MUME 2016 - The Fourth International Workshop on Musical Metacreation, ISBN #978-0-86491-397-5

Audiovisual Synthesis with ABSTRACT/CONCRETE

Victor Shepardson
Dartmouth College
Hanover, NH
victor.r.shepardson.gr@dartmouth.edu

Abstract

ABSTRACT/CONCRETE is presented, a software synthesizer for audiovisual music with a focus on interactivity and generativity. The project is motivated through the lenses of computational creativity and audiovisual music. Its basic principles are described, and some details of its implementation and performance practice are given. The derived work *AVFB#3* is presented as a fixed video piece which illustrates the system's emergent behaviors.

Introduction

ABSTRACT/CONCRETE is an audiovisual synthesizer with an emphasis on generativity and interactivity Shepardson (2016). It couples a high-resolution digital video feedback system to an audio rate multi-agent system, with a human improviser exploring the parameter space.

A/C can be situated in the computational creativity field as one of McCormack's 'creative ecosystems' McCormack (2012), exploring emergent interactions of an ecological or geological kind. Rather than solving creativity as an optimization problem or one of imitating humans, this approach observes creativity which emerges from the interactions of many simple elements. As an organizing principle, ABSTRACT/CONCRETE adopts Hofstadter's notions of 'tangled hierarchy' and 'strange loop' Hofstadter (1989), which are understood to mean feedback between levels of abstraction.

A/C can also be considered as interactive audiovisual music of the kind pioneered by Spiegel with her VAMPIRE software, which modified an early real-time computer music system to also control graphical devices. The result was an interactive musical system where light and sound were algorithmically entangled with each other and a human improviser Spiegel (1998).

ABSTRACT/CONCRETE

ABSTRACT/CONCRETE is built around two basic abstractions: the pixel lattice, and the multi-agent system.

The pixel lattice is the common format of digital images; modern graphics hardware and software make it practical to control large numbers of pixels in parallel at video rates. A generative system built from parallel pixels can simulate millions of elements interactively, can be understood transparently as image or video, and can be portable to very widely available computing platforms. Likewise, the PCM time series is standard for digital audio. Audibilizing emergent patterns as sampled audio can expose periodicity and fine variation across many time scales, and the infrastructure for real-time audio generation is ubiquitous.

ABSTRACT/CONCRETE uses pixels and audio samples as the raw material of a large scale, transparently perceptible, real time, and portable generative system. Complementary modalities are related by a multi-agent system with the speed of audio and spatiality of video: agents moving at audio rate through a video space.

Related Work

One instance of such a relationship between digital sound and image is *wave terrain synthesis* Roads (1995). An audio-rate agent traces some orbit over a two-dimensional terrain, with the elevation of the agent mapped to audio signal. In the formulations summarized by Roads, orbits are fixed or parametric and the focus is on predictable results. James (2005) investigates variations of wave terrain synthesis such as using generative video for terrain and making orbits terrain-dependent James (2005). A terrain-dependent orbit begins to resemble an autonomous agent which navigates by mapping its immediate environment to a behavior.

Agent systems can also act in the space of digital images. Driessens & Verstappen's *e-volver* (2006) uses a heterogeneous population of agents which react to their immediate visual environments, moving through an image and altering it as they go Driessens and Verstappen (2006). ABSTRACT/CONCRETE combines Driessens & Verstappen's drawing agents with James's audio-rate orbits through dynamic environments. Jame's work focuses on predictability and utility as a an instrument, and Driessens & Verstappen's images are constructed over days rather than minutes. A/C is designed to be richly generative (not require very specific intentions on the part of the performer) while enabling a rapid, interactive search of its parameter space.

System Design

A/C agents move through two-dimensional space while observing the environment at their position, i.e., sampling colors from the video. Their movement is influenced by those colors—they make a mapping between color and direction. This can be as simple as interpreting any two color components in some color space C as a vector in some 2D coordinate system S. ABSTRACT/CONCRETE is designed as a platform for exploring various such rules. How does a cylindrical color space differ from rectangular one? What if color gradient replaces color? What if agents incorporate a memory of previous colors or velocities?

Agents draw their paths (leaving snail-trails) to a buffer which can be incorporated in the video feedback environment, allowing them to interact with it and with each other. Since the agents vibrate at audio rates, they can also be listened to as sound: their time-varying properties can be sent directly to the sound card at 48kHz. Those properties include position, velocity, environment color, and combinations thereof. Listening to agents as audio invites the application of classic synthesis techniques to their behavior: if velocity can be interpreted as a stereo audio signal, then e.g. a comb filter can be applied to velocity. Likewise a momentum parameter can make itself audible as a low-pass filter.

This kind of mixing of metaphors is at the heart of ABSTRACT/CONCRETE. In this deterministic, digital system, notions such as 'color,' 'position' and 'direction' are abstractions made over the underlying lists of numbers (themselves abstractions over bit strings, et cetera). So long as numbers representing e.g. colors are treated that way *consistently*, there is a stable hierarchy of representation. When mappings are made between things of different logical types, like color and direction, that hierarchy is tangled in the Hofstadterian sense. Suddenly concepts can't be understood in isolation (yet they *demand* to still be understood via their transduction to light and sound). The relationship between 'color' and 'number,' for example, is perpetually renegotiated as each is meddled with by 'direction.'

Mingling color with space in computer graphics goes back at least to Perlin's texture synthesis technique of spatial distortion of coherent visual noise to produce structured images resembling clouds and marble Perlin (1985). ABSTRACT/CONCRETE applies this technique to pixels in the video environment in much the same way that agents map color to movement. The A/C environment is a patchable video feedback system. The primary video environment with which agents interact can respond to any number of ancillary buffers. The drawing operations relating those buffers are modular and extensible: one can make a temporal low-pass filter, feed it back negatively to form a high pass filter, use it to spatially distort a different buffer, map between color spaces, and so forth.

This multidimensionality is an essential property of ABSTRACT/CONCRETE; a goal is to convey a sense of the process as not just sound and image but the "single integrated supersensory structure" Spiegel (1998) of a many-layered, many dimensional volume. Rapidly alternating between the various layers and representations allows appreciation of the way they interact: relationships between layers begin to animate, giving rise to new appearances and motions.

Implementation

Implementation is in C++ using the openFrameworks[1] library collection. A/C provides a framework for building computational graphs of fragment shaders and automatically exposing parameters to OSC control. It also incorporates an implementation of wave terrain synthesis on a real-time, irregularly sampled video volume. Agent trajectories are stored and and drawn to a texture on the GPU using OpenGL commands, allowing them to move at audio rates while drawing at high resolutions. Conversely, frames of video are downsampled on the GPU before being read back to main memory and assembled into a video volume which can be interpolated at audio rate. The software runs smoothly on a graphics workstation with several 4K resolution video layers and a population of several 48kHz agents. It also runs at lower video resolutions on a laptop with integrated GPU. Source code for the openFrameworks project and requisite addons is available from Github[2].

Demonstration

ABSTRACT/CONCRETE has been developed as both a space for live navigation and a source for fixed material. For a performance, one can first define a patch by editing the configuration files and shader sources which specify the computational graph. Various parameters are available to interactive control by OSC messages. A typical performance is a search for behaviors which will play out themselves or are rich enough to take in for a few minutes, perhaps coupled to a few formally structuring gestures. A simple interface such as mouse and keyboard with Max/Pd-style number boxes works well. Precision and an understanding of how each parameter behaves are more important than dexterity or low latency.

Composition of fixed pieces from gathered material is an opportunity to select the most compelling sounds, images and behaviors. It's also an opportunity to balance and condense a sense of A/C's qualities into a digestible chunk of time. There is a spectrum to be explored between the transparent or didactic and the inscrutable or complex.

AVFB#3[3] an early exemplar of the system in action, recorded live with minimal intervention or post processing. The patch and parameter settings for *AVFB#3* are described below to give a sense of how A/C can be used.

The patch consists of four visual layers: a primary color feedback buffer, a secondary buffer which maintains a low-pass filtered version of the primary buffer, a mapping of the color buffer to monochrome, and an agent buffer visualizing a population of six agents.

The six agents move with constant speed, rotating at each step by an angle related to hue. They tend to trace nested

[1] http://openframeworks.cc/

[2] https://github.com/victor-shepardson/audiovisual-feedback

[3] https://vimeo.com/164777442

curlicue shapes, sounding unstable tones at very high frequencies. The color palette is cyan, magenta, and yellow, drawn on a black field. The agent layer gradually blurs and fades to black such that there is a visual focus on the newest material, a mixing of colors in older material, and a balance of dark ground with saturated figure. The agent layer contributes directly to the time derivative of the primary layer. Color is bipolar: black is negative, medium gray is zero. Because zeros are uncommon in the agent layer, it is always driving change of some kind in the primary layer, which in turn means that agents trace ever-varying paths.

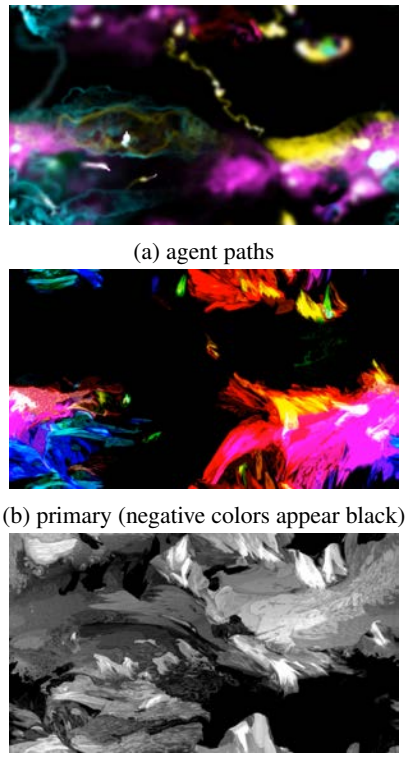
Color flows across the primary buffer in a direction influenced by both hue angle and color gradient. The patch for *AVFB#3* also includes a very low (spatial) frequency, spatially-parametrized component to the primary time derivative which is also spatially distorted by the color of the low-pass filtered layer. This component contributes to large scale structure in the primary layer. The mapping to monochrome complements the color layer by revealing fine structure which is hidden in the blacks and blown-out colors.

The processes at play in *AVFB#3* are initially opaque, the relationship between sound and image unclear. As the piece goes on, forms emerge with a consistent logic inviting comparison to the flow of glaciers. Near the halfway point, flicker between the monochrome, color, and agent layers begins. Soon it becomes apparent that bursts of sound coincide with new color forms in the agent layer, and that elements of the other layers grow from those forms. Eventually, the strobing slows and stops. The viewer gets a good look at the three layers, and can appreciate that two are alternate colorings of the same image. They are left to contemplate that relationship as the view settles on the color layer, and may even discern a relationship between color and direction of flow.

Future Work

Development of ABSTRACT/CONCRETE is ongoing. Future work may include performance improvements to allow larger rates, resolutions and agent populations on cheaper hardware; redesign of communication with the audio thread, to reduce chance of audio glitches; a more robust and intuitive system for patching shaders together; an automatically generated GUI for controlling parameters; a system for saving and loading complete snapshots (the current source code, parameter settings, and system state); and support for live coding fragment shaders and agent behaviors.

There is still much to explore within the video feedback/agent system paradigm of A/C. However, future work may also incorporate different algorithms. Machine vision and listening algorithms which extract abstract features are of particular conceptual interest. For example, a semantic segmentation algorithm might be used to hallucinate forms, allowing regions of the environment to differentiate and evolve separately.



(a) agent paths

(b) primary (negative colors appear black)

(c) monochrome

Figure 1: Strobing layers in *AVFB#3*

References

Driessens, E., and Verstappen, M. 2006. E-volver breeding units. `https://notnot.home.xs4all.nl/E-volverLUMC/machine.html`. Accessed: 2016-05-01.

Hofstadter, D. R. 1989. *Gödel, Escher, Bach: an Eternal Golden Braid*. Random House.

James, S. G. 2005. Developing a Flexible and Expressive Realtime Polyphonic Wave Terrain Synthesis Instrument Based on a Visual and Multidimensional Methodology. Master's thesis, Edith Cowan University, Perth, Western Australia.

McCormack, J. 2012. Creative ecosystems. In *Computers and Creativity*. Springer. 39–60.

Perlin, K. 1985. An image synthesizer. *ACM Siggraph Computer Graphics* 19(3):287–296.

Roads, C. 1995. *The Computer Music Tutorial.* The MIT Press.

Shepardson, V. 2016. ABSTRACT/CONCRETE: an Audiovisual Synthesizer. Master's thesis, Dartmouth College, Hanover, NH.

Spiegel, L. 1998. Graphical groove: Memorial for the vampire, a visual music system. *Organised Sound* 3(3):187–191.