UDU: A live audio-visual performance

Dr Grayson Cooke, PhD, BA(HONS)
Faculty of Arts, Business, Informatics and Education, CQUniversity Australia
www.youtube.com/civilizedmachine
e-mail: g.cooke@cqu.edu.au

Abstract

Udu is a live audio-visual performance, underpinned by generative art principles and the tradition of "visual music" or "colour music." It uses audio-reaction to produce transformations in abstract visuals. It aims towards the generation of a hybrid or synaesthetic audio-visual "substance," a sense in which audio and visual materials are perceived together and in some elemental relation or commensurability. In this performance, I produce music by playing an Udu (a Nigerian clay-pot drum), which I then sample, loop and process live using Audiomulch. This music then drives abstract, coloured visuals in VVVV, where an FFT analysis captures the amplitude and frequency of the sound source and renders this as a data-stream, which is used to drive visual transformations.

1. Introduction

Udu is a live audio-visual performance. All sounds and images are produced live and on-the-fly, such that the only materials that pre-exist the performance are the "patches" in the software applications used to generate and process the audio and visual materials. In this performance, I produce music by playing an Udu, which I then sample, loop and process live using the patch-based audio software Audiomulch. This music is used to produce abstract, coloured visuals in the graphical-programming application VVVV, which captures the amplitude and frequency of the incoming sound source and converts this signal into numerical data. This audio data, along with the data-streams produced by a series of low-frequency-oscillators, is then used to drive transformations of a series of visual parameters. MIDI controllers are used to switch and tweak parameters within VVVV, and trigger and mix audio processes in Audiomulch.

As an essentially abstract, non-narrative audio-visual experience, *Udu* sits primarily within the tradition of "visual music" or "colour music." Visual music is an artistic tradition concerned with translating musical sounds and structure into visual representations [10]. It is a term used to describe a diverse range of projects and productions, from the 18th century colour organs of Jesuit priest Louis Bertrand Castel to Kandinsky's abstract "Composition" paintings and Len Lye's direct films. All of these projects are in some way concerned with producing visual representations that recall or reflect auditory experience and thus auditory sense perception. Visual

music, then, is concerned with the meeting of audio and visual materials and stimuli, and is frequently associated with synaesthesia, in the sense that it is conceived either to communicate to an audience the synaesthetic experience of the artist, or to produce in the audience a kind of simulated synaesthesia, a sense of what Golan Levin has called "audiovisual commensurability," or what Mitchell Whitelaw calls a "cross-modal binding" of audio and visual media and senses [7,13]. *Udu* seeks to produce precisely this sense of "audiovisual commensurability," by producing sounds and music with a strong rhythmic content, that can be perceived to map directly onto visual parameters.

Another intention with *Udu* is to produce an audio-visual performance that foregrounds the performative aspect of the project. As a result of the complex hardware and software infrastructure that frequently underpins audio-visual performances, these performances can often be opaque as to what, exactly, constitutes the "performance"; the resulting audio-visual media stream, or the performer's bodily manipulation of software and hardware, or some combination of the two? [8]. By using a drum to produce sounds which trigger visual transformations, *Udu* seeks to produce a direct relation between image and sound that is manifest in the actions of a performer and is thus perceivable by an audience, allowing that the performance can be constituted *both* by an actual performer and by an unfolding audio-visual experience produced in real-time.

2. Visual Music

Live audio-visual performance is a complex and multi-faceted field, spanning a wide number of contexts, practices, technologies and traditions. VJs are now commonplace in many club environments, where they provide visual accompaniment to the music of DJs, using hardware and software tools designed to mix and process visual "samples" in the same way a DJ works with audio materials. "Live media" or "live cinema" artists use similar tools, but within a context designated not so much for entertainment and the experience commensurate with a club environment, but for artistic purposes and for the experience of audio-visual media as such. Likewise, many theatre and dance groups have taken up the challenge of using digital media alongside live actors and dancers, so tools for the triggering and processing of audio-visual media are utilized in these contexts as well. Additionally, it is worth noting that many audio-visual performers bridge all these contexts, working as VJs but also producing art installations and performance works, so these contexts are really only worth separating for the sake of convenience.

Perhaps because of this multiplicity of contexts for live audio-visual performance, or perhaps because many of the tools used in such performances are only now becoming widely available, with live audio-visual performance being a relatively "new" form of new-media, scholarship on this field is only now beginning to emerge and to constitute audio-visual performance as a field of enquiry in its own right. Thus it is no surprise that within the existing scholarship on the field, a large and disparate number of antecedents and traditions are mentioned, as possible ways to understand the field. Contexts as diverse as the dioramas and panoramas of the 19th century; magic lantern and phantasmagoria shows of the same period; Wagner's

Gesamtkunstwerk; the Italian Commedia dell'Arte; Wassily Kandinsky's abstract Composition paintings; and Andy Warhol's Exploding Plastic Inevitable, are all cited as potential antecedents for the contemporary practices in live audio-visual performance [4, 5, 6, 11]. These contexts and media forms do not all sit side by side, they are a heterogeneous melange reflective of the many contexts audio-visual performance is found in today, and they evince a number of concerns; concern with *immersing* the spectator in a total visual or sensory field, for example, or a concern with *chance* and *aleatoric production*, or a concern with the commingling of different artistic modes of production and different *modes of sense perception*.

As I have already indicated, it is primarily this final concern – the concern with the commingling of different modes of sense perception – that underpins this project, and this concern is most often associated with the tradition of visual music or colour music. Visual music is a long tradition, stretching at least back to the 17th century and Isaac Newton's equation between the seven tones of the Western diatonic scale and the seven discrete colours red, orange, yellow, green, blue, indigo, violet [3]. Even Aristotle is said to have explored the relations between tonality and colour [10]. The equations involved in relating the seven colours and the seven tones of the scale, however, are essentially arbitrary, and have been questioned many times since their proposition, for various reasons: firstly, hue and frequency are continuums that have been broken into discrete portions by perceptual or mathematical convention rather than by intrinsic or "natural" differentiations; secondly, frequency is a linear continuum while hue is circular [3]; thirdly, colour is most frequently perceived spatially whereas sound is generally experienced temporally [4]. Nevertheless, the relations between hue and frequency postulated by Newton and others began a series of attempts to manifest these relations via the invention of "colour organs" allowing the player to play colour the way they would play a piano. For example, Louis Castel's colour harpsichord or *clavecin occulaire* of 1734 was followed by the colour organs of Bainbridge Bishop and Alexander Rimington in the 19th century, and various 20th century colour-generating devices such as Mary Hallock-Greenewalt's "Sarabet," Thomas Wildred's "Clavilux" and Oskar Fischinger's "Lumigraph" [1, 4, 8, 9].

While some of these devices were conceived to create images that *accompany* music, and some were conceived to make images *in the same way* as music is created, a theoretical relation to sound generation and perception spans each of them. This array of devices represents an attempt to formulate a simulated synaesthesia, whereby it would be possible to "see sound" or "hear colour." Synaesthesia is a neurological condition characterized by the involuntary cross-bleeding of sense perceptions, most frequently between the visual and auditory senses, but sometimes also across touch, taste and smell as well. The synaesthetic experience is consistent across time, and one synaesthete's experience of sensory cross-breeding will generally differ from another's, so for example, while one synaesthete may find a certain musical key to be "red," another may find it to be "golden," and another, indeed, may find it to be "pointy" or to "taste like sherbet." In this context, modern visual music practices inherit not only the practical colour-sound combinations of the colour organs, but the painterly and musical experiments in producing or communicating synaesthetic experience conducted by artists like Wassily Kandinsky and Alexander Scriabin. Of course, with all of these experiments

and media forms, it is not possible to produce the actual neurological *experience* of synaesthesia in non-synaesthetes. What is possible, however, is to produce a sense of some kind of audiovisual commensurability for the audience, that is, produce a tight and integrated mapping of audio and visual media – a "cross-modal binding" – in order that the audience member experiences them together, and perceives their tight integration.

We are, of course, accustomed to visual and audio media being associated in our everyday lives, just as we combine our sight and our vision all the time. For example, it is completely normal to see a sound source and expect to hear the sound generated by the source, for example, a person talking and the sound of their speech. Likewise, we experience this same synchronicity when watching this event on television or in the cinema. Additionally, we watch television and cinema and are accustomed to hearing music that "accompanies" visual media, and in the genre of music video, we see visuals that "accompany" the music. Each of these examples indicate a certain kind of relation between audio and visual media and senses. But, they all tend to privelege one form of media over the other, hence the use of the word "accompanying" to indicate that one form is primary and the other secondary. They also tend not to draw attention to sense perception in and of itself, subordinating sensory perception to the experience of narrative entertainment, or the consumption of pop culture. Visual music, however, tends to be non-narrative, and frequently abstract, meaning that in visual music the "reading" strategies required of the audience for traditional narrative cinema do not hold; there is no over-arching narrative structure to subordinate the image and sound tracks to, and there are no representational images to anchor an audience's perceptions within mimetic realism. Rather, the phenomenological fact of sensory stimulation comes to the fore; audio and visual media are used not as means for the transmission of stories, ideas or ideologies, but as means for sensory stimulation, and frequently, stimulation of multiple senses at the same time and in the same manner.

While visual or colour music is the primary context within which *Udu* has been created, another key consideration has been the role of aleatoric or chance production in the performance. As I have indicated, live audio-visual performances call upon a wide array of potential antecedents and traditions, many of which are invested in the role of chance in artistic production. Influenced by Freud's distinction between the "manifest" and "latent" content of dreams, the Surrealists, for example, were interested in chance-based creative processes as a method of getting at the unconscious or latent content of thought. A game like "The Exquisite Corpse" manifests a version of "thought" which is extra-individual and collective, it is a process of making "meaning" from arbitrary juxtaposition, and moreover, a meaning that originates not in the mind of an individual understood to be able to express their intentions, but across individuals, giving us an image of thought, meaning and intention that cannot be tied to a thinking "subject." The cut-ups of Brian Gysin and William Burroughs likewise operated on the principle of latent content emerging through chance juxtaposition. In more performance-based contexts, the Happenings of Allan Kaprow, many of John Cage's works, and the aleatoric music of Pierre Boulez and Karlheinz Stockhausen, are all potential influences on the role of chance in audio-visual performance, where generative algorithms based on random events are used to produce arbitrary and involuntary, and frequently un-intended, audio and

visual transformations. The notion of *not* being in control, of opening the performance up to chance events, the accident and the arbitrary, is celebrated in many accounts of live audio-visual performance, suggesting that live media performance is premised on a dialectic between control and uncontrol, or a "balance between intent and accident" [2].

3. *Udu* as a Performance

Udu uses Audiomulch and VVVV to pursue the goal of fostering an experience of "audiovisual commensurability" in an audience. Both of these software applications are "patched" environments, meaning that they provide the user with an environment in which to develop their own audio or visual program or "patch" by stringing together "nodes" that perform certain tasks.

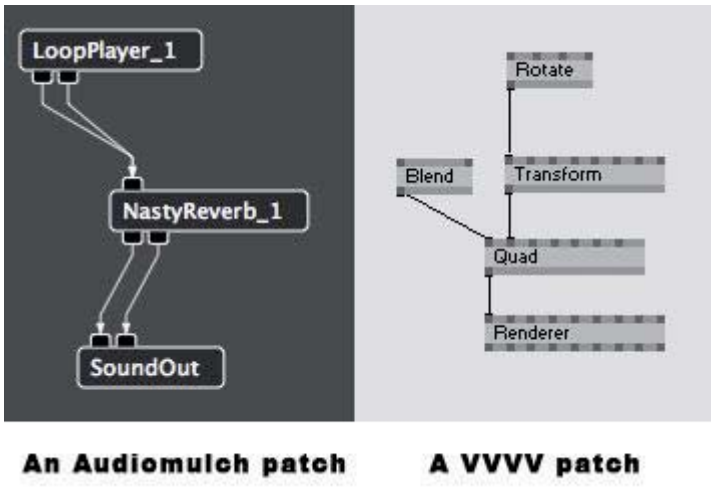

Figure 1: patched environments

Patched environments use nodes that either *generate* some kind of signal (e.g., a low-frequency oscillator, a frequency analyser, a file player), that *process* some kind of signal (e.g., a reverb unit, a 3D transformation node, a blend node), or that *output* a signal (e.g., a sound-out node, a visual renderer).

In Audiomulch, my Udu playing is firstly simply amplified and passed through the software, and later, it is recorded using a loop node. This recording is then used as raw "signal" from which a number of different "channels" of audio are generated, as the looped audio recording is sent through a series of different effect chains. For example, one chain consists of a granular synthesizer, which reduces the audio signal to a series of rhythmic "grains" or sound fragments, and a series of delay effects, ultimately producing a rhythmic and semi-random pulsing. Other effect chains use granular synthesis and VST plugins designed to "glitch" and re-cut the audio signal randomly, again producing a random series of rhythmic pulses. Other nodes are used to impart a harmonic tonality to the signal, to add phasing to it or to apply equalisation to the frequencies. All the rhythmic effects in Audiomulch are set to synchronize to the in-built clock, ensuring the resulting music is all synched to a common beat, at 115 bpm. The separate audio channels are linked to an 8-track

mixer node, which is manipulated using a MIDI controller, allowing me to fade between different tracks as required.

As I have indicated, the only substantive materials that pre-exist the project's performance are the patches in each software application. Each patch awaits a "signal" of some kind in order to produce a result, and without a signal, there is no project. The only visual element in VVVV that pre-exists the project's performance, is a single transparent .png file with a white circle with a gradient fade around all sides. This generic image is multiplied and manipulated to produce the visual side of the project. Within VVVV, a 16-channel FFT analysis (fast Fourier transform) is performed on the audio signal produced in Audiomulch, captured by the internal microphone on the computer. The data-stream from each of the frequency bands is then sent, as a 16-channel data-stream or "spread" in VVVV parlance, through various nodes to multiply and smooth the values, and ultimately, it is used to drive transformations in scale of the visual elements. If there is no audio input, there is nothing to see, because the scale of the visual elements is reduced to zero, meaning that the visual side of the project is entirely generated by the audio side of the project. The same FFT stream is also used to generate a spread of values applied to the hue, saturation and brightness of the visual elements.

Because the audio signal is constantly changing, and is characterized by a mixture of rhythmic and random events, the concomitant visual responses are likewise randomly generated within a loosely determined field of specificity. At the same time, random events are being generated using a low-frequency-oscillator (LFO), and these events, after multiplying and smoothing the resulting numeric values, are used to trigger translations of the visual elements on the X, Y and Z axis. Other LFOs are used to ramp between values in various nodes, producing constantly changing visual effects. These are the essentially "generative" and aleatoric aspects of the VVVV patch, as they are automated processes that, once set in play, will produce endless permutations which, while maintaining a certain visual consistency, will never be exactly the same and can never be exactly predetermined.

Alongside the generative aspects of the project, however, are a number of controlled and intended elements. A MIDI controller is used to control the colour range, for example, allowing rough control of hue, brightness and saturation. The MIDI device is also used to switch a virtual camera on and off, and toggle between various camera movements, allowing for varying 3D perspectives of the visual elements to be produced. Between these generative and intended aspects of the project, lies a vast potential for the aspects of cross-modal binding to be explored. As the project progresses, numerous permutations of the relation between sound and image are explored, as a range of harmonic and rhythmic audio events produce a range of different visual transformations in colour, shape, scale, position and movement, and as the performer switches parameters in real-time, frequently responding to the rhythm of the music, to heighten the audience's perception of audio-visual commensurability.

Finally, it is worth noting how the overall structure of the *Udu* project is designed not only to produce a sense of audio-visual commensurability, but to produce it in the context of a performance that *feels* like a performance. Audio-visual performers

generally conceive their work *as* performance; they produce images and sounds live and on-the-fly, in real-time, and in that sense what they do is a performance comparable to what any other performer does. However, the question arises as to whether the performance is constituted in and by the audio-visual media, or by the artist's actions, or by some combination of the two. One of the frustrations expressed by many audio-visual performers or VJs is that while they know that they are producing the images (and sounds) that the audience perceives, the audience is not necessarily aware of how or even whether this occurs, rendering the performative aspect of the performance problematic [8]. The audience receives the result of the performer's actions as an audio-visual stream comparable to cinema or television, and, since these media forms are frequently produced ahead of time, in the studio or editing suite, the audience may easily receive the impression that this performance, too, was "pre-edited." Likewise, audio-visual performance is frequently underpinned by software applications and hardware devices, the manipulation of which rarely bears any direct perceivable relation to what is seen and heard. That is to say, watching a performer tweak knobs on a MIDI device is neither the most riveting spectacle nor the most transparent.

The point of using a drum to trigger sound, is that the performance begins with a concrete image of performance, where seeing the performer playing a musical instrument directly translates to seeing the production of visual imagery; for the audience, there is a direct relation between sound and image production. Because drums produce short, staccato sonic forms, as opposed to forms with lengthy sustains or durations, they lend themselves particularly well to visual representation, as they produce sudden spikes in amplitude that fall away quickly. When rendered as data-streams and mapped to visual parameters, these sudden spikes in amplitude produce immediate and dramatic visual transformations. Once the udu sound has been sampled, looped and processed, that direct audio-visual connection is replaced by a more malleable and shifting field, but the project begins, at least, by stating that connection "up front," providing the audience with a "ground" for the experience that is to come.

References

[1] Akiyama, Mitchell. "Projections: The Development and Politics of Digital Media in Audiovisual Performance." *Offscreen.* Vol. 11, noa.8-9 (2007). http://www.offscreen.com/biblio/pages/essays/akiyama_digitalmedia/ accessed 28/3/09.

[2] Bucksbarg, Andrew. "Vjing and Live A/V Practices." *VJ Theory: Texts*. 2008. http://www.vjtheory.net/web_texts/text_bucksbarg.htm accessed 11/02/09.

[3] Collopy, Fred. "Color, Form, and Motion: Dimensions of a Musical Art of Light." *Leonardo.* Vol. 33, no. 5 (2000): 355-360.

[4] Galeyev, Bulat M. “The Fire of ‘Prometheus’: Music-Kinetic Art Experiments in the USSR.” *Leonardo.* Vol 21, no. 4 (1988): 383-396.

[5] Grau, Oliver. *Virtual Art: From Illusion to Immersion*. Trans. Gloria Custance. Cambridge MA & London: MIT Press, 2003.

[6] Joseph, Branden. “’My Mind Split Open’: Andy Warhol’s Exploding Plastic Inevitable.” *Grey Room.* No.8 (2002): 80-107.

[7] Levin, Golan, and Zachary Lieberman. "Sounds from Shapes: Audiovisual Performance with Hand Silhouette Contours in "The Manual Input Sessions". Proceedings of NIME '05, Vancouver, BC, Canada. May 26-28, 2005.

[8] Lew, Michael. “Live Cinema: Designing an Instrument for Cinema Editing as a Live Performance.” Proceedings of the International Conference on Computer Graphics and Interactive Techniques, ACM SIGGRAPH (2004).

[9] Makela, Mia. “Live Cinema: Language and Elements.” MA Thesis, Helsinki University of Art and Design, 2006.

[10] Moritz, William. “The Dream of Color Music, And Machines That Made it Possible.” *Animation World Magazine*. 2:1, 1997. http://www.awn.com/mag/issue2.1/articles/moritz2.1.html

[11] Turco, Marina. “The Sudwestern Saga: Live Media and Modular Narratives Between Commedia dell’Arte and Digital Hypertexts.” *VJ Theory:TEXTS*. 4th July 2007. http://www.vjtheory.net/web_texts/text_turco.htm accessed 30/3/09.

[12] Van Campen, Crétien. “Synesthesia and Artistic Experimentation.” *Psyche.* Vol. 3 (1997). http://journalpsyche.org/ojs-2.2/index.php/psyche/article/viewFile/2358/2290 accessed 30/3/09.

[13] Whitelaw, Mitchell. “Synaesthesia and Cross-Modality in Contemporary Audioviduals.” *Senses and Society*, Vol. 3, issue 3 (2008): 259-276.