Soma: live performance where congruent musical, visual, and proprioceptive stimuli fuse to form a combined aesthetic narrative

Ilias Bergstrom

UCL

A thesis submitted for the degree of

Doctor of Philosophy

2010

I, Ilias Bergstrom, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Artists and scientists have long had an interest in the relationship between music and visual art. Today, many occupy themselves with correlated animation and music, called 'visual music'. Established tools and paradigms for performing live visual music however, have several limitations:

Virtually no user interface exists, with an expressivity comparable to live musical performance.

Mappings between music and visuals are typically reduced to the music's beat and amplitude being statically associated to the visuals, disallowing close audiovisual congruence, tension and release, and suspended expectation in narratives.

Collaborative performance, common in other live art, is mostly absent due to technical limitations.

Preparing or improvising performances is complicated, often requiring software development.

This thesis addresses these, through a transdisciplinary integration of findings from several research areas, detailing the resulting ideas, and their implementation in a novel system:

Musical instruments are used as the primary control data source, accurately encoding all musical gestures of each performer. The advanced embodied knowledge musicians have of their instruments, allows increased expressivity, the full control data bandwidth allows high mapping complexity, while musicians' collaborative performance familiarity may translate to visual music performance.

The conduct of Mutable Mapping, gradually creating, destroying and altering mappings, may allow for a narrative in mapping during performance.

The art form of Soma, in which correlated auditory, visual and proprioceptive stimulus form a combined narrative, builds on knowledge that performers and audiences are more engaged in performance requiring advanced motor knowledge, and when congruent percepts across modalities coincide.

Preparing and improvising is simplified, through re-adapting the Processing programming language for artists to behave as a plug-in API, thus encapsulating complexity in modules, which may be dynamically layered during performance.

Design research methodology is employed during development and evaluation, while introducing the additional viewpoint of ethnography during evaluation, engaging musicians, audience and visuals performers.

Acknowledgments

Firstly, I wish to thank my supervisors, Beau Lotto and Anthony Steed, for taking on my project, and for helping me through with the work presented in this thesis.

I am further grateful to have been surrounded by very interesting people at the Lottolab. Without Lefkothea Andreou, Daniel Hulme, David Malkin, David Corney, Erwan LeMartelot, Martina Wicklein, Rich Clarke, Madi Boyd, Emma Byrne, Sam Walker, and Katherine Bash, the time spent working in the lab would have been far less stimulating.

Many thanks are due to Nadia Berthouze, Nick Bryan-Kinns, Alan Blackwell, Marcelo Wanderley, Katherine Stevens, Nigel Jenkins, Olivier Ruellet, Mauritius Seeger, Mat Hourteillan, Alison Hardcastle, Anne-Marie Preston, Illi Adatto, Dor Wand, the members of Accretion and of the Gathering, for their invaluable advice and assistance towards completing the work presented here.

I owe my gratitude to my family, without the support of which I would never have had the chance of writing these lines.

And last but not least, Giulia!

Table of Contents

Table of Figures

"Colours may mutually relate like musical concords, for their pleasantest arrangements, like those concords, mutually proportionate". Aristotle.

1. Introduction

In history, several machines have been constructed to explore the relationship between music and visual art, and prove/disprove theories on the topic, the first known mention being the *"Clavecin oculaire"* (1734), proposed by Louis-Bertrand Castel, implementing a modified version of the note to colour correspondence proposed by Isaac Newton. Many have since followed, made to either accompany music with colour, or provide a form of visual music - named "*Lumia*". The term lumia was coined by pioneer Thomas Wilfred, developer of the "*Clavilux*" colour-organ (1922). Wilfred, rejected the notion of an absolute correspondence between sound and image, and concentrated on his art form of controlled colour, form and motion compositions, meant to stand alone, without musical accompaniment.

The immediacy with which music can communicate emotion has through time been envied by visual artists, most notably pioneer Wassily Kandinsky, who set out to try and recreate it in painting (Dabrowski, 1997). A great inspiration of his was the composer Scriabin, a possible synaesthete (Galeyev & Vanechkina, n.d.), and likely the first to write a musical piece from the beginning intended to be accompanied by lumia. Wagner in turn was to coin the term "*Gesamtkunstwerk*", universal art where multiple forms of artistic narrative are used simultaneously in a performance, to be experienced in combination.

As technology has progressed, so too have the tools that allow the practical exploration of this relationship. Today, artists in many disparate fields occupy themselves with producing animated visual art that is correlated with music, often referred to as visual music. Visual music is defined by Evans (2005) as: "*Time-based visual imagery that establishes a temporal architecture in a way similar to absolute music. It is typically non-narrative and non-representational (although it need not be either). Visual music can be accompanied by sound but can also be silent".*

The fusion of musical and visual art, has led to the inception of *Audiovisual Composition* (Grierson, 2005), an "*artistic form which takes as its starting point the cognitive actuality of multisensory audiovisual experience*". Actual and visual music combined, fuse to become a third art form, where one is not simply accompanying the other, but they are both to be experienced as an inseparable whole.

In modern times, analogue video synthesizers, laser shows and more recently computer graphics have all been employed to accompany music in real-time. Most frequently, there is a direct coupling between music and image, with musical input being processed to partly or wholly control live procedural computer graphics. Popular examples of software functioning in this manner, which can be found on most personal computers, are the Nullsoft Advanced Visualization Studio included with the Winamp media player (http://www.nullsoft.com/free/avs), and the Apple iTunes Visualizer (http://www.apple.com/itunes/).

The most widely established live performance practice today is that of VJing. Predominantly, VJ's (Video/Visuals Jockeys) perform by mixing pre-recorded video clips together, while manipulating their playback, and applying real-time video effects. Modern hi-end VJ software also facilitates including real-time *Procedural* computer graphics, however such features see comparatively limited use as they are both restricted and complicated to invoke in relation to performing using clips of video. The term Procedural is used to refer to graphics that are generated algorithmically, rather than having been created manually or been otherwise sampled (Ebert, 2003); alternatively the terms *Generative* and *Algorithmic* are often used in place of Procedural. A common practice in VJing is to partially control the parameters of the video playback, effects and procedural graphics with values derived from a stereo audio signal. The connection between visuals and music however is limited, because the data that can be derived from this process is limited to only overall beat events, amplitude and tempo. The narrative of a VJ's performance is therefore still constructed manually by the VJ throughout the performance's duration, rather than in any way being controlled through the musical input.

The process of making connections between incoming control data, and control parameters of visual/audio synthesizers, is referred to as *Mapping* (Miranda & Wanderley, 2006). Almost exclusively in established audiovisual performance practice, and not only in VJing, the musical control data is derived from a stereo mixdown of the music, which is processed to derive the music's overall tempo and amplitude, as well as detect individual beat events. This data is then almost exclusively statically associated to control parameters of the controlled visuals.

Despite the intense interest in live audiovisual art, and advances in its associated technologies, current tools/practices are limited in multiple ways:

(i) The mappings between music and visuals are highly constrained because they are limited in complexity, and remain static over time, thus necessitating that the correlation between visuals and music always remains limited, when in fact there is much evidence that increased correlation results in a stronger experience, as will be reviewed later in this chapter.

(ii) Virtually no user interface exists, that allows controlling the performance of visual music/audiovisual art in real-time, with a level of expressivity comparable in nature or extent to that attainable in live musical performance.

(iii) The process of preparing or improvising live procedural visual music/audiovisual performances is overly complicated, compared to live musical performance, as it almost exclusively requires that artists engage in software engineering.

(iv) Collaborative performance, a very common practice in most live performance art, is mostly absent in live visual music/audiovisual performance, largely not due to artistic choice, but to technical limitations.

1.1 Limitations with current practice

In current practice, the mapping between sound and visuals is almost exclusively reduced to using the beat and amplitude of the music, statically associated to controlling parameters of the visuals.

It can however easily be argued, that increasing the detail of the correlation between visual and auditory musical events, to a higher level of synchronization than that achievable with current practice, will further encourage the unified experience of music and image. It is known that the human perceptual system is apt at detecting correlated stimuli across modalities, and fusing these into a single percept before their interpretation (Marks, 1978), (Larsson, 2005). Michel Chion et al. (1994) argue that synchronized music and/or sound provides *"added value"* to a visual narrative, defining *Synchresis* as: *"(...) the spontaneous and irresistible weld produced between a particular auditory phenomenon and visual phenomenon when they occur at the same time"*. It has been shown experimentally, that there is significant positive correlation between how closely discreet auditory and visual events are synchronized, and the perceived effectiveness of the combined audiovisual stimulus (Lipscomb, 2005): *"(...) the manner in which salient moments in the auditory and visual domains are aligned results in a significantly different perceptual response to the resulting composite"*. The term effectiveness was used by Lipscomb without further clarification, purposefully leaving it to be interpreted as widely as possible by the experiment subjects. Experiments conducted using Magnetic Resonance Imaging (MRI) scanning corroborate Lipscomb's findings, and show that the temporal window of such audio-visual interactions is approximately 100ms (Shams, Kamitani, & Shimojo, 2002).

Additionally, the benefits of allowing the mappings to be varied during the course of the performance are clear: such a development will allow the use of suspended expectations, thus also facilitating tension and release in the aesthetic narrative developed through the mapping, both known to be crucial aspects of aesthetic narratives. The cognitive process of fulfilled and suspended expectations on the contents of future perceptual stimuli is widely accepted to be involved in generating emotional states, particularly so in musical narratives (Steinbeis, Koelsch, & Sloboda, 2006). Furthermore, Steinbeis et al experimentally confirmed the effect of harmonic expectations and violations in a musical narrative, using physiological measurements. A specific use of suspended expectations is that of building up tension in the narrative, and subsequently resolving it. This is a fundamental component in all aesthetic narratives, and has also been discussed specifically within the context of visual music (Evans, 2005). Both are crucial aspects of aesthetic narratives, and very difficult to employ in the synchronization between music and visuals, if mappings between incoming control data, and control parameters of visuals remain fixed throughout the performance, as is the case in related current practice.

Furthermore, visual music performances are currently controlled using interfaces on computer screens and/or external hardware controllers (with knobs, sliders and buttons, etc., here referred to as *non-musical controllers*). The control over the performance is akin to that used in musical *conducting gesture performance*, where the original signal sources are controlled indirectly, such as when a conductor directs an orchestra with his baton, or an audio mixing engineer manipulates multi-track audio using his mixing desk. This is in contrast to direct instrumental performance, using a controller through which the performer's musical gestures are immediately translated into sound, as with the interfaces of traditional instruments. Both performance modes are of great relevance to this work, and will be further discussed in chapter 2.2, as well as throughout the thesis, but they needed to be introduced here in order to clarify a further limitation of current live visual music performance practice: when direct manipulation performance takes place, the interface used is still the same non-musical controllers, which however are highly unintuitive in this context. It can be shown how these interfaces allow for limited expressivity, due to the limited control complexity they provide, in comparison to the expressivity which actual direct manipulation instruments afford. The level of control complexity that a digital musical instrument provides is said to be a precondition to the expressivity of the instrument in question (Dobrian & Koppelman, 2006). Control complexity can furthermore directly be gauged through comparing musical controller devices (Miranda & Wanderley, 2006). It is easily derived that the great majority of non-musical controllers, provide a more limited control complexity than traditional musical instruments in direct manipulation performance. For example only two knobs can be controlled at one time by a single performer, one for each hand, thus altering only two values. In comparison, it is easy to recall several traditional instruments which allow far greater control complexity, for example controlling the pitch and velocity of striking the key on a piano, or the greatly complex interaction that a bow and strings facilitate on a violin. This distinction is relatively clear, because of the great differences between the control complexity afforded by traditional instruments, in relation to that of non-musical controllers. It would of course not be as easy to derive which instrument allows for the greatest expressivity when the differences in control complexity are smaller, such as between the violin and the piano used in the above example. Following from this argument, one can derive that non-musical controllers, due to their reduced control complexity, do not allow performers to develop as advanced a level of virtuosity as currently established musical instruments, and thus afford a much reduced level of expressivity, in comparison to that which musicians have with their instruments. When playing musical instruments, musicians exhibit the use of advanced *Enactive Knowledge*, which, as we have seen, can only occur to a much more limited extent when non-musical controllers are used. The term enactive knowledge refers to knowledge that can only be acquired and manifested through action (Varela, E. Thompson, & Rosch, 1992). Examples include dance, painting, sports, and performing music.

Not employing advanced enactive knowledge also has consequences for the audiences' experience. It has been found that: *"(...) the perceiver watching, listening to and experiencing another's motor performance, simulates the actions of the performance within the range of their own motor capabilities"* (Rodger, Issartel, & O'Modhrein, 2007). Therefore, when performers are not using controllers that demand advanced enactive knowledge, the total experience for audiences is as a direct consequence more limited. This extends to the ancillary body movements of the performer (Nusseck & Wanderley, 2009), i.e. body movements that do not directly contribute to producing a sound from the instrument, but that nonetheless *"Have an intrinsic relationship with the music, representing a link between the music and the expressive intent of the musician"*. The position that audiences are consequently more engaged in a musical performance where musicians use advanced enactive knowledge, is widely held in the music research community, albeit no experiments have been identified that confirm this position. For example, in Armstrong's PhD thesis (2006), this position is central, and is extensively reviewed and added to.

With current practice, artists preparing for a live procedural visual music performance necessarily have to engage in creating the visual instruments themselves, through engaging in software engineering at some, usually high, level of complexity. Such programs are predominantly limited in their usability, as they necessarily embody the aesthetic goals of their creator (Collopy, Jameson, & Fuhrer, 1999). Software created by one performer is thus prevented from being used by other performers to achieve their individual aesthetic goals, without extensive modification, necessarily through software engineering practice. It is furthermore not currently possible to combine the visual output of multiple pre-existing visual instruments, without again engaging in software engineering, making the above limitation further significant. Presently available tools thus unavoidably require artists to engage in software engineering in order to achieve expressing their individual aesthetic goals in real-time audiovisual performance. This is a very significant obstacle to overcome, and thus a major limitation in usability.

Finally, with current practice, virtually all live audiovisual performance allows little to no scalability in the number of performers involved, despite collaborative performance being a common practice in other artistic forms. Clear separations, where a single performer is responsible for the visuals, and another performer or group are responsible for the music, are common. But cases where two or more are responsible for the visuals are very scarce, and even scarcer are cases where audiovisual artists perform collaboratively. Only for live VJing has one previous research effort been identified (Engström, Esbjörnsson, & Juhlin, 2008), however the paradigms and designs developed in that work do not generalize to performance of procedural graphics, or to audiovisual performance. Collaborative performance is a very common occurrence in all artistic performance practice, and so can safely be assumed to be missing from the context of live audiovisual performance due to technical limitations, and not artistic choice. There are significant benefits to draw from collaborative performance that allows for *Mutual Engagement* (Bryan-Kinns & Hamilton, 2009)*: "The point at which people spark together, lose themselves in the joint action, and arrive together at a point of co-action 'where you are when you don't know where you are'* (Tufnell & Crickmay, 1990)*"*.

1.2 Proposed system

Here a system is presented for the live performance of procedural visual music / audiovisual art. It was created with the purpose of greatly improving on and extending established practice. The objectives for its creation were to increase mapping complexity, allow the mappings to vary during the performance, allow greater expressivity during the performance, to lower the technical barrier to creativity relative to current practice, to allow for improvisation, and allow for collaborative performance. Following is a detailed description of how the system introduced here is intended to address these objectives. Note that there is no one-to-one correspondence between the ideas based on which the system is designed, and the previously detailed limitations. Instead, the developed ideas work in combination towards addressing those limitations.

1.2.1 Increasing expressivity

It is proposed that musical instruments are used as the primary source of control data for the performance.

In live musical performance, a significantly greater amount of data is generated by the instruments than simply a stereo audio mixdown. First of all, separate audio signals are generated from each individual instrument, allowing the tracking of the amplitude, tempo, and detected beat events of each instrument, rather than just of all instruments lumped together in a stereo audio signal of the music, as is predominantly used in established practice. The data from processing the multichannel audio is then most importantly used in conjunction with very detailed digital procedural control data produced separately by each instrument, provided it is suitably equipped, be it either outputting data following the MIDI protocol (Miranda & Wanderley, 2006) or OSC (Matthew, 1997). These data sets are much richer and include the onset, offset, pitch and amplitude of each individual note, played by each instrument, alongside a plethora of additional data, the nature of which varies depending on the type of instrument in question. For example, on an electronic keyboard, this additional data may include the pressure by which each key is held down, *after* it was initially pressed (aftertouch); whether foot-pedals attached to the instrument are depressed or not; the amount by which the pitch-bend wheel is turned; the value of each of the knobs and sliders on the synthesizer, intended for controlling the instruments sound synthesis engine; and more.

Through using this much richer source of data, the *Musical Gestures* of the performers are therefore more accurately encoded. In this context, the term Musical Gestures refers to the actions carried out by a musician during performance (Miranda & Wanderley, 2006). Musicians are, from this increase in control data bandwidth, expected to be able to better transmit the intent of their musical gestures, and thus to a greater extent be able to usefully take advantage of the advanced embodied knowledge they have of their instruments, thus achieving an increase in expressivity over established practice.

As we saw in section 1.1, the level of control complexity that a new digital musical instrument provides is said to be a precondition to the expressivity of said instrument (Dobrian & Koppelman, 2006), although high complexity does not necessarily guarantee expressivity. Traditional/established musical instruments however can be assumed to have reached maturity in their development, their high control complexity thus guaranteeing a correspondingly high level of expressivity.

When live computer graphics (visuals) are controlled using a signal derived from live musical performance, the musician(s) performing are controlling the visuals with their instruments. A primary question that present research seeks to answer is the following: will taking advantage of the full complexity of data generated by musical instruments, instead of only the beats, tempo and amplitude deduced from a stereo audio signal, give rise to a considerable increase in the expressivity of said instruments also for the task of performing live visuals?

The subjective and highly elusive experience is desired, of the musicians to some extent controlling the visuals, as opposed to the visuals reacting to their playing.

1.2.2 Simplifying preparation for performance

To allow artists to prepare live visual music / audiovisual art performances of their own, without having to engage in software development, it was necessary to develop a new application, as suitable software was not available. The developed application has been dubbed *Mother*.

With Mother, the paradigm of mixing multiple layers of moving graphics is retained from VJing. However, these graphics are not pre-rendered video clips, but are instead the output of real-time procedural *Visual Synthesizers* (*synths*) that run in parallel within the main host application. Each visual synth is a program that renders a particular visual effect, the control parameters of which are all accessible during a performance, so that the appearance of the visual can be controllably altered over time.

Synths can be created as Processing *"sketches"* (Reas & Fry, 2007), using a Processing library provided that enables the sketch to work within the Mother host application. Alternatively, synths can come from a pre-existing library of sketches. In this way the performance of advanced real-time computer graphics by non-programmers is facilitated, through encapsulating the programming complexity in flexible modules that are easily exchanged and managed.

Mother further allows artists to forward digital control data to these visual synthesizers so as to finely control what each displays, and to dynamically rearrange, add and remove synths during the course of a performance. The parameter space of control input for the visuals therefore varies constantly, as synths are added and removed.

Preparation for performance is thus simplified on two levels. Firstly, users of the Mother software may use a combination of synths developed and distributed by other artists. Through the selection and combination process, they thus achieve a higher level of artistic control over the visual content than in previous practice. Secondly, if artists choose to engage in programming to achieve even higher artistic control, this model allows them to use the Processing programming language in a modular manner, thus still reducing the amount of effort necessary in relation to previous practice. The use of the Processing language also signifies a considerable reduction in complexity, since Processing was from the ground up created to reduce the difficulty with which visual artists can program procedural computer graphics.

What is conceptually novel about this approach is the re-adaptation a programming language intended for artists to instead behave as a plug-in API, thus further increasing the language's flexibility and ease of use, by extending its usefulness to a context for which it was originally not intended.

As visual synthesizers are housed in modules that can be interchanged during a performance, it is even possible for an artist to improvise to a much greater extent than in previous practice, provided he/she has access to a sufficient number of such modules. This is already possible in contemporary VJ performance, as video clips can be dynamically loaded during the course of the performance. It has however to date not been achievable with procedural graphics to the same extent.

1.2.3 Increasing achievable mapping complexity and variability

In previous practice, in which a very limited amount of input control data is available, defining the mappings between the incoming control data from the musical performance and the control parameters of the live procedural computer graphics is relatively simple. After all, there are limited uses for beats, tempo and amplitude from a single signal. As such, these mappings are hardcoded in the implementation of virtually all software packages that use data derived from audio processing to control live procedural computer graphics, such as VJing applications. With the approach presented here however it is neither feasible nor desirable to hardcode the mappings, because of the significantly greater amount of input and output control data available. As the type and number of instruments employed is varied, and as visual synthesizers are added and removed, the parameter spaces of input and output control data will vary significantly from one performance to the next, making a hardcoded solution impossible. When there are more input and output parameters to work with, there are also many more options of what connections to make. The only desirable solution is therefore deemed to be one that allows direct control over the mappings.

To address these issues a novel artistic conduct has been devised: that of gradually creating, destroying and altering mappings between the two parameter spaces of input and output control data, both before and during the course of a performance. Furthermore a software application has been developed, to be used for this conduct. The conduct has been christened *Mutable Mapping*, while the software created for facilitating the task was dubbed the *Mediator*.

With mutable mapping, the constraints of previous practice are lifted as both high mapping complexity can be achieved, and altering the mappings over time is made possible. Creating a narrative in the mapping between music and visuals, using tension and release, and suspected expectation, is therefore made possible.

1.2.4 Allowing for collaborative performance

Collaborative performance is very common in most forms of live performance art. In live musical performance, it is common practice that groups of musicians, sometimes in conjunction with conductors or live mixing engineers, perform together. In present work, musical instruments are employed as sources of control data, and mutable mapping is a practice much akin to live music mixing as well as conducting. The same benefits that allow collaborative live musical performance may thus translate to also allowing collaborative audiovisual performance, which may easily scale to large numbers of participating individuals. As a consequence, the benefits of collaborative performance, such as mutual engagement, will allow for a heightened experience for performers and audience.

1.2.5 Soma: a new artistic practice

Mick Grierson (2005) describes how we can create art which capitalizes on the scientific evidence that congruent musical and visual events result in stronger reactions for the perceiver. He thus defines a new art form, named audiovisual composition, where music and visuals are experienced as an inseparable whole, more captivating than similar music and visuals would have been if presented without strong congruence between them.

However, it is known that humans fuse more percepts than just auditory and visual; these are processed in combination also with tactile and other sensory stimulation, depending on the level of congruence between the stimuli presented to the different modalities (Wozny, Beierholm, & Shams, 2008).

It is further known, that *"(...) the perceiver watching, listening to and experiencing another's motor performance, simulates the actions of the performance within the range of their own motor capabilities"* (Rodger et al., 2007), and that as a consequence, when performers exhibit advanced enactive knowledge, audiences perceive a richer sensory experience. Many music researchers assert that this richer sensory experience has as a direct consequence that audiences are more engaged in the performance (Armstrong, 2006).

Through the combination of the above knowledge, the theoretical foundation for yet another art-form emerges. In this, tri-modal sensory stimulus is combined: music is performed alongside congruent visuals, involving performers that employ highly advanced embodied motor knowledge in their performance. Thus audiences perceive the resulting performance on three congruent modalities: audiences subconsciously simulate the perceived actions, thus activating their own brains' motor capabilities, while simultaneously experiencing congruent musical and visual stimulus. The experience is similarly strengthened for the performers enacting the physical gestures. Note that Soma is not intended to be a universal art form, as is for example Wagner's "Gesamtkunstwerk", instead it is a direct continuation of the visual music artistic tradition, which in this thesis is detailed in section 2.1.1.

Although the premise is tri-modal congruence, it is recognized that for tension and release and suspended expectation to be possible in the narrative, it is also necessary to allow for narratives that can transition between states of a high level of congruence and *"Binary opposition or total incongruence"* as also Grierson recognizes is necessary (2005). Such narratives are here achieved through employing mutable mapping. Congruence between abstract stimuli to multiple senses is by Grierson both defined as temporal congruence, where events co-occur on in multiple stimuli, and as structural similarity:

> *"The extent to which an abstract image can be understood as being similar to a sound is entirely to do with relative structural aspects of the sonic and visual components. That is to say, the structure of the sound can be shown to be similar to the structure of the image. For example, a sound which has a high ratio of unrelated overtones to harmonic ones can be described as 'noisy'. Likewise, an image which exhibits less order, less self similarity, and more incoherent complexity could also be described as noisy. In this way, we can begin to think about sounds and images in a similar way with respect to texture (or shape) and timbre. This provides us with a method for discussing the congruence of sound and visual events in the context of abstract audiovisual composition" (Grierson, 2005).*

In keeping up with the longstanding tradition established by the many predecessors of this work, of inventing and naming new artistic practices, the art-form has been given the name *Soma*. Partly because it is the ancient Greek word for body and partly because it is the state sponsored drug administered to all citizens in Aldus Huxley's *"Brave New World"* (1932).

1.2.6 A new performer role

The tasks a visuals performer engages in when using the system detailed here are manifold. Firstly, he/she assumes the role of operating the Mediator software. Secondly he/she is responsible for gradually adding/removing/rearranging visual synthesizers. Finally, he/she may also directly alter the visual synthesizer parameters using a controller device with knobs and sliders, for controlling additional changes, not connected to the musical performance. These changes, depending on the performance at hand, may add to the narrative, but may not be intuitively controllable from the instruments. As an example, this could include zooming in/out on the image, altering the colours of the visual synthesizers, or rearranging the spatial location of elements on the screen.

In summary, the *Trinity* system consists of three software applications: The *Input Processor*, responsible for gathering and processing audio and discreet control data from instruments, *Mother*, responsible for hosting visual synthesizers, and finally *Mediator*, which provides the functionality and user interface for controlling the mutable mapping.

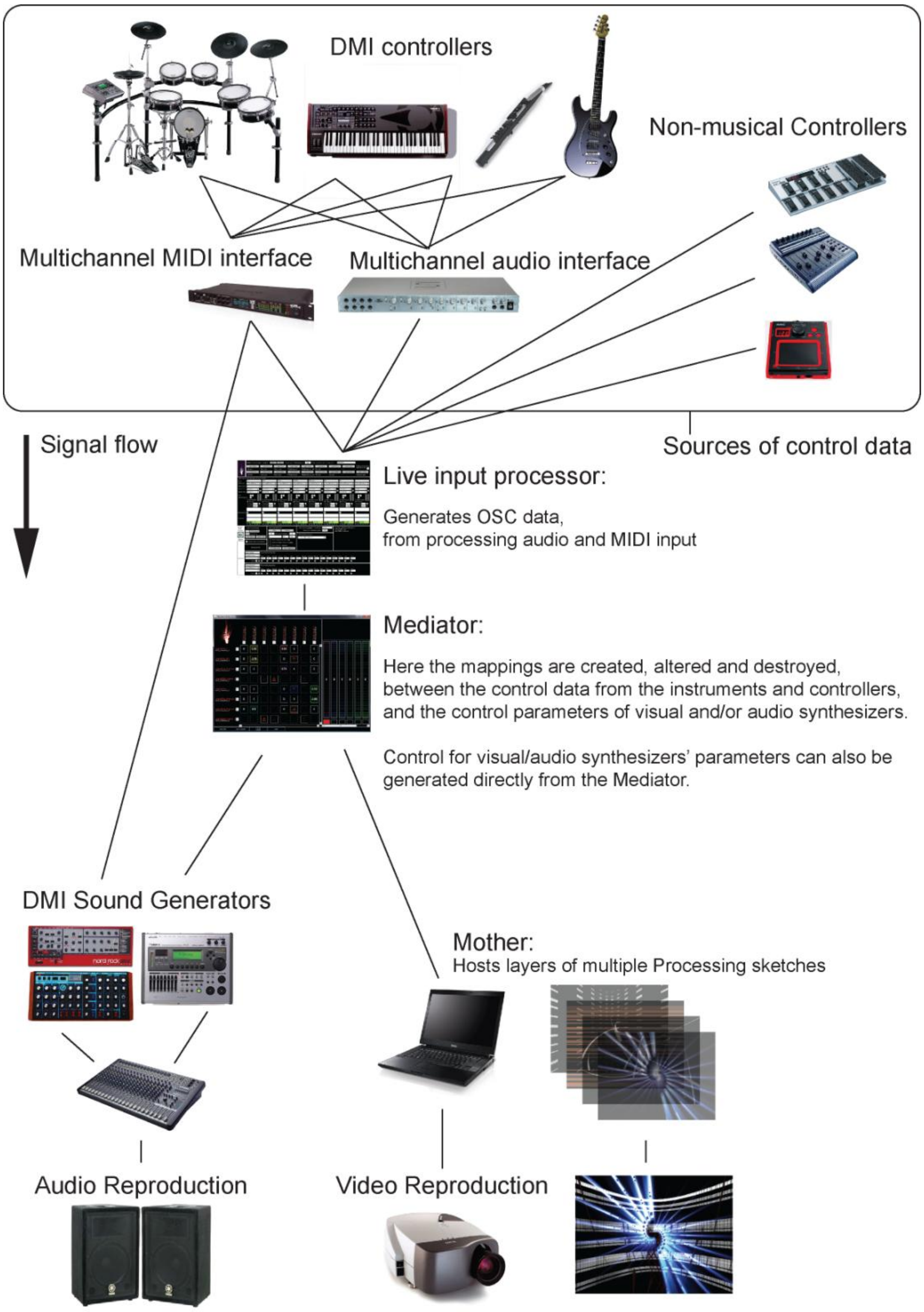


Figure 1: Illustration of signal flow in proposed system. Please refer to this diagram for gaining an understanding of how the different elements of the Trinity system are connected. It is also clear from this diagram, how the signal flows for generating audio and video are independent from each other, with both being controlled from the gestural control data, rather than one from the other (product images in illustration are promotional photographs, as contained in the press-pack from the respective device manufacturer's website).

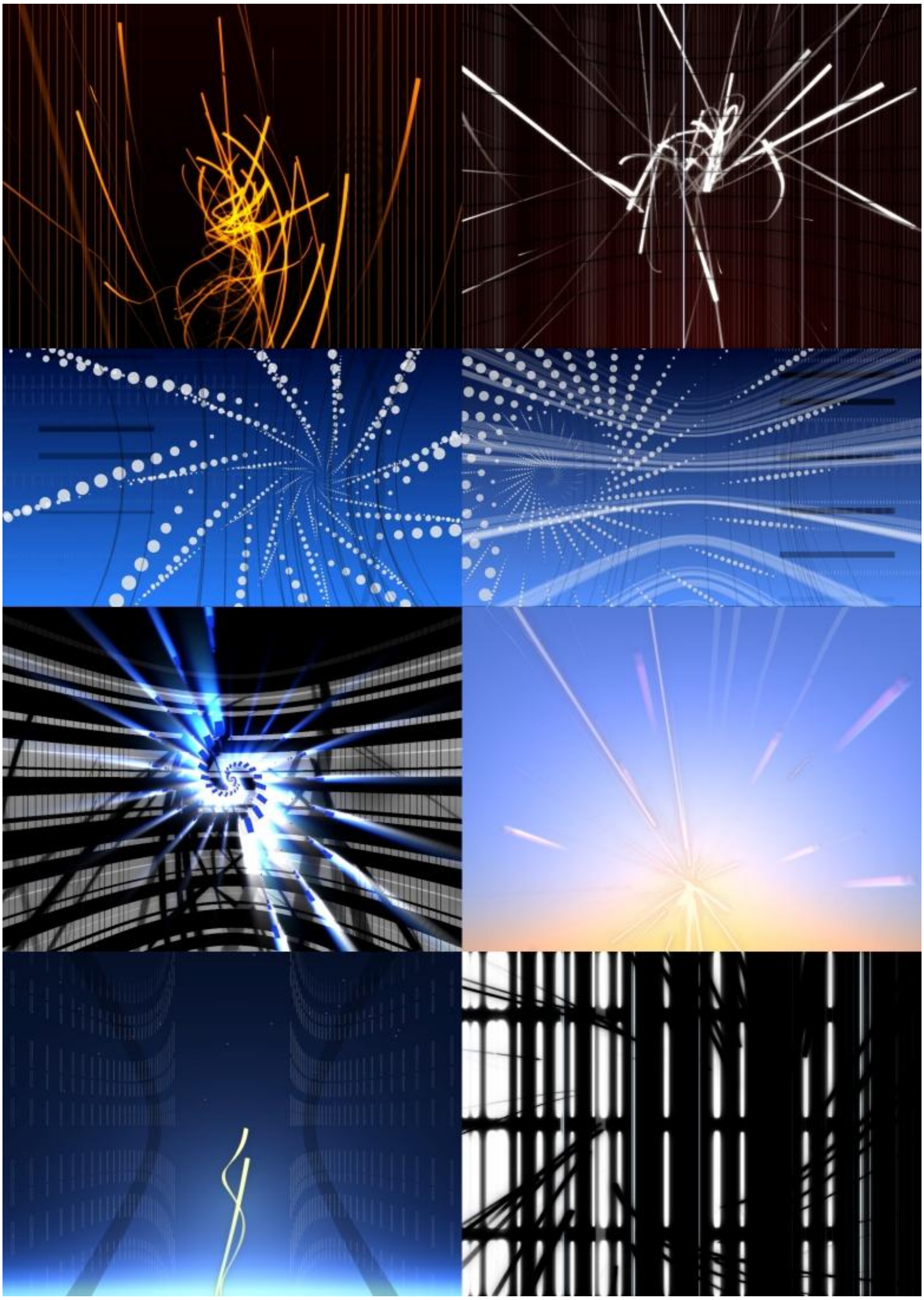


Figure 2: Examples of the visual output possible when using the Trinity system.

1.3 Methodology

Identifying what research and evaluation methodology should best be used for this work was far from straightforward. It was here necessary to extensively review the arguments in the research community, and make an informed choice, with precious few pre-existing case studies to draw from. As such, the review and selection of methodology is, together with the subsequent case study embodied in the research carried out, a research contribution also on its own merit, to the research community's on-going discussion on methodology.

From the reviewed literature, it is clear that present work is best approached as a transdisciplinary (Nicolescu, 2005) music/multimedia technology and Human Computer Interaction (HCI) research project, mostly related to the body of work sometimes referred to as the third paradigm/third wave of HCI (Harrison, Tatar, & Sengers, 2007). The project is best carried out employing design research methodology in its development and evaluation, while also introducing the additional viewpoint of ethnographic method during the evaluation stage. The work will be approached as design research, in that an artifact will be produced, created with the goal of producing knowledge, as opposed to designing it as a commercially viable product (Fallman, 2007) (Zimmerman, Forlizzi, & Evenson, 2007) (S. Greenberg & Buxton, 2008).

The developed system is clearly experience-centred, since it is not used for achieving a particular end goal, as are task-oriented systems. Instead the value in its use lies in the dynamic and continuous experience, for all actors involved, performers and audience. Its context of use is highly situated, the experience being entirely different if the system is used in for example a laboratory setting, compared to its intended context of use; that of live performance in an appropriate venue and with a live audience, alternatively a music rehearsal room, developing new material in preparation for a live performance. As such, it cannot be evaluated with quantitative methods, since the measure of its usefulness is entirely subjective, in a manner similar to other experience-centred work, as well as much other research falling under the third paradigm/third wave of HCI.

As the system is going to be evaluated in an uncontrolled, real-world setting rather than a laboratory, it also needs to have reached a sufficiently advanced state of development to cope with the demands of that environment (Edmonds et al., 2005). This is both in terms of its technical completion - making sure it functions properly, and its aesthetic completion - being able to produce sufficiently intricate and aesthetically pleasing output to keep the artists and audiences involved genuinely engaged.

Because the paradigms and the system developed arguably are novel inventions, as has thus far been presented in this chapter, rather than an iterative refinement of pre-existing technology and practice, there is little reason or possibility to perform usability evaluation at this early stage of its development. For now, the focus therefore needs to be on developing the *"right design"*, leaving eventual usability evaluation, of getting the *"design right"*, to a possible future refinement (S. Greenberg & Buxton, 2008).

The development process is shaped according to the model of iterative development (Dix, Finlay, & Abowd, 2004). An early prototype is created, put to use so that it can be evaluated, and the results from its evaluation are fed back into the development of a further refined prototype. When a sufficient number of development iterations have been completed the specification of a finalized prototype version will be formulated and implemented. The intermittent evaluation stages are due to time constraints necessarily informal, and carried out by the developer of the system. In accordance to the evaluation criteria discussed by (Zimmerman et al., 2007), this development process is documented and presented in chapter 5.

A combination of evaluation methods are used, with the primary methodology being both that of design research and of ethnographic method. The evaluation criteria for HCI design research described by Zimmerman et al (2007) will be taken into account so that it is made sure they are adequately addressed. Further evaluation criteria, as discussed in (S. Greenberg & Buxton, 2008), (Sengers & Gaver, 2006), and (Ippolito, Blais, Smith, S. Evans, & Stormer, 2009), will also be addressed.

For observing the experience of the musicians involved, an ethnographic study of the rehearsals is performed, while for eliciting feedback from audiences, an online questionnaire is employed, used in conjunction with videos of pre-recorded material created using the Trinity system. Finally, live visuals performers and VJ's are interviewed, to elicit their view on the concepts and implementation created here. The ethnographic study is intended to be a valid outcome in its own right - as discussed by Dourish (2006), and as was exemplified by Gaver, Bowers et al (2004) – The study should be read in its entirety, and not simply be reduced to a list of conclusions, or implications for design. Of course conclusions drawn will too be presented and discussed, but this will occur alongside the ethnographic study and not in place of it.

The evaluation process is concentrated on assessing whether the limitations of current practice, as detailed in section 1.1, are addressed by the suggested novel solutions detailed in section 1.2. It is also geared towards gathering feedback on effects that the developed ideas, designs and applications may have had, that had not been predicted, as is discussed in appendix H.

Following design research methodology, the evaluation is conducted according to the discussion presented by Zimmerman et al. (2007): *"(...) The final output of this activity is a concrete problem framing and articulation of the preferred state, and a series of artifacts—models, prototypes, products, and documentation of the design process".* In other words, the implemented design itself, and the process it was arrived at, serve as testament to whether the identified limitations are sufficiently addressed by the conceived ideas.

In addition, employing ethnographic method, the following trials are to be conducted:

From the rehearsals and performances with groups of musicians, feedback is gathered on the overall experience of Soma performance. The potential benefits of using a high bandwidth of control data from their instruments are observed. It is gauged whether the musicians feel they are to a greater extent controlling the performance, compared to a simulated previous practice scenario, thus suggesting that the increase in control data bandwidth may allow them to take a greater advantage of the embodied knowledge they have of their instruments. It is further gauged, whether the combined experience influences the aesthetic content of the music they perform.

Live visuals performers familiar with established practice are engaged, continuously during the project, and more formally in the end, to elicit feedback on the concepts embodied in the system, as well as their implementation. This process is viewed in the light of ethnographic method, through gathering feedback from established practitioners by means of interview and observation, but it may equally validly also be viewed in the light of design research, as the practice of expert critique.

An on-line survey is finally conducted to experimentally gauge whether increased accuracy in the temporal correlation between audiovisual stimuli in a pre-recorded visual music piece, results in a more aesthetically engaging experience for audiences. The survey also serves a second purpose, of providing audience feedback, on how engaging/enjoyable they find it to watch a video created using the system. The video will necessarily not be able to be viewed as Soma, since there is no live performance involved, but it is still a valid question to ask, if an audiovisual composition piece created using the system is enjoyable for audiences to watch.

Although there are strong arguments in recent literature for the relevance of selected methodologies in HCI research, there are few published accounts on their application on evaluating larger scale work. Following the argument in (S. Greenberg & Buxton, 2008), it was elected to make the methodological choices that were deemed the most appropriate for the project, rather than attempt to rephrase the research objectives to make them compatible with traditionally established methodology. This was expected to be a significant challenge, given the comparatively small body of previous practice examples to draw from. The benefit however is that the opportunity presents itself of contributing to the creation of a body of previous practice, and thus help making the methodologies chosen more accessible to future researchers.

1.4 Summary of contributions, and how each is evaluated

As has been noted, there is no one-to-one correspondence between the ideas based on which the system is designed, and the previously detailed limitations. Instead, the developed ideas work in combination towards addressing those limitations. Furthermore, there is also no one-to-one correspondence between the evaluation steps taken, and each individual contribution made. How the identified limitations, the contributions made, and the evaluation effort undertaken all interrelate, is hopefully sufficiently addressed in this section, where each is briefly repeated for the sake of clarity.

a) Through using the full bandwidth of control data that the musical instruments can produce, the musical gestures of the performers are encoded in much higher detail than only beat events, amplitude and tempo from a stereo audio signal, as is common in established practice. This allows for a much higher mapping complexity than is used in established practice, which is immediately technically demonstrable.

b) Following from the above, the hypothesis is formulated, that from the increase in control data bandwidth, musicians are expected to better take advantage of the advanced embodied knowledge they have of their instruments, in the performance of live visual music/audiovisual art, and therefore achieve an increase in expressivity. The subjective and highly elusive experience is desired, of the musicians to some extent feeling in control of the visuals, as opposed to the visuals reacting to their playing.

c) The novel conduct of mutable mapping is introduced. Mutable mapping refers to live performance through gradually creating, destroying and altering mappings between the two parameter spaces of input control data from a musical or other performance and output control data of the visual/musical/other synthesis devices, during the course of a performance. The hypothesis is that mutable mapping may facilitate complex, dynamic, and improvised mappings in live visual music / audiovisual art, which may vary over time, thus allowing for improvisation, tension and release, and suspended expectations, in the narrative constructed using mappings.

d) A hypothesis towards reducing the difficulty inherent in the preparation towards the performance of live procedural computer graphics is introduced and tested. The proposed solution is the encapsulation of the complexity of programming visual synthesis algorithms in a programming language for artists, into flexible modules that are easily exchanged and managed. It is expected that artists may then dynamically layer the output of such modules during the course of a performance to create a narrative, in a manner analogous to VJing, thus greatly simplifying the preparation and improvisation of a performance.

e) Because musicians are already accustomed to performing collaboratively using their instrument, it is expected that their experience can directly translate to also allowing collaborative performance in the context of live visual music/audiovisual performance, as is facilitated using the system designed and implemented here.

f) The novel art form of Soma, in which correlated auditory, visual and proprioceptive stimulus is used to form a combined narrative, is introduced. Soma builds on research findings that humans are more engaged in a performance, when performers' exhibit advanced motor knowledge, and when congruent percepts across modalities temporally coincide. To allow for live Soma performances, all previous contributions are used in conjunction.

The suitability of all the above proposed ideas used in combination, is addressed following design research methodology, through designing and implementing a system which demonstrates how they may work together in practice. This demonstration is achieved through detailing the design and implementation of the system itself. It is further achieved through using the system to create a series of audiovisual pieces, which serve to show that the system functions technically, and to demonstrate the output of the system to musicians for the purpose of engaging them in collaborations. The above created material is further used in two solo performances, serving to technically test the system as well as the developed visual material in a live setting, before actual bands of musicians are engaged. Finally, the system is employed in a context also involving groups of live musicians, where all its constituting elements are demonstrated in use, both in rehearsals and in live performances.

The Mother application implementing the concept employed towards simplifying preparation and improvisation described in (d) is released as free open source software, to engage the live visuals community with the ideas it engenders.

The conduct of mutable mapping (c), the concept employed towards simplifying preparation and improvisation (d), and the concept behind Soma (f), are further evaluated through engaging live visuals performers, to elicit their feedback, and gauge what role these ideas may play in their current and future artistic practice.

The concepts behind the expected increase in expressivity (b), of Soma (f), and of improving the practice of collaborative performance (e), are also evaluated through conducting live Soma rehearsals, involving groups of live musicians, who are interviewed and observed to document their experience. The hypothesis (b) is in this work addressed through an existence proof, demonstrating there is a detectable difference in achieved expressivity. No attempt is however in present research made to furthermore also objectively measure the achieved increase in perceived expressivity.

Finally, an online survey is conducted, to quantify part of the benefit inherent in Soma performance, as is facilitated with the increased mapping complexity (a) possible with the system developed here. The survey is also used to gauge how engaging/enjoyable viewers find it to watch a pre-recorded piece created using the system.

1.5 Thesis structure

Chapter 2 contains a background review, to set the context for this work:

Section 2.1 discusses the history of relating visual arts to music, so as to historically ground the work presented, and also make explicit the limitations that have been identified in current practice, which this thesis seeks to address.

Section 2.2 details the modern history of the development of new musical instruments, concentrating on the interface with which these are played. This is to illustrate the close relation between the technologies used for performing actual and visual music, and frame the argument for why the choice was made in this work to employ existing musical gesture controllers, as the primary source of control data for the live performance.

Section 2.3 presents a brief discussion on what constitutes at the most universal level a narrative in absolute music, and building on that, how the definition is extended to apply to visual music and audiovisual art.

Section 2.4 discusses the practice of programming as a means for artistic expression. The visual output from the system introduced here, is unlike the majority of established practice generated programmatically, building on the modern tradition of programming as art, in which the artistic medium is program code. It is therefore necessary to briefly detail the history, practices and technology of this field, to better explain the rationale behind allowing artists to create their own procedural graphics for the system introduced in this thesis. To further ground the feasibility of this choice, section 2.5 proceeds to detail the present state of development, for programming environments intended to be used by artists with no formal computer science training.

Section 2.6 presents a review of all research and evaluation methodology that is deemed immediately relevant to present work. Identifying what methodology should best be used was far from straightforward: although there is much discussion on methodology in related research communities, and many proposals for adopting methodologies from other disciplines have been made, there still is little conclusive agreement. It was therefore here necessary to review the arguments in the research community, towards rigorously grounding the subsequent methodological choices made for conducting the work presented here. As such, the review and selection of methodology is, together with the subsequent case study embodied in the research carried out, a research contribution to the community's on-going discussion, also on its own merit.

Finally in section 2.7, related research that immediately precedes present work is detailed, along with what shortcomings it has been found to have, and how it differs from the work detailed in this thesis. This brief review serves to clearly define the state of current practice's cutting edge, and thus to render explicit the extent of the contributions that are made with the work presented in this thesis.

Chapter 3 explains the new concept and envisioned practice of mutable mapping in detail. Although mutable mapping was initially conceived of for it to be used as part of the Trinity system, it quickly developed to an idea that generalizes to all digitally mediated artistic performance / live new media art. Because of this wide-reaching generality, it is worth discussing the concept of mutable mapping on its own, before discussing how it also plays a central role in the here introduced practice of Soma, as performed using the novel Trinity system.

Chapter 4 details the Trinity system, as formed by the three software applications developed as part of this research: The Live Input Processor, responsible for gathering and processing audio and discreet control data from instruments, Mother, responsible for hosting and displaying the mixed output of several visual synthesizers, and finally Mediator, which provides the functionality and user interface for mutable mapping performance.

Chapter 5 contains an account of the development process followed in this research, in fulfilment of the evaluation criteria for design research presented by Zimmerman et al (2007), as is reviewed in section 2.6.2.5 of this thesis.

Chapter 6 details what experiments were conducted. Rehearsals and performances were undertaken with groups of musicians, to gather feedback on the overall experience of Soma performance. Live visuals performers familiar with established practice were engaged, to elicit feedback on the concepts embodied in the system, as well as their implementation. Finally, an online audience survey was conducted, for providing audience feedback on how engaging/enjoyable audiences found it to watch a video created using the system, and to gauge whether increased accuracy in the temporal correlation between audiovisual stimuli results in a more aesthetically engaging experience.

In chapter 7 the evaluation outcome of the work is presented, drawing on the data detailed in sections 5 and 6, including both the viewpoints of design research and of ethnography, as well as a technical evaluation of the developed software.

Chapter 8 details what potential future work has been identified: what further experiments may be conducted, what additional software development could be undertaken, and what work could be carried out to further the aesthetic output achievable with the Trinity system.

Finally, chapter 9 restates the contributions of present research, and offers the final conclusions discussion for the thesis.

2 Background

This research is highly transdisciplinary, integrating knowledge from music, visual and multimedia art and performance, music technology, programming as art, software engineering, cognitive science, psychophysics and finally human computer interaction, which in itself is a strongly interdisciplinary field. Because of this, the review section is necessarily long, for readers who do not have expertise in all the above fields to still be able to follow the argument presented.

2.1 Relating visual arts to music

The history of creating abstract moving images, meant to be presented alone or be accompanied by music, was also briefly summarized in the introduction. However, it is so closely connected to the work presented here that it warrants reviewing in greater detail. After this overview of the history leading up to the current state of the art in the related technology and practice, the limitations of current practice identified in section 1.1 will be clear.

This topic can be further broken in to two main subtopics, depending on whether the means of producing the work are real-time or not. Perhaps surprisingly not all early work was created frame by frame, using traditional animation techniques, as one might assume; many instruments capable of real-time performance have also been created, which are now collectively referred to as ‘Colour Organs’.

There have been many attempts to find an analogy between the spectrum of visible light and of audible sound. Isaac Newton had observed a correspondence between the proportionate width of the seven prismatic rays and the string lengths required to produce the musical scale D, E, F, G, A, B, C (Peacock, 1988). Though searching for such a correspondence is by many seen as oversimplifying, researchers have continued to publish theories and findings on this topic well into the previous century. Throughout these early developments however, a widely accepted theory for the correspondence between colour and sound has despite the many attempts made, failed to emerge. For example, the musical note C has by three different pioneers of the field been assigned as diverse colours as yellow, red and blue.

2.1.1 Visual Music, Intermedia, Abstract film, and Synesthetic Art

The immediacy with which music can communicate emotion has through time been envied by many artists, probably most famously Wassily Kandinsky, who, fascinated with the unparalleled emotional power music can bestow on its listeners, set out to try and recreate it in painting (Dabrowski, 1997). Kandinsky frequently used musical terms when describing his works, referring to them as compositions and giving them opus numbers, much like classical composers christened their concerts. Working with abstract painting, he worked with concepts borrowed from music, such as harmony, rhythm, dissonance, etc., attempting in his paintings to create a form of visual music. Interestingly, he was also very much inspired by and interested in the work of Scriabin, the composer who most likely was the first to write a musical piece having from the beginning the intention of it being accompanied by visual music: Prometheus. Kandinsky was for a while part of the German school of Bauhaus, where many of his colleagues shared his interest in fusing the different art forms (Droste & Williams, 2002).

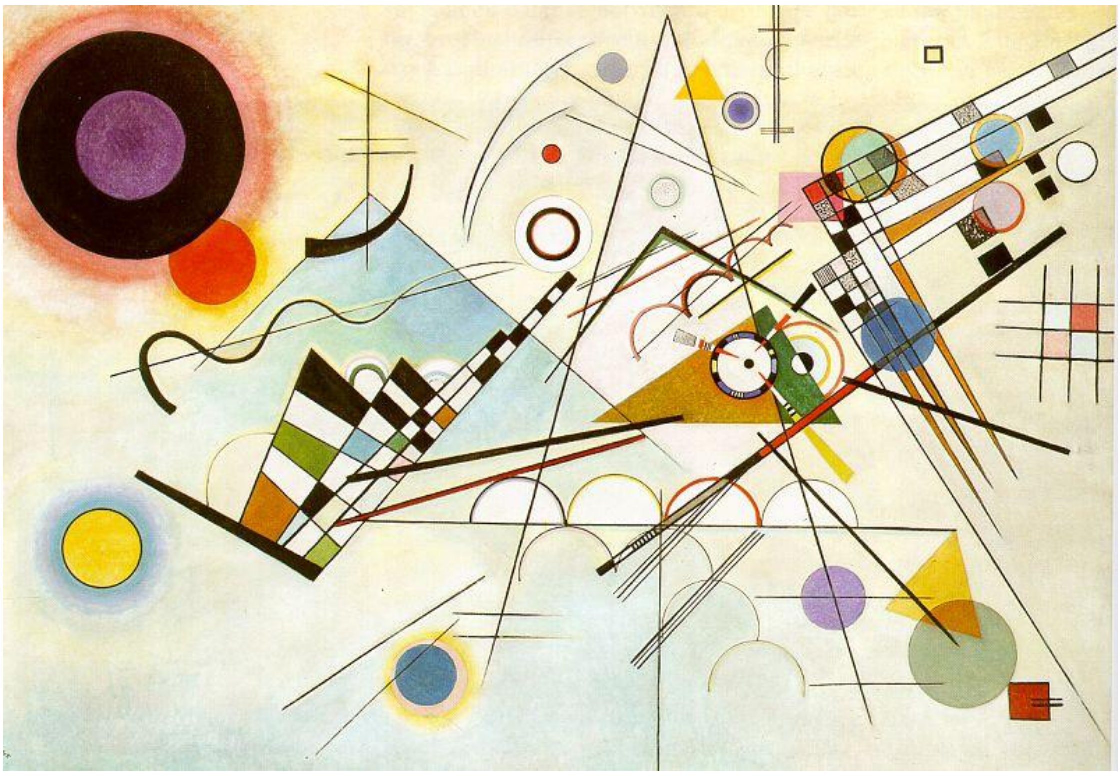


Figure 3: Wassily Kandinsky, Composition VIII, 1923 (Solomon R. Guggenheim Museum, New York).

The real explosion in activity came with the invention of animation techniques, allowing for the first time the creation of images that change over time, with sufficient control over the end result, and few restrictions imposed by the equipment in use in terms of the possible visual output. With music being a time based art form, the advent of animation sparked a great increase in the interest of fusing music and visual art, with artists finally being able to create what they previously could only theorize about.

It is still a relatively young artistic direction, and goes under many names and varying definitions, as variations have been discovered again and again by various groups of artists. Castel and Kastner used the term 'colour music' to refer to their work, while Thomas Wilfred coined the word *Lumia* to refer to his silent colour form and motion performances. In the meanwhile, the earliest experimentations with animation techniques to produce visual music had begun by artists such as Walter Ruttman, Viking Eggeling, Hans Richter, Oskar Fishinger, Len Lye, John and James Whitney, and many more (Ox & Keefer, 2006), sparking a tradition that can be followed to the present day.

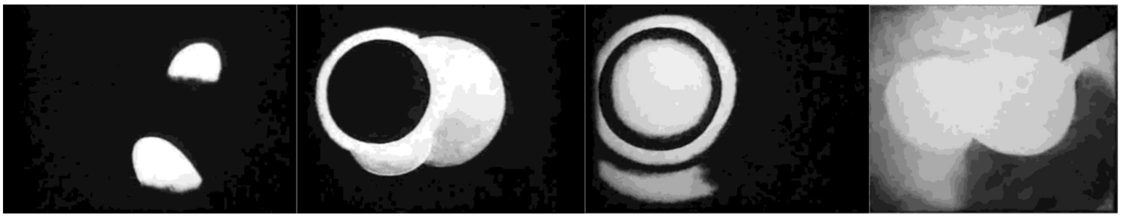


Figure 4: Still images from Walter Ruttmann's Opus II (1921) (images in public domain).

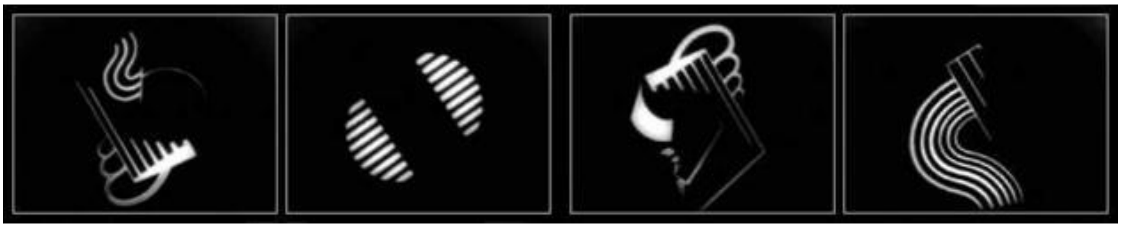


Figure 5: Still images from Viking Eggeling's Symphonie Diagonale (1924). It is by now no longer protected by copyright, and is available to view online at ("U B U W E B - Film & Video: Viking Eggeling," n.d.)

Note that work that could just as well be referred to as visual music, has through time also been created under many other names, such as *Abstract Film*, *Abstract Animation*, *Synesthetic Art*, *Rhythmic Light* and *Video Synthesis*, to name a few (http://www.iotacenter.org/about). Though closely related, the definitions of these terms do not overlap completely, as abstract film for example need not be visual music, and also visual music need not always be abstract. In the 1960s, Dick Higgins chose term *Intermedia* (Higgins, 2001) to refer to the interdisciplinary art that had emerged, and that did not fit into the existing art genre definitions.

Visual music has no one rigid definition; Brian Evans (2005) defines it as:

> *(...) time-based visual imagery that establishes a temporal architecture in a way similar to absolute music. It is typically non-narrative and non-representational (although it need not be either). Visual Music can be accompanied by sound but can also be silent.*

Jack Ox and Cindy Keefer (Ox & Keefer, 2008) further define visual music as anything matching one or more of these four descriptions:

- *A visualization of music which is the translation of a specific musical composition (or sound) into a visual language, with the original syntax being emulated in the new visual rendition. This can be done with or without a computer. This can also be defined as intermedia.*
- *A time based narrative visual structure that is similar to the structure of a kind or style of music. It is a new composition created visually but as if it were an aural piece. This can have sound, or exist silent. Theorist/inventor Adrian Klein wrote in 1930: "...somehow or other, we have got to treat light, form and movement, as sound has already been treated. A satisfactory unity will never be found between these expressive media until they are reduced to the same terms.*
- *A direct translation of image to sound or music, as images photographed, drawn or scratched onto a film's soundtrack are directly converted to sound when the film is projected. Often these images are simultaneously shown visually. Literally, what you see is also what you hear. (An early example is filmmaker Oskar Fischinger's Ornament Sound experiments c. 1932). There are many examples in Visual Music film of this process, e.g. McLaren, Spinello, Reeves, Damonte, Neubauer and other contemporary filmmakers (...). This method has been called a "pure" type of Visual Music.*
- *A visual composition that is not done in a linear, time-based manner, but rather something more static like a 7' x 8' canvas. However, as in Klee, the movement of the painted elements can and have achieved a kind of Visual Music, serving as an artist's visual interpretation of specific music.*

Artists and scholars have through time worked on further defining visual music, as well as on formulating its theory, analogously to the various existing formal theories for musical composition and harmony. Detailing this work however would be beyond the scope of this thesis. There are extensive resources available on the topic one may refer to: The Computer Music Journal has dedicated an issue to the subject of visual music (29:4, 2005), with detailed further reviews, discussions and references. One can also refer to the very extensive cinematography and bibliography resources provided by the Center for Visual Music (http://www.centerforvisualmusic.org/) and Iota Center (http://www.iotacenter.org/), two organizations dedicated to documenting, preserving and promoting the art of visual music.

2.1.2 Colour Organs and Lumia

The first known instrument of this kind was the Clavecin Oculaire, which was proposed in 1734, a construction devised by Frenchman Louis-Bertrand Castel (1688-1757) (Collopy, 2000). This was an instrument of five octaves, where each of the twelve notes of one octave corresponded to a colour, according to a slightly modified version of the note to colour correspondence proposed by Isaac Newton, the first publicized correspondence of this kind. Moving higher up the octaves, the colour produced was more luminous than when playing the lower ones. It is not certain whether Castels intent was to use this instrument to accompany music, or if it was to be used on its own, as visual music. Its construction was simple, involving candles, mirrors and coloured paper. No evidence exists that such an instrument was ever completed, but since its construction certainly had begun, it is nonetheless considered to be the first known instrument of its kind.

The first instrument to have both musical and visual output was the Pyrophone, constructed in 1873 by Englishman Frederick Kastner (Collopy, 2000). This used gas jets, known to produce a controllable, musical sound, and visible flames, contained in crystal tubes that lit up when notes were played.

Most famous and elaborate of the early colour-organs was Thomas Wilfred’s Clavilux (Figure 6), developed around 1922 (Peacock, 1988). He had, unlike most of his predecessors, rejected the notion that there is any absolute correspondence between music and visual art, and instead concentrated only on light, naming this new art form of controlled light compositions Lumia. Wilfred separated the elements to describe a lumia performance into colour, form and motion, in contrast with the earliest pioneers of the art form, who only considered colour in their constructions and compositions. Taking nearly a decade to complete, his instrument was a complicated construction involving six projectors, each with an arrangement of prisms placed in front of it, each prism of which could be rotated 360 degrees around all axes. The instrument was controlled through a control desk consisting of banks of sliders. Wilfred toured with the instrument around most of the world, giving very well received recitals, where he performed complete lumia compositions. These were treated by him very much like musical compositions for the eye, each having had its own Opus number assigned to it. He is arguably the most prominent pioneer of the field of live visual music performance, in that not only did he conceive of and create an instrument, but he also created widely accredited work with it. The work of his that remains today, in the form of smaller systems, is still enjoyed not only on its historical significance but also purely on its artistic merit, with a system of his being permanently exhibited at the Museum of Modern Art in New York, and a documentary about his work having just been released (http://www.lumia.tv/).

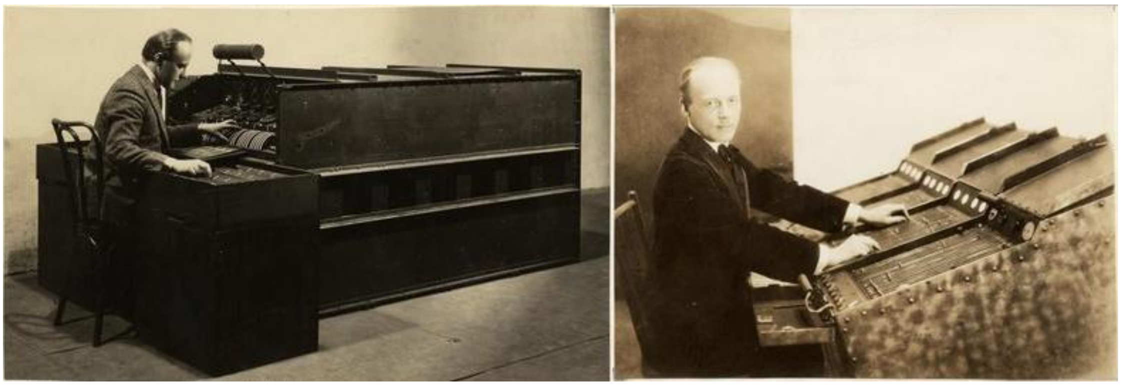


Figure 6: Thomas Wilfred, and his Clavilux Models B (Left) and E (Right) (Images from the Yale University Library).

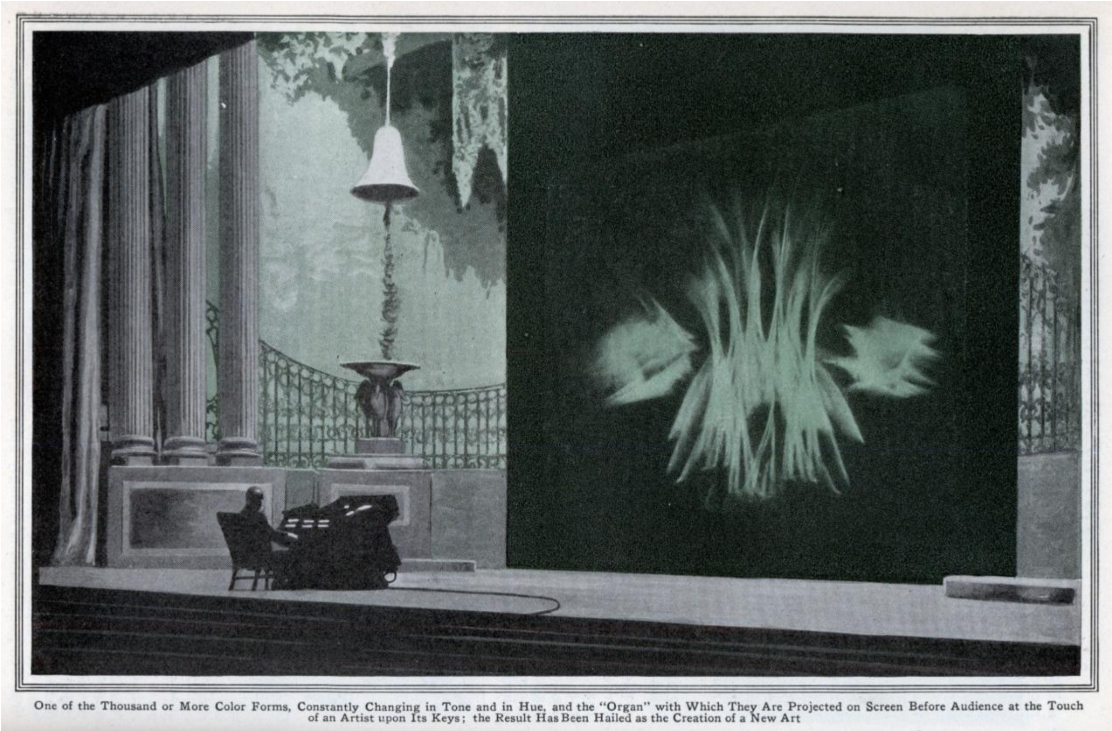


Figure 7: Thomas Wilfred performing on a Clavilux (Popular Mechanics, April 1924, image from the Yale University Library).

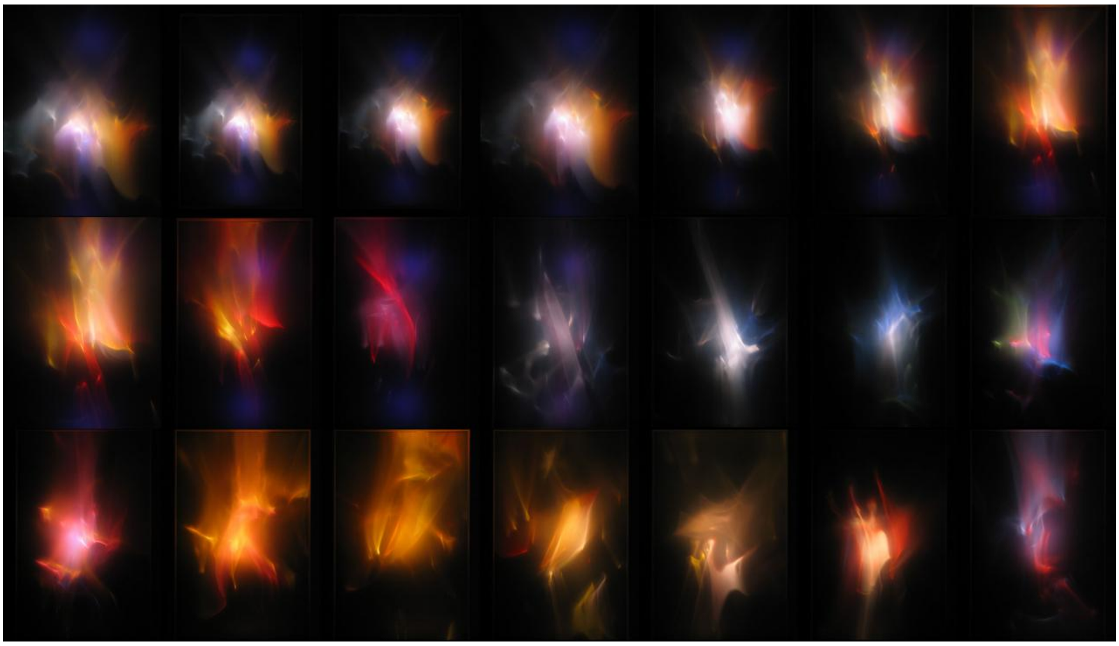


Figure 8: Series of still shots from Wilfred's Opus 161, as produced by a smaller Clavilux, not built for live performance, but meant to be installed as a permanent exhibition piece. It continuously displays the one composition that is built into it. The images are from a restored machine, owned by Eugene and Carol Epstein (www.thomaswilfred-lumia.org).

The first known modern device that allows the control of live visual art from musical input was Gordon Pask's Musicolour system (Pask, 1971). Musicolour was most famously featured in the groundbreaking Cybernetic Serendipity exhibition in 1968, although the systems' first incarnation was created much earlier, in 1953 (Pask, 1971). Pask created a series of devices along the same concept. The signal that controlled the visual performance was derived from processing a monophonic audio signal. Rapid changes in amplitude (beats) were detected, and the overall amplitude of the signal was tracked as well as the amplitude of individual frequency bands. The visual projection capabilities have varied significantly throughout Musicolours' various incarnations. In its most advanced form, its output was produced through projecting light through painted discs of glass, and later through rotating discs containing fluids subjected to electrochemical reactions.

During the 60's and 70's there was an explosion of activity in live visuals performance. The popular culture of the period made audiences very receptive to the art form, making it a common occurrence during popular music concerts, to also have visuals accompaniment. Predominantly the performance would be using a combination of analogue techniques, mainly involving powerful overhead projectors on which artists would project slides, draw, and mix coloured liquids in bowls. The term *liquids* was predominantly used to refer to this style of performance (Spinrad, 2005).

During the same period, the great advancements in electronics and computer technology gave birth to a new breed of systems, most notably the analogue video synthesizer and the computer controlled laser show, both of which were used for creating visual music, and music visualization; then followed the creation of computer systems with the purpose of generating graphics.

Analogue video synthesizers emerged as a development inspired by the invention of the musical analogue synthesizer, the first commercially available systems of which came out in the mid-late sixties. Like their musical counterparts, they use analogue oscillators, that create basic waveforms (sine, ramp and square waves), which are then additively or subtractively mixed together, and filtered. This could be loosely paralleled to the output of an oscilloscope. By adding the effect of video feedback, the ability to move, rotate and scale the image, and the mixing of moving images together from different sources, an instrument was created that could in real-time create fairly diverse and complex effects.

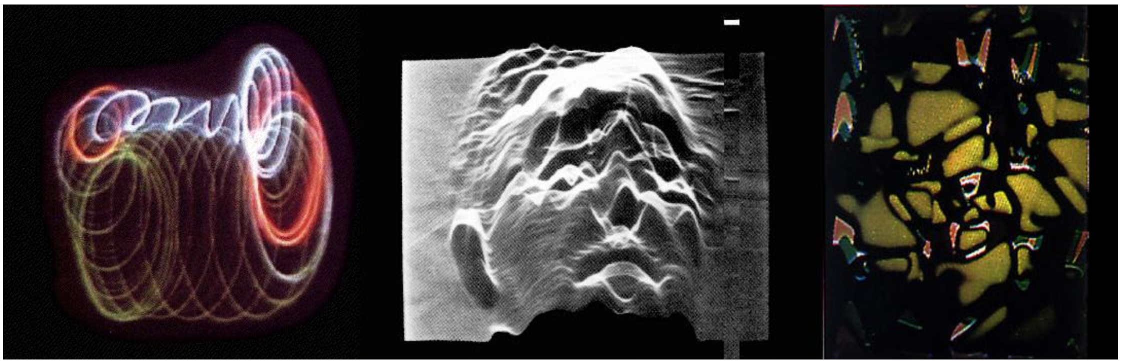


Figure 9: Example output from three different analogue video synthesizers: Bill Hearn's Vidium, 1969 (Top Left), Rutt Etra Video Synthesizer, (Vasulka, 1972) (Top Mid), and Stephen Beck's VSynth, 1973, image originally from (Davis, 1975) (Top Right). All images sourced from audiovisualizers.com

These machines were always very expensive and unstable due to their complexity, and weren't ever really commercially viable, but they still saw use in the cinema and even the mainstream TV industry, as well as the academic arts community. Though it was not their only purpose, analogue video synthesizers have been known to be used for accompanying music at various points in time. Little information is left about them today, and few examples of their output can be found.

In the late seventies, laser lights were connected to computers that could control the path of movement of the light beam with great speed and precision, and so a new interest in accompanying music with real-time controlled synthesized images arose. For a period laser shows were a popular accompaniment to music concerts, and have frequently also taken centre stage position in performances (Spinrad, 2005). Their popularity waned however, arguably because the visual output is limited by the nature of laser light to simple coloured line drawings.

Computer graphics have too been employed to accompany music in real-time. The first computer system made with this intent is the '*Vampire*' (1970), created at Bell labs by Max Matthews as an extension to '*Groove'*, the successor of the first ever computer music application, *'Music 1'* (1957) (Roads & Mathews, 1980). A vast multitude of such systems now exist, the majority of them controlling the visuals through data derived by performing beat and amplitude detection on the stereo mixdown of the music. Popular examples that are both highly likely to be found installed on most personal computers are Nullsoft's Winamp Advanced Visualization Studio, and the Apple iTunes Visualizer.

Most recently, music is frequently performed together with live visuals. At live music concerts, and at clubs with music played by a DJ, there often are projected live graphics performed by what has come to be known as a VJ (for Video Jockey). Initially, a modern VJ's performance could hardly be referred to as live in the strictest sense, since all a VJ did was mix pre-recorded video clips together while altering parameters affecting their playback. VJing technology has since developed, and a more advanced mode of performance has emerged, where VJ's also apply real-time effects to manipulate the video clips' contents. Popular examples of applications used for this purpose are Arkaos (http://www.arkaos.net/), Modul8 (http://www.garagecube.com/modul8/), and Aestesis (http://aestesis.eu/), though there are many more.

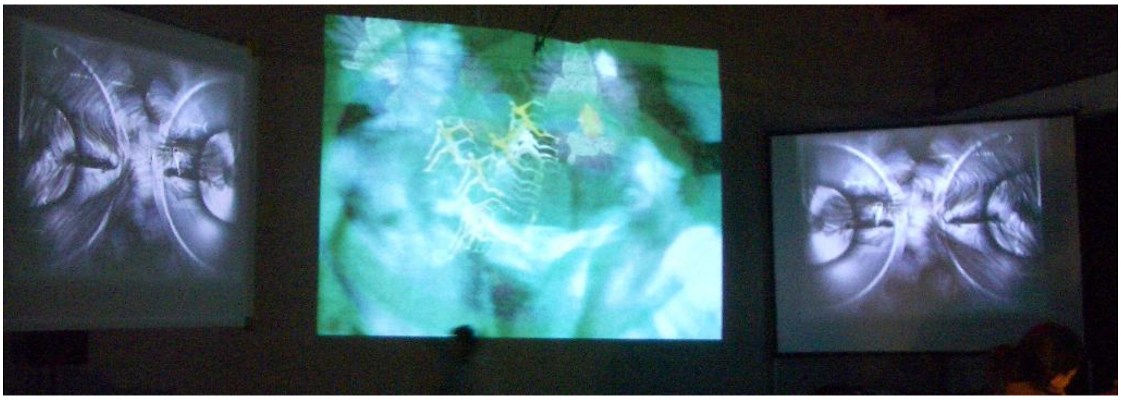


Figure 10: Image from Vizual Contakt Lab 2006 VJ Festival (Copyright Le Collagiste, http://blog.lecollagiste.com).

Figure 11: VJ Vello Virkhaus performing (www.vixid.com, copyright Justin Misch).

Figure 12: VJ Vello Virkhaus accompanying the live group Red Hot Chilli Peppers (www.vixid.com, copyright Justin Misch).

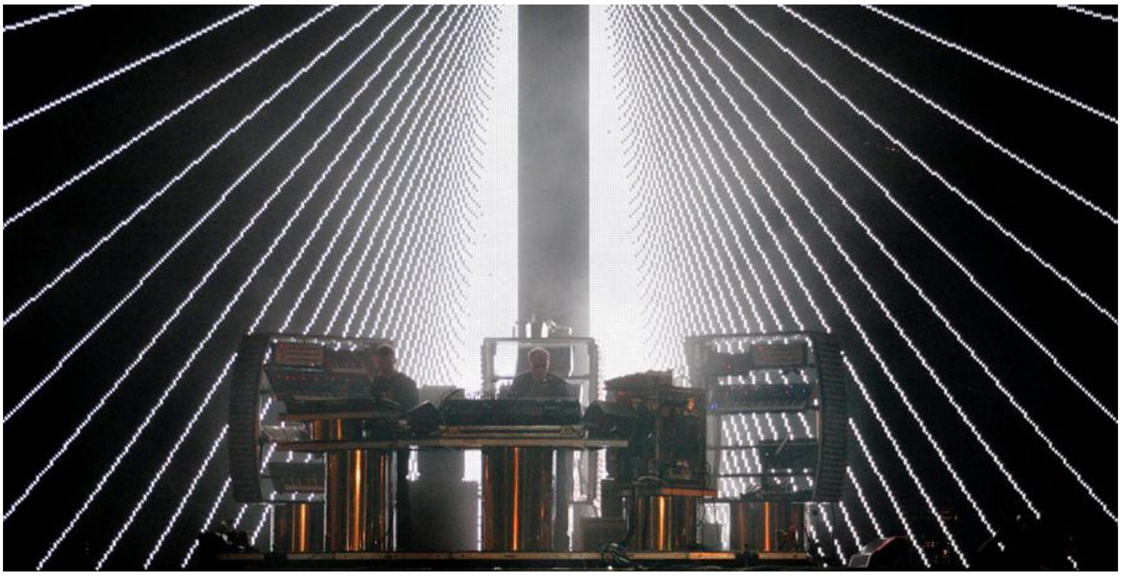


Figure 13: United Visual Artists accompanying the Chemical Brothers performance at Trafalgar Square, 2007 (copyright UVA, www.uva.co.uk).

2.1.3 Audiovisual Composition

Although the premise of the audiovisual composition art-form is perhaps in part implicitly inherent in most audiovisual art, such as visual music with musical accompaniment, or synaesthetic art, it was first explicitly identified and named by Mick Grierson (2005).

Because the work detailed in this thesis largely continues from where audiovisual composition left off, the ideas and arguments underlying audiovisual composition will be presented in more detail. The argument behind audiovisual composition also forms the starting point for the argument behind the art form of Soma, introduced in this thesis.

Grierson defines audiovisual composition as an *"artistic form which takes as its starting point the cognitive actuality of multisensory audiovisual experience"*. Actual and visual music combine, fuse to become a third art form, where one is not simply accompanying the other, but they are both to be experienced as an inseparable whole. Although the premise is bi-modal congruence, it is recognized that for composing a narrative over time, it is also necessary to allow for transitioning between states of congruence and *"binary opposition or total incongruence"* (Grierson, 2005). Arguments for the legitimacy of the art-form follow two paths; firstly, there is significant historical precedent in art, where although not defined as a separate art direction, the premise of audiovisual composition is of important significance. Secondly, there is significant research from the areas of psychophysics and neuroscience on multimodal perception, which corroborates the artistic suppositions of audiovisual composition.

2.1.3.1 Precedent in art

Partly, this topic has already been reviewed in the previous two sections of this chapter, detailing the history of colour organs, visual music, abstract film and synaesthetic art. Further, more current discussion is reviewed by Grierson, detailing John Whitney's differential dynamics theory on audiovisual composition (Whitney, 1980), and the analogous discussion on audiovisual relationships by Adriano Abbado (1988). Michel Chions' widely established concept of synchresis features prominently; Chion argues that synchronized music and/or sound provides *"added value"* to a visual narrative, defining synchresis as:

> *"(...) the spontaneous and irresistible weld produced between a particular auditory phenomenon and visual phenomenon when they occur at the same time" (Chion et al., 1994).*

Chions' ideas on synchresis and added value, are now part of the core of the literature on sound for cinema, and are often employed in the creation of mainstream cinema soundtracks. Chions' own treatment of the topic however is not limited to cinema, but is intended to apply also on audiovisual art/entertainment in general. To further clarify Chions' ideas, Grierson references the following quote:

> *By stating that there is no soundtrack I mean first of all that the sounds of a film, taken separately from the image, do not form an internally coherent*

> *entity on equal footing with the image track. Second, I mean that each audio element enters into simultaneous vertical relationship with narrative elements contained in the image (characters, actions) and visual elements of texture and setting. These relationships are much more direct and salient than any relations the audio element could have with other sounds. (Chion 1994, p40)*

Another effort of creating an audiovisual instrument for live performance, strongly influenced by Chions' synchresis and added value concepts, is Moody's development of the Ashitaka instrument (2009).

2.1.3.2 Related research on multimodal perception

The second path of argumentation for the legitimacy of the audiovisual composition art form followed by Grierson, follows the path of reviewing research from the areas of psychophysics and neuroscience on multimodal perception.

It is already since long established that the human perceptual system is apt at detecting correlated stimuli across modalities, and fusing these into a single percept before their interpretation (Marks, 1978). There are *"vigorous interactions among sensory modalities"* (Shimojo & Shams, 2001).

It has more recently been shown experimentally, that there is significant positive correlation between how closely discreet auditory and visual events are synchronized, and the perceived "effectiveness" of the combined audiovisual stimulus:

> *"(...) the manner in which salient moments in the auditory and visual domains are aligned results in a significantly different perceptual response to the resulting composite" (Lipscomb, 2005).*

Experiments conducted, have also confirmed that we process audiovisual material as a unified percept, and moreover that the temporal window of such audio-visual interactions has been found to be approximately 100ms (Shams et al., 2002). Shams et al reached these results, through the identification of a visual illusion induced by sound: *"(...) when a single flash of light is accompanied by multiple auditory beeps, the single flash is perceived as multiple flashes"*.

More recent research that has followed after Grierson's review, provides evidence that there is fusion of more than just auditory and visual percepts, and that these are processed in combination also with tactile stimulation, depending on the level of congruence between the stimuli presented to the different modalities (Wozny et al., 2008), (Occelli, Spence, & Zampini, 2009). A thorough review on the subject of multisensory processing is presented in (Alais, F. N. Newell, & Mamassian, 2010).

Following this discussion, initiated by Grierson, and corroborated by further findings following his review on the subject, there is strong scientific evidence, that the close correlation between visual and auditory musical events shapes a more effective experience in audiences, to use Lipscombs choice of terminology. As a consequence, an art form concentrated on exploiting the effect congruent/correlated audiovisual stimuli have on their audience, is lent similarly strong validity.

Finally, a body of research that cannot go without a brief mention is that on synaesthesia:

> *"Synaesthesia (Greek, syn = together + aistesis = perception) is the involuntary physical experience of a cross-modal association. That is, the stimulation of one sensory modality reliably causes a perception in one or more different senses" (Cytowic, 1995)*

It is a phenomenon that has been researched extensively, with findings clearly showing that synaesthesia is a neurological condition. Subjects that have the condition, involuntarily experience certain sensory stimulation on more than one modality. Many variations exist; most relevant to the present discussion is the experience of sound as animated colours and/or shapes. Individuals who experience synaesthesia each have their own mappings between perceived auditory stimulus, and the corresponding visual percepts. Furthermore, the people that normally experience synaesthesia form only a small fraction of the general population; although purportedly, also a non-synaesthete can experience synaesthesia, through ingesting psychedelic drugs (Shimojo & Shams, 2001).

At present, it is not feasible to draw conclusions from synaesthesia research, on the specific functions of multimodal perception of ordinary, non-synaesthete individuals. What synaesthesia research was arguably first in contributing to this argument however, is a strong indication that in all individuals and not only synaesthetes, percepts from multiple modalities are not processed as separate streams of information, but are fused early on in the brain, forming a single percept. If this was not the case, a condition such as synaesthesia would not have been possible. This has inspired more recent research that corroborates the assumption with scientific evidence, as is thoroughly reviewed in (Alais et al., 2010), and as was briefly discussed earlier in this section.

2.2 Musical Instruments

Musical instruments and performance fulfil an integral role in the work presented here, though the aim of this thesis is not concentrated solely on contributing to the body of research related to creating new musical instruments. The work presented, is relevant to musical instrument and performance research, despite the fact that the output under control is visual, because the technology used to generate and manipulate the control data is clearly best described as music technology. In fact, one can easily imagine the here introduced practice of mutable mapping, being used in a purely musical performance context.

The practices of developing new instruments for musical performance and for audiovisual performance, share very similar objectives and methodologies to those of present work, warranting the examination of how research that deals with the development and evaluation of musical instruments is carried out.

Following this section, it will be clear why in this work the choice was made to use existing musical instruments, rather than create a wholly new interface for the purposes of capturing expressive musical gestures of performers. Towards this argument, the emphasis of the discussion will be placed on the interface of musical instruments, i.e. how these are controlled by their performer, and what it is in the design of instruments that permits a greater level of virtuosity. It will also be demonstrated that the establishment of a new musical instrument interface is a rare occurrence; in fact most interfaces still widely used today have a history of many centuries.

A development crucial to this work, is the separation between controller and sound generator, a consequence of the development of digital musical instruments, which has engendered the here central notion of mapping between controller and sound generator. It is solely this development that has allowed the initiative taken in this work, of using the control data generated from musical instruments, to instead control the output of visual synthesizers.

It will further be explained, why increasing the complexity of the control data gathered from the live musical performers, is expected to result in the performers to a greater extent being able to employ their advanced enactive knowledge of their instruments, and achieve a higher level of expressivity, in comparison to previous practice.

The role of the audio mixing engineer will be detailed, as it has directly inspired the novel concept of mutable mapping.

Finally, the benefits of the collaborative performance experience will be examined in the context of live musical performance, with the objective of seeing how these may in this work also translate to the context of live visual music / audiovisual art performance. A final reason why the technology and history of new musical instrument controller development is relevant to this research is because music technology has almost exclusively been the forbearer of developments in live visual performance technology. In the current state of the art of live visuals performance, the technology used is most often either directly borrowed from music technology, or used in an adapted form.

2.2.1 Electronic musical instruments and the interface using which these are played

Instrument creators have always been working at the cutting edge of their time's technology, and interest in the development of new instruments has never shown decline. During a time period of little over a century however there have been two big technological breakthroughs that have brought about changes of a magnitude never witnessed before to the musical instruments we use: the advent of electric musical instruments and sound amplification, and the creation of electronic musical instruments, which will be the main focus of attention in this review.

As early as the turn of the twentieth century, in 1896, what is argued to be the first electronic musical instrument was constructed, Thomas Cahill's Telharmonium (Bode, 1984). Predating the radio as well as sound amplification, the electromechanical telharmonium used the telephone network for transmitting sound to customers who would subscribe to remote live performances. To facilitate public performance, several telephone receivers would be used simultaneously, so as to achieve louder sound reproduction. A single telharmonium installation would require several floors of a building specifically adapted for it, and weigh up to two hundred tonnes. Two decades later, in 1917, Leon Thermen's Theremin, an electronic instrument played without the musician ever touching it, gained widespread attention and was to be the first electronic musical instrument to be produced in any significant quantity (Roads, 1996). This was followed a decade later, in 1928, by the Ondes Martenot, Built by Maurice Martenot. The Ondes Martenot was to be similarly successful and widespread as the Theremin, with it still being taught at music conservatories to the present date. Note that later models of the Ondes Martenot had a keyboard added, but they still also retained the original control method of moving a ring attached to two strings at each end, along the instruments front edge. From then till now, a vast number of technological innovations have been made, and as a consequence music itself has been through a great number of transformations (Roads, 1996).

Figure 14: Leon Thermen playing his invention the Theremin (left, public domain image), and the Ondes Martenot with speakers at The Atlier Jean-Louis Martenot in Neuilly (near of Paris) (Right, image under Gnu Free Documentation License, from the Wikimedia Commons).

Despite this rampant progress, the interfaces using which music is predominantly performed have remained largely unchanged. With very few exceptions, all new instruments created during the last century have had an interface virtually identical to that of a pre-existing instrument. Arguably the first electric or electronic instrument to enter the mainstream, the electric guitar developed during the 1920's, to the present date still remains very close to its acoustic counterpart in terms of its interface. After the electric guitar, many electric, electromechanical and electronic instruments have entered the mainstream, eventually leading up to the advent of the synthesizer. Still, the interface with which all were controlled was far from new.

Even synthesizers, being arguably the most different and technologically advanced family of instruments to have been created, began gaining popularity and eventually started becoming part of the mainstream first when Bob Moog decided to fit one of his designs with a piano type keyboard. These first synthesizers were analogue and used the method of Subtractive synthesis, though digital synthesizers later followed, along with many more sound synthesis methods (Jorda, 2005). Instead of developing and adopting new interfaces for the use of the various types of synthesizers that exist, the interfaces for most historically established instruments have been adapted to the use of controlling synthesizers, either by the addition of a tracking system on the traditional instrument, or by the creation of gestural controllers that closely resemble their traditional counterparts. Among the controllers that are made to resemble existing instruments, virtually all existing instruments are now represented.

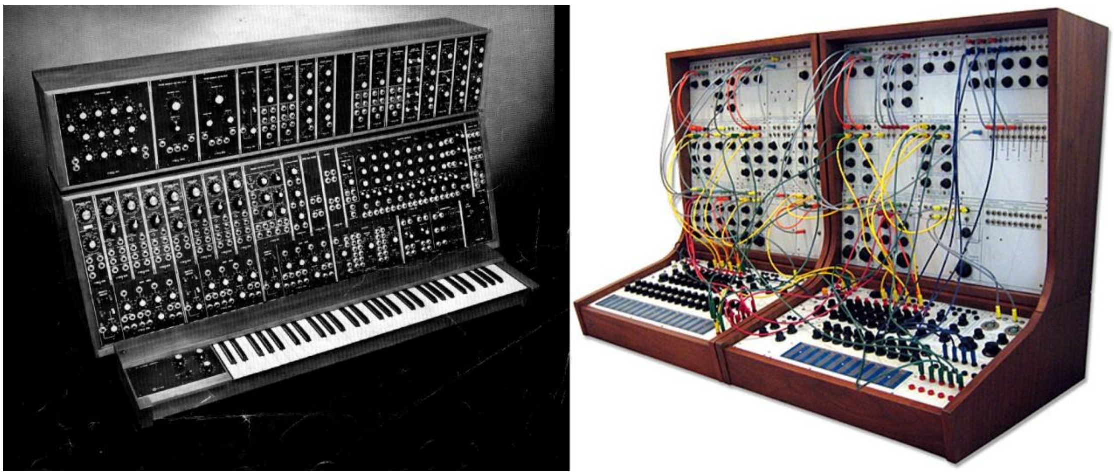


Figure 15: A Moog 55 analogue modular synthesizer rack with keyboard controller (Left, image from original Moog promotional brochure, www.synthmuseum.com), and a Buchla 100 system (Right, image from www.vintagesynth.com). Although the Buchla was developed concurrently to the Moog, and was argued by many to be a better instrument, Don Buchla’s reluctance to fit one with a standard keyboard controller meant that it would eventually be significantly outsold by the Moog. At the bottom of the image one of Don Buchla’s non-standard musical control input designs can be seen.

Throughout this period, the only instruments with a novel interface to gain anything resembling widespread acceptance were the previously mentioned Theremin and Ondes Martenot, and much later the turntable, which when combined with a mixer, was repurposed as a musical instrument with the development of the scratching technique during the 1970’s and 1980’s, becoming by far the most popular new musical performance interface to emerge during the period examined (Jorda, 2004). Anecdotally, the single most successful new instrument of the past century, the turntable and mixer combination, came about entirely by chance, in defiance of all efforts by a great number of researchers to produce a new instrument with similarly widespread appeal.

It is worth noting that even if many new instruments use the same interface as their acoustical counterpart, new techniques for playing these instruments have been developed, as a result of their different tonal characteristics. For example an electric guitar is often played with a technique very different from an acoustic one, and although they all are keyboards instruments, the Hammond organ, the Rhodes piano and the synthesizer may each be played with greatly varying techniques.

Also existing interfaces have not remained entirely unchanged, as there have been numerous additions; these are however better regarded as extensions to the main interface, since the primary interface using which the musician plays still is the same as that of a traditionally established instrument. To name a few examples, with the Hammond keyboard and electronic organs came the addition of switches, knobs and drawbars with which the musician could drastically alter the sound timbre of their instrument. Following the introduction of the synthesizer, means of making the keyboard more expressive also appeared. Joystick controllers, touch strips or wheels were placed near the keyboard, allowing the musician to produce pitch bends, glissandos and vibrato, all impossible with traditional keyboard instruments. Also with the synthesizer, the number of controls using which the musician could alter the sound timbre was greatly increased, giving them the ability to produce emulations of most natural instruments, as well as create many new, unique sounds.

The fact that so few new gestural controllers have gained any sort of widespread appeal is very striking if one considers how many new, novel interfaces are being proposed every day. It is sufficient if one reads through the proceedings of a related conference or journal, such as the New Interfaces for Musical Expression (NIME) conference proceedings, or the journal Organized Sound, to get an idea of the multitude and diversity of new instruments developed every year. It is even more striking if one considers how many new techniques for synthesizing sounds have been developed, giving musicians a previously unimaginable possibility to modify the timbre of the sound produced, even during performance. Together, these facts are seen as indicative of the difficulty of creating a new successful controller for musical performance, that doesn't resemble an existing, established design. Following this argument, in present work, it is chosen to employ already existing gestural controllers, over developing novel ones specifically designed for our purposes.

2.2.2 Analysis of Digital Musical Instruments

Following is the most crucial development in music technology, for the research presented in this thesis. With the advent of synthesizers, the instrument was physically as well as conceptually decoupled into two main components, the *gestural controller*, and the *sound generator*. The gestural controller (often in short referred to simply as the *controller*) may be any different combination of devices that form the part that the musician physically interacts with, such as a piano-style keyboard, a velocity sensitive surface resembling a drum, or even a knob, a button or a slider. All these generate control data, which then needs to be forwarded to the sound generator for it to interpret and produce actual sounds. Instruments that allow this are often referred to as *Digital Musical Instruments* (DMI) (Miranda & Wanderley, 2006), to stress this distinction over acoustic or electric instruments. This separation of the instrument into gesture controller and sound generator was further cemented with the advent of MIDI in 1981, an event that sparked a great increase in development activity for creating alternative controllers (Jorda, 2004).

Wanderley and Depalle (2004), separate the gestural control of sound synthesis into four parts, which are then in turn each analysed separately before they can be analysed as a whole:

- Definition and typologies of gesture
- Gesture acquisition and input device design
- Synthesis algorithms
- Mapping of gestural variables to synthesis variables

In order to design a gestural controller, they argue that it is first necessary to study the musical gestures themselves. Data describing these gestures is treated as a form of signal, which can be recorded, transformed, reproduced and synthesized. While studying musical gestures, they mention that it is of great importance to also study the multimodal feedback that the performer receives, as it forms an integral part of how the performer perceives and interacts with the instrument.

When the characteristics of the musical gestures are known, it is possible to devise a gestural controller capable of capturing them. There are three ways in which this may be achieved (Miranda & Wanderley, 2006).

- *Direct acquisition*, where sensors are employed to capture performer's actions, each capturing a separate feature, such as pressure, displacement or acceleration.
- *Indirect acquisition*, where information about gestures is instead derived from the sound produced by the instrument, by the use of signal processing techniques.
- *Physiological signal acquisition*, such as neural signals, heart and breath rates, perspiration, or muscle tension.

The device that uses one or more of these methods of acquisition to capture the musical gestures is referred to as the gestural controller. Depending on their design, gestural controllers are classified into four categories (Miranda & Wanderley, 2006):

- *Instrument-like controllers*, which in detail reproduce the design of an existing instrument.
- *Instrument-inspired controllers*, which are inspired by an existing instrument, but are not intended for a use identical to that of the instrument from which inspiration was drawn.
- *Extended instruments*, where existing traditional instruments are extended with additional sensors.
- *Alternate controllers*, that are entirely new designs, not based on any pre-existing instrument.

Although all aspects of DMIs have been analysed in great detail by the research community, the one that is of the greatest relevance to this work is that of mapping and will therefore be detailed further in the following section. Giving remaining aspects the same amount of attention however would be beyond the scope of this report, and so readers are instead advised to refer to (Wanderley & Depalle, 2004) and (Miranda & Wanderley, 2006) for recent and comprehensive reviews.

Figure 16: Examples of commercially available DMI controllers, replicating existing instruments to varying degrees (product images from each respective manufacturer's website).

Figure 17: Generic, commercially available non-musical controllers the output of which may be mapped to any parameter of a DMI sound generator (product images from each respective manufacturer's website).

2.2.3 Mapping

In the context of DMIs, the term mapping refers to the pairing of input streams of data, coming from the controllers, to the parameters of the sound generator. This need not be fixed, like in traditional instruments, where for example striking a key on a piano inevitably results in striking one particular string, with a force that is analogous to that applied to the key and a sound duration depending on how long the key is held down. With a DMI, this is only one mapping, out of a multitude of possible permutations; there's no technical reason why the sound's pitch isn't instead dependent on how hard the key is struck, and its amplitude dependent on the horizontal placement of the key on the keyboard.

Mapping can initially be categorized into three general groupings (Hunt, Wanderley, & Paradis, 2002):

- One-to-one
- One-to-many
- Many-to-one

One-to-one mapping is what is most often seen implemented in DMIs. A single input parameter is linked to one single parameter of the sound generator, such as: the force applied to a piano key being mapped to sound amplitude. However, this is not the most natural mapping; rarely does one see examples of simple one-to-one mapping in traditional acoustic musical instruments. Furthermore, it has been suggested that one-to-one mappings may be less satisfactory, compared to more complex mappings (Hunt et al., 2002), (Wanderley & Depalle, 2004), though to verify this it may be necessary to perform more comprehensive experiments.

In one-to-many mapping (also referred to as *divergent mapping*), one single control variable may affect many different parameters of the produced sound. For example, the velocity with which a string is struck, not only affects the volume of the resulting note, but also its duration, and its timbral characteristics.

Similarly, in many-to-one mapping, (or *convergent mapping*), one single parameter of the sound generator may be affected by many different parameters from the controller input. For example, the amplitude of a piano's sound depends both on how hard the key has been struck, but also on whether the string damper foot-pedal has been depressed or not.

Many different approaches to mapping have been discussed over time. A mathematical formulation of mapping has been given in Nort, Wanderley, & Depalle (2004), and complex implementations utilizing highly abstract models have been proposed, such as principal component analysis (Bevilacqua, Müller, & Schnell, 2005), neural networks (Cont, Coduys, & Henry, 2004) and hidden markov models (Kolesnik & Wanderley, 2005), to name but a few examples. Also, mapping need not be linear; many non-linear implementations have been suggested, as discussed in (Bevilacqua et al., 2005), (Cont et al., 2004).

In a DMI, the implemented mapping plays a very significant role; two different mappings, between the same controller and sound generator, have the potential of producing instruments with drastically different qualities. What could be a successful musical instrument using one mapping could just as well be a failure when using another (Wanderley & Depalle, 2004). The qualities defining whether a particular mapping strategy is more successful than another are highly elusive, which has contributed to many drawing the conclusion that defining the mapping may well be regarded as an act of artistic expression in itself (Casciato & Wanderley, 2007). In fact, musicians using DMIs to some extent often do define the mappings themselves, and frequently alter these depending on the needs of the particular performance at hand. Both of these ideas are highly supportive of the artistic performance role envisioned for the concept of mutable mapping introduced in this thesis.

Although a very recent area of systematic research, there has recently been a great increase in research activity on the topic of mapping, the exhaustive discussion of which would require a much more extensive review than that presented here. For further details on the topic, please refer to (Miranda & Wanderley, 2006), (Steiner, 2006), (Wanderley, 2002) and (Wanderley & Battier, 2000).

2.2.4 Embodied / Enactive knowledge, and music virtuosity

Learning to play a musical instrument is a challenging endeavour, which may take several years for an individual to master. It is established that, to achieve virtuosity, performers need to have accumulated roughly ten thousand hours of practice, while simply reaching an average level of proficiency with the instrument requires around five thousand hours of practice (Woody, 2004). A significant portion of what a performer achieves with practice is the improvement of his/her embodied/enactive knowledge of the instrument.

Not all instruments demand the same amount of practice, or allow a similarly high level of virtuosity and expressivity. The kazoo, while being very easy to learn, is limited in terms of the virtuosity attainable with accumulated practice. The violin on the other hand, has a very steep learning curve, but is virtually limitless in the level of virtuosity it allows. A long standing challenge in the community of instrument creators, that has yet to be fulfilled (and may perhaps never be), is to create an instrument that demands a *low entry fee with no ceiling on virtuosity* (Wessel & M. Wright, 2001).

As has previously been discussed in chapter 1, the level of control complexity that a musical instrument provides for direct musical gesture performance, is a precondition to the expressivity of said instrument (Dobrian & Koppelman, 2006), although high complexity does not necessarily guarantee expressivity. This is also why achieving a low entry fee with no ceiling in virtuosity may not be possible; control complexity cannot easily be increased, without making the learning curve steeper.

Besides control complexity, to achieve high expressivity the instrument also needs to be designed in such a manner as to facilitate virtuosity, in other words, to allow the player to take advantage of his capacity for embodied / enactive knowledge. So for example, a collection of knobs, buttons and sliders, while allowing for a significant level of control complexity, does not allow the same expressivity in direct musical gesture performance as a series of strings on a fretboard, because it is not well adapted to a human performer's physical abilities in relation to the task at hand.

For the sake of clarity, it is worth noting here that the inverse is also true in the context of conducting gesture performance: a series of strings on a fretboard would arguably serve the performer quite badly in comparison to an established controller for conducting gesture performance, such as a conductors' baton, or an audio mixing desk.

When performers employ musical interfaces that allow a high level of virtuosity, audiences have been found to be more engaged in the performance (Rodger et al., 2007), also reviewed in (Armstrong, 2006).

A further benefit is that the performers are free to perform other cognitive activities while playing (Hunt & Kirk, 2000), which allows them to more effectively collaborate with other musicians in a group, read a score while performing, or watch a performance which they are musically accompanying.

2.2.5 Collaborative performance, mutual engagement and group flow

The term mutual engagement is in (Bryan-Kinns & Hamilton, 2009) defined as: *"the point at which people spark together, lose themselves in the joint action, and arrive together at a point of co-action 'where you are when you don't know where you are'* (Tufnell & Crickmay, 1990)*".* It is analogous in the benefits it presents, to the state of *flow,* which is by Csikszentmihalyi defined as:

> *"A sense that one's skills are adequate to cope with the challenges at hand in*

a goal directed, rule bound action system that provides clear clues as to how one is performing. Concentration is so intense that there is no attention left over to think about anything irrelevant or to worry about problems. Self-consciousness disappears, and the sense of time becomes distorted. An activity that produces such experiences is so gratifying that people are willing to do it for its own sake, with little concern for what they will get out of it, even when it is difficult or dangerous." (1990)

Playing a musical instrument is an activity that can induce a strong flow sense to the performer, the strength of which varies depending on how challenging the instrument and the performance at hand are for the individual.

Sawyer subsequently identifies the concept of *group flow* (2007), as the equivalent of Csikszentmihalyi's flow concept in the context of group activity, recognizing that the act of working in a group may in itself be a flow-inducing experience. This then helps explain why musicians may often experience an even stronger sense of flow, when performing in a group with their instrument, in comparison to them engaging in solo performance.

2.2.6 The mixing engineer

The concept of mutable mapping introduced with this work, is largely a development inspired by the existing role of the audio mixing engineer. It is thus useful for the argument behind the mutable mapping concept, to present a brief description of what the role of mixing engineer entails.

With the advent of sound amplification, and the previously detailed development of electric and electronic instruments, the need for a new role emerged in the context of live musical performance: that of the mixing engineer. This is detailed here because a similar role inspired by that of the mixing engineer is proposed in the context of this project, stemming from the inception of the mutable mapping concept.

With amplification came the ability to have previously very quiet sources sound arbitrarily loud, allowing a whisper from the singer to be audible over the playing of an entire orchestra, as well as allowing for previously quiet instruments to become established solo instruments in their own right. Later it also became possible, through employing audio signal processing, to significantly alter the sound coming from the instrument, adding reverberation, echo, and a great multitude of other effects. Because all sound now had to be reproduced through loudspeakers however, it became much more difficult for the musicians in an orchestra to balance their individual playing so that the combined music has the desired overall balance. This is because loudspeakers project sound only to a specific direction, towards the audience and away from the orchestra; as a result the musicians were unable to hear their own contribution as part of the whole. Since each instrument could now be arbitrarily loud, the need emerged for someone situated opposite from the orchestra and the loudspeakers, thus having an ideal position for hearing the performance as a whole.

Initially, this was the only role of the mixing engineer: to sculpt the whole of the group's musical performance, by balancing the level of each individual musician's playing. That position turned out to be very advantageous, and quickly the mixing engineer's role became more involved: besides balancing the levels of the whole performance, he/she also became responsible for sculpting the overall sound timbre, by applying effects processing to the sound output of each instrument, as well as to the whole band. This could encompass effects that recording engineers also used in the studio, such as equalization, compression, delay and reverberation, but later came to encompass the entire range of available audio effects.

In the end, several cases emerged where the mixing engineer had a role that was acknowledged as being equally important and creative as the rest of the members in an orchestra, and even frequent occurrences where he/she was as acknowledged to be the main artist, much like is the case with singers or soloists, and their accompanying band.

2.2.7 Performance using Conducting Gesture DMIs

In this section, a mode of performance equally valid and applicable in the context of audiovisual performance using the system introduced in this thesis is presented. It warrants detailing, for the reader to grasp the full breadth of possible contexts of use for the proposed paradigms and systems developed here.

Uniquely for DMIs, because their sound generators can be disconnected from the gestural controller, it is possible to control the sound generator's output by other means than only direct manipulation of the instrument. A sound generator can equally well be controlled using data that is pre-recorded or algorithmically generated. If pre-recorded control data is used, this will be played back from a software or hardware music sequencer, and many parameters may still be alterable in real-time, such as the speed of playback, and the timbral characteristics of the sound generator. If the control data is generated algorithmically, this can be achieved using a great multitude of different methods, ranging from the simplest case of the arpeggiator, to the control of highly advanced algorithmic composition systems that can generate entirely new, complex musical pieces, based on the input of a human operator.

This mode of performance, where the creation of music is not controlled using direct manipulation interfaces, but instead is the result of the performers controlling the parameters of a mediating system, that then in turn generates the final control data, will here be referred to as *Conducting Performance*. The instruments that are intended to be used in such a performance context to control the parameters of the mediating system are referred to as *Conducting Gesture Systems* (Miranda & Wanderley, 2006). A multitude of various conducting gesture systems have been created to facilitate the control of discussed systems (Miranda & Wanderley, 2006), such as baton controllers similar to a conductors baton, from which expressive information such as tempo and amplitude may be derived. A mixing engineer can easily perform in this way, using a variation of his/her usual tools.

A significant amount of music today cannot be performed live in the traditional sense, as it is created entirely in the studio, following an offline process that often does not involve the live recording of even a single physical musical performance. While this type of music was previously often impossible to perform live, its creators now tend to increasingly adopt a conductive mode of performance, thus being able to perform their music to a live audience. It is worth noting however that many argue this mode of performance is not as involving for the audience as it would be to watch a group of musicians perform using direct manipulation gestural controllers (Armstrong, 2006), as was briefly discussed in chapter 1 of this thesis.

Ableton Live (http://www.ableton.com/live) is at the time of writing likely to be the most widespread tool that musicians employ for live conducting performance. It is a software music sequencer that supports many in this section previously discussed ways of working, and there are many different hardware controller options using which a performer can control the application much more directly than if they were using mouse and keyboard.

An interesting final note is to stress how with these developments, the distinction between what constitutes DJing and what constitutes performing live music is blurred to a significant extent. Often, DJs now prepare for their performances from beforehand by sampling loops and short musical phrases from their record collections, and may then mix a multitude of these in real-time, thus blurring the distinctions between DJing, and live conducting musical performance.

2.3 Musical Narrative

Since the art-form of Soma introduced in this thesis builds on the tradition of visual music and audiovisual composition, always involving the element of actual music, it is needed to here clarify what constitutes a narrative in these art-forms, and by extension, in Soma. Most crucially, it is therefore necessary to first briefly understand what absolute music is, and thus what the term narrative refers to in absolute music, before then also extending the use of the term in audiovisual art.

2.3.1 Absolute music

It is not currently established why humans have evolved the facility for music. Many attempt to explain its emergence through biological evolution, arguing it afforded individuals certain reproductive and survival advantages. An alternative suggestion is that facility for music instead emerged through exaptation: *"(...) a structure or attribute whose current use differs from its originally evolved function"* (Livingstone & W. F. Thompson, 2009), drawing from pre-existing facility in humans, e.g. for emotion, motor control, auditory scene analysis, language, etc. Livingstone & W. F. Thompson (2009), propose that music originated when humans developed the facility for forming a *Theory of Mind*: the ability of an individual to *"recognize the emotional and mental state of conspecifics"*. More specifically, through an advanced form of ToM, *Affective Engagement*, specifically associated with human behaviour and cultural practice.

Music has historically been unique among the arts and other human expression, in that nearly exclusively its content is not representational, but self-referential (Besson & Friederici, 2005). A musical narrative does not set off to convey stories as a series of events, real or imagined, that bare any correspondence to the real world, as would a narrative in the sense that the term is used in for example literature, theatre or cinema.

Very few universals have been found to be common to human musical experience, across all musical traditions, and these are restricted to very basic elements (Nettl, 2000). Without at all attempting an exhaustive review, the primary identified universals in absolute music will be detailed: Music is shaped from a sequence of musical *pitches* forming a *melody*, the *rhythm* being the duration of musical notes and how these are grouped together, *tempo* refers to the overall speed of the music, and *timbre* is a term that defies exact definition, but is loosely described as that which distinguishes one instrument from another, when they play the exact same score (Levitin, 2006). A musical piece needn't encompass all of these, there are many examples where music is not melodic for example, or doesn't follow a tempo or rhythmical meter. Building on culturally learned expectations of what musical event usually follows another, a musical narrative is shaped, through the musician choosing when these expectations are to be fulfilled or suspended, to suite the musical intent he/she wishes to convey (Levitin, 2006). This allows for the two fundamental elements of any aesthetic narrative to be employed, of suspended expectations, and of tension and release.

Many sets of rules, often very precise, have throughout history been formulated for how a musical narrative is shaped, beyond the above universals, however these are always specific to their time and culture, sometimes sharing very few elements in common: to name only a few examples of such traditions, there is Indian classical music, western classical music, various European folk music traditions, various forms of traditional African polyrhythmic drumming, serialist, electroacoustic, acousmatic music, thousands of very specific sub-genres of popular music, of urban club music, etc.

What has been found to be universal however, are musical behaviours and functions such as dance, use of music in rituals and ceremonies, and music's connection with affect (Livingstone & W. F. Thompson, 2009). In this context, Livingstone & Thompson reference a broad definition of music, by Cross (2003): *"music embodies, entrains and transposably intentionalizes time in sound and action"*. This definition also encompasses actions that follow a musical narrative, such as dance, and suitably includes also the notion of physical performance through musical gestures, that is a central notion to the work presented in this thesis.

Drawing from the above, one possible view emerges, which will serve as the working description of music in this thesis: Music is a form of time-based, and thus narrative affective communication, which allows for the transmission of emotive content differing from what could be represented through language or other art-forms alone, and which therefore cannot effectively be conveyed in any other medium than itself. Music, as presently understood, defies further conclusive formalization. Musical experience, beyond the few identified universals, emerges only through each individual's notion of it, as shaped by his/her previous experiences and cultural heritage.

2.3.2 Narrative in visual music and audiovisual art

Although music has likely always been used in conjunction with other art-forms, the relationship has always been complementary. Music has remained unique in that it has persistently been impossible to transcribe a musical narrative, to any of these other art-forms, with its emotive content still remaining largely identifiable.

As is reviewed in section 2.1 of this thesis, it has been increasingly common in recent times that narratives analogous to actual music are recreated in painting and moving image, and throughout history musical narratives have also been conveyed through dance. The content in all these contexts is never direct transcriptions of specific pre-existing musical pieces, but instead on some level analogous to actual music. When actual music and visual music or dance are presented together, again these work complementarily, not through direct transcription between them.

Just as we saw that actual music largely defies a rigid formalized and exhaustive definition, it follows that also combined audiovisual and other musical narratives, such as Soma, similarly defy formalization. The notion of narrative in these art-forms can only really be given meaning through referring to the corresponding notion in absolute music, as is also discussed in section 2.1 of this thesis. While artist active in audiovisual arts may have formulated frameworks of rules for narrative in visual music and audiovisual art, none of these can constitute an exclusive definition for what is a universal audiovisual narrative, just as is not the case with absolute music.

Recently, Evans (2005) attempts to define the basic elements of what a visual musical narrative consists of, which to this author reads very much like a proposal for what the universals of visual music may be. Fundamental in Evans' argument, is that, just as musicians and audiences have a tacit understanding, from their upbringing and culture, of what is harmonious and inharmonious music, and what the usual sequences of musical events are, so too in visual culture there are heritages of what describes a harmonious and inharmonious image. Therefore, at its most basic level, a visual music narrative is immediately analogous to that of absolute music: *"it is possible to resolve visual dissonance to consonance, and so move a viewer through time in a way similar to tonal harmony in music"* (B. Evans, 2005).

Grierson (2005) then adds, in his discussion on audiovisual composition which is reviewed in more detail in section 2.1.3 of this thesis, how a narrative can be constructed also vertically, between congruent visual and auditory musical events.

Just as suspended expectations and tension and release are used in the visual and auditory musical narratives separately, they may also be applied in the narrative of how actual and visual music correspond. By first establishing the notion of congruence between auditory and visual elements, a narrative in the correspondence between sound and image can then be created, through invoking previous expectations on audiovisual congruence, as well as inducing new expectations. A narrative in the audiovisual relation can then be shaped, in which the artist chooses as time progresses, whether to fulfil or suspend these expectations.

Both Evans and Grierson reference others who have conducted work with a similar goal, perhaps the most extensive treatise on the subject being that of John Whitney (1980), with elements of the writings of Sergei Esenstein (1947) (1949) also being applicable to musical visual and audiovisual contexts.

For the context of this thesis however, the subject of musical audiovisual narrative need not be further explored. Drawn from the above discussion, the most general definition is sufficient to support and contextualize the arguments that follow:

A musical audiovisual narrative builds on previous culturally or congenitally formed expectations on auditory and visual harmony and consonance, and on the vertical relationships between congruent auditory and visual events. The artist then draws from his/her explicit and tacit understanding of these expectations, which guides him/her through the choice of when to fulfil and suspend them. Thus, the universal tools for progressing aesthetic narratives may be employed, of suspended expectations, and of tension and release.

2.4 Programming as art: Creative Code

The visual output from the system introduced in this thesis, is unlike the majority of previous practice not pre-rendered. Instead, it is generated programmatically, synthesized by the computer just before the instant it is going to be displayed. This builds on the modern tradition of programming as art, where the artistic medium is program code. It is therefore necessary for the argument presented in this thesis, to briefly detail the history, practices and technology of the programming as art field, so as to better explain the rationale behind the choices made for present work, of allowing artists to create their own procedural graphics and integrating these in with the system designed here.

If asked what computer art is, most people's immediate reaction would be to mention the computer generated images, animations and feature films that have become a mainstay in today's media, created using the multitude of available graphics software. Besides this large amount of work however, created using existing software, there is also a multitude of work for which the medium of creation is programming itself, where the piece of art, is not *created* using a program, it *is* the program. In other words the material of which it is made is not paint, collections of pixels or plotter graphics on paper, but program code.

Looking back to the history of art, one will see evidence of the use of computational logic and algorithmic symmetry even long before the invention of computers and other calculating machines, a common example being Islamic art. With the advent of the very first computers, artists were quickly involved in using these for creating images, animations and music, often being instrumental in advancing the technology of their time to facilitate achieving their goals.

Programming as a creative medium is unique in that it allows the creation of highly dynamic work, capable of producing continuous output as it reacts to control input. It again uniquely also allows the artist to define his own algorithms with which a piece is to be rendered, thus not constraining the artist only within the limited capabilities of the existing tools at hand.

Ben Laposky is largely credited with being the first to create computer art, when in 1950 he exhibited a collection of photographs entitled "oscillons", realised by controlling an oscilloscope with an analogue computer, and photographing the result.

Herbert Franke's work was strikingly similar in how it was created, and appeared only slightly later in 1952, probably realized in a parallel development to Laposky's. It is John Whitney however that is by most credited to be the pioneer of digital computer art, which he produced during his stay as the first artist in residence at IBM, between 1966 and 1969, although he too had also worked with analogue computer graphics for a long time before that.

Many more artists have followed in their footsteps. Some pigeonholed themselves squarely as artists, and thus took help from computer scientist to realize their ideas. Others, more interestingly, decided to take on both roles, regardless of whether they had a science or art background, and chose to write the software that would realize their ideas themselves, thus effectively making programming their medium for creation. From the very early pioneers to the present date, many more have chosen to use computer programming as their creative medium. A comprehensive repository detailing the biographies and work of these may be found at the website of the Digital Art Museum (http://www.dam.org/).

2.4.1 New Media art

The umbrella term *New Media Art* is used to group together all art that is created using "new media" technology, a term that was in turn employed to refer to the new, digital media that has come to supersede traditional print and analogue broadcast media (Dixon, 2007) (Tribe, Jana, & Grosenick, 2006). This includes but is not limited to the internet, interactive multimedia software such as computer games, as well as all new combinations of these and/or digitized traditional media, at least according to one definition. The term is disputed and relatively loosely defined however. It is mainly used to refer to all artistic work that cannot be viewed or distributed using traditional media, but requires the use of digital, interactive technology for its presentation. The project proposed in this work fits under this broad definition, as do artworks that are for example implemented as internet web sites, interactive software applications, or involve the display of computer graphics as part of a multi-media performance. For the realization of new media artwork programming is often required, and often development environments such as those detailed in chapter 2.5 are used for this purpose by the artists. There are many festivals dedicated to new media art, such as Optronica (http://www.optronica.org, Onedotzero (http://www.onedotzero.com/), Prix Ars Electronica (http://www.aec.at/) and Transmediale (http://www.transmediale.de/), to mention a few.

To present a short example, modern dance performances have undergone a recent, significant transformation, to more and more often include real-time computer graphics or real-time processed video as part of the performance (Dixon, 2007). Video cameras, or sensors attached to the dancer's body, are employed to provide input data to software, which is then used to project moving images on to the stage. Dancers can then interact with the graphics projection, and it is most often made part of their choreography to do so.

The systems and concepts detailed in this thesis, are all highly relevant, not only to the context of live audiovisual performance, but also to live new media art performance, for example involving sensor data from dancers, or theatrical performers on stage, rather than only from musical instruments.

2.4.2 Demo Scene

Arguably the largest community of computer art can be found in the *demo scene* movement (Tasajärvi, Schustin, Stamnes, & Tolonen, 2004), (Scheib, Engell-Nielsen, Lehtinen, Haines, & P. Taylor, 2002), (Polgár, 2000), which uniquely is a strong youth culture movement. It is even possible to draw a parallel between the demo scene's development, and how present day graffiti art emerged, both being spawned in underground, frequently law-breaking youth culture movements (Gonring, 2009). The roots of the current demo scene trace back to groups of software crackers, that accompanied their releases of cracked software with small programs of their own, the functionality of which was to display graphics demos, sometimes accompanied with music, to credit the members of the group. These evolved to become quite elaborate productions of real-time graphics and music, often taking the effort of more than one person, and many thousands of man-hours to produce. From early on, such *demo groups* began to compete with each other on who would produce the best demo, with there now being numerous highly popular recurrent yearly events named *demo parties*, where demos are presented and winners are elected in a number of different categories.

Although still a relatively young subculture, the aesthetics of it have begun to spread outside of the demo community. Many demo groups have advanced to become developers of computer games, often taking the aesthetics from the demos with them and incorporating them in commercial games productions. Popular musicians many times draw inspiration from, or simply sample, the often very characteristic music of demos, and graphics designers, animators and VJ's are too often inspired by the aesthetics of demo productions. If the demo-scene will ever gain widespread mainstream influence in a similar manner to graffiti still remains to be seen however.

It can be argued that the particular aesthetics of demos do not only stem from the fact that they are created by members of a particular subculture. They are also made distinct from the fact that the output is predominantly generated procedurally; a means of expression that allows creating graphics that would not be possible using existing software tools, a point also discussed by John Maeda (2004). Interestingly though, in demos, the additional advantage of real-time execution is never leveraged, to make presentations that are in any way dynamic or interactive.

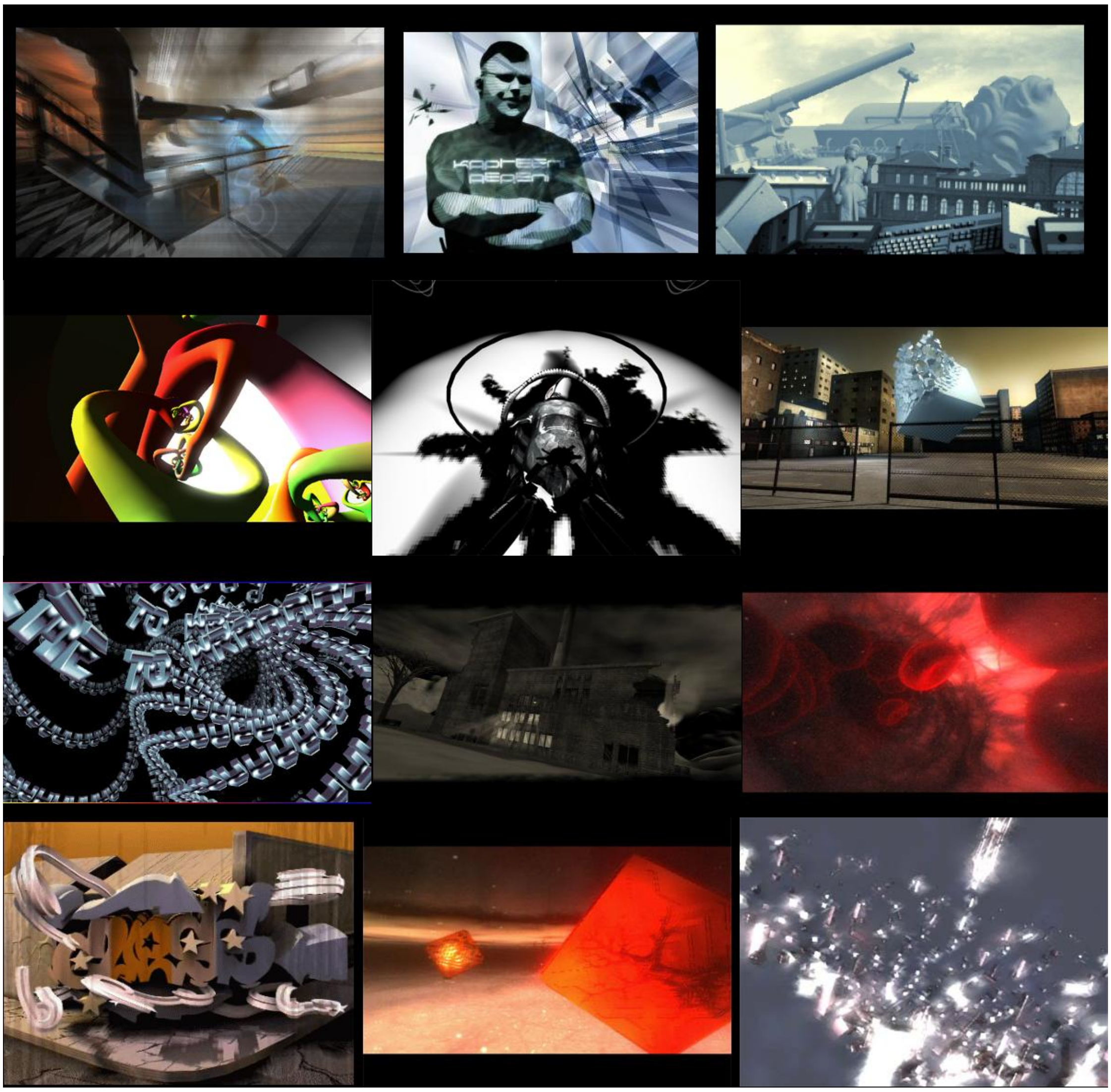


Figure 18: Sample screen captures from 12 different Demos, to illustrate their sometimes very particular aesthetic. Note that a demo is best appreciated when viewed running in real-time on a computer; the reader is therefore encouraged to visit www.pouet.net, where the majority of Demo productions are available for free download (all demo screen captures from www.pouet.net).

2.4.3 Live Coding

Very recently, a practice referred to as *Live Coding* has emerged, in which the artist writes program code as a means of performance, while presenting the output of the written program in conjunction with a projection of the screen that contains the program code (Collins, McLean, Rohrhuber, & Ward, 2003). This practice requires the use of specialized programming environments capable of interpreting the code written on the fly as it is typed, without restarting or recompiling the whole program. It is a practice that is used both for performing music and graphics, sometimes both at the same time, by a group, or both by the same performer. It is briefly mentioned here, partly because live coding forms part of the previously established practice of live audiovisual performance, and because live coding was in an early prototype system considered for its usefulness in defining mappings live during performance, as is further discussed in section 5.2.

2.5 Programming for non-computer scientists

Since it is here proposed that visual artists and other users of the Trinity system, should ideally create the visual synthesizers that they use themselves, a question is raised of whether this is a reasonable expectation to have. Towards addressing this question, the history of making computer programming more accessible to non-experts will be briefly covered, to show what has led to the recent developments that part of this work capitalizes on.

Programming computers is a task that is still by most regarded as a practice reserved to those that are prepared to devote a significant amount of time and effort into learning it, by pursing academic studies or becoming a recluse, spending most waking hours in front of a computer screen. To a certain extent this preconception holds true; fully understanding all aspects of computer programming does indeed take a great deal of effort, to such an extent that it is considered that a computer science degree is only the beginning of a learning process that may take another decade, for the individual to be able to claim that he/she has in the end mastered what is sometimes even referred to as “the black art” of computer graphics programming.

Herein also lays the most common misconception about programming. At the present date, not all that much work is required to learn to program provided one chooses to work using a specialized programming environment for novice programmers. These allow achieving remarkable results, even though it may not be as capable as the full-fledged development environments using which commercial applications are created.

Looking at the history of programming language development, a clear trend is visible over time, of devising higher level languages, which are easier for humans to understand and work with, but more demanding for the computer to execute (Louden, 2003).

A parallel development has been the creation of often highly specialized, domain-specific languages, which at the cost of versatility, often are much simpler to use than general programming languages intended for expert use. Since in the context of the present project we are primarily interested in the feasibility for non-experts to program computer graphics software, this is what the present discussion is going to concentrate on, without it meaning that analogous advancements have not been made also in other fields of computer use.

2.5.1 Visual programming languages

Visual Programming Languages (VPL) are likely to be the most widespread in this context, of which virtually all are implementations based on the *dataflow* paradigm (Johnston, Hanna, & Millar, 2004). Initially the dataflow paradigm was not developed with the goal of ease of use in mind, but was implemented in textually defined languages, with the great benefits towards massive parallel processing that the paradigm affords. It was only much later that visual dataflow programming was made possible, and its significant ease of use was discovered. Using the dataflow paradigm one doesn't need to type to program, instead programs are defined through drawing a directed graph, placing interconnected boxes on a blank canvas, each box of which performs a particular function to the data it receives through its in-connection (inlet), before it sends it on through its out-connection (outlet). Each of these boxes could either be a program itself, described in the same manner, or a program written in a lower level language, as are the building blocks provided in such a languages library. Dataflow languages have the advantage of being very accessible, with a low learning threshold for beginners, while still being able to support the development of complicated programs, which depending on the environment chosen also often execute very efficiently.

The most widespread VPL for multimedia purposes, the Max/MSP/Jitter combination, is a dataflow environment that has a long history of being used by musicians, new media performers and visual artists as their de-facto tool of choice. However there are many more similar applications with varying capabilities, many with wide user bases. Pure Data (Puckette, 1996) was created as an open source successor to Max, written by Max's original creator Miller Puckette, and although similarly popular for music and multimedia applications, its graphics capabilities are not as developed as those available in the Max/MSP/Jitter package. Other notable dataflow environments are VVVV, Touch Designer, Ventuz, Salvation, Quartz Composer and VSX Ultra, all capable of allowing non-experts to program real-time computer generated graphics of varying complexity. Of course the ones listed here are only a few examples of such environments, there being many more.

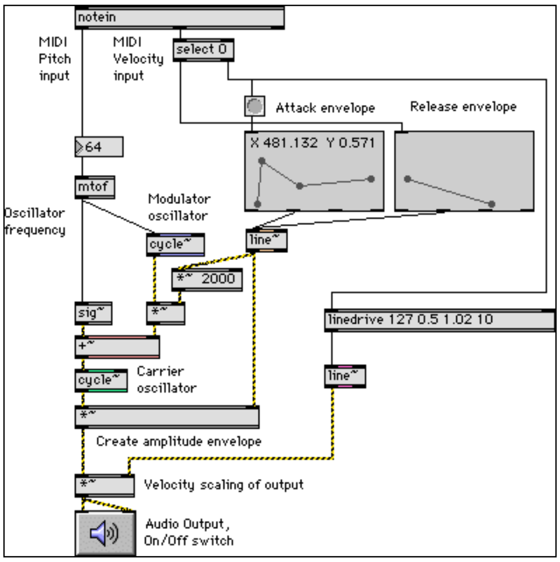


Figure 19: Example of a programs visual representation in the dataflow programming environment Max/MSP

Dataflow languages lend themselves particularly well to programming multimedia applications, as the dataflow paradigm is ideal for the description of Digital Signal Processing algorithms (DSP), either processing audio/video or control messages, all task that are very common in multimedia software production. Many such languages now also allow the implementation of real-time computer graphics, through providing connections to lower level graphics libraries such as DirectX or OpenGL. They are not equally successful in this task however, due to the inherent, great difficulty with which dataflow languages handle the implementation of data-structures and of recursive procedures (Johnston et al., 2004), both very important in the context of programming real-time computer graphics.

These limitations can be circumvented, through the creation of these problematic implementations in the form of plug-in modules using another language, but if one considers it more closely this only brings up one more drawback of dataflow programming; If something that the dataflow environment does not easily support is needed, the only way of successfully implementing this is through implementing it in a lower level language. This means that to implement this functionality, the programmer has no choice but to learn a new, lower-level language, which is considerably harder to use in relation to VPL. To circumvent this limitation, many such environments have also embedded small, more easily learned script languages, using which a programmer can tackle some of the tasks that the dataflow paradigm does not support well. Still, in this case, the inherent simplicity of dataflow languages is gone, since instead of the user programming in one, easy language, he has to juggle between two, and sometimes three languages of different paradigms, in order to achieve the results desired.

Despite these limitations however, if one stays within the constraints of what the chosen dataflow environment supports, it can be a very powerful tool, using which very complex applications can be created with remarkable ease, and with often high computational efficiency.

2.5.2 Procedural languages, and the Processing language

The vast majority of programming languages in use today are procedural languages. The more popular languages are C, C++, Java, C#, Pascal, Ada, with there of course being many more. Since this is the most widespread programming paradigm, it is only reasonable to try to make a simplified procedural language, which may be used to teach non experts programming, while also introducing them to the concepts used in the lower level languages so that there is less of a transition if they later wish to advance into using these. Additionally, the fact that the same paradigm is used has the significant advantage that novices gain access to the vast resources available through bibliography, programming examples and tutorials in their specific domain area. From this, they may draw inspiration and ideas that are easily transferable to the language they use, something that would not have been the case to the same extent if they were programming in a paradigm that was significantly different, such as the two previously discussed.

One such initiative was John Maeda's language Design by Numbers (DBN). He created it to use it as an aid in teaching the *"idea" of computation to designers and artists* ("Design By Numbers," n.d.), to his students at the MIT Media Lab Aesthetics + Computation group. DBN is a highly simplified language that does not attempt to be a general programming tool; instead it only supports a limited set of basic 2D drawing instructions, using which one can in code describe the creation of images and animation. It comes with its own simple programming environment, using which an iterative approach to programming is encouraged through dividing the interface into two windows: to the right the program code is entered, while to the left the visual output produced by the program is displayed. Although DBN has been successful in that it has proved its value as a tool to teach the fundamentals of computer graphics programming to visual artists, its greatest contribution has been that it served to inspire the creation of its successor, Processing.

In what is indicative of DBN's success in teaching programming, two of John Maeda's students, Ben Fry and Casey Reas, were the ones to conceive and create its successor, Processing (Reas & Fry, 2007) (I. Greenberg, 2007), DBN was very limited in what could be created using it, and was only intended to be an introduction to programming from which students would quickly move on. While still being as easy to learn as DBN, Processing has the great advantage of being able to grow with its user indefinitely, in effect achieving the very elusive combination of both having a very low learning threshold, and also virtually no ceiling for how advanced programs may be created with it. It has retained the simplicity of DBN also in that it comes with its own, very accessible programming environment. Since Processing is going to play a central role in this project, it is worth describing it in more detail than previously documented languages.

In the Processing Integrated Development Environment (IDE), programs are named *Sketches*, to emphasize the explorative approach with which users are encouraged to program (Reas & Fry, 2007). Each sketch can also easily be exported either as an executable program, or a java-applet that can be embedded in a web page.

In its simplest mode of use (referred to as basic mode), Processing is very similar to DBN. Where Processing differs however, is it allows the user to transition to another two modes. The first is the advanced mode, in which it is also possible to define functions, access input device events, and even define classes, thus being able to use object oriented programming. It is also possible, within the Processing environment, to use external tools distributed as libraries, all of which are designed to be easy to use and integrate. In its third mode, the programmer abandons the Processing IDE environment, and instead transitions into programming Java, while still using Processing, albeit now as a code library. The advantages of using Processing as a means of teaching programming are evident: students transition from one mode to the next, without discarding any of the knowledge gathered in the previous mode of use, and in the end find themselves programming in a full-fledged Java environment.

Processing today has, despite it being relatively new, a wide user base, as is evident by the continuous activity at the related official and unofficial web resources, primarily the official Processing forum (“Processing Discourse,” n.d.). The fact that it is free and open source has contributed to the creation of what is a very active user community, which has also contributed many libraries to extend its functionality. Also its use as a teaching tool for programming has spread to many more institutions besides the MIT media lab, where it was first used.

Its success has also further spawned similar initiatives, such as one that follows a similar approach but applied to C++ programming, OpenFrameworks (http://www.openframeworks.cc/), and a language that too makes use of the Processing development IDE, and is intended for programming a separate electronic I/O board, named Wiring (http://www.wiring.org.co/).

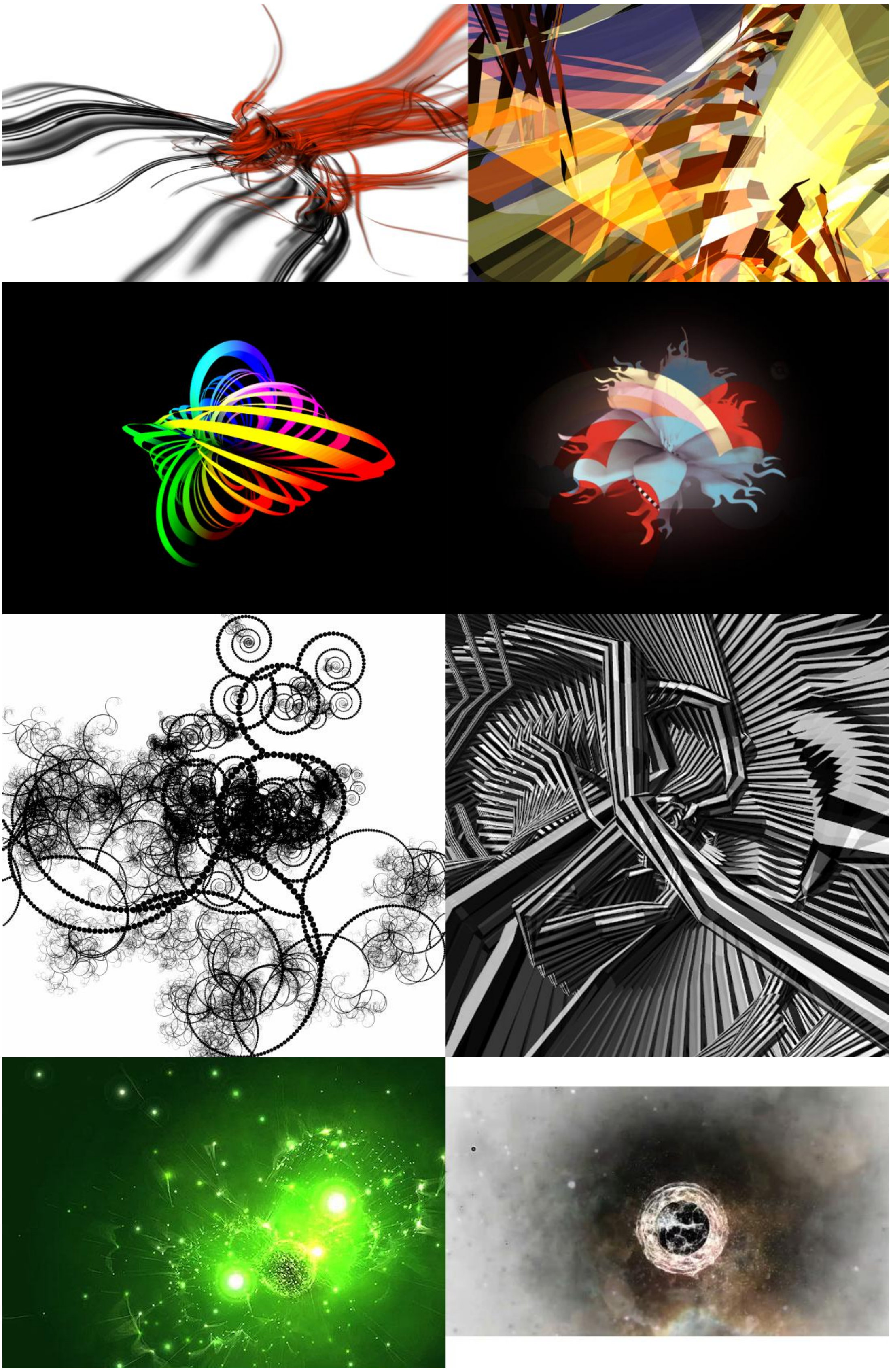


Figure 20: Images from the Processing user group on the image hosting website flickr ("Flickr: Processing.org," n.d.), to exemplify the great variety of output that is possible using the Processing environment. Note that not all of the above can run in real-time as animations.

2.6 Review of related research and evaluation methodology

Identifying what research and evaluation methodology should best be used for this work was far from straightforward. Following an extensive literature review, it was found that although there is much discussion on methodology in related research communities, and many proposals for adopting methodologies from other disciplines have been made, there still is little conclusive agreement. It was therefore here necessary to review the arguments in the research community, and make an informed choice, with precious few pre-existing case studies to draw from. As will be seen in section 2.6.3, six PhD theses with content analogous to this work were reviewed, and in none was a specific methodological choice explicitly identified for the research it presented.

Because of the above, the present section is more than simply a review of the established methodology for the disciplines related to the work presented here. It is also a brief summary of the extensive research conducted to critically review the methodological argument in the disciplines relevant to this thesis, towards the goal of finally concluding what methodological approach to follow for conducting the research in this thesis. As such, the review and selection of methodology is, together with the subsequent case study embodied in the research carried out, a research contribution also on its own merit, to the research community's on-going discussion on methodology.

The project has been found to be highly transdisciplinary, primarily involving the New Interfaces for Musical Expression (NIME) research agenda of the music technology discipline (Miranda & Wanderley, 2006) and human computer interaction (Sears & Jacko, 2007).

The system is relevant to NIME research despite the fact that the output under control is visual, because the technology used to generate and manipulate the control data is clearly best described as music technology. Moreover, the visual output is visual music, a topic which has often been discussed previously both in the proceedings of the NIME conference, and in music related journals such as Organized Sound, and the Computer Music Journal.

As will be discussed in the immediately following 2.6.1 section, NIME research is methodologically highly relevant to the research area of HCI, and as such, also present work is relevant to HCI, for the same reasons as is NIME. Additionally, present work is relevant to HCI, also since it introduces a novel concept for the design of user interfaces: the re-purposing of highly advanced embodied knowledge, learned for one conduct, to a conduct different than that for which it was initially learned.

Present work is most relevant to the Reality-Based Interaction (RBI) area within HCI, as the musical gestures that live musicians perform on their instruments, are adapted from the ‘real’ world of musical performance, and to some extent re-adapted to the computer-enabled world of the live control of visual music. Within HCI, the project is furthermore found to also be clearly falling under the Third Wave, or Third Paradigm of HCI research.

2.6.1 Evaluation of Music Technologies

As argued in the introduction of this chapter, because the work presented in this thesis is in many ways related to the music technology research discipline, it is useful to outline what research and evaluation methodology has been developed for the discipline, to gauge its potential usefulness towards evaluating present work.

The inception of the academic NIME research community is comparatively recent, with the first NIME conference taking place in 2001. As such it has not yet developed a body of research and evaluation methodology suitable to be established as common practice by its members. Orio et al. (2001) argued that, since using a DMI amounts to controlling a task implemented as a computer process under specific constraints, it may be regarded as a special case of the broader subject area of HCI, thus making it possible to borrow tools for the evaluation of DMIs from this discipline. A number of researchers have since proceeded to examine how evaluation methodology from HCI could best be adopted for the evaluation of new NIME research, there now being a number of such efforts published. Several successful attempts have been made, but the methodology applied has predominantly been difficult to generalize, as it in each case applied only to a very particular type of instrument and task (Orio et al., 2001).

Most crucially however, the NIME community seems to have largely overlooked the need also in the HCI community for evaluation methodology that can be applied in new contexts, such as RBI (Christou, E. L. Law, Green, & Hornbaek, 2009), or third wave/third paradigm HCI (Harrison et al., 2007). Because of the great relevance of RBI and the third wave/third paradigm of HCI to this research, as well as to NIME research, they are further discussed in section 2.6.2.2. Furthermore there has been significant disagreement within HCI on evaluation methodologies which has still to be settled. This disagreement is exemplified by the today arguably still relevant ‘damaged merchandise’ controversy (Gray & Salzman, 1998) which is extensively detailed in the (Kaye & Sengers, 2007) alt.chi contribution, as well as by recent discussion, e.g. (Dourish, 2006), (S. Greenberg & Buxton, 2008) and (Crabtree, Rodden, Tolmie, & Button, 2009). The controversy is here detailed in section 2.6.2.1. Discussion on how these disagreements are relevant to NIME research is missing from the NIME effort on adopting HCI methodology. A plausible risk is therefore apparent, that without diligent care, researchers in the NIME community may repeat previous mistakes of the HCI community, or adopt findings from HCI in contexts incompatible to those for which they were intended.

In the related area of Computer Supported Collaborative Work (CSCW), research has just recently begun on using it as a viewpoint for examining music creation, including how such research should best be evaluated (Bryan-Kinns & Hamilton, 2009). This work too is however still in its early stages, and is not yet widely employed, established methodology.

More recent research on creating new musical instruments concentrates on DMIs. The evaluation of such instruments may, besides examining the DMI as a whole, therefore also be broken down to examining each component of the DMI individually, as detailed in section 2.2.2.

Hunt and Kirk (2000) characterise the constraints applying to the simultaneous control of multiple parameters in real-time as being the following:

- *There is no fixed ordering to the human-computer dialogue.*
- *The human takes control of the situation. The computer is reactive.*
- *There is no single permitted set of options (e.g. choices from a menu) but rather a series of continuous controls.*
- *There is an instant response to the user's movements.*
- *Similar movements produce similar results.*
- *The overall control of the system (under the direction of the human operator) is the main goal, rather than the ordered transfer of information.*

- *The control mechanism is a physical and multi-parametric device which must be learnt by the user until the actions become automatic.*
- *Further practice develops increased control intimacy and thus competence of operation.*
- *The human operator, once familiar with the system, is free to perform other cognitive activities whilst operating the system (e.g. talking while driving a car).*

Although HCI proposes many methods for the quantitative measurement of user interface performance, most successfully by the application of Fitt's law and its derivatives (MacKenzie & Buxton, 1992) (Murata & Iwase, 2001), these methods translate poorly to the task of evaluating gestural controllers for musical performance (Wanderley & Depalle, 2004). Successful attempts have been made to apply HCI evaluation methodology to the task of evaluating DMIs, but the methodology applied has been difficult to generalize, as it in each case applied only to a very particular type of instrument and task (Orio et al., 2001). Besides absolute, quantitative measurement, some success has been met in applying methodologies derived from HCI to describe a taxonomy of gestural controllers (Orio et al., 2001), using which it is easier to display the differences between them, although this again does not apply to all types of gestural controllers available.

The advances that have been made in this area now make it easier to compare gestural controllers between each other in terms of how many control channels they offer, what the data resolution of each channel may be, what range of control values they allow the performer to transmit, and other similar, directly measurable quantities. Most notable among this work, is the quantitative measurement of the suitability of particular sensor types for the control of different sound control parameters (M. Marshall & Wanderley, 2006). For facilitating the accurate comparison between completed gestural controllers as a whole, each being comprised of a multitude of different such sensors, there is still a lot more work left to be done however. Although many attempts have been made, the research community seems to be in agreement that there currently still is no means of deterministically comparing the suitability of different, similarly complex, entire gestural controllers for the purpose of musical performance, and therefore no deterministic way with which a proposed new instrument as a whole may be evaluated (Jorda, 2004), (Wanderley & Depalle, 2004), (Orio et al., 2001).

The NIME research community sees it as a crucial necessity to develop such methodology, to facilitate the successful adoption of new designs by performers. Wanderley and Depalle (2004) discuss that, although it is as of yet not possible to systematically evaluate gestural controllers, it is of great importance that means of doing so are developed. Only if this is achieved, can existing designs be directly compared against each other, and informed choices be made during the design of new, novel instruments. By continuing to analyse similar developments in better established fields such as HCI, it is possible that further insight is gained on how to achieve this goal also in the area of developing digital musical instruments. This course is what will hereon be followed.

2.6.2 Human Computer Interaction

Human computer interaction is by now a well-established discipline, with many universities offering education programmes on the subject, and with numerous research journals and conferences dedicated to it. It has its historical foundations in the 1970s work referred to as software psychology (Carroll, 1997), as well as the older disciplines of human factors and ergonomics. With HCI being such a wide subject area, covering the entire field and its history is beyond the purpose of this review, given the very significant amount of work this would require. Instead, only the aspects of HCI that are directly related to the present work will be selectively detailed.

From its inception, HCI has been highly interdisciplinary, involving psychology, computer science, and ergonomics, while more recently, additional disciplines such as sociology, anthropology, linguistics, and various design disciplines have also become important sources for research input (Sasse, 1996). The great difference between how research in these disciplines is conducted has resulted in much discussion over the years, regarding how to best establish a body of knowledge which practitioners from all disciplines can agree upon. A discussion that currently is still very much unsettled, and on-going.

2.6.2.1 Epistemological development and conflict in HCI

Historically, computers have been used predominantly in a work setting, and similarly the majority of early HCI research has concentrated on the use of computers in a work context. Many recent developments however have forced this view to be altered. Partly, use of computers in other contexts than work is now arguably at least as prominent, with computers playing a central role in many other activities, including play, artistic creation, and communication. As a natural consequence, in HCI these new contexts of computer use have begun to also be considered. This development has brought with it many new challenges however, since the pre-existing knowledge that HCI practitioners brought with them, proved to be very hard to adapt.

Even before, there was significant discussion among HCI practitioners about matters as fundamental as the very nature of the discipline, and how research should be carried out and communicated. As research in HCI and the disciplines preceding it has progressed (human factors, ergonomics, software psychology), researchers from new disciplines have become involved in the field, bringing with them their own epistemological approach which has often been incompatible with that of the dominant view in the discipline at that time. See (Jacko & Sears, 2003) for an in-depth review of the evolution of HCI and the paradigm shifts that have followed as practitioners from new disciplines have entered the field.

A clash between different methodologies in HCI, worth detailing to exemplify the nature of the methodological disagreements in the field, is the following 1980s discussion: It was at the time by many seen as necessary to approach HCI as a "hard" science, where research produced from fields that do not fit well into this description (e.g. Psychology or Anthropology), needs to be made to conform to the notion of knowledge that dominates in these disciplines, as suggested by Newell & Card (1985). They argued that psychological contributions would be more useful to HCI if their findings were turned into "hard" science, so that the HCI practitioner is given an unambiguous, formal method that is directly employable in the design process. This approach however has since been criticized as unrealistic by many HCI researchers. Monk & Wright (1991a, 1991b) argued that work in HCI conducted using traditional hypothesis-testing experiments has been proven to lack generality, and that the most promising direction to pursue is to instead approach HCI as a design science. Many more contributions have since been made to the debate, which is arguably, albeit to a lesser extent, still on-going. For a dated but detailed review and discussion on this particular disagreement, see (Sasse, 1996).

Another more recent, arguably similar controversy is the “damaged merchandise” debate that took place in the 1990s, where researchers approaching HCI from diverse disciplines, challenged the usefulness of evaluation methods contributed from disciplines that may have a significantly different methodological approach. The controversy, as many refer to it, started when Gray and Salzman, both cognitive scientists, published a paper (1998) where they questioned the usefulness of discount ethnography usability methods, in favour of instead using formal quantitative evaluation. The discussion sparked was extensive, a great number of researchers disagreeing with their stance. The resulting dialogue is detailed in (Kaye & Sengers, 2007), an alt.chi contribution that reviews the evolution of evaluation in HCI.

Epistemological developments in cognitive science have recently had further, arguably ground-breaking impact on the field of HCI. In stark contrast with the historically predominant view that cognition and knowledge only concern the mind, a new understanding of cognition has emerged over the past century, that of embodied cognition: whereas philosophy and cognitive sciences have traditionally concentrated on the study of abstract, symbolic knowledge, the realization was made that most real world human cognitive activity is in fact situated in an environment, and related to the interaction with and manipulation of external objects (Anderson, 2003). The subject of HCI however has recently been recognized to still be very much rooted in the Cartesian separation between cognition and action (Dourish, 2001), thus not taking into consideration in its practices this recent realization that human cognition is in fact highly embodied.

In traditional HCI, the use model of procedurally manipulating explicitly defined symbolic data has by far been predominant, where the user is expected to separate the task at hand into logical steps, further subdivided into subtasks, and then engage in a dialog with the computer, where each instruction has to be given in turn, followed by a waiting period to anticipate the computers reaction to a given instruction. This dialog is then continued until the computers reaction to the last instruction is satisfactory, and thus the goal has been reached, and the main task completed. Most human activity however, does not translate well into this model, even less so activity that is related to artistic performance. Artistic performance and much other human activity, is instead more in line with the concept of continuous manipulation, where the reaction to user input is instantaneous, as with most activities that involve physical manipulation of objects.

Even though computers now are powerful enough to support tasks that are performed in real-time, very often the user interface for carrying out these tasks is still constrained by the limitations inherent in their originally intended design for a procedural mode of interaction. Furthermore, established HCI research and evaluation methods are mostly incompatible with these new contexts of computer use. Much progress has recently been made, and there are now research areas within HCI which solely address this issue of creating user interfaces compatible with the embodied cognition model, either implicitly, or with the explicit intent clearly stated. Research has begun on *enactive interfaces* (https://www.enactivenetwork.org/), referring to human-computer interfaces that are specifically designed to take advantage of the users enactive knowledge.

It has even been questioned whether the HCI research conducted during the last decades of the millennium really has successfully produced much empirical, generalizable scientific results at all. Indicative of this, is a quote by William Buxton, one of the fields most prominent researchers, where he provocatively describes HCI as a “failed science” (Buxton et al., 2000); it is safe to assume the emphasis is on the word *science*, since he is still a prominent research contributor to the field.

Buxton stresses this is because HCI has yet not succeeded in contributing useful innovative inventions that improve the way in which we interact with computers in any significant way. Although frameworks and methodologies have been created in great numbers, these have failed to contribute to significant progress, and are only useful in assessing the usability of existing designs, as opposed to inventing new, useful ways of interacting with computers. Buxton argues that this is due to the insistence of the HCI community to approach HCI as an engineering discipline, when it would be more useful if it was realized that it should instead be approached as a design practice, similar to architecture, or industrial design.

Note that this quote is taken from the transcript of a round-table discussion Buxton participated in, on the development of new Digital Musical Instrument controllers, and not an academic HCI publication. He follows these statements up a few years later however, in a recent paper co-authored with Saul Greenberg (S. Greenberg & Buxton, 2008). There they discuss what they argue are inefficiencies in the field of

HCI, why incentive for innovative, novel inventions is often stifled, and what needs to be done to remedy the present situation. They emphasize a problematic tendency in the HCI community: researchers try to shape their research questions so that they are compatible to the established HCI evaluation methodology, instead of first formulating a research agenda, and then surveying what methodology is the most appropriate. They argue that: *"the choice of evaluation methodology – if any – must arise from and be appropriate for the actual problem or research question under consideration"*. Because of their great significance for present work, the arguments presented by Greenberg and Buxton will be further discussed in section 2.6.2.3.

Harrison et al (2007) review further criticism to how HCI research is often carried out, stating it is questionable whether it is at all possible to employ true scientific theory in social sciences, a field to which HCI has become more and more related. They raise the question of whether *"scientific criteria are the best and most apt for the field",* and also state that maybe no single methodology is alone a sufficient alternative, recognizing that the highly interdisciplinary nature of HCI necessitates the use of even a combination of methodologies, depending on the task at hand.

2.6.2.2 Resulting new paradigms of HCI

It has in HCI, as is exemplified by the preceding methodological argument, crystallized that entirely different methodologies may be necessary depending on the task at hand, and that no single methodological approach can be employed for all HCI research conducted. Since the turn of the millennium, what is named the third paradigm of HCI has emerged, addressing the phenomena that the previously dominant paradigms regarded as marginal, and had difficulty addressing (Harrison et al., 2007). The research areas of ubiquitous and pervasive computing, non-task-oriented computing work such as ambient interfaces, affective computing and experience-centred design, are all examples of areas that fit under the description for the third paradigm. The third paradigm is also well aligned with the discussion on situated and embodied cognition. This area of the third paradigm is by Harrison et al (2007) termed *situated perspectives*, treating interaction as "(…) *a form of meaning making in which the artefact and its context are mutually defining and subject to multiple interpretations"*.

As Harrison et al also state themselves it is not necessarily so that the only possible paradigm grouping in HCI is the one they propose. Different groupings that are just as valid are also possible, and alternatives have been proposed, such as the Reality-Based Interaction framework proposed by Jacob et al (2008).

The first paradigm, referring to the pioneering work that eventually led to the creation of HCI as a discipline, saw according to Harrison et al interaction as "*a form of man machine coupling",* as inspired by the disciplines of engineering and human factors. The second paradigm followed, which concentrated on treating "*mind and computer as coupled information processors".* This paradigm, has for the last two decades of the millennium dominated the HCI field, as it aligns well with the tendency exhibited during that period to regard "*human information processing as deeply analogous to computational signal processing",* in accordance with Newell & Card (A. Newell & Card, 1985)**.** This view began to show its inadequacies in many contexts early on, but has nonetheless been resilient, because it lends itself well to the desire within the research community to treat HCI as a "hard" science, as discussed in section 2.6.2.1. Although alternative suggestions were made, they were regarded with much scepticism, because they were based on methodologies contributed from disciplines such as psychology and anthropology, which aligned badly with the established methodology at the time, derived largely from the then dominant ideas in the cognitive science discipline. For an in-depth analysis on the emergence of the third paradigm, please refer to (Harrison et al., 2007).

2.6.2.3 Discussion on choosing appropriate research and evaluation methodology in HCI

Being that HCI is such an interdisciplinary field, it is only a natural consequence that there will be significant discussion on what validation methodologies should be used, in the same manner that we have seen is being discussed regarding the definition of what HCI research constitutes in the first place. In "Usability evaluation considered harmful (some of the time)" (2008), Saul Greenberg and Bill Buxton discuss how HCI usability evaluation methodology is too strongly encouraged in the research community, resulting in its use in contexts where it may not be applicable, and where use of a different evaluation methodology would have been more appropriate. This viewpoint is to some extent also echoed in the discussions we have already seen, albeit with a more limited scope than the entirety of the HCI discipline, by Harrison et al (2007), Zimmerman et al (2007), Paul Dourish (2006), and Fallman (2003). The argument presented is particularly relevant to the research project presented in this thesis, and as such a summary of the article's argument is detailed, accompanied with comments of how the article's arguments are relevant to present work.

HCI usability evaluation methodology ranges from laboratory-based user observations, to controlled user studies, and/or inspection techniques. It is within HCI research employed to validate proposed novel ideas and systems, by showing that these facilitate improvements compared to a base set of metrics, or by showing that users can achieve predefined quantifiable goals.

Greenberg and Buxton argue that although usability evaluation has clearly proven its worth in a multitude of situations, it is not by default always the appropriate choice of methodology: *"the choice of evaluation methodology - if any - must arise from and be appropriate for the actual problem or research question under consideration".* Although they stress that usability evaluation is indeed of core importance in HCI practice, its application should not be done blindly, as this then may result in the research producing meaningless or trivial results, misdirecting, or even stalling future design directions.

By observing how in the research community, the discussion and research on evaluation methodology has dwindled compared to past decades, Greenberg and Buxton argue that the community has settled on having a "methodological bias", resulting in researchers often formulating their research questions so that these are compatible with the established methodology. Researchers may oftentimes resort to choosing a method they perceive is favoured by review committees, and then finding a problem to match the chosen method. In education, usability evaluation has been made a core part of the curriculum without teaching alternative methodology to even nearly the same extent. In addition it has become the de-facto evaluation methodology standard for submitted papers to the HCI community's primary conferences and journals.

In other design disciplines, such as architecture, the argument over whether to use objective or subjective evaluation methods has already been considered to a great extent, and in these the value of subjective arguments from experts has come to be considered just as legitimate a contribution as results derived from more objective methods. In HCI however this is not the case, resulting in a stifling of design and engineering innovations because of an excessive focus solely on scientific contribution. Greenberg and Buxton argue that *"depending on the discipline and the research question being asked, subjective methods may be just as appropriate as objective ones"*. They reference a discussion by Snodgrass and Coyne on design evaluation in architecture, which is replicated here to better illustrate the view on evaluation held in design disciplines:

> *[Design evaluation] is not haphazard because the assessor has acquired a tacit understanding of design value and how it is assessed, a complex set of tacit norms, processes, criteria and procedural rules, forming part of a practical know-how. From the time of their first 'crit', design students are absorbing design values and learning how the assessment process works; by the time they graduate, this learning has become tacit understanding, something that every practitioner implicitly understands more or less well. An absence of defined criteria and procedural rules does not, therefore, give free rein to merely individual responses, since these have already been structured within the framework of what is taken as significant and valid by the design community. An absence of objectivity does not result in uncontrolled license, since the assessor is conforming to unspoken rules that, more or less unconsciously, constrain interpretation and evaluation. If not so constrained, the assessor would not be a member of the hermeneutical community, and would therefore have no authority to act as an assessor. (p.123)*

The argument presented this far resonates particularly strongly with the work detailed in this thesis. The contribution of present work cannot be assessed only by addressing a clearly defined question through conducting controlled experiments. The work does to a significant extent, also contribute outcome that is arrived at following design research methodology, for which evaluation methodology such as ethnographic method, and design critique, is arguably far more appropriate than usability evaluation. These methodologies are therefore addressed in more detail in sections 2.6.2.4 through 2.6.2.7.

Continuing the argument of Greenberg and Buxton, often usability evaluation is erroneously applied too early in a design process. When evaluating a sketch-prototype of a novel new interface, it will not have gone through the many iterations of refinement that the established interface it is tested against has benefited from. By comparing the performance a user can achieve using the two, the innovative idea is likely to fare worse, but that needn't mean that it is a bad idea; only that the wrong kind of questions have been asked when evaluating it, through an inappropriate choice of evaluation method. Because usability evaluation does not in fact test for the novel ideas usefulness; it tests only its usability. To quote the authors: *"Usability evaluation is predisposed to the world changing by gradual evolution; iterative refinement will produce more useable systems, but not radically new ones".*

Perhaps more importantly, in the case of developing an innovative user interface, usability engineering fails to take into account the culture of use of designed system, as the system is necessarily examined without considering how the system will evolve and be adopted by a culture over time. In other words, usability engineering focuses on examining the microscopic impact of a systems deployment, when in the case of new inventions, what would be relevant to examine would instead be its macroscopic effects.

Greenberg and Buxton proceed with suggesting a number of initiatives that the HCI community can take to remedy the situation: that the community will need to recognize that usability evaluation is just one out of a toolbox of many methods for performing user-centred design, and that it should always be considered whether it is the appropriate method to apply for the situation at hand, over existing alternatives. Additionally usability evaluation may not be appropriate at all stages of a projects design cycle. If usability evaluation is applied too early in the lifetime of a design it may prove to have a number of negative consequences, and alternative approaches may be more appropriate, while usability evaluation can still be useful at a much later stage.

Because the system developed here clearly is a novel invention rather than an iterative refinement of pre-existing technology, following Greenberg's and Buxton's argument, there is little reason to perform usability evaluation at this early stage of its development.

Greenberg and Buxton conclude that both academics and practitioners HCI community need to recognize that other ways of validation may be just as meaningful, including the detailing of a design rationale, detailing expected scenarios of use, case studies, and practicing participatory critique, as a few examples that may be equally valid depending on the context. Even when usability evaluation is applied, it needs to be accompanied with carefully thought out scientific rigor, while the publishing of studies replicating previous results should be encouraged, as this is very rare in HCI compared to many other disciplines. Finally methods from other disciplines for evaluating design worthiness have to always also be considered. Methods that evaluate cultural aspects of designs need to also be better understood, as has begun within some HCI communities such as CSCW and Ubiquitous Computing (UBICOMP), where ethnographic approaches are an invaluable tool for understanding how technologies are embedded in social groups, and within physical environments.

2.6.2.4 HCI as Design Research

Many HCI researchers have in addition to Greenberg and Buxton begun to reason that HCI is perhaps most usefully approached as a design science. Fallman (2007) notes how an additional influence to HCI from other disciplines escalated from the second half of the 1990s. In particular, social science disciplines such as sociology and anthropology came to gain methodological prominence, helping to address the difficulties stemming from only relying on the epistemology of cognitive science. This was then followed by the additional significant influence on the field from traditional design disciplines (Fallman, 2003) (Wolf, Rode, Sussman, & Kellogg, 2006) (Zimmerman et al., 2007) (Buxton, 2007).

Fallman stresses that a significant difference between design oriented research in HCI and research in natural sciences is that the developed artefacts are necessarily deployed and evaluated in a real world context, and not in an abstract lab setting. In the real world context, people will use the artefacts in uncontrollable, unintended ways, as influenced by the experiences, preconceptions, cultural and societal values and beliefs that the subjects unavoidably will bring with them. Design-oriented HCI research is thus necessarily more of a social sciences discipline, drawing from work in ethnography, phenomenology and sociology, rather than natural sciences.

Regardless of the epistemological approach of any particular HCI research endeavour, design is now recognized as an important part of the process. In contemporary HCI research, one of the fundamental activities is to design and implement new technologies, often in the form of a prototype, through which ideas for novel systems can take concrete shape (Fallman, 2003).

Design in the HCI research process can take on two often conflicting roles: a) design being used to address a research question, with the objective of pursuing an abstract truth and the production of knowledge, and b) being used to create an artefact that addresses requirements for its deployment in a real world setting, with the primary goal being the development of a marketable product. These two situations are by Fallman referred to as design-oriented research, and research oriented-design, respectively (Fallman, 2003). A designer is always involved in both these two conducts regardless of whether he is working in research or a commercial context, so realizing this distinction is important, for the designer to successfully balance the compromises necessary when he chooses to emphasise one conduct over the other. Zimmerman et al (2007) similarly define design research as the design work carried out with the "*intention to produce knowledge and not the work to more immediately inform the development of a commercial product".*

Zimmerman et al (2007) present the result of a comprehensive two-year research project to review as well as invent the necessary methods for interaction design research in the context of HCI research. With this aim, they review the academic publications on the subject, as well as interview leading researchers and practitioners. They propose a model for how research should be conducted, as well as the necessary criteria for evaluating the research outcome.

Unlike design practice, where the work is oriented towards producing a successful product, design researchers create artefacts with the intention of treating them as "*carefully crafted questions"*. Zimmerman et al mention the widely referenced example of the Drift Table (Gaver et al., 2004), as a successful research project where this methodology is employed, while the project also additionally raises "*the issue of the community's possibly too narrow focus on successful completion of tasks as a core metric of evaluation and product success".* Another useful description of design research is that of Nelson and Stolterman, as referenced in (Zimmerman et al., 2007): *"research on a condition that arises from a number of phenomena in combination, rather than the study of a single phenomenon in isolation".*

In the model proposed by Zimmerman et al, the design researchers focus on making the *right* thing: artefacts intended to transform the world from the current state to a preferred state. This is better described by directly referring to the words of the authors:

> *"Using our model, interaction design researchers integrate the true knowledge (the models and theories from the behavioural scientist) with the how knowledge (the technical opportunities demonstrated by engineers). Design researchers ground their explorations in real knowledge produced by anthropologists and by design researchers performing the upfront research for a design project. Through an active process of ideating, iterating, and critiquing potential solutions, design researchers continually reframe the problem as they attempt to make the right thing. The final output of this activity is a concrete problem framing and articulation of the preferred state, and a series of artifacts—models, prototypes, products, and documentation of the design process".*

Zimmerman et al state that in design research, the problems being examined are often *Wicked*. To define a wicked problem, they quote the originators of the term, Ritter and Weber's (1973): *"a problem that because of the conflicting perspectives of the stakeholders cannot be accurately modelled and cannot be addressed using the reductionist approaches of science and engineering"* (Zimmerman et al., 2007).

In conclusion, the contribution of design research to the HCI community is the identification of opportunities for new technology or for advancement of current technology. Artefacts are created that provide concrete embodiments of theory and technical opportunities (Zimmerman et al., 2007). To differentiate the artefacts produced in the context of design research from those produced through design practice, two distinctions are made. First, that the intent of the work is not to produce a commercially viable product, thus taking into account all the parameters that constrain such a design, but to produce *knowledge* for the research and practice communities. Second, the artefact needs to demonstrate significant invention: "*The contributions should be novel integrations of theory, technology, user need, and context; not just refinements of products that already exist in the research literature or commercial markets. The contribution must demonstrate a significant advance through the integration".*

2.6.2.5 Evaluation of design research

In their review, Zimmerman et al (2007) discuss what they have found are appropriate evaluation criteria for design research. They note that much contributed research has been using the design research approach, however no agreed upon standard has been used of what it exactly means to do design through research, or how best to evaluate such a research contribution.

To help formalize the research method, they suggest four criteria or lenses as they also describe them, for evaluating an interaction design research contribution:

Process: While there is no expectation that the process described by the researcher could ever deterministically produce the same results if applied again, it is important that it is described in enough detail to be reproducible. A rationale for why specific methods have been selected needs to also be given, so that the work can then be judged based on this, and on the rigor with which the work has been carried out.

Invention: Interaction design research contributors need to be able to demonstrate that they have produced a *"novel integration of various subject matters to address a specific situation",* that constitutes a significant invention. This is best done by providing an extensive literature review to situate the work, and to detail the aspects that show how the contribution constitutes an advancement of the current state of the art. This needs to also be accompanied with a technical description sufficient to advise engineers on what to build to capitalize on the described research contribution.

Relevance: Given the nature of design research, it is not reasonable to expect that the results produced can be identically attainable by reproducing the development process. This is however often a requirement and evaluation criterion for engineering and behavioural science research, to demonstrate the validity of proposed work. Zimmerman et al instead propose the benchmark of relevance, as derived from anthropology, where the research focus is on "what is real", as opposed to "what is true". In addition, a preferred state for the researchers design attempts needs to be communicated and supported.

Extensibility: The design research work needs to allow the community to build on the resulting outcomes. It also has to have been documented in a way where it is clear how the community can make use of the knowledge derived from it.

2.6.2.6 Ethnography in HCI

In recent years ethnography has gained widespread acceptance in HCI practice as a valuable methodology for observing complicated interactions, where other methodologies have failed to capture the great level of complexity that inevitably arises if a system is put to use by real users in a real-world setting. It can with great usefulness be employed during all development stages of a design project, from capturing user requirements for a system to be designed, through the development process, to recent discussion on its use for evaluating the completed system.

In the "Human Computer-Interaction Handbook" (Jacko & Sears, 2003), a recent and thorough review of the use of ethnographic methods in HCI is presented. A summary from the books' section on ethnographic method is presented here. The use of ethnography in HCI can be traced back to the 1980s, when computers came to be deployed outside of research laboratories into mainstream work settings, thus creating a need for a way to gain an understanding of the everyday reality of people working within these miscellaneous settings. In the late 1990s, computers became just as usual in people's homes as they were at their workplaces, further increasing the recognition of the ethnographic perspective's relevance. However, use of ethnography within HCI has primarily been geared towards the gathering of requirements, before the initiation of the design and development process. Its employment for the evaluation of research outcome is a more recent development. Employing ethnography also in an evaluation context is now nonetheless briefly discussed even in some undergraduate HCI textbooks (Dix et al., 2004), (Preece, Rogers, & Sharp, 2002).

Ethnography has its historical roots in anthropology, but is now an approach used in most traditional and applied social sciences, and interdisciplinary fields such as HCI. It is based on the principle that *"to gain an understanding of a world you know little about you must encounter it firsthand"*. Ethnographic studies rely predominantly on gathering information in the setting in which the activities under observation normally occur, a principle resulting from the recognition that people removed from the social and material characteristics of the environments in which they interact, have a very limited ability to recall what they do and how they do it. Our ability to fully describe what we do is limited due to the tacit nature of the principles that guide our actions (as stated by Polanyi, referenced in (Jacko & Sears, 2003)).

Ethnographic research is further based on the view that activities must be understood within the larger context in which they take place. This is referred to as "holism", meaning that *"studying an activity in isolation, without reference to the other activities with which it is connected in time and space, provides only a limited and potentially misleading understanding of that activity"*. Ethnographic accounts are to the extent that it is possible descriptive, primarily intended to originate an understanding of the events and activities as they occur, while avoiding their evaluation. Although ethnographic studies are often performed to suggest improvements, the conviction held is that conclusions regarding a situation can be drawn only after it has been fully understood as it is.

A primary focus of ethnographers is to *"gain an insider's view of a situation"*, so that the account given is described in the terminology and from the perspective of the people studied, as opposed to those of the research community towards which the study is directed. This is of course performed to the extent that it is possible, since every ethnographic account is by its nature unavoidably located and partial, dependent on the time it was performed and coloured by the perspective of the researcher, among many other factors. The value of an ethnographic study is nonetheless not thought to be in any way reduced by this, as it is still of significant usefulness, provided its unavoidably partial and situated nature is always accounted for. It is worth noting that studies have been performed which show all scientific knowledge production is unavoidably shaped by the larger social context in which the scientific inquiry takes place (Jacko & Sears, 2003), so this is not a characteristic only of ethnographic research.

As has been seen, many research directions within HCI are incompatible with the traditionally established evaluation methodologies. In these directions, the use of ethnographic method has recently become more common, as it has a unique usefulness in allowing the researchers to gain an understanding of how a developed system is being used while its users are situated in the real context the system was intended for.

An in depth discussion on how ethnographic research has been employed in HCI research work can be found in (Dourish, 2001), where the development of a new air-traffic control system (Hughes, O'Brien, Rodden, Rouncefield, & Sommerville, 1995) and the optimization of a print-shop (J. Bowers, Button, & Sharrock, 1995) are discussed at depth. Another more recent, often referenced account can be found in the paper by Gaver et al, regarding the development of the Drift-Table (2004).

2.6.2.7 Ethnographic method

Performing ethnographic research does not involve simply applying prescribed methods selected from a toolbox of techniques. Ethnographic method is instead more a way of looking at a problem, a *"theoretically informed practice"* (Comaroff & Comaroff, 1992). Following is a summary of the methods employed as reviewed in (Jacko & Sears, 2003).

An important start for performing ethnographic research is to formulate a plan of action, describing the research objectives, a strategy for selecting study participants and selecting appropriate research techniques. As a project progresses the strategy formulation may then be subject to change, to adjust for new knowledge gained regarding the investigated area.

The participant *"sampling strategy"* addresses what types of participants, and how many of them are needed for the study at hand. If subjects are carefully selected, it has been shown that it is possible to achieve reliable results with as few as four or five participants, depending on the study, thus disproving popular belief that a reliable study inevitably needs to involve hundreds of subjects.

The observation of subjects in the scene under study is one of the primary methods using which ethnographical data are gathered. An observer may be integrated in this scene to various degrees, his role varying between being a quiet observer, to that of an active participant in the events observed. The choice needs to be made about what, where and when to observe (B. Whiting & J. Whiting, 1970): the observation can either be person focused, event focused, place focused, or object focused. An increasingly important aid in ethnographic fieldwork is the video camera, as it greatly simplifies the task of documenting the events in the scene to produce field notes, even giving the opportunity of repeatedly watching the same events while observing them with a different focus every time.

An equally critical method for ethnographic research is the use of interviewing. It is useful because and despite of the fact there is often a great difference between what people say they do and what they have in fact been observed doing. Use of interviews is imperative, to gain an understanding of the subjects perspective. Ethnographic interviews are often open-ended, especially in the beginning of the study, where this strategy allows the interviewer to learn what questions are important to ask as the interview progresses. It is important to avoid questions that limit the range of possible answers the subjects can give, as the purpose of the interview is to allow the subject to describe his experience with his own words, and not simply to confirm or deny the preconceived hypotheses of the ethnographer. As more knowledge is gained during the course of a research project about what topics are to be pursued, the interviews can progress to becoming more structured. Finally, at the end of a study, a further tool for validating the findings is to use the derived knowledge to structure the interview, so as to confirm that the way this knowledge shapes the question structure and language indeed reflects those of the participants.

Findings from observations and interviews are then combined, to obtain the ethnographic, holistic view, while the combination of observation and interviewing may even be performed simultaneously, if the context permits doing so, a practice referred to as in-situ interviewing. If the study area makes it difficult for the ethnographer to always be present, it is sometimes also useful to employ self-reporting techniques, such as instructing subjects to keep diaries, visual storybooks containing annotated photographs, or more recently also internet based, multimedia blogs.

2.6.3 Review of evaluation methodology employed in recent PhD theses analogous to present work

Besides academic articles, several PhD theses deemed analogous to this research were reviewed to gauge what research and evaluation methodologies were employed. Of the theses reviewed, the following six were examined in depth, because they were found to varying extents of having a goal of designing and implementing a novel instrument for artistic musical/audiovisual performance: (Moody, 2009), (Mark T. Marshall, 2008), (Aimi, 2007), (Collins, 2006), (Grierson, 2005) and (Jorda, 2005). Note that all these theses detail some form of technical evaluation to varying extents, and so this will not be individually detailed in the text below.

Niall Moody, in “Ashitaka: an audiovisual instrument” (Moody, 2009) details the development of his Ashitaka live audiovisual performance instrument. He presents an extended theoretical argumentation for his rationale and reasoning behind the instruments development, but does not perform any trials with audiences or musicians whatsoever. Instead he offers a brief evaluation based on his own experience with using the instrument, and applies an observational evaluation metric that he has developed himself in the thesis.

Mark T. Marshall, in “Physical interface design for digital musical instruments” (Mark T. Marshall, 2008), does not concentrate primarily on developing a single instrument, but instead on developing a broader theory for the making of instruments. As part of his work however, he does also implement the ideas he develops, in a new instrument design. This instrument he evaluates through conducting trials with musicians: a total of five undergraduate music students, each partaking in a forty minute trial session.

Roberto M. Aimi, in “Hybrid percussion: extending physical instruments using sampled acoustics” (Aimi, 2007), develops a novel percussive musical instrument. To evaluate it, he conducts a total of four trials, each lasting a few hours, and each with a different established percussionist. From these sessions, an account of the observed and stated experiences of the percussionists is detailed.

Nick Collins, in “Towards autonomous agents for live computer music: realtime machine listening and interactive music systems” (Collins, 2006), develops several musical and audiovisual systems, intended to be used in live performance. The systems developed are not involved in any trials with musicians, but they are used by Collins in live performances. He states that the fact the systems were useable in live performance, is alone testament that they are at a minimal level useable for the goal intended. Evaluation of the experience of the musicians using the system is listed for future work.

Mick Grierson, in “Audiovisual Composition” (Grierson, 2005), concentrates on developing a theory of a novel art form, referred to as audiovisual composition. A significant part of his thesis however is also the development of several applications, embodying his theory, of which some if not all are intended to be used in a live performance context. Grierson doesn’t however perform any trials involving either members of audience or performers, to evaluate the systems he has developed. Instead, he details the use of these by himself.

Finally, Sergi Jorda, in “Digital lutherie: crafting musical computers for new music’s performance and improvisation” (Jorda, 2005), too concentrates on the thesis’ theoretical contribution, but does also implement several new systems intended for live performance. None of these however are evaluated through involving audiences or performers, instead he applies observational evaluation metrics which he has developed himself in his thesis.

From the review one can immediately observe that the evaluation methodology applied varies greatly. In the cases where trials are at all performed, these are of a comparatively short duration, of a single session lasting a maximum of a few hours. Often the research and evaluation methodology used appears to be predominantly analogous to that of design research, covered earlier in this thesis (2.6.2.4). Note however that nowhere in any of the theses was this or any other methodological choice explicitly stated; instead this conclusion is drawn solely from observing each thesis' description of how the research was carried out.

2.7 Recent related work

Besides the historical predecessors of the Trinity system, the colour organs, there are several recent efforts that have had more immediately analogous goals, and thus warrant a discussion on their design, and what differentiates them from present work.

All identified systems are found to have some combination of the shortcomings also detailed in section 1, to varying extents:

(i) The mappings between music and visuals are highly constrained because they are limited in complexity, and remain static over time, thus necessitating that the correlation between visuals and music always remains limited, when in fact there is much evidence that increased correlation results in a stronger experience, as will be reviewed later in this chapter.

(ii) Virtually no user interface exists, that allows controlling the performance of visual music/audiovisual art in real-time, with a level of expressivity comparable in nature or extent to that attainable in live musical performance.

(iii) The process of preparing or improvising live procedural visual music/audiovisual performances is overly complicated, compared to live musical performance, as it almost exclusively requires that artists engage in software engineering.

(iv) Collaborative performance, a very common practice in most live performance art, is mostly absent in live visual music/audiovisual performance, largely not due to artistic choice, but to technical limitations.

The first implemented modern system to bear a technical resemblance to what has been created here was Gordon Pask's musicolour, as described in section 2.1.2.

Today a vast multitude of software exists that allows creating real-time computer graphics, controllable from musical input. However, virtually all existing implementations are controlled simply by performing beat and amplitude detection on the stereo mixdown of the music, using this information to drive a pre-defined moving graphic. Systems that go further regarding the level of control are very few.

More recently, Ox and Britton created the 21st Century Virtual Colour Organ (2000), a real-time virtual-reality colour organ controlled using MIDI notation as input. Both the visual content and the mappings were hardcoded in the system. There is no mention of the system being used in a context of real-time performance; rather it was controller from pre-recorded MIDI-data.

Taylor et al (2004), detail an implementation where a virtual character, realized using real-time computer graphics, is controlled through a live musical performance. The system is implemented to be controlled from MIDI-data, as well as data derived from audio signal processing, but the mapping and the visual content are both again hardcoded.

Both Jack Ox's and Robyn Taylors' systems are clearly best described as interactive art-pieces, embodying the aesthetic preferences of their creators, neither being at all intended to be useable by other artists. Neither the visual content, nor the mappings can be varied at all.

Mick Grierson, (2005), created several real-time audiovisual software applications / pieces, with the purpose of detailing the *"Technical, theoretical and aesthetic concerns relating to the composition and performance of real-time audiovisual material with particular reference to the connectivity between audio and visual events".* Grierson makes a significant theoretical contribution, by presenting an extensive review and analysis on what he describes as an audiovisual metadiscipline: audiovisual composition, "*An artistic form which takes as its starting point the cognitive actuality of multisensory audiovisual experience*". Grierson does refer to his software applications as instruments, and does intend for them to be used by other artists than him. From observing his detailing of their technical implementation however, each of these instruments is found to have some combination of the limitations of previous practice referred to in the introduction to this section. There is in Grierson's thesis no discussion on how the audiovisual instruments he proposes could be controlled using some alternative to the control paradigms of established practice, such as has been proposed in this thesis. Although some of his applications are intended to be used by other artists, the ones that synthesize procedural computer graphics have their graphics generating algorithms hardcoded. An artist that wishes to adapt these to his own aesthetic preferences will thus need to engage in software engineering practice, and modify these applications. Moreover, Grierson does not discuss collaborative audiovisual performance at all. Finally, although Grierson discusses introducing a narrative in the mapping between music and visuals, he does not detail any concrete design solution employing which such a narrative may be controlled in a performance. This objective is in present thesis addressed with the concept of mutable mapping.

Moody (2009), with the Ashitaka audiovisual instrument, presents a real-time computer graphics system, controlled by musical gestures. Ashitaka is the first relevant system for which it is specifically stated that musical gestures are the source of control for both the visual and the musical output. The instrument used however is an entirely new design created specifically to be part of Ashitaka. As such, it is impossible to deterministically gauge whether it is a design which allows its players to take advanced of their advanced motor capabilities, or if it limits them in that respect. A precondition to expressivity is control complexity however, which can be easily examined. The interface offers six sensors each of which generates a single range of discreet numerical values, and an accelerometer which provides an additional set of three. As such, the interface does afford a control complexity comparable to that of a traditional musical instrument. However, whether this control complexity successfully translates to actual expressivity, can only be determined after players have spent a very significant amount of time practicing on the instrument; 5000 hours of practice are said to be sufficient for achieving a good level of familiarity -but not virtuosity- with a musical instrument (Woody, 2004). As is clear from reading Moody's thesis however, no performer has extensively practiced using his instrument. Furthermore, while in Ashitaka the mappings are hardcoded, the system provides some functionality for varying the visual output, through providing basic support for importing 3d geometry data in the X3D file format.

Although now deprecated, and very dated technologically, the system that is most similar to Trinity in its conception is '*Bliss paint*' (1998). It allows a great level of customization on the presented animation, mapping MIDI to the control parameters, and the option for the user to select one out of 18 different possible mappings. Its greatest limitation however is that only the colours in an image can be animated using it. Its rendering engine is based on the principle of manipulating the colour palette of an imported still image, and thus no animation of shapes is at all possible.

Finally, the possibility of controlling the performance of real-time visuals using musical instruments as the interface has been mentioned before (Gerhard, Hepting, &

McKague, 2004), even in the case of having a group of performers do so (Arfib, Couturier, & Kessous, 2005). However both have only been a brief discussion of the possibility and the presented advantages, without following this up in more detail, even less so with actual research or an implementation.

In this chapter we have established the preceding work that this effort builds upon, from the distant history up to today's cutting edge live audiovisual performance practice, along with what methodology is best suitable for carrying out and for validating the work presented in this thesis. We have seen the history of the artistic practices that relates visual narratives to the narratives of actual music, as well as defined what meaning is given to the term narrative in the context of this thesis. We have examined the history, technologies and practices relating to live performance using digital musical instruments, so as to illustrate the close relation between the technologies used for performing actual and visual music, and frame the argument for why the choice was made in this work to employ existing musical gesture controllers, as the primary source of control data for the live performance. We have reviewed the practice of programming as a means of artistic expression, to explain the rationale behind wanting to enable artists to create their own procedural graphics for the Trinity system. We have reviewed the arguments in the research community for what research and validation methodology is best used for this work, so that the rationale for the subsequent methodological choice made can be firmly grounded: the choice of employing design research methodology, with the added viewpoint of ethnographic method during validation. With the final detailing of current practice's cutting edge, the groundwork is therefore laid for presenting in detail the Trinity system and the concepts behind its design, followed by its subsequent deployment in actual performance, and finally its evaluation, in the subsequent chapters of this thesis.

3 Mutable Mapping

Although the concept of mutable mapping was initially conceived of for it to be used as part of the Trinity system, it quickly developed to an idea that generalizes to all digitally mediated artistic performance / live new media art. Because of this wide-reaching generality, it is worth first discussing the concept of mutable mapping on its own, in separation from Soma and the Trinity system, before discussing how it fits also into those contexts.

As was previously detailed in the review on digital musical instruments, in modern musical performance, the devices serving as interfaces for the performers (controllers) are separated from the devices producing the actual output (sound generators) (Miranda & Wanderley, 2006). The controllers encode the musical gestures of the performers as discreet digital data for the sound generators to interpret, producing the final sound. This is in contrast to traditional musical instruments, where the device used to detect the musical gestures (e.g. the piano keyboard) cannot be separated from the sound generator (the pianos hammers and strings).

This separation between control and output devices can today be witnessed in many more forms of artistic performance. Stage lighting rigs are remotely controlled using purpose-made control devices, which transmit data following the DMX protocol. Visual performance artists such as VJ's, use off the shelf or custom made controllers to generate MIDI and OSC data, which is to be interpreted by VJ software and other real-time computer graphics applications. In modern dance, it is increasingly common that dancers have sensors attached to their bodies, which transmit data to be interpreted by real-time musical and/or visual devices, thus allowing the dancers' movements to influence the music played, as well as the imagery projected on the stage. Finally, in the new media arts movement, some or all of the above mentioned conducts are fused, resulting in interactive multimedia performances and installations, oftentimes allowing also the audiences to be directly involved in the piece/performance.

The fact that different data protocols are employed for different contexts of use has purely historical reasons, as the data transmitted is of a very similar nature in all cases. Due to this very high degree of similarity, today modern systems converge towards using what has emerged as a potential future universal standard: Open Sound Control

(OSC) (Matthew, 1997). Despite it having originally been designed as a protocol for musical control data, OSC has in addition proven to be equally suitable for all above application areas. Existing legacy devices can also be addressed using OSC data, since OSC commands can be translated to and from most legacy protocols with relative ease. In addition, modern developments in signal processing have made it possible to derive digital control data even from analogue sources, thus allowing for example that discreet musical notation data is derived from audio of an acoustic live performance (Miranda & Wanderley, 2006).

Most recently, controllers initially created for one context of use, are often successfully repurposed for another, while a large body of research has been directed towards creating new, novel controllers. Additionally universal controller interfaces are created, agnostic in their design with regards to what context they are to be used in. In all cases, the output of these universal controllers no longer has a default mapping to the parameters of the devices they control. Instead, when preparing for a performance, artists define their own mappings, based on their personal preferences, artistic and pragmatic.

The fact that today all controllers and controlled devices used in the above mentioned artistic performance contexts, may with very little effort be made to communicate, paves the way for a new form of artistic performance: that of gradually creating, destroying and altering mappings between the two parameter spaces of input and output control data.

This novel conduct is inspired by a development which occurred when music began to be electrically amplified and recorded: the emergence of the mixing engineer role, as is detailed in section 2.2.6. Following this development, the established roles of originating the music are now accompanied by live alteration of the signal, with the end result being a joint creation of both instrument players and live mixing engineer, all indispensable members of the group. Today in many cases this development has gone full circle, with mixing engineers taking centre stage as the main artist, while the instrument players – if these have not been entirely replaced by recordings – hold only a secondary role. A contemporary expanded form of the above performance mode is at present embodied in the widely popular Ableton Live software, previously discussed in section 2.2.7. In live, the traditional mixing engineer toolset is augmented with facilities for detailed live triggering of multiple short loops of pre-recorded music. Today many electronic music composers use live in order to perform their music as solo artists.

Analogous to the mixing engineer, the performer responsible for mutable mapping fulfils a similar role, but instead of manipulating the audio signal produced by the sound generators, he/she manipulates the digital control data produced by the gestural controllers.

3.1 Related work to mutable mapping

Although there is extensive past and on-going discussion regarding various mapping strategies, primarily with relation to developing new DMI's, no previous publication has been found on gradually altering mappings as a form of performance. The only work found which is somewhat related to the ideas presented here is that of Malloch et Al. (2007). They have developed the application OSCMapper, intended to be used as an aid in the process of designing new DMI's, through allowing the quick prototyping and alteration of mapping setups. Their software application has functionality analogous to that of the Mediator application, with the significant difference however that it does not allow the gradual altering of mappings, only switching them on/off and introducing discreet, non-animated changes, thus rendering it unsuitable for use in a mutable mapping performance.

3.2 The mutable mapping performer

In summary, mutable mapping can be defined as artistic performance through gradually creating, destroying and altering mappings between two parameter spaces of generated control data from the artist(s) performance gestures and received control data of the synthesis devices, during the course of a performance. Mutable mapping allows for novel forms of performance, both within only a single artistic conduct (e.g. music), as well as performances where multiple art forms - for example music, visual art and dance - are fused.

The role of a mutable mapping performer can vary greatly depending on the specific context. Moreover the role is bound to evolve significantly once a culture of use is established, as it adapts to the particular idiosyncrasies of each context. Arguably, new not yet envisioned forms of use are bound to appear. Following is a non-exhaustive list of examples detailing possible use cases.

In musical performance, the control data originating from DMI controllers is subsequently routed through the Mediator software. The performer then alters the control data and selectively redirects it to the control parameters of one or many sound generators on the receiving end. Examples of what possible alterations may be introduced:

- Arpeggiators/automated accompaniment may be gradually applied to the notes from an individual instrument
- Control data may be gradually rerouted between different sound generators
- Procedural modifications may be applied, so that for example note length or pitch are altered

Combinations of the above allow for complex "layering" techniques, through combining the sound from multiple sound generators into one intricate evolving instrument resulting from their fusion.

It is also possible that the mapping engineer acts as a solo performer, in which case the control data is either pre-recorded - possibly performed using Ableton Live, or generated algorithmically.

Beyond solely musical performance, in a context where live music is accompanied by visual art, for example VJing, real-time computer graphics or light shows, the parameter space addressable by the mutable mapping performer grows significantly. Control data originating from DMI's can now also be mapped to the control parameters governing real-time generated computer graphics, VJ software, or the lighting rig.

The novel advantages stemming from such interconnection are significant: it facilitates creating a narrative in the visual output that is closely correlated to that of the music, because both music and visual performances are primarily controlled using the same control data. The ability to gradually re-route control data in this context is imperative, enabling performers to construct a non-linear, dynamic and free-form narrative in the correlation between music and visual accompaniment.

In all contexts of use, either a single performer can work alone, or multiple may work collaboratively, each responsible for his/her own group of source and/or destination parameter spaces. Conversely a solo audiovisual performer can map pre-recorded or algorithmically generated control data both to audio synthesizers and to visual synthesizers and/or VJ software.

Note also that a solo performer need not be restricted to the task of controlling mutable mapping. He/she can also simultaneously perform as an audio mixing engineer, as a VJ, as a lighting artist or even as a musician using conducting gesture controllers (Miranda & Wanderley, 2006), or software such as Ableton Live.

Like the role of live audio mixing engineer, the role of mutable mapping performer is not singularly defined, and varies greatly depending on context. He/she may be an indispensible member of the group, or just a technician, depending on the level of initiative and intervention he/she has in augmenting the performance of the musicians. The mutable mapping performer may furthermore be partly or wholly a content creator for the performance, such as when performing using pre-recorded material of his/her own creation as a source of control data, or when control data initially generated by musicians concentrating on its' musical use, is used to control live visuals.

Having now examined the concept of mutable mapping in isolation, we have seen how it is envisioned to be useable in a wide range of digitally mediated artistic practice. In the following chapter, we will see how it is for this work implemented in the Mediator software, as well as see the other two software applications together with which the Mediator forms the Trinity system, using which Soma performance is made possible.

4 The Trinity system

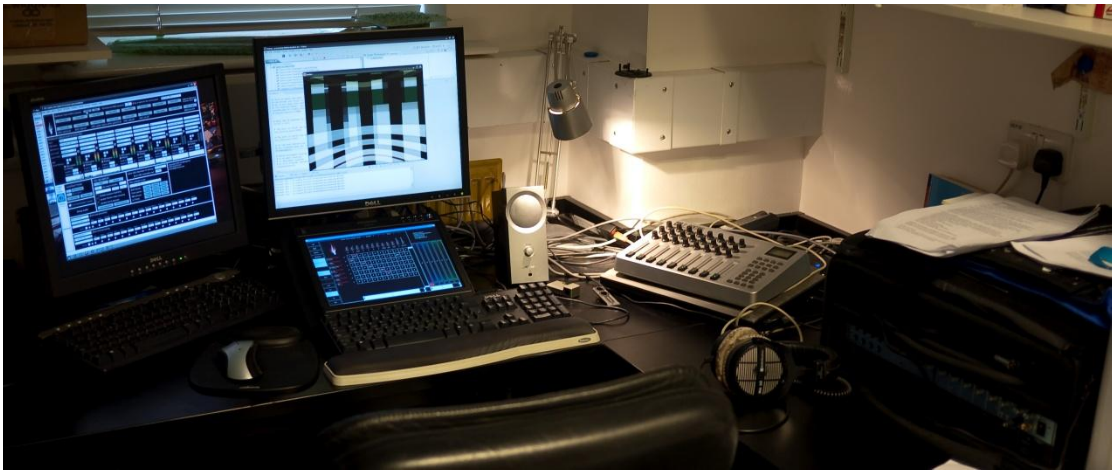


Figure 21: The Trinity system hardware

The Trinity system consists of three software applications: The *Live Input Processor*, responsible for gathering and processing audio and discreet control data from instruments, Mother, responsible for hosting visual synthesizers, and finally Mediator, which provides the functionality and user interface for controlling the mutable mapping.

4.1 Mother

Mother integrates the principles of using VJ-software, and of working with multimedia programming environments such as Max/MSP and Processing.

Although the principle of mixing multiple layers of moving graphics, as in VJing, is retained, here these graphics are not from pre-rendered video clips, but are the output of real-time visual synthesizers ('synths'), running in parallel within the main host application. Each synth is a program that renders a particular visual effect, the control parameters of which are all continuously accessible live during a performance, so that the appearance of the visual is animated over time. 'Synths' are created as Processing 'Sketches' (Reas & Fry, 2007), using a Processing library provided (named *Foetus*) that enables the sketch to work within the Mother host application. Users can either create their own synths, use example synths from a collection provided with the host, or from other users. To see a simple example of a Processing sketch converted into a visual synth which can be used within Mother, please refer to Appendix E. The steps performed in the example presented, are the only ones necessary, irrespective of the size of the sketch.

The system therefore comprises of two modules: the main host application (Mother) and the Processing library (foetus). Mother manages the execution of the multiple synths, mixes their individual visual output into the final output, and handles the forwarding of received control information to individual synths. The Foetus library resides within each synth, and handles its integration with the host. Because it encapsulates the boilerplate functionality for allowing this integration, it also provides an easy mechanism for programming new synths.

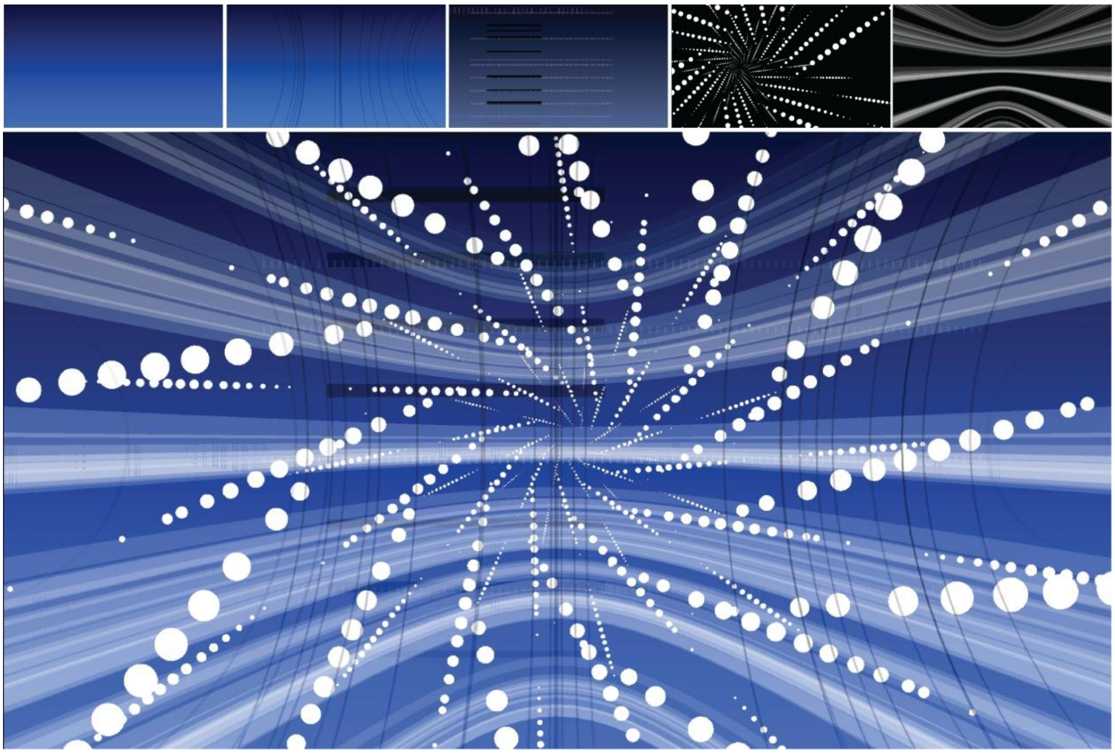


Figure 22: Illustration of how several visual synths are layered to produce a single complex output.

Synths can be of two categories: (i) full-frame synths, covering the whole screen area, thus being most suitable for use as the background layer, (ii) synths partially covering the screen, leaving a transparent background. To view a non-exhaustive list of the synths created for this work, please refer to Appendix D.

The communication of control data to the host and to individual synths is managed solely using the OSC, protocol. Currently there are many systems that allow one to relatively easily define a user interface capable of transmitting OSC control data, both in the form of hardware controllers and of software applications. A comprehensive, up to date list is maintained at http://www.opensoundcontrol.org.

The choice to completely separate the synth host from any user interface was made because there currently is no established generally applicable user interface paradigm for controlling the performance of real-time computer graphics. Users of Mother are instead encouraged to adapt or create their own user interface, depending on what best suits their particular context of use. In this manner, Mother can be used in many more contexts than the one initially intended for it, of being part of the Trinity system.

4.2 Live Input Processor

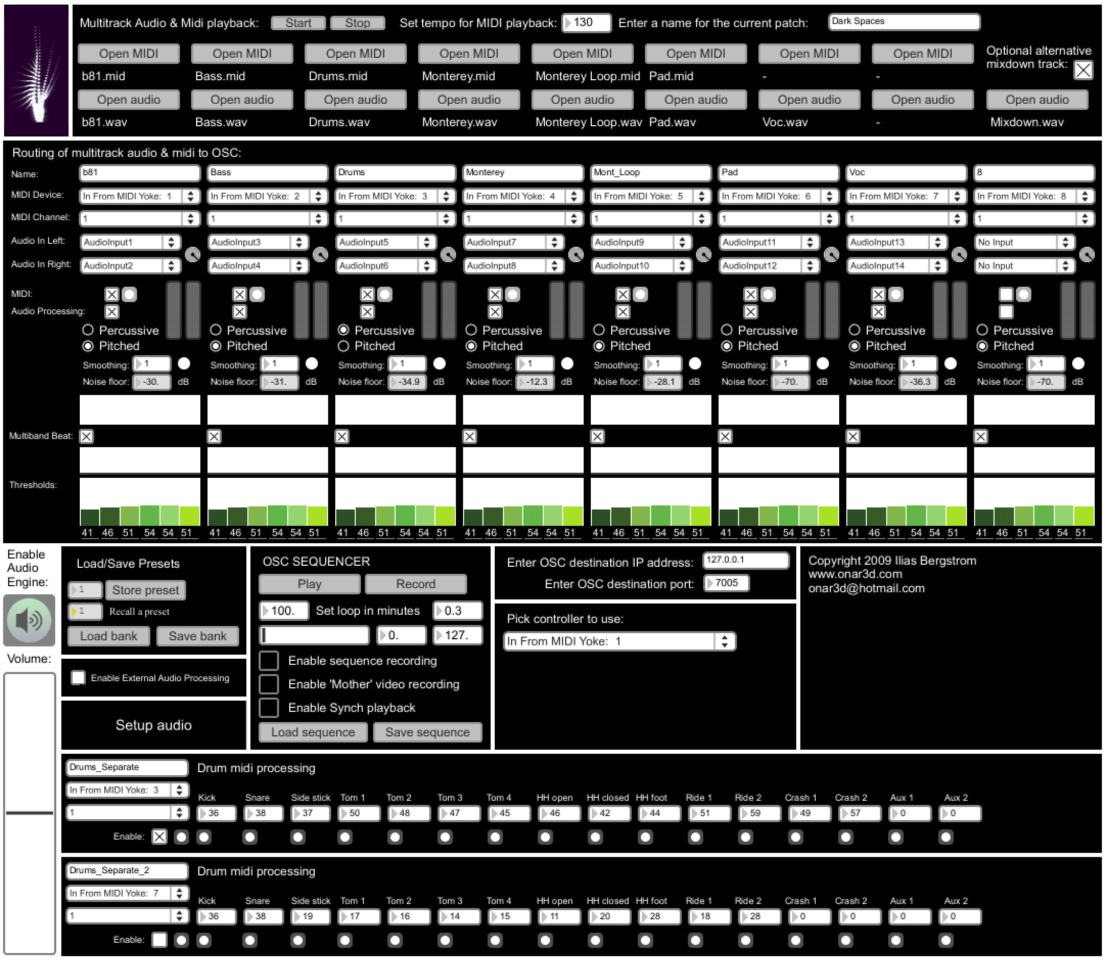


Figure 23: the Live Input Processing software

The Live Input Processor (LIP) application is responsible for the handling of control data from musical instruments. The application was developed in the Max/MSP visual programming language. It is used for the following tasks:

- Receiving incoming MIDI data from musical instruments, translating it into OSC, and retransmitting it in real-time.
- Performing low-latency pitch tracking, amplitude tracking, and beat detection on incoming audio signals, and transmitting the derived data over OSC in real-

time.

- Performing high-pass filtering and/or smoothing on incoming audio, before it is processed for generating control data. This is necessary to clean up artefacts in noisy live signals, so that the audio's musical content can be tracked more accurately.
- Performing low-latency beat detection separately over multiple frequency bands of incoming audio signals, and transmitting the derived data over OSC in real-time.
- Receiving incoming MIDI data from non-musical controllers (such as the Evolution UC33), translating it into OSC, and retransmitting it in real-time.
- Playing back pre-recorded multi-channel audio along with corresponding MIDI data which too can be processed in the same manner as data being received live.
- Recording and playing back OSC data generated from the above processes at varying speeds. Note however that due to limitations encountered in the Max/MSP programming environment, which could not be circumvented, the recording process is in the present version of the software not possible to perform in real-time.

Global settings can be saved and recalled, as can the recorded OSC data sequences. The current implementation supports a total of eight pairs of audio and MIDI channels, and thus a maximum of 8 musical instruments. This is not due to any technical software limitation, as the application would with ease be able to support many more. The hardware audio interface used for the project however has only eight channels of simultaneous audio input, and thus there was no immediate benefit in implementing functionality in the software to support more than eight.

4.3 Mediator

To support the mapping engineer role, a user interface paradigm has been developed, which has been implemented in a prototype application dubbed the *Mediator*. The only control data protocol supported is OSC, selected because it is universally compatible with all domains where the Mediator would be useful. OSC is also with little effort compatible with most legacy protocols, as their vast majority may be translated to it.

The Mediator’s ideal implementation is envisioned as a standalone hardware controller with a multi-touch screen, the device for which would not be unlike the JazzMutant Lemur controller (http://www.jazzmutant.com/), Figure 24. The current incarnation however is as a PC software application controlled using a single pointing device, presently the touch screen of a tablet notebook computer. This was necessary because at the time the software was implemented, there was no commercially available compact multi-touch device, except for the lemur, which was unsuitable because it cannot be used to run user-created software.

Figure 24: The JazzMutant Lemur ten finger multi-touch controller (www.jazzmutant.com).

The central feature of the Mediator user interface is the *Mapping Matrix*. Each of its rows corresponds to the source of a single control parameter, while each of its columns corresponds to a single destination. Each row or column has a data type assigned to it: floating point, integer or text value. At the intersection of each row and column appears a cell, thus forming a matrix. Rows and columns can be dynamically added and removed at runtime, and their assigned OSC address can be altered.

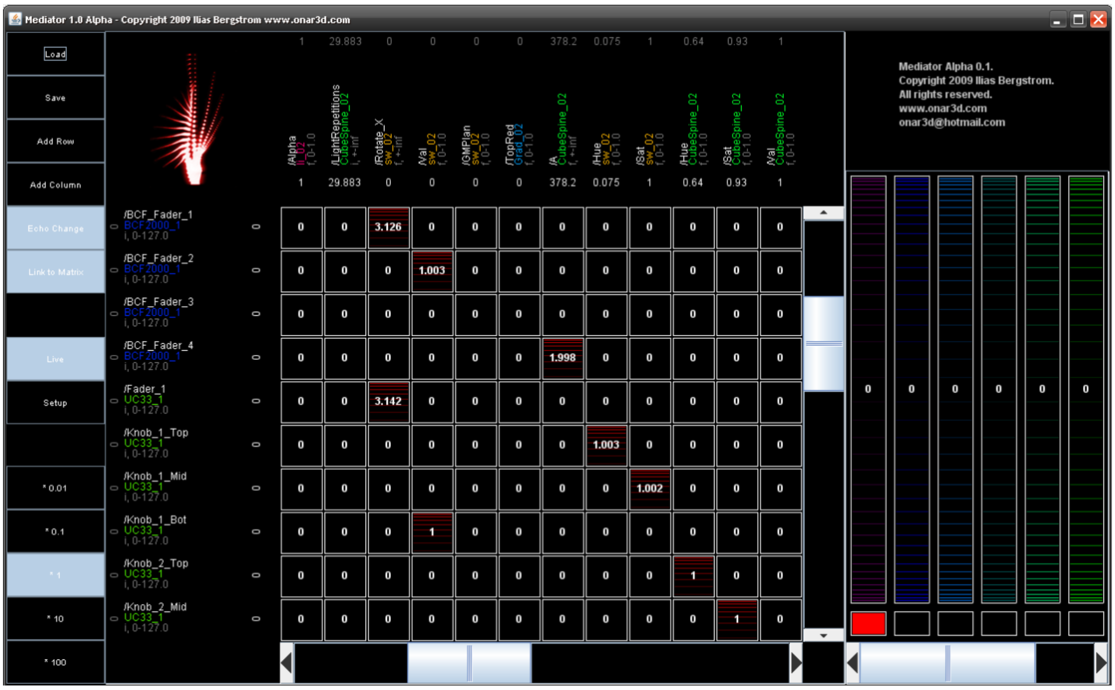


Figure 25: The Mediator user interface: main screen

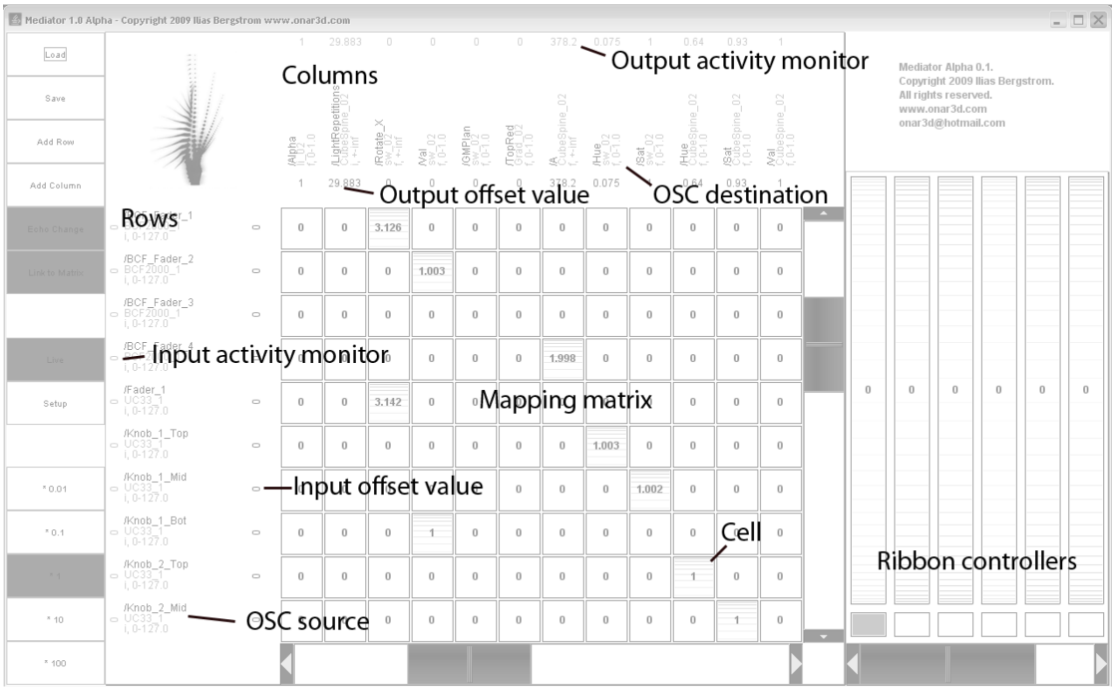


Figure 26: Main screen legend

Figure 27: Input row & cell detail

The predominant data type handled is floating point and integer numerical values. For these, unlike a patch panel, in which the cell connecting source and destination may only be on or off, the mapping matrix is novel in that it is weighted. Each cell holds a floating point value, which is used as a factor with which incoming numerical values are multiplied before they are forwarded to their destination. A value of 0 means incoming values are not forwarded at all. Changes in connections may thus be introduced gradually, by increasing or decreasing the cells' multiplier values, allowing for much more detailed performance than just the on/off states possible using a normal patch panel. Cells' values are altered through dragging, and thus on a multi-touch screen the paradigm allows for the values of multiple cells to be altered simultaneously.

When a cell's corresponding row and column both are assigned to handle text values, the cell changes behaviour to normal patch panel-style on/off forwarding of the unaltered data. Finally if the source row's type is numerical and the destination column's type is text or vice-versa, the values are incompatible and thus their corresponding cell is greyed out, supporting no form of forwarding at all.

In addition to the mapping matrix, numerical control values can be further altered using input and output offset values: a number that is added to the incoming value prior to multiplication with the cell values, or added to outgoing values after multiplication with cell values. These offset values are set individually for each row or column, by dragging the numerical field appearing to the right/bottom of the row/column identifier.

Finally, numerical cells can be assigned to groups. Those which appear with a white outline and red centre are not assigned to any group, while those which have the outline and centre coloured with the same colour are assigned to that colour's corresponding group. A numerical cell is assigned to a group through tapping/clicking it.

Each group is represented by one out of eight ribbon controllers located to the right of the mapping matrix. By dragging the ribbon controller up or down, the user can increase/decrease the multiplier value of all cells assigned to the group simultaneously, by the amount corresponding to how far the ribbon controller is dragged. Multiple ribbon controllers can be manipulated simultaneously, thus allowing the concurrent modification of a large number of cell multiplier values.

Under each ribbon controller appears a button, which is coloured red if it is the currently active controller. When a cell is tapped, it will be assigned to that ribbon controller. With no ribbon controller active, tapping a cell clears its group assignment.

A second screen, referred to as the control screen, is also available in the Mediator software. In this interface, two different tasks can be undertaken.

First, OSC entities can be added, removed and reordered. An OSC entity is an instance of any type of sender and/or receiver of OSC control data. After an entity has been created, its input and output parameters can be assigned to columns and rows respectively, in the mapping matrix. When the entity list is manipulated, the changes are echoed by the Mediator software as OSC messages. Any collaborating software can then listen to these messages, and thus adapt to the changes that have taken place in the Mediator's internal data model. At present, this is how visual synthesizers are added/removed/reordered in the Mother software.

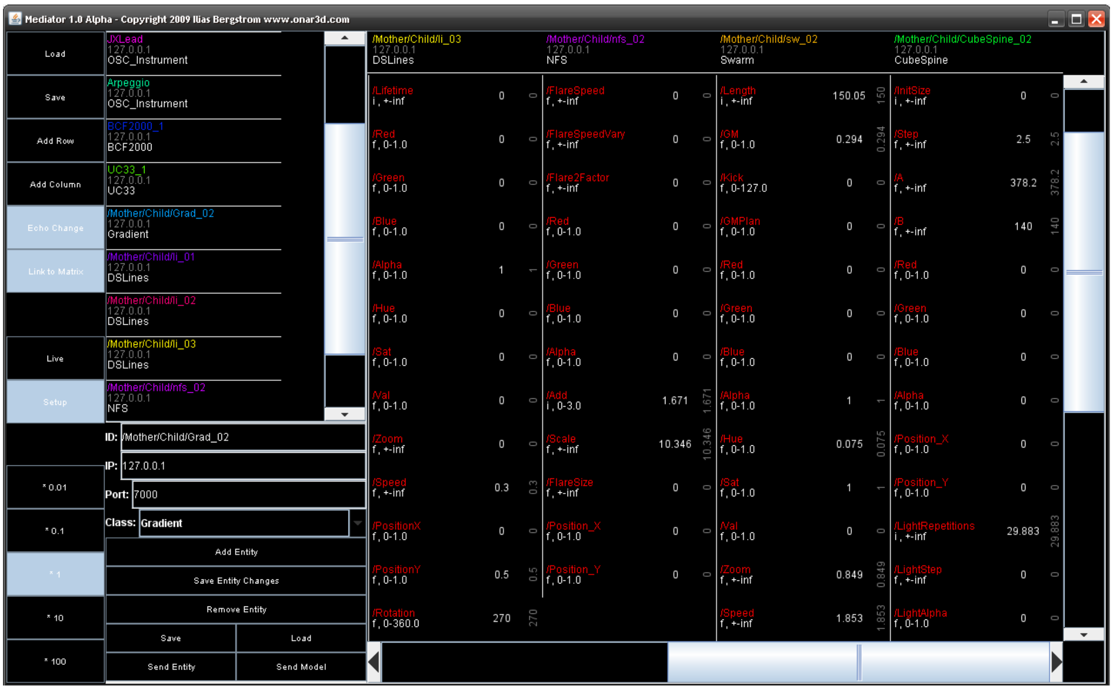


Figure 28: The Mediator user interface: control screen

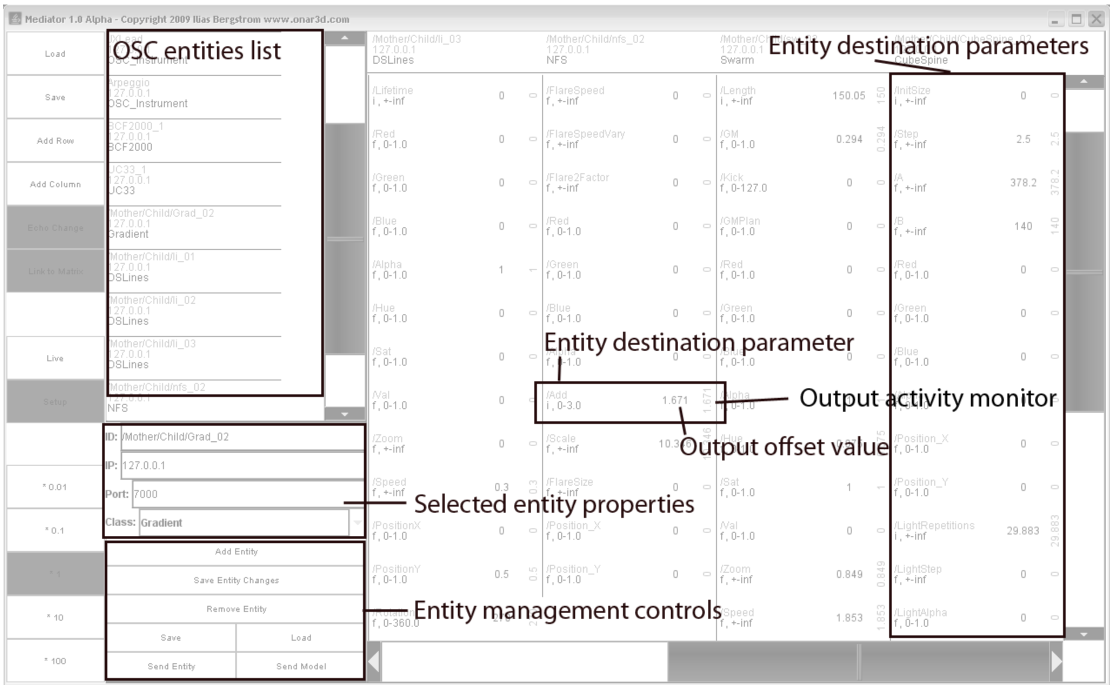


Figure 29: Control screen legend

Secondly, an alternative interface is presented for sending values to the input parameters of OSC entities. This, in the present context of use, would be the control parameters of visual synthesizers. The output offset values from the previously described main screen, are here presented detached from the mapping matrix, so that they may be manipulated regardless of whether they correspond to a column in the mapping matrix or not. A performer can use this screen, to manually control the visual synthesizer parameters that are not mapped to any control input, as well as to alter the offset values of those visual synthesizers that are. Two modes of operation are available: one where the incoming values from the mapping matrix are acknowledged, and one in which the mapping matrix is ignored, in effect rendering the application a dynamic equivalent of a non-musical midi controller, such as the JazzMutant Lemur, or the evolution UC-33.

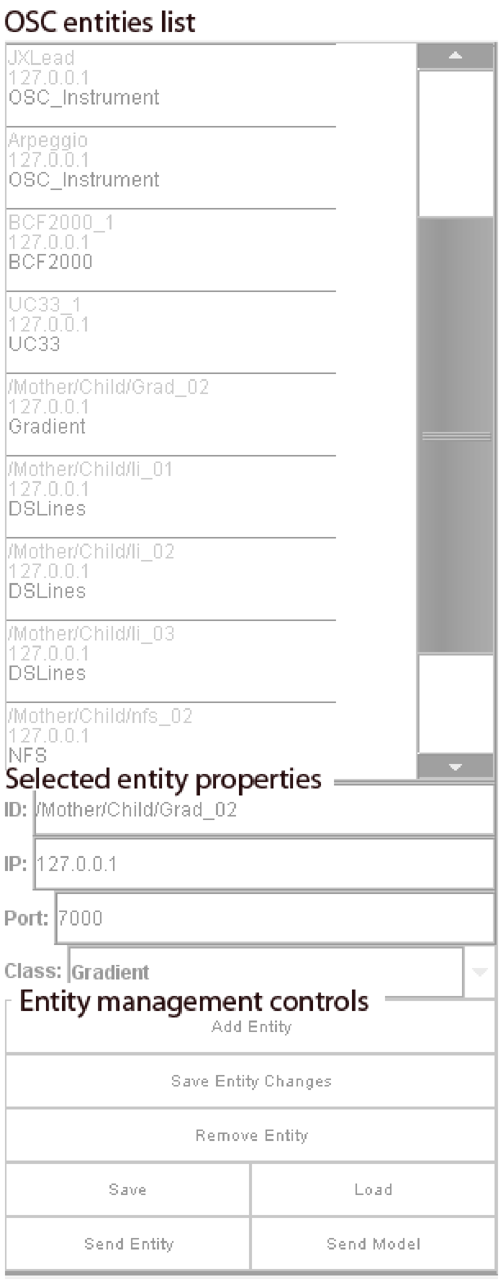


Figure 30: Entity management features

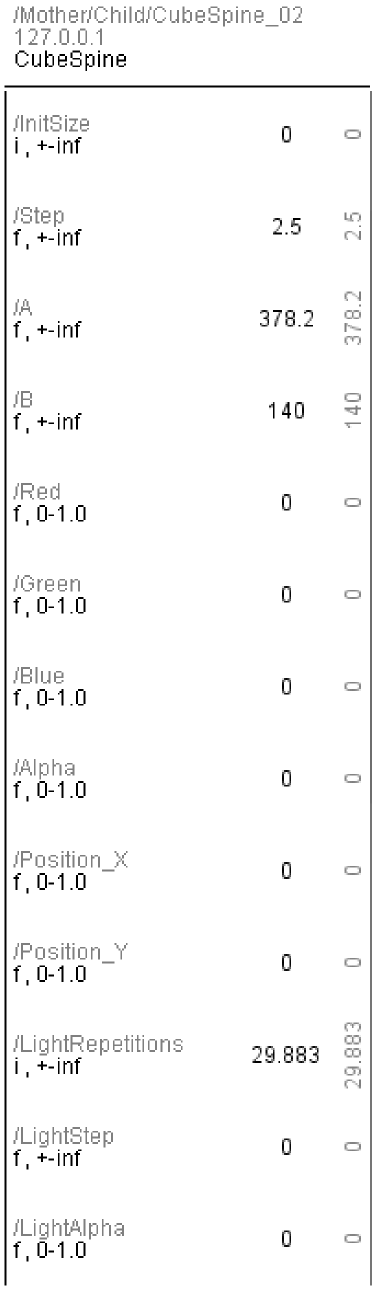


Figure 31: Entity destination parameters

The Mediator's mapping matrix supports the creation of a majority of possible desirable mappings, through allowing a multiplier and two offsets, one before and one after multiplication. Having said that, it is a simple exercise to describe a mapping that the current implementation does not support; for example, any mapping that requires a non-linear value transformation, such as spline interpolation, is currently not feasible.

4.4 Hardware

The Trinity system also needs an extensive array of hardware devices, in conjunction to the three software applications detailed, in order to function. In this section, the devices used for the present incarnation of the system are listed. Of course a similar system can be set up using a multitude of different devices of the types detailed here, but their functionality would still have to be analogous.

The incoming live audio signals are first connected to the ART S8 splitter, which echoes the signal to two outputs. One output goes to the audio PA, providing the amplified sound for the performance, while the other is sent to the Trinity system, through the Focusrite Saffire 10 I/O Pro firewire audio interface. As the signals processed by the splitter are likely to be line-level, it is imperative that the audio interface has a microphone preamplifier for each separate audio input it handles, as does the 10 I/O Pro. The LIP software application then interfaces to the 10 I/O Pro through low latency ASIO drivers, gaining access to 8 separate low-latency live audio channels.

Figure 32: ART S8 signal splitter (www.artproaudio.com).

Figure 33: Focusrite Saffire 10 I/O Pro Firewire audio interface (www.focusrite.com).

MIDI data from the musical instruments is input to the LIP software using the MIDI input of the 10 I/O Pro, and the M-Audio Midisport 4x4 MIDI interface, giving a total of 5 MIDI inputs. In the cases where an acoustic drum set is used, MIDI data is derived from its drums (but not cymbals), using the Roland TMC-6 Trigger-to-MIDI interface, together with a snare drum trigger (RT10S), a kick drum trigger (RT10K), and three tom triggers (RT10T).

Figure 34: M-Audio Midisport 4x4 USB MIDI interface (www.m-audio.com).

Figure 35: Roland TMC-6 Trigger-to-MIDI interface, and RT10S, RT10T and RT10K drum triggers (www.roland.com)

Figure 36: Evolution UC-33 USB MIDI controller (www.m-audio.com)

The Evolution UC-33 MIDI controller is used to facilitate assigning physical knob and slider controls to additional parameters of the visual synthesizers. The MIDI control data generated from the UC-33 is again routed to the visual synthesizers using the Mediator mapping matrix.

Figure 37: Dell Latitude XT tablet PC, and M4400 notebook workstation (www.dell.com).

The Dell Precision M4400 notebook workstation is used to execute the LIP and Mother applications. All audio and MIDI devices above are connected to the M4400. It was chosen because it offers a relatively high performance OpenGL graphics accelerator, and a capable CPU, both features necessary for the tasks assigned to it. The projector used for visual output is also connected to the M4400.

The Dell Latitude XT tablet PC is connected to the M4400 through a crossover Ethernet cable, and is used to run the Mediator software.

The setup time for the system is quite extensive, roughly one hour, depending on the instrumentation of the band involved. This proved to be a problem during live performances, as the time is often higher than the setup time of most musicians.

4.5 Supported modes of performance

To fully exploit all advantages of proposed system, it needs to be used by a group of musicians, each performing with a musical instrument, accompanied by a visual mixing engineer. The system however retains many of its benefits also when used in other modes of performance, such as using conducting DMIs as opposed to instrumental, and even to some extent when using traditional DJ-style performance.

When using conducting DMIs, although the advantages that instrumental DMIs provide are not available, the same amount of control data may still be produced, and thus the system can be used unmodified, in exactly the same manner. The difference will of course be that the control over the visual content will be less immediate, to the same extent that the control over the musical content is less immediate in this mode of performance, compared to the direct manipulation alternative. This is of very significant importance, as it allows the use of the system in this increasingly popular context.

To a limited extent, the system is also useable in the context of accompanying a DJs performance. If one considers the most limited mode of performance, that of the traditional DJ setup involving two record players and a mixer, one still has four audio channels to work with, a stereo pair from each record player. If stereo recordings are used, it is possible - but not implemented here - to separate the vocal from the recording (Xiang & Dubnov, 2005), thus having a separate mono channel for the vocal, and stereo for the music, from each record player. If different frequency bands in the music are in addition separated, these too can be used as separate sources of control information. So although in a much more limited manner, it is still to some extent possible to draw benefit from using proposed system when accompanying a DJs performance, compared to current practice.

The fact that the system is compatible with all of these different modes of performance is of great importance, because it is clear that it will with benefit scale from simple to more complex performance situations: from one-man performances, where the same person is responsible for both visuals and music, to performances involving a small group, up to performances encompassing a full orchestra.

In this chapter we established in detail, what each of the Trinity system's three comprising applications is made to achieve, how each is technically designed to do so, and how all three are intended to work together in different modes of performance. The Live Input Processor, responsible for gathering and processing audio and discreet control data from instruments, Mother, responsible for hosting and displaying the mixed output of several visual synthesizers, and finally Mediator, which provides the functionality and user interface for mutable mapping performance. In the subsequent chapter, we will follow the development process from which the here presented designs were arrived at.

5 Process

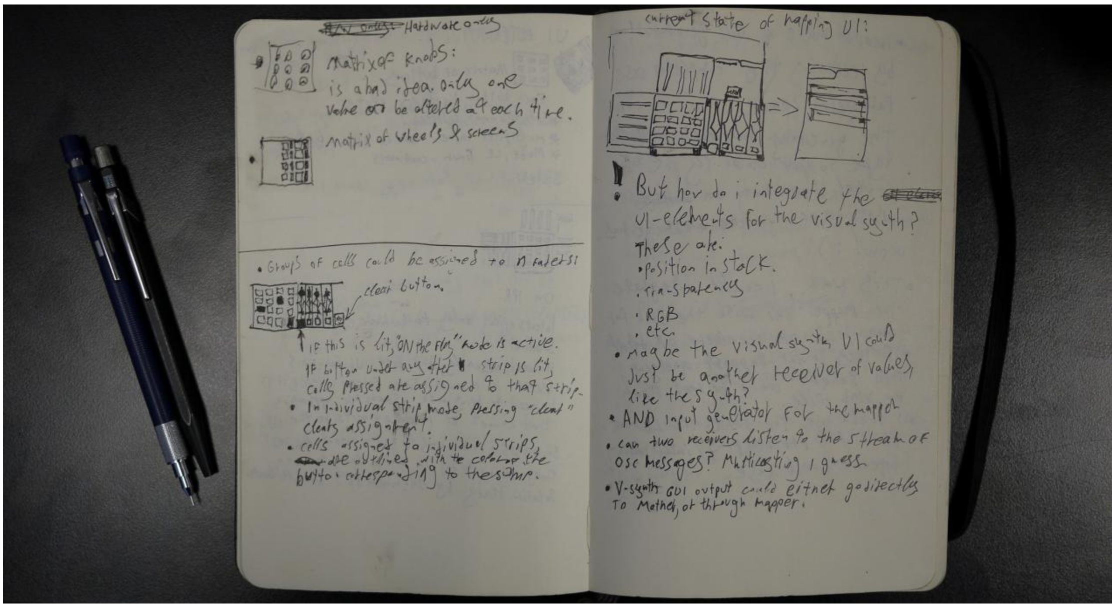


Figure 38: Notebook

In accordance with the evaluation criteria for design research discussed by Zimmerman et al (2007), the development process followed for the project is detailed here.

Throughout the process of developing the software, extensive notes were taken, both in a digital logbook, and in a paper notebook. During periods of brainstorming for new ideas, the paper notebook was found to indeed be as indispensable a tool for advancing a design, as is advised in literature on design practice (Preece et al., 2002).

5.1 First prototyping stage

The first software implementation made was a test-bed to gauge the extent of the advantage in having musical notation control visuals, instead of the previously established practice of using beat and amplitude detected from a stereo mixdown. This was conducted towards this author's MSc thesis project (Bergstrom, 2003). The hypothesis at the time was an analogous premise to that which was later theoretically grounded in far greater depth by Grierson in his PhD research (2005). It also helped in the identification of what technical difficulties may appear in the implementation of such an application. It was implemented in C++ and OpenGL, receiving MIDI data from a standard music sequencer application. No attempt was made to make it useable by other artists than the implementer, rather it embodied a single artistic piece; not a tool using which other artists can express their own aesthetic ideas. Nor was any mapping editing functionality implemented at this stage, working instead with a mapping fixed in the software's implementation.

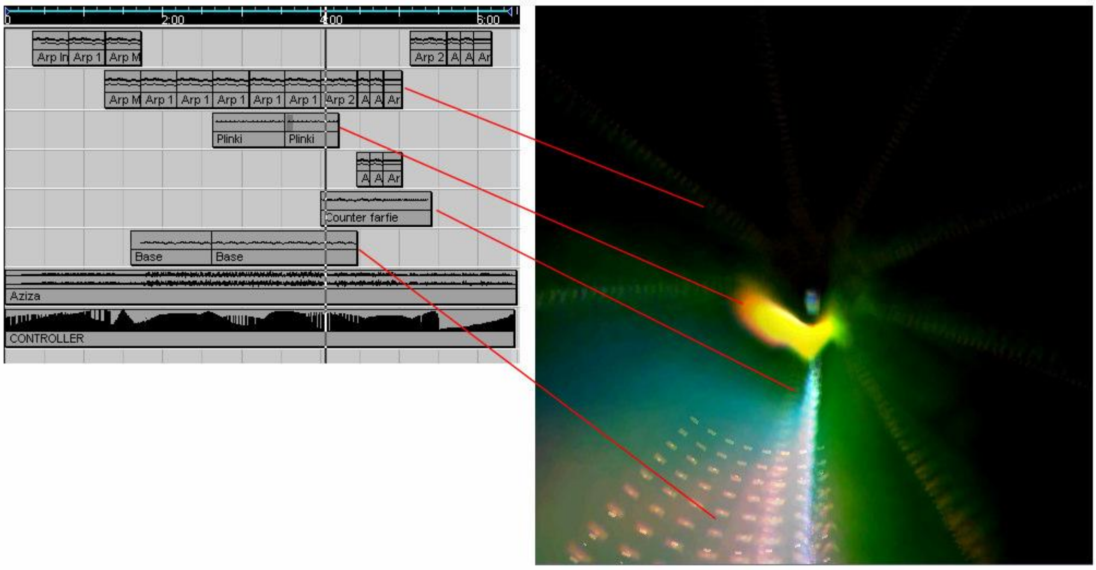


Figure 39: The different MIDI parts of the music, and their corresponding visual elements.

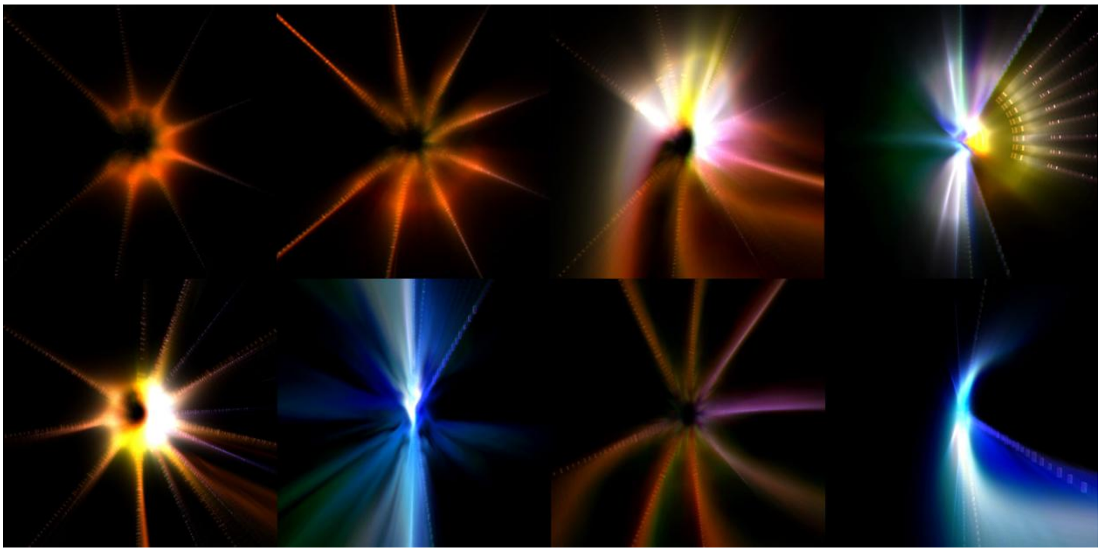


Figure 40: Sequence of still images taken from an animation produced using the prototype

This first prototype was used in a small scale study, comparing it to the then established practice, represented by the Winamp AVS visualization software (Bergstrom, 2003). 10 subjects were involved in total. The subjects were presented with the same music twice, once accompanied with the prototype application, and once with the Winamp3 AVS software, and were subsequently asked to complete a questionnaire. The subjects were informed that they would be asked to compare the two applications, and given no further information besides the questionnaire, shown to them after the screening had taken place.

The study was far too informal and of too small a scale to produce any generalizable results. Nonetheless, within these constraints, it did fulfil its purpose of resulting in an indication that non-expert audiences do notice and point out the conceptual difference between the prototype software and the previously established practice example. A majority of subjects (7/10) indicated that the prototype system did produce visuals that subjectively corresponded to a greater extent to the music. When asked to score the extent that the visual accompaniment enhanced the musical experience on a scale from 0 to 10, the prototype was given an average score of 8.7, while Winamp AVS got 7.1. From the open ended question replies, the study also indicated that the prototype software was able to keep subjects interested for a longer duration of time than the previous practice example.

5.2 Second prototyping stage

The second stage of prototyping involved attempting to gauge to what extent it was feasible to create software that is modular, so that it may be used by artists in pursuing their own aesthetic goals. Furthermore, it involved experimentation with varying mappings between musical control data and visuals, to gauge whether it was possible to define what subset of mapping strategies would be most suitable for the purpose intended.

First, several multimedia programming environments were evaluated to gauge their suitability for the project; these primarily include Cycling 74's Max/MSP/Jitter, Processing, Pure Data, VVVV and Derivative Inc.'s Touch Designer, although many more were also tested less extensively.

Several open source and/or free game engines were evaluated for their suitability to serve as visual synthesizers, with Unity3D and the Blender engine being primarily concentrated on.

Following this process, the program code from the first prototype was discarded in its entirety. Instead it was replaced with varying pieces of software written in the multimedia programming environments Max/MSP/Jitter, and Processing.

Max/MSP/Jitter was found to be a very capable programming environment, highly optimized for implementing signal processing software for audio and control data. It was thus chosen to serve for the prototyping of the input processing software, as well as for prototyping the mapping functionality.

Visual programming languages were deemed unsuitable altogether for the purpose of implementing the visual synthesizer functionality, for the reasons discussed in section 2.5.1. To reiterate the criteria for this choice, it was made because in VPL's it is very difficult to define complex data structures or implement recursive procedures, both crucial in procedural graphics programming. Because circumventing these limitations requires programming in a low-level declarative language, the advantage of using a VPL is largely negated, as users are required to learn two languages anyway. Other declarative language alternatives besides Processing, such as game engines, although very capable, were too deemed unsuitable. This was because using them would mean that artists wanting to themselves extend the visual capabilities of the system developed, would have to be able to program in C++ and OpenGL, a prohibitive obstacle.

Processing was chosen to serve as prototyping software for implementing the visual synthesizer functionality. The criteria for this choice were: firstly, because it uses the procedural programming paradigm, it allows for the vast pre-existing online source code resources to be easily adapted into implementing new visual algorithms, which is not the case for any of the visual programming language alternatives. Secondly, it has a comparatively big community of users supporting the software and sharing their source code. Third, it is a free open source package, which means that it is infinitely customizable in case the need presents itself. Finally, code written in the Processing language could with great easy be ported to Java. Furthermore Processing itself integrates seamlessly with Java, allowing software written in Processing to very easily be incorporated into any software.

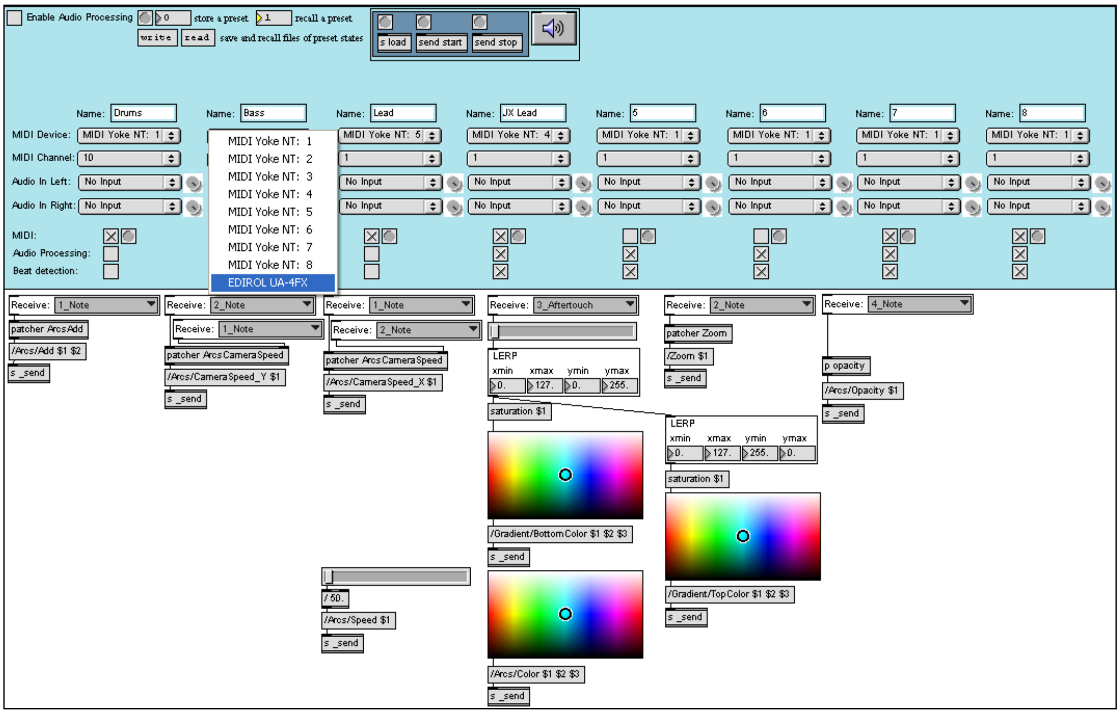


Figure 41: Screen capture of the input processing module prototype, as it appears when set up with a patch for a specific performance (Top). The mapping module prototype implementation can also be seen, set up for the same patch (Bottom).

Using Max/MSP, a prototype of the live input processor was implemented, a screen capture of which can be seen in Figure 41. Most of the functionality of the final version was present also in this early prototype.

It supported the real-time concurrent input from up to eight instruments, from which MIDI data was fetched, as well as stereo audio. Monophonic pitch detection, beat detection and absolute amplitude detection was performed on the audio input of each separate instrument, to support the use of non-MIDI instruments, as well as provide the flexibility of augmenting the MIDI data should this be desirable, since, as previously discussed, some information cannot be accurately conveyed using MIDI, most importantly audio amplitude. The limitation of eight instruments was deemed to be sufficient, partly because comparatively few groups of musicians have eight separate instruments, and because while eight-channel audio interfaces are relatively inexpensive, devices with a higher number of inputs require a significantly higher investment.

To allow experimenting with creating various mappings, a set of tools was developed for the purpose in Max/MSP, which could be used in a live-coding manner to quickly implement new mappings. Through informal experimentation by the author, this process led to the definition of a mapping vocabulary which involved the following basic elements:

a) linear interpolation

b) positive or negative offset of the value

c) multiplication of the value

More involved mappings were also experimented with, but were deemed to be too complicated to be intuitively used in a live-coding context.

A visual synthesizer application was developed in Processing implementing a variety of graphical effects, which could be controlled, as well as switched between, using OSC control messages. The facilities for implementing of new visual effects by new artists was at this stage implemented as a programming API rather than a plug-in API. The source code was needed for the whole application, and the inclusion of a new visual effect required the recompilation of the complete application. Therefore, although this was a step in the right direction, through providing tools for creating a live visual performance environment in Processing, an easy to learn programming language intended to be used by artists, the significant limitation of having to engage in software development in preparation for each performance still remained.

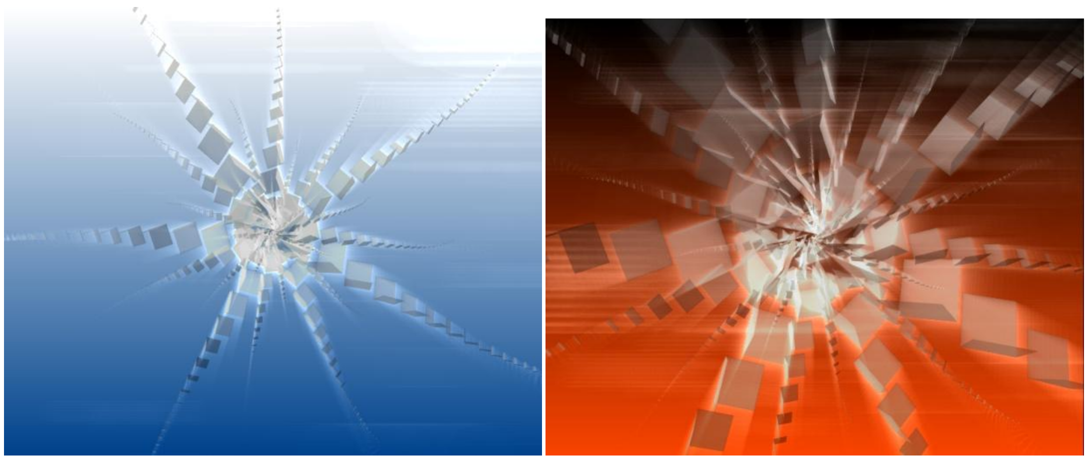


Figure 42: Images to illustrate output possible with prototype at the second development stage.

5.3 Third prototyping stage: the birth of Mother

From the third prototyping stage and on, the system as it developed was on several occasions brought to the regular in person meetings of the members of the www.vjlondon.org online forum. The purpose of these meetings is for the attending

VJ’s to demonstrate their technical setups to each other, exchange ideas on how to use the tools, and also organize events with live visual content. Several professional VJ’s attend the meetings, along with semi-professionals and amateurs. These meetings have been an ideal forum for observing the practice of VJ’s so as to draw ideas for present work. They have furthermore been invaluable in providing continuous feedback on the ideas and system being developed in this work, from demonstrations to and discussion with VJ's attending these meetings.

The LIP software was after testing its technical performance with a variety of pre-recorded and live musical control data input, deemed to be suitable for the task it was created for, and so little development was devoted to it beyond this point, with exception to incremental refinement of the implemented feature set.

For the mapping process, the live-coding paradigm was deemed to be unintuitive and thus unsuitable for the task. In seeking for alternatives the OSCMapper application was found and evaluated for the purpose (Malloch et al., 2007). OSCMapper is intended to be used as an aid in the process of designing new DMI’s, through allowing the quick prototyping and alteration of mapping setups.

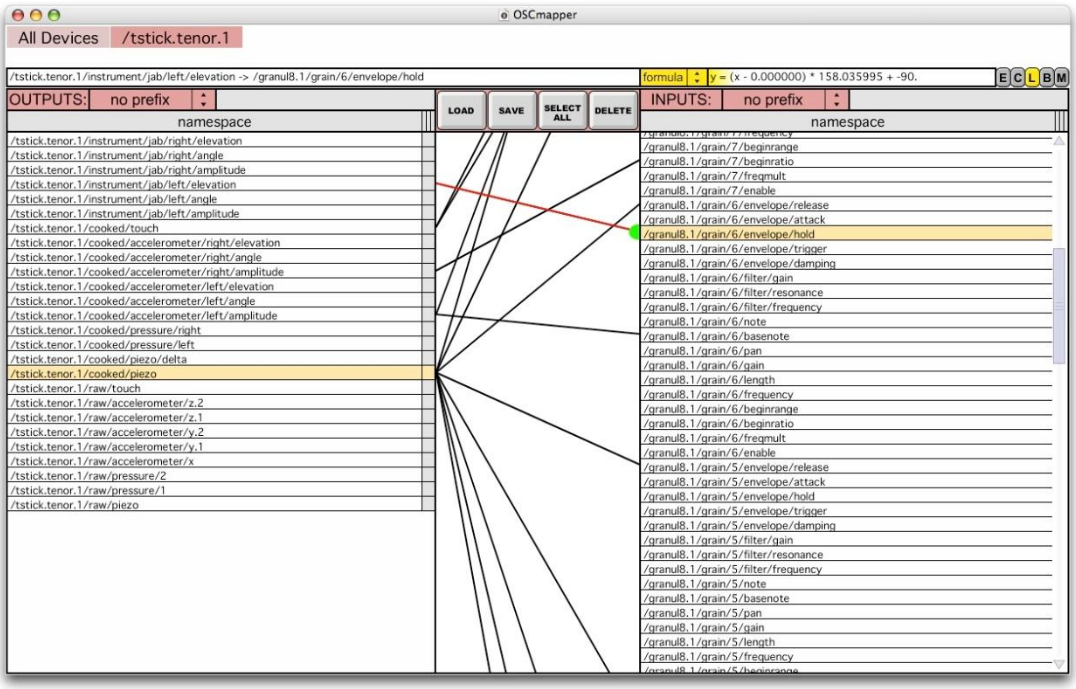


Figure 43: The OSCMapper application (www.idmil.org/software/mappingtools).

For the visual synthesizer module, the existing Processing codebase was ported to java, and was used to create the application later christened Mother, along with a series of plug-in visual synthesizers for it. Mother implements all the functionality necessary for artists to prepare a visual performance without engaging in programming, while still retaining the ability of shaping the performance based on their own aesthetic preferences. As the Mother application is detailed extensively in section 4.1, it will not be detailed further here.

For transmitting control data directly to the visual synthesizers, without it being in any way mapped to incoming control data from musical instruments, the live-programming paradigm was still employed. This was still also used for adding/removing/rearranging visual synthesizers.

At this stage, the software was used for the first time in a live performance context in a public venue, accompanying live electronic music.

5.4 Fourth prototyping stage: Mutable Mapping and the Mediator

What still remained to be addressed was designing and implementing the capability for the system to usefully handle the high mapping complexity between the parameter spaces of incoming control data, and control parameters of the visual synthesizers. The literature on DMI design was surveyed for complex mapping strategies, and all such mapping strategies located were considered, albeit in the end dismissed.

At this stage the notion crystallized that the mapping in itself is, in this context, a form of artistic expression in its own right, and that it makes little sense to use a strategy with less than full control over the details of the mapping. It was considered to be highly advantageous if the mapping was allowed to gradually vary during the course of the performance.

With this notion at hand, several designs were sketched out for a suitable paradigm, which would allow for complex mappings to be gradually altered. All such design ideas were documented, and the advantages and disadvantages of each were considered. The design for a weighted matrix was at this point selected as the most successful idea, and work progressed to design the user interface that would best implement this paradigm. Again several variations of the user interface were detailed and compared. The process eventually led to the basic design of the Mediator application, as is detailed in section 4.3.

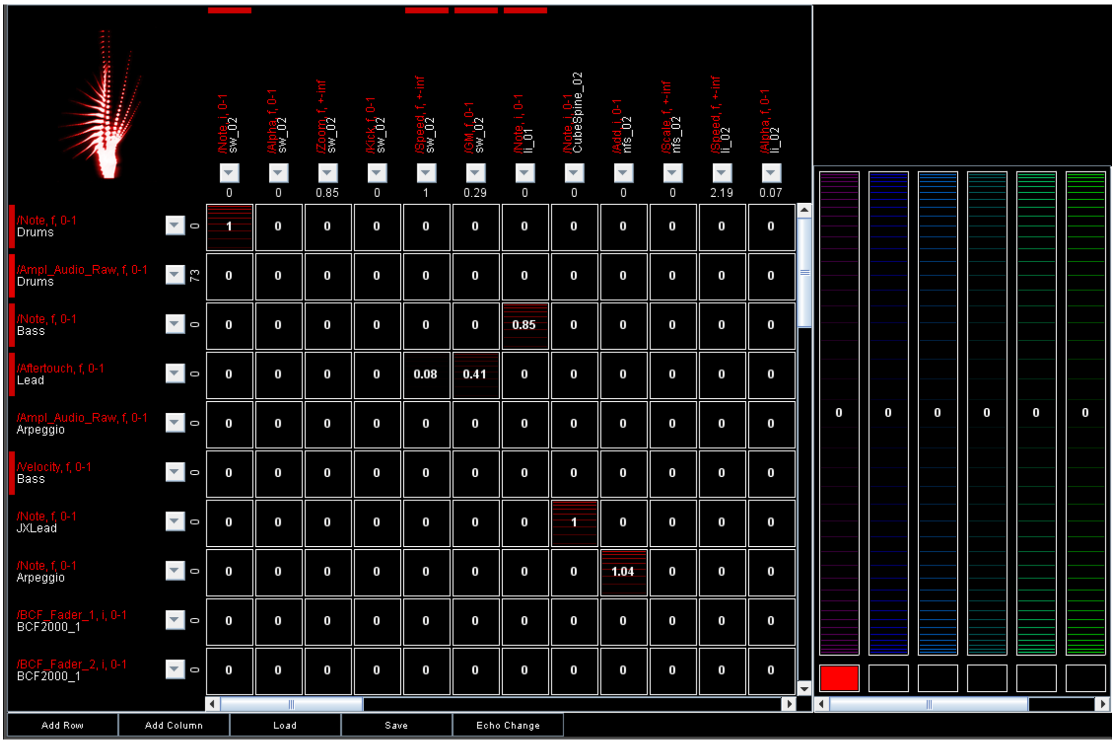


Figure 44: Screen capture of intermediate Mediator implementation.

Adding/removing/rearranging visual synthesizers, and directly transmitting control data to them that is not mapped to incoming control data from musical instruments, was still accomplished using the live-coding paradigm.

All the while, the LIP and Mother applications were undergoing only smaller refinement, as their basic implementation was in practice found to be suitable for the purpose intended. Several more visual synthesizer modules were implemented for Mother however.

At this stage of the software implementation, several more public live performances were made, but still always only using pre-recorded musical data, and not live musical input.

5.5 Fifth prototyping stage: non-musical control

The last extensive software development effort for the project was the development of the "control" screen of the Mediator application, to entirely replace the last remaining live-coding element.

The development was initialized after the realization that the offset values used in the Mediator's mapping matrix, could usefully serve as a replacement for the comprehensive control screen previously used in a live-coding fashion.

Following this addition, only smaller scale development was undertaken, of refining the existing applications, and creating more visual synthesizer material.

It was also at this stage that the system was used in live rehearsals and performances with groups of musicians.

The state of completion at which the system was at this stage is detailed in its entirety in section 4.

From this chapter we have described the development process followed for the creation of the Trinity system, to address the process evaluation criterion for design research proposed by Zimmerman et al (2007), as is in this thesis reviewed in section 2.6.2.5. In summary, their argument for detailing the process is that while there is no expectation that it could ever deterministically produce the same results if applied again, it is important that it is described in enough detail to be reproducible. A rationale for why specific methods have been selected needs to be presented, so that the work can then be judged based on this, and on the rigor with which the work has been carried out.

6 Experiments

The results of the experiments detailed here, are presented and analysed in chapter 7.

6.1 Online audience survey

The survey is employed to address two questions.

Firstly, whether it is it more or less engaging / enjoyable for the audience if the degree of correlation is varied between the music played and the graphics elements derived from it. This question relates to gauging the advantages of increasing the mapping complexity between music and visuals, in a simulated scenario, rather than a direct comparison to established practice. Note however that in this study, no mutable mapping performance is used.

The second question addressed is to quantify the extent in which viewers of a piece created with the system, find it engaging/enjoyable.

To facilitate discussion regarding the first question, the term *composition-level correlation* has been coined, to refer to correlation between visuals and music where the individual instrument parts are reflected, but not the individual notes. In composition-level correlation, when instruments enter or leave the mix, this change is reflected both in the music and in the visuals. The term *note-level correlation* in turn, refers to individual note events, and other short, fine-detailed events, being reflected both in the visuals and the music. With the Trinity system, both composition-level and note-level correlation are easily achieved.

For the sake of clarity, although this survey does not directly compare the Trinity system to previous practice, it is nonetheless worth explaining that in previous practice the only correlation that can be automatically achieved is on a beat and loudness level, entirely ignoring the existence of multiple separate instruments, and the notes each produces. Previous practice thus neither fully achieves composition nor note level correlation, unless additional manual control is introduced to the performance. The differences in the synchronization detail between musical and visual events achievable employing previous practice, and the system developed here, are manifold:

- In previous practice, false positives and false negatives are frequent occurrences, due to the imprecision in the method in which beats are detected.

- In previous practice, detected beat events cannot be associated to individual instruments, and so it is impossible to create associations between individual elements in the visuals, and individual instruments in the music.

Although several studies have taken place with an analogous goal to the one conducted here, as reviewed in section 1, these have not been undertaken in a context where the audiovisual content is specifically created with aesthetic intent. The experiments have rather been examining human perception in a general context. It was therefore deemed interesting to experimentally gauge whether what we know from Lipscomb's (2005) study, where congruent audiovisual stimuli are by subjects perceived as more effective, further translates to the stimulus also being experienced as more aesthetically pleasing. This would experimentally corroborate Chion et al.'s theoretical argument (1994, p. 63) about synchresis, which both Grierson (2005), and Moody (2009) hold as fundamental in the arguments presented, for the audiovisual composition art form, and the Ashitaka audiovisual instrument respectively.

A smaller scale study has been previously conducted, in which the first prototype of what became the Trinity system is directly compared to previous practice (see section 5.1, and (Bergstrom, 2003)). In this -small scale and informal- study subjects reported a preference for the prototype system, employing both detailed note and composition level correlation, over the previous practice scenario, which was highly lacking on both. Composition-level correlation is uncommon in current live performance practice, such as VJing and live procedural graphics, It is however very commonly occurring in pre-rendered animated visual music, and so it can be safely assumed that to a great extent the reason it is not used in a live context is simply because the technology with which it can be readily achieved has not been available. Clear note-level correlation on the other hand, is not as frequent an occurrence, neither in live nor in pre-rendered animated visual music, despite the arguments presented in section 1, that such correlation causes percepts to be perceived as more effective. Although it is common, animations without note-level correlation are arguably equally common.

It was for the present study therefore more interesting, not to simply perform a direct comparison to previous practice, but to also assume that composition-level correlation is a given, and gauge whether more accurate note-level correlation alone will be sufficient to trigger audiences to notice a conceptual difference, and perhaps even a preference. To achieve this technically, composition-level correlation in the previous practice scenario is manually recreated.

Using the software developed for this project, two videos were created of the same piece of music using the same visual synthesizers and mapping complexity. The only difference between the two videos was the nature of the incoming control data used to generate them. The music used was a pre-recorded, pre-existing piece created by the author prior to this research project.

The first video was created using the full bandwidth software under evaluation. Control data included MIDI, and some additional data from a Novation UC33 MIDI controller. The additional data from the UC33 is performed directly by moving the UC33's knobs or sliders, and is for controlling additional changes by altering visual synthesizer parameters not directly connected to input from instruments. For the video described here, these parameters controlled the rotating of a visual synthesizer, and the alteration of some visual synthesizers' colours over time. The second video derived its control data from beat and amplitude detection over a stereo mix-down of all of the instruments of the music combined, to simulate previous practice. Again the same additional control data as on the first animation will come from the UC33 controller. The composition-level narrative of instruments entering and leaving the mix is manually recreated to emulate the same sequence of events in the first video, by altering the transparency of the visual synthesizers using the UC33 controller, as a live performer would have been able to do using tools available in previous practice.

As a consequence, the composition-level correlation between events in the music and events in the visuals is virtually identical in both videos. The only difference between the two videos is therefore the note-level correlation. It is of course finally worth repeating here that achieving composition-level correlation employing previous practice, although possible, requires the continuous attention of a performer largely dedicated to the task, while with our system, composition-level correlation can be achieved automatically.

The questionnaire used is an adaptation of the Audience Response Tool (ART) survey (Glass, 2006), adapted for the context of this work, by removing the open ended questions. The ART questionnaire was initially designed to elicit the emotional response of a dance performance's audience, but could with little modification be adapted to the purpose of this study. Although this is the first time ART is employed by researchers other than its creators, it was still deemed to be the most appropriate choice. The only other questionnaire option in existence would be the AttrakDiff questionnaire (Hassenzahl, Burmester, & Koller, 2003), which is however specialized towards evaluating interactive experiences, and thus less appropriate for the purpose of present work. The free AttrakDiff survey is also much less customizable, making its use cumbersome for this work, in comparison to the here chosen free online survey creation tool, Survey Gizmo (http://www.surveygizmo.com). An important feature which the free survey gizmo service allowed implementing is the randomization of the order in which the two animations and their corresponding questionnaires are presented, so as to avoid bias.

The videos embedded in the questionnaire are hosted on the Vimeo video hosting service (http://www.vimeo.com). The quality of service of a paid for Vimeo account was deemed high enough for the purpose of the survey, with a reasonable trade-off between video quality and transfer bandwidth. Several other competing services were tested and failed to provide comparable quality.

Following from the data gathered, and since the ART survey was designed specifically for the task, it is trivial to use the results gathered for only the first video, to address the second aim with the survey: gauging the extent to which viewers of a piece created with the system, find it engaging/enjoyable, through examining their responses to the likert-scale questions.

The survey used is available to read in appendix A, along with all the data from the completed responses.

6.2 Rehearsals and live performances with groups of musicians

The ethnographic study was carried out in a series of rehearsals with groups of musicians. The subjects needed for the trials are clearly very scarce, and it was from the beginning realized that it would be difficult to find suitable groups of musicians, therefore necessitating the careful deliberation of a suitable sampling strategy.

It was deemed necessary that in every session the musicians always played in a group, as that is by far the context in which the benefits of the Trinity system are most apparent, in comparison to previous practice. The musicians furthermore had to have some previous experience with playing together, so that they are used to each other and compatible as individuals and as musicians. The music played however had to be improvised to the greatest possible extent, otherwise it would be difficult to gauge what influence the experience has on their altering their performance. A majority of the musicians in each group have to play instruments that are compatible with the system, such as electric or electronic instruments, alternatively acoustic instruments that are easy to ‘mike up’ to produce a separable audio signal. For example, woodwind instruments emanate sound from a multitude of points and in many directions, making it much harder to capture sound from them in isolation than say an acoustic guitar, for which the primary source of sound is its sound-hole. Finally, both rehearsals and live performances were pursued.

A significant amount of effort was concentrated on contacting as many musicians as possible, with the hope that a sufficient number of musicians would agree to participate. In brief, these efforts were primarily concentrated within a five month period during which:

- The music students of five universities were emailed in London, through the mailing lists at their universities and through contacts at each respective institution. See appendix C for the exact text that was used to contact them.
- Several musicians were contacted online, after they were deemed to be suitable candidates, as judged by the music and information they had posted on online social services for musicians, primarily http://music.myspace.com, and http://www.last.fm.
- Additional musicians were also contacted in person, after they were deemed to be suitable candidates, as judged after their live performances were attended.
- Calls for collaboration were posted on numerous online message boards for musicians and sonic artists.
- Three music projects for young adults were contacted to gauge whether collaborations were possible with their participants.

The musicians were always contacted through an invitation for artistic collaboration, either only in the form of rehearsals / jam-sessions, or working towards a live performance, and not with the primary pre-text that the sessions are motivated as being academic experiments. This strategy was deemed a crucial aspect of the evaluation, as then the whole process would be a genuine artistic collaboration, not a laboratory attempt at simulating the conditions of one. It was nonetheless always mentioned that the session would also be useful in contributing to research, purely for the ethical reason of having fully disclosed all underlying motivation for the collaboration.

The process of contacting the musicians was initiated only after the system under investigation had been developed sufficiently for it to be useable in a real world context, as judged after following an iterative development process. The system would otherwise have been useable only in a context with a controlled experiment very limited in scope, and not the desired 'naturally occurring' true artistic collaboration.

Before the musicians were contacted, the system was used to produce a series of animations, accompanying music composed by the author, so that the system and its output could be demonstrated to them.

6.3 Interviews with VJs and live visuals performers

Towards eliciting feedback from performers employing currently established practice, interviews were conducted with VJ's / live visuals performers. As a preamble to these interviews, to connect with the live visuals performance community, and to get continuous feedback on the system as it developed, the regular in person meet-ups of the members of the www.vjlondon.org online forum have been frequently attended.

The purpose of these meetings is for the attending VJ's to demonstrate their technical setups to each other, exchange ideas on how to use the tools, and also plan events with live visual content. Several professional VJ's attend the meetings, along with part-time professionals and amateurs. These meetings have been an ideal forum for observing the practice of VJ's so as to draw ideas for present work. They have furthermore been invaluable in providing continuous feedback on the ideas and system being developed here, from discussion with attending VJ's, to which the system at varying stages of development has been demonstrated.

These interviews may be viewed in the light of ethnographic method, as a process of gathering feedback from established practitioners by means of interview and observation, but they may equally validly also be viewed in the light of design research, as the practice of expert critique discussed in section 2.6.2.

Following the completion of the system’s final prototype version, live visuals performers that have shown an interest in the work, were contacted towards engaging them in one-to-one interview sessions, to gauge their feedback on the completed version of the software.

The process of these interviews was the following:

First, a summary of the limitations detected in current practice was described to the subjects, as it is detailed in section 1.1 of this thesis, followed by an explanation of how the ideas behind the Trinity system are intended to work in addressing these limitations, as detailed in section 1.2.

The system's three applications were then demonstrated, both how each is meant to function on its own, and in conjunction with the other two.

If the subject is a user of Processing, and thus has his own Processing sketches, he/she was guided through the process of porting these to Mother, and incorporating them with the Trinity system.

The subjects were then allowed some time to experiment, and improvise a performance, until they felt they have grasped the use of the system.

During and particularly after this process, the subjects were interviewed. Provided they agreed, they were also filmed during the interview stage, so that the discussion and their feedback exist in recorded form for future analysis.

During the final interview stage, the following questions were asked of the subjects, if they had not yet been sufficiently addressed in the discussion with them:

- Would this system as a whole, or parts of it, be useful in their current or future artistic practice?
- If so, would it be the ideas, or also the particular implementations of these ideas created for present work?
- What would they imagine using it for, that differs from the intended use?
- What usability challenges did they experience, and what usability improvements would they recommend?
- What other criticism, suggestions and observations do they have?

In this chapter we have in detail established what experiments were carried out towards the validation of the ideas and systems presented in this thesis, and what rationale lies behind each of these. Rehearsals and performances were undertaken with groups of musicians, to gather feedback on the overall experience of Soma performance. Live visuals performers familiar with established practice were engaged, to elicit feedback on the concepts embodied in the system, as well as their implementation. Finally, an online audience survey was conducted, for providing audience feedback on how engaging/enjoyable audiences found it to watch a video created using the system, and to gauge whether increased accuracy in the temporal correlation between audiovisual stimuli results in a more aesthetically engaging experience. In the subsequent chapter, the results of these experiments are detailed.

7 Evaluation

7.1 Results from online survey

The survey attracted a total of 30 fully completed responses. The survey was promoted through online communities on visual music, for musicians and for visual performers, so the respondents were expected to all have a significant familiarity with music, visual arts, or both. Out of the respondents, 18 were musicians, with an average of 14.35 years of experience, 17 visual artists, with an average of 13.56 years of experience, and from these groups, 10 were both visual artists and musicians. Tests were performed between these sub-groups to gauge whether each subgroup submitted significantly different responses to the survey in comparison to the others, but no such difference was found.

The responses for each sub-group of questions (as defined by the ART surveys creators) were summed into a score, and the median values were compared for each question sub-group, between the two identical questionnaires. From this comparison, the results of the survey were positive. For all groups of questions, the great majority of subjects reported that the video created employing both note and composition-level correlation was more stimulating, enjoyable and emotionally engaging to watch, compared to the composition-level correlation only video, albeit with the caveat that the reported difference was quite small, as is evident from the following results.

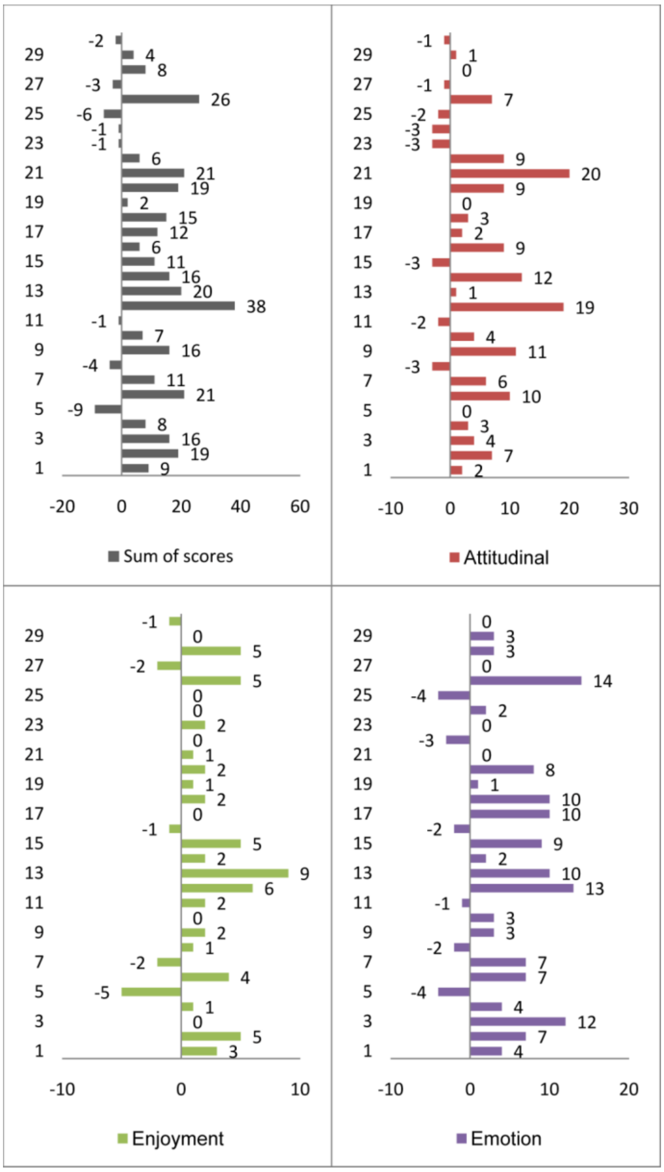


Figure 45: Differences between summed and individual scores of all subject responses. Positive values mean preference for video by Trinity system, negative for previous practice simulation.

The ART survey uses primarily ordinal Likert-scale input, so a Kruskal-Wallis test was performed (Harvey, 1999). The results from the test failed to show statistical significance. For the survey's attitudinal rating scale, the test produced a P value P = 0.273, for the enjoyment rating scale P = 0.4195, and for the emotion rating scale P = 0.192.

To summarize the results, here the average and median of the differences across all subjects will be presented, on a scale of -7 to 7, positive values representing a preference for the more detailed note-level correlation, and negative values a preference for the less detailed note-level correlation. For the attitudinal scale the average is 0.178 and the median 0.288, for the enjoyment scale the average is 0.313

and median 0.2, and for the emotion scale the average is 0.429 and the median is 0.333.

The fact that the survey failed to achieve statistical significance may be because of the relatively small number of responses. It most certainly also is a result of the choice made to only vary the detail of note level correlation between the two videos, because the scenario where both vary was, based on existing knowledge, deemed to be too trivial. Had composition-level correlation also been varied, the differences would likely have been much greater. The survey would moreover have been a clear comparison between the capabilities of the Trinity system, and of previous practice, rather than a comparison between taking advantage of only one of the benefits of the Trinity system, or both of them. The two advantages referred to being high note-level correlation and composition-level correlation, neither of which can be readily achieved in live performance employing established practice.

What these results therefore tell us, is that, although a majority of the audience members used in this study have a preference for more detailed note-level correlation, over having only clear composition-level correlation, from the current study that preference is quite small, and is not clear enough to warrant drawing conclusions about what audiences in general prefer.

Addressing the second question, of gauging to what extent viewers of a piece created with the system, find it engaging/enjoyable, the average mean, median and mode for the attitudinal, enjoyment and emotion scales for the first animation are the following, on a scale from 0-7:

Attitudinal:	mean: 4.233	median 4.357	mode 4.428
Enjoyment:	mean: 4.573	median 4.7	mode 5.8
Emotion:	mean: 4.125	median 4.166	mode 5.111

The complete response data is available to read in appendix A.

7.2 Results from rehearsals with musicians

2.7.1 Setup of conducted rehearsals

Several musicians expressed their interest after seeing the proposal for collaboration, out of which members of nine different groups were interested enough to meet for a discussion and a demonstration of the software. Out of these, collaborations went ahead with four of the groups. One of the groups was located at a very long distance, and so although visual material was produced accompanying recorded multi-track music of theirs, a rehearsal and accompanying performance did not take place. With the remaining three groups however, a total of roughly ten hours of rehearsal took place at three separate occasions, followed by two live performances at public venues. In total, seven individual musicians have been connected to the system, with one musician playing in two of the groups involved. Since at the live performances the musicians necessarily face the audience and not the projection screen, the following discussion will only concern itself with the experience of the musicians from the rehearsals.

A great variety of instruments were used during these rehearsals. Connected using individual audio feeds to the system were an electric guitar, an acoustic guitar and a singer. Furthermore, an acoustic drum kit was connected, using a set of Roland drum triggers and a TMC6 trigger-to-MIDI converter transmitted MIDI data for each individual drum. A Roland Handsonic electronic percussion synthesizer was connected both using direct audio and MIDI. A synthesizer keyboard transmitted MIDI data. Finally, a Native Instruments Maschine device, which operates using a conducting gesture interface principle (Miranda & Wanderley, 2006), was connected using direct audio. Therefore, out of lucky coincidence, all major groups of instrument types have been represented at least once. During each rehearsal only one researcher was present, thus assuming multiple roles:

a) To act as a Mediator operator and visuals performer, by connecting the control data from the musical instruments to the control parameters of the Mother visual synthesizer software, and adding/removing/rearranging visual synthesizers as previously described in this paper (section 2).

b) To act as an external observer of the musicians' as they gain experience in using the system under evaluation.

c) To film parts of the rehearsals with a digital video camera for later observation.

d) To interview the participating musicians about the experience, during breaks, and directly after the rehearsal.

Directly following each rehearsal, a report was written recollecting the experiences and observations drawn.

As the evaluation process also involves a comparison to previously established practice, during each rehearsal, without the musicians being warned beforehand, a simulation of previous practice was performed. This was done to elicit whether the musicians notice a difference between that state and the assumed more advanced state facilitated by the system proposed here. The previous practice simulation was achieved by mixing together all incoming control data into a single channel, instead of treating them as separate streams, thus emulating the case where only a stereo signal input is available. In the cases where the musicians also had some independent experience with previous practice, they were additionally asked to reflect about the experience with the Trinity system in comparison to that.

The two primary questions that were deemed interesting to answer through these trials were: first whether the musicians experienced the subjective sensation that they were controlling the visuals projected, and, if yes, to what extent in comparison to previous practice, both from their own experience of previous practice if they had any, and from the simulated previous practice scenario as recreated during the trials. Second, whether the visual representation of their playing influenced the playing itself, and if yes, how it did so. Initially it was attempted to elicit the answers to these questions from each participating musician without directly posing them, through informal discussion, however if at the end the answers had not been mentioned, the questions were posed directly.

7.2.1 Summary of conclusions from rehearsals

All musicians were observed to watch the projection screen for extended periods of time. For long periods of time they also looked at the other musicians in the band, or concentrated on looking at their instrument. They sometimes also distanced themselves from actively interpreting visual information, shutting themselves off from the visual world and concentrating only on the music performed. Some reported that the very fact that there was a visual accompaniment altered their experience of playing, as they were not accustomed to concentrating on the visual world while performing music. As such, sometimes a conflict was experienced, in that at times the visual information distracted from the musical performance, while again at other times visual information was experienced as an integrated part of the performance as a whole. In the latter state, musicians reported that they found it difficult to look away from the projected graphics, there being a unity between their playing, the resulting music, and the resulting visuals.

Four out of seven musicians reported that at times they did indeed experience the subjective sensation of being to some extent controlling the visuals, instead of the visuals just reacting to their playing. However this was not a continuous sensation, instead it varied according to the complexity and nature of the mapping in operation in each particular instant. Three out of seven of the musicians however reported that they did not experience this sensation at all. These were all in the same group, the only one of the three groups that had a body of songs they were refining during the rehearsal, rather than completely improvising. This could possibly have caused the musicians to concentrate less on the influence their playing had on the visuals, and more on observing how their compositions developed. Given the very small number of subjects however, deterministically gauging whether this is the case is impossible, and will require future study involving larger numbers of subjects.

When asked whether the musicians noticed a difference in comparison to simulated previous practice however, six musicians stated that they did indeed observe there being a stronger connection between the music played and the projected visuals, when employing the system evaluated here. Some musicians observed this independently, and did notice the altering between the two states during the rehearsals, while others only stated they realized this difference after it was pointed out to them.

Four of the musicians reported that they were influenced to alter their playing as a direct result of being connected to the system. Some stated that this was as a result of the aesthetics of the projected visuals, while others stated it was more to try to figure out what the mappings were between their instrument and elements of the graphics. In the latter case, musicians reported to lose interest after having figured the mappings out, and returned to performing music without being directly influenced by the visual accompaniment. Finally, two of the musicians stated that the visuals did not at any time during the rehearsal directly influence their playing, although they did enjoy the accompaniment.

The full report from all rehearsals and performances is available in appendix B.

7.3 Results from interviews with VJs and live visuals performers

Of all relevant VJ's and live visuals performers that were contacted, four agreed to take part in the one-to-one interview sessions. Following are the reports from each of these sessions in turn.

Figure 46: Mauritius Seeger, AKA Dr. Mo, and the three applications comprising the Trinity system.

7.3.1 Olivier Ruellet, AKA Ctrl-N

Olivier is one of the very few performers of live procedural visuals who regularly accompany improvised performance by a group of live musicians. He creates all the software used in his performances himself, implemented using the Processing language, but he is also familiar with the established practices of VJing, which he sometimes employs in parallel to his procedural visuals during performance. The software he has created uses a stereo signal from the band playing as input to control the visuals, while he additionally controls his visuals software with non-musical controllers. The graphics algorithms are built into his software, and as such the application also embodies the aesthetic preferences of Olivier, its creator, and is not an instrument that other performers can use in their own practice without significant modification. His work can be viewed online at his website: www.ctrl-n.net.

His artistic practice also goes beyond live visuals performance. He creates traditional narrative animations, as well as visual music animations. His studies are in visual communication, and he is since several years employed as a teacher of digital animation and of digital arts at Thames Valley University. He additionally works as a freelancer, producing digital art and animation.

Olivier has followed the development of the here presented project for over a year, and has at several occasions been updated in detail about the developments of the system, and the ideas it engenders. He has from this process become familiar with the use of Mother, and has experience in porting his own Processing sketches so that they may be used as visual synths within Mother. As he had already completed the development of his own performance software before learning about Mother however, he has not transitioned to using Mother in his performances, but continues using the software that he has developed himself.

The session with Olivier lasted approximately four hours, during which the procedure detailed in section 6.3 was followed. He agreed to have the interview portion of the session filmed, and so the text below is a summary of his feedback as captured on the video recording.

After having explained the concepts behind the Trinity software to him, Olivier showed particular enthusiasm about the concept of Soma, proclaiming that he is indeed a Soma performer, as the idea agrees strongly with what he has in mind for his own performances. He was additionally positive that he saw significant benefits in the increased control complexity that the Trinity system allows from using the full bandwidth of possible control data from the instruments of the musicians, and of the concept of mutable mapping during the course of the performance. He further added that he would indeed see himself using a system analogous to the Trinity system in his practice, or Trinity itself, provided its development matured a bit further.

Although he did see himself using the system, he was deterred by the necessary setup time, and said that for informal performances of shorter duration, he would be apprehensive of using a system of this complexity, simply because setting it up and tearing it down can take up to an hour or more, depending on the composition of the band of musicians. He would however not see the setup process as a deterrent for high-profile performances of short duration, or longer performances, where he is part of the headlining act of the evening.

On discussing the design of the Mediator software in particular, he would indeed see himself using the actual Mediator application in performance employing mutable mapping, after it had been developed further to address a number of smaller usability difficulties that he identified. His background in studying visual communication, allowed him to provide comments on how the mapping matrix may visually be represented more clearly in the software, so that the information that is more important is more prominent to the user, while less important data is subdued in its visual representation to make it less distracting.

He stated however that ideally, he would want the Mediator functionality to be implemented in a physical, graspable interface, such as a matrix of physical knobs instead of controllers represented on a touch screen. He agrees though that such an implementation is difficult to create at present, since it would require hardware that is not readily available, specifically a large number of infinite rotary knobs, each with its own numerical display to reflect its value.

With regards to the Mediator application's functionality rather than its visual appearance, he believed the core feature set was sound, but would want a more detailed implementation of loading and morphing between groups of pre-sets for the various settings in the application. He suggested allowing the loading of rows and columns of values into the matrix, and not just the setting for the whole software, as is the current functionality. He suggested using a double-tap on the cells of the value to reset it to zero, instead of resting the finger on it for two seconds as the current functionality is. He further suggested having an additional form of visual feedback of the value in each cell, besides the numerical value, although he agreed this would be difficult to achieve since the values are unlimited, and so there is no fixed value range to represent.

A feature he missed, but which he agreed would be difficult to implement, was to have a live visual preview of what each visual synth looks like on its own, above its corresponding column in the Mediator, both for preview purposes, and for easing identification. That would also reduce the generality of the Mediator software, as in its current implementation it is agnostic as to what the actual elements of its rows and columns represent.

The functionality of the live input processor application he found was sufficient for the task it was designed, and did not need any significant further improvement for him to use it.

Finally, he also agreed that the state of current live procedural visuals performance software is far from as advanced as would be ideal, and in that context, Mother is a good choice of software in a live performance context. He does however lament the general lack of more advanced software for this purpose.

7.3.2 Mauritius Seeger, AKA Dr. Mo

Mauritius is a prominent figure in the London VJ scene. He is central in organizing VJing events, of which the most known is the recurrent Electrovision night (www.electrovision-cinema.com). Several of London's prominent VJ's have performed at Electrovision events, while also one of the two solo performances conducted using the Trinity system, took place on an Electrovision night. Mauritius is also often central in organizing the VJ London meetings.

His academic studies are in physics, followed by a PhD in machine vision. Professionally, he works as a VJ, and as a freelance creator of interactive installations, running his own creative studio. He has worked for several high profile clients, both with VJing and with his practice in creating interactive installations. His work can be viewed at www.morishuz.com.

Mauritius too has followed the development of the here presented project for over a year, and has at several occasions been updated about the developments of the system, and the ideas it engenders. His own live visuals performance however is employing established VJing practice, in which he only uses pre-recorded videos mostly of his own creation, and no live procedural visuals. As such, he does not himself have any practical experience of what creating content for Mother involves. He does however in general have experience of programming procedural computer graphics, from his work in creating interactive installations, and so does have a general understanding of the nature of the work involved.

The session with Mauritius lasted almost four hours, during which the procedure detailed in section 6.3 was followed. He agreed to have the interview portion of the session filmed, and so the text below is a summary of his feedback as captured on the video recording.

He began with repeating what he has also been often telling me when we meet: that clients approaching him for work often mention Mother to him, asking if it is something that he can make use of in the work he makes for them. To date he has not used it however, as it has not been relevant to the projects for which he has been commissioned.

Having discussed the ideas behind the system with him, and after seeing the system, he did agree that the concepts in and of themselves, definitely are relevant and useful in the context of live visuals performance, and he is positive that the paradigms, separately or combined, are likely to eventually find their way into mainstream use. After describing the ideas of Soma and of visual music to him, he stated that he finds his work to be conceptually related to visual music, although he was previously unaware that the term existed, or that it had the history it does.

Furthermore, he does see himself using mutable mapping, and likes the paradigm. He also does see himself using complex input, from several instruments, audio and midi. However, in his personal performance practice, he would most likely still not use procedural graphics, but instead replace the Mother software with one that layers multiple pre-rendered video clips, like in VJing, but which allows more complex external control, and the use of large numbers of simultaneous video layers, neither of which is currently supported by existing VJing software. He envisions that each video layer, is connected and controlled by a single instrument, using complex control data from each, channelled through the Mediator software. He mentions this is an idea that he has had since before finding out about present work, but that he found was not achievable using existing systems.

The subsequent discussion with Mauritius concentrated on the mutable mapping paradigm, and the Mediator software. He states he really likes the fact that the Mediator is functionally completely separable from the other two applications, and could work in any performance context, not only in connecting instruments and procedural visuals. He further sees that an application analogous to the Mediator should soon be considered ubiquitous in creating and performing sophisticated concert visuals.

With regards to how the mutable mapping idea has been implemented in the Mediator software, he stated he liked how the controls implemented for altering the values in the matrix cells function. Following his testing of the software, he did however also have usability improvement suggestions to make.

He suggests visualizing the signal path that values follow, so that lines trace their path on the screen through the cells, making for a much more dynamic visualization of what data is being routed through the matrix. He also suggests not showing cells that have a zero value, to clear the screen up, so that the cells that do have values are more immediately visible.

Furthermore, he comments that currently, a performer using the Mediator needs to be familiar with the external entities connected to it, and how these react to changes made to their parameters, as from looking at the screen it is not immediately apparent what each element represents. This is amplified because at the moment the Mediator GUI is very textual. It would be more intuitive if more graphics were used to represent values, and icons to identify entities, instead of only representing both with numerical values and plain-text names.

Finally, he suggested implementing layers of cells on the matrix, so that groups of matrix cells could be enabled and disabled, loaded and saved together.

7.3.3 Mat Hourteillan, AKA Matsai, VJ Om boy

Mat is too a rare occurrence in the VJ world, in that he is both an established VJing performer, and an experienced, released musician. His VJing practice strictly follows established methods, and does not involve procedural visuals, only pre-recorded video clips. Most notably, he has long been a live collaborator with the in audiovisual performance ground-breaking group Coldcut, who belong to the very first to bring live audiovisual performance to the stage in a popular music context. Mat's released music is electronic, predominantly sample based, and has been published on the Ninja Tune label. He has also recently forayed into live musical performance in which he sings and plays traditional instruments, in his yet unreleased project Matsai. He holds an MBA, and works with two companies in the field of live visuals, to promote their equipment and services: Vixid, and Musion. More information on his VJing work can be found at: www.myspace.com/vjomboy.

Mat too has followed the development of the project over the duration of roughly a year, and so arrived to the meeting already having a detailed understanding of what the project involved, from previous discussion.

The session with Mat lasted almost three hours, during which the procedure detailed in section 6.3 was followed. He agreed to have the interview portion of the session filmed, and so the text below is a summary of his feedback as captured on the video recording.

He agrees that the system and the ideas it engenders, are relevant to the practice of live visuals performance and VJing, and he stated he could indeed see himself using an analogous system in the future. In particular, he was interested in the Mediator software, and the mutable mapping paradigm. He further states that he indeed does find the ideas presented to be novel, in relation to his knowledge on established practice. He comments that should the software be refined, he expects there would be sufficient interest for it, even to justify its commercial release, based on his knowledge with promoting analogous systems in his consulting work.

He showed particular interest in live audiovisual performance using Ableton Live as a source of control data, allowing for him as a solo performer to control the whole performance using conducting gestures, employing a hardware controller for the Ableton Live software.

When asked what usability problems he found, he stated that there were no big problems that he could identify, but of course that the usability of any software could always be improved. He finally comments that the ideas presented would indeed be most interesting in a live performance context where there is a mutable mapping performer involved in the performance.

7.3.4 Nigel Jenkins, AKA Nebulus

Nigel too is unique as a live visuals performer. This is in part due to also him being one of the very few who routinely accompany live bands with his visuals, and also due to his very different background to most VJ’s, of being a very highly experienced computer graphics programmer. His academic studies have been in computer science, following which he worked for Silicon Graphics during the era when the company defined the high-end of live procedural computer graphics technology. He subsequently worked with developing installing and maintaining virtual reality cave systems for Trimersion, and now works with developing, installing and maintaining cutting edge digital planetarium systems.

In parallel to his work with planetariums, over the recent few years Nigel has developed his own live procedural visuals performance software, reserved for use by himself only, which is unique in many ways. From the extensive review of relevant work conducted within and outside of academia, Nigel’s is the software that is the most analogous in conception and technical implementation to the Trinity system. Most uniquely, Nigel’s software is created so that its output can scale to cover the whole dome of a planetarium, through utilizing several computers.

Similarly to the Trinity system, it uses visuals plug-ins, interchangeable modules each of which defines a procedural graphics algorithm, the output from which is layered in a manner analogous to VJing. Unlike Mother however, these modules are implemented using an API specific to his software, and need to be programmed using C++, thus making content creation for the system inaccessible to non-expert programmers. Nigel’s software also does not allow the visual synths to be combined in as complex a manner as Trinity. While Trinity allows for an unlimited number of visual synthesizers to be used simultaneously, Nigel’s software always only has a maximum of two visual synths active at the same time.

Also similarly to the Trinity system, Nigel intends for his system to be used primarily in a live musical performance context. However, there is no complex mapping between the control data coming from each individual musician’s instrument to the visuals, instead the established practice method of processing a stereo audio mixdown for beat events and amplitude is predominantly used. Nigel has not implemented any mapping variability in the software at all; in fact audio reactivity, based on processing a stereo audio input, is always on. Each visual synth has its own user interface hardcoded, and Nigel states that in some synths he has implemented the option of switching between what aspects of the synth react to the audio input. He has further implemented rudimentary MIDI functionality, but again with an entirely hardcoded mapping for how each visual synth reacts to MIDI data, implemented within each synth itself.

Nigel has further implemented his own non-musical hardware controller, using which he can control his software more intuitively than were he to use only mouse and keyboard, through mapping the buttons, faders and knobs of the on-screen interface, to physical controls.

Nigel performs using his system in various collaborations with groups as well as alone, both in planetariums and in more typical concert settings. He has further joined the band Autorotation, with whom he also participates in rehearsals, to further reinforce his connection as a member of the group. This is in contrast to what is normally the case with visuals performers, of only participating in live performances.

Also Nigel has followed the development of present project over more than a year, and so arrived for the meeting already having a detailed understanding of what the project involved, from numerous previous discussions.

The session with Nigel lasted almost three hours, during which the procedure detailed in section 6.3 was followed. He agreed to have the interview portion of the session filmed, and so the text below is a summary of his feedback as captured on the video recording.

After having been demonstrated the Trinity system, Nigel reflects that it is more complicated to use than the software he has created himself as well as more complicated than most established VJ software. Similarly the learning curve is quite a bit larger in comparison. He further states that with his own software, the goal with its design has been to cater for non-expert use, where the threshold is very low, so that the software can be learned to be used as immediately as possible.

Using the MIDI functionality implemented in his own system, he has experience with his software being connected to a live percussionist. Since the software supports no mapping at all, the percussionist had instead to adapt his mapping by switching the output values of his Roland percussion controller.

Nigel states he really likes the idea of allowing the mapping of all possible control data derived from musical instruments, to create a complicated mapping. He thinks the fact that it is possible to achieve very low granularity is very desirable, albeit with the caveat that there should also be an option to not always be at the lowest level of mapping detail, a solution to which would be implementing facility for presets of mappings in the Mediator software.

He agrees that the system allows the performance to be more closely connected to the music; having the narrative of the music being reflected in the visuals is in his own words a beautiful functionality.

He further agrees, also from his own previous experience with the digital percussion player and his own visuals software, that using instruments as a source of control data, provides for increased expressivity during the performance, as the musician too, to an increased extent, takes part in the performance of the visuals.

Finally, he does concur that the Trinity system is a good step forward from breaking with procedural graphics performance systems normally being one-offs, to use his words. That is, requiring extensive software engineering practice for creating visual material in preparation for a performance.

From his own practice he identifies a problem in designing visual synths: keeping their control parameters few and meaningful. He states that also the synths implemented to demonstrate the Trinity system suffer from this problem to a certain extent. He stresses that simplifying the number of parameters each synth has, is a goal that although time consuming, is worth pursuing, as the benefits to their usability are significant.

He suggest implementing a virtual camera that is tethered to the centre of the world for all visual synths, as a good solution of allowing three dimensional rotation of the synths, without its control becoming too complex.

With regards to the Mediator software, Nigel too had several suggestions for smaller changes that would improve its usability. He commented that altering the value of a cell in the matrix with its current implementation requires that the user looks at the screen. Reverting to the previous implementation of the functionality, where dragging rather than rotating the finger alters the value, would allow altering values without having to look at the Mediator software interface.

He commented that he would want graphical visual feedback of both current values, and the ranges they have over a recent time window, both in the matrix, and for incoming and outgoing values. He suggested that a control similar to an audio level meter, but with both top & bottom values, would be suitable for this purpose.

With regards to the matrix of cells, he commented that the functionality to re-set all cells to zero would be useful, as well as the functionality to switch cells on/off, without zeroing their value, like a mute function.

It is difficult to quickly improvise a mapping. As a solution to that, he suggested the loading of pre-sets for the matrix only, rather than as the current functionality is, of only having pre-sets that reset the setting globally for the whole software. In this way, there will no longer be any interruption in the visual output, resulting from loading new synths following a global pre-set load command in the Mediator.

He suggested the Mediator control screen view should allow for more types of GUI controls, depending on the type of the values controlled, so that for example colour values have a colour swatch control displayed, etc.

Finally, as inspired from his own software, he suggests implementing “general controls” as he refers to them, across all synths. By that, he means that a specification should be drafted, so that synths that have similar parameters, for example RGB or HLS colour, or XYZ position and / or rotation, should all react to values to these in the same exact manner. Most importantly by his suggestion, the software should then allow the option, quickly from the Mediator software’s interface, to send the same value across to all synths that are capable of receiving them, simultaneously.

7.3.5 Conclusions from interviews with VJs and live visuals performers

In summary from the interviews conducted, the conclusion can be drawn that all four interviewed live visuals performers find the paradigms developed here relevant to their respective practice. Furthermore, they could all envision involving part or all of the developed concepts in their future practice.

However, some commented that the system requires a longer setup time than the current systems they use. They would therefore not consider using it in all performance contexts, as for some brief, low profile performances, the effort of setting the system up and tearing it down, may outweigh the benefits of its use. They would however certainly use it on higher profile, or longer duration performances. It is worth noting here that although the system indeed has quite a long setup time in comparison to established practice visuals systems, of half an hour to an hour and a half depending on the number and type of musical instruments involved, the setup time is analogous to that of several music systems. Setting up an acoustic drum set, or a rig of electronic keyboards, can easily take a similar amount of time to setting up the Trinity system.

All subjects, furthermore, commented on usability improvements that they perceived were desirable for the Mediator software in particular. Reassuringly, the vast majority of these suggestions are very straight-forwardly implemented, and would in combination significantly improve the usability of the system.

Only one suggestion, which all four mentioned, of having a preview of the visual output of each visual synth separately in the Mediator interface, would be difficult to implement. The usefulness of this feature is very clear, but it is technically very complicated to successfully realize, largely due to limitations of presently existing graphics programming frameworks and graphics hardware accelerators. It is perhaps further arguable whether this comment could strictly be applied to the Mediator application, since it is more relevant to how the Mediator integrates with Mother.

With regards to Trinity's other two applications, the live input processor and Mother, the interviewed subjects did not comment negatively about any particular aspect of their design or functionality at all, and confirmed that they found the concepts they embody to be sound.

What is encouraging from all the comments is that none of the four interviewed subjects identifies any fundamental conceptual limitations in the software. On the contrary, all interviewed subjects agreed that this design is sound, and that with straight-forwardly implementable usability improvements, they would consider employing either all or parts of it in their practice.

Following from the discussions with them, a distinction became clear, which can best be expressed with an analogy to contemporary musical practice. While most current VJing practice currently is technically and conceptually analogous to DJing, alternatively conducting gesture performance, the Trinity system embodies a novel approach, in that it is technically and conceptually analogous to live musical performance, both in relation to the complexity of its setup, and the complexity of the subsequent possible visual output.

In musical performance, there is a choice in the level of complexity that a performance will have, from either a simple one man DJing setup, conducting gesture performance using for example Ableton Live, a smaller four-piece group, up to a full classical orchestra. In live visuals/audiovisual performance however, before the Trinity system, there was no possibility to scale the performance so as to involve higher degrees of complexity if desirable. Instead, established practice for live visuals performance has in terms of the complexity involved, both in the logistics involved, and in the resulting output, always been constrained to levels analogous to the complexity of DJing and conducting gesture performance. Only with the Trinity system, is it now possible to take advantage of the benefits in preparing for and setting up a complex collaboration between several performers, like in a band or an orchestra, if so desired.

7.4 Further observed impact of present work

As discussed by Ippolito et al. (2009), it is informative to briefly observe what real-world reactions and feedback the work here has generated.

Two public performances have taken place alongside live musicians, and further two solo audiovisual performances, in which pre-recorded music by the author was used. All performances went without significant technical problems, and were well received.

The Mother software, which forms part of the Trinity system, has been released as open source, and a measure of its relative success is that it has attracted contributions by three independent artists/programmers. Further users of the software have contributed with other means than submitting program code, such as contributing with graphic design, or with writing tutorials for how the software is used. It has, to the authors' knowledge, been used in the context of live performance by another three artists, two of which are also contributors to the project. Finally, it has been accepted to be featured on the official Processing homepage ("Libraries \ Processing.org," n.d.), and prior to that, been independently discovered and featured on several websites prominent in the VJ/Visualist community ("Create Digital Motion » Processing for VJing and Performance, with Mother," 2008), ("Create Digital Motion » Mother 0.2, App for Live Performance with Processing, Updated," 2008), ("VdmoKstati.com : Mother," n.d.). At the time of writing, the program has been downloaded approximately 2100 times since its release on Aug. 2008 (http://code.google.com/p/processing-mother/downloads/list), a number which should of course be viewed in light of the limited number of artists that perform live real-time computer graphics internationally.

7.5 Technical evaluation

The Mother application has gone through several development iterations, during which its stability has been continuously improved. At the present state of development, crashes of the Mother application virtually never occur. Mother has furthermore been implemented to incur minimal computational overhead, both to the CPU and the GPU, in comparison to if all the visual synths had simply been programmed together as one large Processing application. Comparisons of the processor usage and frame-rate of visual synthesizers running as standalone programs, versus running within Mother, have shown no observable difference.

The live input processing software has too reached a state of development where it virtually never crashes/freezes. Although the process of performing live signal processing on eight audio channels simultaneously does leave a very noticeable mark on the CPU usage, it never reaches a stage where it becomes too much for a single CPU core to handle, on the Dell Precision M4400 notebook workstation used. This, when using ultra-low latency ASIO audio drivers for the Focusrite Saffire 10 I/O Pro audio interface, giving a latency of roughly 10ms. The MIDI/OSC processing has a negligible computational overhead and latency.

The Mediator application too has after several development iterations reached a level of stability where crashes/freezes virtually never occur. It incurs a negligible amount of processing overhead. The network bandwidth usage of the OSC UDP packed data communication was measured using the FinitySoft Network Monitor v1.3 software (http://www.finitysoft.com/). The use of UDP packets is common practice in the context of OSC communication, as it incurs lower network latency, and more consistent timing, than when using TCP packets. The Mediator application was being used for the audiovisual piece with the most complex mappings, employing both MIDI and audio DSP derived control data. Even then, the network usage never exceeded 200 KB/s, well below the capacity of even the most modest modern Ethernet network hardware.

To demonstrate that the system functions technically as intended, it has been used to create a series of audiovisual pieces, four using music by the author, and three using music by the signed Bristol band “It’s a Lunken”, who provided multi-track audio recordings for the purpose. To view the material created using the author’s own music, please refer to the DVD accompanying this thesis. Part of the above created material is further used in two solo performances, serving to technically test the system in a live setting, before actual bands of musicians are engaged. The performances took place on the recurrent London events of Immersion (now no longer on-going), and Electrovision (“Electrovision 18th of July 2009 - Roxy Bar and

Screen," 2009).

Please note however that mutable mapping is not demonstrated in any of the videos on the DVD. In any context where the Trinity software was used live, mutable mapping was employed throughout. The reason it is not demonstrated in the videos is purely technical. In order to create and store the video output, it is necessary for the Trinity software to run in a mode that is much slower than real-time, as the process is too computationally intensive. It is therefore subsequently also necessary to record the OSC data generated by the Mediator software in real-time, and then play it back slower than real-time so as to record video, thus requiring a separate OSC sequencer application. No such software exists at the time of writing however, and while an application with this purpose was prototyped as part of this work, it proved to be too complicated to implement with satisfactory performance, within the timeframe of this project. It was necessarily prioritized therefore, to concentrate on the primary purpose of this thesis, of creating a system for live performance, and thus to not fully developing the software necessary for also generating non real-time material using the Trinity system.

7.6 Evaluation discussion

There is already strong argument for why the audiovisual output of a nature similar to that possible with the Trinity system, is a more interesting experience to audiences, than what is possible in real-time with currently established practice. This argumentation was reviewed in section 2.1.3, and is partly also reviewed by Grierson (2005) and Moody (2009). What has been lacking however is a concrete design for a system with which such material may be readily produced and performed in real-time, as opposed to being created through an offline process.

Such a design has here been created, as well as implemented, in the form of the Trinity system. The design and created system as a whole can be shown to fulfil the technical requirements for sufficiently addressing the identified deficiencies in currently established practice.

To address the evaluation points for design research, discussed by Zimmerman et al. (2007):

The **Process** of designing and implementing the system evaluated here has been documented, and is presented in section 5.

Invention has been demonstrated; Clear deficiencies were identified in current practice, followed by evidence-based arguments, of how these deficiencies may be addressed, as detailed in sections 1 & 2.1.3. Based on these findings, new paradigms were developed, and software applications were designed and implemented to support these paradigms. All developed applications have been shown to fulfil the technical requirements for sufficiently addressing identified deficiencies. The ideas they engender have also been acknowledged to be novel, and been positively received, by interviewed practitioners in the field.

Relevance has arguably been demonstrated through the successful use of the developed system in a number of public performances, by the author alone and in collaboration with independent musicians, as opposed to its deployment only in a laboratory setting. Through the process of interviewing practitioners employing current practice, feedback has been gathered for the relevance of the ideas and implementations created with this work. All interviewed current practice subjects confirmed that they did find the work highly relevant to their practice. Furthermore relevance has been demonstrated also by the attention that the Mother software application has received, as detailed in section 7.3.

Further testament both to the novelty and to the relevance of the invention achieved, is the fact that the articles detailing the software and concepts created, have been accepted through the academic peer-review process of their respective conferences and journals (Bergstrom & Lotto, 2009), (Bergstrom, Steed, & Lotto, 2009) and (Bergstrom & Lotto, 2008).

Finally, the **Extensibility** criterion is addressed partly by releasing one of the software applications developed as free open source using the GPL3 license (Mother), and partly by fully disclosing the intent and functionality of Mother and the remaining two applications, live input processor and Mediator, in all related academic publications, as well as in this thesis.

The added viewpoint of an ethnographic study was also employed towards observing the experience of musicians when participating in rehearsals employing the system presented here. The full text of the field report is presented in appendix B, in accordance to the discussion by Dourish (2006), with a summary and conclusions from the experience being presented in section 7.2.

The paradigms embodied in the Trinity system are to a great extent entirely novel, with little precedent in previous work. It is therefore at this stage of development not possible to experimentally achieve objective verification to the extent the paradigms and applications are preferred by artists and audience, over currently established practice. As is stated by Greenberg & Buxton (2008), the true usefulness and success of the developed paradigms and designs can only be gauged after many years have passed, during which these are continuously put to use, so that a culture of use will have had a chance to be established around them.

Only one step can be taken at this early stage of development, towards the experimental evaluation of the developed paradigms and designs: to demonstrate through real world example that it is possible to use them in the manner described, and the context intended. As is evident from the presented material in this thesis, the software has on several occasions publicly, and in a real-world context, been demonstrated to successfully fulfil its intended purpose.

Furthermore, four out of seven musicians during the rehearsal sessions reported that at times, they did indeed experience the subjective sensation of being to some extent controlling the visuals, instead of the visuals simply reacting to their playing. This provides an initial confirmation to the hypothesis, that from the increased control data bandwidth, they are able to better transmit the intent of their musical gestures. They are thus to a greater extent able to usefully take advantage of the advanced embodied knowledge they have of their instruments, finally achieving an increase in expressivity over established practice. To more precisely qualify the extent of the achieved advantage, many more hours of experimentation are needed, over a much longer period of time, and with many more musicians. Such extensive experiments however are very time consuming to organize and conduct, and as such are deemed to be beyond the scope of the present thesis.

Having now seen in detail the results from the validation process for the novel ideas introduced in this thesis, as embodied in the Trinity system, we are in a position of in the subsequent chapter identifying what further effort may be carried out, to continue from where this thesis leaves off.

8 Further work

8.1 Experiments

Most importantly, continuous work with using the now completed Trinity system with as many live groups of musicians as possible, over a period of several years, is necessary. This is to slowly establish a culture for the systems use, the only way in which the successfulness of the paradigms, designs and software applications developed here, can be conclusively determined in an experimental manner (S. Greenberg & Buxton, 2008).

We have seen here, that musicians feel that they are, using the Trinity system, subjectively in control of the visuals through the interface of their musical instruments, to a greater extent than in previous practice. What has not been however gauged with present work is whether that occurs to high enough an extent, for them to have successfully re-mapped the enactive knowledge they have of their instruments, to the task of performing live procedural visuals. An indication has been provided, which, if followed up by further experimental sessions, of longer duration and involving greater numbers of subjects, could lead to more deterministic evidence of such re-mapping taking place.

Furthermore, the mutable mapping paradigm shows a very significant potential of being useful also in the context of performing live music, as well as other live artistic performance, besides the audiovisual context in which it was employed for in this work. It would be highly interesting, to employ the Mediator software in these contexts, and observe how musicians / artists choose to use it. Both to gauge if it is indeed useful in the manner envisioned, as well as to observe whether new uses of the mutable mapping paradigm will emerge, that had not been predicted.

Miranda & Wanderley (2006) present a discussion on the importance of multi-sensory feedback in the context of performing music using a DMI. In this discussion it is argued that a performer is able to use a DMI more successfully if he receives feedback from his performance on many modalities, as is often the case with traditional musical instruments. Potentially, the Trinity system could be used in this context, to gauge whether musicians indeed experience there being an advantage with having the visual feedback possible, in comparison to a scenario where no such feedback is provided.

8.2 Development

The Mother software, while successful, has a number of limitations that would be beneficial to address, so as to improve its usability. Firstly, although it is currently being used by other artists than its creator, these are too few for a community of sharing visual synthesizers to be established. An online repository of such visual synthesizers would be a development that would significantly encourage future artists to use the software.

Secondly, because Processing is implemented using java, Mother is quite computationally intensive, compared to the speed of execution that it could potentially achieve had it been implemented in a lower level programming language such as C++. It is therefore worth considering re-implementing Mother for the OpenFrameworks environment: a very recent development, which seeks to provide a C++ based multimedia programming environment for artists, inspired by the Processing paradigm.

A third identified limitation with the paradigm employed with the Mother application, is that of stability. This is a consequence of the fact that OpenGL is a state machine, which is very sensitive to failing, if it is entered into an invalid state by any of the visual synthesizers running within Mother. Thus, the whole Mother application can potentially freeze, due to a single visual synthesizer having a bug in its OpenGL code. One alternative would be to execute each visual synthesizer in its own OpenGL rendering context, with their final output being mixed together, rather than having them all execute in the same context, as is presently the case. This however would potentially prove to be prohibitively costly in terms of the computational overhead incurred. The benefit would be that a performance need not fail simply because a single visual synthesizer program had a bug.

Finally, the Mother paradigm, regardless of implementation, would greatly improve if it was augmented to also support the real-time manipulation of each visual synth's output using 2d video effects. There is already an established standard for such effects in the VJing community, in *FreeFrame* (http://freeframe.sourceforge.net/), a plug-in format in which numerous effects have already been implemented. Integrating FreeFrame effects into the Mother program's OpenGL rendering pipeline would instantly allow the use of all these existing effects in performances that use Mother.

As visual synthesizer software is developing very rapidly, it is not unlikely that an application will appear, implementing a feature set similar to that of Mother, with which Mother may be replaced. A possible suitable candidate is the new, significantly improved version of Derivative Inc's Touch designer 077 software ("Touch Designer - Derivative," n.d.), currently in public beta testing.

The Mediator software, too, would greatly benefit from further development, to address the usability issues identified in its evaluation. Several small and easily implemented suggestions have been made by the interviewed visuals performers, as detailed in chapter 7.3, which in combination would have a significantly positive effect on the software's usability.

Although the live input processor application developed for this project has fulfilled its purpose successfully, its usability could be improved by re-implementing it. The current version has been fully implemented in Max/MSP, a language with several limitations with regards to the user interface design paradigms it supports. It is planned to fuse the functionality of the live input processor into the Mediator, so that greater immediacy and controllability over the process of connecting to external non-OSC capable devices can be achieved, as a result of the tighter integration of the related parameters between the mapping matrix, and the external data source processing.

8.3 Content

At present, there is a great lack of literature and previous work of procedural graphics algorithms that would work well in the context of live visual music. There is a significant history of visual music and abstract animation to draw inspiration from, but it is not always possible to adapt the ideas from that material to program code that can execute in real-time. Similarly, there is significant literature on procedural computer graphics, but predominantly this literature concentrates on the simulation of real world phenomena and realistic imagery, rather than directly targeting an aesthetic suitable in the visual music / abstract animation context. This literature is to make the matters worse, almost exclusively targeted at expert programmers, and cannot easily be used by an artist-programmer with no computer science training, as are the average Processing users.

Although there is previous practice of procedural, real-time visual music and abstract animation to draw inspiration from, very often the source code and algorithms behind it are not shared by the creators, thus increasing the difficulty with which other artists can draw from previous practice.

This difficulty would be alleviated if an extensive systematic review of real-time procedural graphics algorithms applicable to visual music / abstract animation was undertaken. Suitably documenting these in an accessible language such as Processing, would provide a convenient go-to resource for all artists to draw from.

9 Conclusion

9.1 Summary of contribution

With the work presented in this report, the deficiencies in the current existing tools available to artists for creating and performing live visual music / audiovisual art, have been identified and to a significant extent addressed. These deficiencies are:

I. The mappings between music and visuals are highly constrained because they are limited in complexity, and remain static over time, thus necessitating that the correlation between visuals and music always remains limited, when in fact there is much evidence that increased correlation results in a stronger experience.

II. Virtually no user interface exists, that allows controlling the performance of visual music/audiovisual art in real-time, with a level of expressivity comparable to that attainable in live musical performance.

III. The process of preparing or improvising live procedural visual music/audiovisual performances is overly complicated, compared to live musical performance.

IV. Collaborative performance, a very common practice in most live performance art, is mostly absent in live visual music/audiovisual performance, largely not due to artistic choice, but to technical limitations.

This thesis addresses the above limitations, building on a transdisciplinary integration of findings drawn from a wide range of research areas, including music, visual and multimedia art and performance, music technology, programming as art, software engineering, cognitive science, psychophysics and finally human computer interaction, which in itself is a strongly interdisciplinary field. Novel paradigms have been developed, and based on these a system has been designed and implemented, which is used to demonstrate and, to the extent that is feasible within the timeframe of present work, assess the paradigms it embodies. The bases for the argument and system design presented are the following four theoretical contributions:

A. It is proposed that musical instruments are used as the primary source of control data for the performance. Through using the full bandwidth of control data that the musical instruments can produce, the musical gestures of the performers are encoded in much higher detail than only beat events, amplitude and tempo from a stereo audio signal, as is common in established practice. Musicians are, from this increase in control data bandwidth, expected to be to a greater extent able to usefully re-map the advanced embodied knowledge they have of their instruments, to the performance of live visual music/audiovisual art, and therefore achieve an increase in expressivity. Using the full bandwidth of control data, further allows for achieving an increased mapping complexity. Finally, because musicians are already accustomed to performing collaboratively using their instrument, it is expected that their experience can directly translate to also allowing collaborative performance in the context of live visual music/audiovisual performance.

B. The inception of the conduct of Mutable Mapping: gradually creating, destroying and altering mappings between the two parameter spaces of input control data from the musical gestures and output control data of the visual synthesis software, during the course of a performance. Mutable mapping is expected to allow for complex, dynamic, and improvised mappings in live visual music / audiovisual art, and as a direct consequence also increasing expressivity.

C. From the previous contribution in conjunction, follows the inception of the art form of Soma. In Soma, correlated auditory, visual and proprioceptive stimulus is used to form a combined narrative. Soma builds on research findings that both performers and audiences are more engaged in a performance, when performers exhibit advanced motor knowledge, and when congruent percepts across modalities temporally coincide.

D. A reduction in the difficulty of preparing towards a performance of live procedural computer graphics is attempted, through re-adapting a programming language intended for artists, Processing, to instead behave as a plug-in Application Programming Interface (API). The language's accessibility is thus expected to further increase, through facilitating the encapsulation of all programming complexity in flexible modules, which are easily exchanged and managed. It is expected that artists may in this manner dynamically layer the output of such modules during the course of a performance to create a narrative, thus greatly simplifying the preparation and improvisation of a performance.

All the above ideas have been implemented in the Trinity system. The development of the system has been documented, as have the ideas behind its design, so that the process of creating it can be of use to future developments.

The conclusive experimental evaluation of the work has after a thorough literature review deemed to be unfeasible within the timeframe of the project. Informed from an extensive literature review of the relevant evaluation methodology in the areas of human computer interaction and digital musical instrument design, the evaluation of the work from a design research perspective was instead conducted (Zimmerman et al., 2007). The added viewpoint of an ethnographic field study was also employed towards the evaluation (Dourish, 2006).

As the evaluation methodology for research of a similar nature to present work, still remains a largely unsettled discussion, a further theoretical contribution of this work lies in the review and selection of evaluation methodology, and through the subsequent case study embodied in the evaluation work undertaken.

The practical contribution detailed in this thesis, entails the design and implementation of a system that embodies the above four theoretical contributions. Towards the evaluation of the system and the ideas it engenders, the system was then used in the following contexts:

I. The creation of six audiovisual pieces using pre-recorded musical data, which served to demonstrate that the system functions technically, and to demonstrate the output of the system to musicians for the purpose of engaging them in collaborations.

II. Two solo performances of the above material, serving to further test the system as well as the developed visual material in a live setting, before actual bands of musicians were engaged.

III. One application from the developed system was released as free open source software, to engage the live visuals community with the ideas it engenders. The application has this far seen comparative success, having been covered on the most prominent websites in the community, inviting contributions both to its code and to its documentation, and having seen use in live performance by artists across the globe, on three known occasions.

IV. Three rehearsals and two live performances with a total of seven musicians were performed, to observe them and interview them about their experience. From the rehearsals, feedback was gathered on the musicians' overall experience of Soma performance. The positive benefits of using a high bandwidth of control data from their instruments were to an extent observed. It is found that four out of seven of the musicians felt that they were to a greater extent controlling the performance, compared to a simulated previous practice scenario, thus suggesting that the increase in control data bandwidth may allow them to take a greater advantage of the embodied knowledge they have of their instruments. Four of the musicians further stated the combined experience influences the aesthetic content of the music they performed.

V. Live visuals performers familiar with established practice were engaged, to elicit feedback on the concepts embodied in the system, as well as their implementation. It was assessed what benefit they find in the concept of allowing multiple Processing programs to be dynamically layered during the course of a performance. They were additionally queried about their feedback on what role the concept of mutable mapping may play in their current and future artistic practice.

VI. An on-line survey was conducted to experimentally gauge whether increased temporal correlation between audiovisual stimuli in a pre-recorded visual music piece, results in a more aesthetically engaging experience for audiences. The results from 30 respondents were positive, but did not achieve statistical significance, and were thus not generalizable. The survey also served a secondary purpose, of gauging the extent to which viewers of a piece created with the system, find it engaging/enjoyable. From the survey results, it can be concluded that audiences did find the material to be appealing.

9.2 Closing remarks

Musicians today turn their attention to performing live to a much higher degree than has previously been common, due to very recent developments in the commercial recording industry. Amid speculation that it will no longer be possible for musicians to sustain themselves financially by only selling records - a market that is rapidly decreasing in revenue, musicians have already began to concentrate on live performances as their primary source of income. With the increased competition for audiences' attention, it is likely that a similar increase will take place in the demand for means with which added value can be created for audiences.

This climate is likely to harbour an ever-increasing interest in the use of systems such as Trinity in live performance, signalling the beginning of a significant increase in activity in live visual music and audiovisual art.

Similarly, the very complicated system that is a visual synthesizer has very recently become far more accessible both to afford owning, and to learn using. Although there still is no software available that can currently have an impact analogous to that observed when 3D animation software began appearing for use on desktop workstations, a similar development does not seem too far in the future. The emergence of such software will likely further propel the use of concepts such as those developed here, in live artistic performance.

Appendix A: Online survey questionnaire & responses

Only data from completed responses is included. Note that because the order of the two animations was randomized, it is not certain which animation verbal responses refer to as “first” and “second”, as there is no data in this report on which order they were viewed in.

Intro

This survey consists of two variations of the same animation, each followed by a corresponding questionnaire, and finally a single questionnaire with follow-up questions. It should take a maximum of 10-15 minutes to complete.

Fill in each of the two first questionnaires only after having fully watched the animation corresponding to it.

Note that to avoid any bias, both pages appear identical, but rest assured the animations played in each are different from each other.

Before playing each video, make sure:

You have enabled audio playback for your computer.

You use either headphones or standalone speakers; please avoid using your laptops built in speakers.

The videos' HD playback option is enabled.

The video has buffered enough for playback to be uninterrupted.

The video is maximized to fill the screen.

Your computer is able to play the video without freezing/skipping.

If you are unable to fulfil either of these criteria, please postpone participating in the survey until you can fulfil all.

1. The following items refer to your general level of enjoyment and understanding of the piece you have just seen. To the right of each question is a rating scale for recording your response. Answer each question by circling the number on the 7-point scale that most clearly indicates your opinion. There are no right or wrong responses. Please answer each question as accurately and honestly as possible.

Item	1	2	3	4	5	6	7	Average	Total
1. I enjoyed watching this piece.	3.1% 1	6.3% 2	12.5% 4	3.1% 1	28.1% 9	31.3% 10	15.6% 5	5.0	32
2. My attention was focused on the performance.	15.6% 5	3.1% 1	15.6% 5	12.5% 4	18.8% 6	28.1% 9	6.3% 2	4.3	32
3. I did not find this piece intellectually	12.5%	31.3%	15.6%	9.4%	9.4%	15.6%	6.3%	3.4	32

stimulating.	4	10	5	3	3	5	2		
4. The piece was entertaining.	6.3% 2	6.3% 2	9.4% 3	6.3% 2	31.3% 10	31.3% 10	9.4% 3	4.8	32
5. I did not understand the piece.	21.9% 7	18.8% 6	9.4% 3	34.4% 11	3.1% 1	3.1% 1	9.4% 3	3.3	32
6. I was bored by the piece.	18.8% 6	21.9% 7	9.4% 3	6.3% 2	9.4% 3	21.9% 7	12.5% 4	3.8	32
7. I was intellectually engaged by the piece.	18.8% 6	3.1% 1	25.0% 8	12.5% 4	18.8% 6	9.4% 3	12.5% 4	3.9	32
8. I felt confused by the piece.	37.5% 12	34.4% 11	6.3% 2	12.5% 4	3.1% 1	6.3% 2		2.3	32
9. I found it difficult to concentrate on the piece.	28.1% 9	18.8% 6	9.4% 3	6.3% 2	12.5% 4	18.8% 6	6.3% 2	3.4	32
Average %	18.1%	16.0%	12.5%	11.5%	14.9%	18.4%	8.7%	3.8	288.0

2. The following items refer to aspects of the performance that you may or may not have enjoyed. Indicate how enjoyable the following were to you by circling the number that best represents how you feel. A rating of 1 refers to

Item	0	1	2	3	4	5	6	7	Average	Total
1. Music		15.6% 5	3.1% 1	3.1% 1	15.6% 5	28.1% 9	15.6% 5	18.8% 6	4.6	32
2. Colors		6.3% 2	3.1% 1	6.3% 2	40.6% 13	31.3% 10	6.3% 2	6.3% 2	4.3	32
3. Shapes		6.3% 2	3.1% 1		43.8% 14	15.6% 5	18.8% 6	12.5% 4	4.7	32
4. Animation		9.4% 3	6.3% 2	12.5% 4	12.5% 4	12.5% 4	28.1% 9	18.8% 6	4.7	32
5. All the above as a whole	3.1% 1	15.6% 5	6.3% 2	3.1% 1	21.9% 7	25.0% 8	15.6% 5	9.4% 3	4.1	32
Average %	0.6%	10.6%	4.4%	5.0%	26.9%	22.5%	16.9%	13.1%	4.5	160.0

3. The following items ask you to indicate the extent to which the piece made you feel a particular emotion. After deciding whether the piece made you feel that particular emotion, circle the number on the 7-point scale to indicate your response and indicate why you felt that emotion.

Item	0	1	2	3	4	5	6	Average	Total
1. Happy	37.5% 12		6.3% 2	25.0% 8	9.4% 3	18.8% 6	3.1% 1	2.4	32
2. Sad	75.0% 24	12.5% 4	6.3% 2	3.1% 1	3.1% 1			0.5	32
3. Angry	90.6% 29	3.1% 1	3.1% 1	3.1% 1				0.2	32
4. Energized	25.0% 8	9.4% 3	9.4% 3	18.8% 6	15.6% 5	18.8% 6	3.1% 1	2.6	32
5. Reflective	31.3%	6.3%	6.3%	25.0%	18.8%	9.4%	3.1%	2.3	32

	10	2	2	8	6	3	1		
6. Intense	43.8% 14	9.4% 3	6.3% 2	3.1% 1	12.5% 4	21.9% 7	3.1% 1	2.1	32
7. Fearful	84.4% 27	9.4% 3		3.1% 1		3.1% 1		0.3	32
8. Surprised	65.6% 21	9.4% 3	3.1% 1	9.4% 3	12.5% 4			0.9	32
9. Disgusted	93.8% 30		3.1% 1	3.1% 1				0.2	32
10. Confused	75.0% 24	18.8% 6		3.1% 1	3.1% 1			0.4	32
11. Bored	34.4% 11	21.9% 7	3.1% 1	15.6% 5	12.5% 4	3.1% 1	9.4% 3	2.0	32
12. Excited	43.8% 14	12.5% 4	6.3% 2	18.8% 6	3.1% 1	12.5% 4	3.1% 1	1.8	32
13. Relaxed	34.4% 11	12.5% 4	12.5% 4	18.8% 6	12.5% 4	9.4% 3		1.9	32
14. Anxious	62.5% 20	15.6% 5	9.4% 3	6.3% 2	6.3% 2			0.8	32
Average %	56.9%	10.0%	5.4%	11.2%	7.8%	6.9%	1.8%	1.3	448.0

4. The following items refer to any physical sensations you may have experienced whilst watching the piece. After deciding whether the piece made you feel a particular sensation, please indicate your response by ticking the appropriate column.

Item	Yes	No	Total
1. Heart-racing	21.9% 7	78.1% 25	32
2. Lump in throat	3.1% 1	96.9% 31	32
3. Shivers down spine	6.3% 2	93.8% 30	32
4. Tears		100.0% 32	32
5. Goose-bumps	9.4% 3	90.6% 29	32
6. Laughter	3.1% 1	96.9% 31	32
7. Blushing	6.3% 2	93.8% 30	32
8. Yawning	12.5% 4	87.5% 28	32
9. Sweating		100.0% 32	32
Average %	6.9%	93.1%	288.0

5. The following items refer to your general level of enjoyment and understanding of the piece you have just seen. To the right of each question is a rating scale for recording your response. Answer each question by circling the number on the 7-point scale that most clearly indicates your opinion. There are no right or wrong responses. Please answer each question as accurately and honestly as possible.

Item	1	2	3	4	5	6	7	Average	Total
1. I enjoyed watching this piece.	12.5% 4	9.4% 3	6.3% 2	18.8% 6	21.9% 7	25.0% 8	6.3% 2	4.3	32
2. My attention was focused on the performance.	18.8% 6	6.3% 2	12.5% 4	9.4% 3	18.8% 6	28.1% 9	6.3% 2	4.1	32
3. I did not find this piece intellectually stimulating.	6.3% 2	18.8% 6	12.5% 4	12.5% 4	18.8% 6	6.3% 2	25.0% 8	4.4	32
4. The piece was entertaining.	18.8% 6	6.3% 2	9.4% 3	12.5% 4	21.9% 7	28.1% 9	3.1% 1	4.1	32
5. I did not understand the piece.	18.8% 6	18.8% 6	6.3% 2	40.6% 13	3.1% 1	6.3% 2	6.3% 2	3.3	32
6. I was bored by the piece.	12.5% 4	18.8% 6	6.3% 2	12.5% 4	12.5% 4	18.8% 6	18.8% 6	4.3	32
7. I was intellectually engaged by the piece.	25.0% 8	9.4% 3	15.6% 5	12.5% 4	18.8% 6	12.5% 4	6.3% 2	3.5	32
8. I felt confused by the piece.	31.3% 10	25.0% 8	9.4% 3	15.6% 5	9.4% 3	6.3% 2	3.1% 1	2.8	32
9. I found it difficult to concentrate on the piece.	18.8% 6	18.8% 6	6.3% 2	9.4% 3	18.8% 6	18.8% 6	9.4% 3	3.8	32
Average %	18.1%	14.6%	9.4%	16.0%	16.0%	16.7%	9.4%	3.8	288.0

6. The following items refer to aspects of the performance that you may or may not have enjoyed. Indicate how enjoyable the following were to you by circling the number that best represents how you feel. A rating of 1 refers to

Item	0	1	2	3	4	5	6	7	Average	Total
1. Music	3.1% 1	9.4% 3	6.3% 2	6.3% 2	31.3% 10	15.6% 5	15.6% 5	12.5% 4	4.3	32
2. Colors		9.4% 3		12.5% 4	28.1% 9	40.6% 13	6.3% 2	3.1% 1	4.2	32
3. Shapes		6.3% 2	9.4% 3	9.4% 3	28.1% 9	18.8% 6	18.8% 6	9.4% 3	4.4	32
4. Animation		18.8% 6		12.5% 4	18.8% 6	15.6% 5	28.1% 9	6.3% 2	4.2	32
5. All the above as a whole	3.1% 1	15.6% 5	12.5% 4	9.4% 3	21.9% 7	15.6% 5	18.8% 6	3.1% 1	3.7	32
Average %	1.3%	11.9%	5.6%	10.0%	25.6%	21.3%	17.5%	6.9%	4.2	160.0

7. The following items ask you to indicate the extent to which the piece made you feel a particular emotion. After deciding whether the piece made you feel that particular emotion, circle the number on the 7-point scale to indicate your response and indicate why you felt that emotion.

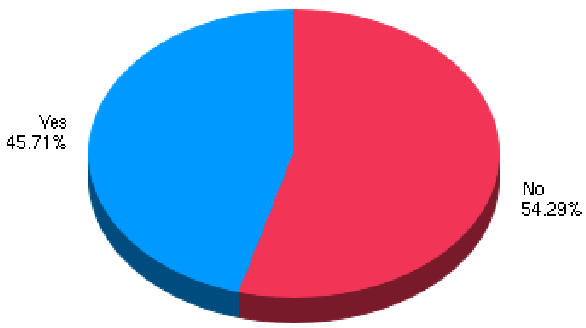
Item	0	1	2	3	4	5	6	Average	Total
1. Happy	53.1% 17		6.3% 2	15.6% 5	18.8% 6	3.1% 1	3.1% 1	1.7	32
2. Sad	78.1% 25	3.1% 1	12.5% 4	6.3% 2				0.5	32
3. Angry	78.1% 25	3.1% 1	6.3% 2	9.4% 3	3.1% 1			0.6	32
4. Energized	37.5% 12	9.4% 3	12.5% 4	15.6% 5	15.6% 5	3.1% 1	6.3% 2	2.0	32
5. Reflective	40.6% 13	9.4% 3	9.4% 3	15.6% 5	15.6% 5	6.3% 2	3.1% 1	1.9	32
6. Intense	46.9% 15	6.3% 2	15.6% 5	12.5% 4	12.5% 4	6.3% 2		1.6	32
7. Fearful	71.9% 23	12.5% 4	3.1% 1	3.1% 1	6.3% 2		3.1% 1	0.7	32
8. Surprised	56.3% 18	12.5% 4	6.3% 2	12.5% 4	6.3% 2	3.1% 1	3.1% 1	1.2	32
9. Disgusted	93.8% 30		3.1% 1	3.1% 1				0.2	32
10. Confused	68.8% 22	12.5% 4	6.3% 2	9.4% 3	3.1% 1			0.7	32
11. Bored	31.3% 10	18.8% 6	9.4% 3	18.8% 6	6.3% 2	6.3% 2	9.4% 3	2.1	32
12. Excited	62.5% 20	3.1% 1		15.6% 5	6.3% 2	9.4% 3	3.1% 1	1.4	32
13. Relaxed	37.5% 12	25.0% 8	12.5% 4	15.6% 5		6.3% 2	3.1% 1	1.5	32
14. Anxious	59.4% 19	18.8% 6	9.4% 3	6.3% 2		6.3% 2		0.9	32
Average %	58.3%	9.6%	8.0%	11.4%	6.7%	3.6%	2.5%	1.2	448.0

8. The following items refer to any physical sensations you may have experienced whilst watching the piece. After deciding whether the piece made you feel a particular sensation, please indicate your response by ticking the appropriate column.

Item	Yes	No	Total
1. Heart-racing	18.8% 6	81.3% 26	32
2. Lump in throat		100.0% 32	32
3. Shivers down spine	6.3% 2	93.8% 30	32

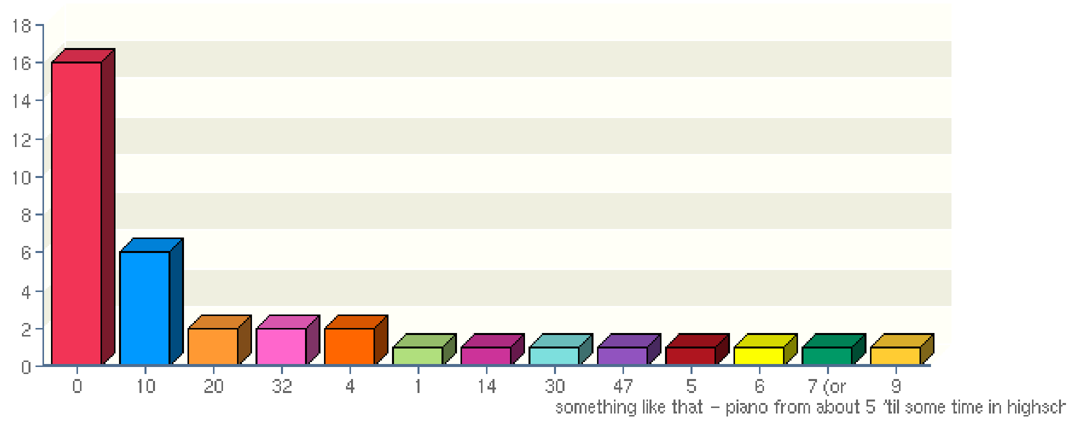
4. Tears	3.1% 1	96.9% 31	32
5. Goose-bumps		100.0% 32	32
6. Laughter		100.0% 32	32
7. Blushing	6.3% 2	93.8% 30	32
8. Yawning	9.4% 3	90.6% 29	32
9. Sweating	9.4% 3	90.6% 29	32
Average %	5.9%	94.1%	288.0

9. Have you previously seen a variation of presented animations?



Item	Count	Percent %
No	18	58.06%
Yes	13	41.94%

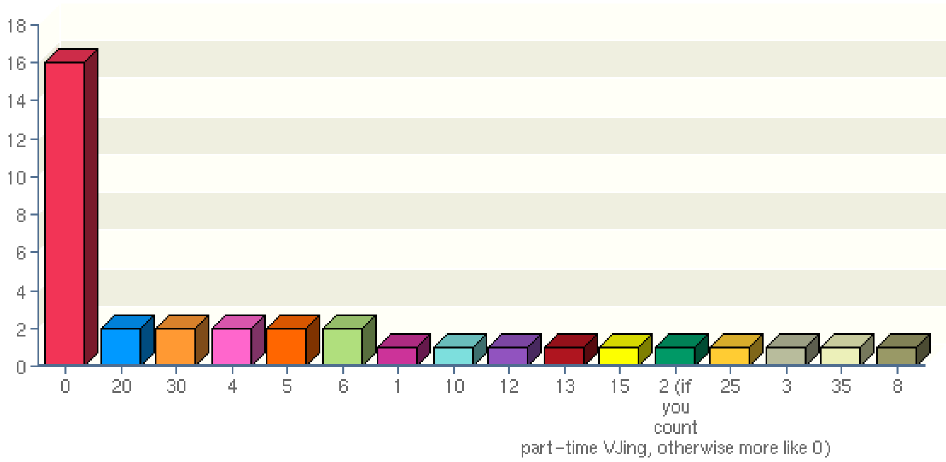
10. For how many years have you played music? (non-musicians enter 0)



Item	Count	Percent %
0	13	41.94%
10	6	19.35%
20	2	6.45%
32	2	6.45%
4	2	6.45%
1	1	3.23%
14	1	3.23%
30	1	3.23%
47	1	3.23%
5	1	3.23%
7 (or something like that - piano from about 5 'til some time in highschool)	1	3.23%

Average: 8.68

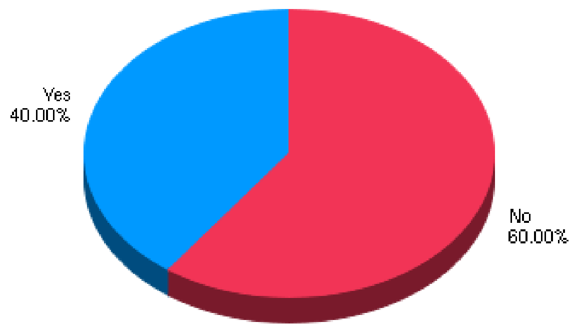
11. For how many years have you practiced visual arts? (non-visual artists enter 0)



Item	Count	Percent %
0	13	41.94%
20	2	6.45%
30	2	6.45%
5	2	6.45%
6	2	6.45%
1	1	3.23%
10	1	3.23%
12	1	3.23%
13	1	3.23%
15	1	3.23%
2 (if you count part-time VJing, otherwise more like 0)	1	3.23%
25	1	3.23%
35	1	3.23%
4	1	3.23%
8	1	3.23%

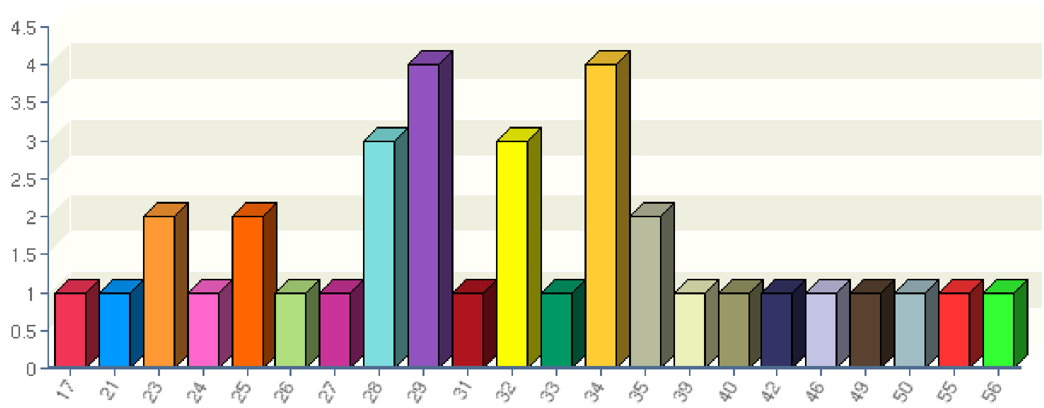
Average: 7.90

12. Is English your first language (tick one)?



Item	Count	Percent %
No	18	58.06%
Yes	13	41.94%

13. What is your age?

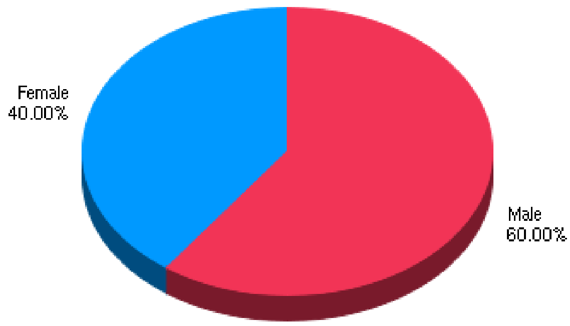


Item	Count	Percent %
29	4	12.90%
34	4	12.90%
25	2	6.45%
28	2	6.45%
32	2	6.45%
35	2	6.45%
17	1	3.23%
21	1	3.23%
23	1	3.23%

24	1	3.23%
26	1	3.23%
27	1	3.23%
33	1	3.23%
39	1	3.23%
40	1	3.23%
42	1	3.23%
46	1	3.23%
49	1	3.23%
50	1	3.23%
55	1	3.23%
56	1	3.23%

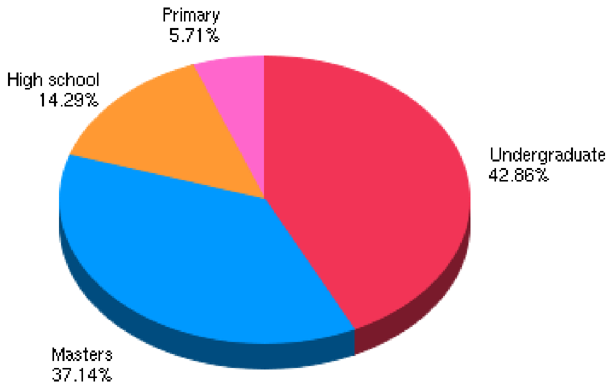
Average: 33.55

14. What is your sex?



Item	Count	Percent %
Male	19	61.29%
Female	12	38.71%

15. What is the highest level of education you have completed (tick one)?



Item	Count	Percent %
Undergraduate	13	41.94%
Masters	12	38.71%
High school	4	12.90%
Primary	2	6.45%

16: Which, in your view, were the differences between the two pieces?

Open Text Responses:
Not much. More sound reactive for the yellow parts in 2nd More grey bands in 2nd
None. Both extremley boring.
mostly the background white strokes, looks much better in second version. Also the orange "snakes" seems to be more dynamic in second one
The first piece was more complex and subtle in the motion, color and tighter sync to the music
The music?
The pace of the "tube tunnels" was more continuous in the second; the second had "racing lines" that then kind of turned into "old time film scratches", the seconds' blue square spirals seemed more nebulous
I am not quite sure
Better sync, visual expression of the musical features and complexity of the second one.
The music piece seems to be the same but the the performances are different, the second one is more vibrant, fast and has more shapes movements.
More intense animation in the second one. Much more intense and immersive as a whole.
I couldn't find the differences,but I felt that the first one is 'happy',then the second one is'sad'.
not too many, I did not watch the second one till the end because it reminded me so much of the first one.
I can see black spots in the second piece.
none

none
Variatons on the central animated elements. Less emphasis on the orange/yellow sparks (very first element on screen), more emphasis on the white/blue spiral in the second version. Different pattern to the white bands (that accompany the electronic click / second element). Slightly more noticeable use of vertical pinstripes in the second version.
2nd relative to first: fewer streamers, less energetic streamers, smaller streamers (without watching again, perhaps only fewer, seeming smaller); descending bands not "clumped" (first pieces clumping seemed a little bland, but less clumping was decidedly less visually appealing); dark spots (from bassline) were more obvious and seemed more frequent; the vertical lines seemed somehow different, but I didn't pay much attention to them to take any particular note of what the difference was, if any; the radial starburst thing had a stronger stream of visible blue squares flying out - I didn't like that so much - and it seemed that it may have been rotated slightly from the first (which seemed to have a more vertical "axis")
The sound appeared more immersive in the second version and the animation appeared more responsive.
the main difference i perceived was the movement of the central shape shifting object became more disjointed in the 2nd video and contaiend less footage of the object(s) forming the double helix shape, i also noticed greyish spots appearing on the black and white bars in the 2nd video. The ending appeared different as well though i can't remember exactly how the first video ended.
the tone: second one appeared darker (visually and musically)
The gray in the back
The second piece had more activity in the graphics - both foreground and background, it was more energetic with more animation of visual elements
I think that it was the same piece.
Second one was sparser. Possibly more synchronized to the music.
The second was much more in sync with the music.
the form of the pieces and colours
number of grey circular lines
The second video showed thicker lines/shapes. The second video has softer & lower tone and lesser intensity/slower rhythm.
Slight differences in animation (slightly different variations of the visualisation patterns)
In 2nd viedo: more interesting shapes, more dynamic in the picture
The biggest difference was in the horizontal lines that move from the top to the bottom of the screen. In the first version they were more syncronized with the music and there were fewer of the. There might have been some smaller differences in the fireworks shapes and timing or in the blue and white swirl shapes timing.

17: Do you have any additional comments or suggestions?

Open Text Responses:
I hope the data is useful! Do you think intellectual response and emotion are quantifiable? I have seen these visualizers previously and think my response was a bit different but not much. I hope you get plenty of non artists responding.
A wider variety of music . . . music is so subjective and so influential over the personal enjoyment of a piece

of animation that offering 4-6 different styles would allow for better accuracy in your survey. For me, this track was just very plain, as were the visuals, so it was difficult for me to feel strongly one way or the other about what I was seeing.

1. the reception of the second one is sort of shaded by the fact that it's fairly similar to what we just saw
2. general note, the blue spirals kind of loose a nice aspiring/climbing/soaring/travelling vibe that the sparks+"tunnel" effect featured, making it feel more like a generic music visualizer - it really hurts the narrative to not have the spiral feel like it's traveling the same way as the sparks.

Introduce more colors to the performances.

MAKE THE SPIRAL ROTATE! On the second video it really felt as if the blue spiral should start rotating. Watching it stand still was like a promise that was never fulfilled or like a girl that got away.

I didn't feel the same while watching the first video. Maybe it is the more "rigid"* feel of the first animation, or the fact that the first spiral is full and seems to radiate from the center, while the second has a void at its center, creating the impression of motion, or that of two half spirals moving together.

* other terms/emotions: "order" (as opposed to chaos), "mathematical"

The second video also reminded me of the classic anime rocket barrage effect, as seen here:

http://www.youtube.com/watch?v=xDw8SHnjXb4

http://www.youtube.com/watch?v=c7GaqoXzZFI

I found the 1:1 corrispondence between some of the elements with the music to be distracting, as if there was too much emphasis on the synesthesia which may be taking place. Disconnecting the visual and auditory elements might lead to a more engaging piece. I found the white horizontal bars (the second element which appears) to be the most interesting part of both pieces.

The starburst thing might be nicer if it rotated slightly - the vertical "axis" is a little disconcerting.
When watching closely, the descending bands causes a little motion-sickness like uneasiness.
I don't especially like the shape/distortion of the descending bars, and the use of grey seems out of place.

The video needs to be miuch more 3 dimentional. At the moment it only moves vertically and whilst thats interesting it would be much more stimulating if the helix moved horizontally or is the camera angle moved horizontally instead to give the illusional of a real 3 dimeniontal environment rather a cheap showreel. All in the all the idea is quite interesting but its a bit too repetetive and it doesnt seem to go anywhere. The shapes aren't that interesting either, the focal point orange things are pretty good and the bars arent bad but the spirals look very tacky.

Ii thought that the spiral shapes would fit better if they were part of the same perspective system as the others.

no

I found the second piece to be more rivetting - really kept my attention because of the animations - it also left me feeling more energetic than the first - it was also intellectually stimulating. The first piece left me feeling more relaxed - even tho the visual elements were so similar it was interesting how the different amount of visual information could affect whether I was relaxed or energetic - however it did not make me feel like getting up and DOING any exercise... Good work

I think that both the music and the visuals are quite bored and commonplace, the video seems even a little bit tacky.

Quack

Interesting. I'm excited to know what's the purpose of this survey....

I question the validity of this survey because the clips are relatively long and uniteresting. I couldn't watch it straight through the second time, so I just scanned it and sampled some parts to see what the differences were between the two versions. The music was a little repetative. But the real problem was in the animation. You really only had enough visual ideas for about 20 or 30 seconds. In my own work I have found that you need a new visual idea at least every 5 seconds. For example the fireworks type things. Even though it is a partical simulation and so it doesn't actually repeat, once you have seen about 5 seconds of it you have a sense of what it's always going to look like. In the interest of full disclosure I create visual music and visual music software, so I have seen a lot of this sort of thing before, and might be a little jaded.

Thank you for taking this survey!

If you are curious to find out more about the project, please visit www.onar3d.com.

This survey is an adaptation of the Audience Response Tool (ART) survey, by Dr Renee Glass, Associate Professor Kate Stevens and Dr Stephen Malloch. MARCS Auditory Laboratories, University of Western Sydney, http://marcs.uws.edu.au.

Appendix B: Reports from rehearsals and performances

The Gathering, St Pancras Old Church

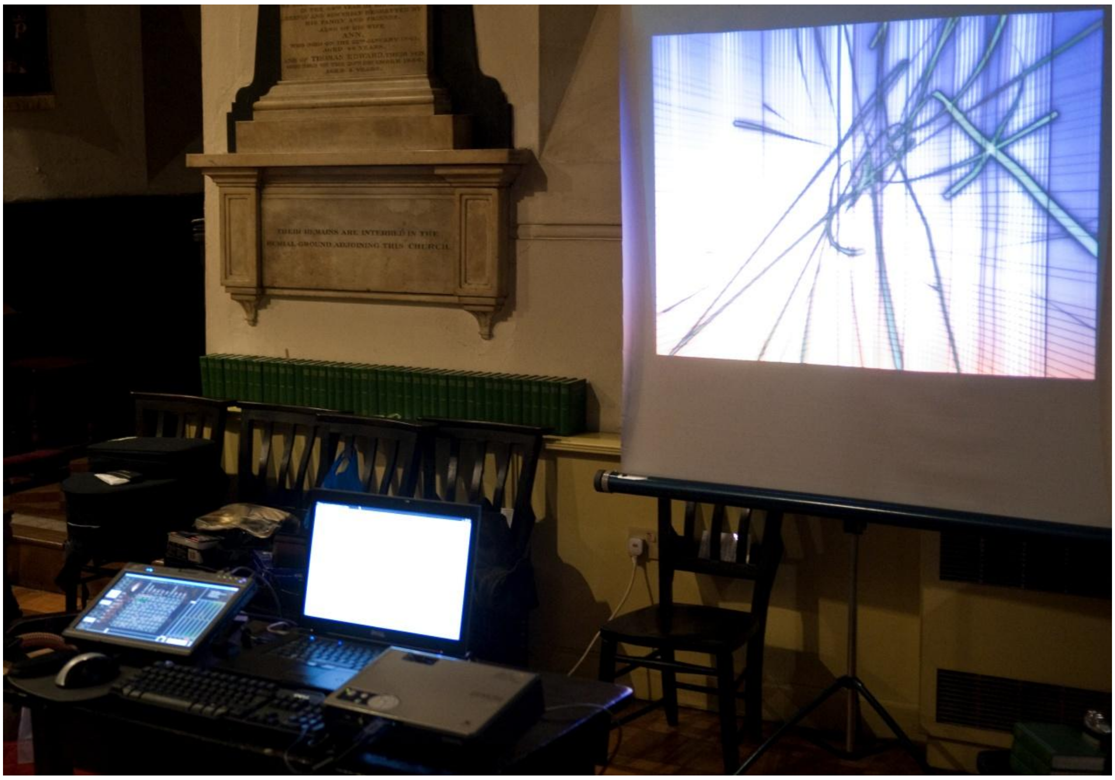


Figure 47: The Trinity system, set up for performance at St Pancras Old Church.

The rehearsal lasted approximately 150 minutes in total, not including setup and takedown times for equipment. The music was improvised in its entirety, by musicians that have long-going familiarity in playing together.

Although there were roughly 10 musicians, only three were connected to the system, as the necessary microphones and midi-devices to connect everyone was not available.

None of the musicians were instructed on the specifics of how the system worked, other than that the signal from each instrument needs to be input separately, and that both MIDI and audio was used.

These three were:

Guitarist 1: Played electric guitar, and was of intermediate level. Direct line audio could be used.

Guitarist 2: Played electro-acoustic guitar, and was of intermediate level. Direct line audio could be used.

Percussion synthesizer player: Used a Roland Handsonic percussion synthesizer from which both MIDI and direct line audio could be used. He too was at least of intermediate skill level, and showed significant familiarity with the particular electronic instrument.

Guitarist 1:

He was sat directly opposite of the projection screen, and thus had an uninterrupted, clear view of the screen throughout the rehearsal. He was observed to mostly look at the screen, but also chose to look away, towards the other musicians from time to time.

After the rehearsal, he chose to walk up and converse with me, stating that he found the overall experience to be appealing and different from anything he had previously seen.

He proceeded to mention that he noticed how throughout the performance, his instrument was connected to varying elements of the visuals, a consequence of my gradually altering the mapping, which is something I hadn't previously revealed that I would be doing.

He reflected on the fact that while normally when he plays music he largely shuts out the visual world focussing his attention on the music only, in this particular instance he was compelled to also take in the visual information from the screen, making for a different experience.

When asked whether he perceived the connection to the visual to be strong enough for him to be controlling the visual, rather than it reacting to his playing, he stated that he experienced both states at varying times during the performance.

When asked whether he believed to have altered his way of playing as a result of his being connected also to my system, he said that he often found himself experimenting with his instrument so as to discover what connections it had to the visual, but that the aesthetics of the visuals did not have any particular noticeable effect on the content of his playing.

The conversation ended with him suggesting that he looked forward to rehearsing again while connected to the system.

Guitarist 2:

She too was sat directly opposite of the screen, next to Guitarist 1, and was observed to look at the screen, but also around her at the other musicians, and at her instrument.

She briefly mentioned that she enjoyed the experience, but I was unable to have a longer conversation with her because she was in a hurry to leave directly after the rehearsal.

My own observation was that she often appeared to be experimenting with the sounds her instrument produced, while watching the screen, leaving me with the impression that she was experimenting with discovering the nature and dynamic ranges of the mappings between her instrument and the visuals.

Percussion synthesizer player:

Figure 48: Roland Handsonic percussion synthesizer (www.roland.com).

He was located next to the screen and did not have a direct view of the projection.

He was however observed to turn slightly so that the projection screen was visible, something he did frequently. It appeared that it was uncomfortable for him, but he repeatedly kept turning around, looking at the screen for sometimes many minutes continuously. He was also often looking at his instrument, and at other musicians.

During a break, and after the rehearsal, he walked up to me and initiated the conversation by stating that he greatly enjoyed the experience. He proceeded to ask whether I would want to come and rehearse with his other band, towards accompanying a performance of theirs in the beginning of January, which I gladly accepted. He promptly got in touch the day after to plan a rehearsal date.

On the night, he proceeded to reflect that he noticed the connections between his instrument and the visuals changed continuously, and asked me to confirm that indeed was the case.

He finally asked me to demonstrate how my system worked, and tried using it himself for roughly 15 minutes.

When asked whether he perceived the connection to the visual to be strong enough for him to be controlling the visual, ha stated that he did experience a varying level of control, depending on the complexity and nature of the mapping at each instant, the mapping being variable.

When asked whether he believed to have altered his way of playing as a result of his being connected also to my system, he said that he definitely strongly noticed a difference between his playing when looking at the screen, in comparison to when he was looking away. In fact he purposefully switched between the two states during his playing.

Accretion, EC1 Music Project

Figure 49: View of the instrument and Trinity system setup, for the rehearsal with Accretion.

The rehearsal lasted approximately 120 minutes in total, not including setup and takedown times for equipment. The music was partly rehearsed material, and partly improvised, by musicians that have long-going familiarity in playing together, of over a year.

None of the musicians were instructed on the specifics of how the system worked, other than that the signal from each instrument needs to be input separately, and that both MIDI and audio was used.

The musicians were:

Guitarist: Played electric guitar, and was of intermediate level. Direct line audio could be used.

Singer: Was of intermediate level. Direct line audio could be used.

Drummer: Used an acoustic drum kit, on which my Roland drum-triggers were mounted, to give MIDI data for each individual drum. No sensors were attached to the cymbals.

Guitarist:

He was located so that to his left he had the projection screen, to his right the drummer, and right in front of him the singer.

During the rehearsal he was often observed to look at the screen, but also at the other musicians and his instrument.

He stated that he enjoyed the rehearsal, and proceeded to ask several questions about the specifics of how my system worked.

When asked whether he perceived the connection to the visual to be strong enough for him to be controlling the visual, he stated that most of the time he felt it was just reacting.

When asked whether he believed to have altered his way of playing as a result of his being connected also to my system, he stated he oftentimes experimented to figure out the mapping, but that the aesthetics of the visual did not influence the aesthetic of his playing.

Singer:

She was located so that to her right she had the projection screen, to her left the drummer, and right in front of her the guitarist.

During the rehearsal she was often observed to look at the screen, but also at the other musicians.

Repeatedly during the rehearsal, she stated that she greatly enjoyed the visual accompaniment, and kept asking me to confirm whether she had correctly understood what mappings were present.

When asked whether she perceived the connection to the visual to be strong enough for her to be controlling the visual, she stated that she only really experienced it as reacting.

When asked whether she believed to have altered her way of singing as a result of her being connected also to my system, she stated she did not alter her way of singing at all.

Drummer:

He was located directly in front of the projection screen, and had an uninterrupted view of the visuals.

He nonetheless often appeared not to be actively using his vision at all, although he was also observed to pay closer attention to the visuals.

During a break of the rehearsal, he stated that he enjoyed the process, and that he would be very happy for it to be repeated, for me to join a rehearsal also of his other band, and of the future live performances of both.

When asked whether he perceived the connection to the visual to be strong enough for him to be controlling the visual, he stated that he only really experienced it as reacting.

When asked whether he believed to have altered his playing as a result of his being connected also to my system, he stated he did not alter his way of playing at all.

Live performance: The Gathering, St Pancras Old Church, 16/12/09

Figure 50: The projection screen on stage, at the Gathering performance.

Figure 51: Part of musicians and stage during the performance.

The performance lasted approximately 20 minutes in total. Only one of the performers had previously used the system, the Handsonic percussionist from the previous Gathering rehearsal. The band consisted of four musicians, a percussionist, a guitarist, a pianist and a singer. All but the singer were connected to the system. The Handsonic was connected using MIDI, while the guitar and piano were connected using a direct signal from the amplifier and a microphone respectively.

Although the performance was received well by the audience, there was no opportunity to ask the performers about their experience since they all necessarily had their backs turned to the screen most of the time, and there was no possibility of using a mirrored screen setup so that both musicians and audience could simultaneously see the projection.

Rehearsal towards live performance at Inspiral Lounge, 08/01/10

Figure 52: View of the Trinity system, the instruments and the performers during rehearsal.

The rehearsal lasted approximately 180 minutes in total, not including setup and takedown times for equipment. The music was improvised following previously rehearsed themes, by musicians that have some previous familiarity in playing together, of roughly a dozen rehearsals.

The musicians were instructed in the workings of the system.

The musicians were:

Percussion synthesizer player, from previous rehearsal, and a player of both a conducting gesture interface - the Native Instruments Maschine device offering a mode of interaction analogous to the well-established AKAI MPC series of samplers, and a synthesizer keyboard.

The percussion synthesizer was connected using individual midi and audio feeds. The Maschine device was only connected using its stereo audio output, since it did not have at the time of the rehearsal a completed MIDI implementation functioning. The keyboard synthesizer was connected using MIDI.

Both musicians were seated so that they would have an unobstructed continuous view of the projection screen. They were observed to at times concentrate at what was shown at the projection screen, at times at their instruments, and sometimes at each other.

Percussion synthesizer player:

The percussion synthesizer player again reported an experience analogous to that in his previous rehearsal with the system.

He felt that his playing did change when he was looking at the visuals, as a result of his ability to influence what appeared in them. He added that he did at times experience there to be a conflict, and purposefully looked away because he found the visuals distracting, while at other times could not look away, because he felt there being a unity between his playing, the resulting music, and the resulting visuals.

When asked to compare his experience with the system to that of an emulated "previous practice" scenario, he reported that he did indeed feel to a much higher extent the subjective sensation of being in control of the visuals, as opposed to them reacting to the music. This, as previously reported, varied depending on the complexity and nature of the mapping used at every instant.

Maschine and synthesizer player:

Figure 53: Native Instruments Maschine (www.native-instruments.com).

The Maschine and synthesizer player had no previous experience in rehearsing with the system, but had previous experience playing alongside a VJ, and of playing with a music visualizer that reacted to the stereo mixdown of the music, which is what is here referred to as “previous practice”.

He too reported that he did experience the visuals influenced his playing, sensing a conflict similar to that reported by the percussion synthesizer player. He also did state that he experienced there being a significant difference in the level of control he felt he had over the visual output, in comparison to his previous experience with visual accompaniment, and to the previously mentioned “previous practice” simulation.

He lamented the fact that his Maschine device did not produce MIDI output, and stated that in comparison to his playing with the synthesizer, he experienced a lesser level of control over the visuals when using the Maschine as a result of there being only a connection over the audio output of the device. He stated that for the next rehearsal/performance, he would research whether he could find a way to enable the MIDI output of the device, for the purpose of better influencing the projected visuals.

After the rehearsal both musicians confirmed that they were enthusiastic about my accompanying their scheduled live performance at the Inspiral Lounge using my system.

Live performance: Inspiral Lounge, Camden Town, 26/01/10

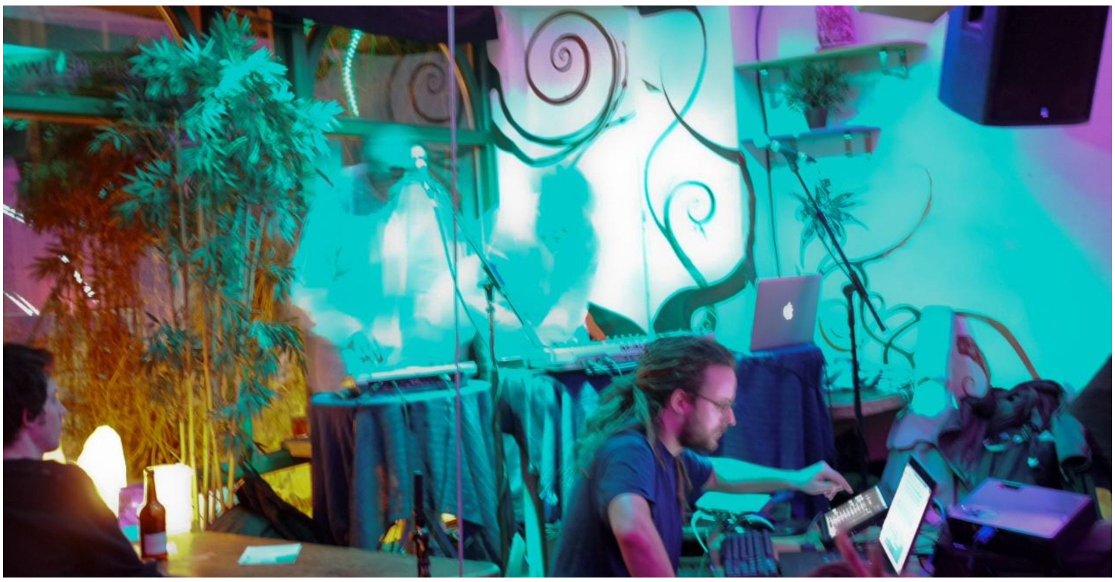


Figure 54: Live performance at Inspiral Lounge.

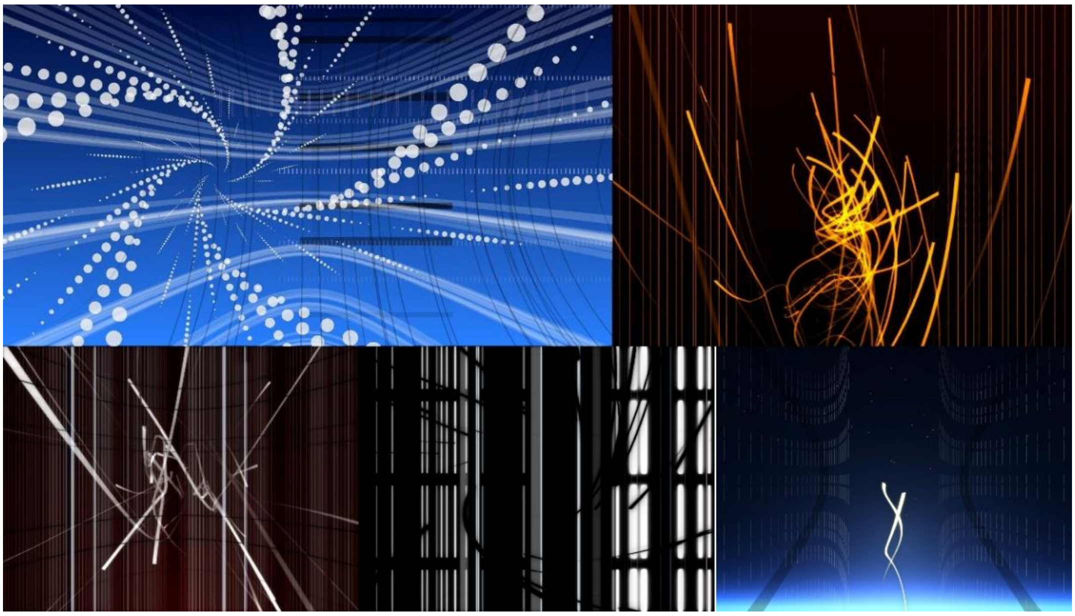
The performers were the same two as during the previous preparatory rehearsal. The instrumentation and technical setup was also identical. Set-up for the performance proved to be difficult, as although we were headlining the night, we were not given time to set up the gear before the nights performances began, but were instead expected to set-up just before our own performance. The result was that the musicians had to start performing without visual accompaniment, as the set-up time for the Trinity system is significantly longer than that of the instruments used.

The musical performance lasted an hour and a half, with visual accompaniment after roughly 15 minutes into the performance.

As was also the case in previous performances however, the visuals were necessarily projected so that the musicians couldn’t view them while performing. The performance was nonetheless well received.

Appendix C: Call for musicians

I am looking for musicians to accompany with live visuals!



I have over the past few years created a system, using which a group of musicians can control live computer generated visuals with their musical instruments, so that the visuals and music can closely interact.

I am now really looking forward to trying the system out with live musicians! If you'd like to try it out, get in touch with me and I'll bring my system to your next rehearsal. Eventually, if you like the results, I would also gladly take part in your live performances.

The only thing I expect to gain from our collaboration is that you agree for me to briefly interview you about the experience, and to partly film rehearsals/performances.

So if you are part of a live group located in or near London, no matter what musical genre, and you find my proposal appealing, do get in touch!

To view videos of my system in action, visit: www.vimeo.com/channels/onar3d/

Ilias Bergstrom: www.onar3d.com, onar3d@hotmail.com, i.bergstrom@ucl.ac.uk

Appendix D: Non-exhaustive list of visual synthesizers created

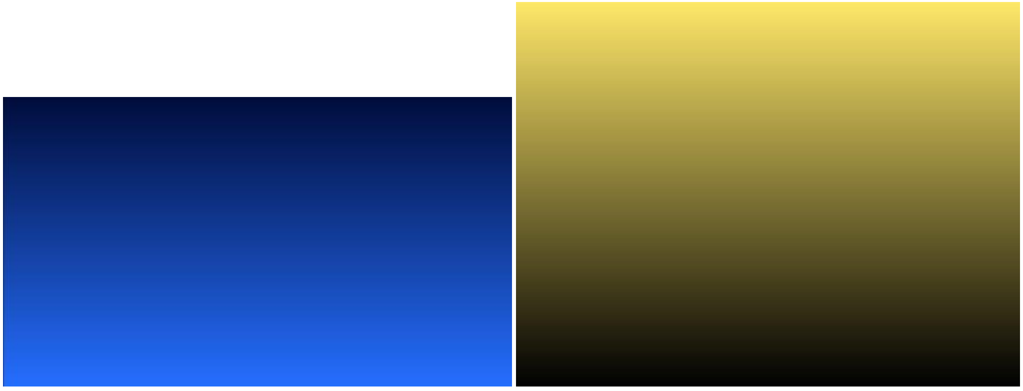


Figure 55: Gradient; Simplest example of visual synth. The code is available in appendix E.

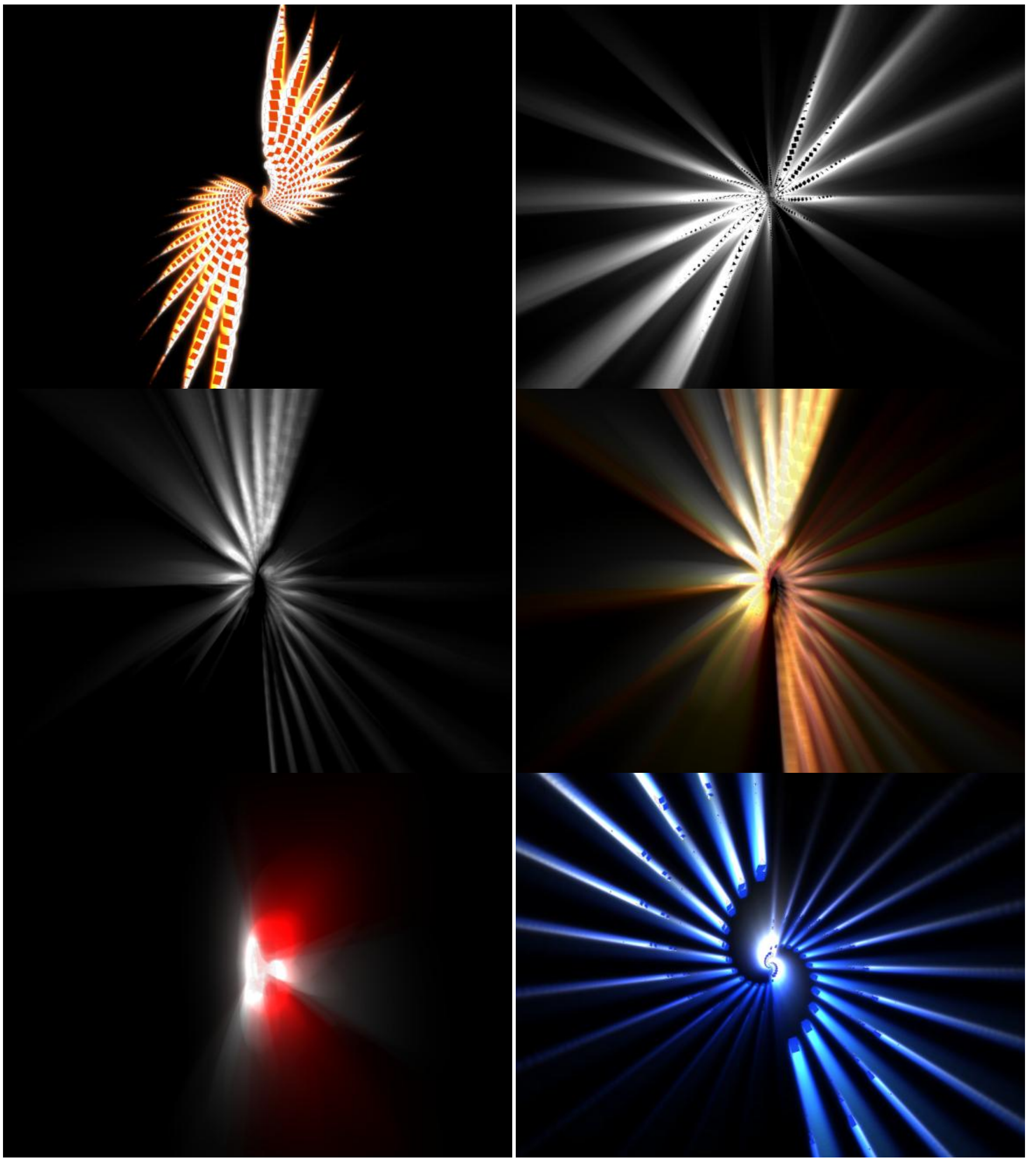


Figure 56: CubeSpine

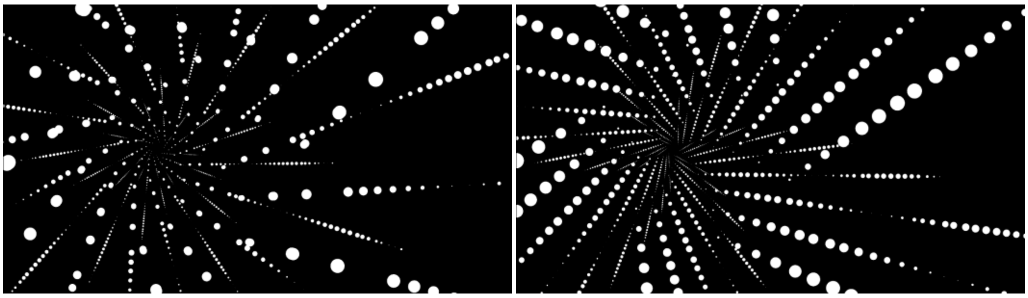


Figure 57: DotSpine, an adaptation of CubeSpine

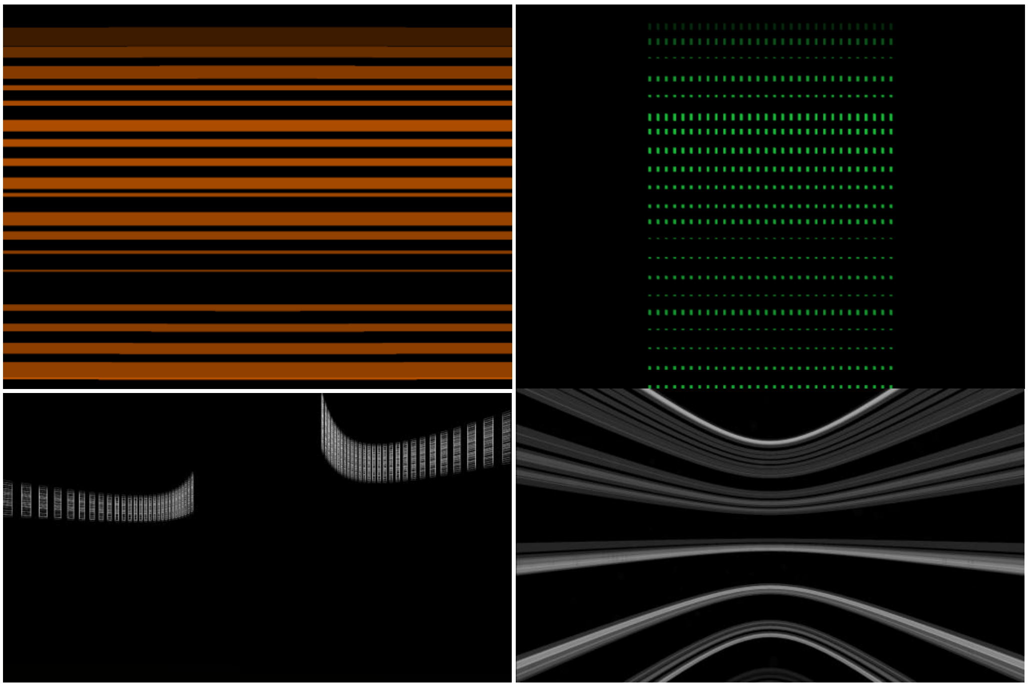


Figure 58: DSLines

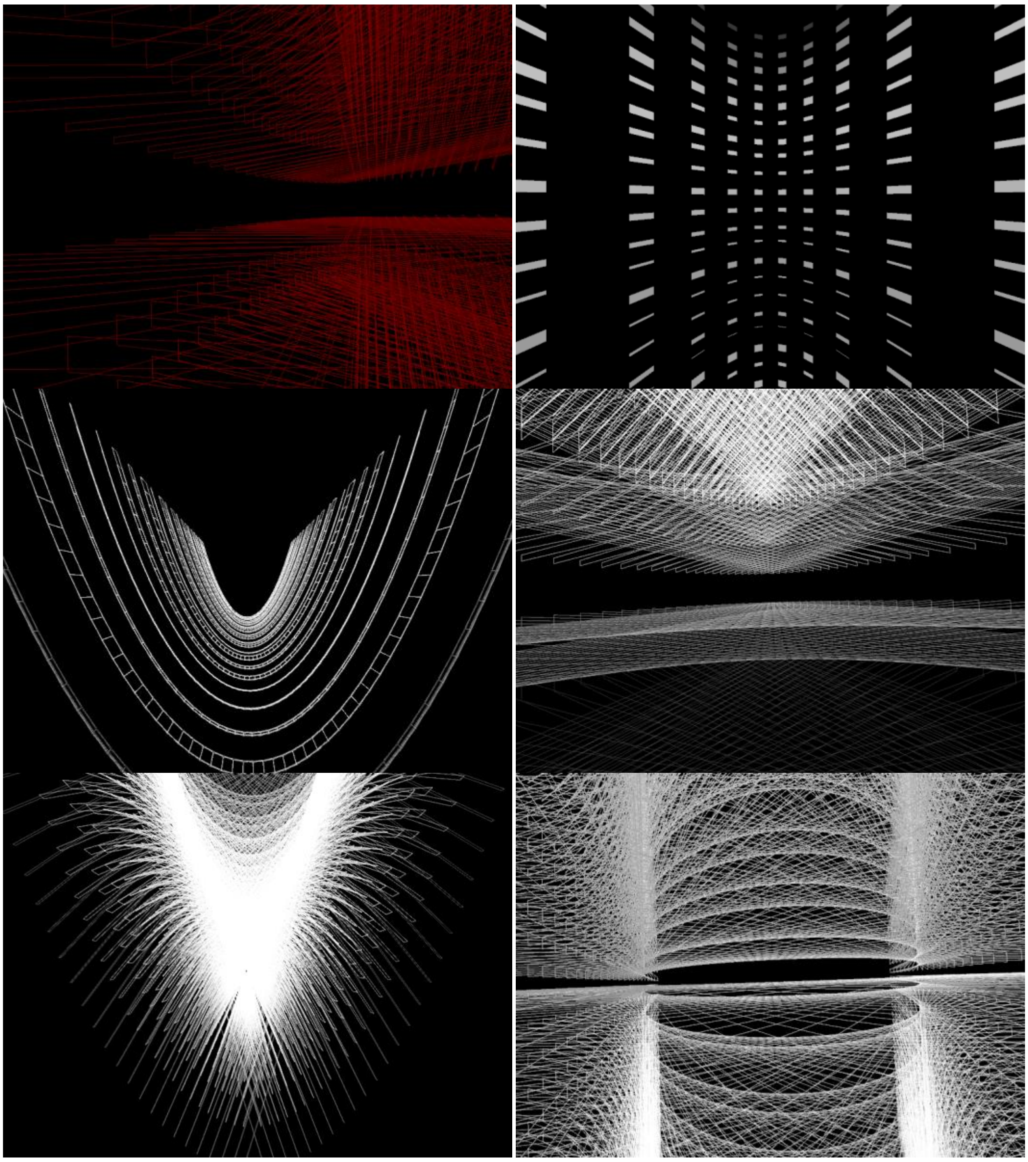


Figure 59: DSLines continued

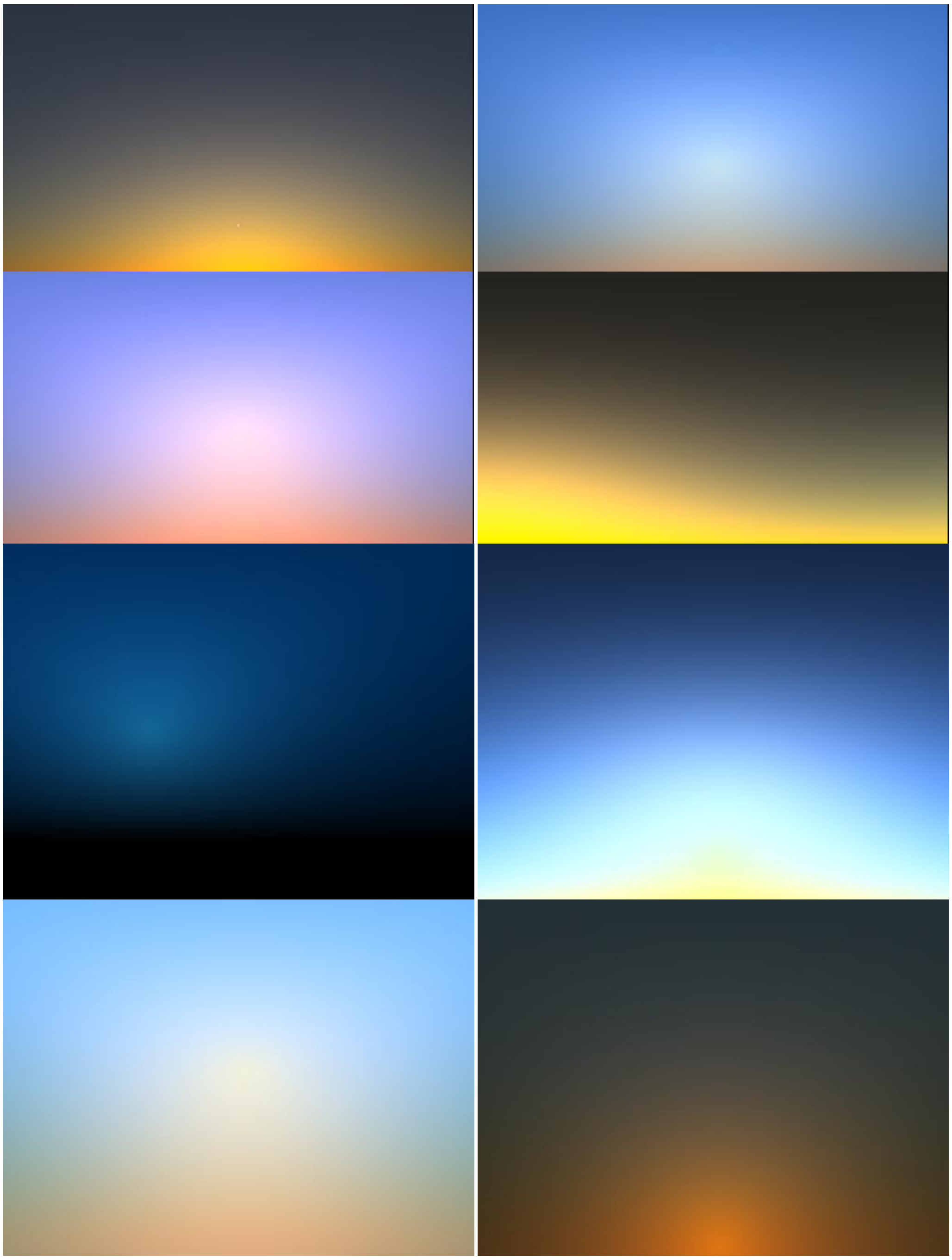


Figure 60: DSky, a non-realistic sky simulation

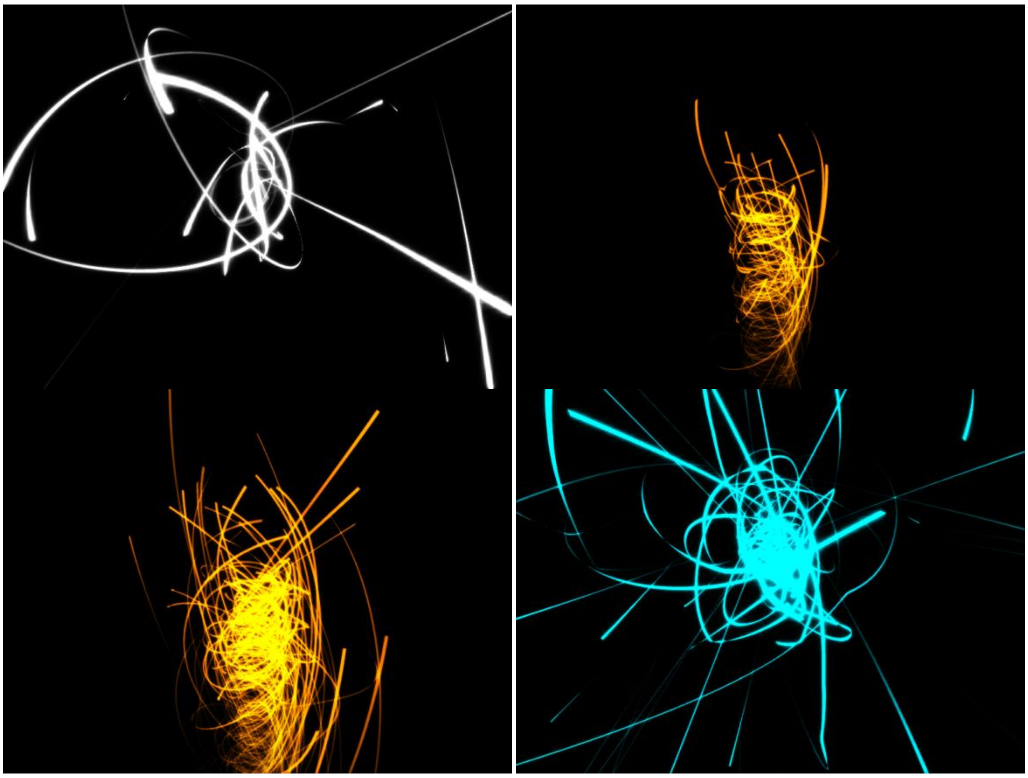


Figure 61: Swarm, a magnetic particle system of ribbons.

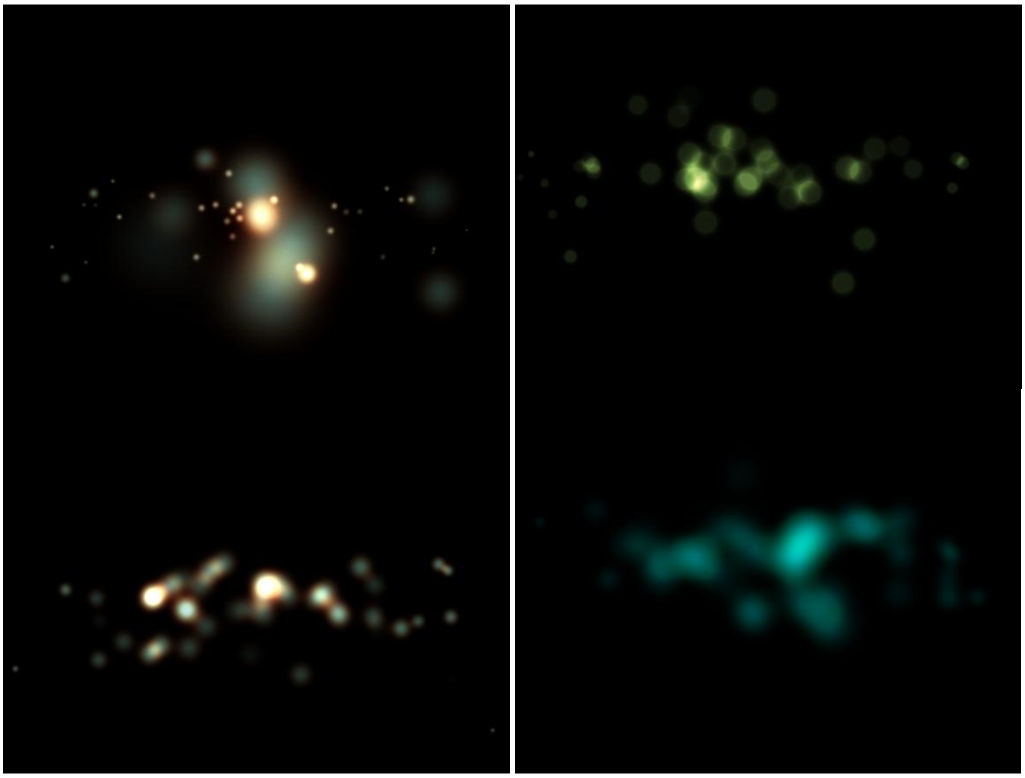


Figure 62: NFS, a bitmap based particle system.

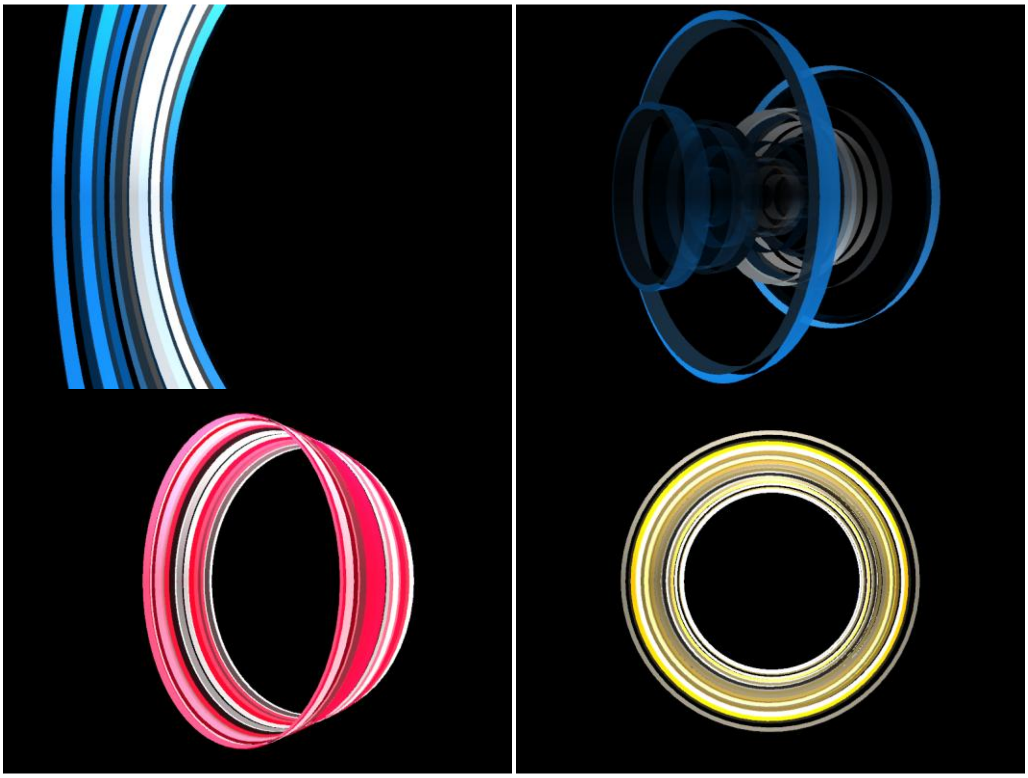


Figure 63: Spire. Particle system of concentric rings.

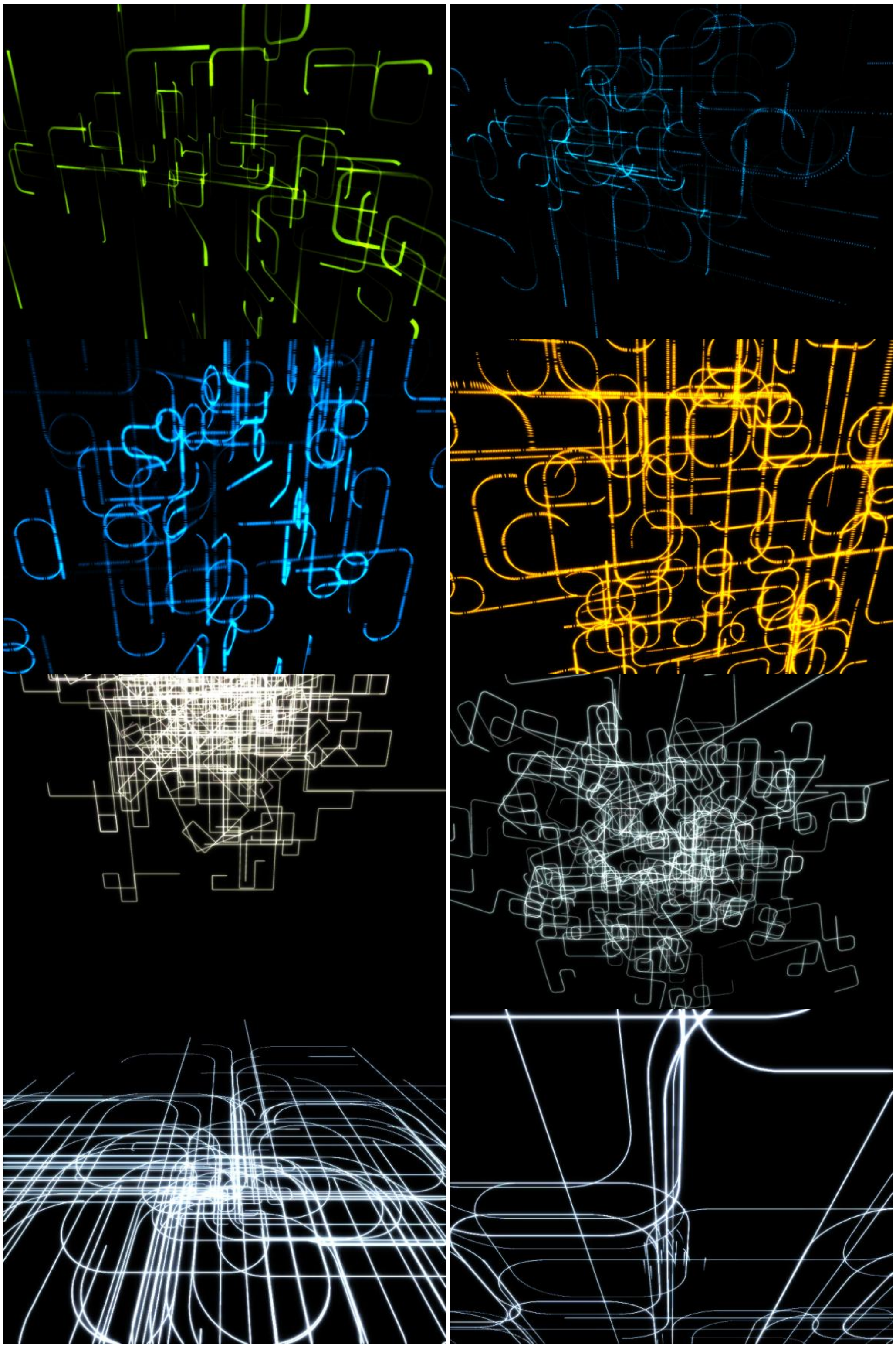


Figure 64: Circuit. Particle system of lines on a 2d/3d circuit

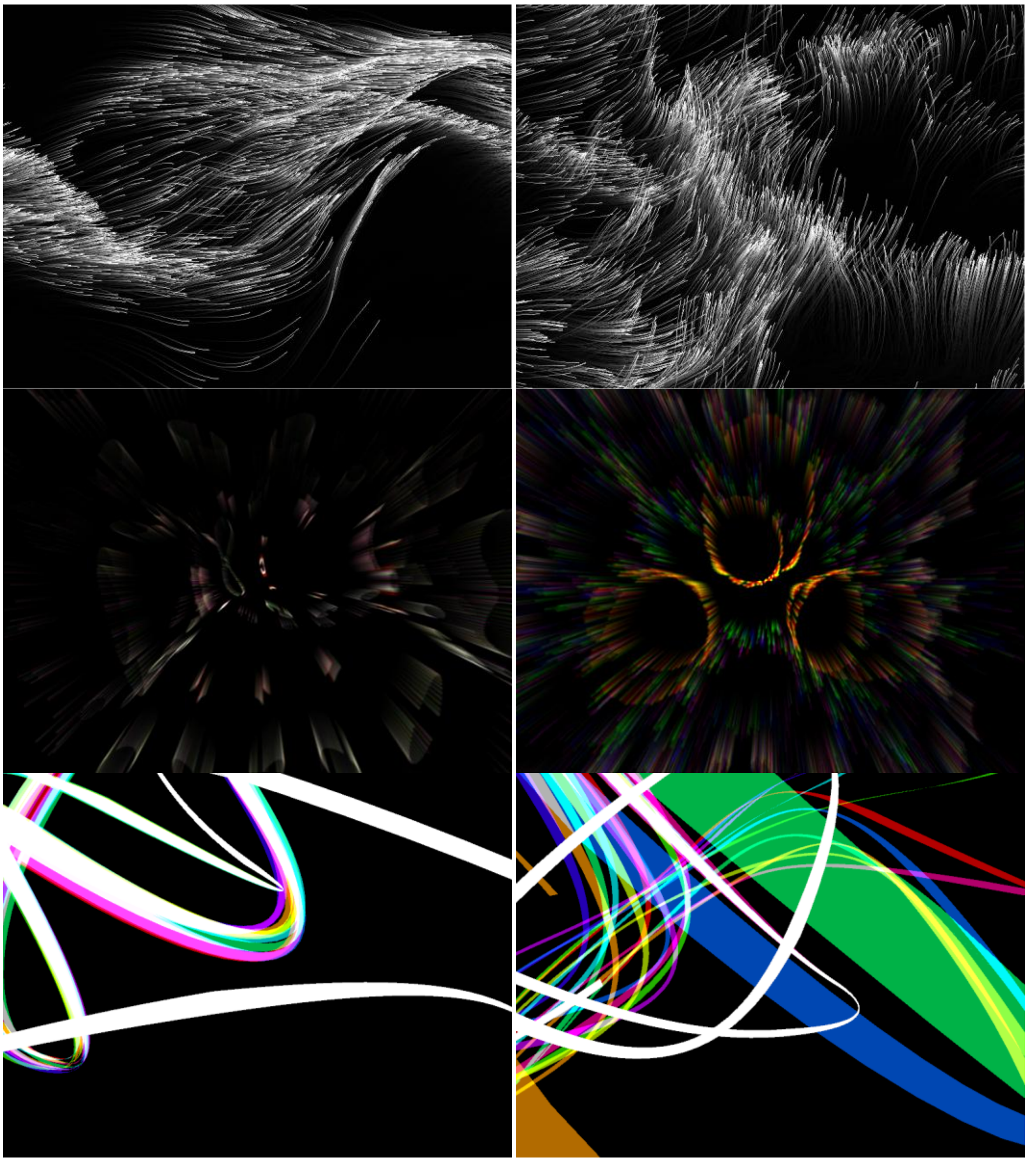


Figure 65: Three Processing sketches from www.OpenProcessing.org, adapted to assess how easily sketches not initially intended for Mother can be integrated.

Appendix E: Example code of Processing sketch turned into visual synths for Mother

Gradient Processing Sketch

```
import processing.opengl.*;
import javax.media.opengl.*;
import javax.media.opengl.glu.*;

float m_TopR;
float m_TopG;
float m_TopB;
float m_BotR;
float m_BotG;
float m_BotB;

void setup()
{
  size(400, 300, OPENGL);
  frameRate(24);

  noStroke();

  m_TopR = 1.0;
  m_TopG = 0.0;
  m_TopB = 0.0;
  m_BotR = 0.0;
  m_BotG = 0.0;
  m_BotB = 0.0;
}

void draw()
{
  pushMatrix();

  beginShape(QUADS);

  fill(m_TopR*255, m_TopG*255, m_TopB*255);
  vertex(0, 0);
```

```
  vertex(width - 1, 0);

  fill(m_BotR*255, m_BotG*255, m_BotB*255);
  vertex(width - 1, height - 1);
  vertex(0, height - 1);

  endShape();

  popMatrix();
}
```

Gradient Visual Synth for Mother

```
import megamu.shapetween.*;
import oscP5.*;
import netP5.*;
import foetus.*;

import processing.opengl.*;
import javax.media.opengl.*;
import javax.media.opengl.glu.*;

public Foetus f;

FoetusParameter m_TopR;
FoetusParameter m_TopG;
FoetusParameter m_TopB;
FoetusParameter m_BotR;
FoetusParameter m_BotG;
FoetusParameter m_BotB;

void setup()
{
  // When running as a synth within Mother, setup() is never called!
  // put the necessary initialization code in a method named
  // initializeFoetus().

  // The size() and framerate() Processing initialization
  // calls are called by Mother, and so should be left out
  // from initializeFoetus().
```

```
  // Finally, for the synth to work as a processing sketch
  // within the Processing Development Environment, call initializeFoetus()
  // from within setup().

  size(400, 300, OPENGL);
  frameRate(24);

  initializeFoetus();
}

void initializeFoetus()
{
  noStroke();

  // Instantiate foetus object here
  f = new Foetus(this);

  // Register messages that synth responds to (see OSC documentation).
  // This is here done automatically by the FoetusParameter constructors,
  // but you need to do this manually if you do not use FoetusParameter.
  m_TopR = new FoetusParameter(f, 1.0, "/TopRed",    "f");
  m_TopG = new FoetusParameter(f, 0, "/TopGreen",    "f");
  m_TopB = new FoetusParameter(f, 0, "/TopBlue",    "f");
  m_BotR = new FoetusParameter(f, 0, "/BotRed",      "f");
  m_BotG = new FoetusParameter(f, 0, "/BotGreen",    "f");
  m_BotB = new FoetusParameter(f, 0, "/BotBlue",     "f");
}

void draw()
{
  pushMatrix();

  beginShape(QUADS);

  fill(m_TopR.getValue()*255, m_TopG.getValue()*255,
m_TopB.getValue()*255);
  vertex(0, 0);
  vertex(width - 1, 0);
```

```
  fill(m_BotR.getValue()*255, m_BotG.getValue()*255,
m_BotB.getValue()*255);
  vertex(width - 1, height - 1);
  vertex(0, height - 1);

  endShape();

  popMatrix();
}

/**
 * This method is called when an OSC message is received by the synth.
 */
void oscEvent(OscMessage theOscMessage)
{
  if (theOscMessage.checkTypetag("f"))
  {
    if (theOscMessage.checkAddrPattern("/TopRed") == true)
      m_TopR.setValue(theOscMessage.get(0).floatValue());
    else if (theOscMessage.checkAddrPattern("/TopGreen") == true)
     m_TopG.setValue(theOscMessage.get(0).floatValue());
    else if (theOscMessage.checkAddrPattern("/TopBlue") == true)
      m_TopB.setValue(theOscMessage.get(0).floatValue());
    else if (theOscMessage.checkAddrPattern("/BotRed") == true)
      m_BotR.setValue(theOscMessage.get(0).floatValue());
    else if (theOscMessage.checkAddrPattern("/BotGreen") == true)
      m_BotG.setValue(theOscMessage.get(0).floatValue());
    else if (theOscMessage.checkAddrPattern("/BotBlue") == true)
      m_BotB.setValue(theOscMessage.get(0).floatValue());
  }
}
```

Appendix F: Presence research

A very recent research area, which is in many ways closely related to the work presented here, is the study of Presence. It is a interdisciplinary area related both to psychology, neuroscience, and computer science, and has varying sub-disciplines, the main ones being those of tele-presence, virtual presence, and social presence (http://ispr.info/). The area that is most relevant to this work is that of virtual presence, and is the one that is primarily going to be defined and discussed in this text. Presence is a default condition that we always experience: it is the sense of being situated in the real world which lies outside the human subject’s physical body. The academic study of Presence however involves mostly simulated situations, where the real world surrounding the subject is substituted with a synthetic one. Presence research concerns:

> *“(...) the successful replacement of real by virtually generated sensory data. Here successful means that the participants respond to the sensory data as if it were real, where response is at every level, from physiological through to cognitive”. (Gillies & Slater, 2005)*

The terms tele-presence and virtual presence are by many used interchangedly, with the only distinction between them being that in tele-presence, the subject is through the mediation of technology made to feel present in an actual, existing environment, that is however located somewhere far from the subject, while in virtual presence, the simulated environment is non-existent in the real world, and is entirely synthesized. Finally, social presence is defined as:

> *"Social presence" (distinct from social *realism*) occurs when part or all of a person's perception fails to accurately acknowledge the role of technology that makes it appear that s/he is communicating with one or more other people or entities. (http://ispr.info/)*

With presence research still being a very young field, there is still a significant amount of discussion regarding what does and doesn’t qualify as relevant presence research (Slater, 2003), with the opinion being voiced that the field is in need of more focus (Slater, 2007a) (Lee, 2004). This is not being made easier by the fact that the successful measurement of presence is to a significant extent still an open research question, and although progress has been made, there is still much work left to be done before presence may be deterministically quantified (Slater, 2004) (http://ispr.info/). It is currently also not possible to detect whether a given situation is one where a subject would experience presence, or if the experience is in fact something entirely different; at the moment the measurement of presence can only be discussed in the context of situations where it is already known that it is a de-facto condition (Slater, 2004).

Because of the difficulties described, it is necessary to be very careful when relating ones work to the field of presence, so that it may be beneficial both for the advancement of related work, as well as the advancement of the presence research field. It would be very easy, because of the early stage in which presence research currently is, to instead partake in diffusing the field.

Relevance to Presence field

If the relation of present work to the presence research field is first to be detailed, the system described can easily be regarded as a virtual reality system, in which performers and audience are immersed, albeit each to different extents. The immersion achieved is of a significant level, since high definition synthesized computer graphics are displayed, with correlated audio, and also correlated proprioceptive stimuli from the live musical performance. Described system has a number of characteristics, which are all well-established within presence research as aspects that contribute to greatly increasing the experienced level of presence:

- It produces a high level of immersion (Slater, 2007b)
- The stimuli presented to the senses are highly correlated between each other (Slater, 2007b)
- There is a high degree of interactivity between the elements in the virtual environment and its users (Lombard & Ditton, 1997)

The problem however, lies in that it is very difficult to state with certainty that present system qualifies as being a virtual environment in which it is at all valid to discuss the topic of presence. An important aspect for the definition of virtual presence is that the virtual environment needs to induce a “perceptual illusion of non-mediation” (Lombard & Ditton, 1997). One cannot easily claim that proposed system could produce such an illusion, since it is not a recreation of a realistic environment, in which one can with certainty state that the subjects experience any sort of "suspension of disbelief". To refer back to Mel Slater's previously provided definition of virtual presence, the participants need to respond to the virtual sensory data *as if it were real*; but how can one then discuss presence, in the context of a virtual environment that presents stimuli which in no way attempts to simulate reality, at any level, as is the case with the system described here? Because presence cannot be detected, and only to some extent be measured, any discussion on whether a user of a system experiences presence, needs to be grounded in a comparison to the baseline provided by a real situation. This is not possible with a virtual environment that makes no attempt at simulating reality. The only comparison that can then be made is between the level of presence that individual subjects would experience if such non-realistic worlds actually exist (Slater, 2003).

To further illustrate this, it is useful to discuss one further description by Mel Slater:

> *"Presence is the response to a given level of immersion (and it only really makes sense when there are two competing systems - one typically the real world, and the other the technology delivering a given immersive system)." (Slater, 2003)*

Again, it would be a difficult case to argue that with the proposed system there is any direct conflict between competing systems. Because although it may be argued that participants may feel much less present in the actual world when using it, this actual world is not replaced with anything in which they may be alternatively present. How could the case be made, that proposed system replaces the current world, and isn't simply a part of it?

To conclude this discussion, the choice has to be made to keep following the work produced by the research community on the subject, so as to re-evaluate whether it is possible to make a connection to the presence research field. Currently however, it is deemed that relating present research to the body of research on presence, might in fact only risk in creating more confusion than it helps clear.

Appendix G: Physiological Measurements for Detecting Enjoyment and Engagement in Video Games, Affective Computing

Much research has been carried out towards observing and measuring the feeling of affect and emotion on human subjects while they interact with various computer systems. Two primary areas related to computer science and HCI have been identified that are concerned with this, but it is likely that similar work has been done also in other areas. Researchers that concentrate on video games have recognized that performing physiological measurements is likely the most promising method for gathering quantitative information of a subjects experience while playing a game. *Affective computing*, a direction within HCI research, uses primarily physiological measurements, to allow computer systems to recognize the emotional state that their users are in, and also communicate emotional states back to the user.

Mandryk and Inkpen (2004) review how researchers have not yet reached an agreement regarding whether emotions can be codified into a number of discrete emotional states, or whether they are instead better understood as existing along multiple axes in space.

Physiological measurements have been used by many to quantifiably observe various aspects of human experience (see (Mandryk & Inkpen, 2004) for a review), however specifically to indicate subject enjoyment or engagement while interacting with a computer, research began only as recently as 2004 (Mandryk & Inkpen, 2004). Mandryk and Inkpen used a multitude of measures, such as galvanic skin response, cardiovascular measures, respiratory measures and electromyography. By performing initial experiments, they found that their measurements could indeed detect changing levels of engagement and enjoyment, which correlated with the reported experience of the test subjects. They followed this study up (Mandryk, Atkins, & Inkpen, 2006) (Mandryk & Atkins, 2007) with a method for modelling user emotional state based on a system that examines the combined physiological measurements. The conclusions from these studies were that although showing great promise, the technique is in need of further development and that physiological responses could not alone be used as a sole evaluation metric, but would need to be used in combination with existing subjective techniques.

In his study Richard Hazlett (2006) used facial electromyography to distinguish between positive and negative emotional valence during play. He found that in an experiment where thirteen boys play a simple racing game, the zygomatic muscle electromyography (that controls smiling) was found to be significantly greater during events through video review identified as positive, compared to events identified as negative. He too concluded however that further research was necessary to validate the technique for more complicated contexts.

Boehner et al (2007), from reviewing the literature on research regarding emotion, find that:

> *"(...) emotion is not thought of as biological, measurable and objectively present because scientists found it to exist in the world that way, but because 19th-century scientists could not imagine studying it scientifically any other way".*

They disagree with the view that emotion may be codified into transmittable form by a discrete computer system at all, arguing that there is no evidence it ever exists in such a form, for it to subsequently be elicited. They propose instead that affective computer systems may be more effective if made to support human users to understand, interpret and experience emotion in its full complexity and ambiguity, where it is not attempted to codify specific emotions into unambiguous, transmittable data. This view is in agreement with the discussion we have followed by Dourish (2001) and others, that human cognition is in fact not analogous to the information processing model supported by cognitive science, but is instead highly situated and embodied.

In the last two CHI conferences from 2007 and 2008, special interest group workshops and panel discussions on emotion research in HCI have been held (Crane, Shami, & Peter, 2007) (Shami, Hancock, Peter, Muller, & Mandryk, 2008). The remaining unanswered questions detailed in the papers announcing these workshops, reflect the conclusion that although at the time of writing it has not been found that physiological measurements can be used alone to detect the level of engagement and enjoyment a user experiences, physiological measurements have nonetheless been shown to be useful tools, if employed with care. They can with benefit be used for providing continuous, high resolution quantitative data that can be corroborated with qualitative studies to provide a more detailed understanding than possible with the qualitative studies alone.

Physiological measurements are used also in Presence research, where results from qualitative studies are usually corroborated with data from the measurements, eg (Slater et al., 2006). The use of physiological measurements in Presence research will not be reviewed here however, since methods for eliciting the level of presence are not directly transferrable to what we are primarily interested in measuring, the feelings of engagement and enjoyment.

Finally, an important note is that identified studies have all been performed using subjects that were relatively still. If the subjects were instead performing more physically demanding motions, using the whole of their bodies, it is possible that many of the measurements performed would be negatively affected. Both because of the possible effect of these motions on the physiological signals, and because the sensors used would need to be attached much more securely to the subjects, so as to avoid noise from the movements performed.

Appendix H: Allowing for multiple interpretations during evaluation process

Sengers and Gaver (2006) discuss developing systems that allow and may even encourage more than a single interpretation of how they can be used. This is in contrast with common HCI practice, where ambiguity is often seen only as a symptom of bad design. Although it is acknowledged in HCI that there are multiple, often conflicting interpretations of how a system should work corresponding to different stakeholders, this is only in order to decide on which of the different interpretations to emphasize as the correct one for the design. Similarly during the evaluation of a design it is examined whether the selected interpretation is correctly communicated by the system to the users, and if usability testing is carried out it examines the system while used according to the one emphasized interpretation.

Especially for systems that are not intended for use in a work context, this focus on only one interpretation can be very problematic. There are many advantages if instead the possibility of allowing for many different interpretations is acknowledged. A system the usage of which can be interpreted differently by different individuals, has the potential of allowing it to be much more adaptable to the various ways in which these individuals work, and to allow for its use in much more diverse and often entirely unpredicted contexts.

This is particularly relevant to the evaluation process of the paradigms and systems presented in this thesis. All are novel conceptions, rather than iterative refinements of pre-existing established ideas and practices, and all have been conceived with a particular set of intended uses. As part of their evaluation, particular care should therefore be taken to also elicit the interpretations about their potential use of the subjects involved, and not just attempt to confirm their usefulness for the purposes they were conceived for.

To provide a concrete example, there is a software application category where this is very strongly evident, that of audio sequencing applications such as Steinberg Cubase, Digidesign Pro Tools, Ableton Live, and many others that follow a similar paradigm. These applications are used in many different contexts: producing electronic music, recording and mixing live music, scoring music for a film, or producing live or pre-recorded radio broadcasts. In all these different contexts, the tools provided in the application are used differently, in a manner that is both dependent on the requirements of the task at hand, and also to a great extent on the user of the application. Even two electronic music producers working within the same sub-genre, may use the same software package in an entirely different manner, to suit each musicians personal way of working.

Sociological and historical studies have shown that the meanings attached to technologies stabilize only after lengthy process of their use, in which all different stakeholders in the use of the technology are involved. During this process, often new uses of the technology emerge that have usually never been considered by its developers. Greenberg and Buxton (2008) detail examples of this occurring throughout history; when Marconi invented the radio, he envisioned it as a means of maritime communication between ships and shore, and did not predict what would come to be its most common use, that of radio broadcasting.

A telling example from the audio sequencing world, occurred when the innovative application Ableton Live was released, one of its main selling points being the novel feature of allowing audio to be sped up and slowed down, while it remained in the same key, or altering the key while keeping the playback speed constant. After it was deployed, a use of it appeared that had not been predicted or designed for by its developers: if audio was significantly altered in tempo, artefacts began to appear, making the sound 'grainy' and alien. Many musicians took a liking of these artefacts, and incorporated them in their music, purposefully exploring the settings of the application that controlled the artefacts, in the same way that they would experiment with the sound-shaping settings of a synthesizer.

Sengers and Gaver (2006) continue discussing how design can be carried out to encourage and discover the possible interpretations that may arise regarding the system design, as well as how to evaluate the outcome of such work. Based on their own experience from research projects such as the drift table, (Gaver et al., 2004), the history tablecloth (Gaver et al., 2006), and other projects referenced therein, they suggest a number of approaches to design for multiple interpretations:

- To design so that while usability of the system is clearly specified, the interpretation of use is open.
- To support a space of interpretations around a given topic.
- To disallowing the interpretations most obviously expected, so as to stimulate new ones.
- To gradually unfold new opportunities for interpretation over the course of interaction.
- Downplaying the systems authority to make space for re-interpretation.
- To prevent any consistent interpretation.

Evaluation while allowing for multiple interpretations

While designing for multiple interpretations may have many benefits, as previously discussed, those benefits come with the price of increased difficulty when evaluating research outcome. Because simply counting the number of different interpretations that a system encourages, is not an evaluation metric by itself, as neither of these interpretations needs to be desirable.

Under the light of this discussion, the meaning of what constitutes a successful system is significantly altered. Should a system that failed to demonstrate usefulness for its initially intended application, but which proved to be greatly useful for a number of other purposes, in the end be considered a success or a failure? What is then the primary purpose of the process of evaluation? To assess the systems usefulness only against the presumptions of how it could prove to be useful, that were held during its inception, or to also identify and test against the other possible interpretations of the systems use?

Sengers and Gaver (2006) discuss that common HCI evaluation approaches within HCI cannot be effectively be applied in this context, since they are virtually all based on the premise that there is a single, authoritative interpretation from which evaluation criteria are derived. They propose that users own interpretations need to be identified and taken into account in the evaluation process, where users may be a selection of many different identified stakeholders, as well as people purposefully selected from varying backgrounds, that are expected to have a differing interpretation of the system.

For the purpose of identifying and recording the different interpretations that may arise, the usefulness of ethnography is apparent, as it is greatly valuable for capturing rich descriptions of people's experiences with new designs. It is also well suited for capturing the way interpretations may substantially evolve over time, as has been demonstrated with the drift table case study (Gaver et al., 2004).

When the varying interpretations that could arise have been identified, traditional HCI usability evaluation methodology can and should of course be applied, if it is deemed to be a useful tool for assessing the systems usefulness in the light of a specific interpretation. Sengers and Gaver (2006) also discuss how evaluation methodologies from other disciplines may be an additional or even alternative tool that may prove beneficial during evaluation. They too discuss the usefulness of expert critiques, as practiced in humanities and arts disciplines, as well as involving commentators from outside of traditional academia, such as journalists, art critics, or as they have done themselves, a documentary filmmaker that was asked to create documentaries on how users were approaching the devices the authors had made. On the subject of evaluation, Sengers and Gaver conclude that:

> *"(...) any form of evaluation in the end relies on our expert peers in the HCI community to judge its effectiveness and success – or lack thereof. This*

suggests opportunities for evaluation which are aimed, not at finding a final answer of what worked and didn't work, but at supplying data in a form which expert readers can interpret for themselves."

Appendix I: HCI, Interaction Design, and present work

There seems to be some ambiguity in how interaction design, HCI, and their relationship are defined in the literature, necessitating a brief discussion on how these terms are used in this thesis. In their book "Interaction Design: beyond human-computer interaction" (2002), Preece et al address the relationship between the two as being a matter of scope:

> "*ID casts its net much wider, being concerned with the theory, research, and practice of designing user experiences for all manners of technologies, systems, and products, whereas HCI has traditionally had a narrower focus, being concerned with "the design, evaluation, and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them" (ACM SIGCHI, 1992, p6)"*. (Preece et al., 2002),

The work presented here is compatible with both these definitions, and when the terms are used the definitions implied are the ones stated here, unless an alternative is specifically described.

In reviewing research papers however, the relation of these two areas does not always seem to be in accordance to this definition. In many, the two terms seem to be used interchangedly, and are not explicitly defined. In their review, Zimmerman et al (2007) do not explicitly define HCI at all, and interaction design is mentioned in the context of it being a craft discipline similar to architecture or graphic design, and often in the paper the relation between interaction design and HCI seems to be approached as if HCI is oriented more towards the description given by Harrison et al (2007) for the first two paradigms of HCI.

There is one possible interpretation for this difference. Since the definition of the third paradigm that Harrison et al presented in their paper from 2008 had not yet been published, Zimmerman et al, who published their review in 2007, instead used the term interaction design to stress a distinction similar to that Harrison et al later made with the third paradigm. If this interpretation is indeed correct, then the discussion by Zimmerman et al is also relevant for our work, since present work is as we have previously noted well aligned with the scope of the third HCI paradigm.

References

Abbado, A. (1988). Perceptual correspondences of abstract animation and synthetic sound. *Leonardo. Supplemental Issue*, *1*, 3–5.

Aimi, R. M. (2007). *Hybrid percussion: extending physical instruments using sampled acoustics* (PhD Thesis).

Alais, D., Newell, F. N., & Mamassian, P. (2010). Multisensory Processing in Review: from Physiology to Behaviour. *Seeing and Perceiving*, *23*, 3-38.

Anderson, M. L. (2003). Embodied Cognition: A field guide. *Artificial Intelligence, 149*(1), 91-130.

Arfib, D., Couturier, J. M., & Kessous, L. (2005). Expressiveness and Digital Musical Instrument Design. *Journal of New Music Research*, *34*(1), 125-136.

Armstrong, N. (2006). *An Enactive Approach to Digital Musical Instrument Design* (PhD Thesis).

Bergstrom, I. (2003). *Visualizing music in real-time using MIDI information as input: Exploring the art of Lumia* (MSc thesis). University College London.

Bergstrom, I., & Lotto, B. (2008). Mother: Making the Performance of Real-Time Computer Graphics Accessible to Non-programmers. In *(re)Actor3: The Third International Conference on Digital Live Art Proceedings* (pp. 11-12). Presented at the (re)Actor3: The Third International Conference on Digital Live Art.

Bergstrom, I., & Lotto, B. (2009). Harnessing the Enactive Knowledge of Musicians to Allow the Real-Time Performance of Correlated Music and Computer Graphics. *Leonardo, 42*(1), 92-93.

Bergstrom, I., Steed, A., & Lotto, B. (2009). Mutable Mapping: gradual re-routing of OSC control data as a form of artistic performance. In *Proceedings of the* *international conference on Advances in computer entertainment technology* (pp. 290-293). Presented at the ACE 2009, Athens, Greece.

Besson, M., & Friederici, A. (2005). Part II: Language and music—A comparison. *Annals of the New York Academy of Sciences*, *1060*, 57–58.

Bevilacqua, F., Müller, R., & Schnell, N. (2005). MnM: a Max/MSP mapping toolbox. In *Proceedings of the 2005 conference on New interfaces for musical expression* (p. 88).

Bode, H. (1984). History of Electronic Sound Modification. *Journal of the Audio Engineering Society*, *32*(10), 730-739.

Boehner, K., DePaula, R., Dourish, P., & Sengers, P. (2007). How emotion is made and measured. *International Journal of Human-Computer Studies*, *65*(4), 275-291.

Bowers, J., Button, G., & Sharrock, W. (1995). Workflow from within and without: Technology and cooperative work on the print industry shopfloor. In *Proceedings of the fourth conference on European Conference on Computer-Supported Cooperative Work* (p. 66).

Bryan-Kinns, N., & Hamilton, F. (2009). Identifying mutual engagement. *Behaviour & Information Technology*, (1), 1–25.

Buxton, B. (2007). *Sketching User Experiences*. Morgan Kaufmann, San Francisco, California.

Buxton, B., Buchla, D., Chafe, C., Machover, T., Mathews, M., Moog, B., Risset, J. C., et al. (2000). Round table: electronic controllers in music performance and composition. *Trends in Gestural Control of Music*, 415–38.

Carroll, J. M. (1997). HUMAN-COMPUTER INTERACTION: Psychology as a Science of Design. *Annual Review of Psychology*, *48*(1), 61-83.

Casciato, C., & Wanderley, M. M. (2007). Lessons from Experienced Gestrural Controller Users. In *Proceedings of the International Conference on Enactive Interfaces (ENACTIVE)* (pp. 333-336). Grenoble, France.

Chion, M., Gorbman, C., & Murch, W. (1994). *Audio-vision: sound on screen*. Columbia University Press.

Christou, G., Law, E. L., Green, W., & Hornbaek, K. (2009). Challenges in evaluating usability and user experience of reality-based interaction. In *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems* (pp. 4811-4814). Boston, MA, USA: ACM.

Collins, N. (2006). Towards autonomous agents for live computer music: Realtime machine listening and interactive music systems. *Unpublished doctoral dissertation, Music Department, University of Cambridge*.

Collins, N., McLean, A., Rohrhuber, J., & Ward, A. (2003). Live Coding in Laptop Performance. *Organised Sound*, *8*(03), 321-330.

Collopy, F. (2000). Color, Form, and Motion: Dimensions of a Musical Art of Light. *Leonardo*, *33*(5, Eighth New York Digital Salon), 355-360.

Collopy, F., Jameson, D., & Fuhrer, R. (1999). Visual music in a visual programming language. In *1999 IEEE Symposium on Visual Languages* (pp. 111-118). Presented at the 1999 IEEE Symposium on Visual Languages.

Comaroff, J. L., & Comaroff, J. (1992). *Ethnography and the Historical Imagination*. Westview Press.

Cont, A., Coduys, T., & Henry, C. (2004). Real-time gesture mapping in Pd environment using neural networks. In *Proceedings of the 2004 conference on New interfaces for musical expression* (pp. 39-42). Hamamatsu, Shizuoka, Japan: National University of Singapore.

Crabtree, A., Rodden, T., Tolmie, P., & Button, G. (2009). Ethnography considered harmful. In *Proceedings of the 27th international conference on Human factors in computing systems* (pp. 879–888).

Crane, E. A., Shami, N. S., & Peter, C. (2007). Let's get emotional: emotion research in human computer interaction. In *CHI '07 extended abstracts on Human factors in computing systems* (pp. 2101-2104). San Jose, CA, USA: ACM.

Create Digital Motion » Mother 0.2, App for Live Performance with Processing, Updated. (2008). Retrieved May 6, 2010, from http://createdigitalmotion.com/2008/11/mother-02-app-for-live-performance-with-processing-updated/

Create Digital Motion » Processing for VJing and Performance, with Mother. (2008). Retrieved May 6, 2010, from http://createdigitalmotion.com/2008/11/processing-for-vjing-and-performance-with-mother/

Cross, I. (2003). Music and evolution: Consequences and causes. *Contemporary music review*, *22*(3), 79–89.

Csikszentmihalyi, M., & Csikzentmihaly, M. (1990). *Flow: The psychology of optimal experience*. Harper & Row New York.

Cytowic, R. E. (1995). Synesthesia: Phenomenology and neuropsychology. *Psyche, 2*(10), 2–10.

Dabrowski, M. (1997). *Kandinsky: Compositions*. Harry N. Abrams.

Davis, D. (1975). *Art and the future: a history, prophecy of the collaboration between science, technology and art*. Praeger Publishing.

Design By Numbers. (n.d.). Retrieved from http://dbn.media.mit.edu/whatisdbn.html

Dix, A., Finlay, J., & Abowd, G. D. (2004). *Human-Computer Interaction*. Prentice

Hall.

Dixon, S. (2007). *Digital performance: a history of new media in theater, dance, performance art, and installation*. The MIT Press.

Dobrian, C., & Koppelman, D. (2006). The 'E' in NIME: musical expression with new computer interfaces. In *Proceedings of the 2006 conference on New interfaces for musical expression* (pp. 277-282). Paris, France: IRCAM — Centre Pompidou.

Dourish, P. (2001). *Where the Action Is: The Foundations of Embodied Interaction*. Mit Press.

Dourish, P. (2006). Implications for design. *Proceedings of the SIGCHI conference on Human Factors in computing systems*, 541-550.

Droste, M., & Williams, K. (2002). *Bauhaus, 1919-1933*. Taschen America Llc.

Ebert, D. S. (2003). *Texturing & Modeling: A Procedural Approach*. Morgan Kaufmann Pub.

Edmonds, E. A., Weakley, A., Candy, L., Fell, M., Knott, R., & Pauletto, S. (2005). The studio as laboratory: combining creative practice and digital technology research. *International Journal of Human-Computer Studies*, *63*(4-5), 452–481.

Eisenstein, S., & Leyda, J. (1947). *The film sense*. Mariner Books.

Eisenstein, S., & Leyda, J. (1949). *Film form: essays in film theory*. Mariner Books.

Electrovision 18th of July 2009 - Roxy Bar and Screen. (2009). Retrieved May 6, 2010, from http://www.electrovision-cinema.com/events/ev7/index.html

Engström, A., Esbjörnsson, M., & Juhlin, O. (2008). Mobile collaborative live video mixing. In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services* (pp. 157-166).

Amsterdam, The Netherlands: ACM.

Evans, B. (2005). Foundations of a Visual Music. *Computer Music Journal*, *29*(4), 11-24.

Fallman, D. (2003). Design-oriented human-computer interaction. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 225-232). Ft. Lauderdale, Florida, USA: ACM.

Fallman, D. (2007). Why Research-Oriented Design Isn't Design-Oriented Research: On the Tensions Between Design and Research in an Implicit Design Discipline. *Knowledge, Technology, and Policy*, *20*(3), 193-200.

Galeyev, B., & Vanechkina, I. (n.d.). Was Scriabin a Synaesthete? Retrieved December 11, 2007, from http://prometheus.kai.ru/skriab_e.htm

Gaver, B., Bowers, J., Boucher, A., Gellerson, H., Pennington, S., Schmidt, A., Steed, A., et al. (2004). The drift table: designing for ludic engagement. *Extended Abstracts of CHI 2004*, 885-900.

Gaver, B., Bowers, J., Boucher, A., Law, A., Pennington, S., & Villar, N. (2006). The history tablecloth: illuminating domestic activity. In *Proceedings of the 6th conference on Designing Interactive systems* (pp. 199-208). University Park, PA, USA: ACM.

Gerhard, D., Hepting, D., & McKague, M. (2004). Exploration of the correspondence between visual and acoustic parameter spaces. In *Proceedings of the 2004 conference on New interfaces for musical expression* (pp. 96-99). Hamamatsu, Shizuoka, Japan: National University of Singapore.

Gillies, M., & Slater, M. (2005). Non-verbal communication for correlational characters. *Proceedings of the 8th Annual International Workshop on Presence*.

Glass, R. (2006). The Audience Response Tool (ART): The impact of choreographic intention, information and dance expertise on psychological reactions to contemporary dance. *Unpublished doctoral dissertation, MARCS Auditory Laboratories, University of Western Sydney*.

Gonring, G. M. (2009). Executable Cinema: demos, screensavers and videogames as audiovisual formats. *Re: live*, 109.

Gray, W. D., & Salzman, M. C. (1998). Damaged Merchandise? A Review of Experiments That Compare Usability Evaluation Methods. *Human-Computer Interaction*, *13*(3), 203.

Greenberg, I. (2007). *Processing: creative coding and computational art*. friends of ED.

Greenberg, S., & Buxton, B. (2008). Usability evaluation considered harmful (some of the time). In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems* (pp. 111-120). Florence, Italy: ACM.

Grierson, M. S. (2005). *Audiovisual composition* (PhD Thesis). University of Kent at Canterbury.

Harrison, S., Tatar, D., & Sengers, P. (2007). The three paradigms of HCI. In *Alt. Chi. Session at the SIGCHI Conference on Human Factors in Computing Systems. San Jose, California, USA*.

Harvey, J. (1999). The LTDI evaluation cookbook. *Learning Technology Dissemination Initiative, Edinburgh*.

Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In *Mensch & Computer* (pp. 187–196).

Hazlett, R. L. (2006). Measuring emotional valence during interactive experiences:

boys at video game play. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 1023-1026). Montréal, Québec, Canada: ACM.

Higgins, D. (2001). Intermedia. *Leonardo, 34*(1), 49-54.

Hughes, J., O'Brien, J., Rodden, T., Rouncefield, M., & Sommerville, I. (1995). Presenting ethnography in the requirements process. *Proceedings of the Second IEEE International Symposium on Requirements Engineering*, 27.

Hunt, A., & Kirk, R. (2000). Mapping Strategies for Musical Performance. *Trends in Gestural Control of Music, 21*.

Hunt, A., Wanderley, M. M., & Paradis, M. (2002). The importance of parameter mapping in electronic instrument design. In *Proceedings of the 2002 conference on New interfaces for musical expression* (pp. 1-6). Dublin, Ireland: National University of Singapore.

Huxley, A. (1932). *Brave new world*. Vintage Canada.

Ippolito, J., Blais, J., Smith, O. F., Evans, S., & Stormer, N. (2009). New Criteria for New Media. *Leonardo*, *42*(1), 71–75.

Jacko, J. A., & Sears, A. (2003). *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*. Lawrence Erlbaum Associates.

Jacob, R. J., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., & Zigelbaum, J. (2008). Reality-based interaction: a framework for post-WIMP interfaces. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems* (pp. 201-210). Florence, Italy: ACM.

Johnston, W. M., Hanna, J. R. P., & Millar, R. J. (2004). Advances in dataflow programming languages. *ACM Comput. Surv, 36*(1), 1-34.

Jorda, S. (2004). Instruments and Players: Some Thoughts on Digital Lutherie. *Journal of New Music Research*, *33*(3), 321-341.

Jorda, S. (2005). *Digital Lutherie: Crafting musical computers for new musics performance and improvisation*. PhD. dissertation, Universitat Pompeu Fabra, Barcelona.

Kaye, J., & Sengers, P. (2007). The Evolution of Evaluation. *Submitted unpublished manuscript*.

Kolesnik, P., & Wanderley, M. M. (2005). Implementation of the Discrete Hidden Markov Model in Max/MSP Environment. In *FLAIRS Conference* (pp. 68–73).

Larsson, P. (2005). *Virtually hearing, seeing, and being: room acoustics, presence, and audiovisual environments* (PhD Thesis).

Lee, K. M. (2004). Presence, Explicated. *Communication Theory, 14*(1), 27-50.

Levitin, D. J. (2006). *This is your brain on music*. Dutton Publishing.

Libraries \ Processing.org. (n.d.). Retrieved May 6, 2010, from http://www.processing.org/reference/libraries/

Lipscomb, S. D. (2005). The perception of audio-visual composites: Accent structure alignment of simple stimuli. *Selected Reports in Ethnomusicology*, *12*, 37–67.

Livingstone, S. R., & Thompson, W. F. (2009). The emergence of music from the Theory of Mind. *Musicae Scientiae*, *10*, 83–115.

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication, 3*(2), 20.

Louden, K. C. (2003). Programming Languages: Principles and Practice. Brooks/Cole, Thomas Learning.

MacKenzie, I. S., & Buxton, B. (1992). Extending Fitts' law to two-dimensional tasks.

In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 219-226). Monterey, California, United States: ACM.

Maeda, J. (2004). *Creative code*. Thames & Hudson London.

Malloch, J., Sinclair, S., & Wanderley, M. M. (2007). From controller to sound: Tools for collaborative development of digital musical instruments. In *Proceedings of the 2007 International Computer Music Conference* (pp. 65-72). Presented at the International Computer Music Conference, Copenhagen, Denmark.

Mandryk, R. L., & Atkins, M. S. (2007). A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International Journal of Human-Computer Studies*, *65*(4), 329-347.

Mandryk, R. L., Atkins, M. S., & Inkpen, K. M. (2006). A continuous and objective evaluation of emotional experience with interactive play environments. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 1027-1036). Montréal, Québec, Canada: ACM.

Mandryk, R. L., & Inkpen, K. M. (2004). Physiological indicators for the evaluation of co-located collaborative play. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work* (pp. 102-111). Chicago, Illinois, USA: ACM.

Marks, L. E. (1978). *The unity of the senses: Interrelations among the modalities*. Academic Press.

Marshall, M., & Wanderley, M. M. (2006). Evaluation of Sensors as Input Devices for Computer Music Interfaces. *Computer Music Modeling and Retrieval*, 130-139.

Marshall, M. T. (2008). *Physical Interface Design for Digital Musical Instruments* (PhD Thesis). McGill University.

Matthew, W. (1997). OpenSound Control: A New Protocol for Communicating with Sound Synthesizers. *Proceedings of 1997 International Computer Music Conference, ICMA*, 101-104.

Miranda, E. R., & Wanderley, M. M. (2006). *New Digital Musical Instruments: Control And Interaction Beyond the Keyboard*. AR Editions.

Monk, A. F., & Wright, P. C. (1991a). Observations and inventions: new approaches to the study of human-computer interaction. *Interacting with Computers*, *3*(2), 204-216.

Monk, A. F., & Wright, P. C. (1991b). Claims, observations and inventions: analysing the artifact. *SIGCHI Bull.*, *23*(1), 52-54.

Moody, N. (2009). *Ashitaka: An Audiovisual Instrument* (PhD Thesis). University of Glasgow.

Murata, A., & Iwase, H. (2001). Extending Fitts' law to a three-dimensional pointing task. *Human Movement Science*, *20*(6), 791-805.

Nettl, B. (2000). An ethnomusicologist contemplates universals in musical sound and musical culture. *The origins of music*, 463–472.

Newell, A., & Card, S. K. (1985). The Prospects for Psychological Science in Human-Computer Interaction. *Human-Computer Interaction*, *1*(3), 209.

Nicolescu, B. (2005). Transdisciplinarity: past, present and future. In *Congresso Mundial de Transdisciplinaridade* (Vol. 2).

Nort, D. V., Wanderley, M. M., & Depalle, P. (2004). On the choice of mappings based on geometric properties. In *Proceedings of the 2004 conference on New interfaces for musical expression* (pp. 87-91). Hamamatsu, Shizuoka, Japan: National University of Singapore.

Nusseck, M., & Wanderley, M. M. (2009). Music and Motion—How Music-Related

Ancillary Body Movements Contribute to the Experience of Music. *Music Perception*, *26*(4), 335-353.

Occelli, V., Spence, C., & Zampini, M. (2009). The effect of sound intensity on the audiotactile crossmodal dynamic capture effect. *Experimental Brain Research*, *193*(3), 409-419.

Orio, N., Schnell, N., & Wanderley, M. M. (2001). Input devices for musical expression: borrowing tools from HCI. In *Proceedings of the 2001 conference on New interfaces for musical expression* (pp. 1-4). Seattle, Washington: National University of Singapore.

Ox, J., & Britton, D. (2000). The 21st century virtual reality color organ. *IEEE MultiMedia*, *7*(3), 9.

Ox, J., & Keefer, C. (2006, August). On Curating Recent Digital Abstract Visual Music. Retrieved February 24, 2010, from http://www.centerforvisualmusic.org/Ox_Keefer_VM.htm

Ox, J., & Keefer, C. (2008). CVM-Ox-Keefer-Visual Music. Retrieved May 18, 2010, from http://www.centerforvisualmusic.org/Ox_Keefer_VM.htm

Pask, G. (1971). A comment, a case history and a plan. *Cybernetics, Art, and Ideas (Greenwich, CT: New York Graphics Society)*, 76–99.

Peacock, K. (1988). Instruments to Perform Color-Music: Two Centuries of Technological Experimentation. *Leonardo*, *21*(4), 397-406.

Polgár, T. (2000). *Freax: the brief history of the computer demoscene. Volume 1*. CSW Verlag.

Preece, J., Rogers, Y., & Sharp, H. (2002). *Interaction design: beyond human-computer interaction*. Wiley.

Processing Discourse. (n.d.). Retrieved April 12, 2010, from http://processing.org/discourse/yabb2/YaBB.pl

Puckette, M. (1996). Pure Data: another integrated computer music environment. *Proceedings of the Second Intercollege Computer Music Concerts*, 37–41.

Reas, C., & Fry, B. (2007). *Processing: A Programming Handbook for Visual Designers and Artists*. The MIT Press.

Rittel, H. W., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy sciences*, *4*(2), 155–169.

Roads, C. (1996). Early Electronic Music Instruments: Time Line 1899-1950. *Computer Music Journal*, *20*(3), 20-23.

Roads, C., & Mathews, M. (1980). Interview with Max Mathews. *Computer Music Journal*, *4*(4), 15-22.

Rodger, M., Issartel, J., & O'Modhrein, S. (2007). Performer as perceiver: Perceiver as performer. *Proceedings of the ENACTIVE conference, 7*, 237–240.

Sasse, M. A. (1996). Eliciting and Describing Users' Models of Computer Systems. *Unpublished PhD Thesis, School of Computer Science, The University of Birmingham*.

Sawyer, K. (2007). *Group genius: The creative power of collaboration*. Perseus Books Group.

Scheib, V., Engell-Nielsen, T., Lehtinen, S., Haines, E., & Taylor, P. (2002). The demo scene. In *ACM SIGGRAPH 2002 conference abstracts and applications* (pp. 96-97). San Antonio, Texas: ACM.

Sears, A., & Jacko, J. A. (2007). *The Human-computer Interaction Handbook: Fundamentals, Evolving Technologies*. CRC Press.

Sengers, P., & Gaver, B. (2006). Staying open to interpretation: engaging multiple meanings in design and evaluation. In *Proceedings of the 6th conference on*

Designing Interactive systems (pp. 99-108). University Park, PA, USA: ACM.

Shami, N. S., Hancock, J. T., Peter, C., Muller, M., & Mandryk, R. (2008). Measuring affect in hci: going beyond the individual. In *CHI '08 extended abstracts on Human factors in computing systems* (pp. 3901-3904). Florence, Italy: ACM.

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, *14*(1), 147–152.

Shimojo, S., & Shams, L. (2001). Sensory modalities are not separate modalities: plasticity and interactions. *Current Opinion in Neurobiology*, *11*(4), 505-509.

Slater, M. (2003). A note on presence terminology. *Presence connect*, *3*(3).

Slater, M. (2004). How Colorful Was Your Day? Why Questionnaires Cannot Assess Presence in Virtual Environments. *Presence: Teleoperators & Virtual Environments*, *13*(4), 484-493.

Slater, M. (2007). Mel Slater's Presence Blog. Retrieved December 30, 2007, from http://presence-thoughts.blogspot.com/

Slater, M. (2007b). The concept of presence and its measurement. Presented at the PEACH Summer School, Santorini, Greece.

Slater, M., Guger, C., Edlinger, G., Leeb, R., Pfurtscheller, G., Antley, A., Garau, M., et al. (2006). Analysis of Physiological Responses to a Social Situation in an Immersive Virtual Environment. *Presence: Teleoperators & Virtual Environments*, *15*(5), 553-569.

Spinrad, P. (2005). *The VJ book: inspirations and practical advice for live visuals performance*. Feral House, Los Angels, Calif.

Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The Role of Harmonic Expectancy Violations in Musical Emotions: Evidence from Subjective, Physiological, and Neural Responses. *Journal of Cognitive Neuroscience*, *18*(8), 1380-1393.

Steiner, H. (2006). Towards a catalog and software library of mapping methods. In *Proceedings of the 2006 conference on New interfaces for musical expression* (pp. 106-109). Paris, France: IRCAM - Centre Pompidou.

Tasajärvi, L., Schustin, M., Stamnes, B., & Tolonen, A. (2004). *Demoscene: the art of real-time*. Even Lake Studios.

Taylor, R., Torres, D., & Boulanger, P. (2004). Using music to interact with a virtual character. In *Proceedings of the 2005 conference on New interfaces for musical expression* (pp. 220-223). Vancouver, Canada: National University of Singapore.

Touch Designer - Derivative. (n.d.). Retrieved December 11, 2007, from http://www.derivativeinc.com/tools/touchdesigner.asp

Tribe, M., Jana, R., & Grosenick, U. (2006). *New media art*. Taschen.

Tufnell, M., & Crickmay, C. (1990). *Body space image*. Virago.

U B U W E B - Film & Video: Viking Eggeling. (n.d.). Retrieved January 14, 2008, from http://www.ubu.com/film/eggeling.html

Varela, F. J., Thompson, E., & Rosch, E. (1992). *The embodied mind: Cognitive science and human experience*. The MIT Press.

VdmoKstati.com : Mother. (n.d.). Retrieved May 6, 2010, from http://www.vdmokstati.com/category/mother/

Wanderley, M. M. (2002). Mapping Strategies in Real-Time Computer Music. *Organised Sound*, *7*(2), 83-84.

Wanderley, M. M., & Battier, M. (2000). *Trends in Gestural Control of Music*. Ircam.

Wanderley, M. M., & Depalle, P. (2004). Gestural control of sound synthesis. *Proceedings of the IEEE*, *92*(4), 632-644.

Wessel, D., & Wright, M. (2001). Problems and prospects for intimate musical control of computers. In *Proceedings of the 2001 conference on New interfaces for musical expression* (pp. 1-4). Seattle, Washington: National University of Singapore.

Whiting, B., & Whiting, J. (1970). *Methods for Observing and Recording Behaviour: A Handbook of Method*. A Handbook of Cultural Anthropology, New York: Columbia University Press.

Whitney, J. H. (1980). *Digital harmony*. McGraw-Hill.

Wolf, T. V., Rode, J. A., Sussman, J., & Kellogg, W. A. (2006). Dispelling "design" as the black art of CHI. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 521-530). Montréal, Québec, Canada: ACM.

Woody, R. H. (2004). The motivations of exceptional musicians. *Music Educators Journal*, *90*(3), 17.

Wozny, D. R., Beierholm, U. R., & Shams, L. (2008). Human trimodal perception follows optimal statistical inference. *Journal of Vision*, *8*(3), 24.

Xiang, P., & Dubnov, S. (2005). Karaoke system with spatial acoustics estimation for vocal or instrumental removal. In *Proceedings of the International Computer Music Conference* (p. 1102). Presented at the International Computer Music Conference, Barcelona, Spain.

Zimmerman, J., Forlizzi, J., & Evenson, S. (2007). Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 493-502). San Jose, California, USA: ACM.