Processing divides a high-definition (1280 by 720 frames) video into two regions (left and right), the amount of movement in each region is estimated by Processing (through the use of a frame-difference algorithm) and sent to Max/MSP as a sequence of numbers that Max saves into a text file and 3) the data in the text file is converted in non-real time into an audio signal which is finally used to control the playback speed of the sound files in the left and right audio channels.

Figure 4. Movement detection on “Miniature 3”. Movement on each side of the screen (white pixels) is used to control the playback speed of sound files in non-real time.

7. CONCLUSIONS

Composing visual music through the use of computers enables the use of digital representations of sounds, images, movement or any kind of data -as well as the analyses that can be made from it- as models for the generation of forms and behaviours in different representational modes. Sounds, images and physical movement may present considerable perceptual differences between them, but represented as symbols, they can attain a certain degree of stability and the open up a space to establish translations (or, more precisely, transductions) between them. By using a transductive approach to visual music composition, the artist uses a representation of something as a model to create something new.

As Simon Emmerson points out (Emmerson, 2007, p. 42), the compositional choices that an artist has to make in using models in his creative process are at least four: 1) the identification of the model and 2) its behaviour, 3) the choice of the variables that will correspond to the chosen model and 4) the definition of the scale of their relationships. However, by choosing to compose digital visual music with a transductive approach, another decision has to be made: the nature of the compromise that will take place between the performativity of the system and the quantity and quality of the materials that will be used in the compositional process. In that regard, a non-real time approach seems to be a good option for the creation of visual music.

8. REFERENCES


9. WEB RESOURCES

All the videos from this article are available at www.carloslopezcharles.com

NOISE AND REPRESENTATION

Adam Collis
Coventry University
Performing Arts Department

ABSTRACT

Technical, aesthetic and philosophical issues around the expansion of computer music practice to include video are considered with respect to the author’s video work Early Evening, Summer, Deganny. An overview of the CD works of Japanese sound artist Yasunao Tone is used to highlight the issues surrounding the technologies for media recording and playback. This paper then describes and advocates a working process that eschews the representation of recorded phenomena in favour of the use of abstracted and noise-like materials in a work to form a presentation. Along with a discussion of the philosophical underpinnings of this work, a framework for understanding the use of ‘noise’ in sonic arts practice is provided.

1. INTRODUCTION

The camera and the sound recorder are designed to capture phenomena so that a representation of those tendencies to associate sounds, whether real or artificial, with a cause. Hyde suggests silence and noise (and their visual equivalents) as phenomena that make reduced listening and visual suspension possible on the basis that both are more readily understood on a conceptual level rather than as an experience of natural phenomena. In Early Evening, Summer, Deganny, sounds and images are thus processed to be more noise-like. What is understood as ‘noise’ though can be more than just the characteristic acoustic or visual phenomena that Hyde implies. To illustrate the various usages of the word, it is useful to consider the relationship between noise and representation that can be seen in the CD works of Yasunao Tone which have been a major influence on this piece. In those pieces, a principle concern of Tone has been the application of data transformation processes.

With his first CD work, Musica Iconologos [11], scanned images representing the individual ideograms of ancient Chinese poems are converted to audio files and processed. In the subsequent Solo for Wounded CD [12], mechanisms for ‘wounding’ a CD of Musica Iconologos are employed so that it skips when played in a CD player. Additionally, with works such as Wounded Soutai Man’yo [13], sound material is derived from Chinese ideograms drawn in the waveform editor of a digital audio workstation using a graphics tablet.

By the process of changing the medium in which information is encoded in these works, Tone strips signs of their signifiers and data loses its semantic meaning. Tone describes the resultant sound as ‘excess’ or ‘noise’, appealing to the French term for interference or background noise in a sound system, bruit parasite or ‘parasitic noise’ [13]. Thus, the sounds generated are noise in the sense of being parasitic on the original message.

What interests Tone is the process of data transformation, resulting in parasitic noise. Yet those processes result in sounds with rapid and seemingly arbitrary fluctuations in amplitude that is characteristic of noise-like sounds. When confronted with such sounds, as discussed by Wishart [15], are more likely to respond to frequency-domain information instead. Thus, these works could be considered as noise from a phenomenological perspective.

There is yet another sense in which the sounds of the CD works of Yasunao Tone can be considered as noise. A noted influence on Tone’s thinking has been Heidegger’s concept of “reduced listening”. An equivalent term for the viewing of images, ‘visual suspension’, is proposed by Hyde [7]. However, he also questions whether reduced listening and visual suspension are actually possible given human tendencies to associate sounds, whether real or artificial, with a cause.
The first stage in the process of creating my work, Early Evening, Summer, Degu Basin, was the generating of the images by abstracting photographs alluded to by the title. The visuals were generated using Apple’s Quartz Composer, part of their suite of developer tools, which provides an object-based programming environment for image generation and manipulation. While Quartz Composer provides standard image-processing objects such as blur, distortion and colour manipulation, there are also programming objects such as ‘Core Image Filter’ that enable the creation of custom image-processing algorithms in the Core Image Language (CIRK). This uses a subset of the OpenGL Shading Language (GLSLang) and this was used to create a routine for the primary processing of the source photographs.

[8] and the issues raised by Tone’s CD works recall Heidegger’s notion of an artwork ‘revealing’ the space in which it is situated [6]. Heidegger illustrated this idea with a temple that suggests its interaction with its surroundings, for example the shadows it forms with the sunlight, high-lighted the characteristics of the environment. Similarly, the acoustic-noise content of Tone’s CD works brings out the acoustic properties of the listening space through the excitation of resonant frequencies. In contrast, when playing mainstream music in the listening space through the excitation of resonant frequencies, a photograph centred at the origin in the Quartz Composer space presents a 3D space represented by a 2D image of pixels at some distance from the viewer or camera. The acoustic properties of the room itself are seen as a problem, causing unwanted tonal changes. Tone’s work, however, is noise also in a teleological sense as it serves the purpose of interacting with the listening space.

3. CREATION OF VISUALS

The basis for the sonic elements in this work comes from the Max patch, Trackpad, which was originally developed for laptop performance with a 48-channel version performed at ICMC2008 at SARC in Belfast [3]. In this patch, a stream of white noise is sampled into a buffer and played back in real time with various performance parameters mapped from the position and acceleration of a cursor controlled from the laptop’s trackpad. Under the control of the user are the length and playback speed of the sampling buffer, filter parameters, delay times and spatialisation. More recently, this patch has been updated and modified, simplifying the sound generation stage by using sine waves as the sound source and removing the filtering section but introducing more complex textures by creating layers using Max’s poly~ object. A performance of this patch was given at Staffordshire University’s Noiseflower festival in May 2012 [4]. For this work, the Max patch was further modified so that, rather than controlled by a performer with the trackpad, it was driven by pixel colour values. The RGB colour components for a single pixel in the video were transmitted via Open Sound Control to the Max patch with absolute values and changes in value of individual components used to change parameters of the patch. Altogether, nine audio layers were created, each driven from pixels distributed around the image frame, and the results mixed down. The aim of the processes undergone here is not a simple data mapping of pixel colour value to a parameter value in the Max patch, but a performing of the Max patch by the video.

5. DISCUSSION AND CONCLUSIONS

The processes outlined above result in a natural coherence between image and sound; changes in pixel colour values result in changes to the sonic texture. Synchronicity between the change in sound and change in image was not, however, a principle aim of the work. Of greater importance was the grounding of the work, both sonically and visually, in non-representation. The aesthetic choices here arise from an aim to appropriate fixed recording media for use in a presentation, rather than the representation of previous events.

The taking of field recordings and processing them into new forms is of course a fundamental part of the practice of an electro-acoustic composer, although the desired goal varies depending on the theoretical framework within which a composer may be operating. Truax, for example, describes the abstracting of sounds within soundscapes not to disguise their origins but to enhance an understanding of that sounds location [14]. In contrast, with her work Kitchen Alchemy, Blackburn works within spectromorphological principles to change the nature of her source material, the ‘sounds of kitchen utensils and appliances’, into something ‘seemingly precious and elaborate’ [1]. With Early Evening, Summer, Degu Basin, it is the images, rather than the sound, that undergo processing to strip them of their signifying characteristics.

My approach, in contrast, is not derived from phenomenological concerns but rather from Rorty’s writings on pragmatic philosophical ideas [10]. Implicit in the technologies behind both the photograph and the sound recording is the ‘realist’ view of ‘truth as correspondence’. Fidelity, faithfulness to the original image, is a desired requirement of audio and visual recording and playback technology. The criteria for evaluating recorded images are therefore established in an objective, natural order. Extrapolating such a ‘realist’ philosophical approach to music or sound-art composition may thus suggest sonic textures dominated by pitched sound material. Absolute, unambiguous frequency values and pitch classes enable the creation of an objective framework, such as even temperament or the major-minor key system, through which structural relationships can be established and work evaluated. In contrast, a pragmatic approach would resist appealing to such objectivity. The consequence of this is a greater use of noise-like materials. Noise-like sounds establish relationships based on relative values, introduce ambiguity and depend on context for their evaluation.

Part of this context is generated by the space where the work is presented. From the preceding discussion of Yasunao Tone’s CD works it was shown that, although
[8] and the issues raised by Tone’s CD works recall Heidegger’s notion of an artwork ‘revealing’ the space in which it is situated [6]. Heidegger illustrated this idea with a temple that suggests its interaction with its surroundings, for example the shadows it forms with the sunlight, high-lighted the characteristics of the environment. Similarly, the acoustic-noise content of Tone’s CD works brings out the acoustic properties of the listening space through the excitation of resonant frequencies. In contrast, when playing mainstream music recordings in conventional domestic settings, the acoustic properties of the room itself are seen as a problem, causing unwanted tonal changes. Tone’s work, however, is noise also in a teleological sense as it serves the purpose of interacting with the listening space.

3. CREATION OF VISUALS

The first stage in the process of creating my work, Early Evening, Summer, Deganyw, was the generating of the im-ages by abstracting photographs alluded to by the title. The visuals were generated using Apple’s Quartz Composer, part of their suite of developer tools, which provides an object-based programming environment for image generation and manipulation. While Quartz Composer provides standard image-processing objects such as blur, distortion and colour manipulation, there are also programming objects such as ‘Core Image Filter’ that enable the creation of custom image-processing algorithms in the Core Image Kernel Language (CIKL).

This uses a subset of the OpenGL Shading Language (glslang) and this was used to create a routine for the primary processing of the source photographs.

\[ z' = \frac{z}{1 + \frac{z}{Z}} \] (1)

where \( x' \) is the new \( x \) position of the pixel, \( z \) is the original \( x \) position, \( Z \) is the \( z \) position of the virtual camera and \( z' \) is the desired \( z \) position of the pixel. A corresponding formula is also applied to calculate the \( y \) position of a given pixel.

By defining a 2D sinusoidal function for \( z \), filters such as bump and pinch distortion or ripple effects that are familiar to users of image-processing applications such as Photoshop can be created. For this piece, however, the \( z \) values that form the basis of the translations were derived from pixel colour data.

4. CREATION OF SOUND MATERIALS

The basis for the sonic elements in this work comes from a Max patch, Trackpad, which was originally developed for laptop performance with a 48-channel version performed at ICMC2008 at SARC in Belfast [3]. In this patch, a stream of white noise is sampled into a buffer and played back in real time with various performance parameters mapped from the position and acceleration of a cursor controlled from the laptop’s trackpad. Under the control of the user are the length and playback speed of the sampling buffer, filter parameters, delay times and spatialisation. More recently, this patch has been updated and modified, simplifying the sound generation stage by using sine waves as the sound source and removing the filtering section but introducing more complex textures by creating layers using Max’s poly~ object. A performance of this patch was given at Staffordshire University’s Noisefloar festival in May 2012 [4].

For this work, the Max patch was further modified so that, rather than controlled by a performer with the trackpad, it was driven by pixel colour values. The RGB colour components for a single pixel in the video were transmitted via Open Sound Control to the Max patch with absolute values and changes in value of individual components used to change parameters of the patch. Altogether, nine audio layers were created, each driven from pixels distributed around the image frame, and the results mixed down. The aim of the processes undergone here is not a simple data mapping of pixel colour value to a parameter value in the Max patch, but a performing of the Max patch by the video.

5. DISCUSSION AND CONCLUSIONS

The processes outlined above result in a natural coherence between image and sound; changes in pixel colour values result in changes to the texture. Synchronicity between the change in sound and change in image was not, however, a principle aim of the work. Of greater importance was the grounding of the work, both sonically and visually, in non-representational choices here; for example, in creating a diegetic space with virtual objects such as a sphere rather than representing extraneous objects. This image processing abstracts photographs through the translation of individual pixels, destroying perspective and thereby obliterating the illusion of a representation of 3-dimensional space. The results of these transformations were bounced down into a video for subsequent image transformations which restore some sense of perspective. In a new Quartz composition, the resultant video was mapped onto a sphere in such a way that in certain parts of the frame the front of the sphere is visible and in other parts the inside of the rear of the sphere is visible. This is achieved by setting the sphere’s diameter so that the whole frame is filled with the edge of the sphere close to the position of the virtual camera and also by disrupting the rendering of the OpenGL depth buffer by setting it to ‘read-only’. The effect of this on the frame can be seen in figure 4.

Figure 3. The distorted photograph mapped onto a sphere.
these works do not feature any kind of representational spatialization techniques such as panning and reverberation, their playback generates a ‘live’ presentation of an acoustic space. Similarly, appreciation of Early Evening, Summer, Degaway results from the viewers experiencing the interactions of the presentation of the complex sonic and visual information within the space. This, despite being a fixed-media piece, by focussing on ‘noise’ and ‘presentation’ the fulfilment of this work is not the video itself, but the interactions between video, viewer and presentation space generated on playback. The piece could therefore be viewed just as much an installation, with sites of site specificity, as it is a video art work. So although the work has a specific length with an implicit dealing with time, the piece is, or could be regarded as ‘timeless’ with no large-scale organisation of materials. In the same way that the absolute relationships of pitched materials are eschewed in this work, linear temporal relationships are avoided to discourage associations with an external narrative. This does raise questions, however, regarding the use of abstraction in fixed-duration pieces that can be explored in further work.

6. REFERENCES

SYNESTHESIA: COMPOSING FROM IMAGES WITH NOISE SHAPING SYNTHESIS

Antonio D’Amato
Conservatorio Statale di Musica “D. Cimarosa”
Avellino - Italia

The artist investigates here the decomposition of the light spectrum in separate components:
La composition reprend le motif iconique du Disque, isolé par Robert sous la forme d’une cible optique en 1913 (coll. part.). Le module géant est démultiplicé en anneaux colorés châtant de toutes les couleurs du prisme, qui absorbent tout la surface, l’unique et annulent toute perspective. Née de la simple observation par l’artiste des transformations des couleurs et des formes imposées par l’éclairage électrique... [3]

2. SOUND FROM IMAGES
The question was how to transform or derive music from images.
A similar process could be adopted in a sound processing algorithm by separating a complex spectrum – i.e. white or filtered noise, or any other synthesized complex spectrum - into separate components.
Furthermore, an inverse process could be adopted in a sound-generating algorithm, by synthesizing a sound from separate components, derived from the content of an image, where different pixel locations in the frame represent different frequencies.
The first part of the composing process was the following: graphically isolating some elements of the painting, and afterwards transforming these into single sounds where their intensity was given by the luminance factor and their frequency from pixel locations.
It could be considered a noise-based synthesis, since it generates original sounds from images as output. It was not a mechanical process because it was clearly possible to experiment variants adopting as source material images modified with any imageprocessing software as the wellknown Photoshop or similar ones. Moreover, it is possible to limit the bandwidth of the noise-shaping process, or even adopt a linear or logarithmic scale in the location/pitch conversion. Consequently, a wide range of different results is possible. A further option in some cases involved a reverse process converting wildly transformed sounds into images again, then processing them again as images and turning them back finally into the audio domain.