audiocel

sonification of images in instrumental contexts

michael hansen and devin hunt
georgia institute of technology
interactive design

Abstract

This paper lays forth the proposal to construct a software system that will synthesize auditory representations of digital images using a method that mimics synesthesic perceptions. The purpose of this task is to analyze the effectiveness of associating discrete color data to discrete sound maps in auditory representation of predominantly visual data and vice versa.
Using this brute force method of direct pixel to tone conversion we hope to uncover a rudimentary method of image sonification that can be used to create simple synesthesic experiences.

Introduction to Sonification and Major Methods

Sonification is the non-speech audio representation of visual information.[i] A key factor of image sonification is defining a set method to extract a dimension of time from the visual data. Sound has an innate dependency on time and a temporal metric must be extracted from an image if it is ever to be sonified in any perceivable sense. The other major factor of sonification is the geometric characteristics; the overall shape and form of the data represented in the image. When both of these characteristics are taken into account – time and geometry – then a rich sonification can be produced. [ii]

Meijer proposed a well accepted method for sonification. His method used sinusoidal oscillators whose amplitude was mapped to the pixel grey level and frequency to the vertical position of the pixel in question. The image was then scanned from left to right producing chords of the constructed oscillators at fixed lengths in time. This was then played back to produce a sonification of the image.[iii] Meijer's system quickly became the basis for the scanning type method of image sonification.

Synesthesia: Visual Sound

Synesthesia is the neurological linking and mixing of the senses. It has a large basis in the arts, where often color represents sound and vice versa.[iv] Synesthesia is not necessarily limited to just the relationship of color and sound, the other forms bear little importance to the paper at hand. A great study has been made of Wassily Kandinsky's works and his interpretation of color and sound. He perceived specific colors as instruments; their size and form change the tone and volume. Reds are drums, yellows horns and blues the violin. The violent swatches of color seen in Composition VII depict a violent orchestral setting.

Our project attempts to utilize the ideals laid out by synesthesics to create a system that can produce an image map that lends itself to a synesthesic experience. We hope to accomplish this by mapping instruments and tones to color, rather then positional information.

MIDI Overview

MIDI (Musical Instrument Digital Interface) is a set of standards used with various electronic musical devices. It came to be in 1983 in order to satisfy demand for a standardized, modular computer music system. A MIDI system is made up of two components: a controller (the physical instrument) and a computer system. The latter interprets messages sent to different channels by the controller and translates them into the desired sound.

A message is a few bytes of data which is decoded by the MIDI system to produce (or halt) notes at different pitches, volumes, and balances. The two most important messages are "Note On" and "Note Off." Note Off stops all sound playing in a channel. Note On defines a new note to be played, with a note number (the pitch) and a velocity (the volume).

Channels are the virtual outputs through which music is played. MIDI has 16 channels, meaning that 16 different instruments can be playing at one time. Channels have a couple properties that we will be using in this project: the instrument (which determines the quality of the sound) and the pan (which determines how far left or right in stereo speakers notes in this channel play).

A sequence is a set of messages in a particular order which define a song. A sequence has a tempo (measured in BPM) and a length (measured in frames).[v]

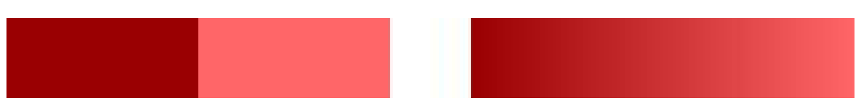
The Proposal

Our project is a software system which will be able to sonify images as well as visualize MIDI music. We will do this by mapping the dimensions of music (pitch, volume, balance, and time) to the dimensions of pixel color (hue, lightness, saturation) and image geometry (size). These metrics have been chosen to better mimic the perceptions of a synesthesic.

As the image will be processed pixel by pixel, we have chosen location to represent time. Time in music and position in an image can be easily represented in a linear manner. When color is deconstructed into hue, saturation and lightness one can quickly draw relationships to sound. Lightness can easily be mapped to the tone of a sound: the darker the lightness the deeper – or darker – the sound. Saturation can be described as the power of the color. This naturally lends itself to the power or loudness of a tone and hence shall be mapped to the volume of a sound. Hue is the position of a color on the optical spectrum and can easily be mapped to the position of a sound in relationship to the listener. In our model, hue will determine the stereophonic position of the note.

Using this system, vibrant and colorful images will yield loud, high pitched sounds, while dark and dull images will yield quiet, low pitched sounds. Images with sharp lines and high contrast will produce sound that jumps around, while smoother images will produce smooth transitions between tones.

Example:



Left: A jump in pitch
Right: A gradual change in pitch

Left: A jump in volume
Right: A gradual shift in volume

Left: A jump in balance
Right: A gradual shift in balance

---

[i] "Sonification." Wikipedia. 2005 <http://en.wikipedia.org/wiki/Sonification>
[ii] Yeo, W and Berger J. "Application of Image Sonification Methods to Music." Center for Computer Research in Music Acoustics. 2005
[iii] Meijer, P. "An experimental system for auditory image representation," *IEEE Transactions on Biomedical Engineering*, vol.39, no.2, pp.112-121, Feb. 1992.
[iv] "Synaesthesia." Wikipedia. 2005 <http://en.wikipedia.org/wiki/Synaesthesia>
[v] The MIDI Specification. 18 October 2005 <http://www.borg.com/~jglatt/tech/midispec.htm>