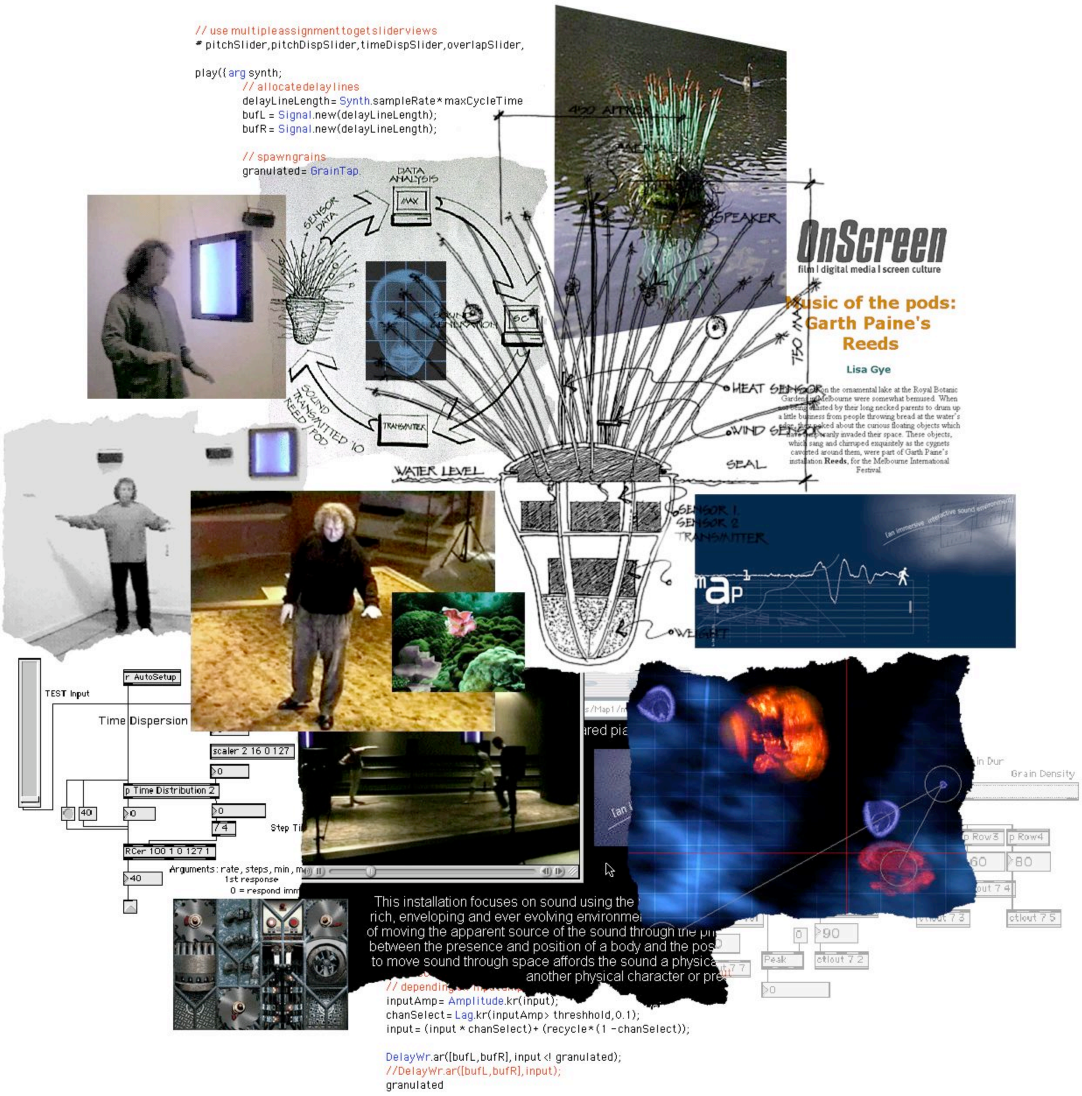
# The Study of Interaction between Human Movement and Unencumbered Immersive Environments





submitted by

**Garth Crispian Paine** BMus


As partial requirement for:

Doctor of Philosophy by Research Project, October, 2002


Centre for Animation and Interactive Media
School of Creative Media
Faculty of Art, Design and Communication
RMIT University, Melbourne, Australia

Supervisors: Sarah Bloomer
David Atkinson

## Declaration

I, Garth Crispian Paine, hereby certify that except where due acknowledgement has been made, this work is mine alone.

The work has not been submitted previously, in whole or in part, to qualify for any other academic award.

The content of this project is the result of work that has been carried out since the official commencement date of the approved research program.

Signed………………………… Dated ……/……/…..

# 1 Introduction

This document outlines six interactive installation works that form a body of research concerned with the development of interactive, responsive installation works that use the gestures of the unencumbered human body as their central activation and control mechanism.

They are therefore an exploration of interactivity, interface technologies and approaches to mapping the sensed data derived from movement in the installation space, onto sound and vision generation schemes. I have conditioned this exploration with a desire to produce art installations; three-dimensional environments that occupy an entire gallery space. The installations were intended to be immersive, and to engage the 'inhabitant' in a direct, visceral and dynamic way. It was intended that the visitor to the installation would require no prior knowledge of the system and, additionally, would require no knowledge of musical practice or the visual arts.

Whilst the installations generate sound and sometimes vision, it was intended that these outputs be engaged with, on an intuitive basis. A great deal of care has therefore been taken in developing interfaces that gather appropriately detailed data about the nature of momentary behaviours. Furthermore, the projects illustrate developments in the mapping of this data on to algorithms that generate audible and visual outcomes that reflect the quality of interaction.

The evolution of approaches, evident between each of the installation works documented here, shows a deliberate movement away from 'triggered' systems to realtime 'streamed' outputs. This evolution also illustrates an intention to gather and refine qualitative data, that is, data that is descriptive of behaviour and movement patterns associated with interactive engagement, rather than the sensing of 'hot spot' or 'collision detection' moments.

In the development of more refined realtime systems, it has also been my intention to remove pre-made or pre-defined content from the installation works, and in so doing, provide as much scope as possible for a unique interactive experience that reflects the fine weights of nuance in ongoing interactions.

The installation works are concerned with gesture. The term gesture is used here both to represent a physical movement, whether choreographed and intentional, or intuitive and exploratory, and the sense of musical or sound based constructions. The gesture of a sound relates to its form, its amplitude envelope, its timbre and the dynamic change in all of these characteristics over time.

I will present much of this document in the first person. Whilst the research is contextualised within the existing body of knowledge and references are made throughout this document to existing literature and previous experiments, these installations are artworks, derived from a personal intention to reflect upon the human condition. The developments illustrated through this series of installations have come about through my desire to:

- increase the intimacy of the interactive experience, and
- remove myself from the role of content creator and dictator of interactive outcomes.

Each installation work takes into consideration the comments made by those who visited the previous installation work, but the way in which these developments are executed and the overall aesthetic and context of the work is driven by personal artistic interests.

To contextualise this work, I will outline three principal areas of interactive practice that have relevance to the objectives of my research:

1. gesture based interactive instruments
2. virtual reality
3. responsive environments

In so doing I will discuss ways in which the work carried out during this project extends or varies the practices of other artists, whose work falls into the above categories

# 2 Context

## *2.1 Extended And Interactive Instruments*

Micro-computers have facilitated an evolving practice of computer mediated art. This research focuses on the role of the computer as a tool for the facilitation and realisation of interactivity in electronic music practice. This chapter outlines a number of approaches to the design of computer mediated electronic instruments relevant to this research. It does so in order to provide the reader with an understanding of the research field, and in so doing, contextualises the projects documented in Chapter Four.

### 2.1.1 Interactive Systems Driven by Acoustic Instruments

Cort Lippe has done a great deal of work at IRCAM[1] using *"the score-following and signal processing capabilities"* of the ISPW[2]. ***Music for Clarinet and ISPW*** (1992) is a good example of Lippe's work with the ISPW. Rowe discusses the processing approach:

> *The signal processing accomplished with the ISPW is quite extensive and includes modules for reverberation, frequency shifting, harmonisation, noise modulation, sampling, filtering, and spatialisation. In fact, all the sounds heard from the computer-part during the course of the piece are transformations of the live clarinet performance. The processing routines listed above can pass signals through a fully connected crossbar, such that the output of any module can be sent to the input of any other.* (Rowe:1993:88)

Lippe states the processing algorithms *"are themselves controlled by every aspect of the clarinet input: the raw clarinet signal, its envelope, the pitch tracker's continuous control information, the direct output of the score follower, and electronic score's event list all contribute to their control"*. (Lippe:1991:2).

Others, like Jean-Claude Risset, took a more direct approach. He used a MIDI keyboard (Yamaha Diskclavier) as input for a system that generates additional parts in realtime, using Cycling 74's Max software. The custom built Max patch *"applies various operations to the human-performed data and then sends out the MIDI commands to play two more hands' worth on the same piano*." (Rowe:1993:82).

[1] IRCAM is the Institu de Recherche et Coordination Acoustique/Musique in Paris
[2] The ISPW, is a hardware and software architecture designed in the early part of the 1990's by the IRCAM Signal Processing Workstation (ISPW) team headed by Eric Lindemann. This workstation, based on NeXT computers with DSP boards, has been widely used in concert situations. Its main software components were the real-time engine FTS (Faster Than Sound) which manages sound processing and a graphical programming environment designed by Miller Puckette (http://www.ircam.fr/index-e.html ).

For instance, "*The section 'Double' plays sequences in response to notes found in the score, and 'Fractals' adds several notes in a quasi octave relationship to each note played"*(Risset 1990).

Robert Rowe (1993) points out that the Risset work, *Duet for One Piano,* uses *"an extensive network of interactive relations…*[however]*, the piece is a completely notated piece of concert music, rather than an improvisation."* The *Duet for One Piano "is usually performance driven but sometimes tracks portions of the human-part more closely, in a score-driven style. In the performance-driven sections, particular pitches on the piano often advance the behaviour of the Max patch from one state to the next. The piece clearly follows a player paradigm; the title alone is evidence of as much."* (Rowe:1993:85)

### 2.1.2 Interactive Systems Influenced by Acoustic Instrument Performance

The trombonist, George Lewis, has designed interactive music systems to take on the role of 'composer' in their own right, performing with no external influence. Within these systems, external sound input causes variations in musical output. Lewis set out to develop a system that would be "*influenced rather than controlled, by material arriving from outside. The system is purely generative, performance driven, and follows a player paradigm. There are no sequences or other pre-composed fragments involved."* (Rowe:1993:79)

The generative materials are based on stored data that specifies scales, note durations, rhythmic patterns, tonality and tempo. Lewis's trombone playing is tracked using a 'listener' module. The characteristics of the trombone music are converted to MIDI data using a Pitch-to-MIDI converter. This MIDI data, then stored, becomes another source for the generation routines, which use probability tables [3] to compose improvised musical performances.

Both the works discussed so far have used acoustic instruments to influence or control the computer performance. Both works extend the acoustic instrument but do not replace them with alternative interfaces.

[3] Probability tables dictate the probability of an outcome based on the input. Five, often used, probability distributions are (a) Linear. (b) Exponential, with a peak in the centre of the table. (c) Bell-shaped with a gentle peak at the centre of the table. (d) Asymetrical. (e) U-Shaped with a broad, gentle low point in the middle of the table. (Roads: 1996)

### 2.1.3 Interactive Systems that Extend Acoustic Instrument Interfaces

Leonello Tarabella's *Imaginary Piano* (CNUCE[4], Italy) extends an existing interface paradigm, building a new interface, fashioned on an existing acoustic instrument, but discarding the original piano. An individual sitting on a piano stool, performs the piano by making gestures in the air in keeping with traditional piano performance practice.

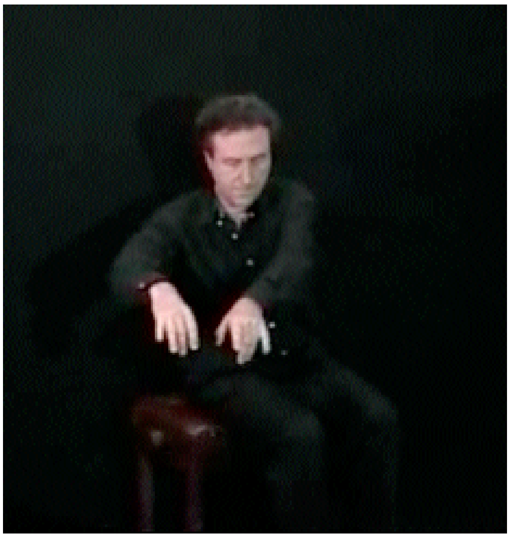


Figure 2.1.3A Tarabella playing the Imaginary piano.

> *The Imaginary piano consists of a realtime image-analysis of video-captured system: here, the interaction 'tools' are, of course, the mere bare hands of a pianist. The pianist is sitting as usual on a piano chair and has in front nothing but the camera, a few metres away, pointing at his hands.*
>
> *There exists an imaginary line at a height where usually the keyboard lays: when a finger, or a hand, crosses the line downward, a specific message (actually a note-on MIDI message) is issued; 'where' the line is crossed states the key number, 'how fast' the line is crossed, states the velocity.* (Tarabella:2000:6)

This interactive work is programmed to perform well-known classical, tonal piano works. Tarabella has achieved considerable nuance in terms of touch (related to the speed of crossing the invisible keyboard line), and good positional resolution. Such a project is useful as a documentation of technical approaches to the tracking of human gesture and movement. It is also striking as a work that is almost completely based on an existing acoustic instrument repertoire and pre-existing performance practice, that defines a gesture repertoire developed for, and refined on another instrument.

One of the principal limitations of this research is the extent to which pre-existing performance practice has limited the scope of both the development of interfaces and new and innovative approaches to performance practice.

---

[4] CNUCE was established as the Computing Centre for the University of Pisa, but is now attached to the National Council of Research of Italy.

### 2.1.4 Interactive Systems that Develop New Interfaces

The Hyperinstruments group at Massachusetts Institute of Technology's (MIT) Media Lab, directed by Tod Machover, is developing some innovative approaches to extended instruments - that is, musical instruments that develop the current acoustic paradigm.

The ***Beatbugs*** by Roberto Aimi and Gili Weinberg illustrates an innovative interface:

> *The Beatbugs are palm-sized digital musical instruments that are designed to provide a formal introduction to mathematical concepts in music through an expressive and rhythmic group experience. Multiple Beatbug players can form an interconnected musical network by synchronizing with each other, trading sounds, and controlling each other's music. Since interaction among players enriches the musical experience, the Beatbugs also encourage collaboration and social play.*[5]

Figure 2.1.4A The Beatbugs.

All the projects associated with the Hyperinstruments group at MIT exhibit an interest in developing existing musical criteria. They explore the generation and analysis of musical material embedded in rhythm, pitch, tonality and other primary characteristics of acoustic music.

The intention for my research work outlined in this exegesis, and illustrated on the accompanying CD-ROM, has been to discover new interfaces and a new performance practice for sound based interactive environments, distinguished by their refusal to be bound by existing music paradigms.

[5] Source: http://www.media.mit.edu/hyperins/projects.html

Sensorband, are an ensemble who have developed new approaches to performance that explore both new interfaces and new approaches to the creation of sound.

> *Sensorband is a trio of musicians using interactive technology. Gestural interfaces - ultrasound, infrared, and bioelectric sensors - become musical instruments. The group, Edwin van der Heide, Zbigniew Karkowski, Atau Tanaka, each soloists on their instruments for over six years, formed Sensorband to create a performance ensemble. Edwin plays the MIDI Conductor; machines worn on his hands that send and receive ultrasound signals, measuring the hands' rotational positions and relative distance. Zbigniew activates his instrument by the movement of his arms in the space around him. This cuts through invisible infrared beams mounted on a scaffolding structure. Atau plays the BioMuse, a system that tracks neural signals (EMG), translating electrical signals from the body into digital data. Together, Sensorband creates a live group dynamic, bringing a visceral physical element to interactive technologies. Sensorband's projects centre around the theme of physicality and human control/discontrol in relation with technology.*[6]

Atau Tanaka uses the BioMuse system as his performance interface.

Figure 2.1.4B  Atau Tanaka using the BioMuse on his arm muscles.

BioMuse is developed and sold by BioControl. Its input source is human bioelectrical signals. It creates a strong connection between the human condition and the generated output. Atau Tanaka uses the BioMuse in live performance with Sensorband. The system is described as follows by its developers, BioControl:

> *In 1992 BioControl introduced the* BioMuse, *a powerful, 8 channel 'biocontroller' that acquires and analyses any type of human bioelectric signal,*

[6] Source: http://www.sensorband.com/root.html

*and then outputs code to control other processor based devices. This unit is intended for use as a development platform for a range of applications.…*

*The BioMuse is a bioelectric signal controller in a rack mountable package (17" X 14" X 3"), which allows users to control computer functions directly from muscle, eye movement, or brainwave signals, bypassing entirely the standard input hardware, such as a keyboard or mouse. It receives data from four main sources of electrical activity in the human body: muscles (EMG signals), eye movements (EOG signals), the heart (EKG signals), and brain waves (EEG signals). These signals are acquired using standard non-invasive transdermal electrodes.*

*The BioMuse amplifies and digitises the biosignals, and then processes the signals using Biocontrol's library of proprietary DSP algorithms. For a given application, the appropriate algorithm then outputs code to control virtually any digitally interfaced device.*[7].

Figure 2.1.4C The BioMuse controller.

BioMuse certainly offers high order interfacing to human responses. However, it requires the wearing of sensors and the associated tethering of the wearer to the control system. It does not therefore provide the freedom of unencumbered interaction which is of critical and artistic interest to this research project. Furthermore, the software API for the BioMuse provides only a limited MIDI interface, restricting its musical application to MIDI note messages and some eight bit MIDI continuous controls.

Atau Tanaka's application of this data, like the Hyperinstruments Group at MIT, has largely reflected the existing tonal music paradigm. The system dictates such an approach by limiting it's MIDI implementation, but the aesthetic decisions are clearly those of the artist. In summary, the interface:

[7] Source: http://www.biocontrol.com/biomuse.htm

1. Offers close dynamic tracking of the human condition through bioelectric signals
2. Requires the user to be encumbered with hardware and cables
3. Provides limited MIDI implementation, suggesting applications in the current musical paradigm.

Another member of Sensorband, Edwin van der Heide, plays the MIDI Conductor with the Hands Interface developed at STEIM in Holland[8]. They send and receive ultrasound signals, measuring the hands' rotational positions and relative distance from each other.

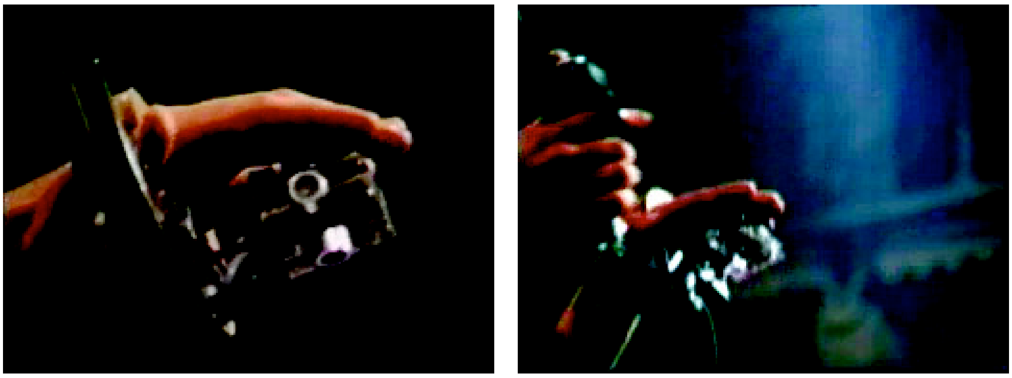


Figure 2.1.4D The Hands Controller used by Edwin van der Heide to control MIDI Conductor.

MIDI conductor is explained as follows by the developer Lorren Stafford

*MIDI Conductor v1.1 is a software performance instrument that accepts MIDI input and maps it to MIDI output in whatever way the user specifies. MIDI Conductor can be mapped in many ways, but here are just a few examples:*

*A single note can trigger an entire phrase (backwards or forwards), select a note randomly from a specified group of notes, play a chord (based on that note), or trigger a looped phrase (backward or forwards).*

*Note velocity (how hard a note is pressed/struck) can be mapped to play a phrase forward/backward or to control parameters such as pitch bend, modulation, etc.*

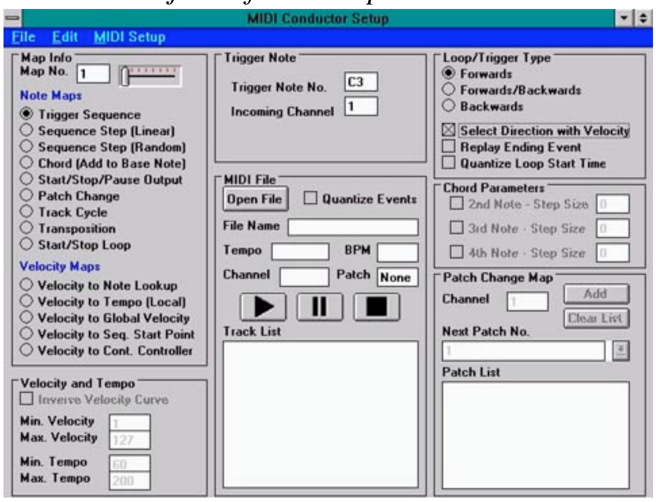



Figure 2.1.4E The MIDI Conductor interface.

*MIDI Conductor is an extremely powerful performance tool in that it allows ONE performer to perform an extremely dense and intricate piece that would not be possible otherwise, (with only one set of hands anyway).*[9]

---

[8] STEIM: http://www.steim.nl/

[9] Source: http://www.winternet.com/~r4c/myself.html

MIDI Conductor is an extension of the concept of triggering. It allows the trigger input to create a pattern of outcomes sequenced to generate as many notes as desired. This approach, while an extension of the single event trigger of the acoustic instrument, is still embedded in the concept of an event causing a fixed outcome.

A pitched note in a musical context is a triggered event. The concept of triggering is in keeping with tonal music practice. It is also common in the current predominant paradigm of interactivity, however , it does not encourage new approaches to musical composition or performance practice for electronic instruments. The BioMuse and MIDI Conductor systems neatly illustrate the different approaches to interface techniques, and establish examples of 'streamed' data input and 'triggered' data input, an important distinction that is dealt with later. It should be noted that whilst the BioMuse system generates streamed data, the data is forced into a MIDI protocol for communication with musical devices, and is thereby restricted to event-based output.

Triggering implies a predetermined, pre-structured output. A trigger starts a temporal sequence, over which the interactor has little or no control.

It is my view that triggering imposes too prescriptive a control on the outcomes of an interactive process. Streamed approaches to interface design, like BioMuse, and video tracking, provide continuous streams of data that reflect ongoing changes in conditions, which, by their very nature, hint at the qualities of engagement the user exhibits. Such a system requires realtime content generation, a requirement that encourages the detailed reflection of nuances in interactive behaviour.

Sensorband have created performance systems that utilise new approaches to gesture based interactive instruments. Their performance systems use both physical limb position and physiological data, such as muscle tension, to create musical output.

Whilst these systems illustrate developments in the gesture repertoire for contemporary musical performance practice, the music making is still very embedded within a tonal music paradigm. It has been a focus of my research to find an approach to sound synthesis that is informed by electroacoustic and acousmatic music. The multi-dimensional morphologies of sound in these genres seem to me to provide a synergy between a human movement gesture and an associated sound response[10].

[10] See Chapter Six for a discussion about the Wishart (1996) concept of Dynamic Morphology and its application to interactive systems.

## 2.2 Virtual Reality

Virtual reality (VR) describes an approach to interactive systems that attempts to create a three-dimensional immersive environment. These systems usually require the wearing of sophisticated head mounted hardware that provides visual and audible feed back to the user.

### 2.2.1 The CAVE As A Realtime Environment System - *Traces*

Virtual-reality systems are typified by the CAVE, developed at the Electronic Visualization Laboratory (EVL) at the University of Illinois at Chicago (UIC).

> *Virtual reality can be defined as the wide-field presentation of computer-generated, multi-sensory information, which tracks a user in realtime. In addition to the more well-known modes of virtual reality, head-mounted displays and binocular omni-oriented monitor (BOOM) displays - the Electronic Visualization Laboratory (EVL) at the University of Illinois at Chicago (UIC) introduced a third mode in 1992: a room constructed of large screens on which graphics are projected onto three walls and the floor.*
>
> *Specifically, the CAVE™ as installed at EVL, is a theatre 10 x 10 x 9 feet, made up of three rear-projected screens for walls and a reflective projection for the floor (the software could support a 6 wall CAVE.) The CAVE library software synchronizes all the devices and calculates the correct perspective for each wall.*[11]

Works using the CAVE system navigate pre-made paths in pre-defined environments.

*Traces* by Simon Penny illustrates an attempt to construct 3D environments in realtime. In discussing *Traces*, Penny comments that, "*The standard interactive paradigm for most immersive projects is constrained to virtual navigation through texture-mapped worlds.*" Penny goes on to say that "*within this paradigm are odd kinaesthetic inconsistencies, for instance: the user can physically turn to face objects of interest, but cannot physically walk toward them. That is to say, the so called 'immersion' is not intuitive*" (Penny:1999:2)

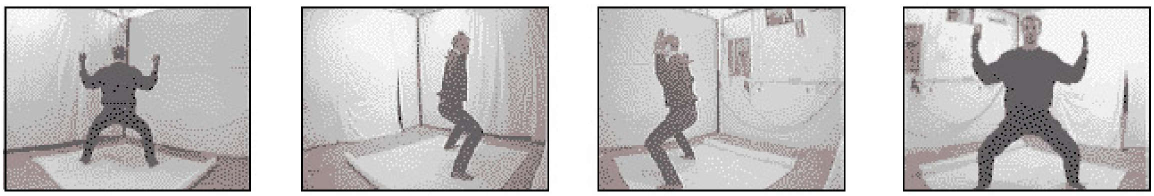


Figure 2.2A Simon Penny in *Traces.*

[11] Source: http://evlweb.eecs.uic.edu/research/template_res_project.php3?indi=161

In this statement Penny acknowledges that the "*The standard interactive paradigm*" is based on pre-determined paths through pre-constructed environments. A paradigm my research sets out to avoid.

Penny's objective in *Traces* was to break free of these 'constraints', to create an environment that was responsive to realtime behaviour. He states:

> *Traces is an unorthodox use of the CAVE, in that there is no 'world' and no virtual navigation. In Traces, the bodily behaviour of the (single) user generates realtime graphics and sound in a limited and static virtual space: a virtual room about double the volume of the physical CAVE. The goal of Traces is precisely not to present a panoptic spectacle for the user, but to turn the attention of the user back onto their own sense of embodiment through time. The movement of the user through the space leaves 'traces': volumetric and spatial-acoustic residues of user movement; which slowly decay. As time progresses, the traces become more active, and in the last stage of the user experience, autonomous entities are spawned by the user, which have complex behaviours of their own*[12].

In working towards these goals, Penny realised that the design criteria for the CAVE system as developed by the University of Illinois at Chicago, had not considered this approach to interactivity. The available software tools were all directed towards the navigation of pre-defined texture maps. Penny comments:

> *Just how different Traces is from a 'regular' CAVE application became clear as we worked on it. Because the system is effectively 'building the world' every time-step based on user movement, the computational processes required for Traces are both demanding and unusual. Many of the standard tools, or major aspects of them, were irrelevant to us. Aside from the CAVE operating system (EVL CAVElib), all of the code in Traces is entirely custom* (ibid)

*Traces* succinctly illustrates a valid step towards an interactive environment that reflects the user's relationship with the environment as the essence of the interactive experience. Penny achieves this by developing a 4-camera machine vision system that divided the space into a three-dimensional grid of 60 x 60 x 45 voxels[13] each a five centimetre cube. If three of the cameras noted activity within a voxel, it was rendered as solid, (see Figure 2.2B). Penny says that *"the use of 'plain' cubes might seem simplistic, [but] note that during the active and passive traces, an average of 2000 voxels are filled and are being displayed."* The demands of such intense realtime computation meant that *"anything more complex than an OpenGL primitive slows the display down unacceptably"*.

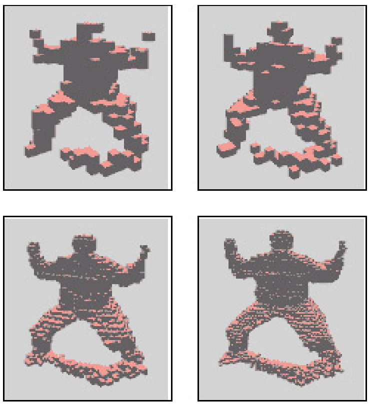


Figure 2.2B *Traces* camera view, Voxel output.

[12] Source: http://www-art.cfa.cmu.edu/Penny/texts/traces/
[13] A Voxel is a three-dimensional volume element, a cube in space.

Penny's use of the CAVE system for *Traces* underlines the predominance of pre-determined interactive experiences. In order to create an experiential work, he had to bypass all the pre-existing development tools, developing entirely custom made interfaces and code. *Traces* remained a single user environment. Limitations in the sensing system, and the CAVE technology constrained the development of a multi-user, or communal experience.

*Traces* does however illustrate two important parallels to my own research.

1. Penny was interested in turning *"the attention of the user back onto their own sense of embodiment through time",* and
2. Penny rejected pre-defined pathways within pre-determined worlds for realtime environments in order to reflect the user's realtime engagement within the system.

### 2.2.2 Engaging the Emotions in Virtual Reality - *Placeholder*

Another innovator in the area of virtual reality is the artist Brenda Laurel. *Placeholder* (1992), produced by Interval Research Corporation and The Banff Centre for The Performing Arts, and directed by Brenda Laurel and Rachel Strickland, explored a new paradigm for multi-person narrative action in virtual environments.

*Placeholder* is relevant to my research, because Laurel and Strickland were particularly interested in how an immersive environment could engage the user on an emotional level. Laurel says:

> *One comes to know a place with all one's senses and by virtue of the actions that one performs there, from an embodied and situated point of view. The mind, observes naturalist Barry Lopez, is a kind of projection within a person of the place which that person inhabits; "Each individual undertakes to order his interior landscape according to the exterior landscape." The environment proceeds to record our presence and actions and the marks that we place there - this is a reciprocal affair.*[14]

[14] Source: http://www.tauzero.com/Brenda_Laurel/Placeholder/Placeholder.html

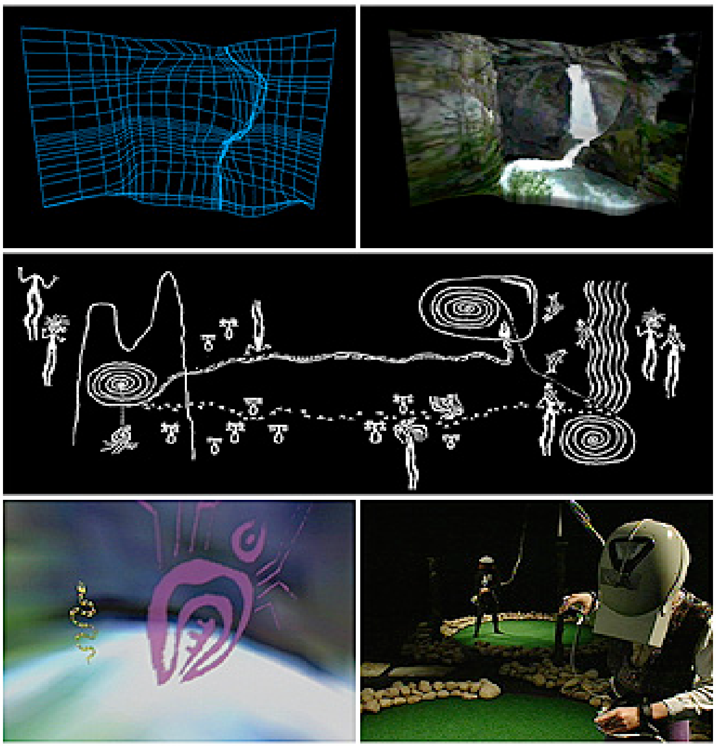


Figure 2.2C  Images from Placeholder.

The cognition of the learned relationship between the human presence and the interactive, responsive environment, has been an important feature of the installation works documented here, and will be discussed at length in Chapter Six. Furthermore, the emotional engagement outlined by Laurel above illustrates a movement from a technology driven activity to a form that engages the human being on a more profound level.

*Placeholder* uses 'active portals' as the mechanism for transition from one virtual world to another. The portals attract the inhabitant with sound from another world. Laurel says:

> *When a person approached, the portal emitted ambient sound from the next environment to which that person might transport. Another person in the same environment might hear the same portal sounds from a distance, but upon approach, might hear the sound of another environment coming through the portal, since the destination of each person was determined individually by random choice.*[15]

The pre-determination of the destination of this journey clearly illustrates the pre-ordained nature of this kind of interactive Virtual reality. *Placeholder* was very innovative in 1992. It pushed beyond much of the technological limitations of the time, and was unique in considering the engagement of emotion. It illustrates a desire for virtual reality systems to free the user from the physical limitations of the human body on an earth where gravity is a predominant force.

[15] Source: http://www.tauzero.com/Brenda_Laurel/Placeholder/Placeholder.html

However, *Placeholder*, like most other virtual reality works, limits the user to the exploration of pre-constructed pathways. Whilst the user has the ability to vary the speed of their journey, and to make choices at points of intersection, the pathways are set, the individual has no scope for the creation of a unique experience directly relating to their current behaviour. The scope for individual cognitive maps, mapping their behaviour onto the evolution of the environment is therefore greatly restricted. The lack of an individualised path, and an associated evolving relationship, must detract from the experience of emotional engagement.

*Placeholder* applied a number of innovations in sensing technology and interface protocols. It employed a *"grippee … a piece of flexible plastic, held in the semicircle defined by thumb and forefinger, which used a sliding variable resistor to measure the distance between the tips of those two fingers, and a Polhemus FastTrak position sensor to define the location and orientation of the hand in three-space."* The *'grippee'* was deployed to both hands of the user allowing them to grasp objects within the virtual worlds. Two user locations were implemented. Each user wore *"display helmets and body sensors"*. There were *"three virtual worlds through which they could independently move. Position sensors (Polhemus 'FastTrak') tracked the three-space position and orientation of the users' heads, both hands, and torsos within a circular stage of about ten feet."* The user took on the avatar of a crow when in *Placeholder*, and moved by flying, literally flapping their wings (arms).

Brenda Laurel summarises the outcomes of *Placeholder* as follows:

> *One was the definition of the medium - thinking about what it was and could be, and how conventions could be used to shape its potential. If VR is to be used as a medium for narrative, dramatic, or playful activity, we should question the appropriateness of conventions derived from computer displays, teleoperations or training simulators. The other issue was the question of the interface - thinking about how people were being sensed and how they were being constrained to behave. Our motto was 'no interface,' expressing our desire to maximize naturalness, to enable the body to act directly in the world, and to minimize distraction and cognitive load.*[16]

While Laurel expresses a desire to *'maximize naturalness'*, and to have no perceivable interface, her subjects were heavily laden with interface technologies. Furthermore, their movement was restricted to stone circles of ten foot diameter and pre-defined navigational paths within the VR world.

My installation work, documented within this exegesis and the accompanying CD-ROM, has specifically set out to develop an approach to interactive, responsive environments that does not require the user to be encumbered with interface technologies. Furthermore a principle impetus has been that the environment the user inhabits is generated in realtime, based on their current movement and behaviour, and as such, embodies the unique individual nuance of gesture of each individual at each point in time.

---

[16] Source: http://www.tauzero.com/Brenda_Laurel/Severed_Heads/CGQ_Placeholder.html

My installations have been conceived of as installation art works. They have been developed from a humanist perspective, with the intention of creating interactive, responsive environments that encourage the user to harness and reflect upon the human condition. They have been designed to offer engagement that reflects the quality of the interaction taking place, not simply a spatial presence. In this sense, the term responsive environments best describes them.

## 2.3 Responsive Environments

Responsive environments are a relatively recent development in interactive systems. They do not require the user to be skilled in the use of an interface, nor do they encumber the human body with sensing or interface technologies. The responsive environment also features realtime generation of audible and visual feedback. The use of realtime synthesis can provide a high level of recognition and response to even the smallest nuance of gesture. Furthermore, there are no predicated paths. The user may explore the environment in any way they wish creating a customised, unique and individual experience.

As the responsive environment does not require the user to be skilled in the interface, or to have direct access to the interface technologies, it affords both individual and communal interaction. It does this by evaluating unique individualised response patterns whilst also including scope for combined activity.

### 2.3.1 Myron Krueger - METAPLAY and VIDEOPLACE

A number of responsive environments that generate both sound and vision exist. These works have a heritage in the early works of Myron Krueger. Krueger developed METAPLAY in 1970. His intention was to develop a piece that contained realtime interaction.

> *Interaction between the participants and the environment was emphasised; the computer was used to facilitate a unique realtime relationship between the artist and the participant. An 8' by 10' rear-projection video screen dominated the gallery. The live video image of the viewer and a computer graphic image drawn by an artist, who was in another building, were superimposed on the screen. Both the viewer and the artist could respond to the resulting image.*
>
> *The artist could draw on the Adage [computer] screen using a data tablet. By using function switches, potentiometers and the teletype keyboard, the picture could be rapidly modified or the mode of drawing itself altered.* (Stiles & Selz:1996:474)

Previously drawn images could be recalled and altered. The existence of the computer made rapid transformations and editing possible. Kruger felt it "*created a far more powerful means of expression than the pencil and paper could provide*." (ibid:475)

Figure 2.3A Krueger interacting with VIDEOPLACE.

Works such as VIDEOPLACE (1975) (which is continuously under development) implement realtime vision generation based on human movement gestures. The ongoing development of VIDEOPLACE illustrates one of the most resolved interface and creation systems of this type. However, it still only works in a two-dimensional plane, and whilst the imagery produced illustrates quite sophisticated texture modelling, the sound produced by the system is very simplistic. It is interesting to note however that the sound production system for this work uses realtime synthesis as distinct from the note based triggering of MIDI sound modules prevalent in so much interactive work.

## 2.4 Summary

Much research, both technological and artistic, is currently being undertaken within the fields of gesture based and extended interactive music systems, virtual reality and responsive environments.

The projects referred to in each category are either exemplary, or, like Simon Penny's *Traces*, are attempts to move the genre towards objectives similar to those reflected in my own research, which has focused on unencumbered interaction, in unrestricted realtime three-dimensional audio visual environments.

None of the projects cited here entirely aligns with my own research objectives. Even the pioneering work of Myron Krueger is presented as a two-dimensional video projection, and employs two-dimensional video sensing. In this sense, the user is never immersed in the work but addresses a two dimensional interface and outcome.

One of my research focuses has been to place equal importance on both the audio and visual systems, designing mappings[17] and algorithms that would be immediately approachable but also provide scope for increased exploration. It was my desire that people would discover new details of interaction, new relationships with the environment upon repeated visits, and that this would serve to encourage further visits. In turn I hoped that this would lead to an ever deepening commitment to the process of interaction and exploration.

This approach required a balance between mappings that were immediately perceivable, and mappings that provided complex, multi-faceted responses.

The extended instruments have provided the best examples of dynamic and evolving mapping strategies. The virtual reality research has provided a platform for the consideration of three-dimensional environments that immerse the user, and through the work of Brenda Laurel, consideration of the importance of emotional engagement in these works. The works of Myron Krueger have provided insights into unencumbered interaction that generate realtime outcomes.

The installation works documented in this exegesis have drawn on each of these practices and have attempted to combine their strengths in the creation of interactive, responsive environments that are immersive and dynamic. A further objective was to design environments that provided the potential for both individual and communal engagement and that presented non-prescriptive, visceral interactive interfaces.

This work has been driven by an artistic intention to reflect upon the human condition. Each of the works encourages the viewer, participant or audience to contemplate various aspects of human life as mediated by their own role within the environment.

In order to achieve these goals the human body is referenced as central to all experience. It is intended that the works stimulate the senses.

A further intention is to contextualise technology based works in the human domain. If a work is to reflect the human condition it must not represent itself primarily as a technical achievement. It must immediately engage the user in a direct and understandable fashion, whilst providing the scope for extended exploration and new discoveries.

Whilst the installations contained in this research do not produce tactile feedback, I will argue that the nature of sound itself provides an immediate and perceivable basis for engagement. As the synthesis output of the works is generated in realtime and responds to small nuances of input, the user perceives a sense of tactile engagement.

Within each of my installation works, the technology is used as a tool to abstract the artistic intention, abstracting both the message and the medium. This is discussed in more detail in chapter three.

---

[17] The term mappings is used to refer to the designed relationships between input data and generative algorithm output, that is, the relationship between perceived system input and output.

The abstraction of the message is achieved through the design of a subtle but engaging interactive experience. The abstraction of the medium is achieved by making the technology invisible to the end user presenting a conduit for experience rather than a technologically mediated activity.

# 3 Developing A Conceptual Framework

## *3.1 Artistic Statement*

As an artist my interest lies in reflecting upon the human condition. There are many aspects of our lives, and many facets of international relationships between differing cultures, governments and philosophical persuasions that we struggle to make the time to consider in our day-to-day lives. Addressing these issues on some level is, however, vitally important to the continuing co-existence of the many disparate parties that form the global community, and are integral to the further development of understanding and insight into the relevance of differences within the global community.

I see art as an appropriate platform for the consideration and expression of these issues. Visual artists, writers, dancers and musicians have explored these issues for centuries with profound results. One only has to look at the prestige attached to the great cultural institutions of the world to see that the product of this artistic endeavour has communal value. It might be argued that the value attached to these institutions is purely financial, that the value placed upon the works they contain is a product of contrary economic principles. Of course, the economics of historical value play a part in their financial value. However, it would be far too cynical to contribute their communal worth solely to the financial market place. If that were the case, the works would be housed in private environments for private enjoyment, not in public institutions, accessible by all.

If it is agreed that these works have a communal worth, then it is also true that artistic endeavour is of value to society. The public institutions attest the worth of the traditional art forms, writing, music, dance and visual arts (painting, sculpture and photography).

In *Theories and Documents of Contemporary Art – A Source Book of Artist's Writings*, Peter Selz discusses the place of the artist after World War II.

> *Characterised by an intensely personal and subjective response by artists to their own feelings, the medium, and the working process, it was an art in which painters and sculptors were engaged in the search for their own identity. In the universe described by existentialists as absurd, the artist carried the romantic quest for the self, and for sincerity and emotional authenticity, into a world of total uncertainty.* (Stiles, K & Selz, P.:1996:10)

He goes on to comment that the *"mechanized mass culture with a plethora of facile and easily accessible public media added to the artist's sense of alienation and the need for individual expression."* (ibid:11)

These pressures of the changing world led to an art that was driven by self expression, an art that focused on *"the personal psychology of the artist rather than on the phenomenological world"*. (ibid:11)
This development grew out of the main thrust of the avant-garde movements of Europe between the world wars.

> *The surrealists' desire for unpremeditated spontaneity held the promise of true creative freedom ... They worked in a realm of ambiguity and communicated through their gestures an aesthetic of incompleteness. At times this exploration turned towards new and unexpected figuration, as in the work of Alberto Giacometti, Jean Dubuffet, and the northern CoBrA[18] Artists, while at others it manifested itself in gestural abstraction. The existential act of making the work was an essential aspect. Even more then previous manifestations of modern art, the dialogue between the maker and the consumer of the work became a necessary element for its completion.* (ibid:12)

The interactive, responsive environment seems a logical and appropriate extension of the desires of the Surrealist artists. It claims the processors of mass media on the one hand, and the sacred space of the art gallery on the other. It combines these apparently opposing forces in an art-work that requires the viewer to be the conduit for the creation of a momentary experience. The viewer takes the role, not only of spectator, but simultaneously of creator; where their behaviour creates the environment, and the environment conditions their behaviour. In so doing they find themselves in a position of contemplation, a position where it is necessary to develop a cognitive map of the relationships between behaviour and environment, between action and reaction, between individual and communal.

Such a position is perhaps unique to the responsive, interactive environment. It reinvents the basis for humanist considerations; considerations that extend from the physical (behavioural, gestural) to the more profound; the spiritual, the sense of connection to environment, to community, to religion, culture, society and all the other structures that form a broader sense of belonging.

A sense of being immersed in the experience became of paramount importance for contemporary artists. Mark Rothko (1903-1970) for instance spoke of his large canvases:

> *I paint very large pictures. I realised that historically the function of painting large pictures is painting something very grandiose and pompous. The reason I paint them, however - I think it applies to other artists I know - is precisely because I want to be very intimate and human. To paint a small picture is to place yourself outside your experience, to look upon an experience as a stereopticon view with a reducing glass. However you paint the large pictures, you are in it. It isn't something you command.* (Rothko:1951:104)

---

[18] CoBrA was a short lived revolutionary northern European artists group of the 1940's. The name is derived from the three capital cities of the countries of it's members; Copenhagen, Brussels and Amsterdam.

Peter Selz comments:

> *… after a period indebted to surrealism and searching for a meaningful mythology, he [Rothko] found his own style of vibrating colour planes. … he explained that his painting needed to be that large in order to place the viewer intimately into the picture space itself. At the end of his life, Rothko completed 14 large paintings from an ecumenical sanctuary in Houston. Eliminating all references to subject matter, but retaining the triptych shapes for his almost monochromatic dark paintings, he succeeded in evoking undefined yet universal meaning and emotion.* (ibid:14)

The sense Rothko had of being immersed in his paintings heralds one of the qualities of an interactive, responsive environment. Sound is perhaps the best medium with which to achieve a sense of immersion. It presents as a homogenised sound field, but may contain points of spatialised information, points of interest that seem separate, dynamically mobile and yet part of the whole. In order to achieve these objectives, the sound generation algorithms must be designed in such a way that the position of the spectator is considered in the sound spatialisation, and that the aesthetic of the sounds reflect an organic and approachable quality.

An artist working in New York at the same time as Rothko, Robert Motherwell, discusses the role of the artist in defining an aesthetic that explores and concentrates emotional experience:

> *The aesthetic is the sine qua non for art: if a work is not aesthetic, it is not art by definition. But in this stage of the creative process, the strictly aesthetic - which is the sensuous aspect of the world - ceases to be the chief end in view. The function of the aesthetic instead becomes that of a medium, a means of getting at the infinite background of feeling in order to condense it into an object of perception. We feel through the senses, and everyone knows that the content of art is feeling; it is the creation of an object for sensing that is the artist's task; and it is the qualities of this object that constitute its felt content. Feelings are just how things feel to us; in the old-fashioned sense of these words, feelings are neither 'objective' nor 'subjective', but both since all 'objects' or 'things' are the result of an interaction between the body - mind and the external world. 'Body - mind' and 'external world' are themselves sharp concepts only for the purpose of critical discourse, and from the standpoint of a stone are perhaps valid but certainly unimportant distinctions. It is natural to rearrange or invent in order to bring about states of feeling that we like, just as a new tenant refurbishes a house.*
>
> *The passions are a kind of thirst, inexorable and intense, for certain feelings are felt states. To find or invent 'objects' (which are, more strictly speaking, relational structures) whose felt quality satisfies the passions - that for me is the activity of the artist, an activity which does not cease even in sleep. No wonder the artist is constantly placing and displacing, relating and rupturing relations; his task is to find a complex of qualities whose feeling is just right -*

> *veering toward the unknown and chaos yet ordered and related in order to be apprehended*. (Motherwell, R:1946:38-39)

While Motherwell is not intending to comment on New Media Art, or more explicitly interactive, responsive environments, I think his commentary is particularly pertinent.

He speaks for instance "*of an infinite background of feeling*" being "*condensed into an object of perception*". It is exactly my intention that the interactive, responsive environments, create "*an object of perception*" by focusing the visitor's attention on the inter-relationship between their behaviour, movement patterns and the quality of environment. The installation work itself "*is the creation of an object of sensing*". Its entire purpose is to reflect the sensitivity of relationship in such a way that the small intimate gestures of each individual are acknowledged; in such a way that every participant is intensely aware of "*an interaction between the body - mind and the external world*". However, as in the case of Rothko's large paintings, it is an "*external world*" in which they are completely immersed.

Motherwell's comment (ibid) that "*it is natural to rearrange or invent in order to bring about states of feeling that we like...*", describes well the desired nature of engagement with my interactive, responsive environments. The inhabitant is engaged in a constant, fluid and dynamic series of streams of engagement, of response, and evolution of environmental qualities. This collection of simultaneous experiences, of "*relational structures*" establishes an architecture of experience "*whose felt quality satisfies the passions*" of those involved in the momentary interaction.

This complex and multi-faceted stream of experience must be designed, or at least established as a potential outcome, by the artist. The artist must set out to establish "*a complex of qualities whose feeling is just right - veering toward the unknown and chaos yet ordered and related in order to be apprehended*." (Stiles, K & Selz, P.:1996: 27)

Each of my projects set out to further this aim and illustrate a line of development that represents:

- greater levels of interaction,
- more simultaneous streams of response,
- dynamic orchestration, and
- multi-faceted, conditional response behaviours that reflect more and more intimately the weight of gesture and the quality of behaviour of each individual engaged with the interactive, responsive environment.

## *3.2 Cybernetics – The Causal Loop*

The relationship between the physical space of an exhibition, the technology used to execute the work, and the human movement and behaviour patterns that form the basis of the engagement is critical in the development of responsive environments.

One of the principal concepts of cybernetics is the causal loop. A closed causal loop is one in which each of the elements contained in the loop act upon the others in a constant and varying fashion to maintain equilibrium. The only influences on a closed causal loop are the elements it contains.

An interactive installation exhibits the qualities of a closed causal loop. Human movement and behaviour patterns act upon the technology. The sensing system collects information about the nature of the human movement, the weight of the gesture, the speed and direction of movement. The data is fed to audio and video algorithms that respond in whatever fashion the artist has designed. The response of the system is presented in the physical space. It takes the form of changing sound patterns and variations in video or animation projections. In this way the technology acts upon the space, altering the architectural and energetic nature of the exhibition area. These changes in the physical space caused an alteration of behaviour by those that inhabit the exhibition. This alteration of behaviour, be it one of excitation or placation, will be driven by an intention to bringing the system to equilibrium, or drive it into an unsteady or chaotic state.

The human response to the alterations in the environment forces the closed causal loop into a further iteration. The input to the technology will be varied, the output of the technology will vary and the physical space will in turn be changed, generating a new and distinct response from those within it.

Interdependent relationships are formed between:

- the technology that mediates the installation response
- the human condition, the behaviour, emotions and relationships that are exhibited in the exhibition space
- the definition and experience of the physical space.

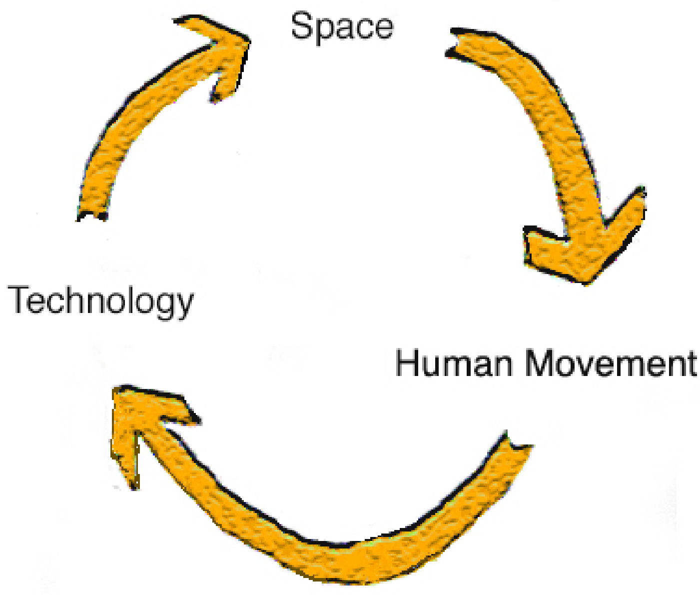



Figure 3.2A a Closed Causal Loop

I have taken particular inspiration from a key proponent of Cybernetics, Norbert Wiener. He conducted extended research into the application of cybernetic principles to the organisation of social systems.
In 1948 Wiener wrote:

> *It is certainly true that the social system is an organisation like the individual, that is bound together by a system of communication, and that it has a dynamic in which circular processors of a feedback nature play an important role.* (Wiener, N.:1948:24)

These words indicate that a well-designed responsive environment may represent patterns of social interaction, and in so doing provide a basis for the consideration of aspects of the human condition.

In 1996 Fritjof Capra wrote

> *... the discovery of feedback as the pattern of life, applicable to organisms and social systems... (helped)... social scientists observe many examples of circular causality implicit in social phenomena,... the dynamics of these phenomena were made explicit in a coherent underlying pattern.* (Capra, F.:1996:62)

So too are the patterns of relationship in an interactive, responsive environment made explicit and coherent through many iterations of the closed causal loop discussed above. Each one rendering the nature of the relationship with greater detail.

In 1998, the virtual reality and interactive installation artists Christa Sommerer and Laurent Mignonneau expressed similar thoughts when discussing the development of the interactive digital arts:

> *. . .the art work . . . is no longer a static object or a pre-defined multiple choice interaction but has become a process-like living system.* (Sommerer, C. and Mignonneau, L:1998:158)

One of the pioneers of interactive arts, the American video, and interactive, responsive environment artist Myron Krueger expresses a similar sentiment when discussing his early interactive video works:

> *In the environment, the participant is confronted with a completely new kind of experience. He is stripped of his informed expectations and forced to deal with the moment in its own terms. He is actively involved, discovering that his limbs have been given new meaning and that he can express himself in new ways. He does not simply admire the work of the artist; he shares in its creation.* (Krueger:1976:84)

Here, Krueger, as an artist, draws the same parallels expressed by Wiener and Capra, as outlined on the previous page. He indicates that the experience of engaging in a responsive environment involves an active engagement with each moment, and that each moment of engagement contributes to the creation of the art work. The participant does not have an option of taking the stance of a detached spectator; they are inherently part of the process, part of the artwork.

## 3.3 Hiding the Technology

The technology involved in all the installation works in this exegesis has been designed to be as transparent as possible. The principle objective has been to immerse the user in the experience of engaging with the work in a way that ensures the technological tools used are not an explicit part of that experience.

I arrived at this view in 1995, prior to commencing this PhD, while presenting a workshop on interactive systems at the Green Mill Dance Festival in Melbourne, Australia. During the presentation, I demonstrated a simple interactive system with the help of dancer Sally Smith.

The demonstration occurred twice:

1. Six touch sensitive floor pads were fixed on the floor and covered by dance flooring so that the pads were invisible to the audience. Sally Smith then performed a dance sequence that generated a musical score based on the triggers generated by the dance choreography.
2. The dance flooring was removed so that the floor triggers were visible to the audience and the dance sequence was repeated.

After each iteration of the dance sequence the audience were quizzed as to what the focus of their experience was. During the first demonstration, the audience remarked that their primary experience was of the dancers movement, and the symbiotic relationship between the choreography and the sound score. During the second iteration of the demonstration, the audience almost entirely focused on the technology. Their focus was to work out which floor pads triggered which sounds and, if the sound-trigger relationship changed, (as it did) why it changed and what conditions generated their expected outcomes.

It was clear that the second test precipitated a focus on the technology at the expense of the dance. This was undesirable in a work that was designed to engage the audience in an artistic experience.

As the works discussed in this exegesis have the two principle goals:

1. artistic expression, and
2. technological development that serves the artistic expression by engaging the audience in an environment that increasingly reflects their individual nuances and provides an immersive sense of engagement.

# 4 The Projects

The projects in this chapter are documented in chronological order.

Each project is outlined by:

1. the original program notes from the first exhibition,
2. the technology developed for the project and the interface design intentions,
3. the mapping of the interface data to the outcomes in the installation environment, and
4. a summary of the strengths and weaknesses of the project and the development it illustrates in response to public feedback for previous projects.

## *4.1 MQM - Moments of a Quiet Mind*

MQM was first exhibited at Linden Gallery, Melbourne, Australia. It generates a changing sound and video environment based on the dynamic of movement of people within it. The following program notes describe the artistic intention of the work.

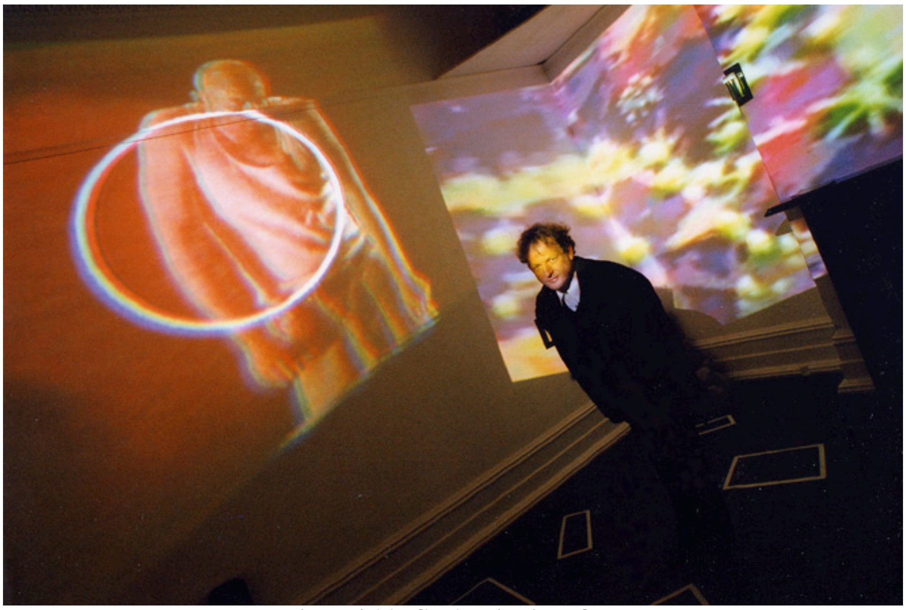


Figure 4.1A Garth Paine in MQM.

### 4.1.1 Program Notes

*MQM is an interactive audio and video virtual environment (IVE) sensitive to human presence and movement patterns. MQM is an immersive experience, a world where individual behaviour and the behaviour of others has a cumulative effect on the immediate environment. The participants are encouraged to work together to create the environment they wish to inhabit.*

*The MQM installation is an 'organic space', developing its own expression of equilibrium when not inhabited. The presence of a <BODY> generates additional responses. The challenge is to explore the environment in order to find a symbiotic existence, cognisant of personal presence within an ever-affected system. The choices are many and varied, being dependent upon the way(s) in which the <BODY(S)> are present. Each experience being unique, a direct relationship between the speed of movement and the quality of environment is established. MQM explores the inter-relationship between people, their actions and their environment. It provides a platform to consider the relationship between the <perceived> and <actual> consequence of behaviour.*

*It is intended that inhabitants of the space make decisions to modify their behaviour in order to create a desirable environment. Further, when there is more than one <BODY>, group decisions need to be made both about acceptable behaviour and the qualities of a preferential environment. MQM considers the effects of events, negotiating a continuing relationship with the terrain surrounding them: the greater an entity's immediate autonomy, the greater the number of feedback loops required for its apparent individualised system.*

*MQM utilises custom software and hardware developed to convert external triggers into data commands that regulate the simultaneous playing of pre-constructed sound bites on six CD-ROM drives attached to an Apple Macintosh computer and video from three PC's projected into the space via five video/data projectors. The MQM software is both sophisticated and intelligent. The input module analyses the nature of the trigger inputs and derived information about:*

*Where the <BODY(S)> are in the installation,*
*How fast the <BODY(S)> are moving within the installation, and*
*In what direction the <BODY(S)> are moving.*

*The software interprets this data, altering the playback response of the system thereby altering the environment as a direct response to the <BODY(S)> behavioural characteristics. MQM creates an integrated relationship between the presence of a <BODY> and the supra-natural environment it inhabits. One cannot exist without the other. This is an important progression from the technology-external controller paradigm of the touch screen and dumb triggering technologies.*

*Each trigger has a number of possible outcomes dependent on the analysis of previous and current activity, challenging the perception of the: <REAL> <IMAGINARY>, <NATURAL> <IMPOSED>*

### 4.1.2 The Technology

Touch-sensitive floor pads and infra-red light beams formed the interface triggers for MQM. The floor pads were dispersed around the gallery space.

The floor pads were grouped in order to provide areas of greater activity and areas of lesser activity. The areas were differentiated using large amounts of riverbed stone to create mounds, and alcoves where exhibition visitors could sit and contemplate the installation.

The floor pad triggers were covered by dance flooring (tarquet[19]), so that the triggers themselves were not visible. The active areas were defined using thin white tape on the back/grey flooring, as can be seen in Figure 4.1A. In this way the active areas were defined, but the technology was hidden. This design decision facilitated the abstraction of the technology from the process of engagement. A primary focus of the design was to concentrate on the experience of engaging with the work, and not on the technology. The work sought to convey an experiential environment. The fact that the environment was mediated using triggering technologies and computers was not intended to be a primary focus of the users' experience or awareness.

The floor pads provided a triggered signal to the computer software. The interface between the triggers and the computer was developed using a standard computer ASCII keyboard, with the addition of circuitry to provide latched timing pulses when a trigger was received, and also to isolate the keyboard circuitry from magnetic interference associated with the long cable runs between the interface and the floor pads. The timing circuits were constructed using 555 integrated circuits and relays. The circuitry was

[19] Tarquet is a vinyl like flooring cover, used by dancers to form a continuous, non-slip floor.

developed specifically for this installation. The additional circuitry and the circuit boards associated with the computer keyboard were housed in a rack mountable casing. The interface circuitry was connected to a Macintosh computer using the Apple Desktop Bus (ADB), and the trigger information was delivered as ASCII key numbers to a Max[20] patch (detailed later).

The benefits of this design solution were:

1. a large number of triggers could be interfaced to the Macintosh computer using existing interface protocols, and
2. each trigger could be identified by an ASCII key number.

It was therefore possible to identify the position within the gallery space that the trigger came from. This information provided relative spatial data, allowing the calculation of speed and direction of movement. A further consideration for using the ADB port was the requirement for multiple MIDI interfaces for communication with the video playback computers. The development of the ADB interface left both the serial ports free for MIDI communication with the video computers.

The principal form of measuring activity within the MQM installation was to monitor the interonset[21] time, the time between any two trigger activations.

The combination of the ASCII key number input from the trigger and the interonset time provided a mechanism for tracking the area of activity and calculating the speed of movement in between trigger points.

One of the disadvantages of using the floor pad and infra-red light beams was that it was very difficult to discern information about behaviour occurring in-between the trigger points. It was, for instance, impossible to accurately define how the person who was creating the current trigger had arrived at that point. Whilst it was possible to group trigger pads into ranges and thereby calculate the likelihood for the current trigger source to be the same person as another trigger source from that same group, this was by-and-large guesswork. The person creating the current trigger may have come from anywhere in the space, and in fact it was possible to move throughout the entire space without triggering any of the floor pads. People who were not making contact with the floor pads could however trigger the infra-red light beams. The light beams were set at about chest height and criss-crossed the space. There were four light beams in total, one dedicated to the entry and exit point to the exhibition, two diagonally crossing the space, and one crossing the far end of the exhibition space.
The audio and video content for MQM consisted of pre-made sound and video files. Two different audio CDs were made for the project, each containing sixty tracks. The sound files were graded in intensity so that track one was a very gentle and lengthy sound sample whilst track sixty was a very intense and relatively short sound. Within the overall file length grading, CD01 contained longer files, and CD02 contained shorter

[20] Max is a high level-programming environment for multimedia and musical applications. See http://www.cycling74.com. The patch is the overall software structure, contained within the highest level window. Sub-patches may exist within the main patch.
[21] The interonset time is the time period between any two events. It is the time from the beginning of the first event to the onset of a second event.

files. This grading of the sound files ensured a degree of overlap of sound in the installation space. It also guaranteed that the time a CD-ROM drive was committed varied widely from trigger to trigger, which increased the likely-hood of a CD-ROM drive being available for the next trigger input.

The design rationale for the variation in track length was as follows:

- The lower track numbers would be played when there was less activity in the space and subsequently fewer triggers received by the system. The frequency of triggers at this stage of evolution would allow for longer periods of time before all seven CD-ROM players were engaged. The inverse was true at higher levels of activity.

- The creation of two CDs, one containing short sounds, and a second containing longer sound files, was developed in order to assist with the availability of playback resources, (short sound files make efficient use of playback resources). At higher levels of activity triggers are received by the system frequently and it is necessary for the audio clips to be short so that the system is not continuously overloaded. If a high trigger rate did not generate a paralleled rate of environmental change, the interactive system would have been perceived as faulty.

- A further design intention for the two CD approach was to establish continuity in the improvised sound score. A mixture of longer and shorter tracks generated an overlapping sound score with smaller changing points of focus.

Early experimentation showed that if all the sound files were of equal length, the sound environment, composed in response to activity within the space, would not illustrate a sense of continuity. The use of only short sound files created a staccato like score with many holes in the audio content. This kind of sound environment reflected only pointillistic activity. The longer sound files created a pedal point, a baseline on to which the other sounds could be added. The longer sounds also implied a multi-faceted, polyphonic response, where some gestures appeared to be considered by the system for a longer period of time and others responded in a brief, succinct manner. The development of polyphonic structures, where voices moved at different rates, was important in generating a sound environment that presented a sufficient level of complexity to engage the visitor in the process of exploration. The sound environment was designed to create an ongoing relationship as distinct from the use of audio in CD-ROM based interactives, when the audio simply indicates an acknowledgement of user activity.

A similar grading system was applied to the video content, although there were only twenty-four video clips.

The audio was delivered into the space from six CD-ROM drives attached to a Macintosh computer. Each CD-ROM drive contained the alternate CD (disk one or disk two). The software provided a series of gates, if the CD-ROM drive was busy the input trigger was passed through the gate to the next CD-ROM drive, and on down the chain until it reached a CD-ROM drive that was available to play the required audio. If all the CD-

ROM drives were busy at the time the trigger was simply discarded. (See the illustrations below)

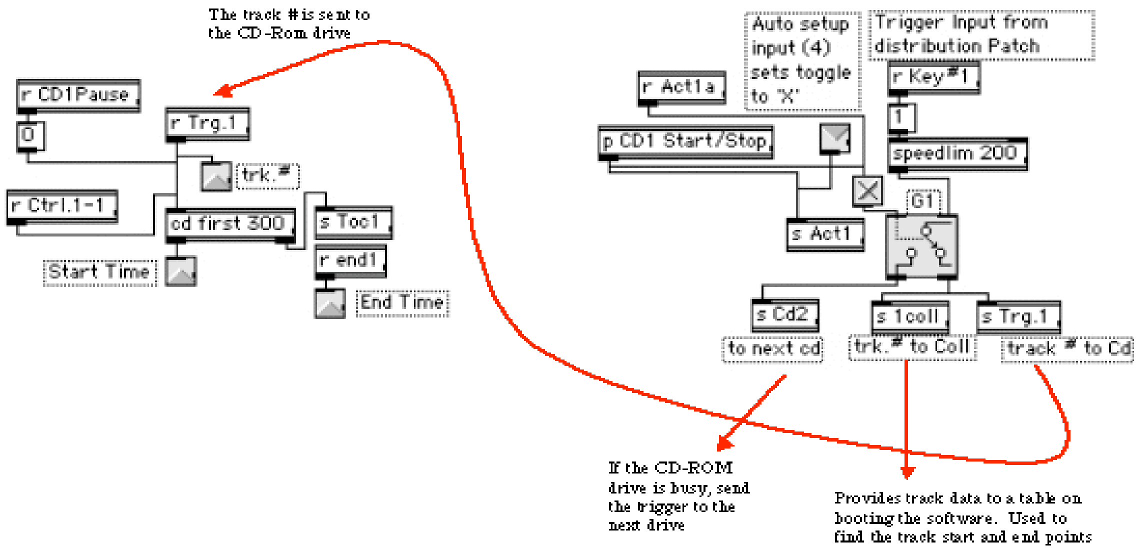



Figure 4.1B The Max patches used to control the audio CD-ROM drives.

The fourteen channels of audio (two per CD-ROM drive) were mixed to eight output channels using a small line-mixer, and fed to loudspeakers in the corners of the room and at the centre points of each wall. The spatialisation of the stereo audio signals were mapped to loudspeakers in diagonally opposite parts of the space. This approach provided a substantial amount of movement in the sound and avoided the situation where holes in the sound field would occur because the CD-ROM drive associated with a particular speaker was inactive.

### 4.1.3 Mapping

The relationship of the interonset time to the audio file number was determined using graphic tables in the Max patch. The relationship curve was adjusted on site to make an allowance for the exhibition space and the positioning of the floor pad triggers. This approach allowed the relationship between activity and sound to be refined in the exhibition space on the basis of experience. The ability to make these refinements proved extremely important to the tactile sense of engagement between gesture and response.

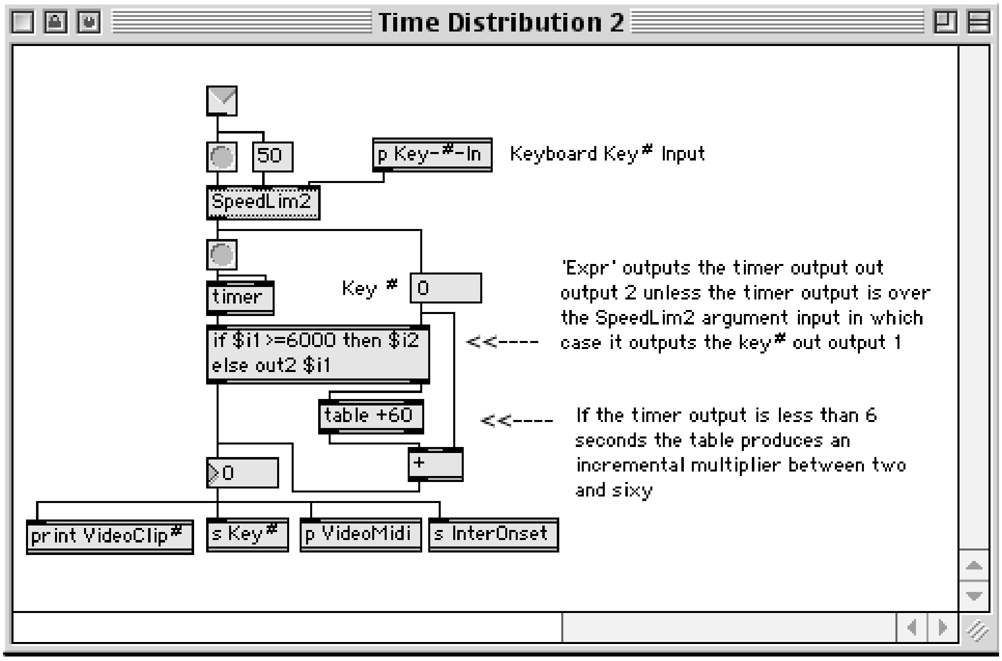



Figure 4.1C The interonset time dispersion patch.

The *timer* object illustrated in the above Max patch calculated the interonset times. If the interonset time exceeded 6000 milliseconds (i.e. no trigger input had been sensed for six seconds), the input key from the trigger was passed through unchanged, otherwise it was directed to the table illustrated in Figure 4.1C. If no activity was sensed in the installation for 90 seconds, the installation was reset to its meditative state.

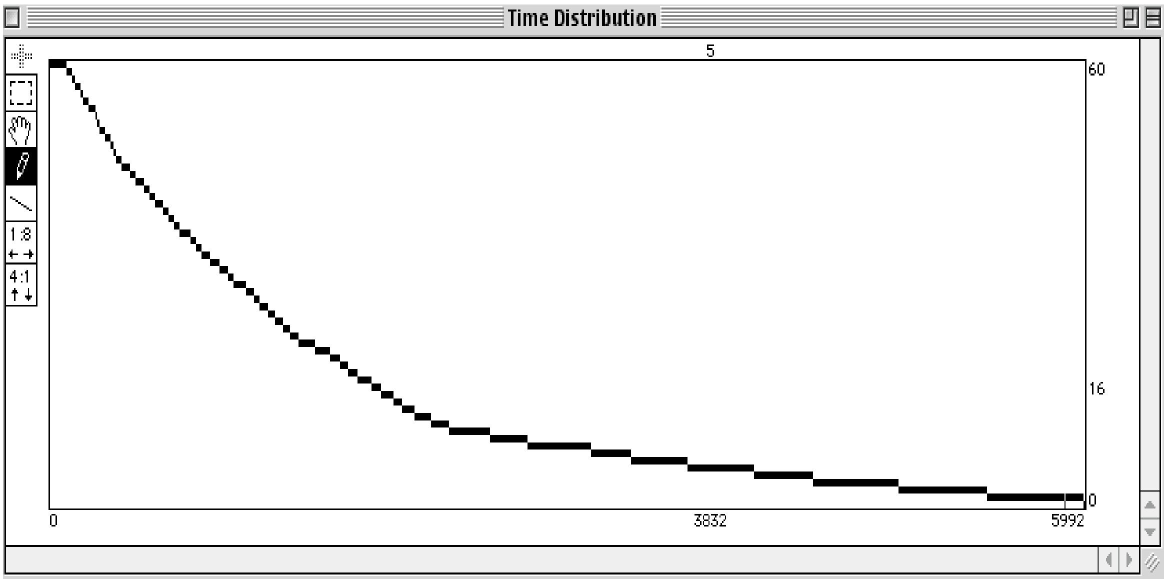



Figure 4.1D The time distribution table.
As interonset times increase (x axis (0 – 6000 milliseconds))
the track number triggered decreases (y axis (0 to 60)).

A similar technique was applied to the mapping of interonset times to video clips. The video was played back using three PCs. Each PC was equipped with a high quality video card, and custom software that related a MIDI note number to a particular video clip stored on its hard drive. The video output was sent to a video projector. Five video projectors were mounted at ceiling height. The three video computer feeds, Video01, Video02 and Video03 were fed to the projectors as illustrated below.

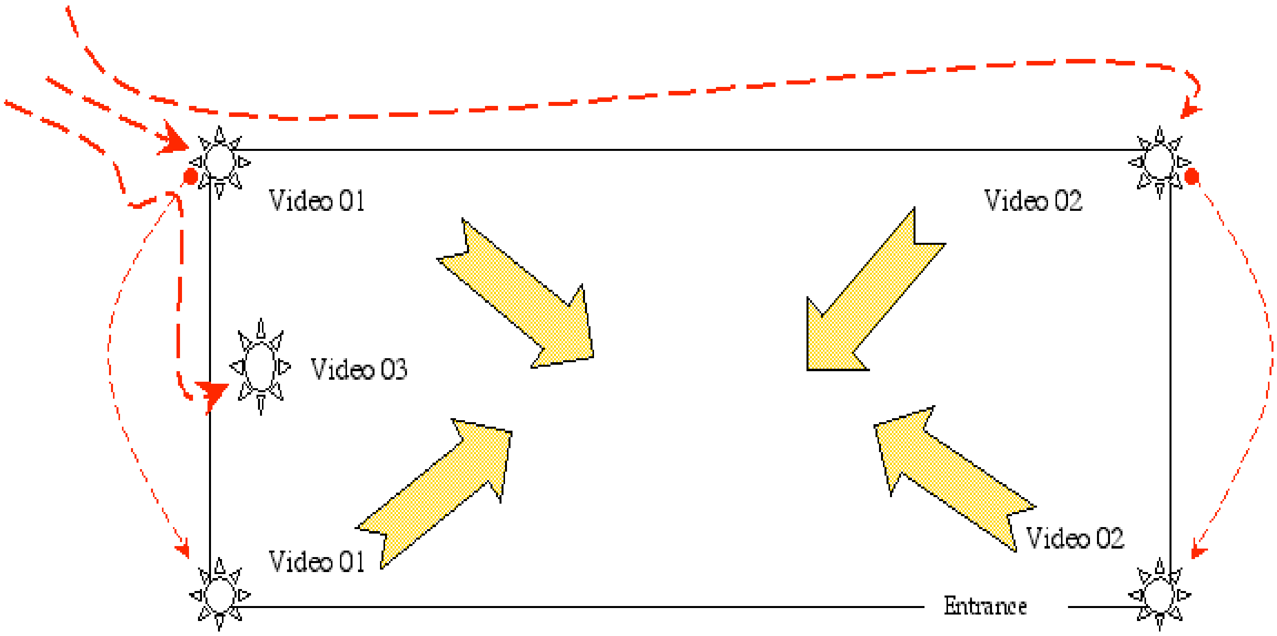



Figure 4.1D The video projection set-up for MQM.

This design approach was applied in order to generate a sense of movement. Some of the video clips rotated across the screen (the gallery walls). The presentation of these video elements in opposite corners of the space, covering large segments of the wall, created an illusion of the physical, architectural space, spinning. When this effect occurred on all four walls simultaneously, the visual movement in the space was very dramatic. The four video projectors focused on the walls were mounted at the diagonally opposite corner of the space to which they were projecting. They were equipped with wide-angle lenses allowing the image to cover substantial amounts of the wall space. The fifth projector was mounted on the floor and projected onto the ceiling. The projection of an image on the ceiling created a three-dimensional immersive environment. The imagery fed to the ceiling projector was different to that used for the wall projections and used different colour gradients from the images projected on the walls. The intention here was to create a sense of height, subjugating the containment characteristics of the architectural space.

### 4.1.4 Summary

Each visitor to the MQM exhibition was able, through a process of exploration, to discover a direct relationship between the dynamic of movement and the intensity of audio and visual material. Once the cognitive mapping had been discovered the exhibition visitor was able to generate a quality of environment that they found preferential. The direct relationship between quality of environment and behavioural patterns presented an interesting point for contemplation:

- should the individual moderate their behaviour patterns in order to generate the environment they preferred? or
- should they accept an environment that is a product of their behaviour patterns?

When there were several bodies in the exhibition the interaction was cumulative. This situation created a communal sense of interaction. If decisions about quality of environment or the moderation of behaviour patterns were to be made they must be made communally.

During the period of the exhibition of MQM at Linden Gallery I observed several interesting communal negotiations. These ranged from some people remaining absolutely still in order to placate an intensity of environment they found displeasing, to a situation where some individuals were moving very dynamically within an otherwise quite group of people in order to generate a more intense environment. On one occasion I observed approximately thirty people in the space all moving very occasionally and observing with intense interest the activity of the other people in the space. Their combined activities generated a gentle ebb and flow of variation that as a group they found pleasing.

MQM represented fairly simple mappings. An increase in the dynamic of behaviour resulted in an increased intensity in the space. As mentioned earlier, a scaling regime was built into the system via a *table object* in Max, which allowed the changes in environment to be mapped non-linearly, so small initial changes in movement resulted in smaller increases in environmental intensity, with an increasing rate of change as the behaviour increased in intensity.

This was done because I found that a logarithmic relationship, as illustrated in Figure 4.1D, was required in order to feel a connection between the intensity of movement and the intensity of the environment. I also found that visitors to the exhibition mostly moved in a range from slow to moderate activity. They wished for higher resolution of mapping (activity to outcome) in this range, but wanted to experience a dynamic escalation of intensity when their activity exceeded the medium activity threshold. Human perception is often logarithmic. Sound pressure (decibels) for instance is measured using a logarithmic curve, and faders on audio equipment implement logarithmic rates of change. The requirement for a logarithmic scaling of activity to outcomes was therefore not unexpected.

MQM achieved unencumbered interactive human engagement within a virtual environment. It represented an environment that responded to the dynamic movement of people within it. In so doing it used relatively linear mappings, where the intensity of the environment increased with the intensity of human activity, and where changes in the quality of environment only occurred in response to human movement within the exhibition space.

The direct relationship between human activity and environmental qualities encouraged participants to consider the relationship between behaviour patterns and quality of environment. The exploration of this relationship was direct and perceivable within MQM.

MQM also raised interesting issues in relation to individual and communal interaction. MQM contained design limitations that treated all trigger inputs as communal. The different groups of trigger pads were not separating out. Additionally, people could not be tracked in between the trigger points. This limitation provided the opportunity for people to remain in the installation without being engaged in the interactive process. This limitation in sensing design reduced the ability of the MQM installation to track direction of movement. Any person coming from any direction within the installation could have caused each trigger input.

These limitations appeared to be inherent in the sensing technologies. The next project attempted to refine the application of trigger sensors to overcome this problem.

## *4.2 GITM – Ghost in the Machine*

The GITM installation was first exhibited at Linden Gallery, Melbourne, Australia. It generates a changing sound and video environment based on the dynamic of movement of people within it.

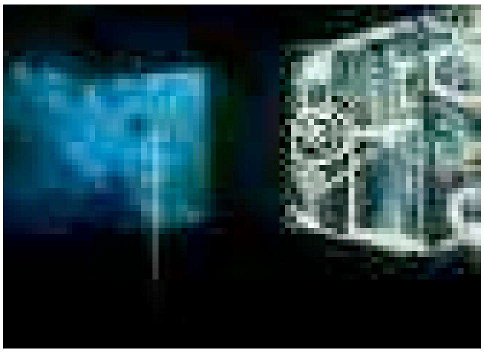


Figure 4.2A A glimpse inside the GITM environment.

In GITM, the installation itself is the controlling force, treating all human presence as an irritant; trying to rid itself of the irritant by evolving through more and more intense environments. The normal interactive process: that is, a system that is responsive to human input, and does nothing unless commanded (i.e. MQM, or any CD-ROM product), is reversed in GITM. Movement within the space, greater than a certain threshold, placates the installation and reverses an automated evolutionary process, encouraging the installation back towards a meditative state of equilibrium that is its resting state, represented by an animated organic single cell environment. Whenever the human movement is acting upon the installation to cause a calming of the environment, a small ghost like human figure is added into the animated environment. The small figure becomes involved in working to remove the more complicated elements of the current environment, and acts as a direct symbolic presence for the human interactive input.

The reversal of the predominant interactive paradigm proved a worthwhile experiment in human-computer interaction.

The following program notes describe the artistic intention of the work.

### 4.2.1 Program Notes

*The GITM installation uses audio and animation/video to create a three-dimensional virtual environment that is sensitive to human presence and movement patterns. Twenty sensors (floor pads and light beams) transmit information about how many people <BODY(S)> are in the room, their position, direction and speed of movement.*

*A Macintosh computer acts as the central controller, collecting the trigger input information, analysing that data and controlling, according to the behavioural patterns within the space, a network of 4 animation computers projecting their output into the environment through 4 LCD projectors. It also controls six CD-ROM drives providing interactive audio. A six-*

*loudspeaker set up is used with the addition of a sub-woofer. The trigger input uses a custom designed ADB interface and proprietary software developed by Garth Paine.*

*Initially, the environment is in a primitive, green, hushed state. Each dimensional plane is constructed of a grid of individual images, which initially creates one image. Each person (<BODY>) entering the room acts as an irritant to the space. As the <BODY(S)> move throughout the space individual image elements morph into other animated elements and evolve before making way for a transition to another world. In this way the inhabited environment evolves through the serene to the mechanised to the chaotic. The environment is soothed by movement. In order to maintain a serene state, the <BODY(S)> must move more and more rapidly. The environment reverts to a serene state when uninhabited.*

*The sound evolves and accentuates the visual environment in symbiosis with the animation, whilst also providing a sense of spatialisation within the room.*

*This installation is designed to reflect on the paradox of the human/machine relationship: we use technology to avoid the dangers of the natural world, but as life becomes less dangerous the human population explodes and we need more technology to sustain us. Thus we are enslaved by the machine and we have to work harder and harder to develop and build new technologies.*

*The development of immersive interactive environments is an attempt to go beyond the screen-based paradigm of so much multimedia, which, I suggest, alienates the audience from the content and the experience.*

### 4.2.2 The Technology

As mentioned above, the sensing system for GITM employed twenty floor pad triggers. The floor pads were positioned to form a grid. The floor pad grid was organised so there was no free space between the pads. Unlike the positioning of the floor pads in MQM, this design approach ensured the system was aware of any horizontal movement across the floor within the installation.

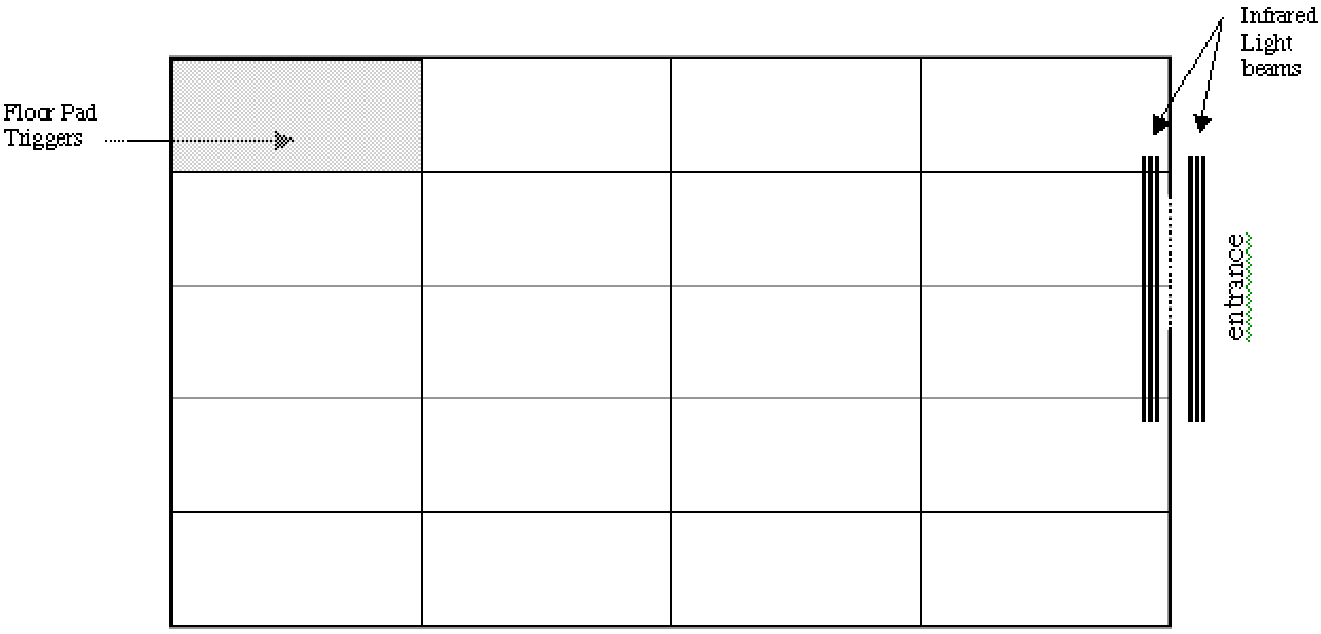



Figure 4.2B The trigger configuration for GITM.

Having inverted the interactive roles, it was important to be able to sense and track all human presence in the installation space. It was imperative that visitors to the installation could not be invisible to the system. Such a scenario would have allowed people to watch the environment return to its state of rest as if no irritant existed. If this were possible the principle of the interaction between the visitor and the installation would have been ambiguous.

When creating an interactive relationship that is unknown to those that engage with the installation, it is vital that the relationship, no matter how complex and multifaceted, not be obscure or indeterminate. The person engaging with the installation must be able to determine the nature of their input and gauge how that input has changed the installation. They must be able to pursue the development of a cognitive map, a mapping of input, in this case gesture, to environmental change, and find that map reliable when tested. Once such a cognitive map exists, the person is able to command the environment, to play the instrument, to make selections in the evolutionary processes of the environment with which they are engaged.

The floor pad triggers were covered by dance flooring (tarquet), as in MQM, so that the triggers themselves were not visible. The active areas were defined using thin white tape on the back-grey flooring. In this way the active areas were defined, but the technology was hidden. This generated a grid of white lines on the floor that represented the conceptual basis of engagement with the installation rather than the pragmatics of the technology.

As previously, the decision to hide the trigger pads facilitated the abstraction of the technology from the process of engagement. A primary focus of the design was to concentrate on the experience of engaging with the work, and not on the technology. The work sought to convey an experiential environment. The fact that the environment was mediated using triggering technologies and computers was not intended to be part of the user's experience or awareness.

Two light beams were also used in this installation. They were used to keep a record of the number of people within the installation. One infrared light beam was placed outside the entrance (1) and one inside the entrance (0). An additional piece of software was written to keep track of the order of these triggers. A trigger sequence of [1, 0] indicated an additional person had entered the installation. A trigger sequence of [0, 1] indicated a person had left the installation. A constant tally of the number of people within the installation space was maintained in a buffer.

It was important to track the number of people in the space in order to scale the threshold at which sensed activity would cause a calming influence on the installation. If this were not done and the interactive threshold were static, it would have been difficult for a small number of visitors to cause a calming of the environment, but inversely very easy for a large number of people within the space to bring the installation back to a resting state. This arrangement would have made the interactive relationship very ambiguous, because as people entered and left the installation, the relationship between the inhabitants and the environment would have continuously altered. The dynamic variation of the interactive threshold ensured the relationship between the activity in the space, whether individuals or communal, and the degree to which human interaction was reflected in the evolution of the environment, was maintained.

### 4.2.3 Mapping

The interactive sound system for GITM was designed along the same lines as MQM, however, as mentioned above, the computer controlled the environment so that changes in the audible and visual activity were controlled centrally. The trigger inputs were used to gauge the level of placation taking place, and as such were detached from directly triggering the audio or visual output. The aesthetics of the imagery and the artistic intention of the work required the composition of new sound material.

As with MQM two CDs were prepared, both containing 60 audio tracks. One CD contained short audio files and the second CD contained audio files of a longer length. The sound files were organised in a graded manner with track one containing meditative material through to intense material on track sixty. The length of the audio files were also graded, with longer files being used for the meditative sounds and short files for the more intense sounds. The design rationale for the variation in track length was as per MQM.

The visual content for GITM was developed using a completely different approach to that applied in MQM. MQM used pre-made video clips, with clips being selected and played back according to activity levels within the space. The visual content for GITM was developed using Macromedia Director. The Macromedia Director projector consisted of many sprited[22] animations. The animation elements were added or removed from the environment according to patterns of behaviour within the installation and not simply momentary activity. This approach facilitated the evolution of the visual environment from one stage to another in gradual steps. In much the same way that the polyphonic qualities of the sound score were carefully developed, this approach to the visual environment provided a similar complexity and flexibility. The gradual evolution of the visual environment provided continuity that was not available in MQM; furthermore, it provided the possibility of introducing small ghost-like human figures into the visual environment as a way of reinforcing the influence human activity was imposing upon the installations evolution.

### 4.2.4 Summary

The developments in both audio and visual interaction in GITM, whilst still using trigger inputs, facilitated a move away from triggered momentary response, to a system that

[22] Sprited animations are individually controllable. Whilst they may occur within a scene, they can be removed, or played at differing speeds, reversed etc. without effecting any other element in the scene in which they reside, or changing the other sprited animations in any way.

created collages of material that more accurately reflected the gesture and movement level of those within the installation space.

When observing interaction in MQM, it became apparent that momentary response to triggered input did not reflect individual nuances of gesture. The only output possibility for the interactive system was a gross response of pre-made material. In this situation the entire clip, be it audio or video was played in full, a kind of glorified disk jockey and video jockey response. The response therefore appeared to be dominated by the initial design decisions. The intention in GITM was to move away from such gross response patterns towards patterns of response that more closely reflected the visitors' experience, and the individuality of their interactive patterns. The development of a polyphonic audio environment, and the use of individually controllable animation elements was a step in this direction.

The reversal of the standard interactive action/response relationship acted as a platform for the consideration of the basis of interaction itself.

- Would exhibition visitors still discover a sense of engagement with the installation?
- Would they be able to develop a cognitive map of response and, if so, would it be reliable when tested?
- Would it remain reliable with different numbers of people in the space?

The answers to these questions are many and varied, differing for each person. Some people could not get past the unexpected reversal of the interactive process to discover their role. Some found the new challenge rewarding and developed a clear understanding of their scope for control, change and interaction. These people found the dynamic nature of the engagement and the fact that the engagement was initiated by the environment to be rewarding because they were offered something immediately, rather than having to initiate the interaction.

This experiment illustrated that people act upon a previously understood paradigm, in this case, the action-response paradigm of screen-based interactivity, and find it hard to move beyond that, even in a new environment (literally in this case).

The reversal of approach did however provide a more stream like sense of interaction. Variations in the environment occurred in small gradations, creating a gradual evolution. This evolution was multi-faceted, and occurred whether human input was sensed or not, in other words, whether the installation was being placated, or trying to rid itself of the written body.

This approach heralded the streamed design of MAP1.

The evolutionary states of the visual environment of GITM:

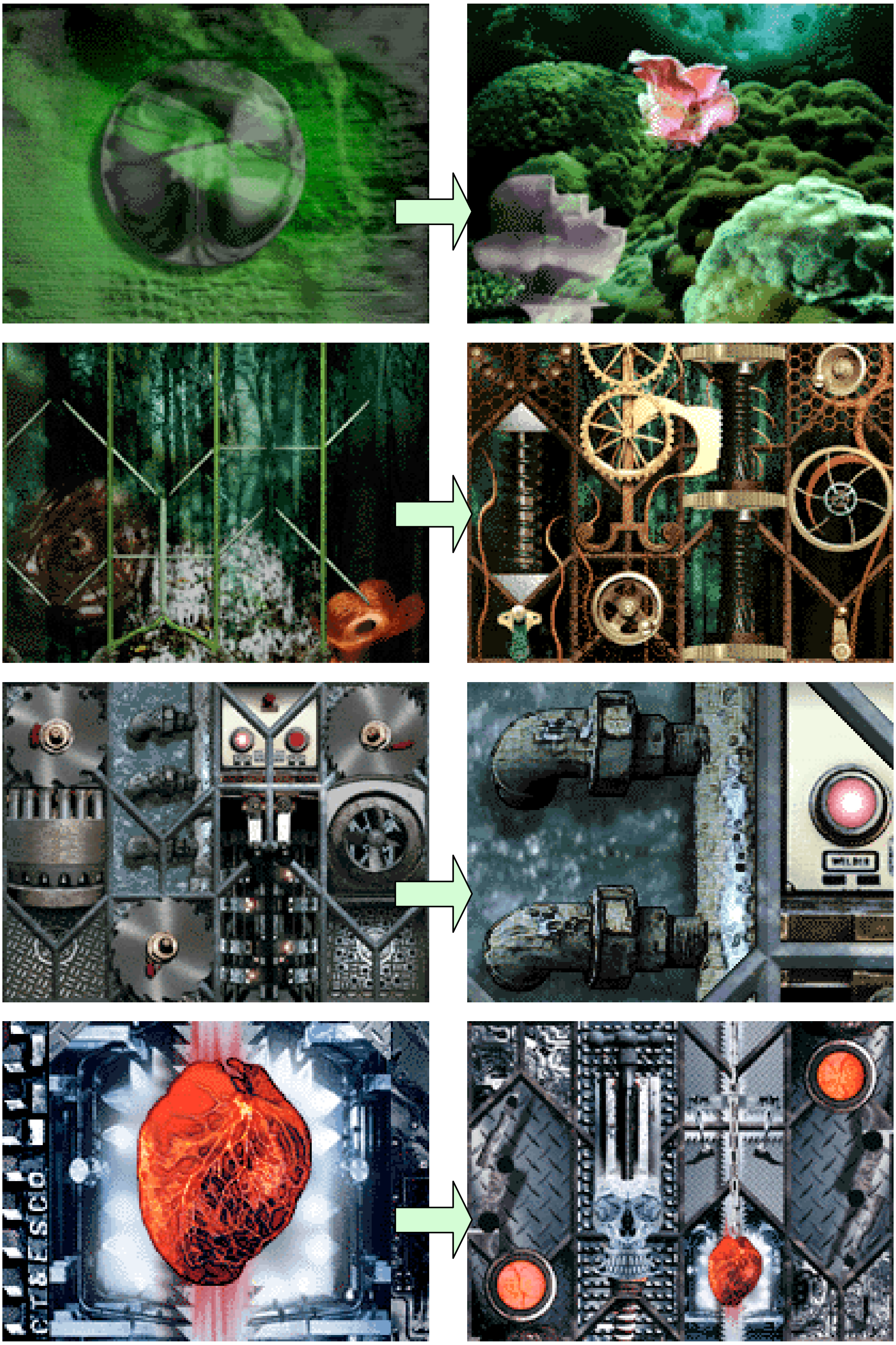


Figure 4.2C Images from the different worlds of GITM, created in Macromedia Director by Rebecca Young.

## *4.3 MAP 1*

MAP1 was first exhibited at Span Galleries, Melbourne, Australia. It explores realtime sound synthesis, driven by video sensing of activity in the exhibition space. MAP1 marks a change from triggered systems to realtime analysis and content creation. It therefore moves away from pre-made content in search of a method for reflecting, in realtime, the smallest nuance of human movement and behaviour within the interactive space. It also reflects a desire to more fully engage the visitor by using the sounds they produce as the source for the synthesised audio environment, something that could only be done in realtime.

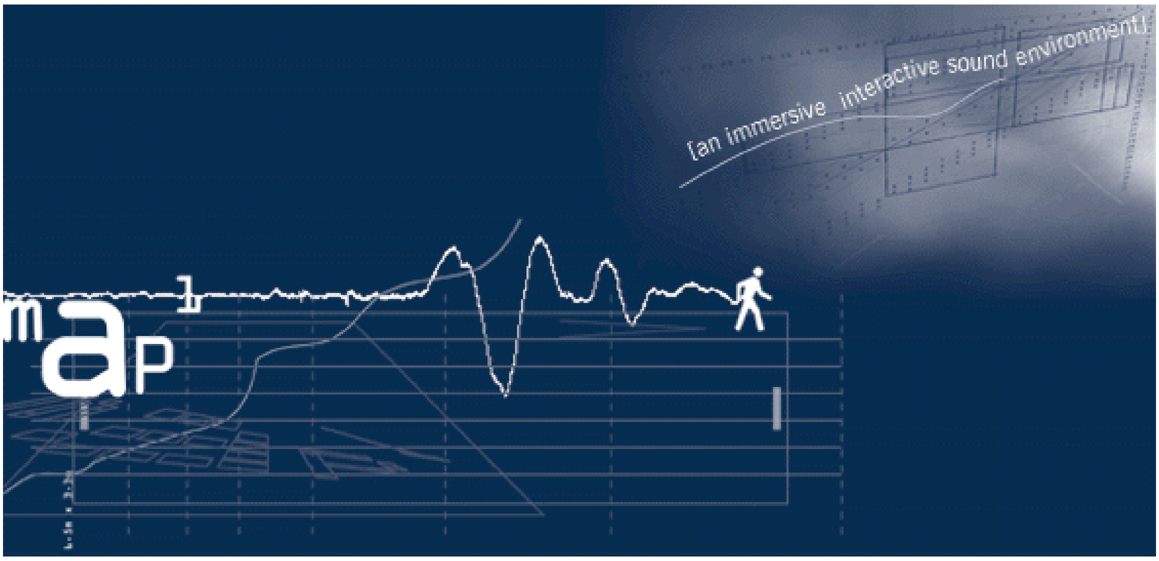



Figure 4.3A The invitation image for MAP1 by Kat Mew.

### 4.3.1 Program Notes

*Map 1 explores ways in which humans develop and re-evaluate cognitive mappings of personal relationships with their environment. Human expectations, frustrations, desires and experiences are usually expressed to the outside world as a physical or aural response.*

□

*Sophisticated sensing technologies in MAP1 observe the physical movements of people within the installation providing information about the mass, dynamic and direction of movement within predefined, independent regions. The movement data is fed to artificial intelligence software that has been trained with a huge range of possible outcomes.*

□

*These control data are used to drive a realtime sound (granular synthesis) process, the source material being the sounds made by those within the installation. James McCartney's SuperCollider*[23] *software is used to do the granulation. Other audio, including spoken text*

---

[23] SuperCollider is an object oriented programming language for sound synthesis, developed by James McCartney (now open source) and available from http://www.audiosynth.com

*elements and prepared piano, are introduced into the space according to visitor movement patterns.*

☐

*This installation focuses on sound using the immersive, fluid and emotive qualities of the medium to generate a rich, enveloping and ever evolving environment. The sounds are made more fluid by the use of a system capable of moving the apparent source of the sound through the physical environment. This creates a dynamic relationship between the presence and position of a body and the position and movement trajectory of the sounds. The ability to move sound through space affords the sound a physicality and, in so doing, the sense that the sound becomes another physical character or presence within the installation.*

*Those within the installation sense a physical interaction with the sound. A wide range of different aural qualities are mapped in qualitative groupings to different regions within the installation space, generating a plethora of aural textures and densities, chosen on the basis of the quality of movement of the body within that region of the exhibition*

*I extend a special thank you to James McCartney for his extensive assistance with the Granular Synthesis algorithms used in this project.*

MAP1 illustrates a dramatic step towards realtime interactive systems. It uses video sensing techniques to track the movement and behaviour patterns of people within the installation space in realtime. GITM illustrated an attempt to avoid non-interactive space by placing 20 floor pad triggers adjacent to each other, covering the entire gallery floor. GITM also displayed developments in polyphonic complexity of the visual and audio content in response to interaction. MAP1 takes the development of interactive sound much further. The application of the realtime video sensing to provide and track complex data about movement within a three dimensional space using the Very Nervous System[24], and the generation of the sound content using realtime synthesis, provide a system that creates fluid responses to changes in interactive states. It used a stream of input data to generate a streamed audio response. This movement from triggered to streamed design criteria caused a profound change in the responsiveness of the system, and the scope for variation.

### 4.3.2 The Technology

GITM and MQM used floor pad triggers for positional information, and infra red light beams to gather height information. They both used trigger inputs, and therefore knew nothing about the behaviour that occurred between trigger points, a characteristic of triggered systems. The application of video sensing techniques in MAP1 provided a continuous stream of information about behaviour and movement patterns anywhere within the cameras view. MAP1 had applied this technique successfully within a two-dimensional sensing schema that provided information about the change in activity between the current and previous video frame within each defined region of the video image. A bi-product of defining regions in the sensing map was that the system also provided positional information.

[24] The Very Nervous System was developed by David Rokeby. See details at http://www3.sympatico.ca/drokeby/vnsII.html

In order to achieve continuous streams of data that represented small nuances of movement in the installation space, I had to find a new way of gathering the data. Video sensing was chosen, because it reports every video frame (30 milliseconds), and provides the flexibility to divide the camera view into a number of independent sensing regions. It also provided qualitative data, describing the dynamic of movement between the current and previous video frames. I chose to work with the Very Nervous System (VNS). It's inventor, David Rokeby describes the VNS as follows:

> *[The VNS is] a device that allows you to extract motion from live video images in realtime.*
>
> *The … VNSIII analyses one video input at 60 fields per second or two synchronized cameras at 30 fields a second each. The image resolution is 128 (horizontal) by 240 (vertical) per field. The grey scale resolution is 8 bits (256 grey levels). For a one camera system, any camera will do, as long as it can produce internal synch. For a two camera set-up you must synchronize the cameras.*
>
> *The basic form of analysis is frame subtraction, meaning that each frame is subtracted from the previous frame, pixel by pixel and the differences are interpreted as motion. You can also compare each frame to a frozen or slowly adapting frame to sense presence. Or you can read the light levels. The resulting information in the normal mode is very dynamic, and muscular, relating most closely to velocity and acceleration. The dynamic range of the system is about 24 bits (motion values from 0 to 16,000,000) although only about 22 bits of that are actually physically possible to generate. The sensitivity is very high. A moving finger can be sensed in an image covering an area of about 12 x 15 feet (varying up or down depending on light levels, contrast etc.) The system does not need a special backdrop of any sort for normal operation. Cameras can be pointed out windows into public spaces. Some higher level analyses may suffer somewhat in complex environments, but the basic operation is very forgiving.*
>
> *The image can be divided into up to 240 regions, each of any shape and size. Each region can be analysed separately, and can trigger and/or modulate a unique event or set of events. (The character of the interaction is completely up to you. Notions of triggering and modulating are a little simplistic) Different maps, and different analysis processes can be applied to each of the two fields per frame (with one camera) or to one field of each camera (with two cameras)*[25].

MAP1 is the first installation work documented here to use video sensing. Before selecting the VNS as the tool of choice, I surveyed all readily available, moderately priced video sensing systems. I would like, at this point, to indicate a few features that made the VNS particularly appropriate for MAP1.

---

[25] Source Rokeby, D.: http://www3.sympatico.ca/drokeby/vnsII.html

1. The VNS adjusts automatically to variations in background light levels. The differencing technique employed in the VNS allows for changes over time in the light level of each pixel. Because background light level changes are generally small, and gradual, they are not reported as dynamic activity. By contrast, BigEye[26] from STEIM in Amsterdam and the VideoIn[27] object for Max, require a reference frame be stored at the beginning of each sensing session. Each subsequent video frame is compared to the reference frame. This technique indicates all changes in light levels from the original reference frame as activity in the space, and does not therefore compensate for changes in ambient light levels.

2. Each standard video frame (a nominal 1/30$^{th}$ of a second in NTSC format video and 1/25$^{th}$ of a second (25 frames per second) in PAL format video) contains two interlaced fields. This is a standard part of the Video protocol, employed as an error correction technique. The VNS provides the option to split the fields and apply separate analysis maps to each. This function facilitates the simultaneous analysis of two distinct maps using one camera. In MAP1, separate sensing maps are applied as follows:

   - The first field is divided into 64 regions that cover the entire floor area and are used to control the realtime granular synthesis algorithm.
   - The second video field is mapped into four regions, one in each corner, and is used to sense people near the loudspeakers.

3. The VNS is a self-contained video digital signal processor. It places very little load on the host computer, leaving the host computer's CPU free to perform all the calculations required for mapping the sensed video data onto sound and vision algorithms. The VNS connects to the host computer using the SCSI protocol. Both BigEye and VideoIn use the host computers CPU, something that is relatively easily accommodated now, but required the most powerful Macintosh computer available at the time MAP1 was first exhibited. Even with current Macintosh G4 computers, host based software video sensing can consume up to twenty five percent of the CPU processor cycles, greatly restricting the scope for realtime synthesis on the same machine.

The VNS is the only video sensing technology in its price range that offers the facility to apply more than one map to each video frame, an invaluable facility in MAP1. It is also the only video sensing device that allows two video cameras to be synchronised to provide a three-dimensional sensing environment, locking the horizontal regions to vertical regions.

---

[26] BigEye - http://www.steim.nl/bigeye.html
[27] VideoIn, now substantially updated to Cyclops - http://www.ericsinger.com

MAP1 uses a single CCTV[28] video camera, placed in the ceiling of the exhibition space, looking straight down, covering the entire floor space of the gallery. Eight loudspeakers were placed:

- four in each corner, carrying sounds triggered by close proximity to the corner, and
- four loudspeakers, equidistant from each corner of the gallery, carrying the granular synthesis output.

### 4.3.3 Mapping

The interactive environment space was divided into two camera maps. As outlined above, the first video map divided the floor space of the gallery into 64 regions. The 64 regions were organised as four rows of 16 regions.

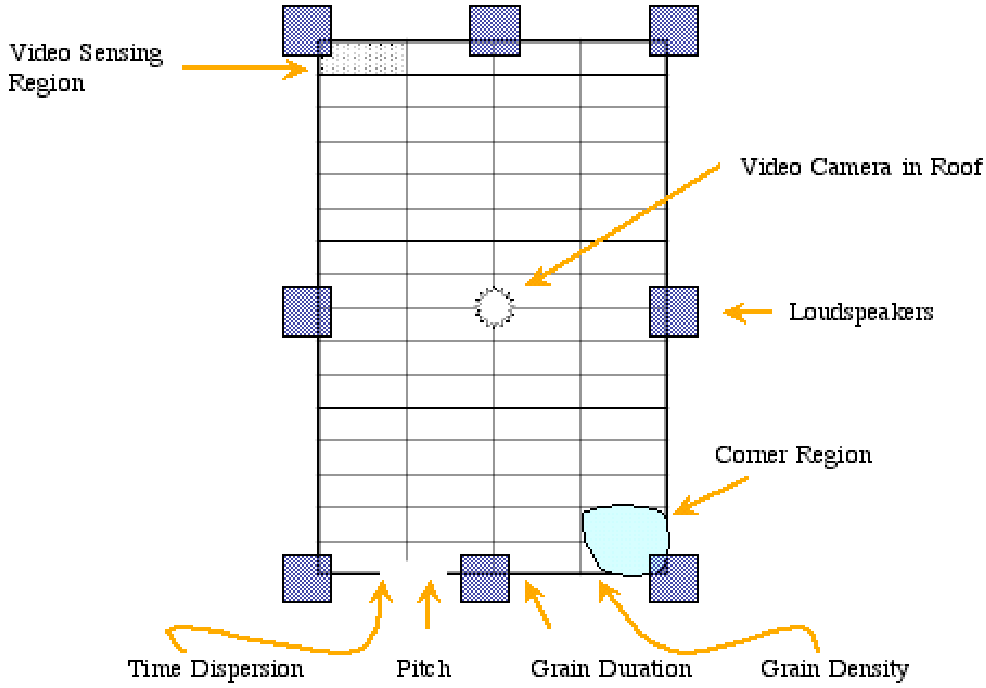



Figure 4.3B MAP1 video sensing layout and loudspeaker placement.

In addition, a second video map was defined to look for activity in close proximity to the corner loudspeakers.

The corner regions were used to trigger additional intimate sound material when the interacting body was close to the corner loudspeaker. This sound material was delivered

[28] CCTV stands for Closed Circuit Television. CCTV cameras are primarily used in security and monitoring situations. The camera bodies are very compact, the video quality is adequate for video sensing applications, and the cameras take a variety of lenses making it possible to recreate the original setup in different gallery spaces.

using two Roland AR100[29] flash card audio players. The AR100 can be programmed to play the left and right channels independently. One group of tracks (odd numbered tracks) was allocated to the left channel and a separate group of tracks (even numbered tracks) to the right channel. Each channel was triggered independently using MIDI note numbers. The channels were allocated to a particular loudspeaker within the associated sensing region.

The VNS reports movement activity every 30 milliseconds. The data being delivered in the form of an integer array, one integer per region, in order, from the top left of the video image to the bottom right.

| | | | |
|---|---|---|---|
| 1 | 2 | 3 | 4 |
| 5 | 6 | 7 | 8 |
| 9 | 10 | 11 | 12 |
| 13 | 14 | 15 | 16 |
| 17 | 18 | 19 | 20 |
| 21 | 22 | 23 | 24 |
| 25 | 26 | 27 | 28 |
| 29 | 30 | 31 | 32 |
| 33 | 34 | 35 | 36 |
| 37 | 38 | 39 | 40 |
| 41 | 42 | 43 | 44 |
| 45 | 46 | 47 | 48 |
| 49 | 50 | 51 | 52 |
| 53 | 54 | 55 | 56 |
| 57 | 58 | 59 | 60 |
| 61 | 62 | 63 | 64 |

Figure 4.3C The VNS region map for field one of the video signal.

The integer is generated by the VNS by calculating the difference in light levels per pixel between the past video frame and the current one. These *difference figures* are then averaged for each defined region. This process generates a large number when the change in activity is substantial and a smaller number in accordance with less movement. The integer reflects the velocity, acceleration and size of the moving object.

It is possible to define the position and direction of movement by tracking the pattern of change of region numbers showing activity. Similarly, the inertia of the moving object can be determined by analysing the rate of acceleration through contiguous active regions. This data is derived rather than provided directly by the VNS.

Four rows of 16 regions each, acted as very large controllers (sliders), as indicated in Figure 4.3B. They were allocated to the following variables within the granular synthesis patch running in SuperCollider[30] on a Macintosh:

[29] Roland AR100 audio player information at http://www.roland.co.jp/worldwide/products/AUDIO/audio_recorders/AR-100.html
[30] SuperCollider is a realtime sound synthesis programming environment authored by James McCartney. See http://www.audiosynth.com

| Controller Name | MIDI Continuous Controller & Channel Numbers |
|---|---|
| Time Dispersion | MIDI Channel # 1, Controller # 7 |
| Pitch | MIDI Channel # 3, Controller # 7 |
| Grain Duration | MIDI Channel # 4, Controller # 7 |
| Grain Overlap | MIDI Channel # 5, Controller # 7 |
| Grain Amplitude | MIDI Channel # 7, Controller # 7 |
| Pitch Dispersion | MIDI Channel # 2, Controller # 7 |

Figure 4.3D MIDI Continuous Controller Allocation in SuperCollider.

The setting for each of these variables is determined either by the region of greatest current activity, or by the last point of activity within its associated regions. A person moving through the space will enter the regions associated with each of these controllers at different points. The controller will move towards the point at which the greatest activity is occurring within that controller row, at a rate determined by the dynamic of sensed activity. The value of the controller will remain at the point of last activity within the row, until new activity is sensed.

This mapping allows the MAP1 installation to be played as an instrument. The body(s) within the space can choose when to alter each of the variables associated with the floor rows, and via the dynamic of their activity, can adjust the rate at which the controllers move to the current body(s) location, the controllers destination.

During an exhibition of MAP1 at the Australasian Computer Music Conference in Wellington, New Zealand in 1999, I witnessed several people in the installation passing parameters from one to another. They did this by placing two people per row, one at each end, and attracting the focus of the controller by varying their dynamic of movement. When either of the participants within the row generated a greater dynamic of activity than their collaborator, the variable would start to move in their direction at a rate determined by the sensed dynamic of movement.

This behaviour allowed several people to ‘perform’ the space simultaneously. During this occasion in New Zealand, nine people performed the installation, one person stationary under the microphone creating sounds used for the granular synthesis process, and two people at either end of each of the four rows.

Two further controlling factors are derived from the video sensing data. The global activity level varies the overall amplitude of the sound environment. The global activity being a cumulative measurement of all movement within the installation. The amplitude of the audio signal increases with an increase in activity level.

The field of greatest activity within the entire installation determines the *grain dispersion*. The *grain dispersion* is therefore constantly varying over a range of 64 (0 - 63). This variable allows for subtle changes in the sound in a manner that is not directly related to specific, conscious movement activity within the installation.

The mapping approach outlined above facilitated the simultaneous control of multiple independent variables. This functionality provided a much broader scope for variation in the sound content and thereby an equally extended range of response to interaction. The provision of multiple independent variables also provided scope for multiple independent users to interact simultaneously with the space.

A further outcome of this mapping approach was that it was possible to create sounds through a combination of the control variables. The users of the installation could experiment with the variation of a single control input without changing the others, and then expand their experimentation to include the addition of other variables.

A further development in MAP1 was that the primary sound source is derived from the utterances of those within the exhibition.

Granular synthesis algorithms usually process an incoming audio stream. In MAP1 the granulation input is an audio buffer of approximately two seconds duration (see Figure 4.3I).

When the granular synthesis algorithm is instantiated, two delay lines are generated, one for each side of the stereo audio signal path. The delay lines are the same length as the audio buffers. The audio buffers are refreshed every cycle. The source for the audio buffer is selected based on the amplitude of the incoming signal. An amplitude follower analysers the amplitude of an input signal from a microphone in the exhibition space. Whilst the audio input is under a specified threshold, the audio buffer is filled from the delay line. If the audio input is equal to, or exceeds the threshold the audio input is recorded into the audio buffer. The granular synthesis algorithm is documented in Figure 4.3I.

This technique presented a dynamic buffer as the input to the granulation algorithm. In practice it meant that people could explore the installation, using the sounds others had entered before them. When they felt sufficiently confident with how they could influence the interactive process, they could enter their own sound material as source for the granulation process.

The way in which sounds were entered into this buffer presented substantial scope for individualised experience. For instance, once exhibition visitors realised that only sounds over a certain threshold were recorded into the buffer, they could experiment with the temporal structure of both the buffer, and the subsequent output, by adding sounds at different points in the buffer. They could in effect create rhythmic patterns, and layers of input by adding material on each iteration of the buffer cycle.

As a musical instrument, a musical composition, or an interactive artwork, MAP1 exhibited no pre-defined musical/sonic/compositional form.

One of the focuses for the development of this work was to move away from pre-defined interactive paths. The use of realtime synthesis as the primary environmental quality, provided a fluid and viscous evolutionary form, that was determined by the interactive behaviour of those within the installation in realtime.

The provision for exhibition visitors to add his or her own audio source material allowed them to personalise, and subsequently change the inherent nature of the entire environment.

These developments removed the limitations of fixed interactive paths, and whilst establishing an overall aesthetic (through the synthesis approach), the installation provided very little pre-made content (the exception being the audio material triggered near the corner loudspeakers), and therefore a high level of separation of the final sound environment from that determined by the artist.  The freeing up of the determination of the nature of the experience within the installation was an intentional step to move the outcome decisions more within the visitors control, thereby shifting ownership of the experience onto the user, and shifting the artist away from content determination.

The MAX software developed for MAP1 is shown in Figure 4.3E.

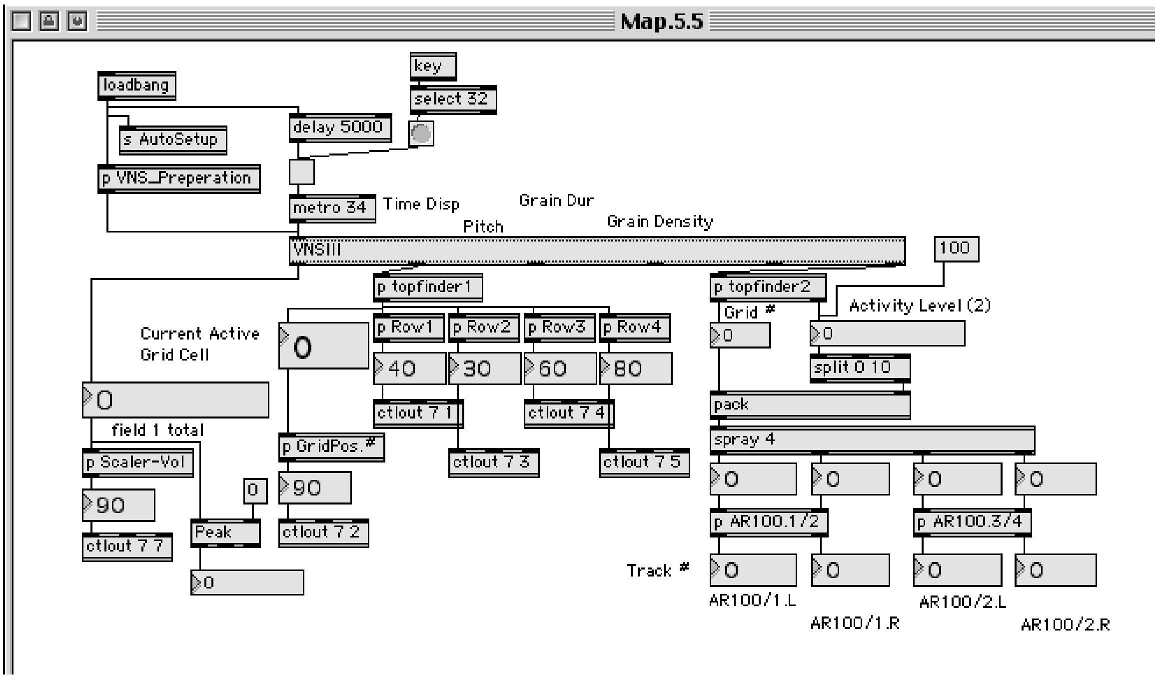



Figure 4.3E.  The top level Max patch for MAP1.

Each row (defined by 16 regions within the field one video map) is controlled by the `Row` sub-patch (see Figure 4.3E).  The `Row` patch looks for activity in regions associated with the specified row.  If activity occurs in one of the specified regions, the region number is fed to a `Change` object that checks to see that the region number contains new activity.  If the region contains new activity, the region number is passed as the destination to the `Scaler-1` sub-patch (see Figure 4.3F).

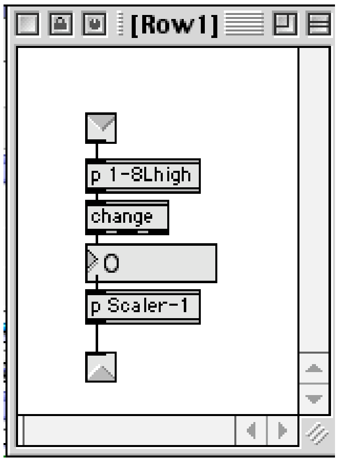



Figure 4.3F The Row patch.

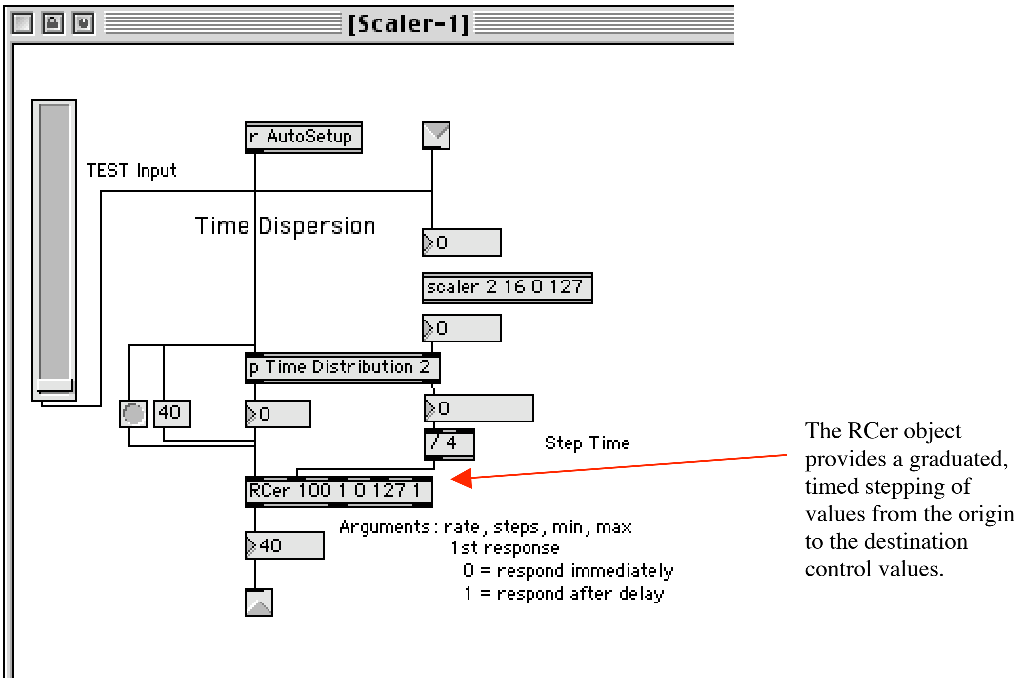



Figure 4.3G The Scaler patch used to manage the rate of change of control output to SuperCollider.

The `Scaler` patch manages the rate of change of the control output to SuperCollider. In order to do so, the 0-15 range is scaled to a full 0-127 MIDI continuous controller range. The scaling up of the 0-15 range is done so that the `RCer` object moves through finer gradations in its journey from set point to new destination. This achieves a smoother transition, and also provides more scope for variations in rate of change and step size variables.

The `Scaler` patch, illustrated in Figure 4.3G, uses a time distribution table to scale the interonset [31] time between sensed activity in the regions allocated to a row, and combined this with the sensed activity level to generate a rate of change for the `RCer` object. It also

[31] The interonset time is the time period between any two events. It is the time from the beginning of the first event to the onset of a second event.

sets the step size argument being fed to the `RCer` object. The `RCer` object illustrated in Figure 4.3G, generates a smooth transition for one destination to another at a rate and at a step size determined by input argument variables. All variables can be changed dynamically.

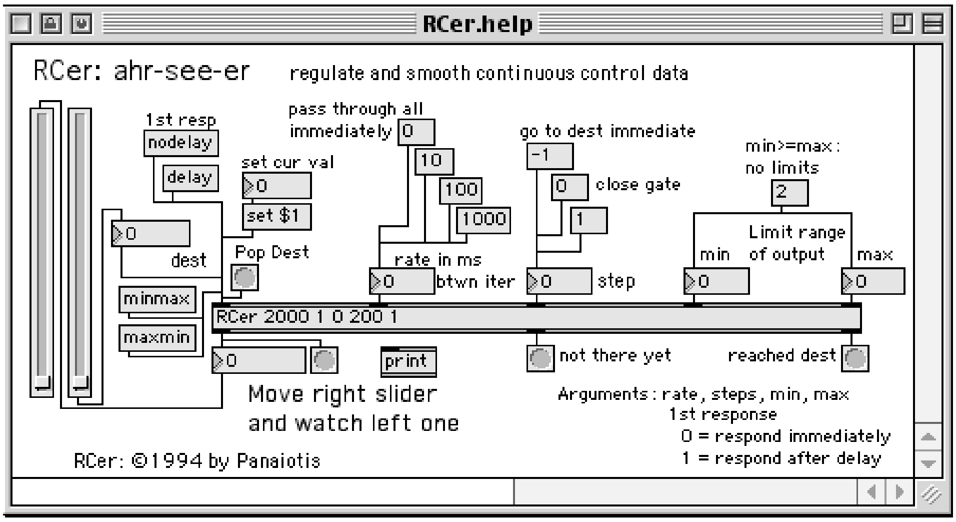



Figure 4.3H The RCer help file.

The scaled, continuous transition from the `RCer` object is fed to a `ctlout` object, which sends the MIDI continuous controller data to SuperCollider as control inputs associated with variables in the granular synthesis patch. The MIDI continuous controllers are all set to controller number seven, but allocated different MIDI channels.

This set-up proved reliable and responsive. It translated varying dynamics of movement into varying rates of change of the granulation variables, which equated to an increased sense of urgency when the movement was more dynamic, and a more meditative rate of change in association with slower movement. The smooth transitions created by the `RCer` object imbued the sound environment with a viscous, fluid quality.

### 4.3.4 Summary

MAP1 illustrates a number of developments over the previous project, Ghost in the Machine.

It responds, in realtime, to the dynamic of sensed movement in a number of different ways. There are multiple simultaneous controls, which vary their behaviour in accordance with the dynamic of activity within the interactive space. The response of the installation is a result of a combination of interactive controls. This provided the scope for multiple people to engage with the space simultaneously, and for the installation to be '*performed*' as an instrument.

MAP1 exhibited no pre-defined musical/sound/compositional form. The sound environment could evolve through an extremely broad range of aesthetic qualities. Those engaged in the space could dramatically alter the aesthetic by entering their own sound material. The use of the sounds made by those within the installation space, as source material for the realtime synthesis process, provided a further layer of engagement; an additional reason to own their individual experience.

The move away from pre-made content, and pre-defined pathways was driven by a consideration of the difference between triggered and streamed outcomes. It was a direct outcome of my own intention to remove myself, as much as possible, from imposing an aesthetic; a collection of outcomes, a form, or a prescriptive experience. I wished for inhabitants of the installation to mould their own individual experience, and I saw any kind of prescriptive system as a hindrance to that outcome.

In working towards this approach, I was inspired by a number of Fluxus artists and composers who addressed these same issues in their musical compositions. Among them, composers such as John Cage, Morton Feldman, Christian Wolff, George Brecht and Earle Brown explored approaches to composition and performance that introduced freedom of form and content to their musical composition.

Earle Brown wished to keep his music ‘plastic’. Whilst he acknowledged the composer's responsibility to establish a structure, a point of focus, a form of guidance, he wished the work to be defined at the point of creation. He wished for each performance of that work to be different, to be influenced and defined by those involved in its performance.

Alexander Calder’s kinetic art, especially his mobiles, caused Brown to consider spontaneity and open form mobility. Brown’s particular fascination with the Calder works, and also the spontaneous techniques being developed by Pollock, is born out by the following reflection:

> [the]… *creative function of ‘non control’ and the ‘finding’ aspects of the work within the process of ‘making’ the work, the integral but unpredictable ‘floating’ variations of a Calder mobile and the contextual rightness of Pollock’s spontaneity and directness in relation to the material and his particular image of the work. Both show an awareness of the ‘found object’ tradition as well as established unique and personal conditions of control of the totality. The momentary resolution of this dichotomy seems to me to be the ‘subject’ (as distinct from object) of today’s art, common to all the arts.*
> (E. Brown in Nyman:1974:56)

Brown comments “*what interests me, is to find the degree of conditioning (of conception, of notation, of realisation) which will balance the work between the points of control and non-control ... there is no final solution to this paradox ... which is why art is.” (*E. Brown in Nyman:1974:56)

Brown’s interest is clearly in the flexibility of systems, the composition of a work in such a manner that its ‘performed’ form would draw from the practices of ‘found object’ art and spontaneity. He sees "*form as a function of people acting directly in response to a described environment … it seems reasonable to consider the potential of the human mind as a collaborative creative parameter.*" (ibid:57)

Brown is very clear, though, that the composer brings something to the work. The work is not simply a collection of spontaneous events; it is embedded in the foundation of the composers outline.

> *… there must be a fixed (even if flexible) sound content, to establish the character of the work, in order to be called 'open' or 'available' form. We recognize people regardless of what they are doing or saying or how they are dressed if their basic identity has been established as a constant but flexible function of being alive.* (ibid:70)

The considerations of Earle Brown are in reference to the composition of acoustic performance works. Composers working with electronics were exploring similar ground. Nyman (1974) states how Gordon Mumma elucidates a

> *… case for the use of electronic systems in live performance; … they present or exploit in some way the qualities of variability, instability or unpredictability - things which may arise of their own accord or are in some way beyond the immediate control of the composer or the operator.* (Nyman:1974:91)

These are some of the many inspirations from which I drew my approach to fluid, evolving, and realtime interactive, responsive environments. MAP1 was the first step in this direction. It successfully presented a very fluid, visceral interaction with the sound environment; however, some people found the multiple independent controls confusing. They found it hard to develop a clear cognitive map of their relationship to the environment.

In developing MAP2, my objective was to develop a clearer interface, but maintain the immediacy, and dynamic engagement offered by MAP1.

```
(
var bufL, bufR, maxCycleTime, input, recycle, threshhold, inputAmp, chanSelect, dummy;
var granulated;
var maxPitchShift, maxPitchDisp, maxTimeDisp, maxOverlap, maxDur;
var delayLineLength;
var pitchSlider, pitchDispSlider, timeDispSlider, durSlider, overlapSlider, ampSlider;

maxCycleTime = 2;    // buffer loop time
maxDur = 0.2;
maxPitchShift = 4;            // maximum upward transposition ratio
maxPitchDisp = 1;
maxTimeDisp = 0.5;
maxOverlap = 8;
threshhold = 0.05;    // noise gating threshold

// create window
if ( not( w.isKindOf(GUIWindow) ) or: { w.name.postln != "Granulate" }, {
        w = GUIWindow.new("Granulate", Rect.new( 128, 64, 528, 264 ));
        SliderView.new( w, Rect.new( 17, 16, 227, 35 ), "SliderView", 1, 0, maxPitchShift, 0, 'linear');
        SliderView.new( w, Rect.new( 17, 39, 227, 58 ), "SliderView", 0, 0, maxPitchDisp, 0, 'linear');
        SliderView.new( w, Rect.new( 17, 62, 227, 81 ), "SliderView", 0, 0, maxTimeDisp, 0, 'linear');
        SliderView.new( w, Rect.new( 17, 85, 227, 104 ), "SliderView", maxOverlap, 0.1, maxOverlap, 0, 'exponential');
        SliderView.new( w, Rect.new( 17, 108, 227, 127 ), "SliderView", 0.02, 0.02, maxDur, 0, 'exponential');
        SliderView.new( w, Rect.new( 17, 131, 227, 150 ), "SliderView", 0.5, 0, 1, 0, 'linear');
        StringView.new( w, Rect.new( 233, 16, 361, 36 ), "Pitch");
        StringView.new( w, Rect.new( 233, 39, 361, 59 ), "Pitch Dispersion");
        StringView.new( w, Rect.new( 233, 62, 361, 82 ), "Time Dispersion");
        StringView.new( w, Rect.new( 233, 85, 361, 105 ), "Grain Density");
        StringView.new( w, Rect.new( 233, 108, 361, 128 ), "Grain Duration");
        StringView.new( w, Rect.new( 233, 131, 361, 151 ), "Grain Amplitude");
        w.onClose = { w = nil; }; // upon closing window, set global variable 'w' to nil
},{
        w.toFront;
});

// use multiple assignment to get slider views
# pitchSlider, pitchDispSlider, timeDispSlider, overlapSlider, durSlider, ampSlider = w.views;

play({ arg synth;
        // allocate delay lines
        delayLineLength = Synth.sampleRate * maxCycleTime + synth.blockSize + 2;
        bufL = Signal.new(delayLineLength);
        bufR = Signal.new(delayLineLength);

        // spawn grains
        granulated = GrainTap.ar([bufL, bufR], durSlider.kr, pitchSlider.kr, pitchDispSlider.kr,
                                timeDispSlider.kr, overlapSlider.kr, ampSlider.kr);

        input = AudioIn.ar([1,2]);
        recycle = TapN.ar([bufL, bufR], maxCycleTime);

        // select between rerecording previous delay line data or new input
        // depending on input amplitude
        inputAmp = Amplitude.kr(input);
        chanSelect = Lag.kr(inputAmp > threshhold, 0.1);
        input = (input * chanSelect) + (recycle * (1 - chanSelect));

        DelayWr.ar([bufL, bufR], input <! granulated);
        //DelayWr.ar([bufL, bufR], input);
        granulated
})
)
```

Figure 4.3I  SuperCollider granular synthesis algorithm developed for MAP1.

## *4.4 MAP2*

MAP2 is a development of MAP1. It continues the exploration of realtime video sensing using the Very Nervous System (VNS), but extends it into a two camera, three-dimensional sensing schema. It develops dynamic orchestration that changes both the synthesis instrument being controlled and the timbre on the basis of the level of activity in the installation. MAP2 also implements multiple independent sensing regions that allow several visitors (up to five) to play the work independently. As with MAP1, MAP2 is a realtime interactive, responsive sound environment.

MAP2 was commissioned by the Staatliches Institut für Musikforschung (State Institute for Music Research) (SIM), Berlin as part of the Festival of Culture (KUNSTFEST) for the millennium. It was developed in collaboration with Dr Ioannis Zannos[32] at SIM and exhibited at the Musical Instrument Museum in Berlin from December 1999 to January 2000.

Like the works that preceded it, MAP2 explores the interaction between human movement and the creation of music. It attempts to create an environment in which people can consider the impact they make on their immediate environment and the causal loops that exist between behaviour and quality of environment. Our personal aesthetic leads to decisions about preferential behaviour patterns and in turn preferential environmental qualities; one conditions the other.

Figure 4.4A MAP2 in Berlin.

[32] Dr Zannos assisted with the writing of the SuperCollider structure for dynamic allocation of synthesis algorithms depending on sensed activity levels and the dynamic audio filters used to constantly vary the timbre of the instruments. Garth Paine wrote the instrument algorithms and the Max code for controlling the VNS and communicating with SuperCollider. Garth Paine was also responsible for all the concept development, sound design and video sensing approach.

### 4.4.2 Program Notes

*MAP2 is an interactive immersive sound environment installation by the installation artist and composer Garth Paine, developed in collaboration with Dr Ioannis Zannos at The Staatliches Institut für Musikforshung in Berlin. It was first exhibited in the Musical Instrument Museum, Berlin as part of the Millennium celebrations, 1999/2000.*

*MAP2 is a three dimensional space which can be entered and encountered, played with and played. It is a virtual musical instrument using the movement of those within it as its raw material for composition and sound development.*

*People enter MAP2 to compose music and sound by using their bodies to solicit responses from the custom developed computer software, generating a rich, enveloping and continually evolving sonic environment in realtime. A range of different aural qualities is mapped in different qualitative groupings to different regions within the installation space generating a multitude of aural textures and densities. These are based on the quality of movement of bodies within that region of the installation. The visitors' responses become part of a dynamic system of development and experimentation.*

*MAP2 grew out of development work on Garth Paine's previous successful installation MAP1.*

*The presence of a <BODY> generates a stream of responses. The challenge is to explore the environment in order to find a symbiotic existence, cognisant of personal presence within an ever affected system. The choices are many and varied, being dependent upon the way(s) in which the <BODY(S)> are present, each experience being unique. A direct relationship between the speed of movement, the position of movement, the mass of movement, the direction of movement and the quality of environment is established. MAP2 explores the inter-relationship between people, their actions and their environment. It provides a platform to consider the relationship between the <perceived> and <actual> consequence of behaviour.*

*MAP2 utilises custom software and hardware developed to map the realtime video tracking information onto realtime sound synthesis algorithms.*

*Go and compose, play and enjoy being immersed in the rich viscous sounds of MAP2*

### 4.4.3 The Technology

MAP2 used the same technology as MAP1. It continued my exploration of video sensing, employing the Very Nervous System (VNS). MAP2 also used the realtime sound synthesis language SuperCollider, which I first used in MAP1. The application of both these tools was extended in MAP2.

As discussed previously, whilst MAP1 was successful as an interactive instrument, there was some confusion about the interface. Rather then spatially separating the interactive controls, as in the large floor sliders in MAP1, I wanted to combine a range of variables in a more spatially homogenous way.

It was clear that this would require a different mapping strategy from that applied in MAP1. The space used for the exhibition of MAP2 at the Musical Instrument Museum in Berlin was a small auditorium with a raised performance/dance platform. The physical architecture of the auditorium meant that a horizontal (side-on) camera could be mounted some distance from the interactive space, and would therefore be able to 'see' the vertical area of the exhibition space quite well. The pragmatics of this situation encouraged me to explore the potential of sensing both horizontal and vertical movement characteristics.

Two video cameras were used in MAP2 to create a three-dimensional video mapping of the exhibition space. The VNS has the facility for two video camera inputs. This function is similar to the dual map video analysis approach used in MAP1. The VNS capability of splitting the single video frame into separate fields, is redirected in MAP2 to simultaneously analyse two separate video inputs.

In order to develop the three-dimensional video sensing approach, it was necessary to use video cameras that could synchronise to an external input. Two VICAM VA810 B/W CCTV[33] video cameras were used. A video distribution amplifier was used to feed the video output from camera one to the synchronisation input on camera two whilst simultaneously feeding the video output from camera one to the VNS. The video output from camera two was fed to the VNS independently (see Figure 4.4B). This ensured that the video signal received by the VNS from both cameras was synchronised to single frame accuracy. Frame accurate synchronisation was a requirement for the VNS as the two fields of analysis, although independent, are treated as though they are the un-interlaced components of a single video input (as per MAP1).

A good deal of information can be derived from the data output of the VNS. As discussed previously, the direction of movement, the size of the object, inertia, and relationships between moving objects are higher order outcomes. I thought that if these approaches were applied in the vertical plane then it might be possible to derive the height of the visitor, or to map the point of greatest activity on the vertical plane to control pitch or amplitude, or some other synthesis parameter.

[33] CCTV stands for Closed Circuit Television. CCTV cameras are primarily used in security and monitoring situations. The camera bodies are very compact, the video quality is adequate for video sensing applications, and the cameras take a variety of lenses making it possible to recreate the original setup in different gallery spaces.

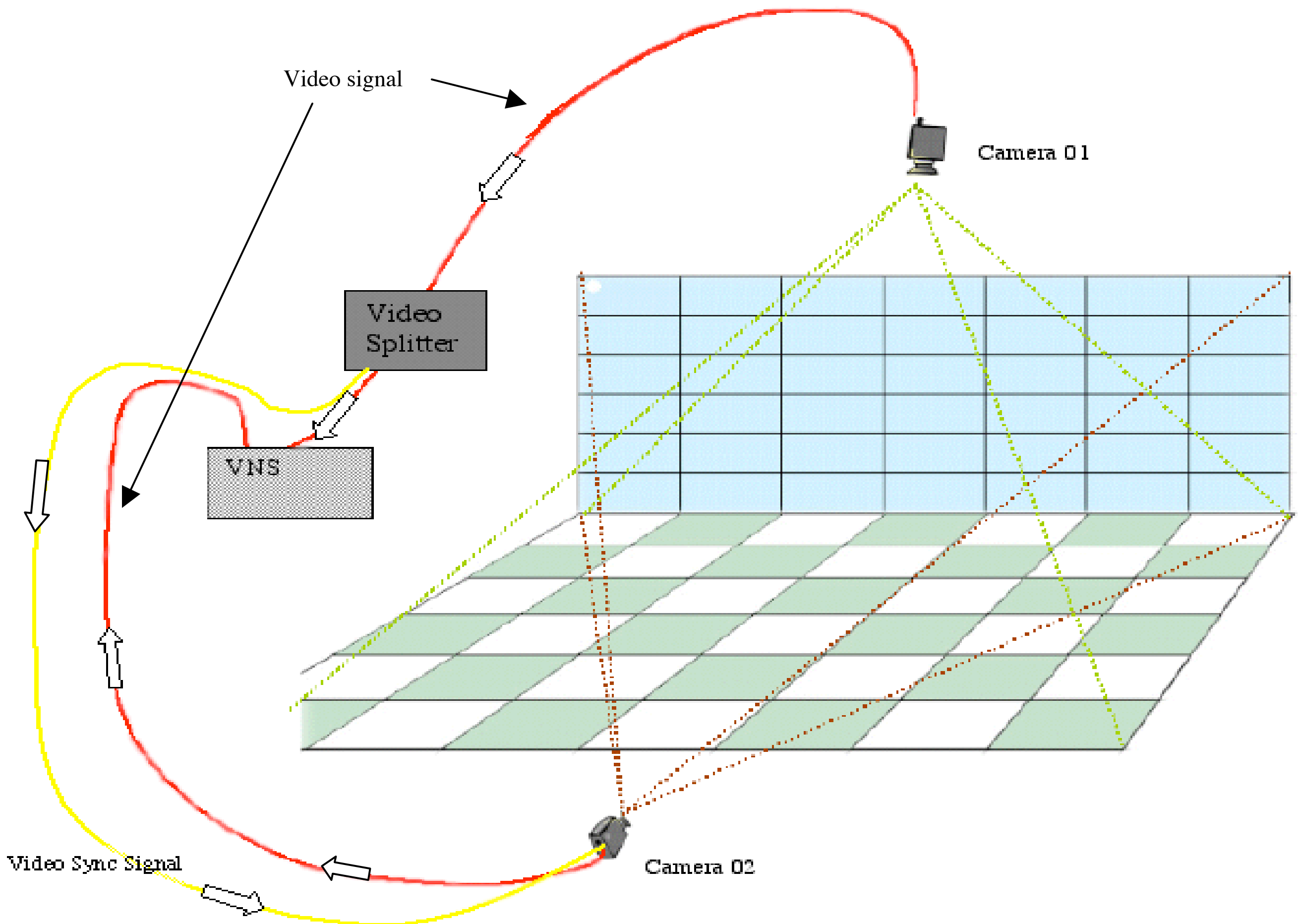



Figure 4.4B  Camera set up for MAP2.

Camera one was mounted in the roof of the exhibition space, approximately five metres above the centre of the exhibition floor space.  It was fitted with a lens that allowed it to view the entire exhibition floor area.  Camera two was mounted on a side wall approximately twelve metres from the centre of the exhibition space, at a height that equated to the head height (approximately 2 metres) of someone standing on the raised performance platform that was the interactive exhibition space.

Video sensing presents a number of challenges.  The Data reported by a video sensing system includes a number of distortions that are inherent in the use of camera lenses that are not of an infinite focal length.

A camera reports the most accurate image when the subject is in the centre of the cameras field of view.  Figure 4.4C, and 4.4D illustrate the distortion associated with the image reported.  This data error was a challenge in both MAP1 and MAP2.  Lens manufacturers and photographers are well aware of these issues; the degree to which the edges of an image are distorted is a common photographic lens test.  The associated problem here though is the shadowing of additional sensing regions.

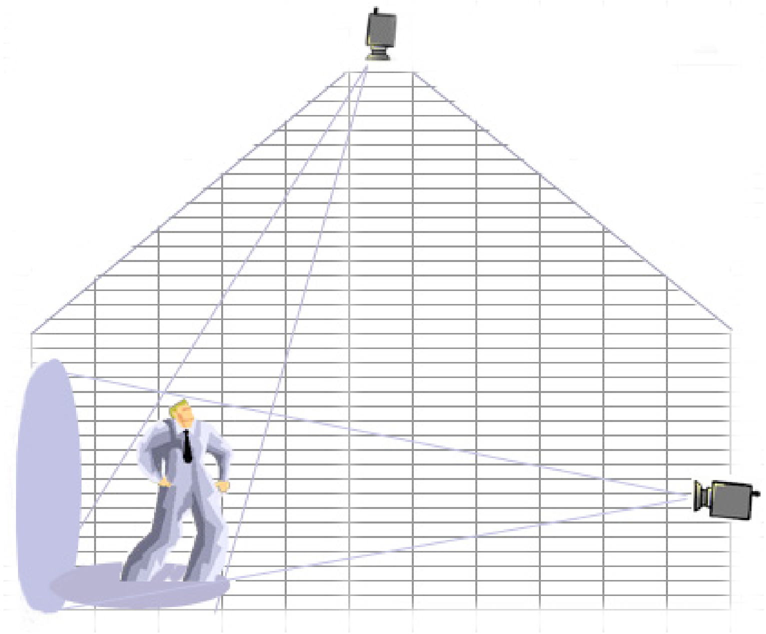


Figure 4.4C  Horizontal and vertical tracking error caused by camera lens

It is clear from Figure 4.4C and 4.4D that the parallax error, inherent in any camera lens generates an image that corresponds to movement in regions where none is present. The floor image shows ghosting, which indicates movement in three columns rather than the two the character is occupying. This problem is more pronounced the closer the object gets to the camera lens (see Figure 4.4D). If the camera were directly overhead, and the subject were infinitely small, it would report the area of occupation correctly. Similarly, if the camera lens had an infinite focal length (i.e. if the camera could be mounted infinitely far away from the subject) then the distortion of view would no longer occur. Of course lenses are not of infinite focal length, and so some consideration had to be made for the inaccuracies in the technology.

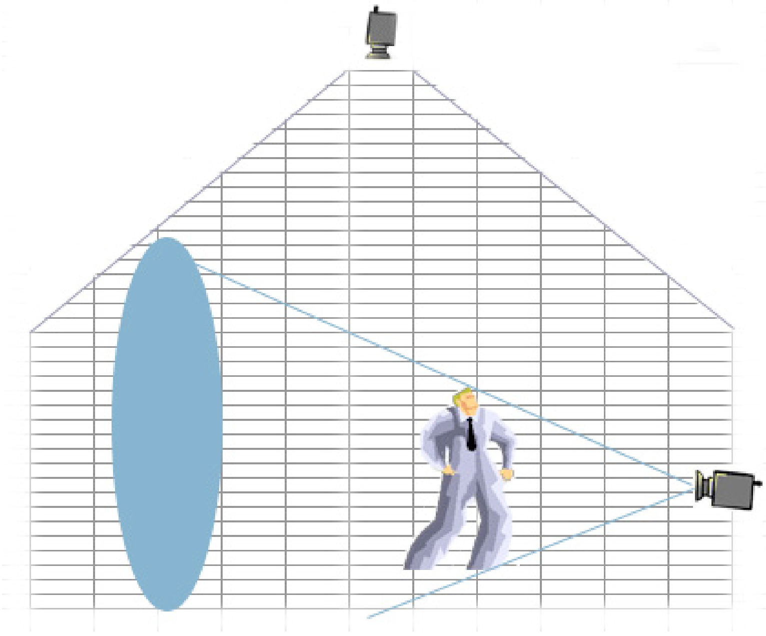


Figure 4.4D Vertical camera parallax error.

One of the principal developments in MAP2 was that the video sensing data was parsed, filtered and collated in SuperCollider. This was done because the object oriented nature of SuperCollider made it possible to very quickly collate the data into zones. The floor area (the image seen from the overhead camera) was divided into four zones, each of which contained 16 regions. As the data output by the VNS is an array of integers, one per defined region, in order from the top left corner to the bottom right corner, it was necessary to gather the regions associated with each zone into a new array. A *Mapdata* class[34] was defined in SuperCollider. It contains methods to collect the zone data into separate arrays as follows:

```
zone1 = #[0, 1, 2, 3, 8, 9, 10, 11, 16, 17, 18, 19, 24, 25, 26, 27].collect({ arg i; cam1kr.at(i).poll; });
zone2 = #[4, 5, 6, 7, 12, 13, 14, 15, 20, 21, 22, 23, 28, 29, 30, 31].collect({ arg i; cam1kr.at(i).poll; });
zone3 = #[36, 37, 38, 39, 44, 45, 46, 47, 52, 53, 54, 55, 60, 61, 62, 63].collect({ arg i;  cam1kr.at(i).poll; });
zone4 = #[32, 33, 34, 35, 40, 41, 42, 43, 48, 49, 50, 51, 56, 57,  58, 59].collect({ arg i; cam1kr.at(i).poll; });
```

Each zone could then be analysed and auralised independently. The zones were defined clockwise from the top left corner.

The camera maps were defined as eight row by eight column grids and gathered into four zones as indicated in Figure 4.4E.

[34] Object Oriented programming languages work by defining ‘blue prints’ for required functionality. These are called *classes*. When the functionality associated with the *class* is required in the software, an *object* is created based on the *class*. This approach allows for the dynamic creation and disposal of *objects*, which in turn allows the software to dynamically change its structure during runtime.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
| 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 |
| 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 |
| 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 |
| 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 |
| 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 |

Figure 4.4E The VNS map for camera one.

The data from the VNS was sent to SuperCollider as MIDI Continuous Controller (MCC) arrays. A Pyrite[35] script in Max examines the integer array from the VNS, and maps the VNS region value to a value on an associated MCC. The MCCs were numbered as per the region numbers illustrated in Figure 4.4E. By way of optimising the data being sent from Max to SuperCollider, a new MCC message was only sent for regions with new data. A MCC on a separate channel was used as a flag in SuperCollider to initiate analysis of new data. MIDI channels were allocated as follows:

| *Data* | *MIDI Channel* |
|---|---|
| Camera One region data | 1 |
| Camera One new data flag | 2 |
| Camera Two region data | 3 |
| Camera Two new data flag | 4 |

Each of the four zones operated independently; they sourced the same synthesis algorithm classes, but had slightly differing filter algorithms on their output. Each zone addressed two audio channels and two loudspeakers, thereby creating a stereo audio field in each zone. This created movement in the zone according to the region of highest activity, and also created the effect of passing the sound from one zone to another as a visitor moved through the space. Brief movements from one zone into another and back created a sense of the sound being 'thrown' down the installation space and back again. Many visitors to the exhibition found this characteristic and enjoyed playing with it.

An important development in MAP2 was the dynamic allocation of synthesis algorithms on the basis of the current maximum activity within each zone. I termed this 'Dynamic Orchestration'. All the installation works prior to MAP2 had a fixed synthesis schema, which was initialised when the program was booted up. MAP2 made use of the object oriented programming approach of dynamic object creation as required, directing this functionality at the synthesis algorithms and providing the functionality whereby instruments were introduced, or disposed of on the basis of activity within each zone.

[35] Pyrite is a script based programming language that runs inside Max. Pyrite was developed by James McCartney for list intensive applications.

A threshold was set for the introduction of a new instrument algorithm. Three algorithms were available (Layer01, Layer02, and Layer03), with the new algorithm augmenting the orchestration whilst activity was within its threshold.

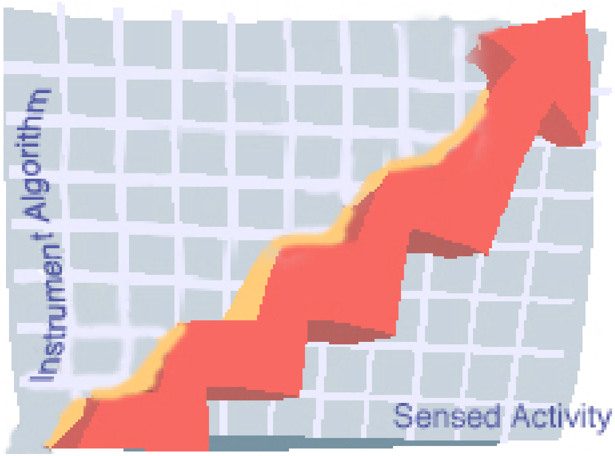



Figure 4.4F Dynamic orchestration.

Each of the instruments was created in a *layer* class and instantiated in a *Main* method as shown in Figures 4.4G:

```
Main : Process {

      startUp {
            super.startUp;
            false.trace;
            currentEnvironment = Environment.new;
            this.initLib;
            ParamSpec.default;
            this.run;

      }

      initLib {}

      run {var map;
            map = MAP2.gui;

            map.addItems(

            Layer01.button,
            Layer02.button,
            Layer03.button
      );

            MAP2.start();
      }
}
```

Figure 4.4G The Main method for MAP2.

Each of the layer classes are included at the end of this chapter, in Figures 4.4J, K, L.

The Camera two map (the vertical image space) was defined as illustrated in Figure 4.4H.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
| 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 |
| 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 |
| 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 |
| 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 |
| 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 |

Figure 4.4H the VNS map for camera one.

The vertical camera was set to be sensitive in one row only. Depending on the installation venue, this has been either row two or three, as indicated by shading in Figure 4.4H.

### 4.4.3 Mapping

As mentioned above, two cameras were used for MAP2, each with its own mapping strategy and audio output. Both cameras used a sensing map that defined sixty-four regions in an eight column by eight row grid.

Each of the 4 horizontal zones contained three instruments that responded to different levels of sensed activity. These instruments created a dynamic orchestration that was designed to provide a continuous and direct change in the audio timbre in relation to the dynamic of movement. Each time the movement exceeded the set threshold for the introduction of the next instrument, the new instrument would come into play in addition to the existing instrument. This created an overlapping of sound textures, which produced a graduated change rather than a sudden change, which is what would have happened if the lower threshold instrument were superseded by the new one.

In addition to the dynamic orchestration approach, filter banks were placed on the output of each audio channel. The filters were designed to change their central frequency in accordance with the region indicating the most activity.

For instance, the Layer01class places a `CombL` comb delay line on each stereo output. The signature of the `CombL` class is as follows:

```
CombL.ar(in, maxdelaytime, delaytime, decaytime, mul, add)
```

In the Layer01class, `CombL` is used as a resonator. The resonant fundamental is equal to the reciprocal of the delay time. It is defined as follows:

```
CombL.ar(basicTextureSound, 0.6, Mapdata.zonefilter1, 10 -
(Mapdata.maxActiv1 * 10))
```

It can be seen here, that the filter uses the basicTextureSound generated for Layer01as its input. A delay line is established at 0.6 seconds in length, and the actual momentary delay is set by the zonefilter1 variable in the Mapdata class. The decay time for each echo through the filter is set by the equation, `10 - (Mapdata.maxActiv1 * 10)`, and is therefore constantly varying with the change in the dynamic of activity on Layer01.

This use of `CombL` creates a varying resonant filter, driven by the sensed activity in the installation. This technique was applied in all three layers, and generated a viscous, visceral sound environment that fluidly mapped the gesture patterns in the installation.

As I mentioned earlier, the four zones were completely independent, and applied slightly different filter settings, therefore producing differing timbres. Four people could therefore play the installation simultaneously (one in each zone), or groups could play as an ensemble.

The layer sounds were as follows:

1. Layer01 consisted of a drone like sound that was constantly being modulated by the resonant filter discussed above. This was the only sound audible when no one was in the installation, and, in that state, slowly changed with different resonant frequencies being prominent before the focus moved to another frequency. This layer responded more dynamically when engaged interactively, because the filters were then driven by the interactive process.

2. Layer02 was a more dynamic and more defined sound. It moved through greater pitch ranges and its filter outputs changed more dramatically in relation to interactive input within its threshold range.

3. Layer03 was a bubbly, bouncy sound that used a varying echo to accentuate the dynamic of activity that was needed to add this layer to the orchestration.

The vertical active space (camera02) was set at just above head height and defines a row at that height through the middle of the installation space. This row was mapped to play a physically modelled guitar sound; one pitch of a scale[36] associated with one region of the row. The camera02 data was mapped in this way for the following reasons:

- During the development time for MAP2, I could not find a use within the sound design approach for the data that was generated in the lower rows of this map (see Figure 4.4H above) when people walked through the space. As the entire body height was being reported, including the movement of arms and legs (the most active parts of the body), the data was too consistently full. When several people were in the installation this camera map could be consistently active in almost all regions.

[36] I used a modal scale to avoid references to tonal music, and associated tonal harmony issues related to resolution of dissonance etc.

- The physically modelled guitar sound was synthesised in realtime, but it was essentially triggered by movement into a new region. Consistent activity in one region did not re-trigger the associated note; activity had to be sensed in a new region for a new note to sound. This accentuated the sense of spatial movement, (you had to move to play a new note), and also mirrored a sense of triggering an acoustic instrument. MAP2 was initially exhibited in the Musical Instrument Museum in Berlin, about ten metres from the original keyboards of J.S Bach, and surrounded by an extraordinary collection of European acoustic music heritage. I felt I had to make reference to this context, and it seemed like a good opportunity to place a triggered sound output (the vertical – camera02) alongside a streamed sound output (the horizontal – camera01).
- The row associated with playing the guitar sound was placed just above head height so that people had to reach up to find the additional sound layer, thereby encouraging them to explore the spatial aspects of the installation in three dimensions.

#### 4.4.4 Summary

MAP2 introduced a number of innovations:

- it defined multiple interactive zones (4) that could be performed simultaneously but independently,
- it extended the concept of video sensing to a three-dimensional schema that mapped different sound algorithms in the vertical plane to those mapped in the horizontal plane,
- it implemented dynamic orchestration, which introduces or disposes of synthesis algorithms on the basis of momentary activity levels,
- it extended and refined the idea of ‘streamed’ audio output in response to the ‘streamed’ sensed activity reported by the VNS.

MAP2 produced the most sophisticated and resolved interface between human movement and sound timbre so far. It presented a much more immediate interface than MAP1, but maintained a very wide scope for individualised experience.

The exploration of object oriented programming in MAP2 produced a platform for dynamic orchestration, a methodology that I felt represented a more flexible approach to interactivity than I had been able to achieve previously, and one that I have continued to develop in recent works.

The dynamic orchestration I developed for MAP2 raised a number of questions with regards to the composition of multiple interactive streams of sound.

I felt that Map2 offered the potential for the combination of completely random, and chaotic sound elements in a single sound environment. The four independent horizontal zones had created a small orchestra of instruments that were, however, not under the direction of a single conductor. At no point during the exhibition of MAP2 did I feel that the outcome from people performing the four zones independently was unpleasant, or inharmonious.

I wondered therefore what it would be like to establish a system that;

1. had as its input, multiple streamed data components,
2. was free to create any truly chaotic combination of the input data streams,
3. would use as its data streams input phenomena that would not display synchronised variations.

The above data input requirements appeared to be met by the sensing of weather conditions.  I chose therefore to create an installation that used weather data as its control input in order to explore truly chaotic relationships as controllers for sound synthesis environments, REEDS was the outcome.

```
Mapdata {
	classvar <>cam1kr;
	classvar <>zone1, <>zone2, <>zone3, <>zone4;
	classvar <>zone1xy, <>zone2xy, <>zone3xy, <>zone4xy;
	classvar <>layer1FilterTuning;
	classvar <>cam2kr, <>cam2,  <>maxCam2, <>maxIndCam2;
////////////////////////////////////
	classvar <>maxActiv1, <>maxActiv2, <>maxActiv3, <>maxActiv4;
	classvar <>maxInd1, <>maxInd2, <>maxInd3, <>maxInd4;
	classvar <>maxInd1kr, <>maxInd2kr, <>maxInd3kr, <>maxInd4kr;
	classvar <>zonefilter1, <>zonefilter2, <>zonefilter3, <>zonefilter4;
//////////////////////////////////////////////////////////////////////////////

/* DECLARATION OF ENVIRONMENT. THIS CREATES GLOBALLY AVAILABLE DATA STRUCTURES
AND STORES THEM INTO GLOBALLY AVAILABLE VARIABLES.   ALSO, IT STARTS A NUMBER OF
MIDICONTROLLERS THAT READ DATA INPUT VIA MIDI.  THESE ARE ALREADY .kr UGENS, SO SYNTHESIS IS
ALREADY RUNNING HERE!*/

*init {		var row, col;
		var filter1xData, filter1xDataReversed, filter1yData, filter1yDataReversed;
		filter1xData = #[0.002, 0.004, 0.006, 0.008];
		filter1xDataReversed = filter1xData.reverse;
		filter1yData = filter1xData.copy * 2;
		filter1yDataReversed = filter1yData.reverse;
		cam1kr = Array.newClear(64).collect({ arg n, i; MIDIController.kr(1, i); });

		maxActiv1 = Plug.kr(0.1);
		maxActiv2 = Plug.kr(0.1);
		maxActiv3 = Plug.kr(0.1);
		maxActiv4 = Plug.kr(0.1);

		maxInd1 = maxInd2 = maxInd3 = maxInd4 = 1;

		maxInd1kr = Plug.kr(1);
		maxInd2kr = Plug.kr(1);
		maxInd3kr = Plug.kr(1);
		maxInd4kr = Plug.kr(1);

		zonefilter1 = Plug.kr(0.01, 0.5);
		zonefilter2 = Plug.kr(0.01, 0.1);
		zonefilter3 = Plug.kr(0.01, 0.001);
		zonefilter4 = Plug.kr(0.01, 0.0);	// Tanbur - like.

		layer1FilterTuning = 0.5;				// 5;

		zone1xy = Array.newClear(16).collect({ arg a, i;
			col = filter1xData.at((i % 4).floor.asInteger) + ((i / 4).floor * 0.0005);
			row = filter1yDataReversed.at((i / 4).floor.asInteger) + ((i % 4) * 0.0005);
			[col, row] }) * layer1FilterTuning;

		zone2xy = Array.newClear(16).collect({ arg a, i;
			col = filter1xDataReversed.at((i % 4).floor.asInteger) + ((i / 4).floor * 0.0005);
			row = filter1yDataReversed.at((i / 4).floor.asInteger) + ((i % 4) * 0.0005);
			[col, row] }) * (layer1FilterTuning * 1.5); /// * (layer1FilterTuning * 0.5);

		zone3xy = Array.newClear(16).collect({ arg a, i;
			col = filter1xData.at((i % 4).floor.asInteger) + ((i / 4).floor * 0.0005);
			row = filter1yData.at((i / 4).floor.asInteger) + ((i % 4) * 0.0005);
			[col, row] }) * (layer1FilterTuning * 1.25); ///  * (layer1FilterTuning * 0.25);  // Tanbur - like strum


		   zone4xy = Array.newClear(16).collect({ arg a, i;
			col = filter1xDataReversed.at((i % 4).floor.asInteger) + ((i / 4).floor * 0.0005);
			row = filter1yData.at((i / 4).floor.asInteger) + ((i % 4) * 0.0005);
			 [col, row] }) * (layer1FilterTuning * 2);

		cam2kr = Array.newClear(64).collect({ arg n, i;  MIDIController.kr(3, i);  });

		cam2 = FloatArray.newClear(64);
		maxIndCam2 = Plug.kr(1, 0.16);
		maxCam2 = Plug.kr(1, 0.16);
```

```
            }

/* STARTUP OF A LOOP TO READ THE DATA
COMING IN VIA MIDI CONTROLLERS CREATED IN THE ENVIRONMENT AND
COPY THEM ONTO ANY REGIONS OR SHAPES THAT NEED TO BE ACCESSED
BY VARIOUS SOUNDS */
*poll {
            var indexLookupThreshold = 0.00000001;
            var waitTime = 0.034;           ///  1;
            var max;

            Task({
                  loop({

                        zone1 = #[0, 1, 2, 3, 8, 9, 10, 11, 16, 17, 18, 19, 24, 25, 26, 27].collect({ arg i; cam1kr.at(i).poll; });
                              zone2 = #[4, 5, 6, 7, 12, 13, 14, 15, 20, 21, 22, 23, 28, 29, 30, 31].collect({ arg i; cam1kr.at(i).poll; });
                        zone3 = #[36, 37, 38, 39, 44, 45, 46, 47, 52, 53, 54, 55, 60, 61, 62, 63].collect({ arg i;  cam1kr.at(i).poll; });
                        zone4 = #[32, 33, 34, 35, 40, 41, 42, 43, 48, 49, 50, 51, 56, 57,  58, 59].collect({ arg i; cam1kr.at(i).poll; });

                        max = zone1.maxItem;
                        maxActiv1.source = max;
                        if ( max > indexLookupThreshold, {
                                    maxInd1kr.source = maxInd1 = zone1.indexOf(max);
                                    zonefilter1.source = zone1xy.at(maxInd1).at(0);
                        });

                              max = zone2.maxItem;
                        maxActiv2.source = max;
                        if ( max > indexLookupThreshold, {
                                    maxInd2kr.source = maxInd2 = zone2.indexOf(max);
                                    zonefilter2.source = zone2xy.at(maxInd2).at(0);
                        });

                              max = zone3.maxItem;
                        maxActiv3.source = max;
                        if ( max > indexLookupThreshold, {
                                    maxInd3kr.source = maxInd3 = zone3.indexOf(max);
                                    zonefilter3.source = zone3xy.at(maxInd3).at(0);
                        });

                              max = zone4.maxItem;
                        maxActiv4.source = max;
                        if ( max > indexLookupThreshold, {
                                    maxInd4kr.source = maxInd4 = zone4.indexOf(max);
                                    zonefilter4.source = zone4xy.at(maxInd4).at(0);
                        });

//                      [maxActiv1.poll, maxActiv2.poll, maxActiv3.poll, maxActiv4.poll].postln;
//                      [maxInd1, maxInd2, maxInd3, maxInd4].postln;

                        cam2kr.do({ arg c, i; cam2.put(i, c.poll) });
                        max = cam2.maxItem;
                        maxCam2.source = max;
                        if ( max > indexLookupThreshold, {
                                    maxIndCam2.source = cam2.indexOf(max);
                        });

                        waitTime.wait;
            });
        });
    }
}
```

Figure 4.4I  The Mapdata class from MAP2.

```
Layer01 {

      *button {

^SoundSwitch.new("Sound Layer 01",
      {     var n, d, basicTextureSound;

            n = 3;      // n*3 components in each channel
            d = 5.0;    // beat freq deviation

      basicTextureSound = XFadeTexture.ar({
            var z, p, q;
            p = Array.new(3*n);
            q = Array.new(3*n);
            n.do({ var freq;
                  freq = (48 + 30.rand).midicps;
                  p.add(freq);
                  p.add(freq + d.rand2);
                  p.add(freq + d.rand2);
            });
            n.do({ arg i;  var freq;
                  freq = p.at(3*i);
                  q.add(freq + d.rand2);
                  q.add(freq + d.rand2);
                  q.add(freq + d.rand2);
            });

            z = [`[p, nil, Array.rand(3*n, 0, 2pi)],
                  `[q, nil, Array.rand(3*n, 0, 2pi)]
                  ];
            Klang.ar(z, 1, 0, 0.01/n);
      }, 4, 4, 1);

      // the audio output array, generates 8 channels of sound, driven by each zones activity level
      [ Pan2.ar(CombL.ar(basicTextureSound, 0.6, Mapdata.zonefilter1, 10 - (Mapdata.maxActiv1 * 10))
            , 0, Mapdata.maxActiv1 + 0.3),
        Pan2.ar(CombL.ar(basicTextureSound, 0.6, Mapdata.zonefilter2, 10 - (Mapdata.maxActiv2 * 10))
            , 0, Mapdata.maxActiv2 + 0.3),
        Pan2.ar(CombL.ar(basicTextureSound, 0.6, Mapdata.zonefilter3, 10 - (Mapdata.maxActiv3 * 10))
            , 0, Mapdata.maxActiv3 + 0.3),
        Pan2.ar(CombL.ar(basicTextureSound, 0.6, Mapdata.zonefilter4, 10 - (Mapdata.maxActiv4 * 10))
            , 0, Mapdata.maxActiv4 + 0.3)
      ].flat;
      });
    }
}
```

Figure 4.4J  The layer one SuperCollider sound algorithm for MAP2.

```
^SoundSwitch.new("Sound Layer 02",
	{
		var voiceActivationThreshold = 0.1; //  play only above this. 0.2 better?
		var soundLevel02Zone01;
// modal space
	soundLevel02Zone01 =
	[
	Pause.ar(
		{ SinOsc.ar(
			(DegreeToKey.kr(
				FloatArray[0, 2, 3, 5, 7, 9, 10, 12] + 5, // dorian scale
				[Mapdata.maxInd1kr, Mapdata.maxInd1kr - 12] * 0.5,	// Zone maxActivity index into scale
				12,			// 12 notes per octave
				1,
				65
			)).midicps,
			0,
			0.1);
		}
		, Mapdata.maxActiv1 > voiceActivationThreshold,
	)
	Pause.ar({ SinOsc.ar(
			(DegreeToKey.kr(
				FloatArray[0, 2, 3, 5, 7, 9, 10, 12] + 7, // dorian scale
				[Mapdata.maxInd2kr, Mapdata.maxInd2kr + 4] * 0.5,	// Zone maxActivity index into scale
				12,			// 12 notes per octave
				1,
				65
			)).midicps,
			0,
			0.1);}, Mapdata.maxActiv2 > voiceActivationThreshold,
	)
	Pause.ar({ SinOsc.ar(
			(DegreeToKey.kr(
				FloatArray[0, 2, 3, 5, 7, 9, 10, 12] - 5, // dorian scale
				[Mapdata.maxInd3kr, Mapdata.maxInd3kr + 8] * 0.5,	// Zone maxActivity index into scale
				12,			// 12 notes per octave
				1,
				65
			)).midicps,
			0,
			0.1);}, Mapdata.maxActiv3 > voiceActivationThreshold,
	)
	 Pause.ar(
		{ SinOsc.ar(
			(DegreeToKey.kr(
				FloatArray[0, 2, 3, 5, 7, 9, 10, 12], // dorian scale
				[Mapdata.maxInd4kr, Mapdata.maxInd4kr + 6] * 0.5,	// Zone maxActivity index into scale
				12,			// 12 notes per octave
				1,
				65
			)).midicps,
			0,
			0.1);}, Mapdata.maxActiv4 > voiceActivationThreshold
	)
	].flat;
	CombN.ar(soundLevel02Zone01, 0.31,
		0.31,
		2, 1, 0) //Use CombN as effect on sound total
})
}
)
```

Figure 4.4K  The layer two SuperCollider instrument algorithm for MAP2.

```
Layer03 {

	*button {

^SoundSwitch.new("Sound Layer 03 (bubbles)",
{
	var pitch1, pitch2, levelThreeZone01Sound;
	var clockRate, clockTime, clock, centerFreq, freq, panPos, patch;


	levelThreeZone01Sound =
	// sample and hold liquidities
	// mouse x controls clock rate, mouse y controls center frequency

	clockRate = Mapdata.maxIndCam2;
	clockTime = clockRate.reciprocal;
	clock = Impulse.kr(clockRate, 0.4);

	centerFreq = MouseY.kr(100, 8000, 'exponential');
	freq = Latch.kr(WhiteNoise.kr(centerFreq * 0.5, centerFreq), clock);
	panPos = Latch.kr(WhiteNoise.kr, clock);
	patch = CombN.ar(
			Pan2.ar(
				SinOsc.ar(
					freq,
					0,
					Decay2.kr(clock, 0.1 * clockTime, 0.9 * clockTime)
				),
				panPos
			),
			0.3, 0.3, 2
		);
	patch

}
)}
}
```

Figure 4.4L The layer three SuperCollider instrument algorithm for MAP2.

## *4.5 REEDS*

The Reeds project was commissioned by the Melbourne International Festival of the Arts in 2000, and first exhibited in November and December of that year on the Ornamental Lake at the Royal Botanic Gardens, Melbourne.

Figure 4.5A The Reeds exhibition at the Royal Botanic Gardens Melbourne as part of the Melbourne International Festival of the Arts in 2000

Reeds marks a slight diversion from the main topic of this exegesis. It explores the potential of multiple, random, continuous control inputs that have an internal chaotic structure. By this I mean that the relationships between the various controllers is in a constant state of flux. In MAP2, I was interested that the independent zones had provided a chaotic output, one where the four zone outputs were not interlinked, or under the control of a common influence. The combined sound was not unpleasant to listen to, even though there were up to twelve independent sound elements occurring together. These sounds were of course drawing from a common algorithmic pool, and therefore a common aesthetic base, but there were differences in the filter set up for each of the zones, which caused marked timbrel differences, representing substantially different spectra[37], which seemed to work together very well.

These experiences with MAP2 encouraged me to find a multi-dimensional control source that embodied chaotic inter-relationships, and would change rapidly, slowly and unexpectedly but within a defined overall range.

---

[37] All complex sounds are made up of a collection of frequencies called spectra. That collection of frequencies is defined as the sounds spectrum.

The weather seemed the most suitable candidate. Reeds was driven by two weather stations sensing:

- Wind Speed
- Wind Direction
- Temperature
- Solar Radiation

The sound synthesis was driven by both the direct input from each of these sensors, and the difference between the two weather stations, which although not far apart (10M) often reported markedly different conditions.

### 4.5.1 Program Notes

*A weed, so easily crushed underfoot, can push its way up through a tarmac path, creating a sizeable fracture in what appears to us to be an impervious surface.*

*One might postulate that if it could see the bigger picture, it might have decided to grow two feet to the left in the flowerbed or the grass. There is clearly an analogy here to our own birth, over which we appear to have little or no say (depending on one's religious beliefs).*

*It is exactly this chaotic behaviour of the natural world that informs the Reeds project. Whilst civilisation tries to harness or tame the chaotic in nature, or to explain it in terms of quantum theory and fractals, humanity cannot perceive a truly chaotic state. The forces of nature that dictate the growth of plant life fall into this category. It is not possible for us to predict with certainty the meteorological conditions from day to day, let alone year to year, and certainly not on the micro scale of the weed in the footpath. It is precisely these chaotic variations that are used in Reeds to conduct the sound score - to control and dictate the output of the realtime synthesis process.*

*Of course, the software design process predetermines the general structure and aesthetic of the sound, but the momentary output is unique. It is unlikely that the combination of wind speed, wind direction, solar radiation, and temperature that occur in this instance will be precisely replicated in any other moment. This chaotic variation is the very source of diversity, which I propose is the structure that creates such beauty in nature.*

*Reeds uses the relatively static external facade of the sculptural form as a way of representing the paradox observed in organic plant life, where in contrast to the apparently static external face of the plant, is the hidden, dynamic activity of photosynthesis and nutrient gathering that keeps the plant alive, and drives it's growth.*

*The Reed pod sculptures, appearing as lifelike presences on the Ornamental Lake at the Royal Botanic Gardens Melbourne, support two remote weather stations, gathering wind speed, wind direction, temperature, and solar radiation data. The meteorological conditions, vital to the plants' life processes, are transmitted back to a land-base where the data is transformed into eight channels of musical sounds that are broadcast back out to the Reed pods. These sounds give a voice to the secret activity of the inner life processes of the plant.*

*The viscous and fluid aesthetic of the sound material is an attempt to capture something of both the dynamism of the processes that maintain life and the ever-changing, silken thread that is the presence of life, the life force itself. The fact that the sound material is generated on the basis of meteorological conditions is a way of drawing as tightly as possible the bond between the processes of nature and the processors of the Reeds installation. The sound material cannot then be avoided, being the voice of the processes of nature.*

*Sound and music is in many ways a unique media, for it is not an external artefact. Sound literally penetrates the body. It is also impossible to concretely tie composed sound or music to a representation of anything beyond a communication of emotional states and journeys.*

*As an artist my interest lies in exploring ways of contextualising digital art processes within the natural organic environment. I have little interest in the purely synthetic, that is the synthesis of sound or images from a purely academic or theoretical viewpoint; but prefer, as is illustrated in the Reeds project, to take a fundamentally organic source as the basis for the synthesis process. In so doing, I hope that some quality of that organic material will permeate the work, thereby bringing the synthetic output at least a small way towards the organic world, and therefore within the human context.*

### 4.5.2 The Technology

The technological structure of the Reeds installation is based on the concept of the organic life cycle. This design approach mirrors the closed causal loop, applied as the foundation for the interactive, responsive environments discussed earlier in this document. The weather data forms the generative seed of the life cycle. Gathered by small remote weather stations (installed in the pods floating on the lake), the momentary weather characteristics are transmitted to a land-based computer, which having analysed the incoming data uses the results to '*conduct*' the sound synthesis software.

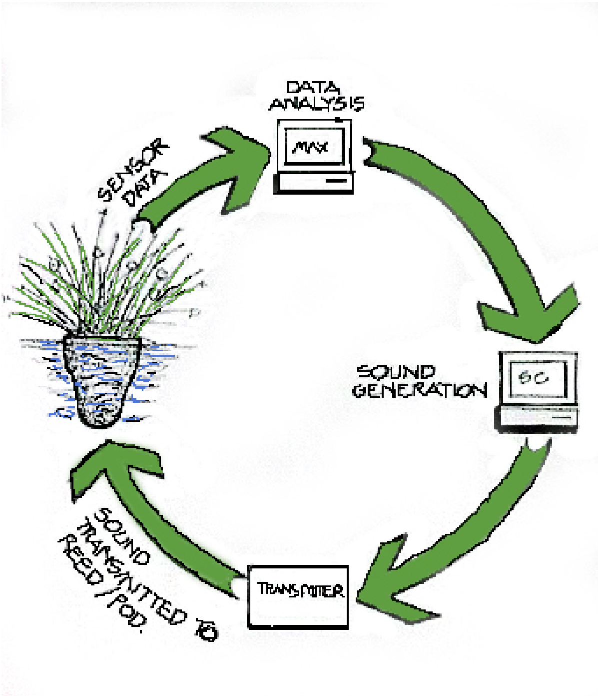



Figure 4.5B The Reeds life cycle.

The sound synthesis software consists of a number of algorithms that generate music in realtime, producing eight channels of digital audio. These eight channels of sound (which emanate from the reed pods) are then broadcast back out to the reed pods using Sennheiser EW300 in-ear monitoring systems. These systems allow the broadcast of high quality stereo audio. The return of the audio signal to the reed pods, and its dispersion to the listener/observer completes the life cycle. This cycle is illustrated in Figure 4.5B.

To flesh out this technological life cycle, I will separate the process into stage headings as follows:

## Collection of Weather Data.

Each weather station reports four pieces of weather data:

- Wind Speed
- Wind Direction
- Temperature
- Solar Radiation

This data is collected using custom built weather stations comprising sensors manufactured by Davis[38] corporation in the USA, and data processing, transmission and reception units designed specifically for this project by Microscan in Adelaide, Australia.

The weather sensors output a sliding voltage scale, which represents their current state, with the exception of the wind speed sensor, the output of which is calculated on the basis of the number of rotations per 1.25 milliseconds (one rotation equals 1.00615 meters of air movement). A data processing board inside the reed pod, converts this data to an ASCII[39] data set in the form:

```
Battery Voltage, Temperature, Solar Radiation, Wind Direction, Wind Speed
```

This data set is transmitted by the weather station once every ninety milliseconds to a land based receiver, which pipes the data into a Macintosh G4 computer as RS232[40] data.

## Weather Data Analysis

The weather data is fed into a software application, developed in Max (see Figure 4.5D), that analyses the incoming data and dynamically scales it, before passing the result in the form of MIDI Continuous Controller messages to a SuperCollider patch, containing six audio synthesis algorithms.

A sub-patch (see Figure 4.5E) of the Reeds Max patch (see Figure 4.5D) polls the serial port every ninety milliseconds to collect the incoming serial data from the weather stations. The weather data is transmitted as machine units, which have a range of 0 to 4095. The weather sensor manufacturers set this range.

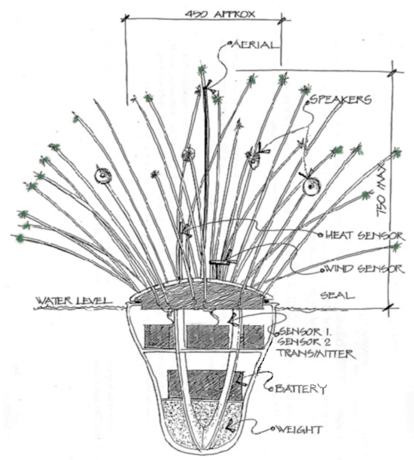



Figure 4.5C Plan view

I felt that I needed to convert the machine units into weather measurements in order to get a clear understanding of the range of activity. The conversion of the data to weather units had the added benefit of providing minimum and maximum measurements to the user

---

[38] Davis Corporation weather station and sensor product information can be found at http://www.davisnet.com/weather/products/index.asp

[39] ASCII is an acronym for the ***A**merican **S**tandard **C**ode for **I**nformation **I**nterchange*. It is a code for representing English characters as numbers. ASCII makes it possible to transfer data from one computer to another. (Webopedia)

[40] RS232 was developed in the 1960s and specifies a serial data communication protocol.

interface. Having the minimum and maximum ranges displayed assisted me in setting the range of variation the sound synthesis software should address, in order to get the best resolution of change in the audio timbres.

The sound synthesis algorithms were mostly based on Fast Fourier Transform[41] (FFT) and Inverse Fast Fourier Transform (IFFT) algorithms.

[41] FFT stands for Fast Fourier Transform. A Fourier Transform captures information about the frequency spectrum of a sound and stores it as time domain data. Inverse Fast Fourier Transforms are an approach for turning time domain signals into frequency spectra, and hence sound. (Roads: 1996)

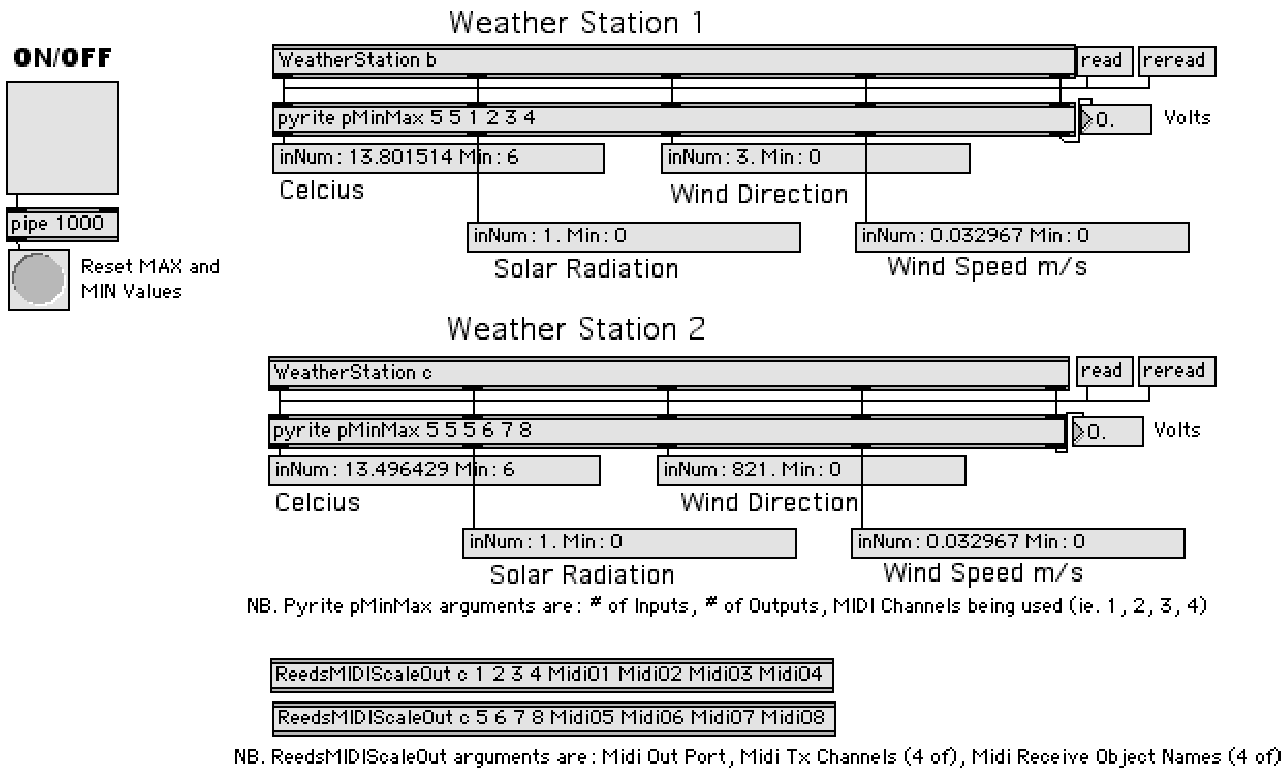



Figure 4.5D The top level Max patch for the Reeds project.

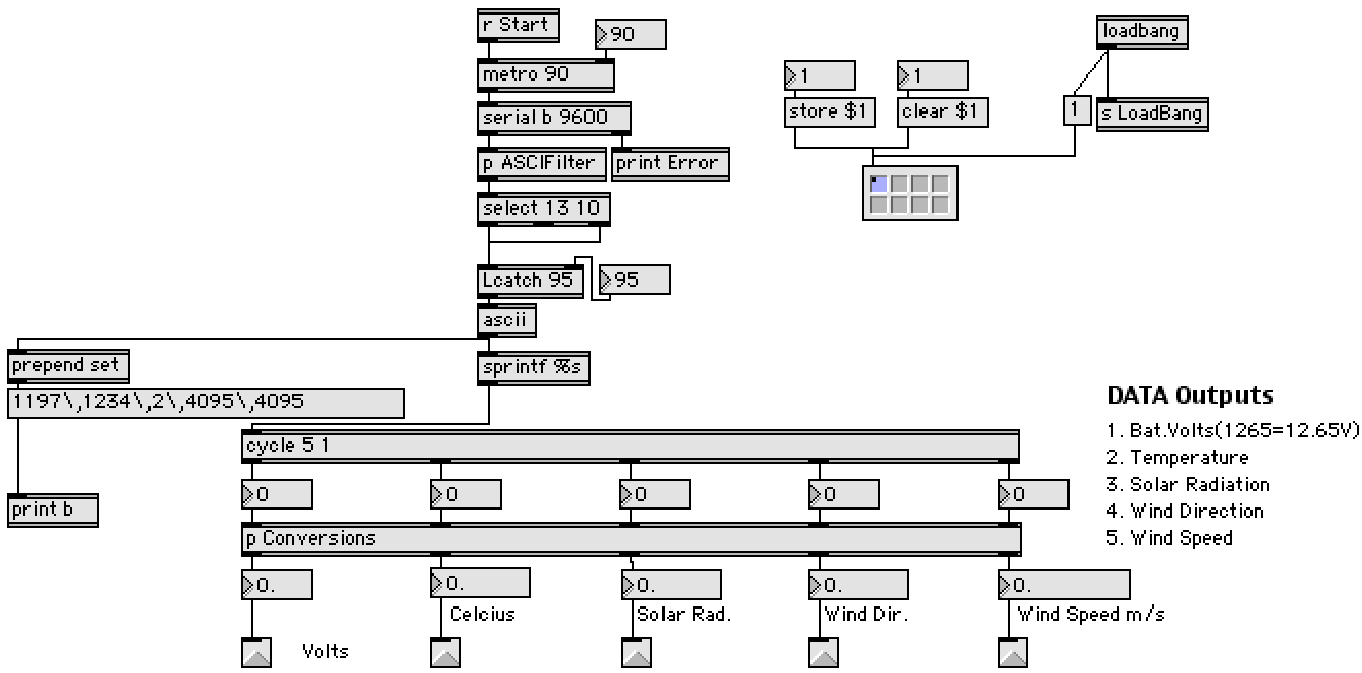



Figure 4.5E The RS232 data input sub-patch, polling the weather stations every 90ms and converting the incoming machine unit (0 to 4095) figures to units of weather.

## Sound Synthesis

A SuperCollider patch creates the audio using FFT and IFFT techniques. A detailed explanation of Fast Fourier Transform techniques is beyond the scope of this exegesis. In brief, a Fourier analysis produces a report of the frequency makeup of a sound, and the associated amplitude of each frequency. An Inverse Fourier technique applies a Fourier analysis to an additive synthesis process, where oscillators recreate the sound using the frequency and amplitude information in the analysis file.

Each of the instrument algorithms is allocated one or more of the weather data streams (i.e. instrument 1 uses wind speed and solar radiation from weather station number 1) which control variables within the algorithm, thereby changing the pitch, texture or intensity of the sound. Instruments one and two (sig1, sig2) produce a stereo signal that is panned across the installation, whilst the other four instruments (sig3, sig4, sig5, sig6) produce a single audio channel. The audio is directed to the eight analogue audio outputs of a Digidesign DIGI001[42] digital audio interface.

Each of the sound algorithms produces differing timbres. They are designed to augment one another, and to produce a range of timbres from gentle, water drop like sounds to roaring wind like sounds. The sound is generally mapped to make the density of sound follow the wind speed, and the weight of sound (pitch for instance) follow the solar radiation readings.

See **Mapping**, below for more detail on the sound mapping approach.

## Broadcast

The audio signals produced by the SuperCollider software were fed to four Sennheiser EW300 In-Ear Monitor[43] transmitters. The EW300 transmitters each broadcast stereo audio of high quality. Sennheiser EK300 stereo receivers, installed in six of the reed pods receive the broadcast signal (each receiver has its own reception frequency matched to one of the four broadcast frequencies). The stereo signal of each receiver is then separated into its two mono components, which are fed to the two adjacent reed pods. See Figure 4.5F for details of the channel allocation and spatialisation.

## Dispersion

Each sounding reed pod contains a battery powered 40-Watt amplifier. The amplifier feeds five loudspeakers: one ten inch full range (20Hz – 20KHz) Misco waterproof loudspeaker, (built into a hat on top of the Reed pod) and four small 40 mm speaker drivers (clipped to the reed stems) which are fed via a crossover, to ensure they only receive signals over 2000Hz. The main speaker carries the full range signal, whilst the smaller 40mm drivers carry the high frequency material that give the crisp edge to the

[42] Digidesign DIGI001 - http://www.digidesign.com/products/digi001/
[43] The Sennheiser EW300 is designed to transmit foldback signals directly into the ear of a musician for the purposes of stage monitoring. See http://www.sennheiser.com

sound. The speakers are placed in a position that allows the sound to bounce across the water surface.

The eight channels of sound were dispersed across the reed pods so that panning a stereo signal (sig01, sig02) would cause the sound to travel across the installation space.

Two of the software synthesis instruments output a stereo signal, which was allocated in each case to loudspeakers in different reed pod clusters. Spreading the stereo signal across the installation created a sense of movement, of sound journeying across the installation space, and also helped to establish homogeneity of the sound world of the entire installation. The other four synthesis instruments output mono channels. They were grouped according to their timbrel quality so that the texture of the soundscape varied as one walked around the shore from one side of the installation to the other.

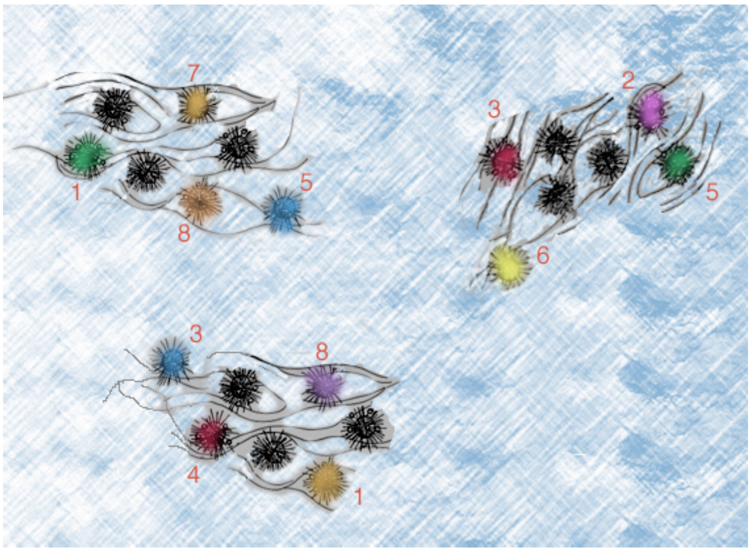



Figure 4.5F The sound dispersion allocation for Reeds
(Colour dots = pods with sound system: like colours = same audio channel)

The coloured dots in Figure 4.5F show the sounding pods, and their audio output allocation, and the table below, Figure 4.5G indicates the allocation of synthesis and audio channels.

| **Synthesis Instrument** | **DIGI001 output** | **Broadcast Channel** |
|---|---|---|
| Sig01 | Output 1, Output 2 | 1, 2 |
| Sig02 | Output 3, Output 4 | 3, 4 |
| Sig03 | Output 5 | 5 |
| Sig04 | Output 6 | 6 |
| Sig05 | Output 7 | 7 |
| Sig06 | Output 8 | 8 |

Figure 4.5G Synthesis and Broadcast Channels.

The amplitude of the audio signals was set so that the sounds of the Reeds installation were embedded within the overall soundscape of the site. The natural soundscape for the lake in the Botanic Gardens, Melbourne, consisted of a wide range of birds sounds (calls, landing on water, fighting etc), the sounds of people at the cafe and walking around the lake, children playing, trucks and cars on the expressway on the other side of the Yarra river, small boat horns, overhead aircraft, and other less frequent momentary sounds.

The reed pods that contained audio equipment were fitted with light sensitive switches so that they turned on at dawn and off at dusk, thereby conserving battery power, and allowing the wildlife the tranquillity of the night.

### 4.5.3 Mapping

The weather data was mapped to the six sound synthesis algorithms as follows:

```
// weather station 1: temperature; UNASSIGNED!
in1Args = [1, 10, 2000, 'exponential', 1];

// weather station 1: solar radiation; mapped to sig1: centerFreq;
in2Args = [2, 100, 8000, 'exponential', 1];

// weather station 1: wind direction; mapped to sig1: clockRate;
in3Args = [3, 1, 200, 'exponential', 1];

// weather station 1: INVERSE wind direction; mapped to sig2: amp3 - varies amplitude envelope
in3Args1 = [3, 20, 1, 'exponential', 3];

// weather station 1: wind speed; mapped to sig2: src1
in4Args = [4, 2, 1000, 'exponential', 1];

// weather station 2: temperature; mapped to sig3: src
in5Args = [5, 10, 100, 'exponential', 1];

// weather station 2: temperature; sets the amplitude in sig3 in inverse proportion to impulse frequency
in5Args1 = [5, 8, 4, 'exponential', 1];

// weather station 2: solar radiation; mapped to sig4: src
in6Args = [6, 10, 10000, 'exponential', 1];

// weather station 2: wind direction; mapped to sig5: src
in7Args = [7, 20, 160, 'linear', 1];

// weather station 2: wind speed; mapped to sig6: Impulse.ar(freq);
in8Args = [8, 0.2, 2.0, 'linear', 1];
```

All of the sound algorithms used Fast Fourier Transform (FFT) or Inverse Fast Fourier Transform (IFFT) synthesis methods. The idea was to auralise the time domain information coming from the weather stations, i.e. current wind speed, solar radiation, wind direction and temperature. This mapping converted the changes in weather conditions into the impulse rate used as the source for the IFFT stage. The impulse rate is the main characteristic in defining the IFFT output. It is defined as follows in sig3 (the third instrument):

```
// inverse transform
     out = IFFT.ar(fftsize, 0, cosineTable, nil, window, src, 0);
```

where `src` is provided by the temperature sensor on weather station 2.

The weather characteristics were chosen both for their importance to plant life (they are the main characteristics that define the growth rate of plants, with the exception of rain), and for their compositional value.

Wind speed and wind direction change dynamically over large ranges in very short periods of time. Small gusts can swing through 360 degrees, and range widely in speed.

Solar radiation changes in a much more gradual way. I had expected the solar radiation to be the most placid variable, changing very gradually over the duration of the day, remaining high during the peak sunshine hours and then diminishing. I had thought that the solar radiation would define the form of the entire day, creating a kind of pedal point for the other more dynamic variations. I was greatly surprised to find that the solar radiation was in fact one of the most dynamic variables. Clearly the human eye constantly adjusts for variations in solar radiation, something a sensing instrument does not do. So, cloud movement, and other changes in the weather saw the solar radiation constantly sliding up and down its range. Obviously, some days were sunnier than others, and therefore the range of movement was generally higher on those days than when the skies were overcast, but the dynamics of change were consistent in all weather conditions. Unlike the wind characteristics, solar radiation never jumped from one point to another, it moved gradually up and down in a step like manner.

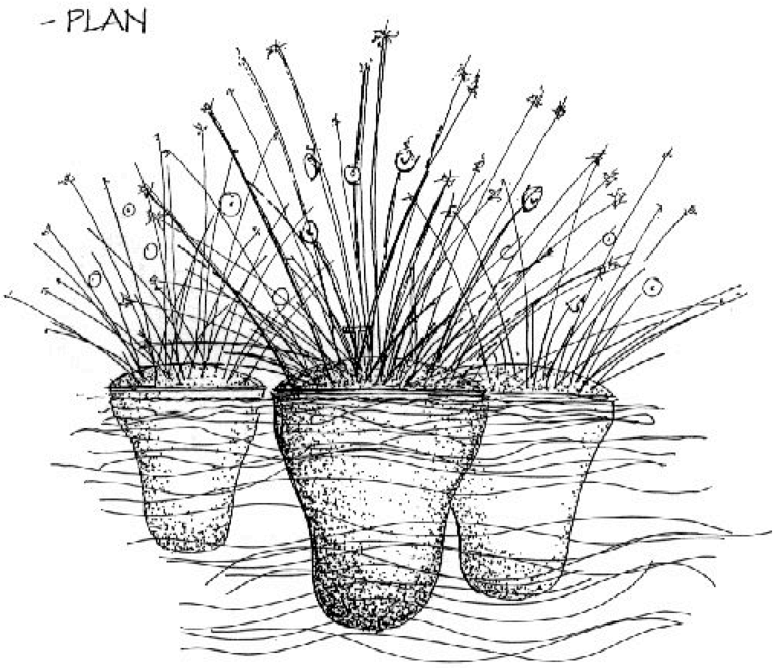



Figure 4.5H One of the original concept drawings for Reeds.

Temperature, as I had expected, changes gradually, moving up and down in small sequential steps. Temperature was the most graceful of the weather characteristics being sensed for the Reeds project, generally forming an envelope shaped by cool evenings, followed by rising temperatures during the day as the sun heated the earth, followed by a reduction in temperature leading towards dusk and into the night.

As indicated above, the weather characteristics were intended to have different temporal structures, so that the more gradual, step like features (temperature and solar radiation) would provide a gradually varying underscore on which the faster changing features (wind speed and direction) would add points of interest and orchestration dynamics.

#### 4.5.4 Summary

Reeds continued my exploration of organic, natural controllers for interactive sound installations. It was a small diversion from my central focus of unencumbered human interaction, but did provide a platform for the testing of multiple, chaotic controllers, working simultaneously to generate a soundscape which continued to draw on the dynamic orchestration principles I had developed in MAP2.

I visited the Reeds installation every day, and as the Melbourne Botanic Gardens exhibition occurred during the season of spring, I observed a wide range of weather patterns. I noted, for instance, situations where the sunshine and temperature were high, but the wind speed was almost non-existent; by contrast I noted similar temperature and sunshine levels on days when the wind contained strong gusts and violent changes of direction. All of these provided very different outcomes, but noting that the weather is naturally scoped, (it is unlikely the temperature will ever get to 60 degrees Celsius in Melbourne, nor is the wind speed likely to exceed 120 mph), the input variables for the synthesis process followed suit. Observing this phenomenon helped me realise that this was also the case within my previous installations using human movement and behaviour patterns as the control input. A human being can only jump so far, or run so fast etc.

From a musical composition perspective, it was useful to consider the different temporal structures associated with each of the weather characteristics in Reeds. I found the different rates of change of each of the weather inputs helped create homogeneity in the soundscape. It was aurally pleasing to have some elements that changed rapidly and others that evolved slowly over time. I started to consider how this approach might be applied in combination with the dynamic orchestration technique to generate a more richly evolving soundscape within my interactive, responsive environments.

Reeds illustrates my ongoing interest in making the technology invisible. I went to great lengths to develop battery-powered amplification systems, and broadcast systems, so that the Reed pods were entirely self-contained, and displayed no obvious source (except for the weather stations and the small high-frequency loudspeakers) for the sounds they emitted.

It was obvious, however, if one sat and watched the installation for even a short time, that the changes in the weather stations (you could easily see the anemometer) caused variations in the musical output.

Figure 4.5I The pods for Reeds being assembled on the edge of the lake for the exhibition on the Ornamental Lake at the Royal Botanic Gardens, Melbourne,

Figure 4.5J The pods for the Weather Station reeds being assembled on the edge of the lake, for the exhibition on the Ornamental Lake at the Royal Botanic Gardens, Melbourne.

The SuperCollider software developed for the Reeds installation is documented in Figures 4.5.K to 4.5L

```
var     TempMAX = 4095, TempMIN = 0,                    -- temperature
        solarMAX = 4095, solarMIN = 0,                  -- Solar Radiation
        windDirMAX = 4095, windDirMIN = 0,              -- Wind Direction
        windSpeedMAX = 4095, windSpeedMIN = 0,          -- Wind Speed
        scaleMin = 0, scaleMax = 16383,
        oldVolt = 0;
-- the scale of 0 to 4095 is the 12 bit high resolution pitch bend range

var Out1, Out2, Out3, Out4;

-- Set all the reporting window min and max figures
bang {          TempMIN = 6, TempMAX = 40,                      -- temperature
                solarMIN = 0, solarMAX = 2400,                  -- Solar
                windDirMIN = 0, windDirMAX = 4095,              -- Wind Dir
                windSpeedMIN = 0, windSpeedMAX = 300,           -- Wind Speed
                oldVolt = 0; }

-- initialise all the MIDI outputs channels 1 2 3 4 WS1 || channels 5 6 7 8 WS2
init { arg inName1 inName2, inName3, inName4;
                Out1 = ((\Midi0).spell $ inName1.spell).unspell;
                Out2 = ((\Midi0).spell $ inName2.spell).unspell;
                Out3 = ((\Midi0).spell $ inName3.spell).unspell;
                Out4 = ((\Midi0).spell $ inName4.spell).unspell; }

-- Methods for checking input data and sending out updated messages when new data arrives
inlet1 { arg inNum;                                             -- Volts
                IF inNum !== oldVolt THEN
                oldVolt = inNum;
                oldVolt.out(5);
                END.IF }

inlet2 { arg inNum;                                             -- Temperature
                var a;
                a = inNum.asInt;
                a.map( TempMIN, TempMAX, 2000, scaleMax).out(Out1);
                [\set 'inNum:', inNum, 'Min:', TempMIN].out(1); }

inlet3 { arg inNum;                                             -- Solar Radiation
                var a;
                a = inNum.asInt;
                a.map(solarMIN, solarMAX, 3500, scaleMax).out(Out2);
                [\set 'inNum:', inNum, 'Min:', solarMIN].out(2); }

inlet4 { arg inNum;                                             -- Wind Direction
                var a;
                a = inNum.asInt;
                a.map(windDirMIN, windDirMAX, scaleMin, scaleMax).out(Out3);
                [\set 'inNum:', inNum, 'Min:', windDirMIN].out(3); }

inlet5 { arg inNum;                                             -- Wind Speed
                var a;
                a = inNum.asInt;
                a.map(windSpeedMIN, windSpeedMAX, scaleMin, scaleMax).out(Out4);
                [\set 'inNum:', inNum, 'Min:', windSpeedMIN].out(4); }
```

Figure 4.5K  The pyrite script used in the Reeds Max patch

```
/* NB!!!

1. the "range" of sig2 has been adjusted from 1600 to 1000. To fine tune it further, use the 3rd
(maxval) argument to the variable in4Args.

2. "reverb" in sig2 is the comb decay time. this has been adjusted from 3secs to 2secs. A variable
(sig2DecayTime) has been added for fine tuning the reverberation time.

3. The minval and maxval args in the variable in5Args1 provide amplitude response adjustments in sig3.
because this is mapped in inverse proportion to the impulse frequency, minval should be a higher value than maxval.

*/

var maxAmp1, maxAmp2, maxAmp3, maxAmp4, maxAmp5, maxAmp6;
var sig2DecayTime;
var in1Args, in2Args, in3Args, in4Args, in5Args, in5Args1, in6Args, in7Args, in8Args, in3Args1;
var w, amp1, amp2, amp3, amp4, amp5, amp6;
var in, sig1, sig2, sig3, sig4, sig5, sig6;

// peak amplitudes for each instrument (adjust these values until you get the right balance of signals)
maxAmp1 = 0.5; //0.25;              //maxAmp1 = 0.3; //0.25;
maxAmp2 = 24;
maxAmp4 = 0.5;                      //maxAmp2 = 24; //40//30;
maxAmp3 = 0.05; //12;               //maxAmp3 = 0.08; //12;              in5Args
maxAmp5 = 24; //60;                 //maxAmp5 = 24; //60;
maxAmp6 = 12; // 20;                //maxAmp6 = 12; // 20;

// decay time of both comb filters in sig2
sig2DecayTime = 1;

// set the effective ranges and response characteristics for each controller in the variables inArgs1-inArgs8
// parameters are midi chan, minval, maxval, warp, lag time
// values from the orignal patches are listed in the comment lines

// weather station 1: temperature; UNASSIGNED! - where do you want it? channel switching?
in1Args = [1, 10, 2000, 'exponential', 1];

// weather station 1: solar radiation; mapped to sig1: centerFreq; args are (100, 8000, 'exponential', 0.1) in original
in2Args = [2, 100, 8000, 'exponential', 1];

// weather station 1: wind direction; mapped to sig1: clockRate; args are (1, 200, 'exponential', 0.1) in original
in3Args = [3, 1, 200, 'exponential', 1];

// weather station 1: INVERSE wind direction; mapped to sig2: amp3 - varies amplitude of signal;
in3Args1 = [3, 20, 1, 'exponential', 3];

// weather station 1: wind speed; mapped to sig2: src1; args are (2, 2000, 'exponential', 0.1) in original
in4Args = [4, 2, 1000, 'exponential', 1];

// weather station 2: temperature; mapped to sig3: src; args are (10, 1000, 'exponential', 0.1) in original
in5Args = [5, 10, 100, 'exponential', 1];       //[5, 10, 1000, 'exponential', 2]

// set these values to map amplitude in sig3 in inverse proportion to impulse frequency
in5Args1 = [5, 8, 4, 'exponential', 1];

// weather station 2: solar radiation; mapped to sig4: src; args are (10, 10000, 'exponential', 0.1) in original
in6Args = [6, 10, 10000, 'exponential', 1];

// weather station 2: wind direction; mapped to sig5: src; args are (20, 200, 'linear', 0.1) in original
in7Args = [7, 20, 160, 'linear', 1];                     //in7Args = [7, 20, 200, 'linear', 2];

// weather station 2: wind speed; mapped to sig6: Impulse.ar(freq); args are (0.2, 2.0, 'linear', 0.1) in original
in8Args = [8, 0.2, 2.0, 'linear', 1];
```

```
// GUI
w = GUIWindow.new("mixer", Rect.newBy(632, 109, 179, 133));
amp1 = SliderView.new(w, Rect.newBy(20, 15, 16, 93), "SliderView", maxAmp1, 0, maxAmp1, 0.01, 'linear');
amp2 = SliderView.new(w, Rect.newBy(45, 15, 16, 93), "SliderView", maxAmp2, 0, maxAmp2, 0.01, 'linear');
amp3 = SliderView.new(w, Rect.newBy(70, 15, 16, 93), "SliderView", maxAmp3, 0, maxAmp3, 0.01, 'linear');
amp4 = SliderView.new(w, Rect.newBy(95, 15, 16, 93), "SliderView", maxAmp4, 0, maxAmp4, 0.01, 'linear');
amp5 = SliderView.new(w, Rect.newBy(120, 15, 16, 93), "SliderView", maxAmp5, 0, maxAmp5, 0.01, 'linear'
amp6 = SliderView.new(w, Rect.newBy(145, 15, 16, 93), "SliderView", maxAmp6, 0, maxAmp6, 0.01, 'linear'

// input
in = {
	arg inArgs;
	MIDIPitchBend.kr(inArgs.at(0), inArgs.at(1), inArgs.at(2), inArgs.at(3), inArgs.at(4));
	};
```

continued on the next page …..

```
// synths
sig1 = {

	var clockRate, clockTime, clock, centerFreq, freq, panPos, patch;

	clockRate = in.poll(in3Args);
	clockTime = clockRate.reciprocal;
	clock = Impulse.kr(clockRate, amp1.kr * in.poll(in2Args) / 8000);
	centerFreq = in.poll(in2Args);
	freq = Latch.kr(WhiteNoise.kr(centerFreq * 0.5, centerFreq), clock);
	panPos = Latch.kr(WhiteNoise.kr, clock);
	patch = CombN.ar(
			Pan2.ar(
				SinOsc.ar(
					freq, 0, Decay2.kr(clock, 0.1 * clockTime, 0.9 * clockTime) ),
				panPos ),
			0.3, 0.3, 2 );
	patch
};

sig2 = {

	var numffts, fftsize, fftoffsets, window, cosineTable, dur1, dur2;

	numffts = 2;		// number of overlapped FFTs
	fftsize = 1024;		// length of FFT buffer

	// make offsets for staggering overlapped FFTs
	fftoffsets = Array.series(numffts, 0, fftsize/numffts);
	window = Signal.hanningWindow(fftsize);		// make a signal analysis/synthesis window
	cosineTable = Signal.fftCosTable(fftsize);	// make cosine table required for FFT

	dur1 = fftsize / Synth.sampleRate;
	dur2 = fftoffsets.at(1) / Synth.sampleRate;

	{	var src1, src2, src3, ifft, out;

		src1 = Dust2.ar(in.poll(in4Args), amp2.kr);			//(in4Args)
		src2 = CombN.ar(src1, dur1, dur1, sig2DecayTime) * WhiteNoise.ar;
		src3 = CombN.ar(DelayN.ar(src1, dur2, dur2), dur1, dur1, sig2DecayTime) * WhiteNoise.ar;

		// inverse transform
		ifft = IFFT.ar(fftsize, fftoffsets, cosineTable, nil, window, [src2,src3], 0);
		out = [Mix.ar(ifft.real), Mix.ar(ifft.imag)];
		out.real

	}.value
};

sig3 = {					//weather station 2: temperature

	var fftsize, window, cosineTable, dur, src, out;

	fftsize = 512;		// length of FFT buffer

	window = Signal.welchWindow(fftsize);		// make a signal analysis/synthesis window
	cosineTable = Signal.fftCosTable(fftsize);	// make cosine table required for FFT
	dur = fftsize/Synth.sampleRate;
	src = CombN.ar(Impulse.ar(in.poll(in5Args), in.poll(in5Args1)), dur, dur, 3, (amp3.kr * in.poll(in3Args1))); //in5Ar

	// inverse transform
	out = IFFT.ar(fftsize, 0, cosineTable, nil, window, src, 0);
	out.real;
};
```

```
sig4 = {

      var fftsize, window, cosineTable, src, out;

      fftsize = 512;        // length of FFT buffer
      window = Signal.hanningWindow(fftsize);         // make a signal analysis/synthesis window
      cosineTable = Signal.fftCosTable(fftsize);      // make cosine table required for FFT
      src = Impulse.ar(in.poll(in6Args), amp4.kr * (in.poll(in6Args) / 10000));  // weather station 2: solar radiatior

      // inverse transform
      out = IFFT.ar(fftsize, 0, cosineTable, nil, window, src, 0);
      out.real;
};

sig5 = {               // weather station 2: wind direction;

      var fftsize, window, cosineTable, src, out;

      fftsize = 512; // length of FFT buffer
      window = Signal.hanningWindow(fftsize); // make a signal analysis/synthesis window
      cosineTable = Signal.fftCosTable(fftsize); // make cosine table required for FFT
      src = Impulse.ar(in.poll(in7Args), amp5.kr * (in.poll(in2Args) / 8000));
      out = IFFT.ar(fftsize, 0, cosineTable, nil, window, src, 0); // inverse transform
      out.real; // out
 };

sig6 = {

      var impulse, a, b, c, d, e, f;

      impulse = Impulse.ar(30, amp6.kr);

      a = IFFT.ar(64, 0, Signal.fftCosTable(64), nil, Signal.hanningWindow(64), impulse, 0);
      b = IFFT.ar(128, 0, Signal.fftCosTable(128), nil, Signal.hanningWindow(128), impulse, 0);
      c = IFFT.ar(256, 0, Signal.fftCosTable(256), nil, Signal.hanningWindow(256), impulse, 0);
      d = IFFT.ar(512, 0, Signal.fftCosTable(512), nil, Signal.hanningWindow(512), impulse, 0);
      e = IFFT.ar(1024, 0, Signal.fftCosTable(1024), nil, Signal.hanningWindow(1024), impulse, 0);
      f = IFFT.ar(2048, 0, Signal.fftCosTable(2048), nil, Signal.hanningWindow(2048), impulse, 0);

          TSpawn.ar({
              var n, env;

              n = [a, b, c, d, e, f].choose;
              env = Env.linen(0.01, rrand(0.1, 2.0), 0.05, 1); // grain envelope
              EnvGen.ar(env, n.real); // out

          }, 1, nil, // no. of channels, no. of repeats
          Impulse.ar(in.poll(in8Args)));
};

// 8 channels out
{
      [
           sig2.value.at(0), sig1.value.at(1),
           sig2.value.at(1),sig6.value,
           sig1.value.at(0),
           sig3.value,
           sig5.value,
           sig4.value
      ]
}.scope;
w.close;
)
```

Figure 4.5L  The last four pages contain the SuperCollider code for sound synthesis in Reed

## *4.6 GESTATION*

Gestation is the final project to be documented in this exegesis. It returns to the exploration of human movement as the controller for interactive, responsive environments. In so doing it develops the dynamic orchestration principles I established for MAP2, and draws on experimentation in Reeds with multiple independent control inputs forming a multi-dimensional soundscape.

Gestation is presented in two separate but interlinked galleries. One gallery contains an interactive, responsive sound environment, which acts as the generation source for outcomes in both galleries. The second gallery is a visual environment, containing large projected images of foetuses, visualised using ultrasound.

Sound, as a medium for defining experience, is the common thread between the two galleries. Movement and behaviour patterns in the sound gallery generate an environment that gives birth to 'new life' in the second gallery. The 'new life' being made visible through the application of sound (ultra-sound).

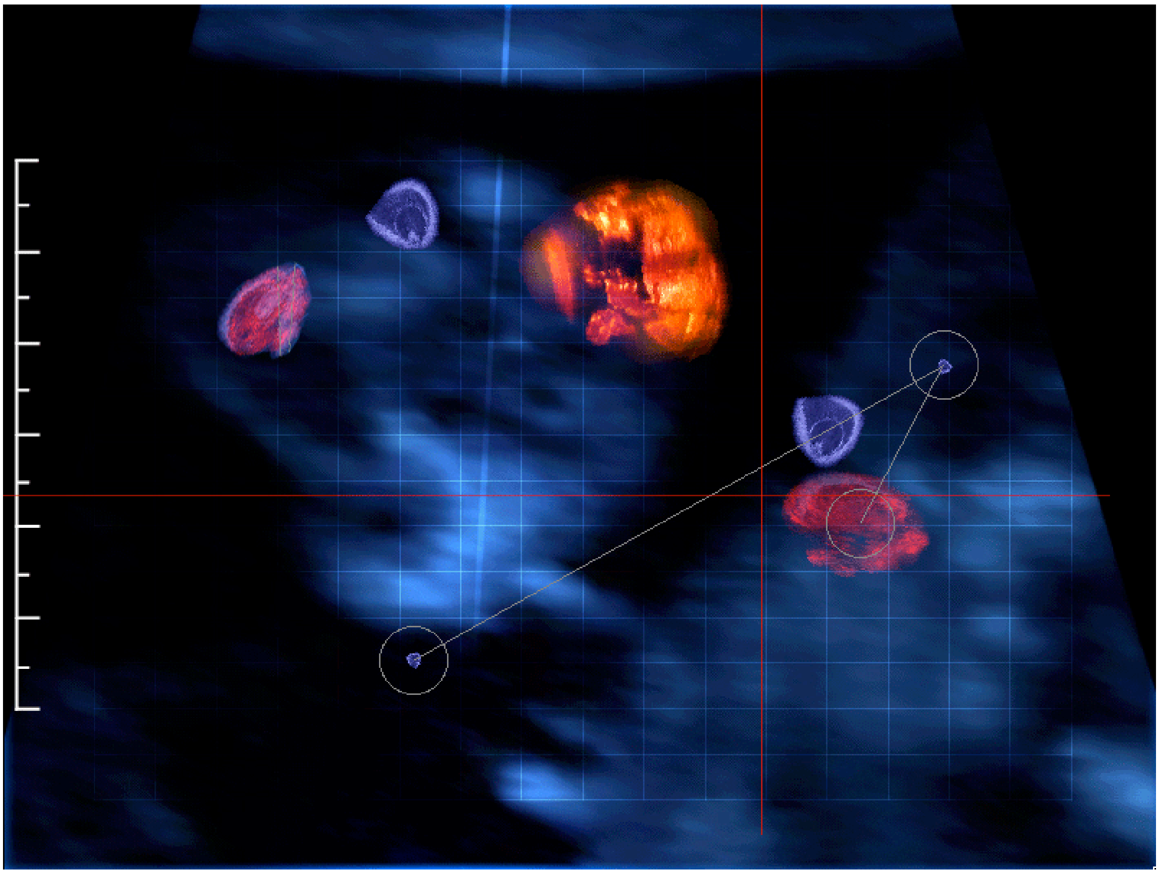


Figure 4.6A A screen shot from Gestation, developed in collaboration with Kat Mew.

### 4.6.1 Program Notes

*The computer has opened up whole new contemporary art genres where primary composition material can be drawn from any source, and once digitised, becomes a fluid and viscous medium.*

*My interest lies in placing the exploration of the potential of these technologies within an organic and human framework.* ***Gestation*** *focuses on creating an immersive environment that responds to the movement and behaviour patterns detected within it. The body becomes the controller. The organic process of human exploration, cognition and response becomes the central influence in defining the output of the interactive process. As with my previous installations,* ***Gestation*** *is inspired by the work of Myron Krueger, who comments:*

*In the environment, the participant is confronted with a completely new kind of experience. He is stripped of his informed expectations and forced to deal with the moment on its own terms. He is actively involved, discovering that his limbs have been given new meaning and that he can express himself in new ways. He does not simply admire the work of the artist; he shares in its creation. (*Krueger:1976:84*)*

***Gestation*** *is an interactive responsive environment. It contains two integrated spaces. One gallery contains a surround sound field, generated in real time using video sensing equipment (visible to visitors only as a small CCTV video camera in the middle of the roof) that maps the behaviour and movement patterns of the visitors on to real-time audio algorithms providing a tight gestural relationship with their movement and behaviour patterns. No pre-recorded material is being used in the generation of the sounds, they are all generated algorithmically in realtime, creating evolving streams of sound.*

*In the second gallery, a large projected image represents the development of new human life in response to activity in the first gallery. The image background represents a sea of life forming cells. Additional layers are formed by the development of new foetuses. Each foetus starts to grow at the point at which particularly dynamic activity is sensed in the first gallery.*

*The aesthetic of the sound environment is a carefully tended intimately textured sound. It is intended to create a viscous, fluid environment for the 'making of life'. The qualities of this sound change in relation to the direction and speed of movement and number of people within the space. In addition to the underscore sound, more contained points of interest are tied to the creation of each new foetus, and are associated with the position within the gallery space at which that activity is sensed. The growth sounds express the qualities of life forming: the binding of cells, the development of human form, and the growth of the foetus.*

*Over the last five years I have collected ultra-sound videos from friends and acquaintances that have had children. The videos are all of their first-born children and form the basis of the moving images. The cells begin growth at a point in the two-dimensional grid associated with the sensed movement in the sound gallery, and grow at a rate associated with the dynamic of that activity. Varying rates of growth are associated with thresholds of activity.*

*Participants in the sound gallery cannot see the visual element without leaving the gallery space. They can make life, but not observe it at the same time.*

*The two galleries are detached to illustrate the hidden outcomes of our activities. This approach also allows the visitors to be more deeply engaged in the details of the sound environment, in the hope that they will more consciously engage with the fluidity and variability of the sounds.*

### 4.6.2 The Technology

Gestation continues the evolution of techniques developed in MAP1, MAP2 and Reeds. It implements real-time sound synthesis, controlled by the output of a real-time video tracking system.

Video sensing was once again chosen as the appropriate technique to track the movement and behaviour patterns of people within the exhibition space because it provides a stream of data describing the quality of movement in the installation at all times. On this occasion however, I used an extended `videoIn`[44] object within the Max programming environment in place of the Very Nervous System (VNS). This change was brought about by the fact that the VNS hardware broke down only two weeks prior to exhibition date. It was not possible to repair the VNS prior to the exhibition date, so I resorted to the `videoIn` object.

The structure of the output of the `videoIn` object differs from that of the VNS. Whilst it is possible to define regions in the camera view, the output for the region is not an average of the regional activity, as is the VNS output, but takes the form:

```
zone#   trig   PixelCount   centreX   centreY   Xlo  Xhi  Ylo  Yhi
```

where

- `Zone#` is the zone number associated with the following data;
- `trig` is as follows: 1 = new motion, 0 = continuing motion, -1 = motion ended;
- `PixelCount` equals the total number of changed pixels;
- `centreX centreY` is the centroid of the changed pixels in the zone;
- `Xlo Xhi Ylo Yhi` are the extreme coordinates of changed pixels in the zone;

All X and Y coordinates are local references, that is, they represent a pixel measurement for the top left corner of the associated zone, not the entire camera view.

The nature of the output of the `videoIn` object required a rethink of how I should deal with the video information. While the data outputs of the VNS and the `videoIn` object sound similar, they provide subtly different qualitative information. The VNS output provides positional information, and immediate data about the size of the moving object. It was possible to derive similar information from the `videoIn` object, but the `videoIn` object's output is focused on one point of focus (subject) per zone, rather than per region (of which there may be many when using the VNS). The `videoIn` object provided a centre of activity, which could be the centre of a single human being, or could be a midway point between several human beings in the space. The `PixelCount` equates to the VNS region output, however, here it refers to the total number of changed pixels in the zone, not the difference of light intensity per pixel response which the VNS provides. The `PixelCount` is a gross measurement of change, and whilst it represents the total area of movement since the last video frame, it does not provide the qualitative information of the light level differencing data from the VNS. The VNS output provides qualitative data that describes both the size and speed of the moving object.

---

[44] The `videoIn` object I used for Gestation was an extended version of the original object developed by Peter Elsea, with additional code by Ronald Kuivila from Wesleyan University, USA.

The `PixelCount` and `centreX`, `centreY` data from the `videoIn` object were used to drive the sound synthesis variables.

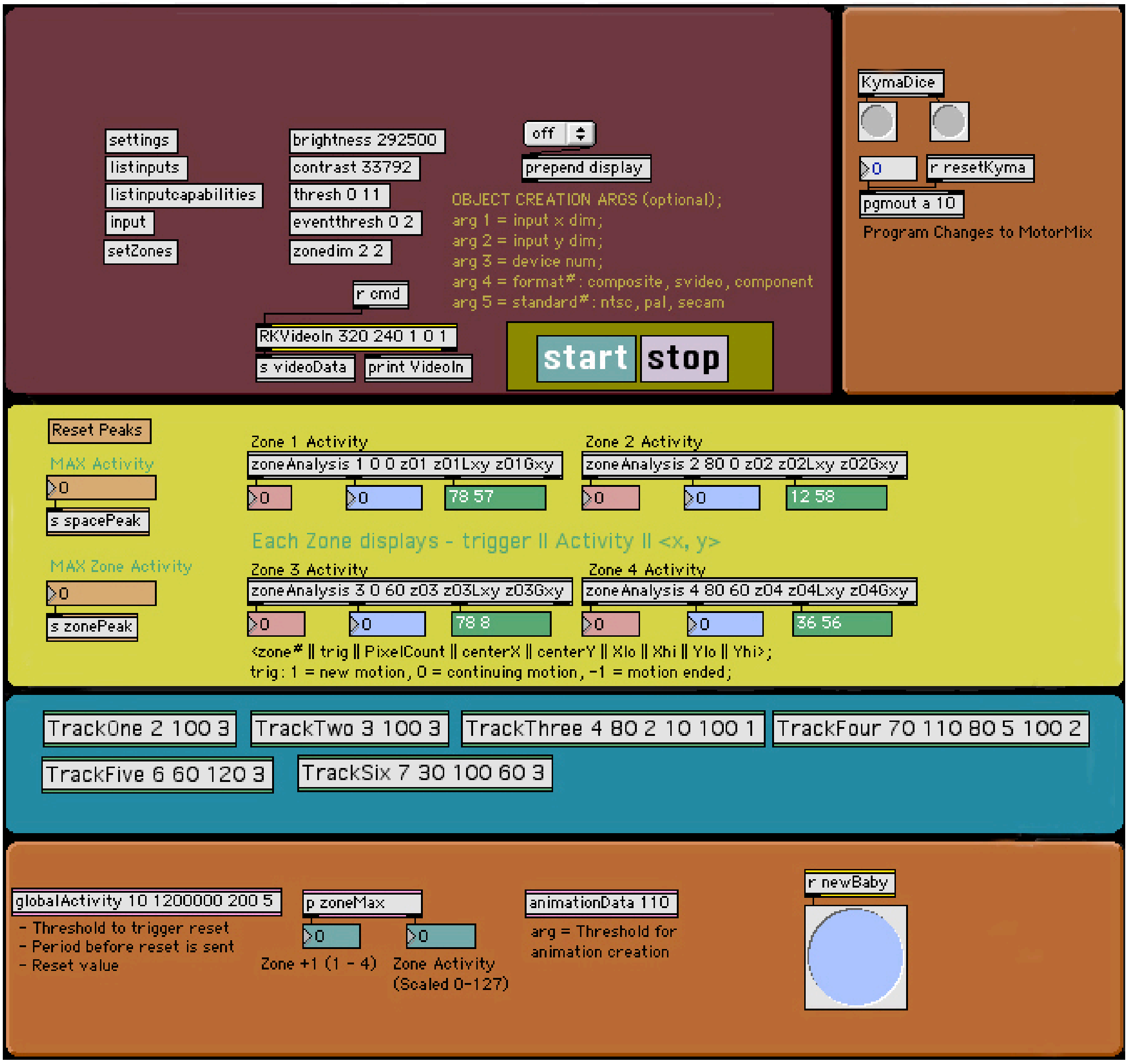



Figure 4.6B The top level Max patch for Gestation.

The sound synthesis for Gestation was achieved using a Capybara 320 sound computation engine system from Symbolic Sound[45]. The Capybara 320 is programmed using the Kyma programming language. The Kyma program is compiled and downloaded to the Capybara hardware at runtime.

During MAP2 and Reeds, I had been frustrated by the processor limitations of even the most powerful Macintosh computers (SuperCollider only runs on Macintosh computers). When dynamically creating synthesis instruments, one can very quickly run out of computing power. I had dealt with this problem in MAP2 and Reeds by limiting the

[45] The Capybara 320 Sound Computation Engine is designed and manufactured by Symbolic Sound - http://www.symbolicsound.com/

number of possible synthesis voices available at any one time, however for Gestation, I felt that limitation was unreasonable because I wanted to more fully explore the ideas of dynamic orchestration. The Capybara system is designed to take additional DSP cards, so it was possible to quite readily commission a system with many times the processing power of a Macintosh computer.

One of my interests common to the projects outlined in this exegesis has been the use of '*found sounds*'[46] as a source for synthesis processes. I feel this helps contextualise the computer music aesthetic within the human experience.

The Capybara system allows any kind of data table to be used as a sound synthesis variable. Whilst the primary variables in the synthesis algorithms were driven by the video sensing system, sounds associated with a child in the womb were used to drive frequency spectra and other characteristics of some of the synthesis outputs. For instance, when a new foetus was created in Gestation, a small baby-like giggle was added to the soundscape. This giggle was in fact constructed from a frequency analysis of a baby girl's giggle, but was re-synthesised in different ways according to the dynamic of the activity that generated the sound. This was achieved by applying the frequency spectra to the resynthesis process at a different rate, and with different frequency weightings according to the initiating activity. There were also delay buffers on the output of the sound which were varied dynamically causing changes in pitch, echo period and feedback or total time for decay to zero.

The Capybara system I used for Gestation had four audio outputs. All sound synthesis algorithms were set up to dynamically pan around the space in accordance with the position of greatest activity. The panning position varied in accordance with the dynamic of activity; some sounds moved away from the subject when they were moving in a very dynamic way while others moved towards them. The spatialisation was dynamically varied in accordance with the level of activity within the installation.

Each of the four zones in Gestation had the same sound synthesis palette. There were in fact four small orchestras of sounds, initiated as required by the sensed activity.

The Kyma programming language provides Smalltalk[47] based scripts as a base level for controlling sounds. It is possible using these scripts to dynamically instantiate algorithms and dispose of them on demand. This process brought me closer to my perceived application of dynamic orchestration than I had been able to get using SuperCollider on a Macintosh computer.

The animated imagery was developed from source videos of Ultra-Sound examinations of pregnant women, which I had collected in recent years. These moving foetus images were digitised and converted to animations, which were placed within a varying background, developed from an Ultra-Sound close up of a child's face. The animated environment ran in Macromedia Director[48]. The position, and growth patterns of the

---

[46] Found Sounds are any sound recorded from the environment in which we live.
[47] Smalltalk is an object oriented programming language developed at the Xerox Palo Alto Research Centre, USA, in the early 1980s
[48] Macromedia Director is an authoring environment for interactive visual applications. http://www.macromedia.com/software/director/

foetuses was controlled using MIDI communication from the sound and video sensing computer, using Ross Bencina's MIDIExtra.

The animations were set up as sprites[49] in Macromedia Director. Seven sequences of growth from a small blue dot to a fully developed foetus were created. The angle of rotation, the rate of growth and the initial position of the foetus were set by the dynamic of activity that created it. Once created, the foetuses are autonomous. The characteristics provided at the time of '*conception*' define its life. A gentle flocking[50] algorithm was introduced to the animation environment so that the foetuses did not remain in a static position, but flocked towards the mean point of the last three points of '*conception*'.

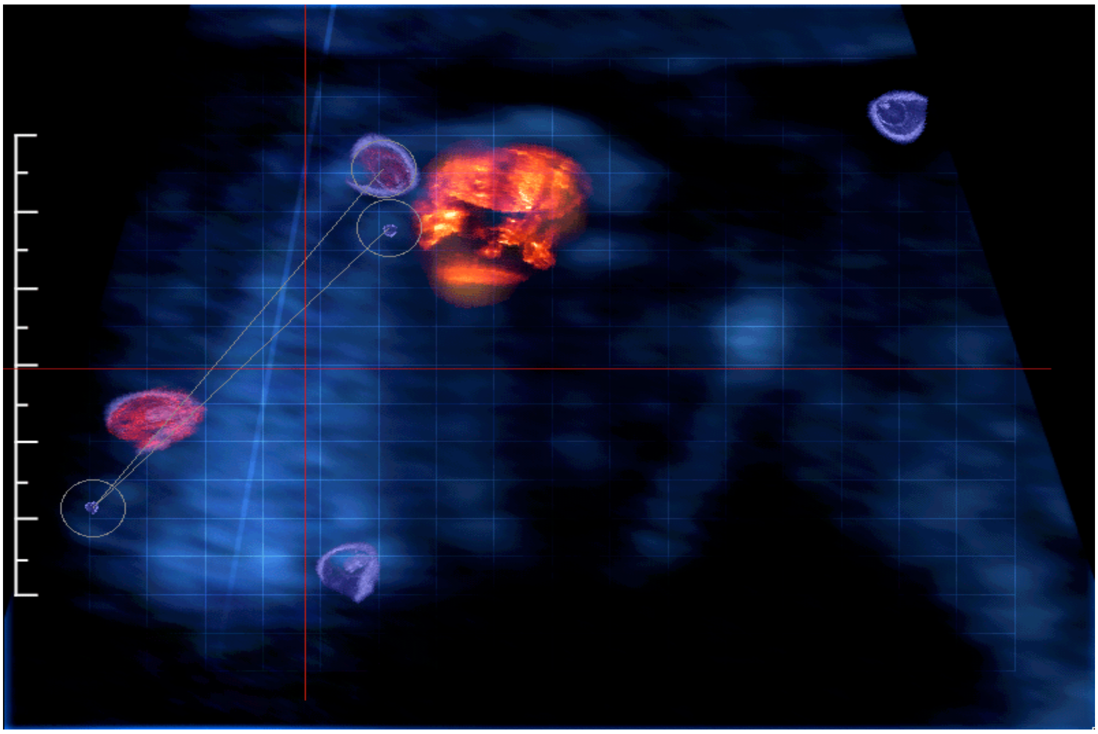


Figure 4.6C Foetus's growing in the Gestation installation.

A scan line, (a light blue line seen in Figure 4.6C), moves across the screen to suggest the ultra-sound technology that provided the images. The cumulative effect of this animation world was a sense of actually seeing the instance of a 'new life' forming, and developing at an accelerated rate of growth in a kind of communal womb.

### 4.6.3 Mapping

Gestation used the dynamic of movement within the installation space to control dynamic orchestration processes, and the position of each movement within a zone (4 zones) to control filter and other synthesis variables. Many of the synthesis variables operated within ranges; the ranges were altered depending on the current dynamic of activity in the installation.

---

[49] A sprite is a separate animation element that can be added to the main canvas and animated in isolation from all other animated elements.
[50] Flocking is a computer algorithm that replicates the behaviour of flocks of animals, such as a flock of birds flying.

The basis for each of six sound algorithms was set-up in a Kyma timeline, which allowed the dynamic control of spatialisation, and the layering of the sounds. The sound algorithms are shown in the timeline in Figure 4.6D with the base sound in track one, and the higher order sounds in tracks two through to six.

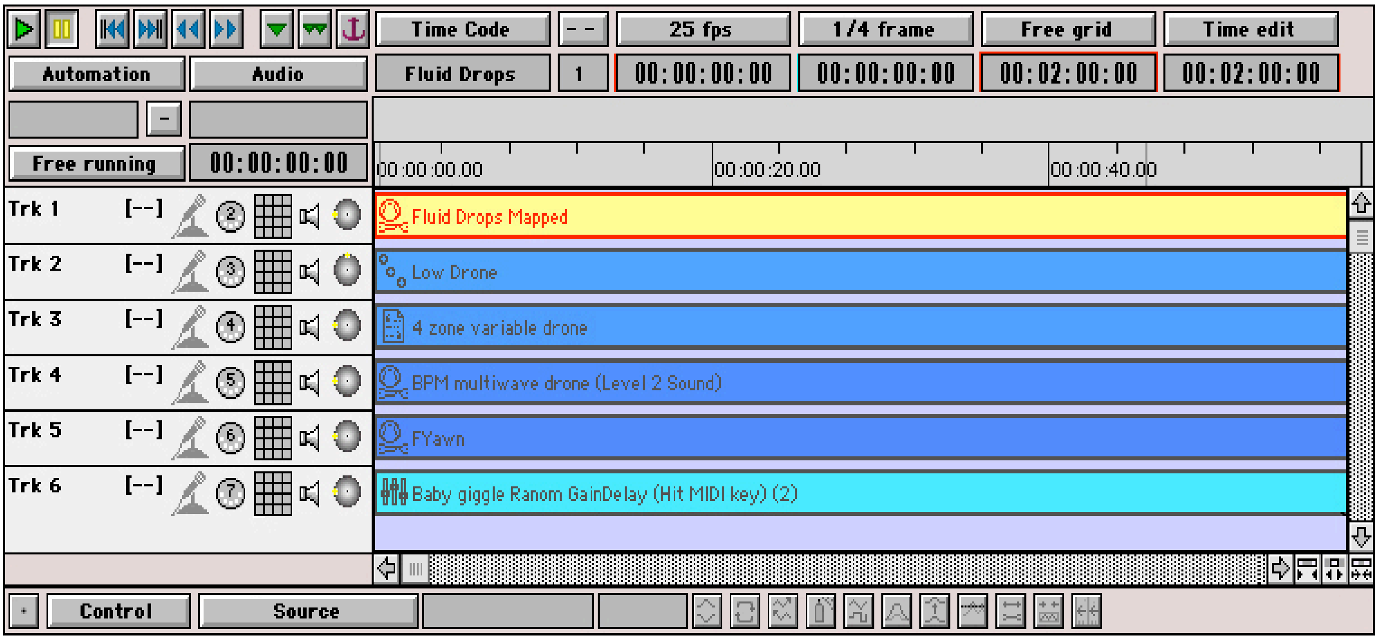



Figure 4.6D The Kyma timeline containing all the synthesis algorithms for Gestation.

The six algorithms developed for Gestation became active at different levels of activity within the installation. The sound aesthetics can be described as follows:

1 A watery drop like sound that created a sense of a fluid, viscous, womb like enclosure.
2 A low frequency bass drone. The position of greatest activity in the installation space caused changes in the sounds texture, while the bandwidth of the signal varied in accordance with the dynamic of activity within the installation.
3 A high frequency drone, which had a female voice like quality. This sound was generated using a vocal sample as an input to an FFT, IFFT additive synthesis process. The spectra of the FFT analysis were varied in accordance with the activity in the installations, generating changes in pitch and texture.
4 A bubbly, dynamic sound formed the fourth layer. Its pitch and dynamic of temporal structure was controlled from the installation.
5 A female breath like sound formed the fifth layer. This sound was based on a granular synthesis process, fed with a female breath sample. The density of grains, the length of the grains, and the temporal structure of the sound output was driven by the video data from the installation.

6 The final level of sound was a baby giggle. This sound was based on a spectral analysis file of a female baby's giggle. The way in which the spectra were used in the resynthesis was changed in accordance with the activity in the installation space. A set of delay lines were also placed on the output of this sound. Activity within the installation set the echo rate, delay feedback, and delay/dry signal mix, causing variations over the duration of the synthesised sound. An envelope for these delay characteristics, and the time over which the delay envelope was executed were defined in the Gestation Max patch, and varied in accordance with activity in the installation.

The top-level interface for defining the level four sound, *BPM Multiwave Drone* is shown in Figure 4.6E.

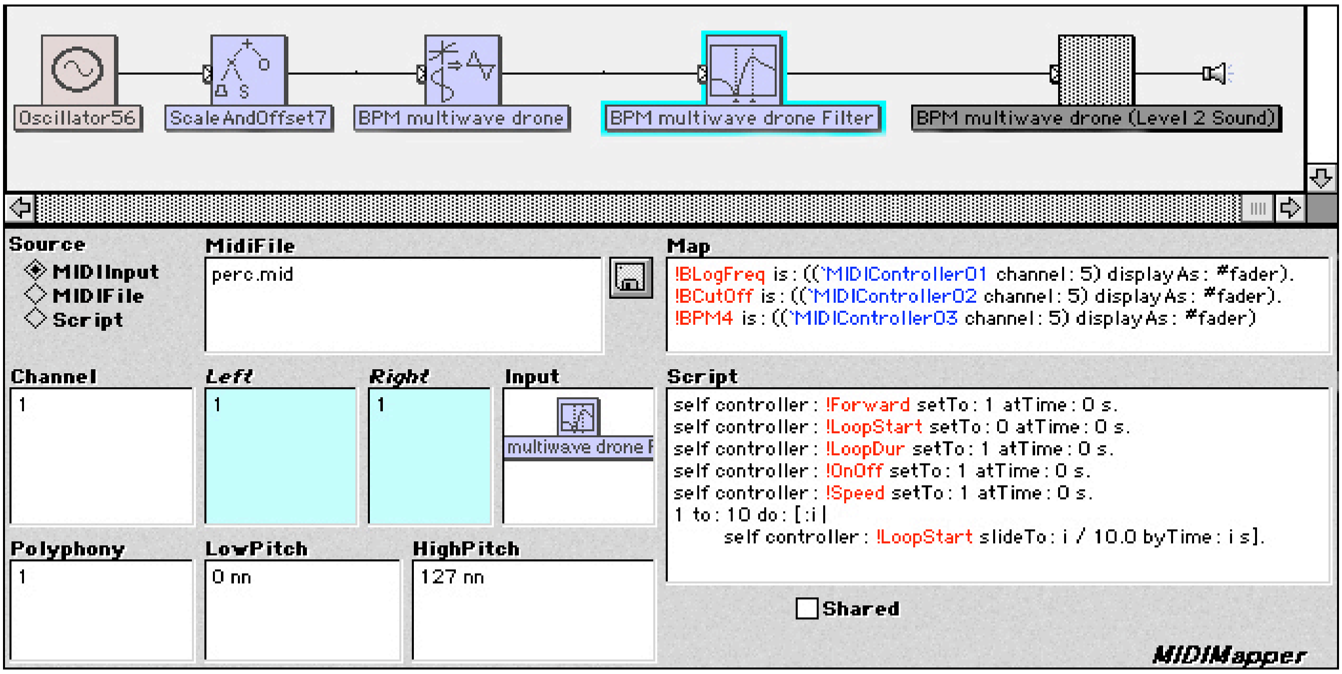



Figure 4.6E The level four sound for Gestation: BPM Multiwave Drone.

The Kyma window in Figure 4.6E shows the overall algorithm flow diagram in the upper part of the window and the detailed variables for the currently selected object (greyed when selected) in the lower part of the window, in this case the MIDIMapper object, mapping the incoming MIDI Continuous Controllers to the variables within the synthesis algorithm.

Each of the sounds was designed in this way. Once the algorithms were defined, they were placed into the timeline (Figure 4.6D) where the dynamic spatialisation and algorithm control could take place.

Scripts could be added to the sounds to change defined characteristic during run time. In addition to the dynamic orchestration approach taken in previous works, I defined differing ranges for the scaling of sensed data onto sound synthesis algorithms based on the overall dynamic of activity within the installation. This approach provided micro-scale adjustments of the timbre states within certain available activity ranges. It also meant that an algorithm could be written that had a much broader scope for variations in output; an algorithm that was much more dynamic and mercurial. Because the variables being fed to the algorithm were being scoped in the Max patch on the basis of current activity levels within the installation, a lot more variation could be achieved in the aural output from a similar number of synthesis algorithms used in the previous installation.

I took the dynamic scaling approach one step further. The output from all video sensing systems varies in scale according to the amount of light in the installation space. I wanted to set the light levels in the sound gallery very low, to create a womb like feeling. (I actually used twelve blue light globes in the ceiling of the gallery to recreate the effect one sees in the video). I also wanted to achieve as much resolution from the range of video output as possible. I achieved this by building a dynamic scaling function into the top level Max patch.

The dynamic scaling function measures the minimum and maximum activity within the video sensing output stream, and scales the video data so that a consistent range of values is sent to the Kyma patch, regardless of the light level in the exhibition space. This function means that the installation's response range is constant even if the number of people or environmental conditions change. This is important, because in MAP1 and MAP2, I had set an overall range which had taken into account the activity generated by one quiet individual at the minimum end, through to the activity levels generated by a large group of people. This setting meant that only a small portion of the overall resolution of the installation was used most of the time; the range associated with the activity of one or two individuals. Establishing a dynamic scaling function meant the installation could scale differently when being used by one person compared with the scaling for ten people within the space. This in turn meant that both groups of people would get similar outcomes for similar levels of individual activity.

The dynamic scaling sub-patch contained a reset function, which reset the dynamic scaling maximum activity figure after there was no activity within the space for two minutes (adjustable). This function ensured that if a group of people had been in the installation, and some minutes later, an individual engaged with the installation, the installation would not retain a memory of the scaling required for the group, but would adjust to the individual. This function provides a consistency of experience over a wide range of different conditions of use. A further advantage of this approach is that the installation self adjusts when installed in different gallery settings.

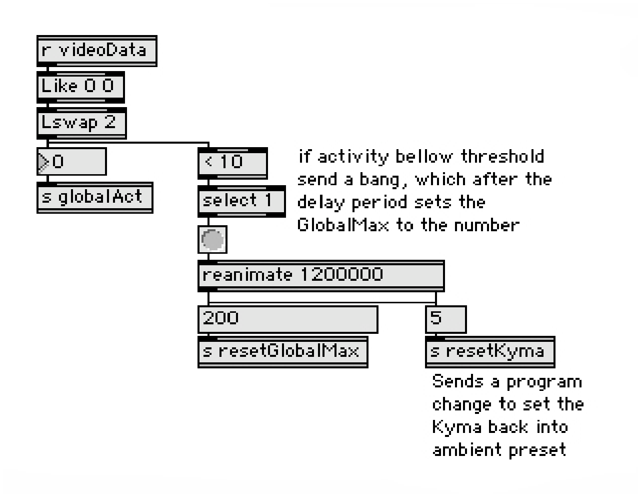



Figure 4.6F Reset scaling and environment qualities.

### 4.6.4 Summary

Gestation clearly states my humanist perspective with regard to conceptualising New Media art within the human domain. It was not intentionally dealing with the issues of sexual politics or bioengineering.

The research and development work carried out during this PhD culminated in Gestation. It exhibits:

1. dynamic orchestration techniques,
2. dynamic realtime scaling of video data to achieve a high-resolution of nuance in mapping movement and behaviour patterns onto audio and visual outputs,
3. spatialisation of the sound driven by activity within the space,
4. multi-variable, independent synthesis schemas (there were more than 30 synthesis variables per zone in Gestation),
5. an animated visual environment driven by the video sensing data,
6. multiple independent, sprited animation elements within a single animated environment, and
7. independent multi-user interaction.

Gestation drew excellent responses for its dynamic sense of interaction. People did feel a fluid, visceral and immediate sense of empathy with the installation.

Gestation presented both streamed and triggered environments simultaneously. The sound environment engaged realtime streamed interaction techniques whilst the visual environment undertook to combine both approaches. It presented triggered events (birth) that continued to evolve over their life span in accordance with the conditions under which it was created.

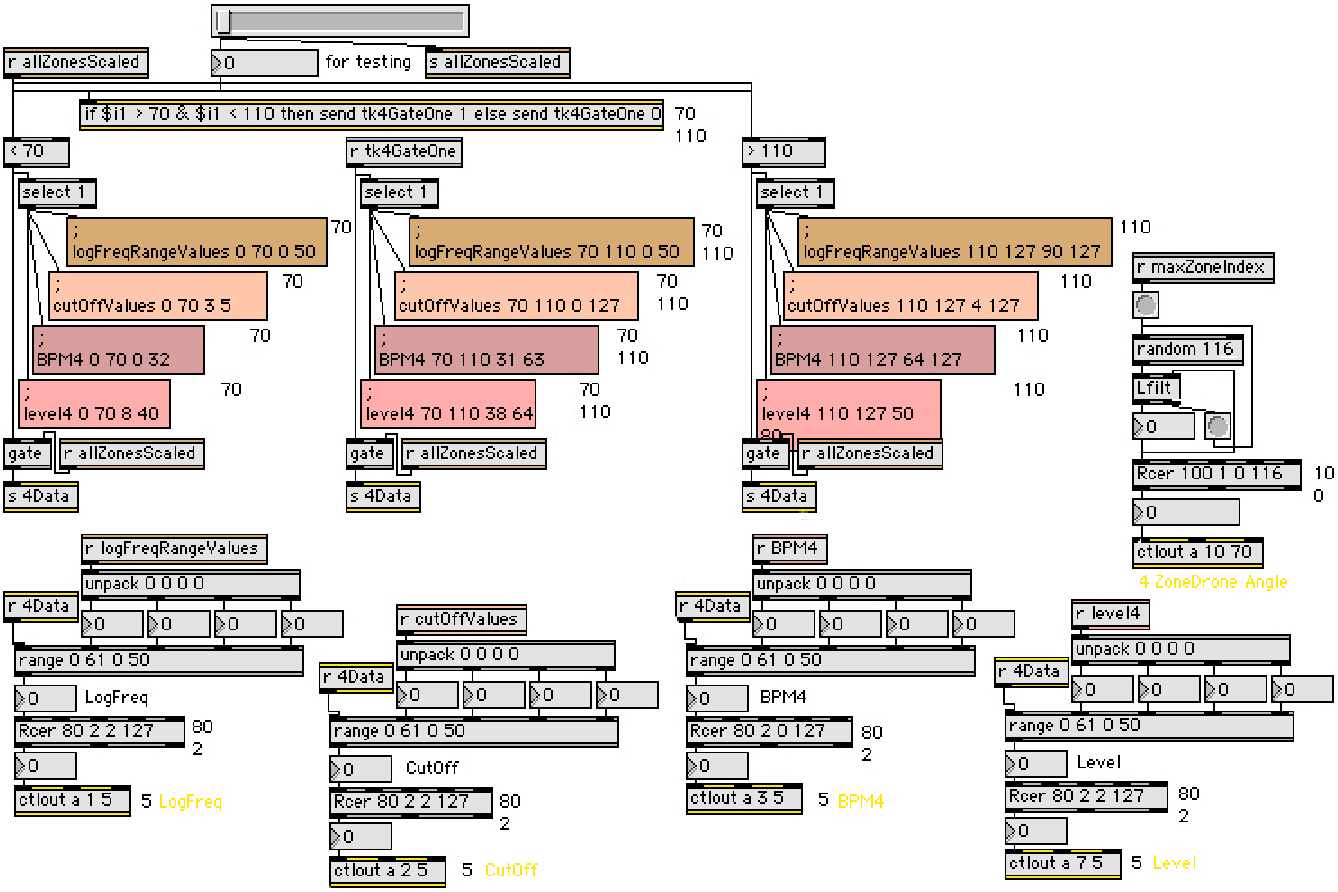



Figure 4.6G  The level four sound control patch in Max.

# 5 Interaction

Digital technologies offer an opportunity to move beyond the temporal structures associated with acoustic musics. The movement from analogue to digital technologies heralded an improvement of sensitivity in interactive systems.

> *In analogue systems, the synthesiser could, in a crude sense, 'listen' to a performer's gestures by means of triggers (pulses emitted when a performer touches a key on a keyboard, for example) and envelope followers (devices that trace the amplitude envelope of an event). Digital technology, with its increased memory and more flexible processing capabilities, provides more sophisticated recognition of external gestures. Computers can create elaborate and context-sensitive responses to the gestures they detect.* (Roads, C.:1996:626)

As Curtis Roads comments, digital technologies provide a platform for complexity. They allow us to run many envelopes, filters and other transformation processors on a single sound source, creating a rich and varied timbrel structure. The flexibility in structural complexity made available by digital technologies has been an important development in realtime sensing. Digital interactive instruments have seen a separation of the sound source from the act of excitation. This additional stage provides an opportunity for the abstraction and re-mapping of the excitation gesture.

> *The separation of the input device from the sound production leaves the opportunity to process and map the information coming from the input device in a variety of ways.*
>
> *Processing shapes the data, by inversion, compression and expansion, limiting, smoothing, quantising (thresholding). The data can be analysed for rates of change, delayed, convolved or distorted by linear or non-linear transformations.*
>
> *By means of mapping, one input device can control different parameters on a synthesiser at various times. A single message received from the input device, such as Key Velocity (how fast a key is pressed on a keyboard), can be mapped to spatial movement, loudness, or other parameters.* (Roads, C.:1996:625*)*

Software provides a relatively fast and extensive way to intervene in the data path, and reshape it in a dynamic fashion.

Once the artist has control of the intricacies of sensed input, the mapping of these characteristics becomes of paramount importance. Approaches to these mappings vary in accordance with their patterns of change. The principle categories of input type are:

- Trigger – a singular binary input (on/off)
- Stream – a continuous input that varies according to changes in its state (i.e. light, heat, etc.). This sensor type usually outputs a variable voltage, which is converted to a MIDI Continuous Controller.

The computer-based arts have been stymied by the complexity of the computer programming required to generate even simplistic and subsequently relatively unengaging images or sounds. Naturally occurring sounds are infinitely complex and variable; they contain intricate combinations of frequency partials[51], which vary over time in elaborate ways. These variations are determined by the many environmental factors present in any momentary event. The modelling of these multifaceted patterns requires very substantial computing power.

Fortunately, technologies like the Apple Macintosh G4 computer, or the Symbolic Sound Capybara 320 present the ability to generate audio in realtime. This relatively affordable computing power has allowed artists to move from *off-line* image and sound generation (that is, the generation of final output in a slower than realtime process) to realtime sound synthesis, animation and image composition. While this may not appear overly exciting, it marks a movement from the predominant approach to interactivity where trigger inputs control pre-made content, to a methodology where continuously evolving sensor data can be mapped to realtime streamed outputs. This change has precipitated a much more profound exploration of the genre of interactivity, and is one of the main foci of the research documented here.

Realtime computation has allowed the interactive arts to create realtime causal feedback loops. As the installation works discussed in this exegesis create sound in realtime, I have found it useful to draw on developments in twentieth century musical composition practice to clarify the movement from triggered events to evolving stream-like interrelationships.

The famous twentieth century music composer, Pierre Boulez, describes composition as a selection of notes derived from a finite predefined set (Nattiez:1993). Trevor Wishart (1996) points out that contemporary composition, especially within the genre of electronic music, goes well beyond *"a finite lattice and the related idea that permutational procedures are a valid way to proceed. . ."*. Wishart proposes a *"musical methodology developed for dealing with a continuum using the concept of transformation."* (Wishart, T.: 1996:7)

The concept of a stream of constantly evolving sound is directly supported by the use of realtime sound synthesis. The ever evolving, (sometimes audible; sometimes not), processes of data driven art, follow the Wishart approach. The Boulez approach is more closely aligned with the commercially prevalent paradigm of interactivity as a response to a defined challenge with a pre-specified finite outcome, such as the triggering of existent sound files in CD-ROM based interactives.

[51] Frequency partials are the individual frequencies that combine to generate a complex audio signal with a perceivable timbre.

Realtime synthesis provides a subtle but profound alternative. The use, within the synthesis instrument, of variables controlled directly by movement gestures (streamed sensor data) provides a way of not simply creating a personalised mix of existent sound samples (much like a DJ does with pre-recorded music) but of creating a completely unique sound stream. The user creates the temporal form as well as the pitch/timbre and '*orchestration*' of the score in realtime. My intention has been to use this freedom of response to create individualised outcomes that so tightly reflect the actions of the user as to be both qualitatively and quantitatively superior to an interactive experience utilising pre-made sound sample events.

An installation work using existent sound samples can be made to reflect user input by varying the playback polyphony, sample choice or small amounts of pitch bend and filter modulation. These relatively coarse reflections of control input do not reflect small intricacies of movement in as symbiotic a manner as the realtime timbre, envelope or modulation variation discussed in the project outlines.

Through my research into interactive, responsive environments, I have come to feel that the paradigm of music composition based on the temporal ordering of sound events is not appropriate within an interactive, responsive environment. The character of this composition process is both clearly understood, and limited.

For example, the note C3 on a piano defines a pitch and timbre that is only slightly variable, exists within a predetermined tuning system, and is probably contained in a sonata, fugue or other predefined temporal structure. It relies on an extensive historical tradition of composition, instrument building, performance, practice and development.

The interactive, responsive environments contained in this exegesis (except Reeds) propose the human body as an instrument. In this context, the instrument is largely an unknown; the movement of the body may be related to its use in the playing of a more traditional instrument. However, these instruments have a well-known and highly developed performance practice that is extremely specific about the way in which the human body is applied to the instrument.

Given sufficient time, a suitable and repeatable performance practice could be developed for a specific installation, the installation being documented and precisely repeated in each incarnation so as to provide a predictable instrument.

This situation is about as divergent as is possible from that of a public exhibition, the way my installation work is usually presented to the public.

The public exhibition allows the imposition of some controlling factors, which may encourage the visitor (the performer/the instrument) to 'perform' in a limited range of ways. This does not however constitute a semblance of performance practice; nor will the visitor be given a pre-structured composition to perform. The parallels therefore between a performance practice embedded in the tradition of chromatic, temporal composition and the experience of exploring, performing, or playing an interactive, responsive sound installation are difficult to find.

So, the questions become:

1. If the tradition of temporal, chromatic, composition is not applicable to interactive sound installations, how does the installation artist develop a framework for sound generation and structuring appropriate to the interactive, responsive installation?

2. How does the installation artist communicate the established sound infrastructure to the visitor in such a way that they are able to quickly come to terms with both the instrument interface, and the scope of variation inherent within the sound production capability of the installation?

These questions have been carefully considered throughout this research. The development of dynamic orchestration within realtime sound synthesis systems has been an attempt to address the first question. In developing this methodology, I have tried to make a sound synthesis structure that had potential for a vast number of different responses driven by realtime input. I have also tried to establish a structure that would serve the interaction, but exhibited as little of my own influence as possible. I have tried to remove myself from the final experience as much as possible; to become as transparent as possible so that the user experience is determined largely by their actions, not mine.

Krueger comments:

> *For the artist the environment augurs new relationships with his audience and his art. He operates at a metalevel. The interactant provides the direct performance of the experience. The environmental hardware is the instrument. The computer acts much as an orchestra conductor controlling the broad relationships while the artist provides the score to which both the performer and the conductor are bound. This relationship may be a familiar one for the musical composer, although even he is accustomed to being able to recognize one of his pieces, no matter who is interpreting it. But the artist's responsibilities here become even broader than those of the composer who typically defines a detailed sequence of events. He is composing a sequence of possibilities, many of which will not be realized for any given participant who fails to take the particular path along which they lie.*
>
> *Since the artist is not dedicated to the idea that his entire piece be experienced he can deal with contingencies. He can try different approaches, different ways of trying to elicit participation. He can take into account the differences among people. In the past, Art has often been a one-shot, hit-or-miss proposition. A painting could accept any attention paid it, but it could do little to maintain interest, one it had started to wane. In an environment the loss of attention can be sensed as a person walks away. The medium can try to regain attention and upon failure, try again. The piece has a second strike capability. In fact it can learn to improve its performance, responding not only to the immediate moment but also to the entire history of its experience.* (Krueger:1976:84)

The second question is more difficult to address. I have judged my success in this area from:

- the comments people have left in a Comment Book outside the exhibition,
- the feedback I have had from gallery staff concerning the public response to the exhibitions,
- my own observations of people's exploration of the installations.
- discussions with friends and acquaintances about their experience of the installations,
- my own experience of using the installation once it is installed in a gallery.

Each of the works documented in this exegesis illustrate developments based on comments and criticisms of the previous works. I have used the feedback from visitors to each exhibition as a guide to their level of engagement, and, with each work, progressed further towards a truly individualistic, enveloping and engrossing experience. (Selected visitor comments are contained in Appendix One).

A further consideration when using physical movement as the instrument interface is the differentiation between spatial and temporal reflections of the interactive input.

An interactive, responsive environment presents interesting challenges in terms of the representation of spatial input. Spatial distinctions are present in traditional instrument design. A piano, for instance, presents the performer with a spatial division of pitch; successively lower pitches to the left of centre, and successively higher pitches to the right of centre. The spatial layout of the keys associated with these pitches follows a pattern, refined over hundreds of years, that allows the performer to move very rapidly up and down the pitch sets, and to cover multiple pitches simultaneously. No such historical precedent exists within interactive, responsive environment installation practice.

A generalisation of all acoustic instruments is that the spatialisation of the interface is almost exclusively applied to pitch. There is little spatialisation associated with timbre, or with rhythm, with some subtle exceptions from the percussion family. It could be argued that the placement of the bow close to the bridge *(Sul Ponticello)* or up the fingerboard *(Flautato)* of a stringed instrument constitutes a spatial relationship to timbre. Indeed it does, but it does so within such powerful constraints that it is not really relevant to the discussion at hand.

The focus of the current discussion is human gesture as a controller within interactive, responsive environments. The nature of such gestures is both dynamic (in terms of its variability-large/small-fast/slow), unpredictable (unless choreographed) and complex (more than one limb can be engaged simultaneously in an asynchronous gesture).

As the sound stream generated by interaction within an interactive, responsive sound environment is almost exclusively improvised, the element of repeatability is not available as a general measure to equate the intent of the user to the quality of the sound output. This makes it extremely difficult for the artist to draw any clear guidelines as to the perceived success of the designed response of the installation.

Subjectivity is another variable that makes such a comparison difficult. Each visitor to the installation brings with them a different aesthetic, a different set of cultural conditionings. An examination of interactivity cannot therefore be based on aesthetic judgements of the visitors to the installation, but should be based on the scope of the installation to engage the visitor, and evolve in concert with their changing behaviour patterns.

The success of each installation has been gauged by the feedback about the degree to which people felt immersed and engaged by each piece, and the sense of individualised gestural response.

So, what does interactivity mean?

In my view, interaction is a widely abused term in computer-mediated art.

The Oxford English Dictionary (2001) describes interaction as follows:

> *inter-, pref. Between, among, mutually, reciprocally.*
> *interact v.i., act reciprocally or on each other*
> *interaction n., blend with each other*

The Collins English Dictionary (1992) contains the following definitions:

> *interact vb. to act on or in close relation with each other*
> *interaction n., 1. A mutual or reciprocal action or influence 2. Physics, The transfer of energy between elementary particles, between a particle and a field or between fields.*
> *interactive adj., 1. Allowing or relating to continuous two-way transfer of information between a user and the central point of a communication system, such as a computer or television. 2. (of two or more person, forces etc.) acting upon or in close relation with each other; interacting.*

The prefix of the word interaction, *inter*, is described as something between, among, mutual, reciprocal. This definition implies that the two parties act upon each other; that the parties exchange something, they act upon each other in a way that is reciprocal. The above Collins English Dictionary definition of interaction outlines an action that involves reciprocal influence. In the field of physics, it leads us to understand that an exchange of energy takes place.

How then do these definitions translate into the area of new media art? Does an exchange of energy occur when one is viewing a CD-ROM? An exchange of information certainly occurs, but an exchange of energy, probably not. The user requests a piece of information, and the computer, through the programming of the CD-ROM, delivers that information to a screen in such a way that the user can comprehend it.

One could argue that a transfer of energy takes place when playing a computer game that requires a racing car driving wheel, gear changer, brake and acceleration pedals to be used (many of these can be seen in amusement parks). In this case the user is directly transferring energy through the interface by turning the steering wheel, changing gears and possibly operating the accelerator and brake pedals. This energy is transferred to the interface. The variation in condition of the interface is transferred as data to the computer program. The computer then draws a scenario to the screen, to which the user responds. There is clearly a causal loop here. (A causal loop being a scenario in which all parties require the other for their survival, and where the interaction of all parties maintains a balanced system). However, does the racetrack alter because of the behaviour of the driver? Is there actually a reciprocal transfer taking place? No, the user is simply attempting to maintain a state that is acceptable to the criteria of the game, i.e. keep the car moving forward and on the track. The computer program defines the conditions to be met and these conditions do not change as a result of the users input. They are the same every time the game is played.

If interactivity is predicated on the ability of both parties to change in a way that reflects the developing relationship or discourse between them, then we have to accept that multimedia systems that do not evolve their behaviour in relation to accumulated patterns of input are not interactive; they are simply responsive.

In order for the system to represent an interaction, it must be capable of changing, of evolving. The process of evolution ensures continually unexpected outcomes. The outcomes are based upon the nature of a response-response relationship where the responses alter in a manner that reflects the cumulative experience of inter-relationship.

For this to be upheld, each exchange must be personal. It must reflect the unique qualities of each particular dialogue.

We experience this kind of interaction every day. A discussion between two individuals illustrates interactivity that is (Winkler:1998:3);

- unique and personal to those individuals,
- unique to that moment of interaction, varying in accordance with the unfolding of the interaction, but is
- maintained within a common understood paradigm (both parties speak the same language).

Within such an interaction the starting point is known by one of the parties engaged in the interaction, but the journey they undertake during their discourse, and the point at which they end, are unknown. This process of interaction is extremely dynamic with each of the parties constantly monitoring the responses of the other and using their interpretation of the other party's input to make alterations to their own response strategy, picking up points of personal interest, expanding points of common interest, and negating points of contention.

If this kind of interaction were applied to the design of interactive instruments, they would need to adopt a structure that recreates the system dynamically in response to interactive input.

The very structure of the system would become dynamic, and would change over time as a result of the interaction.

## *5.1 System Design*

Interactive music research has largely focused on the study and creation of systems that perform tasks appropriate to the creation of chromatic music, where pitch is paramount, harmony is integral, and the beat is a persistent basis for rhythmic definition.

The imposition of musical characteristics on interactive music systems only makes sense when designing a system as an accompaniment or improvisatory application where the other elements of the ensemble are adhering to chromatic music performance practice.

If one accepts that the aim is to create a new instrument, to generate new or innovative music, then the acceptance of criteria associated with chromatic music (i.e. pitch, pitch structures that define tonalities, the recognition of chords or other pitch structures, the recognition of tempi and time signatures etc.) are inappropriate. The field of philosophy defines this kind of criteria confusion as a '*Category Mistake'* (Ryle:1995).

Robert Rowe, in **Interactive Music Systems** (1994) discusses the interpretation of low-level musical signals into structured high-level representations. He comments that interactive systems:

> *… interpret the input by evaluating human musical understanding (Rowe:1993:3) … In their interpretation of musical input, interactive systems implement some collection of concepts, often related to the structure musicians commonly assume. Each interactive system also includes methods for constructing responses, to be generated when particular input constructs are found. As methods of interpretation approach the successful representation of human musical concepts, and as response algorithms move towards an emulation of human performance practices, programs come increasingly close to making sense of and accomplishing an instruction such as* ***'broaden the end of the phase'****.* *(ibid:4). (Rowe's emphasis)*

Rowe outlines three stages of an interactive music system;

1. sensing; which includes pitch and rhythmic pattern detectors,
2. processing; which includes the scheduling of tasks that create musical events in response to the sensed inputs,
3. response; where the constructed audible response is delivered back to the interactive agent.

In a situation where the system is designed to accompany or improvise with a musician, it makes sense to construct responses within an agreed musical aesthetic. However, when the interactive system is exhibited in a public space, as an interactive artwork, the system creator(s) cannot expect those engaging with the system to have any knowledge of musical paradigms. Furthermore, when the input to the interactive system is a human gesture, it is questionable whether a musically constructed response is appropriate.

The appropriateness of response is a central issue in the design of interactive systems. The mapping of sensed input data to processing algorithms is the most complex and subjective aspect of system design.

Why, for instance, would one wish to convert the movement of an arm to a musical chord, which in turn embeds itself within a chromatic structure? What role or relevance does a sense of tonality have to the experience of interaction within such an environment? Equally, one has to contemplate the role of rhythm, melody, and the other equally ingrained aspects of musical composition, musical forms and structure.

When presenting an interactive installation within a public space, the artist cannot assume an existing level of musical skill or interactive aptitude.

In order to address these issues, the interactive input must closely represent natural activities undertaken using the body. Video sensing allows continuous tracking of human movement. It provides a continuous stream of data that represents a qualitative indication of the movement taking place. The qualities of the movement can be defined as follows:

- Direction of movement
- Speed of movement
- Size of a moving object
- Proximity of movement to other moving objects
- Inertia
- Consistency of movement

Whilst direction and speed of movement may be applied to existing musical parameters, they make much more sense when equated to the characteristics of naturally occurring sounds.

## *5.2 Dynamic Morphology*

It is for these reasons that I have chosen to explore dynamic morphology (Wishart:1996) as the principal concept of both sound generation and sensing systems. Dynamic morphology suggests a continuous generation of sound events (audible or silent). The contemplation of existing musical paradigms suggests the creation and collection of individual events. Events equate to single points of action and only become part of a continuum of movement when contextualised within known musical forms.

Trevor Wishart (1996) discusses the difference between fixed timbrel objects (most acoustic instruments) and sound objects with what he calls a Dynamic Morphology.

A dynamic morphology relates to a sound object represented in a "*a three-dimensional pitch-duration-timbre space*" (Wishart:1996:94). This describes a sound altering in a continuous fashion through the characteristics of pitch, timbre and time simultaneously and continuously. A naturally occurring sound has just such a morphology, fashioned by the many facets of its initial sounding source, and the architectural and acoustic space it inhabits. Rodet, Potard and Barrière (1984), point out that our recognition of certain consonant sounds in speech is heavily dependant on the nature of this morphology, so it is vital to our understanding of the world, and our everyday experience of it.

*In general, sound objects with a dynamic morphology can only be comprehended in their totality and the qualities of the process of change will predominate in our perception over the nature of individual properties.* (Wishart:1996:94)

Wishart defines two primary areas of Morphology

1. *Gesture*, which is "*the articulation of the continuum by the agent which instigates the event*"
2. The classification of Morphology used "*in relation to perceived natural phenomena"* (ibid:102*).*

The morphology is only apparent if the sound-object is perceivable as a whole. Wishart comments:

> *we may expect, however, that the category of gesture is in some ways more restricted than that of natural phenomena structures … It is in other respects more extensive than the category of natural phenomena - higher organisms are capable of very subtle articulations of the continuum, which we should only expect to find by chance in the structures of a natural phenomena. The interface of these two types of description may be seen in the relationship between vocal and instrumental music. A musical instrument is a device used to stabilise, through its resonance structures, the pitch and timbrel dimensions of a sound-event. The morphological structure of the sound-event is thus dominated by the characteristics of the natural phenomena of resonance.* (ibid:102)

A human movement is characterised by a smooth continuum. Each movement being made up of many infinitesimally small variations and adjustments. The overall movement is called a gesture, and the nature of the way in which that gesture can be enacted is described as the weight of the gesture. It is not possible to extricate from the gesture individual moments of movement and adjustment that make up the overall gesture. In this sense, human movement reflects Wishart's (1996) statement that "*sound objects with a dynamic morphology can only be comprehended in their totality*".

Computer music is also characterised by a continuance of variation and adjustment. Similarly, a single sound is perceived as a whole, not as a collection of the myriad samples that create it in a digital playback system. Whilst it is possible with a computer to extract a single sound sample it is meaningless to the human ear.

When considering interactive installations that use sound as their principle medium, dynamic morphology needs to be directed at the sound production process and the relationship between the person(s) interacting, and the system response(s). For the installation to respond dynamically, there must be dynamic variation in timbre, pitch, amplitude and associated envelopes.

This discussion has important ramifications for the design of responsive and interactive instruments. It also provides a profound insight when addressing the design of software based performance instruments, prevalent within the 'laptop brigade' of electronic music performance.

For the gesture of a performer to be fully inscribed within a realtime sound output, the sound must fulfil the Wishart concept of dynamic morphology. The sound must evolve and change in such a manner as to correlate with the qualitative development of the gesture, an evolution of momentary events, which is unknown to the system at any point prior to their execution. The morphology of a particular gesture is unknown even at the beginning of the movement. The player may change the direction or speed of the movement at any time, and may alter the position of the limb in both the vertical and horizontal planes at a rate of change that matches no previous event. This will occur even when the player is attempting to perform the gesture in an identical way to a previous movement event. It is only the highly trained dancer, with a spatial awareness developed over years of exacting training who can reproduce spatial positioning, rate of change and horizontal and vertical gesture within sufficient bounds for us to perceive them as repeatable. To the video analysis system, even these highly trained executions alter in subtle ways.

The model for dynamic morphology is embedded within the concepts of Object Oriented (OO) programming. The building block of all OO software design/applications is the *Class* that has a range of functionality. In sound synthesis terms it may look for a range of variables and perform some operation on an audio input stream, or may create an audio stream which is fed into a software based mixer within the program (patch/instrument).

The *Class* will contain some parameters in its description; these parameters must be provided by the process that creates the instance of the class, (each instance of a class is called an *Object*), in order for the class to be become extant. For instance the parameters may be the frequency of the oscillator, or the variable that will be mapped to frequency, or the mixer/bus input into which the generated audio will be fed. The functionality of the class is the domain of the code written within it. A *Class* is essentially a blueprint for an *Object*. Once extant it becomes an object that performs a function(s), as described in the class design, until it is no longer required. At this stage the *object* may be disposed of, and a garbage collection mechanism removes the *object* from the computers memory, thereby providing memory allocation for new *object* creation.

The beauty of the OO approach is that units of functionality can be dynamically created, plugged together, detached and disposed of in such a way as to fulfil the momentary requirements of the global program.

An analogy may be an orchestra whereby instruments are created when required and disposed of when no longer required. To some extent that is the case with a symphony orchestra. A composer will orchestrate a work to use particular instruments to create desired timbrel qualities at points throughout the work. However, the capacity of the orchestra is essentially fixed. If four flutes are required, then the flutes are seated in the orchestra at the beginning of the work and remain until the end. What would happen if a work were being orchestrated in accordance with a responsive/interactive schema, predicated on dynamic morphology? Such a situation may require 20 flutes at one point, and no violins! A software infrastructure can allow for such occurrences within the limits of the processing power of the host CPU, and the limitations of the software design.

In this approach a dynamic morphology would extend beyond the scope of filtering, or otherwise variable synthesised output from one or even a collection of algorithms, (which no matter how the algorithm is designed will have a finite range of aesthetic and timbrel variation), to a dynamically forming orchestration. In such a dynamic orchestration, a new sound object would be created when the morphological scope of the current algorithm reaches its limits. The new algorithm may exist only as long as it is required, and may be augmented by other dynamically created instruments, before being disposed of until the interactive input requires it again. Furthermore, if this instrument were a flute, the type of flute is defined at the point of creation; it is defined by the process that creates it. Hence, the object could be a piccolo, alto flute, bass flute, wooden flute etc as the moment requires and in accordance with the scope of the parameter variables of the class.

Most musical programming languages, although they may implement object-oriented approaches to programming, do not allow for the dynamic creation and disposal of synthesis processes. The synthesis process must be connected to the base synthesiser (the foundation synthesis engine and audio output structure) at the time the program is instantiated. It is not possible even to add a band to a filter in realtime, one must create an entirely new filter object, dispose of the previous object, and plug the new filter into the base synthesiser, which creates a disruption in the audio signal. Lexically separate synthesis processes are required for true dynamic morphology of synthesis processes.

So, while some individual instruments can be created and disposed of dynamically, the mixing infrastructure (audio generation, routing, filters etc) must be created at the point of instantiation of the program, and so all the algorithms run at once, from the time the piece is turned on. They may be audible or not as the input demands, but the processes are all constantly running. Clearly, this is not an efficient use of resources, and is not a good way to encourage dynamic morphology beyond the initial capabilities of the instruments.

Drawing on the previous orchestral analogy, the above situation would equate to the orchestra playing a wind quintet, but requiring all the other members of the orchestra to remain on stage, and worse, to require the entire orchestra to attend all the quintet rehearsals and sit silently in their chairs during the rehearsals and performances.

My responsive interactive environment installation MAP2 explores the object oriented programming approach. The music programming language SuperCollider was used to create an eight-channel instrument.

MAP2 is a three dimensional space which can be entered and encountered, played and played with. It is a virtual musical instrument using the movement of those within it to compose music in realtime. There is no pre-recorded sound material in MAP2. All sound is generated in realtime in response to the speed, direction, position and quality of movement of those within the installation.

People enter MAP2 to compose music and sound by using their bodies to solicit responses from the custom developed computer software, generating a rich, enveloping and continually evolving sonic environment in real time.

MAP2 uses video sensing to map the movement and behaviour patterns of people, within the installation, into sound. The synthesis approach uses a number of algorithms collected into an orchestra. The instrumentation is augmented each time the sensed activity increases above a certain threshold. The horizontal space is broken into quarters, each independently sensed, allowing four 'players' to use the installation simultaneously and asynchronously. Multiple synthesis algorithms are available within each of the four horizontal zones; each zone being an independent instrument. The output of those algorithms is controlled by the position and dynamic threshold of gesture within the instruments zone. Each of the four zones has slight variations in the synthesis algorithms providing a diversity of timbre. The variation is achieved using filter bands that change their makeup based on the sensed activity.

Each zone is allocated a pair of loudspeakers. Eight loudspeakers surround the installation. The **zone >> loudspeaker** relationship means the sound follows the individual through the installation space.

In MAP2, the video sensing was developed to sense the space in 3D, using 2 synchronised video cameras. This approach allowed height information, and vertical position to be applied to an additional sound algorithm.

The approach taken in MAP2 is conditioned by a limitation in processing power of the host CPU, (a Macintosh G4). It is further limited by the programming approach of musical OO packages. Although SuperCollider cannot create lexically separate synthesis processes out of the box, it is possible to dynamically create multiple synthesis processes inside a Tspawn node (Tspawn is a subclass of the class Spawn.) Instances of Spawn spawn new events at timed intervals and mix them to output channels. When it is time for a new event, the newEventFunc function is called with the Spawn instance as an argument. The function should return a graph of unit generators. This graph will be installed into a new Synth instance and evaluated until it reaches the end. Therefore it is important that the sound does have an end. The definition of the Tspawn class is as follows:

**TSpawn.ar(newEventFunc, numChannels, maxRepeats, trig, mul, add)**[52].

The predominant paradigm of music composition is assumed. The composer, programmer must create the resources they expect to need for the entire composition. These resources are created at the beginning of the work. They contain a set group of instruments, with an inherently limited morphological scope. I believe this is a limitation that has no place in interactive electronic music performance.

---

52 SuperCollider is a musical programming language developed by James McCartney, see. http://www.audiosynth.com

## *5.3 Summary*

I have pursued interaction that:

- does not include any pre-defined pathways,
- takes dynamic morphology as its foundation,
- requires dynamic software, built on the object oriented approach to dynamic instrument instantiation and orchestration.

The objective has been to move away from the concept of interaction as defined in CD-ROM and Virtual reality systems where the user experiences a mediated journey along predefined paths, during which they are presented with predefined audio/visual stimulation. My objective has been to develop an experience that is not pre-determined, but instead reflects each individual's nuance of input in a unique and fulfilling manner.

In working towards this point I have come to realise that unless the system changes in response to accumulated user input in the ways discussed above, it is essentially responsive, and not interactive. I have therefore used the term, *interactive responsive environment* to describe my installations, which I believe herald a journey towards true interactivity.

Although the above discussions and observations relate to sound and musical composition, it is clear, given appropriate computing power, that the principles of dynamic morphology and object oriented programming can be applied to realtime creation of image and animation outcomes that reflect high levels user interaction nuance.

# 6 Conclusion

This exegesis has outlined a research journey from:

1. triggered data input that could not avoid interactive 'blind spots' within the installation environment, to streamed data input, that monitors the entire exhibition space at all times,
2. direct, often linear mappings of input data onto sound and vision algorithms, to dynamic, high order mappings including simultaneous, multi-controller, synthesis processes occurring in autonomous interactive zones, and
3. fixed sound algorithms, which in MQM and GITM utilised pre-made content, to dynamic orchestration of image and sound techniques utilising realtime synthesis applications in MAP1, MAP2, Reeds and Gestation.

All the pieces documented in this exegesis are primarily engaged with the mapping of sensing data onto sound and graphics algorithms to create an environment in which the interactor feels directly engaged, and where part of that engagement is the feeling that the environment reflects the quality of their presence and input.

The research carried out during this PhD brought me to understand that there are three overriding imperatives when developing mappings for responsive, interactive environments. They are as follows:

- Mapping for immersion
- Dynamic orchestration
- Approaches to abstracting the technology from the experience.

## *6.1 Mapping for Immersion*

Immersion involves creating a three-dimensional interactive, responsive environment that envelops the exhibition visitor in such a way that they feel engaged and captivated by the exhibition.

This experience is defined as distinctively different from the experience of listening to the sound and music from loudspeakers as a detached observer.

Simon Emmerson (1994) alludes to the importance of a perceivable relationship between performance and the outcome of a performer's gestures. For instance, a mouse connected to a laptop computer can be used to create massive and very fast changes in an electronic music performance. The size of the movement belies the outcome. Emmerson (1996) discusses these issues in relation to the perceived location, or rather 'dislocation' of the sounding source in acousmatic music and much electroacoustic music using multi-channel diffusion techniques.

These issues have a direct relevance to the mapping of human movement to sound and vision creation in an interactive, responsive environment installation. The weight of the gesture must be translated into a sound quality that communicates back to the mover something about the installations ‘perception’ of the movement. If the movement is large then the sound must change in scale with the movement. If the gesture is small then the timbre or texture of the sound must respond in concert with the gesture. In this way, it becomes apparent that the response of the installation is a direct result of not only the movement, but also the quality of the movement, and therefore a reflection of the intent of the person initiating the gesture.

Furthermore, the spatialisation of the sound within the installation must be considered from two perspectives:

1. The creation of a sound field that creates a sense of immersion. Such a sense of immersion is generated when the sound field seems continuous, and when one is not aware of particular loudspeakers as the point source of the sound.

2. Individual elements of the sound field that are directly responding to the current movement should be diffused in such a way that they appear to have a spatial relationship with the source of the gesture in the installation. The sound source should be positioned as if the user has created a disturbance in the sound field. This requires a dynamic spatialisation system, placing each layer of the sound score in relation to the position of sensed movement.

Emmerson outlines the *“Three great ‘acousmatic dislocations’ established in the half century to 1910. These are: (1) Time (recording), (2) Space (telecommunications, telephone, radio, recording) and (3) Mechanical causality (electronic synthesis, telecommunications, recording)”.* (Emmerson, S.:1994:98)

These three categories can be adopted for the generation of sound within an interactive, responsive environment. They can be described as follows:

- **Time**: The speed of response of the environment to the gestural input must be such that there be sufficient immediacy for the user to perceive a direct relationship with their movement. In this usage, there is an attempt to prevent the ‘dislocation’ of time being argued by Emmerson.

- **Space**: Space, as discussed above, must be considered both in terms of the diffusion of sound, and the way in which the architectural space contributes to the construction of the interactive, responsive environment experience.

- **Mechanical causality**: The relationship between the causality and the response within an interactive, responsive installation is of utmost importance. The mapping of response patterns must generate a relationship that is immediately perceivable. The mapping must communicate something about a qualitative relationship, drawing a parallel between the quality of gesture and the nature of the change in response.

Emmerson goes on to say *"The aim is to be clear that in abandoning any reference to these 'links of causality' the composer of electroacoustic music – especially that involving live resources – creates a confusion (even a contradiction) and loses an essential tool for the perspective and engagement between the forces at work."* (Emmerson, S.:1996:1)

This same clarity in intent is vital within an interactive, responsive environment. The mappings between movement and outcome may be many and varied, they may change in relation to the number of people in the space (GITM), or the current dynamic of movement within the exhibition (MAP1, MAP2, Gestation), and so they may offer an ever deepening relationship as the user understands the finer characteristics of the interface, but they must always be clear and immediately perceivable.

## 6.2 Dynamic Orchestration

As discussed in 6.2, the orchestration of each of these installation works was designed to be dynamic, to vary in accordance with variations in the dynamic of movement in the exhibition space. This approach was pursued in order to change the weight of the sound texture in such a way as to reflect the weight of the sensed gesture.

Creating a perceivable link between the weight of gesture and the density of texture, gave a visceral, tactile quality to the interactive experience. This relationship draws on traditional instrument design, where a more intense engagement with an instrument generates a change in timbre that reflects a more complex overtone structure. In general, acoustic instruments also illustrate a relationship between energy input and amplitude of output.

This consideration led to the mapping of movement to intensity outlined in MQM and GITM.

Amplitude is supported in these two works by changes in:

- **The intensity of sound**. For instance, the sounds in MQM and GITM range from meditative drones and slow rhythmic patterns through to loud distorted tones. This change in the character of the sound is what is meant by intensity.

- **The period of sound events**. The sounds in MQM and GITM reduced in length in direct relationship to their intensity. Gentle sounds are longer (08 - 20 seconds); intense sounds are shorter (3 - 8 seconds). The variation in sound file duration brings about a slowing or speeding up of the rate of sound events. This change in the speed of sound events creates a variation in density (meditative/ energetic) and the subsequent increased complexity of orchestration.

MAP1 freed the sound/gesture relationship allowing exhibition visitors to create changes in the sound environment by explicitly varying individual parameters.

The floor space of MAP1 was divided into four large sliders, each controlling a different synthesis parameter (described in detail in Chapter 4.3). This synthesis variable would move to the position of last activity within its zone. The rate at which the parameter moved was determined by the sensed activity in the target field (dynamic activity = fast, meditative activity = slow, and all gradations in between). In this way, MAP1 was played like an instrument.

Other variables were determined by the sensing field (64 fields in four rows of 16) with the highest activity. These variables include a range of variation (jitter) in some variables, i.e. grain length and pitch, which created an increasing rate of change in the focus of the primary setting in direct relationship with the increased intensity of sensed dynamic of movement. The centre of the variable range was, however, set as described above.

Beyond the rate of change of synthesis parameters, MAP1 separated the synthesis parameters from a direct relationship with the dynamic of movement and gesture in favour of precise control of parameters based on spatial plotting within active zones. This approach was continued in the latter installation works.

## *6.3 Abstracting the Technology from the Experience*

In order to achieve the objectives of immersive engagement, the interactive, responsive installations documented here were designed to present interfaces that utilise natural human activity. Considerable effort was put into hiding the technology of the interface in order to encourage the exhibition visitor to engage with the content of the installation and not with its execution. For instance, in both MQM and GITM, the floor trigger pads were hidden underneath dance tarquet. The trigger pads were then outlined with a white line indicating their position and presence under the tarquet. This approach provided a clear indication of the points of interaction, whilst disguising the pragmatics of the technology.

The abstraction of the experience from the technology is vital if the user is to be immersed in their engagement with the environment, and not in a process of inquiry as to how the installation operates.

This abstraction of the experience from the technology was amplified with the use of video sensing as the primary sensing technique. The video camera was hidden in the roof of the installation (MAP1, MAP2, Gestation), remaining completely unnoticed by most exhibition visitors. The gallery visitors were entirely unencumbered by technology.

In developing these works I have contemplated the artistic value of interactive new media art works, and asked myself the following questions:

- What do we take away from them?
- How do they enrich our understanding of the world?

- Do we continue to think about the experience afterwards, thereby developing a deeper appreciation of the ways in which that experience reflects upon our own lives, as one does long after viewing a good film?

I don't suggest that my own work has these outcomes, although, like many other new media artists, I strive to create works that will facilitate these outcomes.

New media art is nowhere near its zenith. There is much work to be done in developing a language that communicates clearly and is sufficiently varied to accommodate the many individual artists working in the medium.

Is the experience of these works simply one of mapping the development of the art form, and in turn the evolution of the technologies or an unrestrained expression of artistic intent?

I think current new media art provides both, however there are still many works that communicate little more than a technical achievement.

If new media art is to be taken seriously as an art form with the capacity to communicate something about human experience, the metaphysical and the spiritual, then we need to lose the technology, the technology that makes the work possible, the hours/weeks/months of programming, the innovative technical development. These aspects of the work, which are often revered as great achievements, need to be translucent, conspicuous by their absence.

The visitor to an interactive, responsive environment should be unaware and unconcerned with the technology creating the experience. They should, however, experience a symbiotic relationship with the work that permits a real sense of freedom of interaction, and an infinite scope for self-expression and exploration.

This is my goal, a goal I hope you see illustrated by the development of the projects in Chapter Four.

Although technology has developed in leaps and bounds in the last decade, affordable computing power is just now sufficient for a level of realtime interactivity. The current state of technology is encouraging for the development of this kind of work, and encourages realtime data driven sound synthesis through fast computing and excellent software tools.

We must shift our focus from technical achievements to a user driven experience. The technology must become both infinitely variable, and invisible to the end user.

The development of virtual reality technologies has shown a distinct partiality to the visual. In my view, sound is a much more direct and affective stimuli. If we can make sound more responsive to individual interaction intricacies, I am sure responsive, interactive sound environments can be shown to be a superior form of immersive experience, whether they include a visual aspect or not.

In summary, I feel the most rewarding outcomes in responsive, interactive environments continue to be achieved through the exploration of realtime sound and/or vision generation. Realtime synthesis processes can reflect the small intricacies of individual interaction, allowing the participant to feel directly acknowledged by an unequivocal reflection of individuality, subsequently encouraging a deep level of commitment to the exploration of the installation's potential.

## *6.4 Future Work*

During 2000, while I was the Australia Council for the Arts, New Media Arts fellow at RMIT University, I began some research that is beyond the scope of this exegesis, but expands upon the work it contains.

Through this research, I have formed the opinion that in order for interactive system responses to be substantially more complex and sophisticated, a level of artificial intelligence must be introduced between the sensing stage and the mapping of the sensed data to the synthesis parameters. A relatively linear mapping of input data to a limited and fixed number of synthesis parameters does not support the evolution of system response as discussed in Chapter Five on interactivity. In order to explore these possibilities, I undertook some research with Dr Dinesh Kant, Senior Lecturer in Biomedical Engineering Research at the School of Electrical and Computer Systems Engineering, RMIT University, Melbourne, Australia.

We set out to explore the possibility that Neural Network computing techniques could be used to recognise gesture patterns so that;

- we could categorise the gesture using subjective human criteria (i.e. violent, loving etc),
- we could identify the individual making the gesture (i.e. distinguish between individuals movement characteristics),
- we could track multiple individuals with a single camera view,
- we could develop evolving mapping strategies that would change, based on subjective analysis by the neural network system (i.e. the characteristics of the current movement gestures).

A number of experiments were carried out which illustrated that it was likely that a Neural Network could be trained to differentiate between two simple activities (drinking a glass of water, and writing a page of notes). There was also some positive indication that with further development, it would be possible to differentiate between the individual subjects.

All of this work was done off-line. A great deal of further research and development work is required before the system produces reliable results, or most importantly before the system can run in realtime. This research continues beyond this PhD.

# 8 Bibliography


Art & Design (1994) *Art and Technology*. London. Art & Design Magazine.

Art & Design, (1993) *Parallel Structures, Art, Dance, Music*. London. Art & Design Magazine.

Bianchini, R & Cipriani, A. (2000) *Virtual Sound*, Rome. Contempo

Boulanger, R.(ed) (2000) *The Csound Book*. Massachusetts. The MIT Press

Bandt, R (2001) *Sound Sculpture,* Sydney. Craftsman House

Brown. R (2001) *Biotica, Art, Emergence and Artificial Life*. London, RCACRD Research Publications.

Buxton, B. (1996) Alias Wavefront, inc. *A Directory of Sources for Input Technologies* Available email: buxton@aw.sgi.com www.dgp.toronto.edu/people/BillBuxton/ImputSources. viewed 28/02/98

Buxton, B.(1996*) Selected Papers*. http://www.dgp.toronto.edu/people/BillBuxton/billbuxton viewed 28/02/98

Capra, F. (1996) *The Web of Life*. London. Harper Collins, London.

Computer Music Journal Vol.22 No1 (1998) The MIT Press. Cambridge, Massachusetts

Collins English Dictionary (1992) 3rd Ed. HarperCollins Publishers, Sydney, Australia

Computer Music Journal Vol.22 No4 (1998) The MIT Press. Cambridge, Massachusetts

Computer Music Journal Vol.23 No3 (1999) The MIT Press. Cambridge, Massachusetts.

Cook. P.R (ed) (1999) *Music, Cognition, and Computerized Sound*. Massachusetts.

Darrell, T.J. & Pentland, A.P. (1996) TR#312: *Attention-driven Expression and Gesture Analysis in an Interactive Environment*. http://www.white.media.mit.edu:80/vismod/ publications/techdir/TR-312-ABSTRACT viewed 28/02/98

Davies, P. (1995) *About Time Einstein's Unfinished Revolution*. London, Viking

Davies, S. (ed) (1995) *Essays in Sound*. Sydney. Contemporary Sound Arts.

De Oliveira, N. Oxley, N. Petry, M.(1994) *Installation Art*. London. Thames and Hudson Ltd.

Dodge, C & Jerse, T.A. (1997) *Computer Music, Synthesis, Composition and Performance*. New York. Schirmer Books. 2nd Edition.

Emmerson, S. (1994) *'Live' versus 'Real-time'* Contemporary Music Review 10:2 pp. 95-101 Harwood Academic Publishers, Amsterdam Netherlands

Emmerson, S. (1996) *Local/field: Towards A Typology Of Live Electronic Music*, Journal of Electroacoustic Music Vol.9 January 1996 pp.10-12, Sonic Arts Network, London UK Also available www.sonicartsneetwork.org/ARTICLES/ARTICLE1996EMMERSON viewed 12/03/97

Frieling, R & Spitz., N(2000*) Current Media Arts*. Germany. Goethe-Institut

Gershenfeld, N. (1999) *When Things Start to Think*. London. Hodder & Stoughton.

Hand, C (1996) *Multimodal Interaction and VR*. www.cms.dmu.ac.uk/~cph/mmvr. Available by email: cph@dmu.ac.uk viewed 13/09/2000

Hillier, B (1996) *Space is the Machine*. Cambridge UK. Cambridge University Press.

Holtzman, S.(1994) *Digital Mantras*. Massachusetts. The MIT Press

Horn, R. (1999) *Concert for Buchenwald*. Germany. Scalo

Huntington, J ( 1994) *Control Systems for Live Entertainment.* Newton MA. Butterworth-Heinemann.

*Hyper-Weirdness by World Wide Web*, Science. Artificial Intelligence, Cognitive Science and Robotics, MIT Press. Available: www.physics.wisc.edu/~shalizi/hyper-weird/ai viewed 16/01/2002

Iturbide, M.R. (1995) *Unfolding the Natural Sound Object through Electroacoustic Composition.* Journal of New Music Research, Vol.24(1995), pp.384-391. © Swets & Zeitlinger

Jeffery, M (1999). *The Human Computer*. London. Little, Brown & Company(UK)

Johnson, T(1996) *Self-Similar Melodie*s. Paris. Editions 75

Katkere, A & Moezzi, S. & Kuramura, D. & Kelly, P & Jain, R. (1996) *Towards Video-Based Immersive Environments.* Available email: katkere@ucsd.edu
http://vision.ucsd.edu/papers/virtv-final/ viewed 14/09/99

Keller, D. (1999) *TOUCH'N'GO: ECOLOGICAL MODELS IN COMPOSITION* Master of Arts thesis. www.sfu.ca/sonic/EcoModelsComposition/Title viewed 06/12/2000

Kittler, F.A (1997) *Literature Media - information systems.* (Ed).John Johnston. Amsterdam. Overseas Publishers Association.

Krueger, M. .*(1976) Computer Controlled Responsive Environments*, Doctoral Thesis, University of Wisconsin, Madison

La Belle, B & Roden, S (ed). (1999) *Site of Sound of Architecture & the Ear*. Los Angeles. Smart Art Press

Leitner, B (1997) *Geometry of Sound*, Stuttgart. Reihecantz

Leonardo Music Journal Vol.6 (1966) Journal of the International Society for the Arts, Sciences and Technology. San Francisco. The MIT Press

Leopoldseder, H & Schöpf, C. (1999) *Cyberarts* 99. Linz, Springer-Verlag.

Lewis, G.E. (1989) *Singing the Alternative Interactive Blues*. Lecture at MIT Media Laboratory, Cambridge, Mass.

Lippe, C. (1991). *Using Max on the IRCAM Musical Workstation: A Musical Study*. Manuscript.

Lovell, R.e. & Mitchell, J.D. (1996). *Intelligent Stage Research Lab / Global Controller*. http://mythos.fa.asu.edu/ viewed 28/02/98

Lucier, A. (1995). *Reflections, Interviews, Scores, Writings*. Köln. MusikTexte

McNabb Software Arts. (2000) *The Java MIDI Kit*. www.mcnabb.com/software/fantasia
Available email: michael@mcnabb.com viewed 03/10/2001

Miranda, E.R (ed). (2000) *Readings in Music and Artificial Intelligence*. Amsterdam. Overseas Publishers Association N.V

Miranda, E.R. (1998) *Computer Sound Synthesis for the Electronic Musician*. Oxford. Butterworth-Heinemann.

Miranda, E.R. (2001) *Composing Music with Computers*. Oxford. Butterworth-Heinemann

Moser, M.A. & MacLeod,D (ed)(1996) *Immersed in Technology*. Massachusetts. The MIT Press

Motherwell. R: (April 1946):*Beyond the Aesthetic* in *Design* 47 #8

Mulder, A. (1995) *The I-Cube System: moving towards sensor technology for artists*. Text and graphics was published in the ISEA 95 proceedings. Available email: amulder@move.kines.sfu.ca

Nattiez, J-J. (1993) *The Boulez-Cage*. Correspondence. Cambridge. Cambridge University Press.

Morman, D. (1988) *The Psychology of Everyday Things*. USA. BasicBooks, HarperCollins.

Nyman, M. (1999) 2nd Ed. *Experimental Music: Cage and Beyond*. Cambridge. Cambridge University Press.

Oxford English Dictionary (2001) 3rd Ed. Oxford University Press

Penny, S./Smith, J/Bernhardt, A. (1999) '*Traces: Wireless full body tracking in the CAVE*'. http://www-art.cfa.cmu.edu/penny/texts/traces . viewed 21/09/2001

Poval, R. (1995) *Compositional Methods in Interactive Performance Environments*. (1995) pp109-120.

Putman, W. & Knapp, R.B. (1996) *Input/Data Acquisition System Design for Human Computer Interfacing*. Available http://ccrma-www.stanford.edu/CCRMA/Courses/252/sensors/sensors.html viewed 28/02/98

Reichardt. J Ed. (1969) *Cybernetic Serendipity* New York. Praeger

Risset, J. C. (1990) *From Piano to Computer to Piano*. In *Proceedings*. Ed. S. Arnold and G. Hair. Glasgow: ICMC Glasgow

Roads, C. (1966) 2nd Ed. *The Computer Music Tutorial*. Massachusetts. The MIT Press.

Rothko, M. (1951) *I Paint Very Large Pictures*, excerpt from *Symposium on How to Combine Architecture, Painting and Sculpture*: *Interiors* 110, #10 (May 1951)

Rowe, R. (1993) *Interactive Music Systems*. Massachusetts. The MIT Press

Rowe, R. (1996) *Incrementally Improving Interactive Music Systems*. Contemporary Music Review. © 1996 OPA (Overseas Publishers Association) Amsterdam BV www.academy.qut.edu.au/music/newmedia/Resources/Rowe1996. viewed 16/06/2002

Rowe, R. (2001) *Machine Musicianship*. Massachusetts. The MIT Press.

Russell, K. & Starner, T. & Pentland, A. (1995) *Unencumbered Virtual Environments*. http://www.-white.media.mit.edu:80/vismod/publications/techdir/TR-305-ABSTRACT viewed 12/05/99

Ryle, G. (*1995*) *The Concept of Mind*, *The Oxford Companion to Philosophy,* Oxford University Press

Salzman, E. (1967) *Twentieth-Century Music*: *an Introduction*. New Jersey. Prentice-Hall inc. 2nd Edition

Scaletti, C & Craig, A.B.(1991) *Using sound to extract meaning from complex data*. SPIE Vol. 1459 Processing, Display Interaction11 Also available email: scaletti@cerl.uiuc.edu viewed 03/08/2000

Smalley, D. (1997) *Spectromorphology: explaining sound-shapes*. Organised Sound 2:2 Cambridge University Press, Cambridge UK

Smalley, D. (1986). *Spectro-morphology and Structuring Processes. In The Language of Electroacoustic Music* (Ed.) S. Emmerson. The Macmillan Press, Basingstoke, UK

Sommerer, C & Mignonneau, L (ed) (1998) *Art @ Science*. Wien. Springer-Verlag.

Stiles, K & Selz, P. (1996) *Theories and Documents of Contemporary Art*. California. University of California Press.

Stocker, G & Schöpf (ed) (1999) Ars Electronica 99. Linz. Springer Wien.

Tarabella, L. (2000) *Gestural and Visual Approaches to Performance*. In *Trends in Gestural Control of Music*, *Musical Performance Practice on Sensor-based Instruments*.
IRCAM - Centre Pompidou..

Tenner, E. (1996) *Why Things Bite Back – New technology and the Revenge Effect*.. Clays Ltd, St Ives plc. UK

Verplank, B./ Sapp, C./ Mathews, M. *A Course on Controllers*.
http://www-ccrma.stanford.edu/courses/250a viewed 03/06/98

Webster, J.G.(ed) (1998) *Tactile Sensors for Robotics and Medicine*. USA. John Wiley & Sons. Inc.

Weibel, P. (ed)( 1992) Eigenwelt Der Apparate - Welt - Pioneers in Electronic Art.Linz. Ars Electronica.

Wessel, D.& Wright, M. *Problems and Prospects for Intimate Musical Control of Computers*
Available email: {wessel,matt}@cnmat.berkley.edu

Whitelaw, M. (1968/1998) *Rethinking A Systems Aesthetic*. ANAT: News #33 Critical Writing.
www.spin.net.au/~mitchellw viewed 17/07/01

Winkler, T. (1998) *Composing Interactive Music, Techniques and Ideas Using Max*. Massachusetts. The MIT Press

Wishart, T. (1994*) Audible Design*. UK . Orpheus the Pantomime Ltd.

Wishart, T. (1996) *On Sonic Art*. Amsterdam. Overseas Publishers Association B.V.

Wren, C.R, (2000) '*Understanding Expressive Action*.' Unpublished PhD Thesis Massachusetts Institute of Technology

X. Rodet, Y. Potard, J.-B. Barriere. (Fall. 1984) *The CHANT-- Project: From the Synthesis of the Singing Voice to Synthesis in General*. Computer Music Journal, The MIT Press

Xenakis, I. (1992) *Formalized Music, Thought and Mathematics in Music*. Hillsdale. NY Pendragon Press.

Zadorin, L (1997) *Granular Synthesis: An Introduction.*
Available email: wicked-mentality@usa.net.
www.Academy.qut.edu.au/music/newsmedia/Granular_Synthesis/granSynth viewed 06/05/2000

Zurbrugg, N (ed) (1994)*Electronic Arts in Australia*. Vol. 8, No.1 Australia. Continuum, The Australian Journal of Media & Culture.