



This work is protected by copyright and other intellectual property rights and duplication or sale of all or part is not permitted, except that material may be duplicated by you for research, private study, criticism/review or educational purposes. Electronic or print copies are for your own personal, non-commercial use and shall not be passed to any other individual. No quotation may be published without proper acknowledgement. For any other use, or to quote extensively from the work, permission must be obtained from the copyright holder/s.

A portfolio of compositions consisting of  
*Sketches, Mutations, Evolution, Emergence,*  
*Insomnia, Xpressions,*  
and *Ut infinitio quod ultra*

Andrew Martyn Willy

December 2013

Keele University

A submission for the degree of Doctor of Philosophy

# Abstract

This thesis builds a conceptual framework that enables the creation of synergetic audiovisual works which is “fundamental to the nature of complex interactive dynamic sound and light systems and their resulting forms.”<sup>1</sup> Utilising electroacoustic Schaefferian principles and applying them to a new media, I propose that objects can be established synergistically as an audiovisual whole, greater than the sum of their parts. The portfolio develops several tools to aid the creation of the video stream and they are further explored to allow for real-time parametric mapped audio and video manipulations. Throughout the portfolio the concept of an associative audiovisual tripartition is created and refined, thus allowing for the exploration of the media as well as a discursive approach to the auditory and visual streams. The set of works offered in this portfolio aims to raise the video medium from mere visual commentary to an audiovisual stream within a new synergetic audiovisual media.

This portfolio comprises seven audiovisual compositions for digital media and a written commentary that describes the aesthetic concerns, technical features, and compositional approaches addressed in the portfolio as a whole and in each work in particular.

The pieces are submitted on USB Flash drive and optical media, along with accompanying software patches. Compositions are organised as follows:

Sketches. Duration: 7 minutes and 38 seconds.

Mutations. Duration: 11 minutes and 24 seconds.

---

<sup>1</sup> Ronald Pellegrino, *The Electronic Arts of Sound and Light* (New York: Van Nostrand Reinhold Company 1983). p 208.

Evolution. Duration: 11 minutes and 53 seconds.

Emergence. Duration: 6 minutes and 26 seconds.

Insomnia. Duration: 5 minutes and 5 seconds.

Xpressions. Duration: 7 minutes and 19 seconds.

Ut infinitio quod ultra. Duration: 13 minutes.

# Contents

Abstract .....	1
Contents .....	3
List of figures .....	4
List of tables .....	4
Acknowledgements .....	5
Chapter 1: Introduction .....	6
Chapter 2: <i>Sketches</i> .....	28
Chapter 3: <i>Mutations</i> .....	46
Chapter 4: <i>Evolution</i> .....	55
Chapter 5: <i>Emergence</i> .....	61
Chapter 6: <i>Insomnia</i> .....	69
Chapter 7: <i>Xpressions</i> .....	78
Chapter 8: <i>Ut infinitio quod ultra</i> .....	86
Conclusions .....	95
Bibliography .....	97
Appendix .....	99

## List of figures

Figure 1 Continuum of gestural audio vs. visual association categories.....	12
Figure 2 Audiovisual tripartition .....	16
Figure 3 <i>Sketches_JitterPatch1</i> .....	32
Figure 4 <i>Sketches_JitterPatch2</i> .....	33
Figure 5 Wake filter settings.....	37
Figure 6 Four main sections of <i>Sketches</i> .....	40
Figure 7 Stars.mov .....	41
Figure 8 walkingJavaFilt.mov .....	42
Figure 9 Jitter MIDI CC mapping and interpolation .....	49
Figure 10 Process 1 flow chart.....	52
Figure 11 Process 2 flow chart.....	53
Figure 12 Waves clip after application of 'Plaster' filter. ....	64
Figure 13 <i>Insomnia</i> filtered and composited source material.....	72
Figure 14 <i>Insomnia</i> interwoven lines.....	73
Figure 15 <i>Insomnia</i> visual texture. ....	73
Figure 16 <i>Insomnia</i> windows.....	74
Figure 17 Surround Setup: speaker location.....	99

## List of tables

Table 1 <i>Mutations</i> : Initial settings for opening phrase.....	50
Table 2 <i>Mutations</i> ' structure with initial video settings. ....	51

## **Acknowledgements**

My sincere gratitude goes to Diego Garro for his help, support, guidance and encouragement in the creation of this thesis. To Rajmil Fischman for his enthusiasm. To Sohrab Uduman for his faith and the opportunities he gave to me. To Cliff Bradbury for his invaluable help and technical guidance. To Kate Cushing for her unwavering support through all the difficult periods, particularly in the latter stages of this endeavour. To Steve Bird for his oblique musings and to Juliet Yates for her friendship and proof reading skills.

Special thanks go to my parents, without whom I wouldn't have got this far, to Hilary, and to Sarah-Jayne Ainsworth for starting me off down this road and 15 years of relentless encouragement thereafter.

A portfolio of audiovisual compositions consisting of *Sketches*, *Mutations*, *Evolution*, *Emergence*, *Insomnia*, *Xpressions*, and *Ut infinitio quod ultra*

## **Chapter 1: Introduction**

My earliest experiments with audiovisual composition were undertaken whilst studying electroacoustic composition as an undergraduate at Keele University, where Dr. Diego Garro introduced me to the world of creative multimedia and the works of Prof. Dennis H. Miller and Adriano Abbado, amongst others. My master's research, supervised by Dr. Sohrab Uduman, focused exclusively on audiovisual composition, investigating combinations of real world and abstract sound and imagery, using compositional skills and techniques derived from electroacoustic music. The Ph.D portfolio builds on these earlier experiments and focuses on the development of various techniques for combining and manipulating audio and video to create audiovisual works that seek out moments of synergy, where new audiovisual objects can be created.

The key aim of this research project is the development of a concise compositional strategy that deals effectively with both audio and video media which guides the composition portfolio. Whilst this approach is firmly based in the electroacoustic methodologies derived from the work of Pierre Schaeffer and the continual development since then, the inclusion of a moving image in the works creates a new set of challenges and requires new tools and techniques to be derived and considered.



## Audiovisual/visual music composition

The term visual music has been adopted by numerous practitioners and analysts to describe multimedia works that elicit an intrinsic link between auditory and visual stimuli. The term was coined by Roger Fry to describe Wassily Kandinsky's paintings: "connecting Kandinsky's non-representational art to the similarly abstract nature of music was a way to explain and interpret this new art form."<sup>2</sup> The term is now used more broadly to define many types of multimedia composition, which Ron Pellegrino attempts to classify in his book *Emergent Music and Visual Music: Inside Studies*. Pellegrino categorises at least 16 visual music 'flavours', while none of them lend themselves entirely to the works presented here, he concludes by suggesting that "to be included in the visual music field, a work's dynamic imagery should be variously informed by the music, integrated with the music, embody the music, and/or grow out of the music."<sup>3</sup> It is this conclusion that offers a much closer insight into the portfolio. Fox-Gieg *et al* suggest similar definitions, including: "Visuals composed as if they were music, using musical structures."<sup>4</sup> This suggests that there need not be any sound at all, a link back to visual music's earliest use. Despite theorists and practitioners striving for an absolute definition, with many using different locutions to describe their works, it is still hard to pin down exactly what is – and what is not – visual music, which only goes to highlight the relative infancy of the genre.<sup>5</sup>

For this research portfolio I describe the works primarily as audiovisual compositions. Andrew Hill uses Emerson and Smalley's definition of

---

<sup>2</sup> N Fox-Gieg, Keefer, C and Schedel, M, "Editorial," *Organised Sound* 17, no. 2 (2012): p 97.

<sup>3</sup> Ronald Pellegrino, *Emergent Music and Visual Music: Inside Studies* (Denton: Electronic Arts Productions, 2009). p 71.

<sup>4</sup> Fox-Gieg, "Editorial," p. 97.

<sup>5</sup> Maura McDonnell discusses further definitions in her paper "Visual Music – a Composition of the things themselves".

electroacoustic music to derive the term electroacoustic audio-visual music, which can be seen as analogous to my works:

[A]n electroacoustic audio-visual music work could be defined as: a cohesive entity in which audio and visual materials are accessed, generated, explored and configured, primarily currently with the use of computer-based electronic technology, in the creation of a musically informed audio-visual expression.

Electroacoustic audio-visual music works explore the possibilities that the combination of their two time-based media (sound and moving image) allow.<sup>6</sup>

In this commentary I take this broader definition to encompass visual music – at least those definitions that represent the presentation of both musical and visual ideas – and assume a position where both audio and video mediums have equal standing in the compositional process. These are then composed within a temporal architecture,<sup>7</sup> building on Fischman's notion that "regardless of any other levels of paradigmatic coherence, music can work at an ultimate abstract level: the structuring of time."<sup>8</sup>

Visual music has been linked, at least historically, with the multi-sensory pitch/harmony-colour associations of synaesthesia. The earliest experiments with visual music were driven by an interest in colour, with Bertrand Castel, A. Wallace Rimington and others creating the very first colour organs. These machines were keyboard based instruments that responded to inputs by playing both note and displaying an associated colour on a screen.<sup>9</sup> Whilst there has been much creative

---

<sup>6</sup> Andrew Hill, "Desarrollo De Un Lenguaje Para La Música Audiovisual Electroacústica: Investigación Sobre Su Comunicación Y Clasificación," in *En El Límite — Escritos Sobre Sonido, Música, Imagen Y Tecnología*, ed. Raúl Minsburg (Universidad Nacional de Lanús, 2010). p 144.

<sup>7</sup> Brian Evans, "Foundations of a Visual Music," *Computer Music Journal* 29, no. 4 (2005): p 11.

<sup>8</sup> Rajmil Fischman, "Music for the Masses," *Journal of New Music Research* 23(1994). p 260.

<sup>9</sup> Fred Collopy, "Color, Form, and Motion," *Leonardo* 33, no. 5 (2000): p 356.

work established on synaesthetic principles, my own research is not synaesthetic, and I make no claim to be able to hear colours or see sounds. To that end I have made no attempt to engage with synaesthetic approaches, and would argue that, as Nicholas Cook notes, whatever the poietic<sup>10</sup> significance of ascribing colour labels to a work, it has no esthetic impact on the listener.<sup>11</sup> What I do make are conscious pictorial connections with regard to form, shape, texture and gesture when listening to electroacoustic music and it is this which informs my audiovisual composition.

## **Compositional approach**

My compositional approach, which is based on the principles of electroacoustic music and particularly those of early pioneers such as Pierre Schaeffer, is clearly at odds with the prime axiom of *musique concrète* – that of reduced listening<sup>12</sup> and the primacy of the ear. Garro suggests that whilst this is clearly the case, “in the hands of many composers, [visual music works challenge the ear’s primacy] while

---

<sup>10</sup> Nattiez uses these terms to separate the process of production (the poietic process) from the process of reception (the esthetic process) via the ‘neutral’ or ‘trace’ – in this case the musical work. The implication is that whatever the codification or structuring process employed by the composer may be, these processes may not be directly perceivable by the listener through the work.

Jean-Jacques Nattiez, *Music and Discourse: Toward a Semiology of Music* trans. Carolyn Abbate (New Jersey Princeton University Press, 1990).

<sup>11</sup> Nicholas Cook, *Analysing Musical Multimedia*, 2 ed. (Oxford: Oxford University Press, 2004). p 32.

<sup>12</sup> One could reinterpret the term for this combined media. Hyde suggests an extension of the reduced listening paradigm to the visual domain can be achieved. His work *Vanishing Point* (2010) deals with notions of noise, treating visual materials much like visual objects, with any narrative associations ignored. Whilst this works on a conceptual level, the use of recognisable video footage creates the same juxtaposition as found in the audio domain, rendering it impossible for a viewer to avoid real world connections.

Hyde, *Vanishing Point* Programme Note, Available <http://www.josephhyde.co.uk/video/vanishing-point/> [Accessed 2 December, 2012].

treasuring all technical and artistic values nurtured during several decades of sonic endeavours, research and practices.”<sup>13</sup>

Whilst I find resonance in much of the aesthetic of *musique concrète*, the idea of reduced listening – the listening mode that takes a sound “...as itself the object to be observed instead of as a vehicle for something else.”<sup>14</sup> – is fundamentally flawed. As Hyde observes, “when presented by unfamiliar and abstract sounds, the listener will tend to imagine a source of origin even if such a source is not apparent.”<sup>15</sup> However, far from being problematic, this failure highlights one possible way in which to introduce and interact with a visual medium. By acknowledging that we seek the source of a sound, whether it is in an attempt to uncover the transformed sound’s source – ‘is that really the sound of waves crashing against rocks?’ – or to establish some other form of associative or causal link, we can build on the listener’s need to unify a sound with its ‘source’ by offering a visual cue/identity to the sonic content. Note that I am not proposing the creation of a visualisation of sonic material via some algorithmic process, nor the simplistic ‘Micky Mousing’<sup>16</sup> of sonic and visual events, but the composition of audiovisual correspondences. It is this proposition that triggered my early experiments.

---

<sup>13</sup> Diego Garro, "From Sonic Art to Visual Music: Divergences, Convergences, Intersections.," *Organised Sound* 17, no. 2 (2012): p 104.

<sup>14</sup> Michel Chion, *Audio-Vision*, trans. Claudia Gorbman (New York: Columbia University Press, 1994), p 29.

<sup>15</sup> Joseph Hyde, "*Musique Concrète* Thinking in Visual Music Practice: Audiovisual Silence and Noise, Reduced Listening and Visual Suspension," *Organised Sound* 17, no. 2 (2012): p 173.

<sup>16</sup> Cook, *Analysing Musical Multimedia*, p 179.

As Piché notes, due to the increasing sophistication of computer workstations audio and video editing tools and techniques have converged allowing the audiovisual composer to work with video in a very similar way to audio<sup>17</sup>:

I was doing a lot of sound design for video artists such as Tom Sherman. We'd go into the studio together and he would do his capturing and processing of images, and I looked at it and realized that this is exactly what I do when I work with sound. It's the same workflow that I use when I'm composing with computer. You have your timeline, your sounds, your processing chops, and with the timeline you organize a coherent discourse.<sup>18</sup>

More than just a tool for compositing clips, digital video editing software allows for myriad transformations (filtering, stretching, reversing, etc.) to be applied to video clips in exactly the same way as working with audio within digital audio editing software. Because of these developments in editing tools, video materials can now be treated in the same concrete fashion as audio; this allows congruous manipulation and transformation of sonic and visual entities into a synergetic whole work. Furthermore, the ability to design bespoke video editing software patches, allows the composer complete freedom to work with combinations of audio and video material, often in real time.

### **Association strategies, parallel streams and audiovisual objects: a tripartition**

This portfolio was created with the aim of producing a set of works that raise the video medium from mere visual commentary to an audiovisual stream where objects can be established synergistically as an audiovisual whole, greater than

---

<sup>17</sup> Moreover, composers are now able to run both audio and video editing packages on their digital audio workstations thanks to the increased computational power afforded by recent computing developments, allowing immediate engagement with the audiovisual media.

<sup>18</sup> Paul Steenhuisen, "Interview with Jean Piché," <http://www.jeanpiche.com/Textes/Interview.htm>.

the sum of their parts, utilising electroacoustic Schaefferian principles and applying them to a new medium. Tools have been developed to aid in the creation of the video stream, and further explored to allow for real-time parametric mapped audio and video manipulations. Throughout the portfolio the concept of an associative audiovisual tripartition has been created and refined, which has thus allowed for the exploration of the media as well as a discursive approach to the auditory and visual streams.

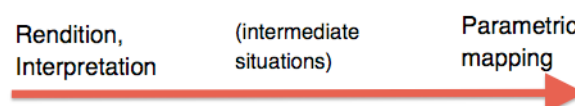
It seeks to build a conceptual framework in three parts to allow the creation of synergetic audiovisual works, which as Pellegrino observes, is “fundamental to the nature of complex interactive dynamic sound and light systems and their resulting forms.”<sup>20</sup>

## Association strategies

At the nucleus of my compositional praxis is an (edited) association continuum.

Garro discusses this concept in his 2005 paper ‘A Glow on Pythagoras’ Curtain’<sup>21</sup>, illustrated in figure 1.

**Figure 1 Continuum of gestural audio vs. visual association categories**



At one end of the continuum audiovisual associations are created by pure interpretation, where audio and video gestures are allowed to evolve with no

<sup>20</sup> Pellegrino, *The Electronic Arts of Sound and Light*, p 208.

<sup>21</sup> Diego Garro, "A Glow on Pythagoras' Curtain: A Composer's Perspective on Electroacoustic Music with Video." (paper presented at the Electroacoustic Music Studies Network Conference, Montréal, 2005), p 4.

obvious causal links. At the other end, associations are based on some identifiable phenomenological precept in the sound and visuals. Intermediate situations develop from interpretive, loosely connected media to more firmly associated audiovisual gestures; these tend towards synchronised sound and visuals as one travels towards parametrically mapped strategies. Synchronised audio and video materials play a particularly useful role in forming associations. Chion describes this as *synchresis*, where there is a “spontaneous and irresistible weld produced between a particular auditory phenomenon and visual phenomenon when they occur at the same time. This join results independently of any rational logic.”<sup>22</sup> This last statement is particularly powerful when working with predominantly abstract material, such as those found in this portfolio; it is the composer who forges the link between auditory and visual stimuli. The term *synchresis* will be used in place of synchronised to infer this newly created phenomenological relationship.

My research revisits and revises the rendition end of the continuum, the development I offer suggests that there need not be any link between the two media and that in fact, there can be complete separation. The lack of any link is, after all, a strategy.<sup>23</sup> Whilst I appropriate the continuum and associative terms, I disregard the continuum’s exclusive concern with gestural audiovisual profiles and instead suggest that it can be used much more freely to associate all audiovisual morphologies and trajectories when used within the tripartition.

---

<sup>22</sup> Chion, *Audio-Vision*, p. 63.

<sup>23</sup> Garro makes a similar alteration to his continuum in ‘From sonic art to visual music’ (Garro 2012: 107), adding separation at the continuum’s extreme and promoting rendition to ‘intuitive complementary’.

## Audiovisual objects

Audiovisual objects are combinations of audio and video streams that are intuitively combined using the edited continuum to form a synergetic whole. They can exist throughout the association range but have a predisposition toward the synchresis/mapped end of the continuum. These objects are not solely fixed relationships based on synchresis, as suggested by Hyde,<sup>24</sup> which tend to be exhausting when used exclusively, creating meaningless visual commentary to an electroacoustic composition. Instead, through careful selection and manipulation of materials, and through profiling techniques based on the entire association continuum, sounds and images can be linked to create true synergetic audiovisual entities. By considering these objects and their morphologies they may also suggest particular associations that can then be explored through the continuum.

To clarify, these objects are not the same '*objet audiovisuelle*'<sup>25</sup> that Garro proposes, as these refer to predominantly gestural articulations, and which is why the gestural function of Garro's continuum is disregarded. Instead, my research develops this idea and suggests that these objects can have both gestural and textural profiles: it is the underlying structural process and temporal trajectory that defines the object.

A further example, by way of metaphor, is that one instance of an audiovisual object can be seen as a house brick: it has shape, form, surface texture. It is relatively small and can be easily moved from one area to another. Another object is a brick wall: it too has shape, form, surface texture but it is a much larger object,

---

<sup>24</sup> Hyde, "*Musique Concrète* Thinking in Visual Music Practice: Audiovisual Silence and Noise, Reduced Listening and Visual Suspension," p 172.

<sup>25</sup> Garro, "A Glow on Pythagoras' Curtain: A Composer's Perspective on Electroacoustic Music with Video.", p 1.



which has (both literally and metaphorically in this instance) structural functions. It cannot be easily moved (although it could be knocked down and replaced by something else!). In this example the brick is representative of a short audiovisual gesture, with a morphology of, say, one second, the wall therefore represents a textural passage of ten to twenty seconds – a drone, or dense granular cloud – that underpins a particular phrase or passage.

## **Parallel streams**

In close correlation to audiovisual objects is a *parallel streams* approach towards audio and video. These streams represent several distinct entities within audiovisual works and are quasi-hierarchical: at the top level they represent each audiovisual medium individually – the audio stream and video stream. This division is useful when conceptualising the initial structure of a work, or section of a piece, in as much as the composer can decide whether the streams should remain static with a particular morphological association, should converge, tending to the right hand side of the association continuum, or diverge and separate, tending to the left of the continuum.

At the second level one can categorise individual streams in each medium separately. There may be many audio streams within a passage or phrase, defined, for example, by frequency – high frequency streams, low frequency streams etc.; by spectromorphological archetypes<sup>27</sup> – stream of granular sounds, or closed attack-decay morphotypes. Similarly visual streams may include particular forms or shapes moving around the screen, for example: grainy

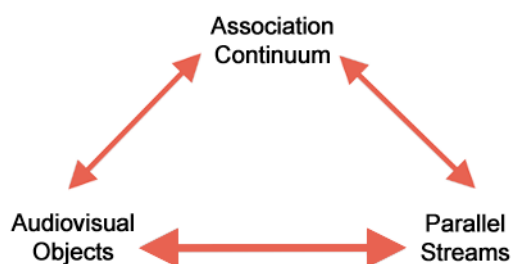
---

<sup>27</sup> Cf. Smalley, Denis. "Spectro-Morphology and Structuring Processes." In *The Language of Electroacoustic Music*, edited by Simon Emmerson, 61-93. New York: Harwood Academic Publishers, 1986.

particles, lines; streams of flashing imagery; planes or surfaces that exhibit a particular texture or colour. Again, these streams can be combined to create audiovisual objects, they can be placed in static association, or separate and agglomerate based on some temporal or morphological trajectory.

The parallel streams paradigm offers both syntax with which to describe audiovisual events with regard to their temporal trajectories and a compositional tool with which to structure these events in time. Furthermore it also allows the composer to work with sonic and visual elements on an equal footing where audio and video streams run in parallel, interacting to create an audiovisual discourse. This third approach shares many similarities with Garro's discursive audiovisual association strategy.<sup>28</sup>

To summarise, the tripartition offers this three-part process to audiovisual design: The association continuum acts to offer strategies with which to join audiovisual elements; by linking audio and video elements, or streams, one can create audiovisual objects which aid in the creation of a synergetic audiovisual whole; a parallel streams approach acts both to raise the visual element of a multimedia composition from mere commentary to a visual stream, and to organise materials in time. A visual representation of the tripartition can be seen in figure 2.



**Figure 2 Audiovisual tripartition**

<sup>28</sup> Garro, "From Sonic Art to Visual Music: Divergences, Convergences, Intersections.," pp. 107-8.

Whilst the earlier works in the portfolio focus on and discuss specific aspects of each of the three parts, the framework has been developed as a holistic approach to the creation of synergetic audiovisual works. It is clear that, although not formalised during the creation of the early works in the portfolio, all pieces display a strong, cohesive structure based on the functions of the tripartition.

## **Technical concerns**

The convergence of audio and video editing tools, with regard to approaches to audiovisual composition, has already been mentioned as part of the discussion of methodology.<sup>29</sup> In terms of technical development, the discipline is one that has, and will continue to, advance with technical developments in hardware and software design. Tools and techniques that were created and used five years ago were in their infancy and have developed and been refined by their authors. This is evident in the portfolio when viewed as a whole, and particularly the works *Sketches* and *Mutations*, which were created working on the very limits of what was possible (visually) at the time. Increased computer processing power and software sophistication has alleviated the majority of technical issues encountered and will continue to do so. This will allow for even more complex processes to be attempted in the future.

The creation of this portfolio was enabled by utilising a variety of tools. At the heart of my compositional praxis is the Adobe Audition audio editor (successive versions)<sup>31</sup>. The majority of the sound design for the portfolio was achieved within this package. Arranging, certain effects processing, and spatialisation was accomplished in Steinberg's Cubase for *Sketches*, *Mutations* and early versions of

---

<sup>29</sup> See above: paragraph 10.

<sup>31</sup> Formally Cool Edit Pro, by Syntrillium Software Inc.

*Evolution*. The final version of *Evolution* and the remainder of the portfolio used Apple's Logic Studio Pro. This switch (accompanied by a change in platform from predominantly Windows to Mac) was primarily due to the inconsistent stability of the former and I have found little difference in the workflow. Several other plugins and software synths were used to manipulate and transform source material: these will be mentioned specifically in the individual commentaries where relevant.

Visual processing was achieved through an equally varied set of tools. U&I Software's ArtMatic was used for initial source material transformation in *Sketches* and *Mutations*. These were further transformed by tools created in Cycling74's Jitter environment. *Sketches* and *Mutations* were assembled and rendered in Adobe's Premiere Pro although I found this to be a less than intuitive program. All further works were created using Sony's Vegas Pro video editing software. I have found working in this environment much more compliant with my overall workflow,<sup>32</sup> with the software offering many filters and processes comparable to audio techniques offered in Audition. I also used Adobe's Photoshop to process some of the video material. Although not particularly well documented, the most recent versions of Photoshop are able to apply filtering effects to video files, effectively batch rendering them as individual images. This allows the audiovisual composer to interact with the various powerful graphics functions afforded by the software.

Video files were rendered as high definition 720p files, using the widely available H.264 codec, and audio files have been mixed down to 24 bit interleaved wave files. These were then combined, or multiplexed, by the multimedia tool

---

<sup>32</sup> The only negative to this approach is the Windows-only compatibility of Vegas. This required working in Apple's Bootcamp, or via a virtual machine in Parallels.

MKVmerge into the MKV file format, which will play on any modern computer using the freely available VLC media player<sup>33</sup>.

## **Surround sound**

From their inception all works have been produced in 5.1 surround sound. The format was originally designed for cinema spaces as a front-orientated sound stage with supportive rear effects or ambiances,<sup>34</sup> however it is most effective in a more musical context. The space afforded by a multichannel system can be controlled in several ways, as suggested by Otondo:

[F]or some purposes one can work with a conventional stereo image reinforced with the sound of the rear speakers and for others one can treat all the speakers of the systems as discrete sources that can be used to create contrapunctual (sic) and independent spatial developments, not necessarily based on a stereo image but on a number of discrete independent sources.<sup>35</sup>

This is how I decided to treat the space in my own work: primarily utilising the front channels as an expanded stereo pair, with the rear channels acting to create an immersive environment by reinforcing particular sonic streams, whilst using a discrete source approach for specific gestural streams. Through trial and error, I have found that often this latter approach works best if sounds are subtly supported using several channels: for example, if a discrete sound in the rear-right channel is desired it may be helpful to allow a small percentage of the signal through to the front-right and rear-left channels also. This method helps alleviate

---

<sup>33</sup> Available from <http://www.videolan.org/>

<sup>34</sup> Francis Rumsey, *Spatial Audio*, Music Technology (Oxford: Focal Press, 2001), p 87.

<sup>35</sup> Felipe Otondo, "Some Considerations for Spatial Design and Concert Projection with Surround 5.1," *eContact!* 8.3(2005). [http://cec.sonus.ca/econtact/8\\_3/otondo.html](http://cec.sonus.ca/econtact/8_3/otondo.html) [Accessed 2 December 2012]

the spatial discontinuity Otondo reports and creates a more integrated surround space.

The decision to use a multichannel audio format was based on the desire for the spatialisation to form part of the compositional process rather than remain a diffusion/performance technique. The audio streams' spatial trajectories are conceived from the very beginning of the compositional process and fixed in the multichannel format, rather than relying on manual diffusion techniques.

From both attending and performing in diffused concerts it became clear that, if specific trajectories are suggested by particular audio streams, it would be advantageous to compose these trajectories from the outset, rather than relying on various (often complicated) fader manipulations on a mixing desk to achieve the same result.

This approach was prompted by and developed from Harrison's 1999 article *Diffusion: theories and practices, with particular reference to the BEAST system*. Harrison suggests a holistic approach to composition that takes performance issues into account from the very inception of a work. He notes that the composer is already sculpting materials in time and space, and that "...diffusion is merely the necessary continuation and enhancement of this process – articulating spatial qualities already possessed by the sound, not arbitrarily imposing spatial behaviour on it."<sup>36</sup>

---

<sup>36</sup> Jonty Harrison, "Diffusion: Theories and Practices, with Particular Reference to the Beast System," *eContact! 2.4* (1999). <http://cec.sonus.ca/econtact/Diffusion/Beast.htm> [Accessed 21 December 2012]

The 5.1 format was chosen for both technical and aesthetic reasons; it is useful for dissemination purposes – works can be encoded and distributed on DVD. It was also chosen due to the predictability of performance in small to moderate size spaces and the availability and straightforward configuration/integration of playback equipment. Larger venues require reinforcement of the rear channels, which can be achieved by duplication of the rear outputs, but preferable is the creation of specific mixes for individual performances. This is easily achieved by adjusting the automation within Logic Studio Pro.

## **Portfolio development**

The portfolio chronicles the experimentation with strategies and development of a compositional framework, in the form of the audiovisual tripartition. *Sketches* (2007, 7:38) deals primarily with association strategies, seeking to explore initial ideas of combinations of streams to create audiovisual objects; it considers the concept of streams in an embryonic state only. Key to the creation of this work was the development of video tools (Jitter) with which to design, create, and manipulate visual material.

*Mutations* (2008, 11:24) builds upon these initial association strategies, and attempts to combine real-time manipulations of both audio and video through the further development of the Jitter software which is combined with audio synthesis tools.

*Evolution* (2009, 11:53) explores filtering approaches in both video and audio domains. A thirty second video sequence was used as the video source material with a limited palette of audio source material. Through treating, effecting, manipulating and stretching the audio and video sources, the material is

developed to the point of abstraction. The resulting audiovisual material is then explored to create congruent audiovisual motifs, which therefore, develops the converging / diverging streams paradigm.

*Emergence* (2011, 6:26) further develops the concrète treatment of materials experimented with in *Evolution* by exploring more cohesive source material in the form of several recordings of waves lapping the shore.

*Insomnia* (2012, 5:05) advances the parallel streams paradigm in its approach to the materials and develops increasingly divergent audiovisual correspondences. The work treats the visuals as a contrapuntal stream to the increasingly complex sonic morphologies.

*Xpressions* (2012, 7:19) explores the primarily syncretic<sup>37</sup> relationship between the audio and video streams. It develops rhythmic movement within the image which is associated in several different ways within the audio stream.

*Ut infinitio quod ultra* (2012, 13:00) aims to be a culmination of each of the three strands in the tripartition, which brings together and unifies them within a complex gesture-rich audiovisual work. It employs the full range of associations, utilising mapped objects as staging points for streams to diverge, before being reinterpreted by the changing visual gestural morphologies – from vapourous, indistinct shapes, to rays with distinct contours; from those contoured shapes to objects with complex structures.

---

<sup>37</sup> From Chion's synchresis.



## Placing the portfolio within current research practices

This portfolio has been developed over the last five years and, as is inevitable with an extended piece of research, in that time an increasing number of composers have begun to engage with audiovisual composition. Research to date explores primarily strategies for composing with audiovisual media, specifically approaches to combining the new media's audio and video streams. This section will examine the work of several composers and compare their approaches to those used in this portfolio.

Bill Alves' practice is highly influenced by the work of John Whitney, having worked together for ten years. Alves derives many of his compositional principles from Whitney's *differential dynamics*, a process of manipulating points of light into symmetrical patterns based on Pythagorean number proportions.<sup>38</sup> A clear distinction can be made between this approach and my own compositional praxis: whilst Alves seeks to unite the audio and video streams via a numerical framework, I use the audiovisual tripartition, mentioned previously, to guide the structuring of streams and objects in time.

Alves' *Static Cling* (2000) uses the POV-ray and Csound programming languages to produce an algorithmically derived work that explores ways in which harmonic patterns can emerge from visual or auditory chaos.<sup>39</sup> It uses just intonation ratios to manipulate the audio stream and provide coordination with the *differential dynamics*-derived symmetrical visual patterns. A structural correspondence is created, whereby points of sonic stability are complemented by relative visual

---

<sup>38</sup> Bill Alves, "Digital Harmony of Sound and Light," *Computer Music Journal* 29, no. 4 (2005): pp. 45-46.

<sup>39</sup> *Ibid.*: p. 51.

symmetry and coherence,<sup>40</sup> similar to the structuring technique of visual consonance and dissonance proposed by Evans.<sup>41</sup> This creates a parametrically mapped work, with particular emphasis on iterations that coincide with “metrically important points”<sup>42</sup>. After an initial predominantly textural introductory section, vertical lines flash in correspondence to interesting polyrhythmic patterns, which are sustained throughout the piece.

There appears to be little textural development of the soundtrack beyond the opening 45 seconds. Noise-based streams of highly distorted vocal sounds gradually resolve into multiple streams of vocal material – local news reports. This resolution also takes place in the video stream, with ‘noisy’ video static developing into a grid shape (ca. 00:00 – 00:45).

My own work treats the visual space quite differently with a reliance on textural underpinning to counterpoint the gestural content. In *Insomnia* emerging lines have a similar initial gestural profile to the polyrhythmic flashing lines in *Static Cling*, however the auditory and visual streams are allowed to diverge. As the number of lines increases, the textural quality of the interwoven mass dominates the screen, creating a metaphorical drone in the visual stream, with audio streams multiplying and increasing in complexity to form a rich sonic counterpoint.

Jean Piché takes a diverging streams approach to the extreme in his work. In *Australes* (2011) there are no causal relationships between the audiovisual media. Piché suggests that any syncretic relationship is distracting and breaks any

---

<sup>40</sup> Bill Alves, "Consonance and Dissonance in Visual Music," *Organised Sound* 17, no. 2 (2012): p. 117.

<sup>41</sup> Evans, "Foundations of a Visual Music," p. 11.

<sup>42</sup> Alves, "Digital Harmony of Sound and Light," p. 51.

higher metaphorical link between the two streams.<sup>43</sup> I have already introduced the proposition that purely syncretic relationships are somewhat simplistic and have garnered the somewhat pejorative term 'Micky Mousing'<sup>44</sup> from some quarters. As part of an association strategy, however, I believe the syncretic paradigm to be a useful one. As mentioned above, *Insomnia* makes great use of creating an initial syncretic object, then allowing the streams to quickly separate, creating a sense of space within the work and this is further developed in the interlude section of *Xpressions*. It is perhaps interesting to note that Piché believes the soundtracks to his works should be able to exist outside of the audiovisual paradigm.<sup>45</sup> However this concept diverges greatly from the synergetic audiovisual whole which I seek to create in this portfolio: the works cannot be considered complete when one stream is missing.

Bret Battey's *Luna* series of works "...explore the potentials of editless composition, in which a single audio or visual process is modulated from beginning to end without cuts or splices."<sup>46</sup> Battey argues that this creates the most beautiful and enchanting, almost hypnotic works, that seem quite at odds with their technical developmental process. *Mercurius* (2007) uses similar theoretical processes to those found in Alves' work, albeit on a much more sophisticated scale. Visuals are generated using his 'Brownian Doughnut Warper' plug-in for Apple's Motion software. This package enables still or video imagery to be broken

---

<sup>43</sup> Jean Piché, "De La Musique Et Des Images," *CIRCUIT: musiques contemporaines* 13, no. 3 (2003): p. 47.

<sup>44</sup> Cook, *Analysing Musical Multimedia*, p. 179.

<sup>45</sup> Steenhuisen, "Interview with Jean Piché."

<sup>46</sup> Bret Battey, *Mercurius* Programme Note (2009). Available at <http://www.vimeo.com/13392276> [Accessed 2 December 2012]

up into points of light, then manipulated in space and time.<sup>47</sup> For audio generation and processing Battey uses his own Pitch Curve Analysis and Composition System. This is inspired by Indian vocal music and shapes glissandi using just-tuned pitched steps of the material to be processed.<sup>48</sup> This combined process is designed to create works of complex isomorphism, based on gestalt principles. The result is a highly textural work, with a constantly evolving ebb and flow – a very different approach to achieving a similar aim to my own works, that of a synergetic whole.

Diego Garro utilises his continuum of gestural audio vs. visual association categories in *Pointes Précaires* (2003) to explore strategies with which to combine audio and video elements. The work explores several association strategies, for example the interpretive opening sequence, highly syncretic granular streams from 3:40, and parametrically mapped objects at 5:30 – 6:00.<sup>49</sup> The work is concerned primarily with gestural entities, with textural elements often linked via synchronised breaks or cuts, at 2:40 for example. This leads to a convincing audiovisual work that demonstrates successfully the creation of an '*objet audiovisuelle*'.<sup>50</sup> Garro's work is a key influence on my own research, which takes his associative continuum and develops it further to include full separation, and extends it to include textural as well as gestural associations. This is a key part of the audiovisual tripartition that I propose which advances association strategies

---

<sup>47</sup> Bret Battey, "Brownian Doughnut Warper Image Processing Demo" (2009). Available at <http://www.vimeo.com/14957896> [Accessed 2 December 2012]

<sup>48</sup> Robert Coburn, "Computation, Manipulation, Performance, Found Object: Four Composers Approach Sound and Visual Image" (paper presented at the Proceedings of Asia Computer Music Project, Tokyo, 2011), p. 12.

<sup>49</sup> This work is discussed in greater detail in Garro, "A Glow on Pythagoras' Curtain: A Composer's Perspective on Electroacoustic Music with Video.",

<sup>50</sup> See footnote 25.

further by including a parallel streams approach and the creation of audiovisual objects in a three part process.

The majority of these composers, myself included, compose using principles of consonance and dissonance, tension and release, as suggested by Evans – either explicitly as part of a compositional strategy or implicitly through a developed temporal understanding of musical harmony, discord and counterpoint. After all, this is the concern of most, if not all musics. It is the extremely varied compositional approaches to all the other factors governing audiovisual works that enable such a wide and varied range of compositions to be developed in such a relatively short time span – just over 10 years. This portfolio aims to add to this body of work and offers seven compositions that explore a new approach to audiovisual design, which is based on a tripartition of association strategies, audiovisual objects and parallel streams.

## Chapter 2: *Sketches*

*Sketches* (2007, 7:38)

### Programme Note

The title of this piece is a reflection on the work's initial compositional development. The imagery was conceived by creating and then refining a selection of 'video sketches' – moving imagery that had been developed as part of my research into video manipulation with the software environment Jitter. These initial sketches were then further explored and developed to create a palette of gestural and textural articulations and trajectories with which to compose the final piece.

The conceptual development was based on the notion of a painter sketching their initial ideas before committing the work to canvas. *Sketches* was specifically influenced by Henri Matisse's gouaches découpées works, particularly his 1953 work *The Snail*, which was created by cutting and tearing shapes from paper that had been painted with brightly coloured gouache, a pigment-based ink. These shapes were then assembled and mounted to a white background.<sup>51</sup> The visual processes and articulations are treated in a similar way to the coloured paper in *The Snail*, in that they are shaped and crafted before being arranged in time, as well as positioned on-screen.

*Sketches* expands this treatment to the audio track, exploring the sonic and visual relationship of the abstract material. An association strategy, initially focused on textural/gestural correspondences, is employed to underpin this relationship. The

---

<sup>51</sup> T Riggs, "The Snail," *Tate Gallery Artwork summary* (1998).  
<http://www.tate.org.uk/art/artworks/matisse-the-snail-t00540/text-summary> [accessed 2 December 2012]

piece utilises parametric mapping and synchronisation to create a strong sense of audiovisual identity throughout the work.

## Background

The composition *Sketches* is the result of extended research with Jitter: a software environment for the design and processing of both live and off-line video data.<sup>52</sup>

Being already familiar with Jitter's core environment Max and its associated sound engine MSP, Jitter was used as a tool to investigate as it offered a large suite of video filters and manipulation effects. Furthermore, it offered the possibility to extend this control to the audio domain via MSP integration, providing a method of parametric mapping of sound to video and vice versa. *Sketches* represents the first stage of this investigation – the creation and execution of video manipulation tools created in Jitter.

The creation of *Sketches* was prompted by a desire to focus on structuring and arranging approaches based on a visual medium. As such, the video track for this piece was created in its entirety before work on the audio commenced. This methodology represented a significant deviation from my previous audiovisual working practices, based on concurrent development of audio and video materials.

The title reflects this initial compositional development of video materials. The imagery was conceived by creating and then refining a selection of 'video sketches' – moving imagery developed in Jitter.<sup>53</sup> These initial sketches were then further explored and developed to create a palette of gestural and textural articulations and trajectories with which to compose the final piece.

---

<sup>52</sup> I use the terms live and off-line to distinguish between destructive, real-time editing (live) and non-destructive post production editing (off-line).

<sup>53</sup> Part of the Max visual programming language by Cycling74.



The conceptual development was based on the notion of a painter sketching their initial ideas before committing the work to canvas. *Sketches* was specifically influenced by Henri Matisse's *gouaches découpées* works, particularly his 1953 work *The Snail*, which was created by cutting and tearing shapes from paper that had been painted with brightly coloured gouache, a pigment-based ink. These shapes were then assembled and mounted to a white background.<sup>54</sup> Visual processes and articulations are treated in a similar way to the coloured paper in *The Snail*, in that they are shaped and crafted before being arranged in time, as well as positioned on-screen.

*Sketches* then expands this treatment to the audio track, exploring the sonic and visual relationship of the abstract material. An association strategy, initially focused on textural/gestural correspondences is employed to underpin this relationship, and utilises parametric mapping and synchronisation to create a strong sense of audiovisual identity throughout the work.

---

<sup>54</sup> Ibid.

## Compositional concerns

The compositional approach and structuring process used for *Sketches* will now be discussed.

### *Jitter software design*

Initial investigations as part of research training undertaken for the PhD led to the generation of several Jitter patches.<sup>55</sup> These were used to apply morphological manipulations to video-data with control upon the various parameters of the Jitter objects<sup>56</sup> within the patches. As the patches were developed it quickly became clear that Jitter would prove to be an extremely powerful tool in the creation of original video output that could then be used within a composition.

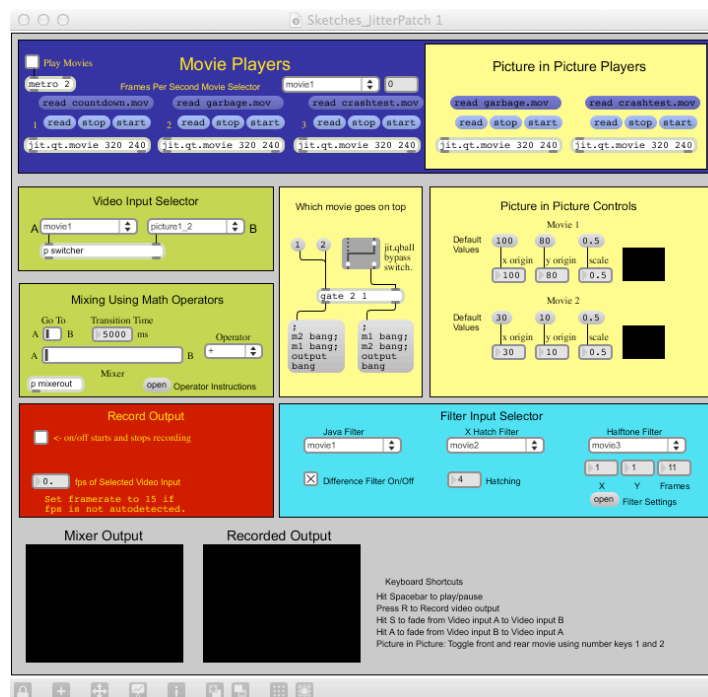


Figure 3 *Sketches\_JitterPatch1*

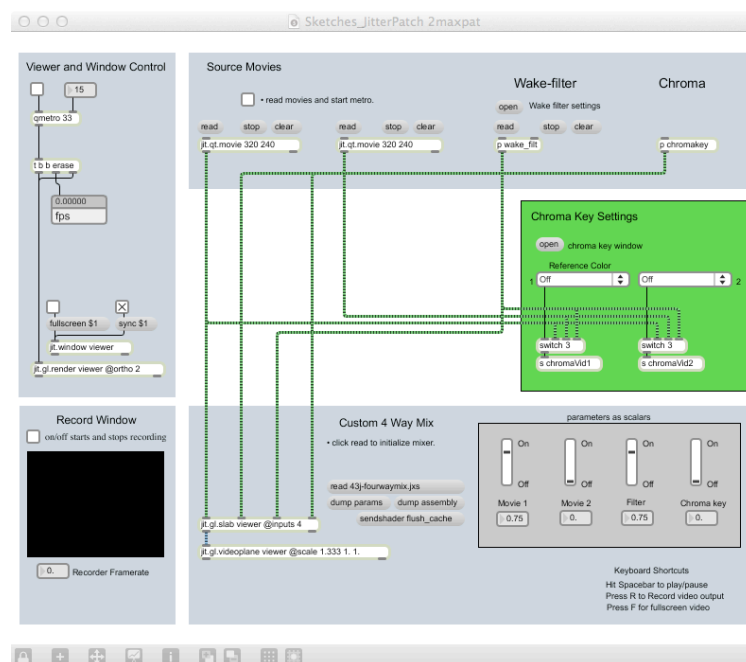
<sup>55</sup> Patches refer to applications built with Max/Jitter software.

<sup>56</sup> Each patch can contain numerous objects, which perform command functions within the application.

Video manipulation in *Sketches* was accomplished using the two patches located in the USB flash drive/optical media accompanying this commentary.

*Sketches\_JitterPatch1* (patch 1) (figure 3) employs four effects and filters: a 'picture-in-picture' effect, a Java filter, cross-hatch filter, and halftone filter. The patch is capable of mixing two feeds; the feeds can also be transitioned between using the mixing tool.<sup>57</sup>

*Sketches\_JitterPatch2* (patch 2) (figure 4) contains two video clip playback objects, a wake filter and a 'chroma key' effect, located in the top panel. Each of the 4 functions is fed to a mixing device, which is then sent to the output screen. The 'chroma key' function can be fed two inputs from either of the two video feeds, or the filter.



**Figure 4** *Sketches\_JitterPatch2*

<sup>57</sup> A full explanation of the patches will be discussed below.

## ***Signal routing***

Patch 1 uses a system of switches and gates to perform routing and mixing within the patch. Video inputs are selected in the top (blue) panel, then selected and routed to effects/output in the video input selector (green panels). The outputs are then mixed via the mixing tool, which incorporates a transition operator, adjusting the proportion/scale of the mix or transition. The mixing tool can be used either as a transition effect, dissolving from one video input to another, or can act as an opacity/mixing effect. The go-to button automates a transition from one input to the other, with the duration of the transition set (in milliseconds) via the number box, the default value being 5000ms.

Patch 2 uses a similar graphical layout to that of patch 1 in that video inputs are initialised in the 'source movies' top panel. However the mixing in this patch is controlled via the Jitter GLSL shader language object `jit.gl.slab` and a Jitter XML Shader (.jxs) file.<sup>58</sup> The object creates an efficient way of combining multiple video inputs and, in practical terms, mixes the video channels via the faders in the grey panel.

Patch 2 also has the option to route each of the source movies to the 'chroma key' effect via an input switch. This effect can then be controlled via the 'chroma key' window, which can be viewed by clicking the open button on the 'chroma key' panel (green).

---

<sup>58</sup> As it is beyond the scope of this commentary, please see the Jitter `jit.gl.slab` help file for more information and explanation of the processing and coding behind this object.

## ***Compositing effects and filters***

### **Patch 1**

This patch employs four effects and filters. First is a 'picture-in-picture' effect that is controlled by the settings in the yellow panels. The two videos used in the effect are selected with the 'picture-in-picture' players (top yellow panel). The imagery can then be manipulated via the 'picture-in-picture' controls panel.

Filters can be applied to each of the three video inputs from the (blue) movie players panel. Drop down menus are used to select the video input to be processed by each filter on the (blue) panel.

The Java filter uses Java classes to manipulate the input video data. In addition to the main filter, there is a difference operator effect that is initialised by default. This filter cannot be modified in real time therefore there are no other controls. The cross-hatch (X Hatch) filter is a standard Jitter filter that has been incorporated into this patch. The number box adjusts the hatching effect.

The halftone filter splits and rotates each of the RGB<sup>59</sup> fields and then uses the halftone filter pattern, which by default is a colour bars picture, to tint the resulting fields. Any image/movie can be used as a halftone pattern, but must be initialised before the filter will work. This can be done in the halftone filter settings window. The window opens by clicking 'open' in the (blue) filter input selector panel.

For convenience the main control settings for the filter are included on the top level of the patch. However, it is useful to use the controls within the filter settings window. The X and Y settings change the size of the halftone screen frames, whilst the frame setting adjusts the number of frames used in the matrix.

---

<sup>59</sup> Red, green, blue

As with all effects and filters, their output will only be visible in the mixer output window when they are selected as a source in the (green) Video Input Selector panel.

## **Patch 2**

This patch has two effects: a wake filter and a 'chroma key' filter.

### ***Wake filter***

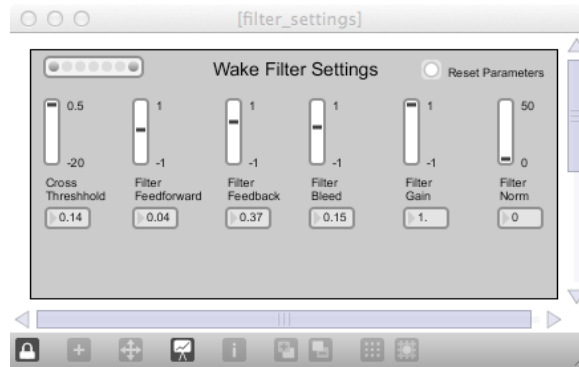
The wake filter, as used in *Sketches*, is actually created using three separate Jitter objects. These are the wake filter object (jit.wake), the RobCross object (jit.robcross), and the Brcosa object (jit.brcosa). The wake filter creates a video feedback of the input source with a convolution stage added at the end of this process. There are various controls for feedback and forward as well as individual manipulation for each colour (RGB) in the matrix. The RobCross object is an edge detection object. The Jitter help file gives this description:

This particular edge detector uses a 2x2 convolution kernel to compute a 2-dimensional spatial gradient of an incoming matrix, brightening features with "high spatial frequency" -- a large amount of change from cell to cell -- and darkening features with less change.<sup>60</sup>

Finally, the Brcosa object allows simultaneous adjustment of the brightness, contrast and saturation of an image. This object is included at the end of the signal path to allow for any adjustments/corrections to the filter's output.

---

<sup>60</sup> Jit.robcross object html help file included with Jitter version 1.6.2 (Mac OS X version).



**Figure 5 Wake filter settings**

A threshold of 0.14 was set for the RobCross edge detection object, whilst the Brcosa object's parameters were set to brightness: 1.5, contrast: 1, Saturation: 1. The RobCross threshold was set first by manipulating the input factor for the optimal aesthetic video output. The brightness was then increased from 1 to 1.5 in the Brcosa object to compensate for a reduction due to the RobCross threshold.

The main wake filter control settings can be found in the filter\_settings window of the patch, as seen in figure 5.

The sliders and number boxes in this window represent the main parameters that operate the filter. The cross threshold has already been mentioned, and operates the edge detection threshold level. Feed forward and feedback are global, controlling all the matrix plains. Bleed sets the levels for the convolution kernel values. Filter gain sets the gain level for all colours (the default is 1) and filter normalisation sets the normalisation level across all 4 plains.<sup>61</sup>

As all of the parameter values are individually controllable, the user retains a high level of filter control. Using the control window also allows for quicker editing of parameters, as well as greater intuitive explorations of video material processed

<sup>61</sup> For more information please consult the Jitter 1.7 reference manual, available at <http://www.cycling74.com/>

by the filter. The settings seen in the window were used to create the opening section of *Sketches*.

## **Chroma**

The 'chroma key' effect is a very powerful tool. It is better known as a film and television 'green screen' effect; in 'green screen' footage is shot on a matt monochrome background (usually green or blue). By applying the effect to the footage the matt colour can be extracted and a separate video feed can be introduced in its place. This is the same process used in 'chroma keying', however the effect used in the patch is able to extract any colour and superimpose a second video feed in its place. For example, in figure 6 a 'chroma key' effect was used to extract a mid-yellow colour from the red footage in section four; material from section one was then been superimposed.

The video inputs are selected using the drop-down menus in the (green) 'chroma key' settings panel. The movie with the reference colour is selected in menu 1. Traditionally this would be video shot against blue/green screen, but it is also possible to use any colour as a reference, so long as it is of a uniform hue and brightness. The effected movie is then selected in menu 2. All other settings for the 'chroma key' effect can be found in the 'chroma key' window, which is accessed by clicking the open button in the (green) panel.

The reference colour can be selected by clicking anywhere within the left video window, titled reference colour movie. Alternatively, the reference colour can be chosen from the colour swatch. Chroma threshold can then be set using the number boxes below the video input windows. The resulting video can then be mixed with the other content via the faders in the grey mixing panel.



## ***Recording***

Patch 1's record function records the output of the patch – the same output that can be seen in the mixer output window. To record the best quality output, the frame rate of the record object must be set to the input movie with the highest frame rate. This is simple to do: in the (blue) movie players panel there is a drop-down menu that will display the frame rate of each movie as it is selected. Once the movie with the highest frame rate is selected, the recorder frame rate will automatically be set.

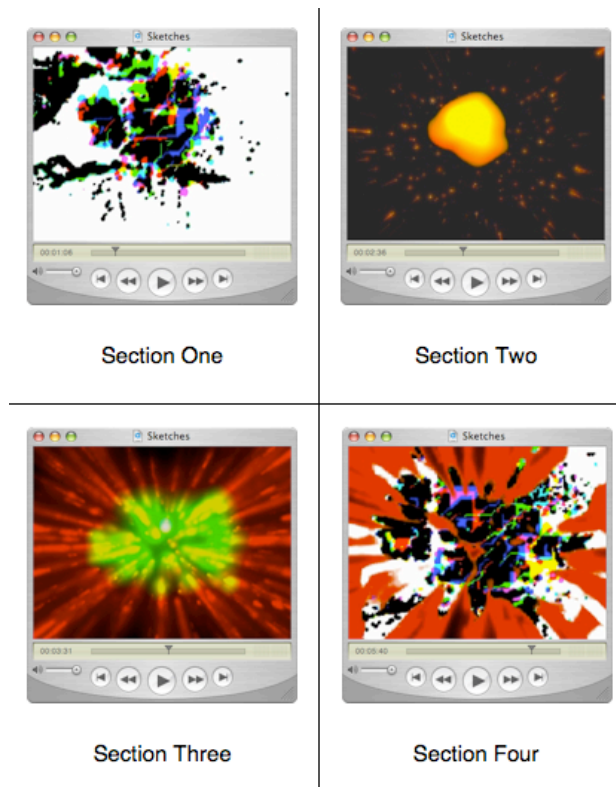
To record, simply click the record toggle in the (red) record output panel or use the keyboard shortcut R. This will open a save file menu. Give the file a name; click save, and the recording will automatically start. To stop recording simply toggle the record switch to off or press R again.

Patch 2's record function is almost identical, except that the frame rate will automatically be detected from the first source movie.

## **Description of the structure: video creation**

This section will refer exclusively to the video track, as it was fixed before the soundtrack development commenced. There are four main sections to the piece plus a coda, illustrated in figure 6. They are:

1. 0:12 – 1:42
  2. 1:43 – 3:04
  3. 3:05 – 5:37
  4. 5:38 – 6:48
- Coda. 6:49 – 7:38



**Figure 6 Four main sections of *Sketches***

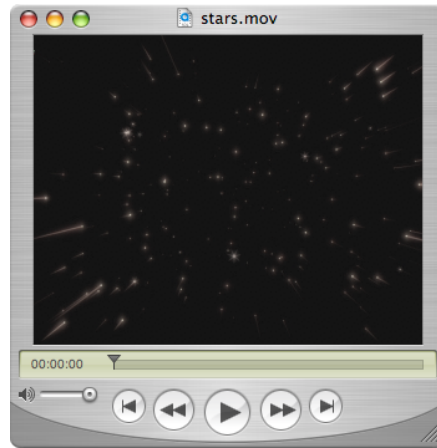
The sections were created separately and derived from each other using the two Jitter patches. The sections were then arranged to form the final video track. Section one used the wake filter with the settings outlined earlier to manipulate a video extract created in ArtMatic Pro. The resulting filtered output created an interesting multicoloured lattice-like structure within a dynamic, black object that expands and contracts from the centre of the screen. This lattice effect was similar to one created with the halftone filter used in patch 1, however the wake effect was chosen as it created greater internal motion within the centre of the circular object. The effect was achieved by manipulating the edge detection settings within the wake filter controls.<sup>62</sup>

Section two was created by reprocessing the video generated for section one with the wake filter, and combining this with a video called stars.mov (see figure 7).

Patch 2 was used to combine the video feeds using the four-way mixing tool. Filter

<sup>62</sup> Cross threshold (fader one in figure 4) was repeatedly manipulated from positive to negative.

settings were set to their defaults,<sup>63</sup> and then manipulated in real time. The fade out at the end of the section (ca. 02:52) was also achieved with the mixing tool, reducing the input levels from 100% to 0% with the mix faders.



**Figure 7 Stars.mov**

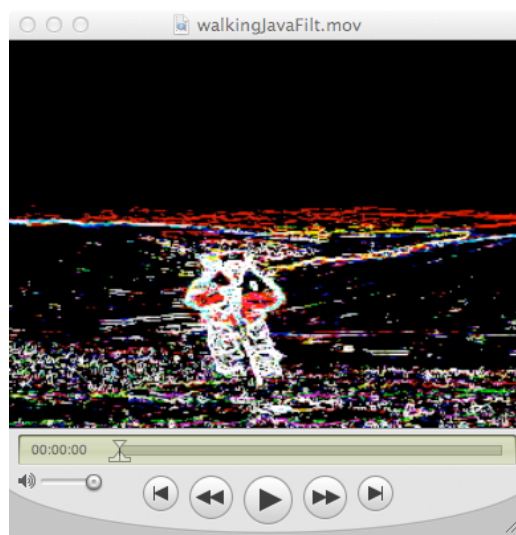
The reprocessing of materials and reapplication of video manipulations establishes a sense of cohesion and underlying structural unity throughout the piece. The progressive alterations and combinations of the two original video input materials served to both reference and repeat gestural phrases to aid this underpinning.

The third section uses stars.mov (see figure 7) processed through the wake filter and, once again, settings were derived from the defaults and adjusted in real time. The feedback and feed forward settings were manipulated to create the movement from the clouds of colour (ca. 03:22) to the streaking star 'warp speed' effect (ca. 03:35). The gain and bleed levels were also manipulated (ca. 04:25) to create the gradient changes witnessed in the final video output.

---

<sup>63</sup> It should be noted that where I have used the term default settings, these settings have in fact been coded into the Jitter application for use with this composition and are therefore specific to the type of effects intended, rather than arbitrary catchall or best-fit parameters.

The fourth section and coda were created in reverse, as the 'chroma key' effect used in section four was based on the coda material. The coda section was created using the halftone filter in patch 1. The halftone filter takes the main video input, in this case the stars.mov file, and splits it into RGBA. The filter then rotates each of the colour fields individually and tints them using a halftone pattern matrix, which can be any picture of video file. The x&y filter settings were set to 1 and the frames setting was 78. The video file imported for the halftone pattern was walkingJavaFilt.mov, a screenshot appears in figure 8.



**Figure 8 walkingJavaFilt.mov**

Section four was then created using the output of the halftone filter, by importing it into patch 2 and passing it through the wake filter. The default settings were used, with feedback and cross threshold settings being adjusted in real time (ca. 05:26) to create the adjusted gradient and star-streak effect.

Once the output of this process was recorded, it was used in the 'chroma key' effect, via the reference colour input (the affect input). The effect (resultant) input was set to the first section video, with a reference colour of approximately 255 232 34 (RGB) set using the colour swatch in the 'chroma key' effect window. This

extracted a yellow colour at the very centre of the star-streak effect and in its place added the first section footage.

### ***Assembling the video track***

Once the final video material had been generated, it was assembled using Adobe Premiere Pro. To maintain the 'Jitter aesthetic' no transformations or effects were applied to the video in Premiere. The sections were assembled using simple cuts, and beginning and end credits were added.

### ***Audio track***

The video track's creation gave clear structural parameters when preparing the audio material. The functional aspects applied to the videos were equally valid in the audio domain, that is to say the use of repetition, reprocessing, and manipulation of filtering techniques.

Association strategies are based on various degrees of synchronisation between audio and video – loose to tight – at the onset of each gesture or textural stream. Audio materials were created to match the morphological profiles of each video stream – there are never more than four visual streams present in the video track, a characteristic which is mirrored in the audio. It should be noted here that the concept of *parallel streams* was embryonic at this stage in the portfolio's development. These initial experiments sought to establish an audiovisual identity<sup>64</sup> by combining visual and auditory gestures and textures at specific points to associate a particular sound or sound set with a particular visual gesture or texture, colour or gradient. This concept of identity has been developed from

---

<sup>64</sup> A concept that will later be developed into the audiovisual tripartition.

Gorbman's 'narrative cueing',<sup>65</sup> using the principles of expectation on the part of the viewer/listener. Sweeping visual gestures require a similar sound set, as this is expected by the listener/viewer. The first section of the video, for example, opens with a visual texture that brings to mind granular clouds in the audio domain. Therefore these sounds were included in the soundtrack to reinforce this sense and create an audiovisual identity that encompasses the swirling textural image and sonic counterpart.

Section one (00:12 – 01:42) of the piece opens with granular sounds that accompany the visual texture. As the image increases in density, so does the granular cloud, becoming more noise based as the visual imagery expands to fill the screen (ca. 00:55). Frequency range also extends as the internal motion of the lattice structure increases (ca. 01:38). The granular clouds were created using the CDP Grainmil application as well as a purpose-built granulator Max/MSP patch. Max/MSP was utilised throughout the audio development as it allowed the 'sketching' of audio phrases in real time, a key element to *Sketches*' conceptual framework.

Section two (01:43 – 03:04) opens with an explosive gestural morphotype constructed from several noise based sources. The liquid sounds ca. 01:55 are constructed from the earlier granular textures, which were then cross synthesised with the enharmonic texture from the opening passage of the piece, then vocoded with the bass pulse that is originally heard at 01:16.

The star field that appears at 02:05 is accompanied by its sonic counterpart. Created using two streams of spectrally stretched electronic glitch-based noises,

---

<sup>65</sup> Claudia Gorbman, *Unheard Melodies* (London: British Film Institute, 1987). Chapter 4.

these were edited into several phrases and processed using the Doppler effect filter in Audition to create open attack-decay morphotypes with a high degree of spatial movement. These files were then added to the master 5.1 session and panned to keep their stereo image. One file was set to pan from rear left to front left, whilst the other was set to pan from rear right to front right. A surround reverb was added to thicken the texture and create a sense of space within the field, creating a convincing star field effect within the 5.1 mix.

Section three (03:05 – 05:37) is dominated by the Korg Legacy Cell synthesiser. There are two pads used, which are loosely associated with the star field and bold contoured colours of the two video streams. As the visual streams evolve – the star field dissipates whilst the bright contoured material expands – the synth pad modulates in unison. As the visual stream's brightness decreases (ca. 04:31), the sonic textures become more dissonant and are replaced by noise-based sounds.

The fourth section and coda (05:38 – end) act as a recapitulation of the previous sonic material. The synthesised textures return, as do variations of the first section's granular material.

## Chapter 3: *Mutations*

*Mutations* (2008, 11:24)

### Programme Note

This work is the second of two pieces that explore new methods of audiovisual creation and manipulation utilising the Max/Miller programming language. Building on previous research undertaken for the composition *Sketches*, Jitter patches have been developed to manipulate audio and video simultaneously via MIDI control data. These data can be used to control both video manipulation and software audio synthesisers and effects, either in real-time, or by offline processing methods.

The composition *Mutations* takes an initial video input source and applies a transformation to it. The data used to create this transformation is mapped to parameters of a software synth and used to create a parametrically mapped audiovisual identity. This process is repeated six times, once for each section of the work. However, in each section the control data is treated differently: the first three sections apply manipulations to the video file and then impose the resulting MIDI control data to an audio synthesiser. The data is then used to process a pool of audio source material via a suite of audio plug-ins. The second three sections reverse this process: six initial pad sounds are created using a software synth (two for each section). The MIDI data is then used to process the original video source. Again, this data is then used to process the same pool of audio source material. As the title suggests, this compositional process creates six mutations of the original material, which is then arranged to form the final composition.



## Background

*Mutations* progresses the audiovisual paradigm constructed in *Sketches* by exploring associative parametric mapping strategies. This composition develops Jitter patches created as part of the previous research by including a scaling MIDI control change (CC) interface. This interface is used to exchange MIDI data with an audio software synthesiser and various audio effect plug-ins to map the creative manipulations from one audiovisual domain to the other.

The piece uses Absynth structures developed from research training undertaken in additive synthesis techniques. Absynth is a semi-modular software synthesiser, created by Native Instruments Inc.<sup>66</sup> It is semi-modular in as much as the framework of the software is fixed; one can combine oscillators, filters and various modulators, but their routings are constrained, much like working with a hardware synth. The software can be used as a standalone application or incorporated into a digital audio workstation (DAW<sup>67</sup>) via the accompanying VST<sup>68</sup> instrument.

Absynth was used to generate spectro-morphologically complex sound textures and gestures via MIDI control data that was then used in real time to control the synthesis of original visual material using Jitter and vice versa.<sup>69</sup>

It is important to state that it is not the audio or video itself – rather the control data used to create and manipulate the audio/video parameters – that affects the

---

<sup>66</sup> Native Instrument's Absynth Version 4.0.1007 was used to create the accompanying examples. Each Instrument created in Absynth will be referred to as a structure. This is to distinguish them from Max/MSP/Jitter patches.

<sup>67</sup> In this document I assume DAW to primarily mean audio software with associated I/O devices. Steinberg's Cubase 5 was used to realise the examples described in this work.

<sup>68</sup> Steinberg's Virtual Studio Technology. Used to integrate software synthesizers and effects with DAWs. See <http://www.steinberg.net/> for more information.

<sup>69</sup> It should be assumed that where the commentary discusses the creation of one medium based on the other's initial MIDI data that the opposite situation is also possible. Furthermore, it is also possible to create the audio and video tracks simultaneously in real time.

creation of the material. This is the core objective of the piece; the audiovisual development is controlled by the applied processing, rather than via some algorithmic generation of one medium based on the analysis of the other. The audio and video elements are intrinsically linked through the use of MIDI CC data as, in essence, they are created from the same material.

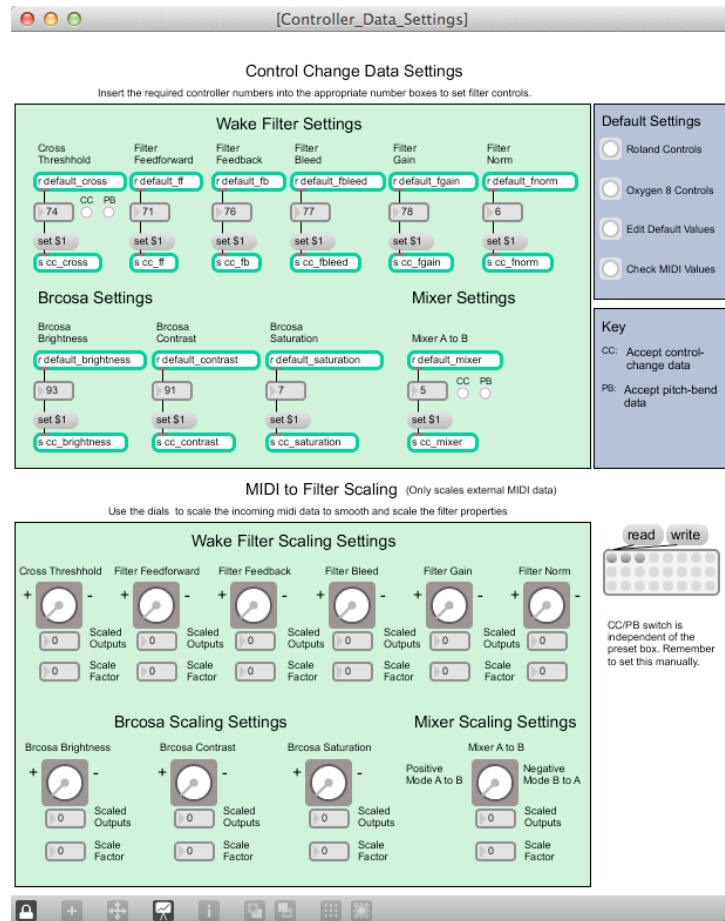
## **Compositional Concerns**

### ***Jitter***

The Jitter patch<sup>70</sup> used for *Mutations* has been developed from previous research materials. The main addition to the patch is a MIDI mapping and interpolation framework, which allows the patch to be controlled by an external instrument/controller, in this case Absynth, or to export manipulation information in the form of MIDI files for further use. The patch employs a variable mapping system that allows the patch controls to be set to the MIDI CC data output by/used to control Absynth which can be seen in figure 9. The ability to assign, store, and recall MIDI CC values to the various controls within the patch was vital due to the number of manipulations required when manipulating both Jitter and Absynth simultaneously. The CC mapping system developed here permitted near instant change to the MIDI CC message number, the scale factor applied, as well as switching the mixer and wake filter threshold controls between CC messages and pitch bend.

---

<sup>70</sup> The Jitter patch and two Absynth structure examples used in the composition have been included in the software folder of the USB flash drive/optical media accompanying this commentary.



**Figure 9 Jitter MIDI CC mapping and interpolation**

These routing options offered by the patch's MIDI CC settings allows the Jitter control parameters to be switched and/or scaled depending on the particular aesthetic and technical qualities of the output. The scaling objects were of key importance in the creation of successfully mapped audiovisual materials. They allowed the interpolation between the MIDI integer data range of 0-127 and the Jitter wake filter settings but, more specifically, scaling the input range gave a greater overall control of the effects process offered by the various software.

For example, table 1 lists the settings for the opening phrase of the piece. Initial experimentation demonstrated that a scale factor of 31 must be applied to the wake filter's cross threshold parameter when used to control Absynth's modulation parameter. This allowed the mapping of the pulsating imagery (ca. 00:20) to be mirrored in the soundtrack. The threshold scale factor was applied to allow the

subtle manipulations required to produce the video effect to be multiplied, which created a much more convincing audio counterpart.

**Table 1 *Mutations*: Initial settings for opening phrase**

<b>Absynth Control</b>	<b>Jitter Control</b>
Modulation (MIDI CC 74)	Filter Cross Threshold (Scaled by factor of 31)
Main Volume (MIDI CC 7)	Brcosa Saturation (Scaled by factor of 32)

### ***Absynth***

Absynth instrument structures are created in a patch window, similar to the one found in Max. The software allows for up to three oscillators, filters and modulation settings and a master channel modulation, filter and effect bank. The prime synthesis technique relied on creating and shaping short waveforms (less than one second), which were imported as oscillator waveforms rather than relying on the included sound bank. Morphing waveforms were also created using non-linear transformations. This created particularly interesting transients as the waveforms morphed.<sup>71</sup> These waveforms were also used to modulate the sounds and act as filters for wave shaping.

The key reason for using Absynth in this composition was the high level of control afforded by the built-in MIDI interface which enabled the structures to be manipulated using standardised signals – MIDI data. Absynth also offered powerful audio tool that required little computer processing power. This was an important consideration when searching for compatible audio software as Jitter

---

<sup>71</sup> For more information on the process see Absynth 4 User Guide, Section 3.4.3, p 30.

requires significant power when using multiple effects. It is this combination of features which makes it possible to create complex sounds based on standardised, repeatable MIDI signals that in turn can be used to map parametrically visual textures and gestures created in Jitter to the audio.

## Description of the structure

The piece is divided into six sections, outlined in table 2. The first three sections (process one) deal with video-led manipulations. They focus on the Cross threshold setting in the Jitter patch. The final three sections (process two) are audio-led manipulations, using the previous three sections' video material which is mixed and re-processed using the Cross, wake filter and Brcosa settings.

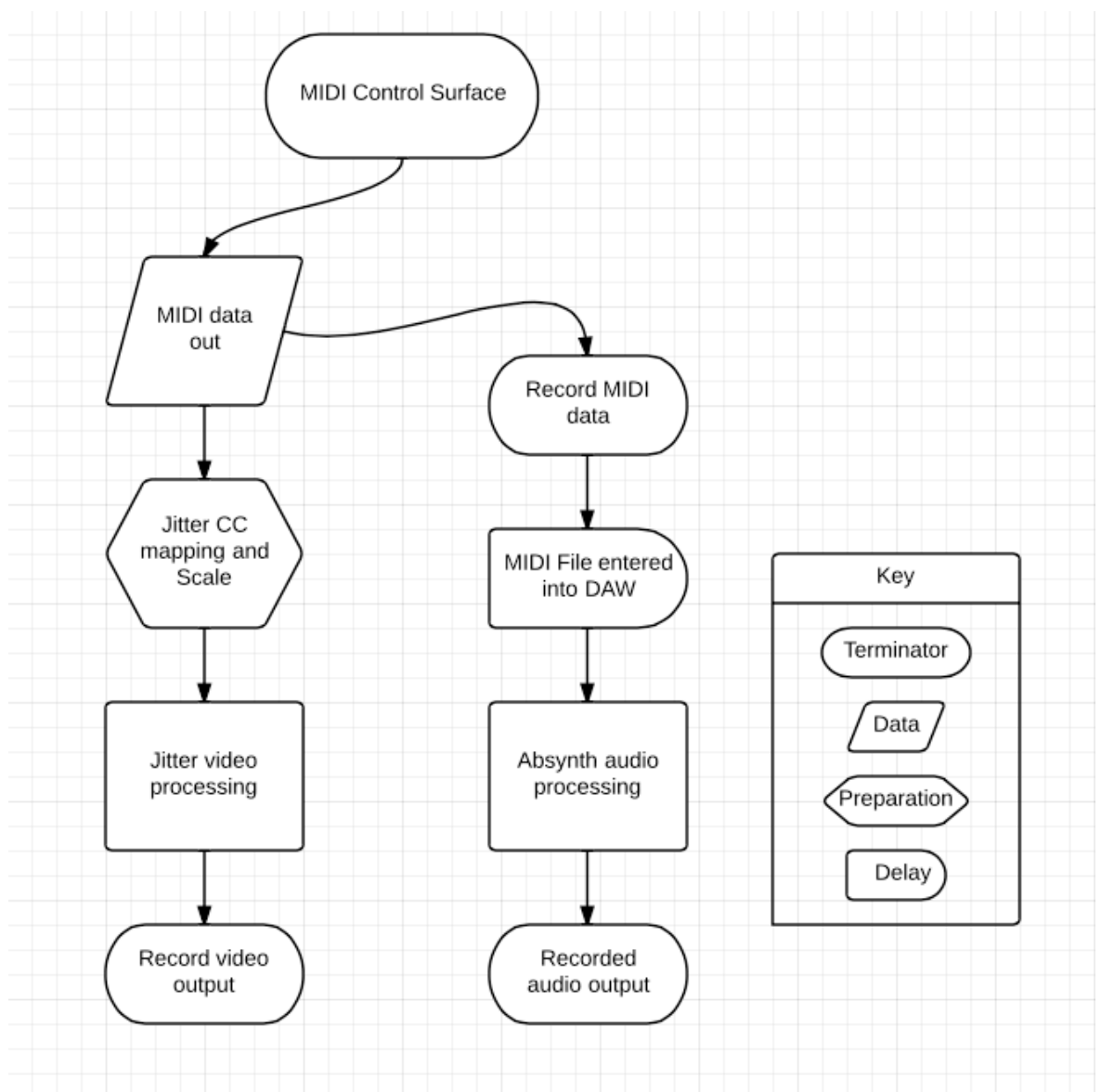
**Table 2 *Mutations'* structure with initial video settings.**

<b>Process one</b> Cross Threshold		
Settings 1 0:00 - 2:20	Settings 2 2:20 - 3:30	Settings 3 3:20 - 5:35
Cross = -0.5	Cross = -14.7	Cross= -0.72
Slow	Fast	Slow
Black/White	Blue	Red

<b>Process two</b> Wake + Cross and Brcosa		
Settings 1 Mix 5:35 - 7:55	Settings 2 Mix 7:55 - 9:06	Settings 3 Mix 9:06 - End
A-B: 800-1000	A-B: 600-1000	A-B: 400-1000
Brc_sat = -0.1	Brc_sat = +0.25	Brc_sat = +0.72

### ***Process one***

Initial video source material was loaded into the Jitter patch and manipulated in real time using a MIDI control surface assigned to the patch parameters. Both the video output and MIDI data were captured using the record object. Speed for each section is controlled using the duration number box in the main patch window. The initial cross threshold settings are indicated in table 2. The MIDI files were then imported into the DAW and used to trigger the Absynth structures.



**Figure 10 Process 1 flow chart**

## Process two

Absynth structures were manipulated via the MIDI control surface to produce the structural underpinnings of the final three sections. The MIDI data was then imported into Jitter and used to generate the final three sections. Some initial settings are given in table 2. Of particular note is the A-B mix settings, which were scaled and assigned to the pitch bend data. These data were constrained to limit

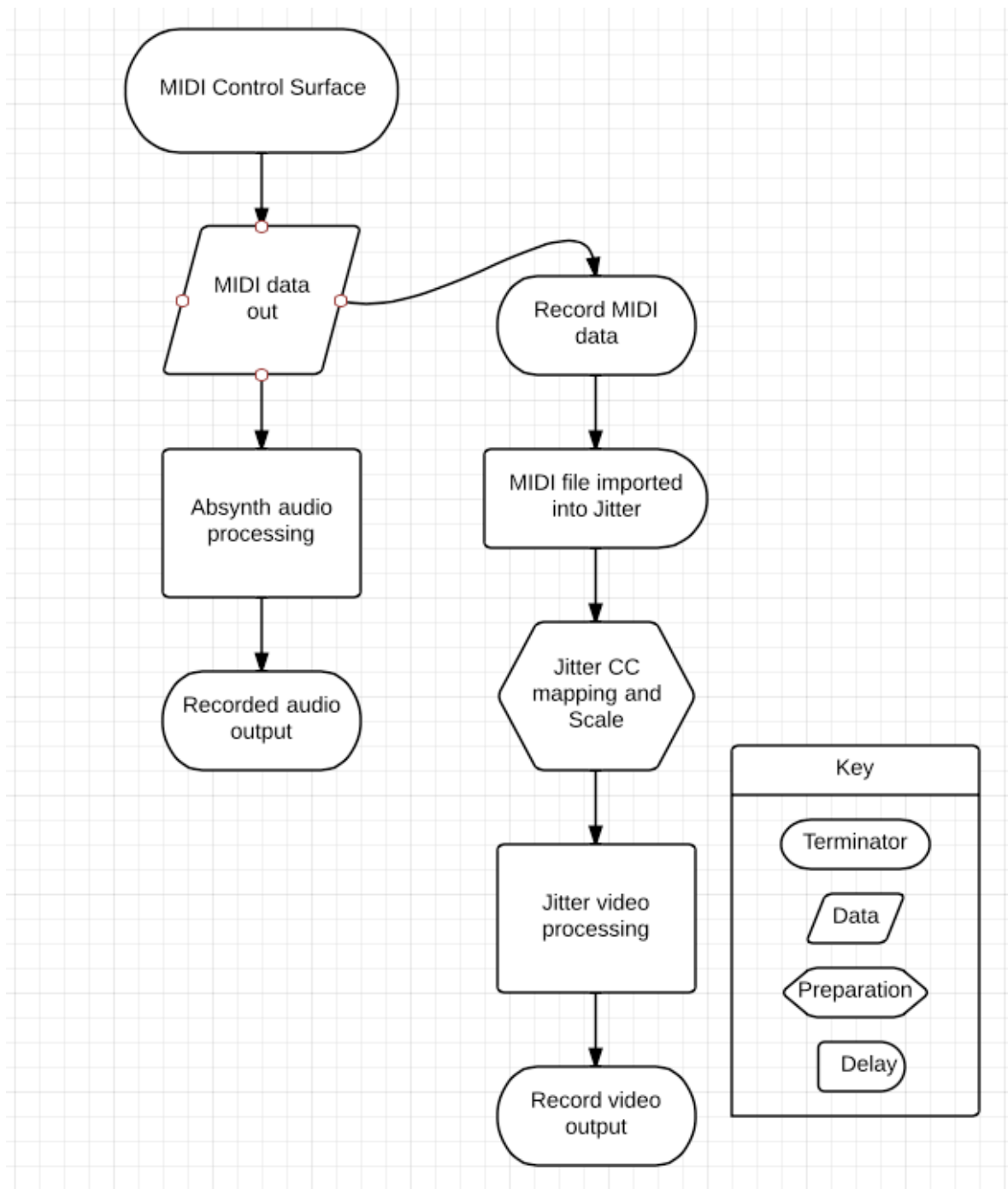


Figure 11 Process 2 flow chart

the mixing effect, creating a wake transparency effect.

These mapped streams of sounds and visuals demonstrated a strong cohesion and a sense of a definite audiovisual identity. Far from the 'Mickey Mousing' discussed in chapter one, a strong audiovisual identity is created without obvious synchronicity. Once this structural framework was completed, the MIDI data was used to aid the processing of the remaining source sounds. The sounds were arranged in the DAW and synchronised to particular visual gestural trajectories. Processes (effect plug-ins) were assigned to each individual track, with settings mapped to the CC MIDI data. This allowed the morphology of the visual and synthesised audio streams to be imprinted on the source sounds. By repeated manipulation and rearrangement of the source material, successful transformations were produced which mapped successfully to the visual stream, again creating strong audiovisual identities.



## Chapter 4: *Evolution*

*Evolution* (2009, 11:53)

### Programme Note

*Evolution* explores filtering approaches in both video and audio domains. A thirty-second video sequence was used as the video source material, with a limited palette of audio source material. Through treating, effecting, manipulating and stretching the audio and video sources, the material is developed to the point of abstraction. The resulting material is then explored to create congruent audiovisual motifs, developing the audio and video streams paradigm to consider converging / diverging approaches to audiovisual associations.

As the title suggests, the piece deals with evolving materials, which avoid rapid changes of state between a point A and point B. Associations of audio and video streams are initially based on the appearance and internal movement of the visual structures. Phrased streams of gestural elements in the audio underpin the visual structures, whilst avoiding obvious morphological links.

This concept is developed in the second section, with the audiovisual streams converging to underpin the textural drive and overall movement (ca. 4:00-4:30) of the work, before diverging to explore a textural motif, which treats the visual texture (ca. 5:20) as one of several streams that underpin the gestural morphology of the audio.

Section three explores converging visual streams, developing material from the previous sections. Imagery is now superimposed to create a new audiovisual identity.

## Background

The composition of *Evolution* was prompted by a desire to explore audiovisual associations predominantly to the left of the association continuum. This was particularly relevant in light of the previous work *Mutations*, which explored parametrically mapped material.

This piece focuses on developing the streams paradigm, mentioned briefly in the previous commentaries, to consider converging / diverging approaches to audiovisual associations. The structure is not derived from the creative process, as in *Sketches* and *Mutations*, but instead uses the streams paradigm as a framework to create congruent audiovisual motifs.

The decision to use a single video clip for this work was based on a previous exercise as an undergraduate, where I created an acousmatic piece from a single sound file. The use of a single thirty-second video source allowed me to focus on developing processes that could create an extended range of material. These processes would then be used in future compositional development.

## Compositional concerns

During the preliminary work experimentation was advanced through the use of several new processing techniques. As well as the Jitter software discussed in the previous chapters, material was processed using filters in Sony Vegas.<sup>72</sup>

The *Mutations* Jitter patch was used to create the linear objects first seen ca. 4:27 by applying the Cross threshold setting of the Wake filter to the source material.

---

<sup>72</sup> As noted in chapter one, Vegas also replaced Premiere Pro as the video assembly tool at this time.

New filtering techniques in Vegas enabled the creation of the remainder of the video materials. For example, inverting and adjusting the levels of the source file create the initial visual material. This was then duplicated and layered, with a burn compositing mode applied. This darkens the background based on the top layer's colour values. In this case it acted to subtract the colour information to create the swirls seen in section one. Another example of this process can be seen at 3:52, where an extreme zoom has also been applied.

An example of the Min/Max filter can be seen at 5:30, where it is used to pixelate the imagery to create a mosaic effect. The Gaussian blur filter is used at 6:10, first with a vertical blur applied, and on a separate file at 6:22 where a horizontal blur is added. The pinch/punch filter is used at 8:10 to reinforce the circular imagery first seen in section two.

Several compositing techniques are also applied throughout the work: the screen mode<sup>73</sup> is used ca. 6:15 to create a multiple layered effect. Layers of horizontal and vertical blurred video are combined to create a hatched effect. A similar process is applied at 8:40 to superimpose the circular imagery with a version of the initial visual motif.

The association strategy employed in the first section of *Evolution* is influenced by Prof. Dennis Miller's treatment of the audiovisual associations in *Residue* (1999). Miller creates an initial association between a rotating cube structure and streams of inharmonic pitches which appear to be unrelated. There is, however, an initial temporal association, followed by further repetitions that create a powerful

---

<sup>73</sup> Multiplies the inverse of the overlay color values with the background color values. This makes overlay colors weaker and less dominant and results in a lighter video image. Due to the black background in each layer, the process created the effect of 'weaving' the lines together.

audiovisual object.<sup>74</sup> *Evolution* exploits this temporal association strategy to link the organic shaped material to the audio streams. The repetitive nature of the streams, combined with loose syncretic associations – the appearance of the initial visual material associated with the open attack decay morphotype at the beginning of the work – at significant points offer a coherent phrasing.

## **Description of the structure**

*Evolution* is divided into three main sections, broken into subsections thus:

Section one.            00:00 – 03:51

Section two.            03:51 – 07:07

2.a            03:51 – 04:27

2.b            04:27 – 05:15

2.c            05:15 – 06:04

2.d            06:04 – 07:07

Section three.            07:07 – End

3.a            07:07 – 09:33

3.b            09:33 – 10:40

3.c            10:40 – End

**Section one**            00:00 – 03:51

Exposition of initial motifs: Association of auditory and visual streams based on gestural audio motif linked to overall movement of visual structures. Loose syncretic relationship between sonic gestures and visual structures'

---

<sup>74</sup> Diego Garro, "A Glow on Pythagoras' Curtain: A Composer's Perspective on Electroacoustic Music with Video" (paper presented at the Electroacoustic Music Studies Network Conference, Montréal, 2005), p 7.

appearance/degradation and movement (both internal and within the screen space), which become increasingly complementary.

**Section two**      03:51 – 07:07

2.a

Textural motif based on converged streams. Interpretive associative relationship allows for coincidental synchronization, whilst underpinning the overall movement.

2.b

Diverging iterative gestures in both audio and video streams.

2.c

Iterative gestures underpinned by converging textural streams.

2.d

Visual stream developed to represent a drone-like texture. ‘Ball-in-a-jar’ sounds, which explore the spatial field are subtly mapped to changes in the visual stream’s brightness and internal movement.

**Section three**      07:07 – End

3.a

Superimposition of visual streams.

3.b

Reiteration and variation of motifs from section one and two b. based on superimposition of streams.

3.c

Recapitulation of original material with multiple stream superimpositions.

## Chapter 5: *Emergence*

*Emergence* (2011, 6:26)

### Programme Note

*Emergence* focuses on concrète principles, taking several recordings of waves lapping against a shoreline at low tide, manipulating them in the studio before reconstructing them based on syntactic rules abstracted from the materials.<sup>75</sup>

The piece investigates the ebb and flow of the organic wave-like materials, which are in a constant state of change – from turbulent tidal motion to tranquil pools of reflected light. Form and shape emerge from the textural currents to create both the phrasing of individual gestures as well as the overall structure of the piece.

---

<sup>75</sup> Simon Emmerson, "The Relation of Language to Materials," in *The Language of Electroacoustic Music*, ed. Simon Emmerson (New York: Harwood Academic Publishers, 1986), p 20.

## Background

*Emergence* further develops the concrète treatment of materials experimented with in *Evolution* by exploring more cohesive source material, in the form of several recordings of waves lapping the Blackpool shoreline. *Emergence* is inspired by Denis Smalley's 1984 work *Tides*, particularly the first movement *Pools and Currents*,<sup>76</sup> in which similar aesthetic concerns are addressed, although in different ways and without the visual element.

This piece represents the conclusion of the first aesthetic theme of the research portfolio – representations of materials that appear somehow natural or organic, possessing characteristics that suggested the materials were not man-made.

## Compositional concerns

The aesthetic concept for the piece was based on the concrète development of the materials. *Emergence* uses only sound and video captured as part of preparatory work, which was then manipulated in the studio before reconstructing the auditory and visual elements based on syntactic rules abstracted from the materials.<sup>77</sup>

These are based on the energy profile of the wave materials, which are in a constant state of change – from turbulent tidal motion to tranquil pools of reflected light. At a micro level, individual gestures are created from the textural video material by exploring the trajectories of short arrangements of waves. Similarly, on a macro level the overall structure is derived from wave patterns, which at first appear random but their repetitive nature creates a motion which is explored and arranged. This is based on a conceptual framework which will be described in the structural discussion in the next section.

---

<sup>76</sup> D Smalley, "Tides," (Montréal: empreintes DIGITALes, 2000), Track 3.

<sup>77</sup> Emerson, "The Relation of Language to Materials," p 20.



One recurring example of this gestural interaction is the association of a green-hued sweeping visual structure to a short mid to high frequency spectrally sweeping audio sound (ca. 0:46).<sup>78</sup> This association, reinforced by its repetition throughout the work, creates a strong audiovisual identity, or object, based on the mutual morphologies of both media. A similar structural association occurs to conclude the first phrase at ca. 0:57, whereby the textural visual pattern recedes momentarily and is accompanied by a low frequency pulse.

The method used to develop the gestural and textural elements for the work was as follows:

- Initial captured source material was split into audio and video files.
- Audio material was divided into shoreline wave sounds, bird/wildlife sounds, waves breaking against pier, waves breaking over rocks. Any vocal/human sounds were rejected.
- Audio sources were processed using the Michael Norris 'SoundMagic Spectral' plug-ins.<sup>79</sup>
- Processed and source files were then manipulated with an edited version of the Reaktor Travelizer.<sup>80</sup>
- Banks of audio motifs were created featuring dense granular textures, spectrally stretched glissandi, and open and closed attack-decay morphotypes – one bank for each of the original source categories.
- Video materials were darkened in Sony Vegas using the levels filter.

---

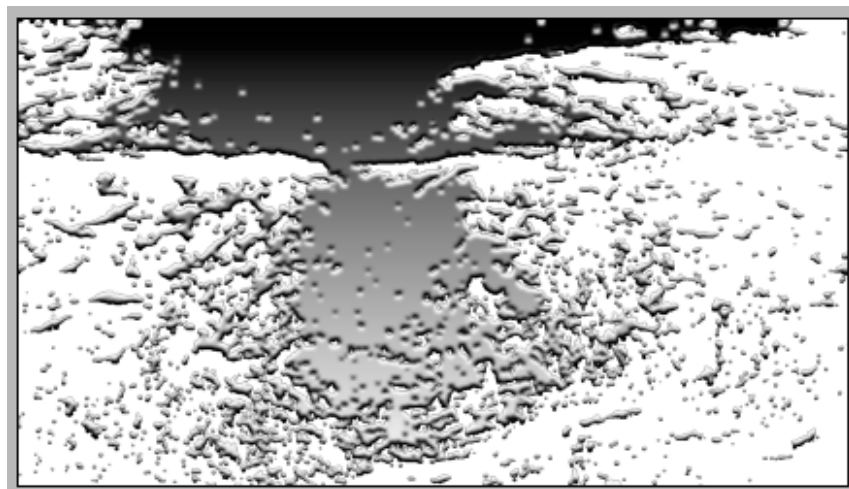
<sup>78</sup> This will be referred to as a green wave in later discussion.

<sup>79</sup> Freely available from the author's website: <http://www.michaelnorris.info/software/soundmagic-spectral.htm> [Accessed: 2 December 2012]

<sup>80</sup> Edited to remove all delay and filtering effects. Travelizer is a Reaktor instrument by Native Instruments.

- Other filtering processes were applied in Vegas, including the find edges filter, creating the visual effect (seen ca. 1:00 – 1:10), and the pinch / punch filter, used to create the circular ‘fish-eye’ effect in parts four and five.
- Source files were also processed using Adobe Photoshop CS5 extended (technique discussed below).

An innovative approach to the video design was to utilize the graphic design software Adobe Photoshop. Although not widely publicised, the ‘Extended’ versions of the software have the ability to edit and render video material. This allowed the use of very powerful filters, more usually associated with photography, to be applied to the source footage. Of particular use was the ‘Plaster’ filter. This creates a smoothing effect, but also creates a gradient background, which is useful for ‘chroma key’ effects or other compositing techniques (figure 12). The process proved somewhat analogous to a noise reduction process in the audio domain. The level of detail was reduced to a point where only the broadest outlines of the wave material were visible, allowing the repetitive nature of material to be emphasised.



**Figure 12 Waves clip after application of 'Plaster' filter.**

This filtered footage was combined with filtered source material in Vegas using the 'burn' compositing mode. The output was used to create part one of the composition, as well as forming the basis of all textural development throughout the work.

Granular synthesis processes are of particular importance within the work.

Granular audio streams offer an obvious correspondence to the textural structures in the visual domain, thus creating effluvial audiovisual objects. These act to underpin the structural nature of the merging wave patterns in both media and drive the piece forward. The passage in part one from 1:00 to 1:35, the beginning of section two, exemplifies this process: multiple streams of granular sounds slowly increase in density, mapped to an increase in colour and brightness in the visual textural stream. This process creates increasing tension that is eventually resolved by the contrapuntal exposition of the green wave object and the low frequency sound heard earlier in the section (0:46 and 0:57 respectfully). This also demonstrates clearly the two separate but complementary roles of the audiovisual object: as a structuring device and a gestural motif.

## **Description of the structure**

*Emergence* was composed, using a conceptual framework based on the anecdotal nature of the source material, in that the composition is influenced and guided by the direct experience/expression of the movement of water in the recorded material. Various aspects of the waves and shoreline are explored, which are noted below. Associations are based on textural elements that are driven by the

motion of the waves. From this textural flow emerge gestural motifs that act as 'currents' to propel the work forward.<sup>81</sup>

**Part 1: Currents** 0:00 – 2:12

Emerging patterns are created through conflicting currents and tides. Complex wave oscillations in the source material are used to create the loose A-B-A structure of this section. Source material was recorded as the tide was going out, when multifarious sets of larger-then smaller-then larger waves were created by the tidal flow. This energy profile was used to phrase individual gestures as well as the overall structure.

Part 1.a 0:00 – 1:30

Establishment of textural underpinning and exposition of emerging initial gestural motifs.

Part 1.b 1:30 – 1:52

Breakdown of textural flow into single iterative streams of audio and video media.

Part 1.c 1:52 – 2:12

Reestablishment of textural motion: green wave audiovisual gesture triggers multiple high spectral glissandi streams that combine with granular texture in both audio and video, thus reestablishing the textural audiovisual object.

---

<sup>81</sup> A notion borrowed from Smalley.

**Part 2: Ripples and pools** 2:12 – 3:30

Part two's conceptualization was based on the receding tide revealing small rock pools. The water in these pools remains still, until being disturbed by a set of waves.

As the visual motion from 1.c stabilizes, forming a 'pool' structure with constrained internal motion, audio streams are reduced.

A gestural audio stream at 2:30 breaks this structure, triggering gestural iterations in the audio and increased movement around the screen in the video. Note that whilst the motion and trajectories of the audio and video streams are similar, there is no intentional synchresis at this point.<sup>82</sup>

The video material slowly evolves into a repetitive pattern with a corresponding low frequency glissando (ca. 2:55). Repetition is formalized at 3:00 by a short audio gesture that triggers a corresponding ostinato associated with the visual 'ripple' effect.

The visual streams slowly coalesce, with high frequency gestures developing as the visual streams stabilize in the top right of the screen space.

**Part 3: Stillness** 3:30 – 4:17

A break section that considers the rock pool once the tide has retreated further down the shore.

---

<sup>82</sup> That is to say, there has been no attempt made to synchronise the streams, however there are several points between 2:35 and 2:50 where the streams suggest a syncretic association.

The streams coalesce into a static audiovisual object, which acts to counterpoint the previous sections. High frequency trills are loosely associated with the ripple movement of the contoured outline of the imagery. These then diverge at 4:00 where internal movement in the visual structure increases and is released in multiple waves with an associated increase multiple rising glissandi audio stream density.

***Part 4: Whirlpools***      4:17 – 5:10

Converging eddies create swirls of water, reflected in the bright sunlight, as the tide begins to turn.

This section represents one large object in constant motion. It serves to recapitulate previous material and functions as an audiovisual object precept. Multiple audio streams converge / diverge over the visual structure, with the reintroduction of the 'green wave' audiovisual object as well as a retrograde of the gestural object found at 0:57.

***Part 5: Low tide: coda***      5:10 – End

The representation of slack water: low tide brings momentary calm to the constant movement of the waves, before the tide advances and the tidal pattern repeats.

This section acts as a coda, reprising materials (now filtered: spectrally in the audio domain, with the Vegas 'pinch/punch' filter in the visual) found in part three. The fade at the end of the piece is used to represent and emphasize the ongoing motion of the wave-like structures, suggesting an infinity to the audiovisual world created.

## Chapter 6: *Insomnia*

*Insomnia* (2012, 5:05)

### Programme note

*"O sleep, O gentle sleep,  
Nature's soft nurse, how have I frightened thee,  
That thou no more will weigh my eyelids down,  
And steep my senses in forgetfulness?"*

Henry IV, Part II, Act III, Scene I

## Background

*Insomnia* is the first of three works based on the second aesthetic theme seen in the portfolio: representations of materials that appear somehow man made, or unnatural. *Insomnia* moves away from anything curved and organic and into the realm of the linear and iterative.

The exposition of the linear material in the piece is inspired by Kapuscinski's *Mondrian Variations* (1992, 2011).<sup>83</sup> In Kapuscinski's work "Visual and musical events are synchronized to create an integrated audio-visual counterpoint."<sup>84</sup> *Mondrian Variations* is propelled by the ever-increasing number of junctions formed by the movement of the linear material and their increasingly complex sonic iteration. *Insomnia* treats the audiovisual events in a similar way, although both the materials and outcomes are very different.

*Insomnia* explores the initial articulation of the material by deliberate synchretic association of only some of the lines and linear intersections within the work. Initial iteration is used to create an expectation for the viewer that all lines/intersecting points will be articulated sonically; however, by deliberately choosing not to articulate some of the intersecting points, a tension/suspense is created.

---

<sup>83</sup> Kapuscinski, Jaroslaw. *Mondrian Variations*, computer animated video 1992. Programme note available at <http://www.jaroslawkapuscinski.com/work-mondrianvariations.html> [Accessed 10 October 2012]

<sup>84</sup> *Ibid.*



## **Compositional concerns**

Insomnia advances the parallel streams paradigm in its approach to the materials and develops increasingly divergent audiovisual correspondences. The work treats the visuals as a contrapuntal stream to the increasingly complex sonic morphologies. Kapuscinski's manipulation of the Mondrian works to create a playful motion between intersecting lines was an important influence in both the creation and treatment of the two streams.

The title and conceptual framework for the piece are based on a loose narrative related to the inability to sleep.<sup>85</sup> Nominally, the textural background represents sleep, whilst the line objects represent interference. A further discussion of the conceptual framework is included in the description of the structure.

Structurally, the work uses initial associations to propel the streams in increasingly divergent trajectories, to the point where the visual streams act as a counterpoint or background. Thereby creating a visual 'drone' to the increasing sonic complexity and density that tends to the foreground.

## ***Video creation***

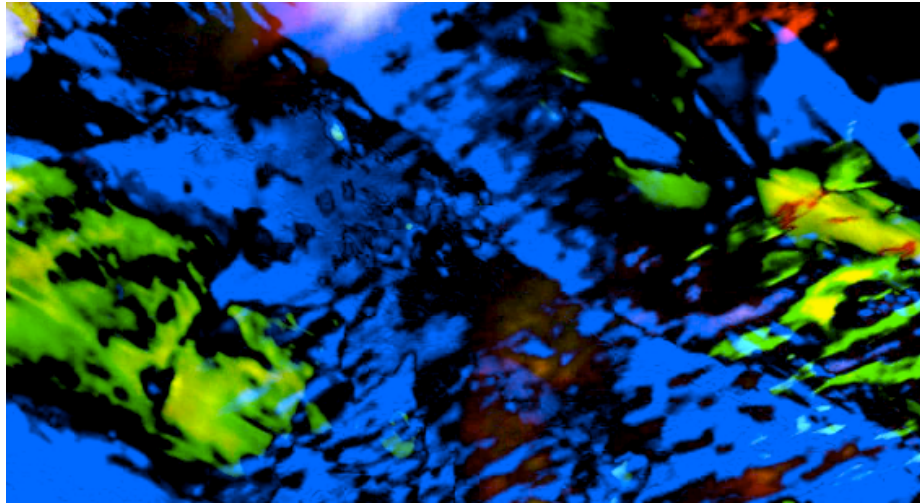
The video track was created exclusively in Sony Vegas. Using a levels filter with its settings pushed to the extreme, two source files<sup>86</sup> were repeatedly over-saturated. They were then superimposed using a 'difference level' compositing

---

<sup>85</sup> Sleep seems completely unobtainable and the more one tries to force oneself to sleep, the more futile the attempts become. When exhaustion finally overcomes the body, and a light sleep descends, one is abruptly woken by a nightmare or disturbance (a loud noise, increase in volume on the television etc.). This conceptual narrative is based on the author's own experiences rather than any structured analysis on sleep disorders. The piece offers no explicit commentary on insomnia, nor does it attempt to document a physiological process.

<sup>86</sup> A slow pan around a dead sunflower and a small waterfall. The aesthetic relation between the material and the piece is the morphological attributes of the video material. But, more importantly, in terms of the overall portfolio theme, the manipulations can be seen as imposing a human influence on the organic material.

setting,<sup>87</sup> with the clip of the waterfall slowly rotated by 180 degrees to allow the sense of direction to change constantly. The resulting output (figure 13) was used to generate the remainder of the visual material.



**Figure 13 *Insomnia* filtered and composited source material.**

The series of interwoven lines was created by masking the manipulated source material with two black bars to create one vertical and one horizontal line. The black bars were animated to ‘scroll’ across the screen, which creates the appearance of the lines travelling around the screen space. This process was repeated several times to create a bank of lines (figure 14). These were then superimposed using a ‘screen’<sup>88</sup> compositing mode to create the increasing iteration and density. This was an extremely time consuming process of manipulation: rendering the output, then reimporting and superimposing more lines. Several of these files were then combined to create the woven texture seen in figure 15.

---

<sup>87</sup> A difference map compares the overlay and background pixel by pixel and subtracts the darker color value from the lighter color to generate a new color value.

<sup>88</sup> Multiplies the inverse of the overlay color values with the background color values. This makes overlay colors weaker and less dominant and results in a lighter video image. Due to the black background in each layer, the process created the effect of ‘weaving’ the lines together. A process first explored in *Evolution*.

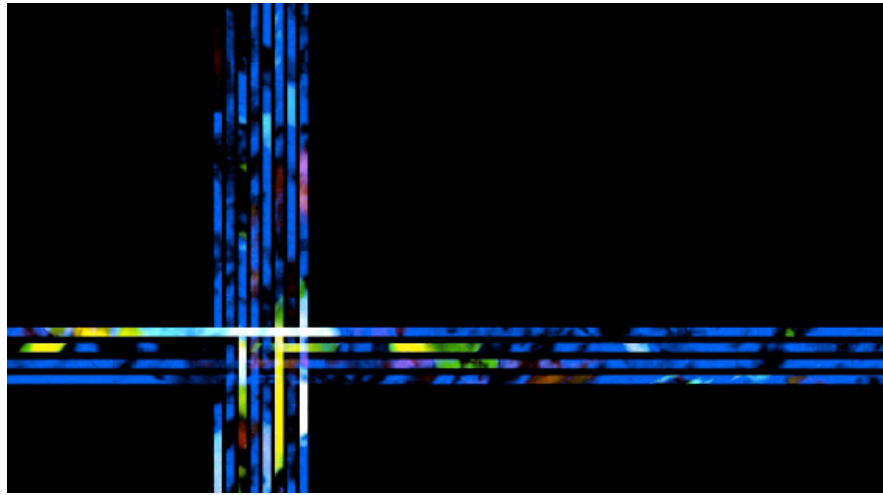


Figure 14 *Insomnia* interwoven lines.

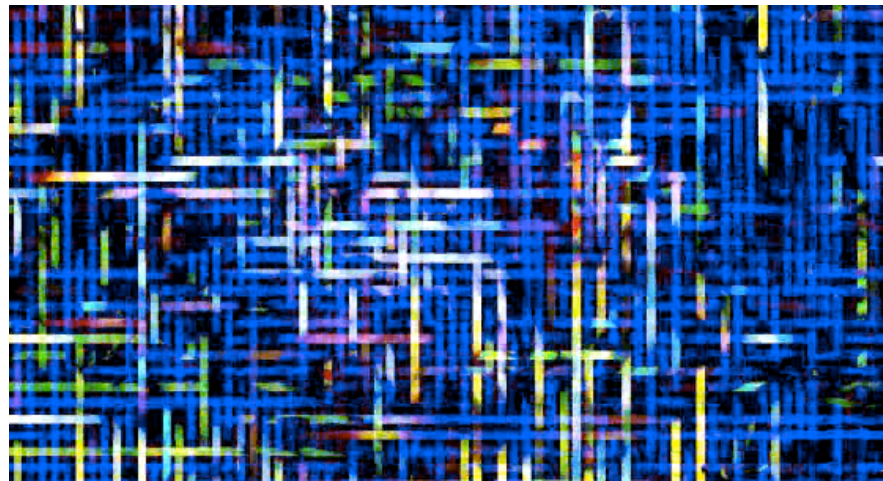


Figure 15 *Insomnia* visual texture.

This process was further expanded to create the dense visual texture seen at 3:21. By repeating the screen compositing process with the woven texture as the ‘top’ layer and the initial processed material as the ‘bottom’, the complex motion of the texture was extended.

These two elements were explored in terms of foreground and background streams by manipulating the converging line material to create ‘windows’ through which the background texture (the processed source material) would emerge. (Ca. 2:10, 3:12, 3:51 for example.)

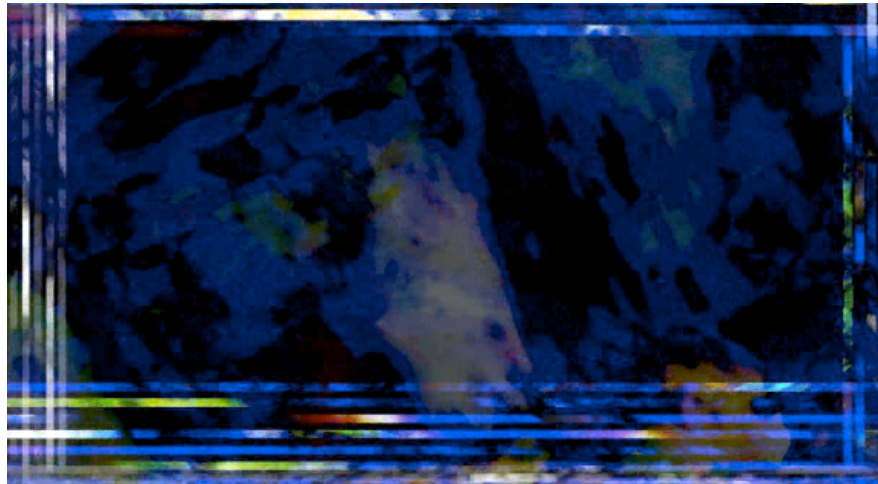


Figure 16 *Insomnia* windows.

### ***Audio creation***

In contrast to the video, varied audio source material was selected for the work.

The sound recorded during the video capture was ignored, for both technical and aesthetic reasons: audio from the sunflower clip was both aesthetically uninteresting and was distorted by wind noise. Water sounds captured in the waterfall recording have been explored extensively in *Emergence*.

Sources<sup>89</sup> were instead selected and divided into the following categories:

- Synthesised sounds
  - Recordings from a selection of KORG software and hardware synths of both sample-based and FM synthesis sounds.
- Vocal sounds
  - A recording of a child's bedtime story.
  - Laughter.
- Instrumental sounds
  - Cymbal and bell sounds.
  - Orchestral motifs.

---

<sup>89</sup> Like the video clips, the aesthetic relation between the sources and the piece is the spectro-morphological attributes of the material. However, the one exception is the vocal sounds, which were chosen to conceptualise the nightmare/disturbance state in part three.

Among the multifarious sound processing techniques used in *Insomnia*, the development of granular streams and ‘spectral drones’<sup>90</sup> are of particular importance. Unlike *Emergence*, where granular streams in both the audio and video were used as to structurally underpin the work, here they are used first as associative streams to represent the internal motion of individual lines (ca. 0:56), and later diverge representing loosely the visual background stream (ca. 2:16).

Spectral drones created from the orchestral material are initially deployed as complementary streams to the individual lines, and work to support the gestural motifs that are syncretically linked at their onsets, but are also allowed to separate as the lines traverse the screen space. Furthermore, they are used as representations of the textural visual stream that is just visible behind the lines.

The drones created from the vocal material used in part three correspond to the visual stream, which is deployed in a similar role (ca. 3:35). As the visual density increases, so does the spectral density of the drones.

Gestures with morphologies of approximately half to one second, and containing high spectral energy, were designed from the various source materials and deployed as discrete audio streams. These were used as syncretic associations with the initial lines motif in part one. Parts two and three extend this initial gestural material and develop longer phrases that separate slowly from the visual streams. The audio streams act to propel the work forward with a gesture-based morphology, based largely on rising and falling motifs.

---

<sup>90</sup> Utilising Michael Norris’ ‘SoundMagic Spectral’ plug-in Spectral DroneMaker.

## **Description of the structure**

*Insomnia* is composed employing a process whereby tension is built via increasing audiovisual density, which is only fully resolved in the final motif. The work is divided into three parts, based on the conceptual framework discussed in the previous section:

Part one: Attempting to sleep	0:00 – 2:00
Part two: Waking and dozing	2:00 – 2:47
Part three: Bad dreams – WAKE UP!	2:47 – End

### ***Part one***

Initial motifs during the title sequence act as a short introduction to the gestural content of part one. These are phrased by the titles and accompanied by an oscillating granular stream.

A device is shaped whereby initial syncretic associations are made with gestural audio motifs and the appearance of the coloured lines in the visual. These streams quickly separate with associations based on the morphological trajectories in each media. This process continues, with streams converging at points where the lines intersect, then diverging as the lines part. As mentioned in the previous section, drones are used as structural underpinning in both the audible and visual streams.

### ***Part two***

This same device is applied to part two's opening gesture, but now the streams remain separated. The device is now applied to a granular stream, this is initially associated with the 'background' visual stream, which is formally introduced.

### ***Part three***

Functional roles of audio and video streams in part three begin to switch. Where previously the streams' mutual morphological trajectories propelled the piece forward, this function is now reserved for the audio streams alone, as they continue to diverge with the video, which now fulfils the role of the audio drones in part one.

Associations are primarily limited to the 'opening and closing windows' visual motif. These are based on an extended ascending stutter theme between 2:48 – 3:20 and open attack decay morphotypes between 3:40 – 4:10.

A final syncretic association is made between the flashing imagery and associated audio stream, which acts as a device to build tension and dissonance before a final explosive resolution. The extended audio tail allows a moment's recovery after the audiovisual onslaught of the final passage.

## **Chapter 7: *Xpressions***

***Xpressions* (2012, 7:19)**

### **Programme Note**

This piece explores materials that have been pushed to the edge of abstraction.

Establishing a theme of the man-made the work deals with materials that at once are familiar and yet abstract and unnatural. The work is based on concepts of construction and the imposition of large engineered works on the landscape: transport infrastructure in general, and civilly engineered structures in particular – bridges, viaducts, tunnels, etc. The use of straight lines is advanced from previous works to create a repetitive motion that emphasises the synthetic nature of the materials.



## Background

This piece explores materials that have been pushed to the edge of abstraction. Developing the portfolio's second theme of the man-made the work deals with materials that at once are familiar and yet abstract and unnatural. The work is based on concepts of construction and the imposition of large engineered works on the landscape: transport infrastructure in general, and civilly engineered railway system in particular – bridges, viaducts, tunnels, etc. The use of straight lines is advanced from the work *Insomnia* to create a repetitive motion that emphasizes the synthetic nature of the materials.

The work stems from a boyhood fascination with the source material used – that of trains – of which I have amassed several hours of recordings over the course of this research project. The footage used includes various types of train, from fast intercity commuters to steam trains in both restored and original condition. Of particular interest are recordings of the *Shinkansen* 'Bullet' train, made whilst travelling through Japan and used to create section three in *Xpressions* (ca. 3:25 – 4:00).

As a child, summer holidays were spent in Norfolk, with regular visits to the Bure Valley Railway, "Norfolk's longest fifteen inch gauge line,"<sup>91</sup> when we would take trips from Wroxham to Aylsham, before cycling back along the path next to the railway line. *Xpressions* attempts to capture the essence of those early trips and explores the rhythmic, repetitive motion of the train.

---

<sup>91</sup>Bure Valley Railway promotional material. <http://www.bvrw.co.uk> [Accessed 2 December 2012].

## **Compositional concerns**

*Xpressions* employs primarily syncretic relationships between the audio and video streams. It develops rhythmic movement within the image which is associated in several different ways within the audio. To that end, the majority of the video track was created before the audio development progressed beyond the preparatory stage.

The video track creation sought to emulate the repetitive nature of the material by employing a repetition of techniques. Source files were processed in Sony Vegas by applying a horizontal Gaussian blur with a range setting of 0.2 which created the striated components throughout the work. This process was applied exclusively to materials for the first two sections. The other filters and effects applied were: a glow effect, a desaturation filter, RGB levels adjustment filter, clip inversion effect. These will now be discussed with reference to specific examples in the work.

Section three repeated the previous Gaussian blur combined with a glow effect which creates a 'halo' around bright objects. The effect allowed for adjustment of the following settings:

- Amount of glow: control of brightness of the halo.
- Suppression: controls the range of the halo.

These settings, along with the amount of Gaussian blur, were controlled dynamically via key frame adjustment. This allowed the blur levels to be decreased, the glow amount to increase, and the level of suppression to decrease

over time. This created the effect seen between 3:25 – 4:10, whereby the bright glow emanating from the bright white object slowly envelops the screen.

Section four uses a similar compositing technique to that employed in *Insomnia*, however this time an additive compositing mode.<sup>92</sup> Two clips were processed with the blur filter; they were then composited together, with the opacity of the top layer being slowly reduced over time. The following procedure was then applied:

- The resultant clip was rendered to a separate file.
- This new clip was imported to a new Vegas file.
- The clip was duplicated.
- The duplicate was reversed and spliced together via a cross dissolve transition.<sup>93</sup>
- The non-reversed clip was duplicated a second time.
- This was then cut into several short sections and filtered to increase its brightness (with the levels filter).
- These brightened sections were then superimposed on the palindrome clip.

The last section was developed using inverted source footage from section one.

This was processed with the blur filter and then adjusted with the 'black and white' filter to reduce the colour information by sixty percent. Opacity was slowly reduced to create the final fade to white.

Audio material was developed to mirror the repetitive morphologies of the video streams. Sounds were created in three categories:

---

<sup>92</sup> This mode combines the colour values in each composite layer.

<sup>93</sup> The visual equivalent of an audio crossfade.

1. Short gestural phrases to be associated with flashes, stutter-like imagery, sudden movement, emergence / decay.
2. Longer phrases with internal spectral motion to be associated with overall movement of the striated visual material. Textural granular streams, drones, etc.
3. Ostinatos to suggest train movement – clatter of wheels on track rails.

Examples of category 1 streams can be found at 1:50, the opening audiovisual gesture of section two. Further examples can be found at 1:56, 2:02, 2:05, and 2:14 where short gestures in the visual stream are associated with these short staccato phrases.

Category 2 phrases are particularly prevalent throughout the first section: high and mid frequency sweeping gestures at 0:30, 0:44, 1:26, and 1:35 for example.

Category 3 ostinato phrases are again prevalent throughout the work, at 2:20 for example. They are used to particular effect in section four, for example at 4:40 and 4:50.

Transformations were applied to the source material in several ways: A Max/MSP patch was used to create the majority of the category three sounds. It was designed with a series of granulator objects, a harmonizer and delay. This was used to great effect in section four. The patch was encapsulated to allow it to run as an Audio Unit plug-in within Logic Pro, which then processed the audio captured with the video source material used in that section. This allowed the

shared morphologies of the audio and video streams to be combined to produce a highly synchronized section.

Category one and two sounds were often treated with a Doppler effect in Audition. This proved particularly advantageous on sweeping gestural streams, reinforcing synthetic movement in the visual stream.

Audition's noise reduction (NR) feature was also utilized as a creative spectral subtractive synthesis tool. The spectral profile of one sound file was captured by the NR tool and then used to process a second sound file. By manipulating the amount of NR applied, the profile of the first sound is removed from the second. This created highly unpredictable but spectro-morphologically interesting results. The transients and artifacts imposed by the tool created varied and interesting effects. An example of which can be heard at 2:27, where the sonic stream has been processed using the stutter morphotype at 2:20.

## **Description of the structure**

The work is divided into five main sections which are based on the visual motivic development:

***Section one:*** 0:00 – 1:48

An exposition of the initial striated motifs, arranged from centre to right, from centre to left, then crossing. Multiple streams of audio are linked syncretically to the visual motion, with streams of sweeping granular textures triggered by iterative visual gestures. Streams of short metallic gestures are associated with the yellow

hued visual streams in the second motif as they ‘pass beyond’ the edge of the screen. These streams then combine to form the third crossing motif.

**Section two:** 1:48 – 3:12

The striated textures are taken to the extreme in this section, with the green textured object occupying the entire screen, followed by the bold streaks of second gesture that quickly evolves into an expanded, retrograde version of the second motif in section one.

Associations in section two are triggered by the onset of the visual motifs, and the quick changes in brightness and contrast that interrupt the visual motion. These are linked to sonic streams with fast morphological profiles and dissipating spectral energy. A process of applying a combination of sweeping filter and Doppler effect is employed to achieve this dispersal.

**Section three:** 3:12 – 4:10

This section serves an important structural function, acting as an interlude to the repetitive nature of the material. Strong associative links are maintained by mapping parametrically filtering techniques in the audio stream to the visual key frame process described in the compositional concerns section. Audio streams are limited to a low to mid frequency drone and a granular texture, that is mapped to the halo effect in the visual.

**Section four:**

4:10 – 6:00

This section is highly synchronized and uses exclusively the audio captured at the same time as the source video to create a strong morphological link. The video creation was discussed in the previous section where it was noted that a palindrome treatment was applied. This treatment is not directly mirrored in the audio, however the overall shape is the same with high frequency granular streams signalling the beginning and end of the section.

The fast motion in the visual streams is complemented with repeating ostinatos and bursts of noise-based streams associated with the flashes / abrupt changes in brightness.

**Section five:**

6:00 – End

The final section combines the coloured textural stream in section two with section one's morphological trajectories in both the audio and video streams. Previous audio streams are transformed via a process of spectral subtractive synthesis utilizing Adobe Audition's noise reduction filter. The relationship to the first section motifs is enhanced by the inclusion of transformed versions of the metallic audio streams first heard at ca. 1:00.

## Chapter 8: *Ut infinitio quod ultra*

*Ut infinitio quod ultra*<sup>94</sup> (2012, 13:00)

### Programme Note

For decades mankind has broken free of the Earth's gravity with the aid of rocket propulsion. Astro and aeronautical engineers have designed rockets that propel men and woman into near Earth orbit and beyond... and yet we are no nearer to reaching the stars – or even the moon – than we were in 1972 when Apollo 17 left the moon's orbit. The majority of rockets now launched into space carry more mundane payloads – satellites. Unmanned delivery systems fulfil their roll and then fall back to earth, burning up on re-entry or else crashing into (hopefully) the sea.

*Ut infinitio quod ultra* explores notions of a rocket's flight into space: Vapourous gasses escaping the craft before lift off; the lights of the tower as the rocket blasts off; its progress through the atmosphere and out into space; the quiet emptiness of space, with stars twinkling in the distance; a fiery re-entry.

---

<sup>94</sup> Or, in the words of Buzz Lightyear, "To infinity and beyond!"



## Background

For decades mankind has broken free of the Earth's gravity with the aid of rocket propulsion. Astro and aeronautical engineers have designed rockets that propel men and woman into near Earth orbit and beyond... and yet we are no nearer to reaching the stars – or even the moon – than we were in 1972 when Apollo 17 left the moon's orbit. The majority of rockets now launched into space carry more mundane payloads – satellites. Unmanned delivery systems fulfil their roll and then fall back to earth, burning up on re-entry or else crashing into (hopefully) the sea.

*Ut infinitio quod ultra* explores notions of a rocket's flight into space: vapourous gasses escaping the craft before lift off; the lights of the tower as the rocket blasts off; its progress through the atmosphere and out into space; the quiet emptiness of space, with stars twinkling in the distance; a fiery re-entry.

Back on Earth, materials for this work were more subdued, captured at a fireworks display in Betley, Cheshire. Subsequent audio recordings of sounds with similar profiles to those of the fireworks – metallic objects being struck etc. were made to create a rich and varied sound bank with which to work with.

## Compositional concerns

*Ut infinitio quod ultra* aims to be a culmination of each of the three strands in the tripartition – association strategies, audiovisual objects, parallel streams – which brings together and unifies them within a complex gesture-rich audiovisual work. It employs the full range of associations, utilising mapped objects as staging points for streams to diverge, before being reinterpreted by the changing visual gestural

morphologies – from vapourous, indistinct shapes, to rays with distinct contours; from those contoured shapes to objects with complex structures.

To explore this multifaceted approach a deliberately rich visual canvas was created with multiple streams of gestural and textural video material. This allowed for extensive and varied use of the association continuum, with particular focus on how to treat and associate the ray material in each section.

The video source material was exploited to its fullest to create the explosive gestural articulations seen in this final work. Rather than the conventional filtering and effecting approach developed thus far, a new technique using the source material as a trigger was employed exclusively in the development of the video track.

Like *Xpressions*, this work uses Sony Vegas for all video manipulations. The source material – a recording of a firework display – was imported and treated exclusively with the ray and starburst filters. Instead of rendering the output to create the final work the video source was used to trigger the effect and then removed from the final output using a technique specific to the filter. These ‘effect-only’ clips were then rendered and used to produce the resultant video.

Section two, for example, was created using a selection of clips filtered with the above process and edited thus:

- A clip of the vapourous material (seen in the top half of the screen ca. 3:20) was further manipulated with the starburst effect.
- The colour was adjusted within the starburst filter to create the blue hue.

- The resulting imagery was compacted using the pan/crop tool to create the linear motif seen in section two.
- A further ray effect was applied to create the shards of light protruding from the linear blue material and to create the iterative motion within the line.
- This process was repeated with the starburst filter to create the interference / star-like patterns in the linear material.
- The original clip – the vapourous material – was imposed over the linear material (ca. 3:17) using an additive compositing technique.

Similar techniques were applied throughout the work to create the various gestural and textural elements, with the additive compositing mode used in the final section to overlay several clips of the contoured lines/rays.

Audio source material was selected to complement the explosive nature of the video source material. In addition to the audio recording made with the source video capture recordings of various impacts were made:

- Objects hit against metal and plastic pipes.
- Objects hit against car brake drum.
- Metal, concrete and plastic objects thrown on the ground.
- Large sheet being hit with tennis racquet.
- Flapping sound of a large sheet being shaken.

The initial rays motif in section one is voiced by granulated transformations of a sheet being hit and shaken.

Brake drum 'hits' were used as the source material for sections four and five.

Individual impacts were time stretched and pitch shifted, before being spectrally stretched. These were then processed using the granular synthesis objects in the

Max/MSP patch used in *Xpressions* to create the inharmonic notes passage (ca. 6:44 – 8:00).

The metallic impacts were of particular use throughout the work. Examples of spectrally pipe impacts can be heard throughout section five and six, where they were used in association with the flashing imagery.

## **Description of the structure**

The work is divided into seven sections, each exploring a particular aspect of the audiovisual tripartition. The shape of the work overall can be described in terms of the associations used: the association continuum is explored broadly from left to right – from separation to mapped relationships – before returning to more intuitive / complementary associations.

### ***Section one:*** 0:00 – 1:54

An introduction of the ray material, initially as indistinct vaporous structures, these are replaced by an emerging lattice of twinkling rays which intersect, creating nebulous radiances.

Streams of ray structures and short audio gestures are linked primarily via a complementary association strategy, whereby audio streams are shaped to match the morphological profiles of the vaporous structures, but are not completely synchronised. This is the case for all but fifteen of the rays which were selected due to their bold gestural movement or brightness. These rays share a synchretic association with their audio streams. For example, at 0:40 a short sweeping gesture is mirrored in both audio and video streams and at 1:00 a sudden, but

subtle, increase in brightness caused by the collision of two rays is linked in the audio by a staccato noise-based gesture.

The nebulous radiances at the end of the section employ the parallel streams approach to associate separate audio and video streams that are related exclusively by their perceived mutual characteristics. The audio streams are divided into an effluvial materials that are designed to follow the movement of the lattice structure, and mid to high frequency metallic textures that were crafted to suggest the bright, radiant features of the intersections. In this instance the combined streams create a structural audiovisual object of the type described in chapter one.

***Section two:*** 1:54 – 4:06

This section explores increasingly syncretic associations between the audio and video. Granular audio streams are initially linked with the blue iterations in the video. A descending gesture is linked syncretically to the blue iterations as they diverge and exit to the left and right of the screen space at 2:00. This process is repeated at 2:08.

Low frequency pulses are also linked syncretically to the bright flashes emanating from the contiguous sections of the linear structure at the centre of the screen, examples can be found at 1:58, 2:11, 2:21 and 2:53. A similar treatment is applied to the bright bars and blue circular entity that appears accompanied by a high frequency texture that decays rather more slowly, and at a much lower amplitude than the low pulses.

The low frequency pulse stream becomes more intuitively associated from 3:17, where a visual stream of vaporous material enters. The audio stream undergoes a transformation whereby it is stretched, delayed and a surround reverb applied to create a low frequency texture functions to underpin the increasingly complex gestural material.

As the visual activity and velocity increases the audio and video streams associated with the linear material begin to diverge. The linked flashes maintain the synchresis, but these are now accompanied by longer phrased audio gestures, that develop somewhat separately from the visual stream.

The rays projecting from the central linear material are not articulated in this section. They are treated as part of, or exuding from, the central material and therefore already associated within the overall phrasing.

**Section three:** 4:06 – 6:03

Section three offers a parametrically mapped exposition of very limited material. The visual stream (ca.4:06) consists of one vaporous band that drifts across the screen from right to left and from bottom to top. The audio stream consists of a spectral drone that is passed through a resonant filter. The resonance is mapped to the changes in brightness of the visual stream. Amplitude is also increased with brightness as a result of the filtering process.

**Section four:** 6:03 – 7:55

The vaporous band in section three is transformed in this section via compositing and use of the starburst effect in Vegas (see discussion in compositional

concerns). The result is an imposition of the star-like visual material that interferes with and emerges from the vapours (ca. 6:11). Two audio streams are associated with the visuals – a low frequency ostinato and repeated short granular gestures – to create a passage of associative audiovisual objects: the predominately textural object of the transformed vaporous material and the star-like gestural objects, observed between 6:03 – 6:45.

The gestural objects paradigm is maintained from 6:45 – 7:53 where the visual streams resolve into bright star objects. A similar association process to section one is employed, whereby preliminary stars are syncretically linked to the audio material. The streams then diverge with associations based on the complementary nature of the sonic spectro-morphological characteristics and visual brightness, texture and morphology.

**Section five:** 7:55 – 9:15

This section explores the increasingly syncretic relationship between audio and video. The video is treated as two streams: the star-like ray material, previously voiced in section one, and the bright impulses. The impulse material is treated as the primary gestural stream, dominating the textural ‘background’ of the rays. A similar textural background is created with the audio, with material from the previous section transformed to act in this new role. Bursts of noise-based material are linked syncretically to the ‘starburst’ impulses, which occur throughout the section in increasing intensity.

**Section six:** 9:15 – 10:34

Highly synchretic treatment of audio and video streams: in contrast to previous sections here all visual motions have been linked to an associated audio gesture, creating a tightly synchronous set of gestural audiovisual objects.

From 10:00 this treatment degrades, with the morphological characteristics of the audio stream transforming to become more suggestive and less representational of the visual streams, hinting at the associational development for the final section.

**Section seven:** 10:34 – End

The work concludes with an increasingly dense, complex textural development of the previous material. Stream separation continues from the previous section to 11:28, where the ray material is now highly contoured and evolves in increasingly fast gestures. These visual streams are superimposed to create the complex textural pattern that continues to increase in velocity to the point of atrophy. A similar treatment is applied to the audio streams, creating a dense textural morphology. Tension is built by the layering of enharmonic tones and spectrally filtered drones, which is finally resolved by the stretched metallic gesture that is associated synchretically with the point of atrophy in the video.



## Conclusions

The objective at the outset of this PhD was the creation of a portfolio of audiovisual works that demonstrated synergy between the audio and video media. Thus I began experimentation with strategies to associate the media based on a continuum created by Garro. These initial investigations led to the identification and development of a streams paradigm, whereby parallel streams of textural, gestural, audio, and video information could be combined utilising the associations strategy, to aid in the development of an audiovisual 'identity'.

The use of Schaefferian principles and their application to both media led to the establishment of an inherently 'musical' discourse which aided the promotion of the video from mere visual commentary to a visual stream.

This identification facilitated the development of the audiovisual object paradigm. To achieve this the notion of identity was expanded from predominantly causal, gestural associations to encompass the whole associative continuum. By treating the created / manipulated audio and video materials as individual streams, audiovisual objects, could be established synergistically by a process of stream association.

These strands were formalised into a conceptual framework which I termed an audiovisual tripartition. The framework gives equal weight to audio and video streams and acts as a tool with which to develop a compositional strategy.

The tripartition offered a three-part process to audiovisual design: The association continuum acted to offer strategies with which to join audiovisual elements; by linking audio and video elements, or streams, audiovisual objects were created

which aided in the creation of a synergetic audiovisual whole; a parallel streams approach acted both to raise the visual element of a multimedia composition from mere commentary to a visual stream, and to organise materials in time.

The results from this investigation demonstrate that utilising this framework can create a synergetic audiovisual composition. The multi-strand approach offers a dynamic range of associations and it is this range that establishes the framework as a success. The research has demonstrated that the stream and object paradigms act to suggest particular associations based on the materials. Vice versa, the association strategies can be imposed via the stream and object paradigms.

Note that quite unlike the work of Jean Piché, which was discussed in chapter one, the tripartition makes it not only unwanted but impossible to remove a stream and maintain its integrity.

### **Further explorations**

The completed portfolio suggests several areas that could be advantageous for further research. The streams paradigm could be expanded on a primary level to include multiple video feeds, projected onto several screens. This could be further extended into the world of 3D, offering the possibility to create truly immersive works. This also offers obvious scope to expand the audiovisual object paradigm.

Work to date using Jitter has been concerned with real-time user control. *Ut infinitio quod ultra's* use of triggering could be expanded into Jitter environment affording new and extended possibilities in real-time and offline audiovisual mapping techniques.

## Bibliography

- Alves, Bill. "Consonance and Dissonance in Visual Music." *Organised Sound* 17, no. 2 (2012): pp. 114-19.
- — —. "Digital Harmony of Sound and Light." *Computer Music Journal* 29, no. 4 (2005): pp. 45-54.
- Chion, Michel. *Audio-Vision*. Translated by Claudia Gorbman. New York: Columbia University Press, 1994.
- Coburn, Robert. "Computation, Manipulation, Performance, Found Object: Four Composers Approach Sound and Visual Image." Paper presented at the Proceedings of Asia Computer Music Project, Tokyo, 2011.
- Collopy, Fred. "Color, Form, and Motion." *Leonardo* 33, no. 5 (2000): 355-60.
- Cook, Nicholas. *Analysing Musical Multimedia*. 2 ed. Oxford: Oxford University Press, 2004.
- Emmerson, Simon. "The Relation of Language to Materials." In *The Language of Electroacoustic Music*, edited by Simon Emmerson, 17-39. New York: Harwood Academic Publishers, 1986.
- Evans, Brian. "Foundations of a Visual Music." *Computer Music Journal* 29, no. 4 (2005): 11-24.
- Fischman, Rajmil. "Music for the Masses." *Journal of New Music Research* 23 (1994): 245-64.
- Fox-Gieg, N, Keefer, C and Schedel, M. "Editorial." *Organised Sound* 17, no. 2 (2012): 97-102.
- Garro, Diego. "From Sonic Art to Visual Music: Divergences, Convergences, Intersections." *Organised Sound* 17, no. 2 (2012): 103-13.
- — —. "A Glow on Pythagoras' Curtain: A Composer's Perspective on Electroacoustic Music with Video." Paper presented at the Electroacoustic Music Studies Network Conference, Montréal, 2005.
- — —. "Pointes Précaires." From the Video DVD anthology accompanying *Computer Music Journal* 29, no. 4 (2005).
- Gorbman, Claudia. *Unheard Melodies*. London: British Film Institute, 1987.
- Harrison, Jonty. "Diffusion: Theories and Practices, with Particular Reference to the Beast System." *eContact!* 2.4 (1999).
- Hill, Andrew. "Desarrollo De Un Lenguaje Para La Música Audiovisual Electroacústica: Investigación Sobre Su Comunicación Y Clasificación." In *En El Límite — Escritos Sobre Sonido, Música, Imagen Y Tecnología*, edited by Raúl Minsburg, 144-65: Universidad Nacional de Lanús, 2010.

- Hyde, Joseph. "*Musique Concrète* Thinking in Visual Music Practice: Audiovisual Silence and Noise, Reduced Listening and Visual Suspension." *Organised Sound* 17, no. 2 (2012): 170-78.
- McDonnell, M. "Visual Music - a Composition of the Things Themselves."  
[http://tcd.academia.edu/MauraMcDonnell/Papers/521371/Visual\\_Music\\_-\\_A\\_Composition\\_Of\\_The\\_Things\\_Themselves](http://tcd.academia.edu/MauraMcDonnell/Papers/521371/Visual_Music_-_A_Composition_Of_The_Things_Themselves).
- Nattiez, Jean-Jacques. *Music and Discourse: Toward a Semiology of Music*  
 Translated by Carolyn Abbate. New Jersey Princeton University Press, 1990.
- Otondo, Felipe. "Some Considerations for Spatial Design and Concert Projection with Surround 5.1." *eContact!* 8.3 (2005).
- Pellegrino, Ronald. *The Electronic Arts of Sound and Light*. New York: Van Nostrand Reinhold Company 1983.
- — —. *Emergent Music and Visual Music: Inside Studies*. Denton: Electronic Arts Productions, 2009.
- Piché, Jean. "De La Musique Et Des Images." *CIRCUIT: musiques contemporaines* 13, no. 3 (2003): pp. 41-50.
- Riggs, T. "The Snail." *Tate Gallery Artwork summary* (1998).
- Rumsey, Francis. *Spatial Audio*, Music Technology. Oxford: Focal Press, 2001.
- Smalley, D. "Tides." from the album "Sources / Scènes.". Montréal: empreintes DIGITALes, 2000.
- Smalley, Denis. "Spectro-Morphology and Structuring Processes." In *The Language of Electroacoustic Music*, edited by Simon Emmerson, 61-93. New York: Harwood Academic Publishers, 1986.
- Steenhuisen, Paul. "Interview with Jean Piché."  
<http://www.jeanpiche.com/Textes/Interview.htm>.

# Appendix

## Surround sound setup and playback

All works in the portfolio have been encoded in 5.1 surround sound, according to the ITU-775 standard.<sup>95</sup> The channels are: left, centre, right, left surround, right surround, low frequency effect (LFE).

## Initial setup

Speaker placement should broadly follow the layout in figure 17, in line with the ITU's guidelines. The main 5 speakers should be on the same horizontal plane, with the LFE subwoofer placed centrally, preferably behind the front three speakers.

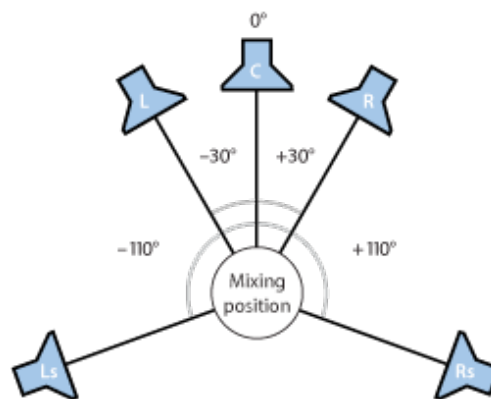


Figure 17 Surround Setup: speaker location.

## MKV file playback

A speaker setup and calibration track - PhD Surround Setup.mkv - is provided with the portfolio to allow for system setup prior to playing the works. Note that this file is only for the calibration of the MKV versions of the works submitted in the portfolio.

<sup>95</sup> International Communication Union Recommendation ITU-R BS.775-2, available [http://www.itu.int/dms\\_pubrec/itu-r/rec/bs/R-REC-BS.775-2-200607-S!!PDF-E.pdf](http://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.775-2-200607-S!!PDF-E.pdf) [Accessed 21 December 2012]

To change the surround configuration on an Apple Mac running OS 10.8:

- Go to /Applications/Utilities/Audio MIDI Setup.app
- Choose the correct sound card in the left panel.
- In the right panel select the output tab.
- Click 'configure speakers'.
- From the new window choose multichannel, then from the drop down box choose 5.1 Surround.
- Assign the correct channels based on the setup track.

On a Windows PC, configuration will most likely be through the Control Panel, and/or the system sound card software.

Channels are usually assigned as follows:

- Left = 1
- Right = 2
- Left Surround = 3
- Right Surround = 4
- Centre = 5
- LFE = 6

## **Speaker Calibration**

The second section of the setup and configuration track has bursts of pink noise to allow for the balancing of speaker outputs. All 5 main speaker levels should be balanced to achieve an identical output.

## **Subwoofer level**

The subwoofer has been used as a low frequency effect channel, to emphasize frequencies between ~ 20Hz - 120Hz, with the five main outputs having a full frequency range of between ~ 20Hz - 20KHz, as per the ITU-775 standard. Note that on some sub-satellite systems, such as the Genelec 8030A/7070A system, the subwoofer acts both to extend the satellites' frequency range and as a discrete LFE channel, which can cause an over-emphasis of the low frequencies. It may therefore be advantageous to adjust either the LFE channel output or the satellite crossover frequency range to remedy this.

## **Playback**

A device capable of playing MKV files with a 720p video track, encoded with the H.264 codec, and a 5.1 24bit 44.1KHz audio track (.wav), will be required. VLC media player has been provided with the portfolio of works, which should install and run on most modern computers (within the last 3 years). Installers for Windows and Mac OS X can be found in the VLC media player folder on the USB/optical media.