually, professional musicians would be invited to the aquarium to accompany the fish in the music-making process[2].

Baldan et al. [3] also discuss a real-time motion-tracking-based aquarium installation in their 2012 paper. The authors use four webcams, the Processing programming language, Pure Data, and Open Sound Control (OSC). The use of Free and Open Source Software (FOSS) makes this project more accessible and is something we would like to consider. Using a single application like MAX/MSP simplifies this process since a single program can be run. Baldan et al. used a number of interesting criteria for music making, such as evaluating the different blobs for color and using this data to influence timbre. They also use blob velocities which is something our system does not currently consider.

This installation, while not entirely novel, stands out due to the fact that it is permanent. Some of the other authors we mention in the literature review intervene in different contexts briefly, in contrast, this project was designed to be viewed by patrons of the aquarium 363 days a year (we close on New Year and Christmas). This means that thousands upon thousands of people will experience this material in one form or another over the course of the year - last year's attendance record for our aquarium was almost half a million people!

2 Technical Design

Talking with Fish relies on a Panasonic AW-HE2 camera mounted across our large kelp tank at Birch Aquarium at Scripps. The video signal is fed to a capture card which lets us use the feed directly in MAX/MSP. Once the video signal is recognized we modify the feed to facilitate blob tracking. To reduce GPU load the frames are resized to a much smaller resolution and these are then converted to a gray-scale before converting to a binary format. In this final conversion, the pixel data is essentially assigned a value of 0 or 1 based on the luminosity of the data point (e.g., bright pixels become white, dark pixels become black).

The centroid algorithm then works by looking for clusters of white pixels. A sizeable mass of grouped white pixels is considered a blob, and these blobs are evaluated moment per moment to determine their position at each frame. Our system currently allows for at most eight voices for the harmonic synthesis section. In order to create a constant sonic environment we built an operator which has a random note duration, above a specified minimum, and an envelope to control its gain. After each operator concludes producing the desired note, it uses the last available frequency value stored for the analysis to play the next chord member.

Before playing the new note, the operator also determines which musical key the system is currently in, to determine the final MIDI value to assign to the voice. As a result of this design the system is often in between musical keys and the harmonic analysis of the music would reveal that it is perpetually modulating around the chromatic scale. In other words, very often, the chords are composed of members of two musical keys, closely related by the number of incidentals. Figure 1 depicts this idea using only

[2] For example during special events.

three notes; notes before the key change are in C major, and notes after the key change are in G Major. At certain moments the two keys are overlaid one on top of the other, making the modulations smoother.

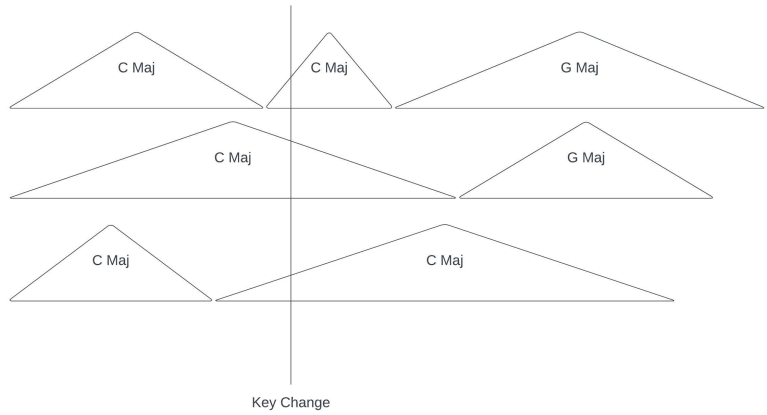

Fig. 1. Diagram showing harmonic structure using only 3 notes.

One of the challenges with this installation was creating something that would be informed by the data from the feed but that would remain active while fish were out of frame. Guests, however, also expected a marked sound to be played when a blob was detected. For this reason, a second instrument was developed digitally to satisfy this criterion. We call this the melodic element because the tones are short and sequenced linearly. The instrument matches the musical key of the chord generator and evaluates the number of columns in our data matrix. The OpenCV algorithm dynamically tracks the number of blobs, thus, whenever the count goes from low to high, we generate a melodic note to signify that a new blob has been detected.

3 Fabrication

The hardware for this project was placed inside a wooden box custom-made to fit the space we needed; the box actually shaped like a parallelepiped due to the topology of the space. Inside the box we hold the computer which runs the MAX/MSP standalone application and a mouse, to start the program each morning. The box also contains the aforementioned video card which turns the Panasonic AW-HE2 HDMI output into a serial stream recognizable by the computer. The audio output of the computer is the built-in headphone jack which connects the computer to a set of PreSonus speakers mounted on the walls of the aquarium. Above the box is a display monitor which shows the original video, the processed video (with green circles around each blob), and information about the piece. A QR code on the GUI links to the project page on the artist's site. Please visit this temporary link for a demo. [3]

4 Future Work

One of the big criticisms of the work currently is that the kelp inside the Kelp Forrest often triggers sound more than the fish does. This is because the OpenCV program we are using does not use a trained model to identify fish, it simply searches for clusters of white pixels[4]. In the next version of this project, I would like to use a database of marine life species (e.g., images) to train a model. Unfortunately, the current hardware we have might not be able to handle this task in real-time, so for now we are sonifying both the fish and kelp. Even more appealing, would be the idea of identifying specific species using machine learning and computer vision. We envision a system that is trained to identify all the various species found in this tank and use a different instrument for each species - for example.

5 Conclusion

This paper has described the installation *Talking with Fish* which generates music material from the video analysis of a kelp forest in San Diego, CA. The MAX/MSP visual programming environment was leveraged to create a custom system that tracks centroids in the field of view. The area and coordinates of these pixel masses are used to drive a live synthesis algorithm that creates harmonic material in all 12 keys of the Western scale sequentially. The resulting program was compiled as a standalone application and copied to a dedicated computer responsible for creating the sound material from the live video feed.

References

1. Walker, Bruce N., Jonathan Kim, and Anandi Pendse. "Musical soundscapes for an accessible aquarium: Bringing dynamic exhibits to the visually impaired." ICMC. 2007.
2. de Londres, Rua, and R. A. E. Macao. "FuXi: a Fish-Driven Instrument for Real-Time Music Performance."
3. Baldan, Stefano, Luca A. Ludovico, and Davide Andrea Mauro. "" Musica Sull'Acqua": A motion tracking based sonification of an aquarium in real time." (2012).

[3] `https://drive.google.com/file/d/17xMwLS4-QyuEB091nsoMBiWkgVycsVFR/view?usp=sharing`

[4] A task made complicated due to the changing lighting conditions of the tank over the course of the day.