2006

The integrated sound, space and movement environment : The uses of analogue and digital technologies to correlate topographical and gestural movement with sound

Jonathan A. Mustard
Edith Cowan University

The integrated sound, space and movement environment:

The uses of analogue and digital technologies to correlate topographical and gestural movement with sound.

Jonathan Mustard
Bachelor of Music
Master of Arts, Creative Arts

Faculty of Communications and Creative Industries
Edith Cowan University
1st July 2006

USE OF THESIS

ABSTRACT

This thesis investigates correlations between auditory parameters and parameters associated with movement in a sensitised space. The research examines those aspects of sound that form correspondences with movement, force or position of a body or bodies in a space sensitised by devices for acquiring gestural or topographical data. A wide range of digital technologies are scrutinised to establish what the most effective technologies are in order to achieve detailed and accurate information about movement in a given space, and the methods and procedures for analysis, transposition and synthesis into sound.

The thesis describes pertinent work in the field from the last 20 years, the issues that have been raised in those works and issues raised by my work in the area. The thesis draws conclusions that point to further development of an integrated model of a space that is sensitised to movement, and responds in sound in such a way that it can be appreciated by performers and audiences.

The artistic and research practices that are cited, are principally from the areas of dance-and-technology, sound installation and alternative gestural controllers for musical applications.

Key words and phrases: music, dance and technology, sound installation, alternative gestural controllers, movement analysis, motion capture, electronic music, computer music, composition, choreography.

DECLARATION

I certify that this thesis does not, to the best of my knowledge and belief:

i. Incorporate without acknowledgement any material previously submitted for a degree or diploma in any institution of higher education;
ii. Contain any material previously published or written by another person except where due reference is made in the text; or
iii. Contain any defamatory material

I also grant permission for the Library at Edith Cowan University to make duplicate copies of my thesis as required.

Acknowledgement

I would like to thank Lindsay Vickery for his encouragement to begin this project and his support for the work. I would especially like to thank Dr Maggie Phillips whose eyes, ears, intelligence and patience were indispensable in keeping the project afloat and on course.
Special thanks must go to Chrissie Parrott whose wonderful choreography; energy and commitment to some of my wilder ideas enabled the birth of some fabulous and unforgettable productions.

Table of Contents

Chapter 1

The integrated sound, space and movement environment: The uses of analogue and digital technologies to correlate topographical and gestural movement with sound.

1.1 Introduction and methodology

The aim of this research is to investigate ways in which whole-body movement can manipulate sound using electronic devices. Two aspects of movement, gestural and topographical, are considered for their potential to yield meaningful data for the manipulation of sound synthesis and musical parameters. Gesture has a substantial history of research behind it as a means of acquiring data for controlling sound, but topography, except as a sub-set of gesture, has not had the same attention. Topography, in this instance, refers to the position of a thing in a space, rather than the features of the space itself. For a definition of gesture, I have referred to the work of Insook Choi – in particular – "Gestural Primitives and the Context for Computational Processing in an Interactive Performance System." [1]. Other movement definitions that provided a background to the research derive from conventional French dance terminology and the writings of Rudolph Laban.[2].

The conventional musical parameters of pitch, tempo and dynamic as well as timbre, portamento and rhythm are explored as the effected elements of sound. Depending on the type of effects and method synthesis, other parameters are also explored as effected components of sound for their expressive potential.

The context for body movement in my practice has been in the areas of music, theatre (including puppetry), and dance. The accompanying DVD shows excerpts from experimental pieces and 'finished' creative work designed to explore the following questions.

- What aspects of sound may form perceptible correspondences with movement, force or position of a body in a space sensitised by devices for acquiring gestural and topographical data?
- At what point do the correspondences between linked movement and sound parameters become imperceptible?

- What are the most efficient technologies for acquiring movement data, transforming the data and synthesizing the sound?

[1] The first chapter surveys the literature on the types of gestural and topographical devices and the history of their application in a number of artworks from the point of view of what makes a “good” device and the debate surrounding this perspective. Also in chapter one, the basic concepts of mapping are explored. Mapping, here, refers to the way in which one or more movement parameters (inputs) are mapped to one or more sound parameters (outputs) within a given system [3].

The second chapter surveys the hardware components used for sensing movement in space. This descriptive survey attempts to answer the question of what the most effective technologies are for movement sensing in live performance and installation contexts. A more detailed exposition of mapping issues is explored in chapter two, as well, in order to put into context the kinds of mapping procedures that are implemented in my experimental pieces and creative works. The purpose of the surveys in these two chapters is to gain a familiarity with the existing research in the area, including the technologies and methodologies involved so as to position my practical work within the broader theoretical framework as well as to engage the practical work in the discourses revealed by the survey.

Chapter three documents the practical implementation of the research, from initial experimentation with different technologies, to a fully-fledged dance performance work incorporating sound, lighting, projection, costuming and dance. The process of learning to use software programs also assisted in gauging the efficiency of one or another technology as the time it takes to learn becomes a factor in determining efficiency. In this way the practical learning process has been organically linked to the methodology of the written research itself.

Chapter four outlines what I term cultural differences between three distinct artistic practices that use technology to mediate between movement and sound in an interactive environment. The question raised here is that if there are fundamentally different reasons for acquiring movement data for sound manipulation, will there be more or less appropriate technologies for data acquisition and manipulation in light of these reasons? In addition, what are the (cultural) similarities between these practices that might point to a common technology? The three distinct disciplines, or practices referred to, are music, dance and installation. The three distinct disciplines, or practices referred to, are music, dance and installation.

[1] The reader may find it useful to view the DVD more than once throughout this text so as to provide a context to the written exegesis, especially the writing concerning mapping and the journal of creative and experimental processes outlined in chapter three.

The methodology then, employs several types of qualitative approaches: interpretive, in that chapter 1 situates the research historically, while chapter 4 situates the research in a cultural frame; critical, in that some of the theories, methods and cases discussed are in need of qualification, rebuttal or reconciliation; and post-modern in that the principal narrative of the research centres around the practical and creative research that is diarised in chapter 3 [4]. While Blaxter and Hughes say that judging critically entails analysing the underlying principle and consistency of a line of reasoning, discovering its faults, weighing up the legitimacy of its supporting information and posing questions beyond the argument itself, this thesis adheres only partially to their definition in most instances. In order to limit the scope of this thesis, I have focussed my analysis on those cases that point to issues of relevance to the research questions without posing too many theoretical and philosophical questions outside those immediately relevant to the thesis questions. Merriam notes that 'information rich' case studies provide *purposeful* data for analysis [5]. The surveys and journals combine to provide an 'information rich' collection of case studies and processes as the basis for a critical analysis of theoretical positions and of contextualising the research questions and practice-based component.

As practice-based research, the methodology also takes into account attributes of the context and commentary models as well as of the research-question model discussed by Milech and Schilo. They examined the relationship between the exegesis and the practical component and came to the conclusion that the 'research-question model' best instantiates the research question that is posed in a post graduate arts degree. "Both the written and creative component of the thesis are conceptualised as *independent answers to the same research question* - independent because each component of the thesis is conducted though the "language" of a particular discourse, related because each "answers" a single research question." [6] While I hope to achieve this outcome, and while I agree with many of the arguments expressed in their article, elements of the context model in which the practice is informed and positioned by the exegesis and elements of the commentary model in which the exegesis comments on the creative production, are apparent, and to my mind useful, in this exegesis. They were useful in terms of extending the practical work in the first instance and in examining the results of the practice in the second.

1.2 Background

I will describe the most pertinent work in the field from the last 15 to 20 years covering the issues of what is a gestural and/or topographical controller, what attributes enable a controller's efficiency and in what artistic contexts do these devices occur.

Academics and practitioners adopt a variety of positions in this area of research and I will discuss key texts that represent the broad spectrum of opinion. This background work and literature places my written and practical research in a historical and theoretical context and has informed the way in which the practical work developed over the two years spanning 2002 to 2004.

The growth and diversity of alternative gestural controllers for manipulating electronically mediated sound sources have spawned a parallel growth in the literature on the subject raising many issues in regard to parameter mapping, interactivity, performance, audience reception and the relationship of electronic music in general to traditional music making.

The link between topographical data and control of sound is perhaps less well served in the literature except where systems for gestural control use this data as a means for transposition of gesture into sound. Topographical data is simply data that defines where an object is in a particular space, perhaps also where it is in relation to other objects and how much space it occupies. The literature and practice of dance-interactive performance, while equally sophisticated in the realm of interactivity, performance and audience reception for example, generally focuses on higher-level structural and aesthetic concerns rather than on the minutiae of parameter mapping. This is not to say that the many examples of topographical movement to sound mapping in dance are in any way undeveloped, simply that the literature does not focus on the finer details in the way that is apparent for writing on alternative gestural controllers in music. The few exceptions to this case are almost always articles presented to music, rather than dance conferences and journals. In some ways this tendency reflects a discourse in dance that is focused more sharply on questions of when and why various technologies are used, whereas in music, substituting one type of instrument with another merely tends to raise questions as to the effective utility of the device as an instrument in the traditional sense or as a convincing example of a novel interface.

Similarly, installation art and its associated literature focus largely on aspects of culture, for example, the interplay of cultural modalities: learning, exploration, play and reflection or the relationships between human and machine systems.

The different emphases in the above three areas of practice reflect fundamentally variant purposes for the acquisition of sensor data and as such present solutions that tend to justify particular aims within each field of arts practice. However, similar aesthetic and philosophical concerns arise across the boundaries of these practices. In particular, the affects arising because there are almost no limitations to effects, vis a vis causes, when the mediating device is a computer. These limitless effects tend to provoke the most problematic and perennial concerns. Unlike analogue devices, the computer also offers the artist a level of control over higher order structure previously unavailable. Output may be significantly displaced in time from time of input, for example, or the nature of the interactivity between 'actors' in the system may be entirely hidden from the audience's perception. One-to-one relationships between inputs and outputs may be bypassed entirely in favour of other more strategic goals.

The author's interest in correlating movement to sound stems from a long association with making music for theatre and dance. Twenty-seven years ago, I began making music for a Sydney theatre company where sometimes purely abstract movement was accompanied by a rich score, lighting, text and costume in what were essentially 'empty' spaces; spaces that contained no props or scenery. Light, sound and human movement in the space were the sole purveyors of meaning and emotion. At this time (1976-78), I had heard about the use of Theramins, but had had no direct experience of them. Later while studying at the N.S.W Conservatorium, I saw a very brief film clip of Cage and Cunningham's exploration of Theramin control of sound and this pricked my curiosity enough to want to explore the possibility of creating sound from movement in space. My curiosity was not confined solely to the use of electronic technologies as the mediating devices, but ranged across a number of approaches to the idea of making sound/music from whole-body movement. These approaches included body percussion and the design of props and costumes that would be made from various materials suitable for sound production.

For the purposes of this thesis, however, I have limited the discussion to the digital/electronic resources available to composers and performers, partly because of their inherent flexibility in terms of sound production, but also due to their adaptability to a range of art practices. This adaptability gives the composer or sound artist access to a broader situational palette than most mechanical means, extending across purely musical performance, through dance, theatre, gaming, installation and virtual spaces for the correlation of movement in space to the manipulation of sound over time.

1.3 Overview of gestural controllers

Since the advent of MIDI (Musical Instrument Digital Interface) in the early 1980s paved the way for a standard means of communicating analogue data to a digital format that could be interpreted by a computer, virtually any device that produced a voltage was fair game for use as a means of controlling sound [7] [8]. This situation has led to the invention of literally hundreds of devices for manipulating sound via a computer or other dedicated digital equipment. Prior to the invention of MIDI there were but a handful of analogue instruments that bore little or no resemblance to traditional instruments and an enormous number of electronic versions of traditional instruments. The most well known of the non-traditional types is of course the Theramin which dates back to the early years of the 20th century and is a well documented example of an alternative gestural controller [9].

Parallel to the growth in the number of gestural controllers, the discourse about these instruments has crystallized around a few key issues.

The definition of what is a gestural controller and what constitutes a good one is a major issue and has led to the formation of an international association - the Gestural Controller Group – devoted to the research and dissemination of information about this body of instruments [10].

The question of how performers and audiences understand the relationship between sound and gesture is another major issue that also contributes to the notion of 'the good controller'. The question of what is a musical gesture and what types of gestures there may be has also provoked significant discussion [11]. And finally the question of how gestures are mapped to sound parameters becomes crucial for the way in which performers and listeners are able to appreciate cause and effect and for the level of expressivity available to any particular device. The mapping strategies for gestural control are also common to topographical control of sound and therefore a common concern in dance and installation works that use, for example, video-tracking technologies [12].

1.3.1 What is a gestural controller?

Any device that may be acted upon by the body to produce sound can be said to be a gestural controller. This includes all traditional acoustic instruments, which are mechanical rather than electronic devices. Electronic versions of traditional instruments are fairly non-controversial in terms of whether they fit the profile of a gestural controller even where these versions may present a number of omissions in the level of expressivity in comparison to their acoustic counterparts. Essentially they are devices that attempt to mimic as much as possible the physical interface of the original and, though the sound output may differ considerably in terms of timbre and the locus of sound dispersion (from a speaker rather than the body of the instrument), the gesture to sound relationship is very similar.

Devices that fit the profile of non-traditional controllers often employ sensors that detect movement, light, heat, angle and force and are often deployed in novel interfaces that may or may not resemble the forms and materials of acoustic instruments. It is these types of instruments that are the main subject of discussion here. The physical interfaces of these instruments range from very simple – a computer mouse [13] – to very complex – The Accordiatron or the Miburi body suit [14] [15] and Midi Dancer [16].

The approach to these interfaces differ also in their iconography, with some interfaces, like the computer mouse and sliders like those found in audio or lighting mixing desks, being adapted from extra-musical contexts.

1.3.2 The 'Good' Controller

An informative and insightful round-table discussion was held by the Gestural Controller Group in the late 1990s and was subsequently published in Wanderley and Battier's *Trends in Gestural Control of Music*, which is broadly representative of the attitudes surrounding the problem of what the attributes of a good gestural controller may be. The differences in opinion expressed seem to derive primarily from conflicting compositional and performative standpoints. On the one hand, many, like Don Buchla, Chris Chaffe and, to a certain extent, Laetitia Sonami, express the view that gestural controllers ought to maintain the majority of characteristics of traditional acoustic instruments and I will discuss their reasoning below. On the other hand, advocates of the alternative view, such as Michael Waisvisz and Tod Machover, claim that qualities inherent in some devices themselves assist in the exploration of alternative modes of performance and composition, and thus ought not to be constrained by traditional performance modalities [17]. Despite these differences in approach there is still much agreement about design elements for these instruments. There is broad agreement that . . .

- There should be a stable and repeatable relationship between gestural behaviour and the resultant sound [18].
- The overall design must be relatively simple even if the technique for playing the controller is highly complex [19].
- In order for such a controller to approach the status of an instrument that might attract the interest and attention of a large number of performers it must be worthy of years of practice.
- The controller must be able to sustain a broad repertoire of music [20].
- The controller must be designed in such a way that the level of input effort on the part of the performer is commensurate with the attributes of the output sounds [21].

In regard to some of the points above, audience and performer perception and understanding is seen to be paramount. In order for musical comprehension to be complete, gestural behaviours must have consistently recognizable outcomes in sound. It is widely recognized that listeners, in the absence of visual cues, will still draw on a learned database or imagined set of physical gestures in order to interpret the heard sounds [22]. Unless there is a strong coupling of physical gesture to sound there is no clear recognition that the actions of the performer have much to do with what is heard. In order for this recognition to occur, actions must be repeated and must result in the same sound each time. An exception to this rule occurs when mapping procedures alter during the course of a performance. In this case, the relationship between sound and gesture must be relearned by both performer and listener in order for a new relationship to be established in a convincing way. This (re-) establishment of relationships has serious consequences for structure in music where a didactic approach to certain sections of the music seems to be necessary in order for new relationships to be revealed. What is meant here by a 'pedagogical approach in music' is that the music itself must, by some means, explain the processes that are unfolding to the listener. This might be possible, for example, by slightly varying a repeating pattern in a way that differences between iterations are highlighted; a common feature of American minimalist composers like (early) Phillip Glass and Steve Reich.

The subtlety of control immanent in traditional acoustic instruments takes many years of training to master and, as well as forming a hierarchy of performer skill from absolute beginner to accomplished professional practitioner, level of control also forms the basis for our appreciation of musical effort. If this effort is not perceivable in performance, we are deprived of a primary source of information about the subtlety and complexity of the music and the level of accomplishment required of the performer in its execution. When effort is absent or not commensurate with output, a traditional notion of what constitutes a performance must be reconfigured [23].

Michel Waisvisz argues that expression is a historical construct based largely on a Western approach to what a performer does when playing a melody and this, he asserts, lacks accurate definition [24]. The attributes above are in many ways closely aligned with a desire for instrumental types that carry the same historically framed mode of expression. Waisvisz goes on to describe how sounds themselves have become more important in a number of genres and that our notion of expression and therefore our approach to gestural controllers must take this fact into account [25]. Broadly this aligns with Godlovich's argument that our definition of performance may need to change and this is especially true at the point where low-level motoric gesture is used to define higher order musical structures than the purely moment-to-moment sounds. Users of both dance and music-oriented controllers are increasingly interested in higher-level control of structure. The desire for perceptibility and higher order structural control presents artists using these tools something of a dilemma; certainly a problem in which the tension between existing instrumental paradigms and the possibilities offered by extended structural control will need to be reconciled.

1.3.3 Types of Gesture

Some kind of reconciliation may be embedded in the way various types and functions of gesture are exploited. Insook Choi describes three types of gesture available to musicians and their instruments.

• Trajectory-based: changes of orientation; (point, grab, throw towards, put there)

• Force-based: gradient (linear) movements; (lean, push, pull, squeeze, bend, twist, blow)

• Pattern-based: quasi-periodic movements; (walk, chew, scratch)[26].

In musical performances, these gestures are independent of particular instruments and independent of specific sounds. In addition, a gesture may have more than one function. A gesture may result in a sound, but not all gestures that are important for a performer to execute and for an audience to observe and interpret affect sound output. The breath a string player may take to punctuate phrases, or the rocking and swaying motion performers make to keep time or communicate to fellow performers has no sonic outcome, but form part of an extensive gestural set for conveying musical meaning. This is quite a different situation in comparison to dance where all gestures function “intensively” - all gestures are primary carriers of meaning and none can be excluded as extraneous to the dance itself. Even an inadvertent scratch or brushing of hair from the face becomes, for the observer, part of a choreographic / theatrical event.

Pattern-based gestures are the obvious choice for extending moment-to-moment sound events into higher-level control of musical structure. Here we have gesture that is already made up from a set of micro-gestures that are repeated in some way and repetition is often the first step in defining any kind of (musical) structure. In computing terms though, the recognition of patterns, like the recognition of shapes, can be a memory intensive exercise often relying on a large database of learned patterns, which must be constantly referred to in order to decide whether incoming data is a match.

Trajectory-based and force-based gestures, when converted to digital data are computationally much easier to deal with. It is easier to demarcate a threshold along single streams of data that may be used as an index for some kind of structural change. In a trajectory-based gesture, demarcation may, for example, be at the limit of extremes for that gesture: extremes that in performance may be easily avoided until maximum or minimal effort is required to execute structural change. Similarly with force-based gestures, it is computationally easier to negotiate change at thresholds. The difference with force as opposed to trajectory is that minimum or zero force is less likely to warrant attention as an index unless perhaps importance is placed on the number of times that zero events are present in the stream of data. This *number of times* though, then becomes pattern recognition at a basic level.

1.3.4 Mapping

One kind of choice that is peculiar to the gesture-to-sound relationship in computer-mediated devices is the manner in which a particular gesture is mapped to the parameters of sound and this furnishes the composer/programmer or performer with rich territory for one type of structural control. The point at which a performer leaves behind one set of affects in order to initiate an exploration of another set can be a powerful structural juncture, both musically and theatrically, but one that presents perceptual difficulties for the audience and performer alike as mentioned above. When gestures whose sonic meanings have been established suddenly start to produce different affects there can be confusion. At least two solutions are apparent in this scenario. One solution is to introduce the new mapping procedure incrementally and the other is to change the style of gestures being used in the new section.

The incremental approach is perhaps the more *softly, softly* didactic approach and may be seen as equivalent to the traditional bridge passage or linking phrase in music that allows key modulation or new thematic material to be introduced without severe disjunction. How it is possible to introduce new mapping procedures incrementally is another problem as sometimes even small changes in mapping can have large effects.

The types of mapping strategies available are now well documented. These are . . .

- One to one: each independent gestural output manipulates a single sound output parameter.
- One to many or divergent mapping: a single input parameter manipulates several sound output parameters simultaneously and
- Many to one or convergent mapping: several input parameters manipulate a single sound output parameter [27] [28].

The one-to-one mapping strategy is the simplest and most direct form of mapping. It is easily perceivable for both performer and audience, but is expressively very limited [27]. My own experience of one-to-one mapping bears out some of this criticism, but all *compositions* operate within certain chosen parameters and often exploit the concept of a thoroughgoing exposition of a limited set of material and processes. Given the situation where any input can affect any output, there is still rich ground for creative exploration using this mapping strategy. Where the criticism is most apt, is the attempt to use this strategy for an *instrument* type that must be able to perform effectively over a range of works in many styles and reward the effort of prolonged practice with mastery.

Both divergent and convergent mapping strategies are the forms most associated with traditional instrumental practice and indeed to many other forms of physical interactivity. Hunt et al cite the example of the influence of the violin bow on the parameters of timbre, pitch and amplitude; a divergent mapping, and the effect on pitch of bow pressure, finger position and finger pressure; a convergent mapping. Convergent mapping is also apparent in the activity of riding a bicycle. Steering with the arms and shifting one's body weight both function to steer the bike and apparently humans are used to this more complex mode of experiencing and interacting with the world and objects in it. "Human operators *expect* to encounter complex mappings, and yet so often engineers provide nothing but 'one-to-one' correspondences (for example a set of sliders, each controlling a synthesis parameter)"[30].

1.4 Topographical controllers

Topographical controllers are those controllers that give information about where a thing is on a plane or in space. The types of controllers are those based around a number of acoustic, optical, electro-mechanical, electromagnetic, and electrostatic sensing devices. The most prolific of these are the optical devices, and within this category the video camera based systems are the most abundant [31]. Many of these sensors, devices and systems have been used to create sound and music at some time, though, like the devices used in new gestural controllers, they were created for use in other contexts and have been adapted by musicians looking for creative connections with other media or as new tools for exploring human movement and its relationship to sound.

> Examples of spatial sensors includes the *Radio Drum* (Mathews), which measures the location of two batons in three dimensions in relation to a rectangular radio receiver; Donald Buchla's *Lightning* (Rich), which uses an infrared signal to locate a performer within a user-definable two-dimensional grid, and several systems that analyze location and movement in space using video cameras, such as David Rokby's *Very Nervous System* (Cooper), and the *Virtual Stage Environment* (Lovell). Inexpensive pressure triggers, light beams, and motion detectors have been used extensively to identify a person's location within a space [32].

Because usually they do not involve physical contact between a moving object and the sensing device they are not used to measure force-based gestures and are not useful for haptic feedback.

Because these devices measure position in space, they are particularly good at measuring trajectory based and pattern based gestural primitives and so are often also used as gestural controllers. In as much as a change in position (that is – topography) can be equated with trajectory, topographical controllers may be said to be a sub-set of gesture. However, I have made a distinction here to avoid a meaning that limits the word *gesture* to a movement event within the boundaries of a single human kinesphere. This expanded meaning allows a more generalized discussion of topographical data and its potential for translation into sound that may include more than one person, dancers, non-musicians, non-humans, and movement made without a musical intent behind it. Topography has also been used in ways other than for the recognition or translation of trajectories and patterns. Simple positional data has been used to trigger or manipulate sound without regard to prior position, notably in works by Burt, Rockeby, Winkler, Coniglio and Paine. See Below. Topographical data is sometimes used as a first step towards gesture recognition and pattern recognition prior to these being mapped to some kind of sonic outcome and this method informs most applications in purely musical situations [33]. There seems to be very few applications of topographical data that are not limited to the movements of an individual or their individual body parts. This is perhaps as it should be because as soon as multiple entities in larger spaces become the subjects of topographical data capture, elements of theatre, dance or installation are evoked almost automatically. Even in my own work "Monody for Coloured Objects", in which a number of objects are moved on and around a table so as to manipulate sound, the theatricality of the situation becomes undeniable and hence must be incorporated with intent for the piece to work convincingly. See also Chapter 3 and 4.

1.4.1 Musician Topography – Music

The most straightforward use of topographical data in a musical setting are those installation-like pieces, sometimes made for children, that track the position of simple shapes on a surface. These are often built as educational fun pieces that have specific learning outcomes associated with them, but they are capable of serious sound and compositional outcomes as well. Tod Machover's light wheel is one example of such an interface. Various coloured squares, circles and triangles are placed on a revolving disk, like a large transparent "lazy susan", under which there are fluorescent lights. An overhead camera captures the image and the position of the shapes determines which notes are played when, how loud, for how long and other timbre-shaping effects.

Another major type of investigation in sound topography has been the placement of sound itself within a given space. This has been going on since Renaissance composers, like Gabrielli and Montiverdi, began placing choirs of voices and instruments in different parts of their church to create antiphonal and echo effects and has continued to the present day in some genres of acoustic music. Spatial articulation in music really took off in the 20th century as the (electronic) technology for sound dispersal grew in sophistication. As a consequence of these developments, the theoretical basis for the practice of spatial articulation, with particular concern for its structural implications, has also grown. Doornbusch and McIlwain confirm Stockhausen's view that sound placement is as structurally important, compositionally, as any other sound parameter, such as pitch, timbre, dynamics, tempo and so on, and go on to argue that compositional and auditory coherence is dependant on the behaviour of a sound in relation to its acoustic, visual and physical setting [34].

Spatial articulation, then, is a different issue to the definition of topography that I have proposed to investigate but is, nevertheless, investigated incidentally as a parameter of sound to be manipulated by movement, and hopefully, in a coherent manner.

1.4.2 Dancer Topography – Music

Some of the earliest works to use video capture of dance for processing topographical data were made in Australia in the 1980s by Warren Burt and dancer-choreographers, Jane Refshauge, Shona Innes and Sylvia Staehli. Entitled *Fair Exchanges*, these pieces form a set of works that use Simon Veitch's 3DIS computer vision control system. In this system, a number of areas on screen can be defined for analysis. The difference in light levels from one video frame to the next are analysed and this information is used for manipulating sound [19]. In several of the works described by Burt in an article in *Writings on Dance*, the defined areas are organized in such a way as to elicit data from discreet segments of the performance area so in a sense the performance space as a whole is fragmented so as to define set roles. These roles are generally played out by an associated dancer or, alternatively, the dancer's role defines the role of the space at any particular time in the work.

Within each of these defined areas are smaller areas or, as Burt describes them, "gangs", that are generally used as on-off trigger points[2] for sampled and synthesized sounds. The samples are variously single sounds, whole sequences or musical excerpts that evolve independently of the dancers' movements. The dancers do not control any other micro parameters in these pieces, but they do effect some larger-scale structuring of the music. Various types of delays are implemented in the work that reduces the predictability of one-to-one gesture-to-sound relationships that tend to become dull over a relatively short time frame. And certain amounts of randomness are built into one piece so that different sounds are likely to occur when any one "gang" is re-triggered.

[2] A trigger point is a pixel or a number of contiguous pixels (area) on a screen. This will correspond to a particular location in a performance space and will not change unless either the camera position is altered or the trigger point's coordinates are redefined.

Many of the issues that the artists encountered in making and performing this work are still very relevant today. Burt and dancers, Staehli and Refshauge, comment that

> composers could not think purely in sonic terms, and choreographers could not think in purely kinesthetic ones. Rather, we found it necessary to surrender the integrity of our specialist art forms in order to evolve a working method that would address both our needs and the capabilities of the system in a more holistic manner[35].

These reflections point to some of the difficulties and anxieties of operating interactive movement to sound systems using a one-to-one approach, which is close to the musical instrumental paradigm.

> One of the problems of working with invisible areas of space as sound triggers in comparison with making music with physical objects is the lack of kinesthetic feedback. Even the most insensitive synthesizer or organ keyboard allows one to feel physical contact when a sound is produced. In the case of percussion instruments, the feedback is even more pronounced, as the physical nature of playing the instrument defines much of the player's movement. In *Inside/Out* [one of the works in *Fair Exchanges*] the contradictions inherent in triggering percussive sounds with non-percussive gestures were explored, making a dance/music which used the seeming contradictions and limitations of the system. This was especially clear in the last section when the sound was suddenly switched off in the busiest part of the piece, leaving Jane to bring the energy level of the piece down, using gestures she had developed to make sound, gestures which now suddenly functioned quite differently as movement[36].

The anxiety associated with performer's and audience's perception of interactivity when the artists operate at a stage or two removed from the one-to-one approach becomes apparent.

> Only if one knew the functioning of the system very well was it possible to follow the logic of Sylvia's movements[37].

False or overstated expectations of how the technology might behave can lead to disappointment, as Refshauge observes.

> I found working with 3Dis system frustrating because sound composition was limited to turning sounds on and off. Whether you walked up and down in a pedestrian fashion or danced an intricate movement phrase the sound generated would be the same[38].

At the same time, the technology's actual capabilities may significantly alter a performer's approach to their primary medium in a positive way.

> What was novel was that the music was articulated in special zones. Floor patterning became the starting point for the moving. That is to say that the movement became a reaching out towards a sound, or, travelling along and between sound events. . . I found this relationship very challenging – to have to 'listen' physically and musically simultaneously[39].

And there are many more pertinent comments from artists that are repeated in the subsequent literature on dance and technology. Later technological developments enabled greater levels of control of the micro-parameters of a sound and these developments had the potential to resolve some of the difficulties and limitations of the systems implemented in *Fair Exchanges*, raising almost as many issues as they solved. With increased sophistication come greater possibilities but also a new level of discourse that must be engaged. In some respects, this discourse had been going on in relation to analogue and digital gestural controllers and it was up to those artists who were engaging in dance technologies to follow through with these issues in their practice.

In 1996, Tod Winkler began working on a solo piece for interactive dance with choreographer Walter Ferrero using David Rockeby's *Very Nervous System* (VNS)[40].

VNS is a system for tracking (human) motion that had been in use since 1983 in a number of Rockeby's installation works (see below). The system bears some resemblance to the 3DIS system and, in some ways, Winkler's approach is similar to Burt's with the notable exception that in many instances the dancer is given more control over the evolution of sounds and musical processes after they have been triggered. Topographical data was used primarily as trigger data, while data relating to light difference was used in manipulating sound evolution and tempo of sequences, processes or melodic fragments [41].

In Winkler’s work, regions, roughly equivalent to the “gangs” in *Fair Exchanges*, forming a 4x4 grid, assume a variety of roles during different sections of the piece and regions may combine to create super-regions for controlling musical parameters such as tempo and mixing of sounds or sequences. This approach enables a single setup to be used flexibly and to be learned effectively by the performers involved. That is, performers become familiar with the location of regions in the space and are more able to make the transition from the character of one section to the character of the next, which might be diametrically opposed, with slow, smooth flowing movements in one and quick staccato robotic motion in another[42].

Questions raised by Winkler in several of his papers include whether or not the instrumental paradigm is an appropriate one for dance-interactive work, how human movement is measured and the way in which captured data is mapped to sound.

With regard to the first question, Winkler states…

> These relationships may be established by viewing the body and space as musical instruments, free from the associations of acoustic instruments, but with similar limitations that can lend character to sound through idiomatic movements . . . Physical constraints produce unique timbral characteristics, and suggest musical material that will be idiomatic or appropriate for a particular instrument’s playing technique. These reflect the weight, force, pressure, speed and range used to produce sound. In turn, the sound reflects, in some way, the effort or energy used to create it. . . .
>
> Each part of the body has its unique limitation in terms of direction, weight, range of motion, speed and force. In addition, actions can be characterized by ease of execution, accuracy, repeatability, fatigue, and repose[43].

The ongoing collaboration of artists in the Palindrome group have been tackling these issues over a number of years attempting a decidedly holistic approach. The artists Robert Wechsler, Frieder Weiss and Butch Rovan, argue for gestural coherence between dance and music using comprehensive data capturing modes and comprehensive mapping strategies, which I will discuss in the next chapter. The system they use is *Eye-Con*; using three or more small web cameras input to a PC, data is sent via MIDI to a *Mac G3/4* running *Max/MSP*. Using this system, they engage with and explore the latest discourse in both gestural control of music and dance interactivity in a manner that audiences and performers alike are able to perceive without sacrificing artistic vision[44].

The kind of topographical data that is captured extends the notion of a simple mapping of 3D space to include a number of real and symbolic spaces. The performance space is of course one element in a complex equation that also comprises the space occupied by a single body and the space between different bodies in performance as well as movement dynamics and degree of left-right symmetry[45]. This understanding of space and corporeal relationships relates directly to the conception of dance as it is traditionally performed and provides an enormous array of data that might be mapped to sonic outcomes. Although force-sensing technologies are not directly used, topographical data is manipulated to some extent to synthesize the perception of forces in their system.

Force synthesis is done by extrapolating positional data from one moment to the next to get information about the speed of movement, the degree of expansion or contraction of the body and its relative position in space and the position of one dancer in relation to other dancers. For example, if the body is tracked from a position of contraction to a position of expansion in one part of the performance space to a far removed sector of the space in a relatively small amount of time, a large amount of force can be assumed to have been expended for this to occur.

1.4.3 Audience Topography in Interactive Installations – Music

The third main area of practice that uses topographical information to manipulate sound is installation art and perhaps the most consistent practitioner in this field has been David Rokeby. Since 1983, Rokeby has created a large body of works involving performer or audience manipulation of sound and image using a system he designed and built himself, the *Very Nervous System* (VNS) [46]. Other practitioners include Garth Paine (who also uses Rokeby's VNS), Thecla Schiphorst, Todd Winkler, Palindrome and Alex Davies to name a very few who have made work in this field.

The VNS, as mentioned above, divides the camera's vision into a user-defined number of areas that return information about variation in light from one video frame to the next. Rokeby generally uses more than one camera to obtain information about an area from different angles, giving a more three dimensional representation of activity within the space, but the precise ways in which human movement manipulates sound is discussed only in very general terms.

> …a program for the Apple II which controlled a Korg MS-20 Analog synthesizer to make sounds in response to the movements seen by the cameras. Movement also controlled the volume of two tape loops of water sounds. The synthesizer and water sounds were mixed individually to 4 seoakers [sic] in a tetrahedron (one on the ceiling and three in a triangle on the floor. The sounds would move around you in addition to responding to your movement [47].

> The computer searched for movements in these images, analysed them and created sounds in response, simultaneous to the movement itself [48].

> In these systems, I use video cameras, image processors, computers, synthesizers and a sound system to create a space in which the movements of one's body create sound and/or music [49].

In fact, he uses quite similar procedures from one work to another, while the materials and subject matter of each work vary widely. There is a predilection for natural sounds like wind, water, rain and man-made environmental sounds like footsteps and conversation mixed with synthetic sounds that more closely resemble traditional “musical sounds.”

Different systems, and different technologies, seem to provoke a variety of procedures and aesthetic approaches to correlating movement with sound. A range of technologies and their attributes are described in the following chapter, while procedural, aesthetic and philosophical approaches are discussed in chapter four.

Chapter 1 References

1. Choi, I. (2004). Gestural Primitives and the Context for Computational Processing in an Interactive Performance System. 1st. 2000. CD Rom. Ircam M.M. Wanderley and M. Battier. p. 194-247.

2. Laban, R. (1966). The Language of Movement: A Guidebook to Choreutics. Boston, Plays.

3. Rovan, J., Wanderly, M., Dubnov, S. and Depalle, P. (1997). Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance. Kansei- The Technology of Emotion Workshop, Genova - Italia. p. 68-73.

4. Blaxter, L., Hughes, C., and Tight, M. (2002). How to research (Second ed.). Berkshire, Open University Press. p. 61 & 238.

5. Merriam, S. B. (1998). Qualitative research and case study applications in education. San Francisco, Jossey-Bass. p. 8 & 211.

6. Milech, B. & Schilo, A. (2004). 'Exit Jesus': Relating the Exegesis and Creative/Production Components of a Research Thesis. TEXT Special Issue No 3 (April 2004). p. 5-7.

7. Dobrian, C. (2001). Aesthetic Considerations in the Use of "Virtual" Music Instruments. Society for Electro-Acoustic Music in the United States(Spring 2003). p.1-4.

8. Winkler, T. (1995). Making Motion Musical: Gesture Mapping Strategies for Interactive Computer Music. Proceedings of the International Computer Music Conference, ICMC. p. 261.

9. Battier, M. (2004). Electronic Music and Gesture. 1st. 2000. CD Rom. Ircam. Wanderley, M. and Battier, M. p. 8.

10. Wanderley, M. and Rovan, J. (1997). Gesture Research in Music: Who, What, and Why.... IRCAM. Available: http://www.ircam.fr/equipes/analyse-synthese/wanderle/Gestes/Externe/people.html. Accessed 18.08.04.

11. Choi, I. (2004). Gestural Primitives and the Context for Computational Processing in an Interactive Performance System. 1st. 2000. CD Rom. Ircam. Wanderley, M and Battier, M. p. 194-199.

12. Wechsler, R and Rovan, J. (2001). … Seine Hohle Form… - Artistic Collaboration in an Interactive Dance and Music Performance Environment. Proceedings of the Cosign Conference. Amsterdam: Cosign. p. 43-47.

13. Hunt, A and Kirk, R. (2004). Mapping Strategies for Musical Performance. 2000. CD Rom. Ircam M.M. Wanderley and M. Battier. p. 391.

14. Gurevich, M. (2001). The Accordiatron: A Midi Controller for Interactive Music. Proceedings of the 2001 New Interfaces for Musical Expression. Seattle, USA: NIME. p. 1-2.

15. Mustard, Jonathan. Correlating Movement in Spaceto the Parameters of Sound. Proceedings of the 2002 Australasian Computer Music Conference. Melbourne: ACMA, 2002. p. 91- 98.

16. Coniglio, M. (2004). Mididancer, Troika Ranch. http://www.troikaranch.org/mididancer.html. Accessed 29.01.05.

17. Buchla, Sonami and Machover. Round Table. 1st. 2000. CD Rom. Ircam M.M. Wanderley and M. Battier2004. p. 639-641.

18. Buchla, D., Sonami, L. and Machover, T. (2004). Round Table. 1st. 2000. CD Rom. Ircam. M.M. Wanderley and M. Battier. p. 424.

19. Machover, M.M. Wanderley and M. Battier and. Round Table. 1st. 2000. CD Rom. Ircam2004. p. 424-5.

20. Buchla, Sonami and Machover. Round Table. 1st. 2000. CD Rom. Ircam. Wanderley, M. and Battier, M. 2004. p. 426.

21. Waisvisz, Michel. (2004). Round Table. 1st. 2000. CD rom. Ircam M.M. Wanderley and M. Battier. p. 425 & 633.

22. Choi, I. (2004). Gestural Primitives and the Context for Computational Processing in an Interactive Performance System. 1st. 2000. CD Rom. Ircam. Wanderley, M and Battier, M. p. 207.

23. Godlovitch, S. (1998). Musical Performance: A Philosophical Study. London: Routledge. p. 109

24. Waisvisz, Michel. (2004). Round Table. 1st. 2000. CD rom. Ircam M.M. Wanderley and M. Battier. p. 631.

25. ibid. p. 631.

26. Choi, I. (2004). Gestural Primitives and the Context for Computational Processing in an Interactive Performance System. 1st. 2000. CD Rom. Ircam. Wanderley, M and Battier, M. p 207-208.

27. Rovan, Wanderley, Dubnov & Depalle. (1997). Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance. Genova: The Technology of Emotion Workshop, 1997. p. 69.

28. Hunt, A and Kirk, R. (2004). Mapping Strategies for Musical Performance. 2000. CD Rom. Ircam M.M. Wanderley and M. Battier. p 387 – 427.

29. ibid. p 390.

30. ibid. p 388.

31. Mulder, A. (1998). Human movement tracking technology: resources. Vancouver, Simon Fraser University. p. 1-17.

32. Winkler, T. (1995). Making Motion Musical: Gesture Mapping Strategies for Interactive Computer Music. Proceedings of the International Computer Music Conference, ICMC. p 2.

33. Modler, P. (2003). An Experimental Set of Hand Gestures for Expressive Control of Musical Parameters in Realtime. Proceedings of the 2003 Conference on New Interfaces for Musical Expression, Montreal, Canada, NIME.

34. Doornbusch, P. & McIlwain, P. (2003). Converging Technologies. Australasian Computer Music Conference, Western Australian Academy of Performing Arts at Edith Cowan UIniversity, Perth Australia, Australasian Computer Music Association. p. 75-80.

35. Burt, W. (1990). Fair Exchanges. Writings on Dance. 5 (Autumn). p. 40.

36. ibid. p 41.

37. ibid. p 42.

38. ibid. p 46.

39. Staehli in Burt. p 47.

40. Winkler, T. (1997). Creating Interactive Dance With the Very Nervous System. Connecticut College Symposium on Arts and Technology, Connecticut, Connecticut College. p. 1-7

41. ibid. p 3.

42. Winkler, T. (1998). Motion-Sensing Music: Artistic and Technical Challenges in Two Works for Dance. Proceedings of the International computer Music Conference, Bucharest, ICMC. p. 1-5.

43. Winkler, T. (1995). Making Motion Musical: Gesture Mapping Strategies for Interactive Computer Music. Proceedings of the International Computer Music Conference, ICMC. p. 1-4.

44. Wechsler, R and Rovan, J. (2001). … Seine Hohle Form… - Artistic Collaboration in an Interactive Dance and Music Performance Environment. Proceedings of the Cosign Conference. Amsterdam: Cosign, 2001. p. 45.

45. ibid. p 46.

46. Rokeby, D. (1983-4). Reflextions. Ontario, mac.com. Home pages. http://homepage.mac.com/davidrokeby/reflexions/. Accessed 19.10.04.

47. ibid.

48. Rokeby, D. (1984). Body Language. Ontario, mac.com. Home pages. http://homepage.mac.com/davidrokeby/bodylanguage/. Accessed 19.10.04.

49. Rokeby, D. (1986). Very Nervous System. Ontario, mac.com. Home pages. http://homepage.mac.com/davidrokeby/. Accessed 19.10.04.

Chapter 2

2.1 Software, Hardware and Mapping

2.1.1 Introduction

This chapter describes a range of technologies that have been or are currently in use as a means of capturing human movement data for either gestural or topographical control of sound. As mentioned in chapter one, there has been a proliferation of technologies enabling the gestural control of sound synthesis. Some of these devices and their component parts are relevant to spatial sensitisation and I will include mention of these particular cases from time to time, but there are far too many gestural controllers of the musical instrument type to discuss all of them in detail. Technologies that deal directly with whole-body movement or movements associated with dance are analysed as the primary focus in this chapter.[3]

The purpose of this overview is to identify technologies that are flexible, accurate, efficient and economically feasible. Further criteria for defining these attributes are, in the case of flexibility, a system or device that is capable of application in a broad range of circumstances. That is, the system or device must be able to satisfy many artistic goals - goals that its creators might not have imagined. Accuracy also implies reliability and minimal latency. To be accurate, any repeated input must deliver the same data at each iteration. The system's efficiency may be gauged by the time it takes to deliver data throughout the system from the onset of any measured input to the onset of output. Efficiency may also be gauged by how much processing power is required for functions over a range of complexity.

Measures of efficiency and economy must also include the time it takes to learn how to use the system and how much time is spent in modifications for each separate artistic application. Of course, there is a great deal of subjectivity involved in judgments concerned with the making of complex systems, productions that, by their very nature, are time-consuming. The monetary cost of any one system is a factor to be measured, but may be offset against time efficiencies. Unfortunately there is not very much information available on the question of how long it takes to learn or apply the knowledge of many systems, so

[3] For a detailed discussion of gestural controllers for solely musical application I would refer the reader to *Trends in Gestural Control of Music* edited by Marcelo Wanderley and Marc Battier[1] and Wanderley's PhD Thesis, *Performer-Instrument Interaction: Applications to Gestural Control of Sound Synthesis* for Bibliographic data and as a starting point for further research[2].

some judgments must be made on information about, for example, the necessity for cleaning up inaccuracies or for modifying various system attributes.

The second half of this chapter deals with the issue of mapping; the strategies that are used to map input data to sound parameter output and the current thinking that tackles the problem of why one would map the data one way or another. Extensive research has been undertaken in the field of gestural control of music synthesis and much of this research is applicable to systems for the whole-body manipulation of sound. The primary differences of opinion between “musical instrument” and “whole-body” paradigms lie in whether one or another strategy is best applied in each case, not what the strategies are.

2.2 Hardware

The hardware incorporated into systems for the capture of human movement range from the simplest types of binary sensors to complex arrays of vision sensing equipment, sonar, ultrasound, radar and magnetic field transducers. In most cases, the intervening devices between data capture and sound / image output is a computer. Seldom in the last fifteen to twenty years are there other dedicated analogue or digital devices that mediate input and output data. This means that, in all cases, the physical forces at play in the world must be converted first of all to a voltage and then to a digital representation of that voltage; re-processed, re-mapped and the output is yet another voltage that is converted into either sound waves or light waves via a speaker, video projector or monitor. In a small number of cases, the output is a mechanical device like a robot, motor or light-switch or light-fader.

The layers of technology required in the translation of one media to another therefore involve a delay from the onset of any monitored event to a result in another medium. Depending on the processing power and sophistication of the hardware, this latency may be as little as a few milliseconds (imperceptible) or as much as several seconds in some “real-time” systems. Other factors that contribute to latency are the amount of information that must be processed. The greater the level of monitoring and/or output parameters, the greater the latency is likely to be.

The latest advances in wireless and USB communication over MIDI and the speed of bulk information processing have expanded to the point where significant, perceivable latency times common two or three years ago have been virtually eliminated. The parallel advances in data manipulation techniques, however, have meant that gains in processing speed have been partly offset by the amount of power required by some new techniques. This is not to say that one must always use these techniques simply because they exist, so the gains made in diminishing latency are real, but where processor-hungry techniques are used, speed may be sacrificed.

2.2.2 Sensors

Sensors are devices that monitor the physical forces in the real world and convert them (usually) into electrical impulses. Although there are a limited number of physical forces, there are many more ways of converting these forces into their electrical representations. This has led to the development of hundreds of types of sensors with varying capacities for delivering information about the world.[4] Since 1998, a number of new products have been developed for commercial and research production. The *Miburi* body suit, *Midi Dancer*[3], *BodySynth*[4], *Digital Dance Suit*[5], *SoundBeam2*[6] and motion capture suits like *Gypsy Gyro*™, *Motion Captor*™, *Gypsy 4*™, *Gypsy Torso*™[7] and *Vicon 8* designed for animation, sport and performance. The *Miburi* suit is now out of commercial production. *Miburi* was produced by Yamaha and, like their *WX11* wind controller, Yamaha decided that these devices were not commercially viable or were likely to be superseded.

In terms of suits that are available at the cheaper end of the market, the disappearance of *Miburi* means that there is not much competition or variety of commercially available devices in *Miburi's* price range (up to US$2,000). Some developments in motion capture suits and the general trajectory of technology becoming less expensive over time may perhaps fill this vacuum. *Gypsy Torso*™ is available for under US$5000 though it captures only torso and head motion. Arm elevation on this suit is partly limited by the machinery on the shoulder making some dance movements impossible. The next level of full motion capture suits start at around US$20,000.

The major differences between dance suits like *Miburi* and *BodySynth* lies in the number, type and sensitivity of the sensors embedded in them and the applications for which they were designed. Before I discuss the differences and similarities, I will provide a brief overview of the types of sensors and the physical attributes that they measure. More technical information is available in the appendix and via the references below.

The most common types of sensors used for human motion[5] detection are:

- Pressure - measuring the amount of pressure exerted on a single point or discreet area. [appendix**i**] The pressure sensor can be used for both gestural and topographical motion capture depending on whether it is placed on the body

[4] An extensive list of sensor systems relating to human movement tracking is available from http://fas.sfu.ca:70/7m/cs/people/ResearchStaff/amulder/personal/vmi/HMTT.add.html, Axel Mulder's addendum to a technical report of July 1994 for the School of Kinesiology, Simon Fraser University and updated in May 1998[8].

[5] Other sensors may be more common, but are used mainly for industrial and scientific measurement of non-human physical forces.

or somewhere in the performance space. The pressure sensor may be a simple binary sensor, but more usually it has a dynamic range.

- Flexion – measures the angle of flex between two points. [appendixii] There are several types of flexion sensors on the market offering one, two or six degrees of freedom. Flexion sensors can really only be used to measure bodily gestures or the angle between objects and have not been used to deliver topographical information.
- Light – sensitive to light levels which can measure the level of ambient light or the level of light within a discreet beam of space. [appendixiii] These are typically used in installations or on instruments for triggering the presence of an object at a particular point in space or at a particular point on an instrument. Photocell sensors that measure the general ambient light can also be used dynamically in installation and performance.
- Proximity – measures the closeness of an object to the sensor. [appendixiv] There are several types of proximity sensor; infrared, ultra-sound, sonar and radar[9]. Other position-sensing devices, by their very nature may also be used to determine the distance between objects. Proximity sensors are most often used in installations, sometimes in instruments and occasionally for dance.
- Angle – measures the angle of rotation between the sensor and an object fixed to it. [appendixv] Used in a wide variety of situations, the potentiometer version of this sensor is most commonly recognized as the volume knob on your average stereo system. Other non-contact angle measuring sensors include a type of Hall effect sensor – see below.
- Hall effect – a general-purpose sensor-technique that is used for a variety of measurements including proximity, position, speed, tilt and rotation. [appendixvi] It is often used in conjunction with other sensors as a means of dealing with magnetic fields.
- Accelerometer – measures the linear or rotary speed of an object. There are two basic types of sensors in this category: magnetic sensors and inertial gyroscopes. Magnetic sensors are used most often for unaided human movement. Some motion capture suits also use inertial gyros.
- Inertial gyros – measure angular rates and six degrees of linear acceleration. [appendixvii] They are particularly good for delivering very accurate, clean data for motion capture and are appropriate for wireless application. As yet no association of whole body movement with sound or music has been

implemented with this technology, being used primarily for gaming, animation, gait analysis and sport.

Miburi

The *Miburi* body suit was released by Yamaha in 1994 and represents perhaps the only digression into alternative controllers by a major electronics company. The Miburi has 8 fingertip pressure sensors and 4 heel and toe pressure sensors, 6 flex sensors at the arm joints, 2 angle sensors at the thumbs and 2 pressure sensors also at the thumbs. Pressure sensors at fingers and thumbs are similar to those found in standard music keyboards delivering binary on/off information at a certain threshold of pressure and initial pressure level or velocity, but no "aftertouch" information[10]. In other words, they are a middle ground between simple binary on/off pressure switches and fully dynamic pressure sensors conforming to the MIDI protocol for note on/off and velocity, where the velocity at which a button is pressed translates to the loudness of a note's attack.

Aftertouch sensitivity would have given the performer the ability to dynamically control effects between the initiation and cessation of an event from the one sensor. This is alleviated to a certain extent by the bend sensors at the thumbs. The limitation with this arrangement is that it is less intuitive for a performer to initiate an event with one sensor, control its evolution with another and then turn it off with the original sensor. Even when a performer masters playing in this way, this still leaves the two thumbs to modulate effects for eight fingers and adds complex programming requirements to the system.

Despite these limitations and others, the *Miburi* delivers stable and reliable data at low rates of latency. The creation of a small repertoire of work from composers and performers such as Lindsay Vickery, Hiroshi Cu Okubo, Saburo Hirano and Susumu Hirasawa across a fairly wide range of contemporary genres has meant that the *Miburi* more closely resembles the definition of a successful alternative gestural controller than many other attempts (See Chapter 1 – 2.2 "The Good Controller").

Because *Miburi* is a MIDI controller, many different software applications may be used in conjunction with it and, by extension, the controlled hardware are also many and varied. Vickery, for example, uses MAX/MSP and Image-ine softwares to map the *Miburi* for sound and image control. He has also used external devices such as the Ensoniq and a MIDI lighting desk as the end hardware for sample manipulation and control of lights.

Yamaha produced its own synthesizer that the *Miburi* could control directly with or without the computer as a mediating device. The synthesizer's sounds in this case were fairly limited and without the intervention of another computer program only one-to-one mappings were possible[11].

Digital Dance Suit

The *Digital Dance Suit* developed by Siegel and Jacobsen at DIEM (Danish Institute of Electroacoustic Music) in 1998 is a much simpler, more open system than *Miburi* consisting of only one type of sensor. Six bend sensors are distributed over various parts of the body and are capable of capturing the angle and velocity of flexion. *Digital Dance Suit* has the advantage of being a wireless system (*Miburi* is [legally] wireless only in Japan) providing plenty of freedom for the dancer to execute movement over a wide area. To date, there are few records of performances using this system with one example of a work by Wayne Siegel performed at the 1999 Columbia University Interactive Arts Festival [12]. In 1998, the suit was available for US$1897. The system included an interface to convert the analogue voltage to a digital signal and thence to MIDI information [13]. No additional software is included in the package except, for example, MAX patches demonstrating sensor monitoring.

Midi Dancer

Midi Dancer consists of eight flex sensors typically positioned at locations like the elbows, wrists, hips, or knees and like the *Digital Dance Suit* is capable of capturing the angle and velocity of joint flexion. Information is encoded and transmitted wirelessly to a receiver 30 times per second. The receiver digitises the information and converts it to MIDI information that represents the position and acceleration of the performer’s limbs[14]. Again because it is a MIDI device, rather than duplicating existing MIDI capable software, the user is assumed to have appropriate music software. *Midi Dancer* is used exclusively by Troika Ranch and was never in commercial production, but a significant body of work has emerged between 1989 and 2003 incorporating this device. Recently Troika Ranch have been concentrating on the possibilities of emerging camera tracking technology and *Midi Dancer* hasn’t been “taken out of the cupboard” for a while, though Coniglio says it may appear again. “The thing that’s great about it [*Midi Dancer*] compared to camera stuff, is that it always works in performance, can be calibrated on the fly, and I have a deep understanding of how to get information from it and use it in an aesthetic way”[15].

BodySynth

BodySynth uses only four EMG sensors to monitor the muscles of your choice while you move. The system includes a twelve-channel processor so that input may be sent to different channels to modulate numerous effects simultaneously. MIDI information that can be sent is note on/off, pitchbend, continuous controllers and modulation wheel. The system may be used with or without a computer, as the control of parameters is adjustable from the MIDI unit's display panel[16].

Perhaps *BodySynth*'s most recognizable artist is the artist named Pamela Z, an American vocalist renowned for her use of extended vocal techniques. Pamela Z frequently uses *BodySynth* as a way of controlling other elements in her performance, often manipulating her own voice either in real-time or in the real-time manipulation of pre-recorded instrumental and vocal samples. Although not a dancer, per se, Z references dance movement in a way that highlights the subject matter of her performance. In her work, *Gaijin* (Japanese for "foreigner"), she references the movement of Geisha dance to explore her experience of living in Japan as a foreigner and the feelings of alienation she experienced there[17].

The four systems above are all representatives of paradigms Axel Mulder describes as "inside-in". All use on-body sensors to return data about the body. Other paradigms are "inside-out" - on-body sensors that return information about external (artificial) sources and "outside-in" - off-body sensors that return information about the body[18]. The current favourite for "outside-in" sensing is easily won by video.

2.3 Video

Since the computerization of video data in the mid 80's, a number of software programs have been developed in order to extract gestural and topographical data from the moving image. Programs like *EyeCon*, *EyesWeb*, the *Very Nervous System* (*VNS*), *Cyclops*, *Jitter* and *Isadora*® all have features that enable one or another or both forms of data extraction. Video by its very nature fits the "outside-in" paradigm and this paradigm is common to the assumptions behind the construction of all the above programs, but the way in which data is processed by these softwares differs in kind and quality.

EyesWeb is a PC based program (though an open platform version is available for beta testing as I write) that was explicitly designed for the "analysis and processing of expressive gesture in movement, MIDI, audio, and music signals"[19]. It has been extensively used by a relatively large number of experimental/contemporary performance ensembles and individuals, including Palindrome Intermedia group, Troika Ranch, Laboratorio di Informatica Musicale, Luciano Berio and many others. EyesWeb's focus on expressive gesture is indicative of the interests of the Laboratorio's principal researchers and artists, in particular Antonio Camurri who has had a long association with research into gesture and alternative gestural controllers. Topographical data can be extracted using this program's trajectory and space analysis functions, though its reliance on the division of the video frame into grids inhibits some forms of spatial analysis (See Cyclops below and Chapter 3, Experimental Modelling).

The ***Very Nervous System***, synonymous with its creator, David Rokeby, has been around in various guises for over 20 years. The first systems used simple hand-built cameras with a very low 8x8 pixel resolution, hand-built processors and digitisers. An *Apple II* computer was used for mapping the pre-processed information to sound. “The sounds were produced by custom software running on a Mountain Hardware Digital synthesizer sitting in the *Apple II*”[20]. Since 1984, continuous improvements on all the components in the system were implemented, enhancing camera, processor and digitiser equipment. Then in 2001, *softVNS* and, in 2003, *softVNS2* were created to do away with all but the camera components of the external equipment. Now virtually any digital camera may be used with *softVNS* software to sense motion in video. A number of *softVNS* tools are available for video tracking including presence and motion tracking, head-tracking, colour-tracking and object following.

The software is fully compatible with *MAX/MSP/Jitter*, allowing any kind of mapping procedure available via *MAX's* object library. *SoftVNS* is stable and reliable for most basic video functions without perceivable loss in quality. The tracking tools are relatively reliable in stable lighting conditions and the software is faster and easier to learn than equivalent functions in *Jitter*.

Since 1983, Rokeby has created over 20 installations in more than 90 locations throughout the world using the *Very Nervous System*, some of which have been active seven days per week, 24 hours per day for over a year with no down time[21]. *VNS* has been used for dance as well as installation work.

Todd Winkler describes several works he has created for dancers that demonstrate the flexibility and accuracy of *VNS* and *softVNS* from 1997 to 2003. Using *softVNS* as the engine of motion analysis,

> moving into specific locations on stage can start or stop various musical functions, trigger specific sounds or cue video events. Continuous data, representing the dancer's overall speed, is used in the audio realm for such things as timbre shaping via filters, sample playback speed or delay. In the video realm, continuous data may be applied to image offset, colour, luminance, or distortion[22] [6]
>
> The live soloist is projected as a composite grid made up of nine video panels; each delayed differently in time. The movement is designed as an ensemble piece, with the live soloist creating dynamic interactions

[6] “Continuous data” in this context refers to a stream of digits representing, for example, scaled changes in levels of light over time as opposed to “trigger data” or “binary data” that return only true or false statements about a thing without any gradation between the two states.

> between the nine projected versions of herself as they appear to move apart, come together, touch each other or disappear out of the frame. Specific audio samples with similar time delays are triggered by location, while speed alters timbral characteristics via pitch shifting and flanging[23].

Priced at US$350, *softVNS2* seems to be a bargain. One just has to bear in mind that prior ownership of *MAX/MSP* (at US $900) is almost a must and at least one decent camera (Au$2000) is recommended. Complex data processing in *MAX* requires years of learning, but can be extremely rewarding.

Winkler's descriptions also seem to imply that a number of mapping strategies are in use simultaneously or at least serially. While there is a preponderance of one-to-one mappings, the one-to-many strategy is also a feature. He says that location triggers the onset of video or audio events, which is a one-to-one relationship, but the morphology through time of this event is subject to a one-to-many strategy – speed alters pitch and flanging, continuous data (speed or light level for example) alters image offset, colour, luminance, or distortion. Though this last example may imply one effect <u>or</u> another, one effect <u>and</u> another is also possible.

Cyclops

Cyclops is a *MAX/MSP* object written by Eric Singer that processes video data. Its functions include both colour and grey-scale analysis of incoming video frames. Each frame may be divided into a user-determined number of rectangles called zones. The difference between grey or colour analysis options in a zone is compared to the same zone in subsequent frames and this data is output as a list. The list may then be used in any way possible in the larger *MAX/MSP* programming environment[24]. *Cyclops* is compatible with *Jitter* in *Mac OS9*, but has yet to be ported to *OSX*[7]. Fairly reliable results were achieved in controlled lighting conditions, but some bugs and the perennial problem of operating in the grid or zone paradigm limits the type of data achievable with this software. *Cyclops* only does video analysis and does not include processing for video output. At US$100 it may return its value in satisfaction, depending on what you want to do. It is relatively easy to learn if you are moderately familiar with *MAX/MSP* application-building rules.
See chapter 3 for a more detailed discussion of *Cyclops* as it was used in my modelling and experimentation.

[7] Apple Macintosh's operating system version ten.

Jitter

Jitter is a group of *MAX* objects for the capture, analysis, synthesis, manipulation and output of video. Like *MAX*'s MIDI and audio functions, its 130 (and rising) objects allow for an infinite variety of discourse between video and other media. The ability to specify an enormous array of possibilities comes with the warning – this is a complex and difficult program to learn from scratch – indeed, it is often difficult for veteran users of *MAX/MSP*.

The advantage of this software's architecture is that for any given artistic goal often only small portions of the software need to be learned. In *Jitter*, live colour-tracking, for example, really only requires learning 5 or 6 objects and perhaps 25 to 30 commands. The kind of data that is returned in colour-tracking enables a wider variety of approaches to the analysis of colour in video than most other systems. Straightforward lists of coordinates are returned for each colour that is tracked, enabling information like height, width, area and centre-positions relative to the overall dimensions to be determined easily. The total area may also be divided into a grid for analysis of individual cells if so desired.

The reliability of colour-tracking in general, that is, in any system, not just *Jitter* is problematic. Digital cameras and computers are objective machines and do not respond or analyse colour in the way that eyes and brains do. While the human brain may compensate for radical changes in light and shade, recognising the colour red as being the same red whether it is in full light or part shadow, a computer will not. Even specifying broad-spectrum light frequencies for tracking will not mimic the subjective operations of the human brain. Add to this problem other variables like changing lighting states in a theatrical situation, or the movement of a performer from light into shadow or the shadow created by the curve of a performer's costume around their body, then calibration of a computer program to track what humans track so effortlessly becomes very difficult.

The results, accepting the computer's mode of "seeing", however, can be rewarding and interesting. See Chapter 3 for details of my work using Jitter for colour-tracking in live performance.

2.4 Motion Capture

Motion capture (mo cap') systems refer to systems that are used to analyse movement for scientific and medical purposes and also to those systems that are used in films and video games for various types of animation. Historically, motion capture systems have been regarded as the "high end" of the market, often costing around ¼ million dollars. This kind of price tag has meant that most artists working in performance have effectively been prohibited from accessing this technology. The exceptions have been limited to short periods of experimental work or pieces that use the cheaper motion capture systems. Only recently has there been experimentation using motion capture to generate or manipulate sound[25].

Probably because of the high cost though, there is fierce competition between the various systems and manufacturers, which has already impacted on price and system sophistication seeing a quantum leap in data accuracy, reliability and a plunging price-tag with full wireless motion capture systems available for around US$80,000[26].

There are three main types of motion capture, distinguished by the type of sensor used to capture the motion data: video, magnetic, and electromechanical motion capture. More broadly, these types fall into two categories; *passive* and *active* motion capture [27]. *Passive* describes most forms of video motion capture where reflective markers are used and *active* refers to any system where the nodes on a body being tracked actively transmit data; whether that is in the form of an electromagnetic field, biofeedback, inertial gyroscopic data (another way of doing magnetic fields) or ultrasonic pulses.

2.4.2 Video Motion Capture

Video motion capture generally involves tracking a number of markers placed on the body by four or more cameras placed around the area in which the body is moving. Usually more than four cameras are used for more accurate data capture. The issue here is that parts of the body hide markers from the camera's view from time to time so more cameras mean less occlusion. Some video motion capture systems use infrared markers and cameras, while others use reflective balls or button-like tabs attached to black jump-suit costumes. There can be anything from eight to thirty points on the body tagged for tracking. Most video motion capture systems can be used in real time, though for fine detailed work using many markers, occlusion and track point confusion in the system is common and usually means that the resulting data needs to be "cleaned up".

Video motion capture works by calculating the position of each marker on the body by triangulation between two or more cameras. The data that is returned for each video frame is in the form of a three dimensional x, y, and z coordinate for each marker. With frame rates of up to 190 per second, this can mean dealing with potentially 17,100 (3x30x190) pieces of information every second. Even a more musically significant frame rate of about 50 frames per second can still yield 4,500 pieces of information every second. Double this figure if nyquist criteria apply[8].

There are many limitations to the use of video motion capture in performance. The capture area is limited to a relatively small space; $3m x 8m^2$ is about the maximum for the best video systems – much smaller than most stages for dance. Set up time and calibration prohibits portability and use in typical proscenium stage situations. Lighting must be bright and static

[8] This is a technical term meaning that the sampling rate needs to be twice the frequency of the most rapid feature of interest.

for reflective systems or dark for infrared systems. Costuming is sometimes limited to the perennial black jump-suit or leotard. One advantage of video motion capture over some other systems is that the number and positioning of markers is flexible. Live performance use of video motion capture systems may best be utilised in telematic or distributed performance situations where an audience might be viewing/hearing a projected/amplified visual/sonic avatar.

Because mo cap' data is not in a format that communicates with music software programs a way of transforming this data in real time is required in order for a music software program such as MAX to process it. In 2003, Christopher Dobrian reported on a project involving the *Vicon 8* video motion capture system and his design of a software system that would translate *Vicon's* data for use in controlling sound.

> The initial design for MCM [Motion Control Music] intended to make as simple—and as simple to use—a program as possible for mapping motion capture data to musical control data. The design allows for the user to select a marker (i.e. a position on the body), a coordinate (*x*, *y*, or *z*), and a range of space in which to track that coordinate, and linearly map that data to any range of MIDI values for any MIDI channel message. The user can specify as many such mappings as desired. The intention was to make the most direct possible way for any motion capture parameter to be used to control a MIDI device. For more complex mappings, it was assumed that an additional program would mediate the transmission between MCM and the MIDI device(s)[27b].

Despite the final statement about more complex mappings, the specifics of the program's implementation seems to limit translation of markers to the many-to-many style of mapping (see 'Mapping' below). There does not seem to be any way in which several markers may influence a single stream of MIDI (many-to-one), an important strategy for damping or amplifying a number of timbral qualities[28].

2.4.3 Magnetic Motion Capture

Magnetic motion capture works by using a sensor to measure the magnetic field created by a source. The source and sensors are cabled together and connected to a control unit, which is then cabled to a computer running a driver program that reads the data and communicates with an animation program[29]. (Ascension Technology Corporation have recently produced a wireless version of their magnetic motion capture system so that the performer is not tethered by cables to a control unit).

Typical human motion capture solutions for animation place sensors at the joints of limbs, but with software filters that can infer joint position, given the length of each limb, fewer sensors still enable accurate motion capture. The types of filters available are *Biped*, by Alias | Wavefront and *Inverse Kinematics* (I.K.) by Kinemation. Data returned from the sensors give both the absolute position and the rotation of joints.

> Because the magnetic systems provide data in real-time, the director and actors can observe the results of the motion capture both during the actual take and immediately after . . . . This tight feedback loop makes magnetic motion capture ideally suited for situations in which the motion range is limited and direct interaction between the actor, director, and computer character is important[30].

Animation artists have in the past held magnetic motion capture in high regard for its cleanness, accuracy and efficiency. It is still the preferred option in many film studios because of this reputation and the level of expertise through extensive use that operators have gained over time. Magnetic mo cap' though, has a limited live performance capability.
Usually only one performer may be 'captured' at any one time because multiple performers increase the likelihood of signal interference and/or "cross-talk"[9]. (Here again, Ascension Technology, claim that their system minimises signal interference and up to five performers can now be tracked) [31]. Performance areas also need to be free from metal, as this will interfere with the signal. Signals also tend to deteriorate towards the edges of the capture volume (performance area).
To date, no known use of Magnetic Motion Capture for manipulating sound has been undertaken.

2.4.4 (Electro-mechanical) Exoskeletal and Gyro Motion Capture

In an Exoskeletal type of set-up the performer must wear a bulky set of metal rods that are attached to their back and limbs. This external skeleton is then forced to move with the movements of the performer and the rotation of each joint is measured using rotation sensors. One advantage of this technology is that there is no interference from magnetic fields or light and any number of performers may be active simultaneously.
The problem with this technology is that there is no measurement of absolute position so a jumping movement for example returns only the movements of the limbs during the action

[9] "Cross talk" is a confusion between two signals where one sensor may be mistaken for another in the signal stream.

and not the lift and fall of the whole body from the ground. The other obvious drawback is that the device itself restricts certain movements. A dancer thrashing about on the floor for example will most likely damage themselves and the device. Extreme extensions of limbs are also inhibited by the limit of travel of the device's rods and joints.

One solution to the problem of absolute positioning is to include one or two markers for optical tracking in tandem with the suit.

Gyroscope motion capture is included under this heading as this device also lacks the ability to return absolute positioning data. The difference in comparison with the exoskeletal device is that there is very little to impinge a performer's movement. Up to 19 inertial gyroscope sensors (small solid state sensors) fit snugly into the joint positions of a lightweight suit and may be worn underneath normal clothing[32].

Other claims by MetaMotion for their *Gypsy Gyro* suit, available for under U.S. $80,000 are that up to 64 actors may be working in these suits simultaneously. The range, or capture volume for these wireless devices is at least the size of a football field. MetaMotion claim that the only (earthly) environment not suited for *Gypsy*'s use is under water and that all the data may be handled in real time on a single laptop computer.

In 2002, Company in Space, a Melbourne-based dance company premiered *The Light Room* using, among other devices, Gypsy's Exoskeletal motion capture suit. In this work the solo performer's movement manipulates images and text, but not sound.

> The LIGHT ROOM was also an interactive installation by day time. The general museum public could play with these devices/interfaces attached to gloves and left on the glass table. Three people could share the control of the projection of this text, by moving the gloves freely in space. The data collected, not only amplified the way the text was projected but also was used to trigger sound, both spoken word and composed score. So by sitting at the table they shared the navigation of this projected and aural environment[33].

I asked David Chesworth, the musical collaborator in this project, whether there were any manipulations of sound by the motion capture performer and whether the link was gestural or location-based in nature.

His response was:

> There wasn't any direct linkage to the motion suit or any video capture. We did discuss it but I have never been interested or convinced in the

> value of mirroring or interpreting actual visual or movement gestures in sound. For me it is too literal a starting point and presents me with a situation where the only creative path forward is to introduce distortions in that relationship to avoid banality, which kind of defeats the original purpose.” [34] (See Appendix for full text [appendixviii]).

I think that there are more ways of dealing with the issue technically, aesthetically and philosophically than Chesworth has explored in his response here. No doubt he has looked at some of them in the context of the above collaboration. Given the company’s interest in manipulating image and text using the mo cap’ suit and the use of the data glove to trigger sound in the installation, I thought Chesworth’s reasoning at odds with the company’s goals in multi-media interactivity.

2.5 Optimal and minimal requirements for gestural and topographical motion tracking.

In order to define a system that effectively monitors the movement of a body or bodies in space, the forces, types of gesture and the spatial relationships that appear to be crucial in defining movement require some clarification.

- Gravity,
- Pressure and tension (of weights, masses and musculature),
- Fine motor gestures,
- Gross motor gestures,
- Position of body in space,
- Position of body parts in relation to one another,
- Position of bodies in relation to one another,
- Position of a body’s parts in relation to another body’s parts

The points above are only a sketch of elements that are strong carriers of meaning in human movement and it appears that none of the systems mentioned so far would, by themselves, be capable of delivering data for all of these parameters. A system that comes close would need to include: a computer model that could simulate the effect of gravity on the body being tracked, pressure sensors and muscle sensors to measure pressure and muscle tension, face tracking and some kind of data glove for monitoring hand and feet movements for fine motor gestures, a sensor to measure the absolute position of each body in space, enough sensors to measure the position and rotation of the limbs of each body without signal confusion or interference.

Such a system implies a combination of optical motion capture, data gloves, gyro' motion capture and numerous pressure sensors and muscle sensors distributed over the body. Such a system barely seems feasible in terms of economics and data management, so what might be the next best thing? Which carriers of meaning are we able to do without? Which forces at play in our sketch of human movement parameters are least important?

I think this will depend on individual situations and styles of dance if it is dance that is being monitored. A dance work whose primary focus is fine motor gesture will need a different system than a dance focused on gross motor movement and positional relationships.

These are somewhat crude distinctions that do not always reflect the reality of many dance pieces, or the way in which an audience perceives dance. Most dance works will combine or juxtapose a focus on one or another parameter within the space of a single work and an audience will always be made up of a variety of people who will interpret the forces at play on the stage with their own predilections as the focal point of meaning and appreciation.

A more suitable question might be: which are the forces that are most apt for a transposition into sound at any particular point in a performance? Making a choice about which kind of data to monitor will assist in managing data in our optimal system and make it a more feasible option. In doing so, we will also be making an artistic choice about what is important to us in the moment under observation. The artists' choices are also what an audience understands as the ultimate referent in any contemporary work of art.

Some examples of minimal systems are ones mentioned above, like *Miburi*, *Midi Dancer* and *Body Synth*, but these focus entirely on gesture without reference to the body's relationship to its environment or to other bodies or objects in the performance area. I would suggest here that a minimal system that is able to monitor these other parameters might be a better instrument for more complete tracking of human movement. A *Miburi* or a gyro motion suit coupled with a video motion tracking system would enable some gestural and some positional data to be extracted from the scene. The parameters of force - pressure and tension - may be inferred from this data by calculating speed and rates of acceleration or deceleration of any given mass. Pressure sensors in the *Miburi*, for example, might also be modified to return continuous data dynamically for a direct measurement of pressure on the feet. Pressure sensors could also replace some flexion sensors in *Midi Dancer* or be added to a gyro mo cap' suit.

Any positional measurement (like video motion tracking) that is reasonably continuous over time can be said to infer the affect of gravity - gravity being a constant force in a scene. However, our perception of gravity is very much heightened in some situations. Where a performer is “flying”[10], climbing a ladder, performing at a height, or simply being lifted high off the ground, an expectation is created that is informed by our own proprioception. To a certain extent, the higher we are from the ground, the greater is our fear of falling. What we do with this measurement of height and how and when it is useful to use, as an indicator of gravity is another variable, rather than a constant.

In a parallel of gravity, the element of relational positioning of bodies on the stage can have a similar theatrical impact. Two people moving towards each other, for example, is often used to heighten dramatic tension and the further apart they were to begin with, the greater the excitement as they come closer together. Whether one has in mind the soft-focus picture of a couple closing in on each other from opposite ends of a beach or the circling movements of mortal enemies about to engage in hand to hand combat or some less clichéd set of relationships, the capturing of positional relationships is useful in reflecting the drama inherent in the topography of the scene.

The period in modern dance in the latter half of the C20th was partly characterized by works that played with various illusions to do with ownership of limbs. Duos, trios, quartets and every other imaginable ensemble of dancers – even solos – have played with the confusion between bodies and their constituent parts, where limbs seem to take on their own separate life and bodies seem to have gained more parts than the usual. [Alwin Nikolais’ *Grotto*, Merce Cunningham’s *Lines in Space*, Ann Halprin’s *Parades and Changes* and just about every other late C20th choreographer[35].

Musically, there are equivalents in the techniques of masking and re-spelling. Roughly speaking - masking is where the attack of one instrument masks the entry of another instrumental timbre that, for example, continues a melodic line and respelling refers to the ambiguous nature of notes in the system of tonality so that any note may be “re-spelled” to a new role in a new key.

All of the elements of movement mentioned above seem to be indispensable capture parameters for a system that attempts a thoroughgoing approach to motion capture; that is flexible for any situation and provides artistic choices. Certainly such a system will be expensive – in the order of $100,000 – and the amount of information to process daunting, but it is not out of the question. Two or three computers would be required to manage the information (for 2 – 3 dancers – more computers for additional dancers in mo-cap’ suits) and a small team of operators and programmers to enable real-time translation of mo cap’ data to a music software program like MAX. Other parts of the system for capture of video and force-sensing data are already compatible and well developed for immediate use.

[10] “Flying” a performer is a stage effect generally using cables and rigging - not literally flying.

2.6 Mapping

Mapping refers to the way in which input data is mapped to output data. The input data is usually an analogue signal of some kind that has been converted to a stream of digits that a computer software program can understand. Since 1984, most electronic musical instruments have used the MIDI (Musical instrument digital interface) format to communicate with either a computer or another MIDI device. The type of parameters that such an instrument might generate are …

- Note on and off,
- Note velocity (how much force was applied to the device for any particular note), Aftertouch (how much force is applied to a note over time),
- Pitch bend (how much a note's frequency is altered over time),
- Modulation (the depth and frequency of vibrato type effects) and a number of controllers that can be used to modify volume, reverberation, chorus effects and portamento to name a few.

If you were to attempt using a MIDI instrument to drive a car for example, the choices you make in determining which of these input parameters would be applied to steering, braking, acceleration and changing gears could be described as a mapping strategy. You would soon discover that the type of input parameters you used and the way in which they interacted and were scaled to deliver the kind of performance necessary for accident-free driving is non-trivial. In addition, you would discover that the haptic feedback that you get from a car's controls (steering wheel, pedals etc…) and the feel of the tyres on the road is a crucial function in driving that will need to be emulated somehow on your instrument.

Of course the reverse is also true. If you were to use a car as a controller for a MIDI instrument, the decision about which inputs (steering wheel, pedals etc…) go to make notes or volume or vibrato is a mapping strategy, and a non-trivial one at that.

Considerable research has been undertaken in the field of mapping input to output parameters in the last seven or eight years, especially in relation to gestural controllers. Some research into mapping as it relates to dance and non-contact control of musical parameters is also beginning to mature. Artists who, inspired by the possibilities presented by new technologies, have come up against problems and questions that require solutions and answers in order to make effective work and extend their initial forays into the artistic applications of the media are doing most of this research.

2.6.2 Mapping Strategies

As mentioned in chapter 1, there are three basic mapping strategies.

- One to one: each independent gestural output manipulates a single sound output parameter.
- One to many or divergent mapping: a single input parameter manipulates several sound output parameters simultaneously and
- Many to one or convergent mapping: several input parameters manipulate a single sound output parameter[36] [37].

These can be graphically represented as

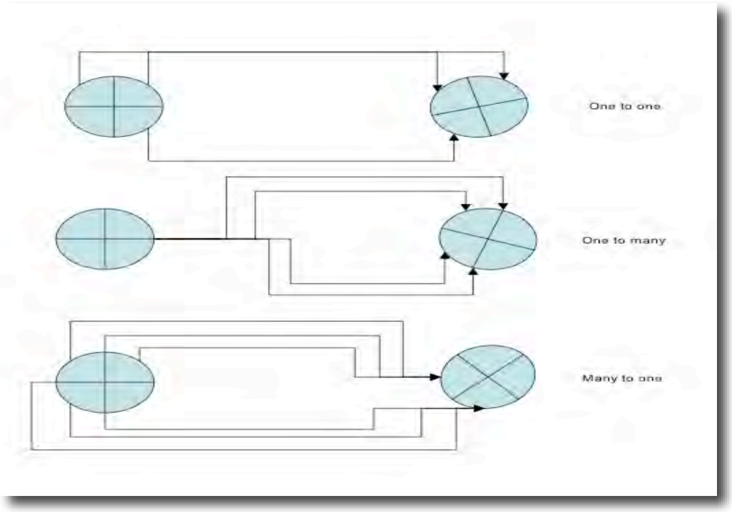


Fig.1 Representation of common mapping strategies

Hunt has posited a fourth strategy, many to many, but as a strategy that is a combination of two or more of the above strategies[38]. I call this mixed mode mapping to emphasise its hybrid nature. The simplest and most direct strategy is the *one to one* strategy. The literature on this strategy suggests that it is the most limited in terms of its potential for expressivity. However, it is noted that the *one to one* strategy is most easily discernable for performers and audiences in the context of a dance-interactive piece. For this reason (perceivability), it is often used in conjunction with other mappings where artists feel that performer and audience perception of input/output connectivity is important.

In works by both the Palindrome Intermedia group and various pieces by Todd Winkler, positional data, that is – a specific point or area in a space – is a designated trigger point and will, for a particular section of a piece, result in the same sound being made audible when a performer comes into contact with this area[39]. In a way, air-space has been divided up like a virtual keyboard whose notes simply await sounding by objects passing through it[40].

Successful works that make use of the one-to-one mapping strategy exclusively are few and its use is well considered both in terms of a compositional limitation (artistic choice) and as an easily discernable referent for performers' and audiences' perception of interactivity in the work. In fact, for dance / music interactive works the opposite is true. There are hardly any pieces I know of that do not use a one-to-one mapping at some point and this is precisely because of this mapping's limitation and perceptual accessibility. The examples cited above are rather one-dimensional in that "contact" with a point in space produces the same sound over and over again.

A more dynamic model is also common where, for example, the horizontal position of a performer controls the panning position of a sound[41]. The accumulation of multiple one-to-one mappings can also be an effective strategy for generating quite complex sounds over short periods. In the modelling experiments described in chapter 3, multiple one-to-one mappings are often used. As a research strategy for determining which parameters translate most effectively to sound parameters, a one-to-one mapping strategy, in my opinion, is an indispensable tool. All experimental research describing mapping strategies employ this tool if only as a kind of placebo for the investigation of something else[42].

The many-to-one and one-to-many strategies can also provide direct correspondences that are readily perceivable when handled with care and given time for practice. Correlating the horizontal position of a moving body to pitch and panning or using horizontal position and horizontal width to determine the bandwidth of frequencies are both easily perceivable mappings. All of these mappings and correlates can become tedious in a relatively short period of time, but when added together, these multiple simplicities can quickly become complex and engaging sonic events. The focus of interest then becomes "not only the direct correlations, but also the counterpoint of mappings" - between the different "voices"[43].

2.6.3 Mixed Mode Mapping

Many-to-one and one-to-many strategies most closely resemble traditional instrumental paradigms. Any wind instrument may provide a case study. Fingering, breath and lips (embouchure) all combine to influence the pitch of a sound. There is also a one-to-many relationship in this example - as fingering, breath and embouchure can also affect the timbre of a tone. Audiences readily perceive the effects of these mappings in the common practice of instrumental music. These are the same complex mappings that

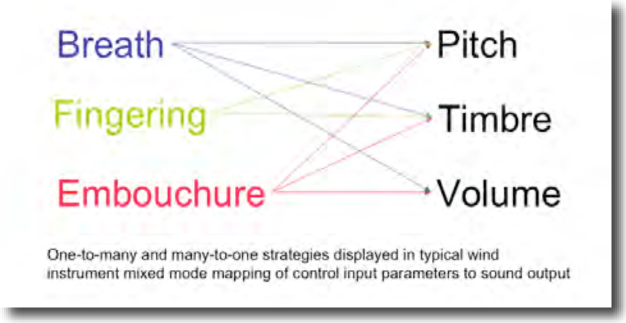


Fig. 2

we use to drive a car or ride a bike, but what happens when these same inputs and mappings are used to drive/ride/play something else? There seems to be no natural complex mapping procedure that easily facilitates audience or performer perception of which inputs are controlling which outputs. This situation arises because output parameters are completely arbitrary, having no relationship to the natural physical world. What is required is the building of an abstract physicality in the mind of the performer(s) and the mind of the audience is. The need for mindfulness of this abstract physical world is the same need that is satisfied by traditional music and dance situations and that is the acknowledgement and recognition of physical effort to produce particular movements or sounds[44].

Over time and with greater exposure audiences might eventually be able to "read" these mappings and thus more fully appreciate the effort that produces the effect. The problem with this position is that each new work takes advantage of the myriad possibilities for mapping one way or another. Each new work becomes its own icon for the mappings involved. In this scenario, there is nothing for an audience to get used to outside of the frame of a single performance. Should a large enough body of work that makes use of cross media translation / transposition ever be broadcast to a large enough audience base, the one thing that will be noted is the need to learn a new abstract physical world for each new work.

This is not necessarily a negative. Parallels can be drawn to the changes and challenges that new works within individual disciplines create for their audiences on a regular basis. All the major developments in art from impressionism, cubism, surrealism to jazz, aleatoric music and music concrete have required that audiences view and listen to the work and indeed the world in new ways. Transmedial performance practices and the mappings that form them ask us to consider virtual worlds where physical laws may be suspended and, by extension, the boundaries and limitations of the body we inhabit.

2.6.4 Choosing Parameter Output

Prior to determining which parameters to control with a dancer's input, the question of which structural level of sound is reckoned to be the most desirable to control must be decided. Will movement manipulate the microscopic details of timbre, the middle ground of pitch, rhythm and dynamics or the macroscopic form that structures the overall shape of the piece? What happens in the case of multiple performers, each of who

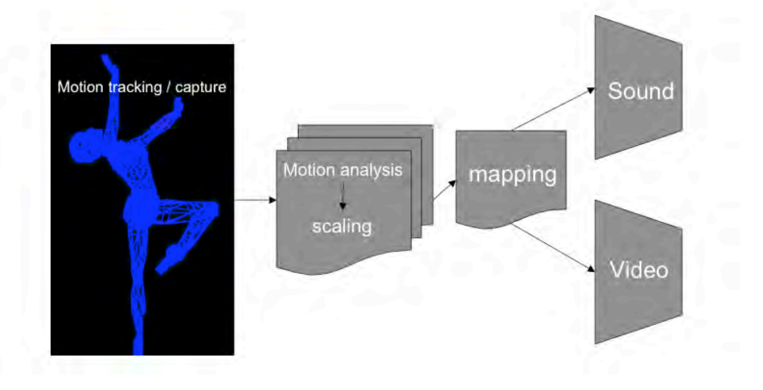


Fig. 3 Staging movement input to an unlike output

may have some role in shaping sound parameters?

In 2002, Frédéric Bevilacqua reported on his research using the *Vicon 8* video motion capture system for gestural control of music in which the quality of gestures were the focus for translation to sound. Two types of analysis were performed on a moving subject to detect or recognize a gesture. Gesture recognition requires the comparison of a performed gesture with a library of gestures stored in the computer and gesture detection relies on changes in the quality of a moving subject's gesture. The general indicators for changes in quality of a gesture were acceleration, speed and change in direction[45]. These indicators were then used as points of departure for mapping gesture to sound one way or another. In Bevilacqua's system, 30 markers were used on a single performer to capture their movement.

In 2002, at the time of writing his report, the way in which the data was used to map to sound parameters was still under investigation but reference is made to triggering events and using continuous data based on gestural segments, pattern recognition and principal component analysis. Gestural segmentation is, crudely, significant changes in the trajectory of a gesture as indicated by (radical) changes in acceleration. And principal component analysis is a way of pattern recognition that identifies the principal components of a gesture or pose that are matched to a library of gestures / poses[46]. The reasoning behind this approach is two-fold. Firstly, the frame rate of the motion capture system, which at 30 Hz or 900 pieces of information per second, must be reduced to a manageable, musically significant size and, secondly, that the performer need only be concerned about the quality of the gesture and not their orientation in space. With this approach, the motion analysis is kept distinct from its transformation (mapping) into sound – in other words a further (abstract) mapping precedes mapping to sound.

Actually, this is no different to any other system. Every system ever made for transposing movement to sound has included this layer of analysis prior to sound parameter mapping. This other layer of data manipulation is not always as complex as Bevilacqua's and may simply be a re-scaling of input data in order to match significant audio scaling (see below). The point is that the question remains the same regardless of the complexity of the system, and the basic structure for movement analysis, mapping and output are very similar across all systems. Bevilaqua's problem is still his choice of what inputs, however modified by his techniques of analysis, to use in manipulating sound. The fact that empirical (rather than deductive) data about movement in 3D space is available to him may add to the problem in comparison with regular 2D video data, but there are plenty of analogies in sound by which connecting lines may be drawn. 3D and 2D parameter space modelling for acoustic properties abound[47] [48].

Whichever method of verbal / graphic description of sound is employed the actual parameters remain pretty constant. Whether the perspective is from "multidimensional parameter space", "phase space" or whatever, essentially one ends up having to deal with the same materials and there are no hard and fast rules as to which one goes with what.

With all these systems, the affluence of data may be as much a bane as a boon. The problem of mapping all of this information to sound parameters is at least as daunting as writing for an orchestra of orchestras.

2.6.5 Scaling Input and Output

Some constants seem to hold true for how input parameters are mapped to output such as that very few musical parameters respond well to straight linear input. Regardless of mapping strategy, a straight line is less likely to render an expressive result in any parameter unless it is hidden in curves manipulating other parameters around it. Straight lines (from 0 to 127 for example) might be expedient in some situations, but, in my opinion, they shouldn't be left exposed.

Almost every parameter in every domain – pitch, dynamic, timbre, panning, flanging and reverberation – responds more expressively when progress between upper and lower limits follows a path that is not straight. This is especially so if any single sound is exposed for a significant duration on its own.

Chapter 2 References

1. Wanderley, M. (2004). Trends in Gestural Control of Music. Wanderley, M. and Battier, M. Paris, IRCAM. p. 1-788.

2. Wanderley, M. (2001). Performer-Instrument Interaction: Applications to Gestural Control of Sound Synthesis. IRCAM – CENTRE POMPIDOU. Paris, UNIVERSITY PARIS 6. p. 1-174.

3. Coniglio, M. (2004). Mididancer, Troika Ranch. http://www.troikaranch.org/mididancer.html. Accessed 08.08.04.

4. Van-Raalte, C. (2002). BodySynth. UK, BodySynth. www.synthzone.com/bsynth.html. Accessed 20.01.05.

5. Birchfield, D. (1999). 1999 Columbia University Interactive Arts Festival. Computer Music Journal. 24. p. 1-8.

6. Carter, C. (2001). Music and Movement. Sound on Sound. October 2001. p. 1-3.

7. MetaMotion (2004). Gypsy Gyro. USA, MetaMotion. http://www.metamotion.com. Accessed 20.01.05.

8. Mulder, A. Fels and Mase (1998). Human movement tracking technology: resources. Vancouver, Simon Fraser University. p. 1-17. http://www.xspasm.com/x/sfu/vmi/HMTT.add.html. Accessed 15.05.03.

9. Raes, G.-W. (2003). Gesture Controlled Musical Instruments. Gent, Hogeschool Gent - Music & Drama Department, Orpheus Institute, Logos Foundation. p. 1-24.

10. Vickery, L. (2002). The Yamaha MIBURI MIDI jump suit as a controller for STEIM's Interactive Video software Image/ine. Australasian Computer Music Conference, Melbourne, ACMA. p. 181.

11. ibid. p 183.

12. Birchfield, D. (1999). 1999 Columbia University Interactive Arts Festival. Computer Music Journal. 24. p. 1-8.

13. ibid. p 1.

14. Coniglio, M. (2004). Mididancer, Troika Ranch. http://www.troikaranch.org/mididancer.html. Accessed 08.08.04.

15. Coniglio, M. (2005). MidiDancer. J. Mustard. New York, unpublished**:** e-mail communication.

16. Van-Raalte, C. (2002). BodySynth. UK, BodySynth. www.synthzone.com/bsynth.html. Accessed 20.01.05.

17. Durling, K. (2000). The Electronic Avant-Garde, Hearing Anew, Moving Ahead. Contemporary Music Review. May 6 2000. p. 1-2.

18. Mulder, A. (1996). The I-Cube System User Manual. T. Sternberg. Los Angeles, Infusion Systems. p. 66.

19. Laboratorio di Informatica Musicale. (2004). EyesWeb Open Platform. Genova, InfoMus Lab. www.eyesweb.org. Accessed 27.01.05.

20. Rokeby, D. (1984). Body Language. Ontario, mac.com. http://homepage.mac.com/davidrokeby/body.html. Accessed 19.10.04.

21. Rokeby, D. (2004). softVNS 2 real time video processing and tracking software for Max. Ontario, mac.com. http://homepage.mac.com/davidrokeby/softVNS.html. Accessed 19.10.04.

22. Winkler, T. (2003). Live Video and Sound Processing for Dance. Video, Technology and Performance Festival, Brown University. p. 1-3.

23. ibid. p 2.

24. Singer, E. (2001). Cyclops. Vers. 1. Computer software. Eric Singer, 2001. Mac, plug in for Max-Msp. p. 1.

25. Dobrian, C. (2003). Gestural Control of Music Using the Vicon 8 Motion Capture System. New Interfaces for Musical Expression, Montréal, Québec, Canada, NIME. p. 1-4.

26. MetaMotion (2004). Gypsy Gyro. USA, MetaMotion. www.metamotion.com. Accessed 20.01.05.

27. Furniss, M. (2000). Motion Capture. Massachusetts, MIT Communications Forum. Accessed 04.04.07.

27b. Dobrian, C. (2003). Gestural Control of Music Using the Vicon 8 Motion Capture System. New Interfaces for Musical Expression, Montréal, Québec, Canada, NIME. p. 2.

28. Hunt, A., Wanderley & Kirk (2000.). Towards a Model for Instrumental Mapping in Expert Musical Interaction. International Computer Music Conference, San Francisco, International Computer Music Association. p. 387 – 427.

29. Dyer, S. Martin, J., Zulauf, J. (1995). Motion Capture White Paper, Windlight Studios, Alias and Wavefront. p. 3.

30. ibid. p 5.

31. Ascension Technology Corporation. (2006). Products - MotionStar Wireless 2. Burlington, Ascension Technology Corporation. http://www.ascension-tech.com/products/motionstarwireless.php. Accessed 4.04.2006.

32. MetaMotion (2004). Gypsy Gyro. USA, MetaMotion. Accessed 20.01.05. www.metamotion.com

33. Sky, H. (2002). Company in Space. Melbourne, www.companyinspace.com. Accessed 24.02.05.

34. Chesworth, D. (2005). Re: your co.inspace collaboration. J. Mustard. Melbourne. unpublished**:** e-mail communication.

35. Highwater, J. (1978). Dance: Rituals of Experience. New York, A & W Publishers, Inc. p. 45.

36. Rovan, Wanderley, Dubnov & Depalle. (1997). Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance. Genova: The Technology of Emotion Workshop, 1997. p. 70.

37. Hunt, A and Kirk, R. (2004). Mapping Strategies for Musical Performance. 2000. CD Rom. Ircam. Wanderley, M and Battier, M. p. 390.

38. ibid. p 391.

39. Winkler, T. (1998). Motion-Sensing Music: Artistic and Technical Challenges in Two Works for Dance. Proceedings of the International computer Music Conference, Bucharest, ICMC. p. 2.

40. Wechsler, R and Rovan, J. (2001). … Seine Hohle Form… - Artistic Collaboration in an Interactive Dance and Music Performance Environment. Proceedings of the Cosign Conference. Amsterdam: Cosign, 2001. p. 43.

41. Mustard, J. (2002). Correlating Movement In Space To The Parameters Of Sound. Proceedings of the 2002 Australasian Computer Music Conference, Melbourne, ACMA. p. 94.

42. Hunt, A and Kirk, R. (2004). Mapping Strategies for Musical Performance. 2000. CD Rom. Ircam. Wanderley, M and Battier, M. p. 392-395.

43. Dobrian, C. (2001). "Aesthetic Considerations in the Use of "Virtual" Music Instruments." Society for Electro-Acoustic Music in the United States(Spring 2003). p. 1-4.

44. Godlovitch, S. (1998). Musical Performance: A Philosophical Study. London, Routledge. p. 97-107.

45. Bevilacqua, F., Ridenour, J, and Cuccia, D. (2002). 3D motion capture data: motion analysis and mapping to music. Workshop/Symposium on Sensing and Input for Media-centric Systems, Santa Barbara CA, SIMS. p. 3-5.

46. ibid. p 4.

47. Mulder, A. Fels and Mase (1998). Human movement tracking technology: resources. Vancouver, Simon Fraser University: 1-17. http://fas.sfu.ca:70/7m/cs/people/ResearchStaff/amulder/personal/vmi/HMTT.add.html. Accessed 10.02.04.

48. Choi, I. (2004). Gestural Primitives and the Context for Computational Processing in an Interactive Performance System. 1st. 2000. CD Rom. Ircam. Wanderley, M and Battier, M. p. 238.

Chapter 3.

3.1 Documentation of real world experimentation

What follows is a journal, or diary of the development of techniques and processes leading to the creation of some fairly substantial creative works. Some of the text derives from notes that were taken at the time the practice was being developed, while other parts of the text were written quite some time after the projects had been completed, so there is a combination of reflections here that are more or less reflexive, depending on their distance in time from the work in the studio. Generally, the greater the distance in time from the practice, the greater is the reflexivity in the commentary on the practice.

The principal method operating in this chapter, then, is the ‘Commentary Model’, which is, in recent Australian history, a fairly standard method that "'elaborates, elucidates and contextualises the […] creative work'"; that "'*present[s] the research framework: the key questions, the theories, the disciplinary and wider contexts, of the project*'"; that "'*tells the story of the research: its aims, its methods, its achievements*'"[1].

The part of the model that is the focus here is the ‘story’ of the practical research that elucidates the development of experimental etudes based on the idea that there may be some mapping procedures that are more or less appropriate for the transposition of topographical data to sound parameters in a broad range of circumstances. It also charts the learning process that flows from the acquisition of new software tools. In this case, the main tools used were *Cyclops* and *Jitter*.

In August 2002 the studio acquired *Cyclops,* a *MAX/MSP* object written by Eric Singer that processes video data. Its functions include both colour and grey-scale analysis of incoming video frames. Each frame may be divided into a user-determined number of rectangles called zones. The difference between grey or colour analysis options in a zone is compared to the same zone in subsequent frames and this data is output as a list. The list may then be used in any way possible in the larger *MAX/MSP* programming environment[1].

My initial concept for correlating the movement of an object to sound parameters was based on the conceit that an object in an empty space displaces the volume of part of that total space. An auditory analogue of this situation would be that an equivalent sonic object inhabits an equivalent volume of the auditory space. As an example of what an auditory space might be I chose white noise as a starting point. White noise is usually recognized as consisting of all sound in much the same way that the entire colour spectrum is represented in the light vibration of white, hence the term “white noise”.

In this conceit then, an object occupying the total space would sound as white noise and an object occupying a subset of that space would be represented as “pink noise”. In other words the object moving in space is analogous to a filter, which ‘reveals’ a portion of this white noise.

From this initial standpoint it is then a matter of determining the correlation of behaviours of the object in space to the behaviours of a filter. The initial conceit or analogy does not offer any clues, a priori, as to behavioural correlation; therefore subjective, cultural correlates must apply to the details of one parameter and another.

An example might be:

- horizontal *dimensions* (width) of an object mapped to bandwidth,
- horizontal *position* in space to stereo disposition of the sound and
- vertical position to frequency (pitch).
- depth position mapped to reverberation amplitude.

Vertical dimension and depth dimension of a three dimensional object in this analogy becomes problematic in relation to pitch and reverberation because, in the case of the former, bandwidth (a spectrum of pitches) is already dealt with by horizontal dimension and, in the case of the latter, depth position already determines reverberation amplitude.

Finding consistent and comprehensible correlations for three-dimensional objects was a primary task of the etudes and the examples below and still is a problem for future work. Throughout the development of the various etudes, slightly different approaches to the “space-to-noise” equivalence have been used. In general, I have modified noise to more discreet sub-sets of noise in the pitch domain such as the chromatic scale and other artificial scales. In most instances, I have tried to retain the notion of bandwidth, which, in stepped pitch space, will be equivalent to a chord or cluster of notes. Due to lack of processing power on the computer or unreliable results, bandwidth was dispensed with in some cases. Polyphony in general seems to remain problematic in on-board computing situations due to both processing limitations and the clumsy programming requirements of some software environments.

The choices above that correlate horizontal width and vertical position to bandwidth and pitch were choices that were self-consciously acculturated. Pitch, in Western European culture is most often associated with height: i.e. how high or how low a sound is in relation to other sounds around it. And bandwidth is strongly associated with keyboard practice: i.e. how wide our fingers and hands might stretch in the horizontal plane in order to play notes simultaneously. This culturally specific kind of correlation I see as being safe. The correlation is safe in the sense that it is more likely to be perceivable and comprehensible to an audience of that culture. So the correlation is not arbitrary in this instance. Other, more arbitrary (less culturally specific) correlations are explored in future etudes, both for their expressive features and to test levels of perceptibility.

In creating and naming *MAX* patches (a collection of interacting functions representing a purpose-built software system), I have usually started with a filename followed by a version number in the form filename.01. Patches that have successive version numbers may have a family resemblance to the original; they may be "bug-fixes" of a previous version or they may be completely new patches that simply operate along similar performance paradigms.

3.2 Experimental Modelling

Cyclops.01

In this patch, the video frame was divided into 100 numbered zones, each with a point at the centre, that reports average grey level for that rectangle. Zones were numbered from the top right to the bottom left as follows.

Table 1 showing zone number configuration in Cyclops.01

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	…						

\- - -

91	92	93	94	95	96	97	98	99	100

The zone data behaved as a filter for a noise generator in the following way. The grey level, a number between 0 and 255 where 0=black and 255=white, was reported by each zone and was used to modulate only the first 100 bands of 255 bands of frequencies. The amplitude of each frequency band measured between 0 and 255 where 0=silence and 255=maximum amplitude.

Variations in light level over a number of zones, usual for any object moving across the sensory field, resulted in groups of frequency bands increasing and decreasing in amplitude in a discontinuous manner. This discontinuity was due to the numbering of the zones.

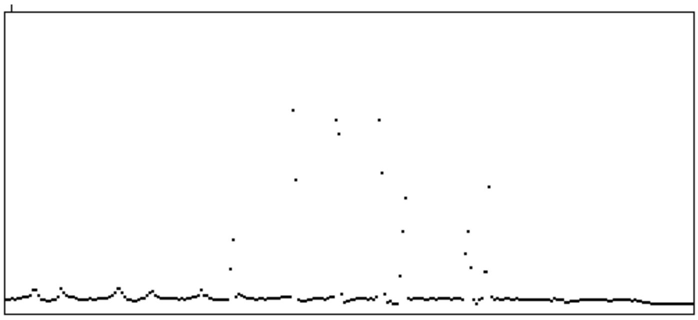


Fig.1 Shows amplitude of non-contiguous filter points against background noise

This particular patch was therefore not true to the analogy because contiguous zone data is erroneously mapped to non-contiguous filter points on the noise generator. However, the connection between image and sound was still perceived to be an analogous one-to-one relationship.

In order to reduce the problem of non-contiguous filter points and to more accurately pair the zone data to the noise generator's filter points Cyclops.02 was made with the video frame divided into 240 rectangles each with a zone point at the centre, that reports average grey level for that rectangle.

Zones were numbered this time from bottom left to top right in a zigzag manner.

Table 2 showing zone number configurations in Cyclops.02

240	239	238	237	236	235	234	233	232	231	230	229	228	227	226
. . .														
31	32	33	34	35	36	37	38	39	40	41	. . .			
30	29	28	27	26	25	24	23	22	21	20	19	28	17	16
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Zone data fed through to the noise generator in this version returned a finer resolution over a broader range of frequencies. 240 zones influenced a frequency spectrum divided into 255 bands.

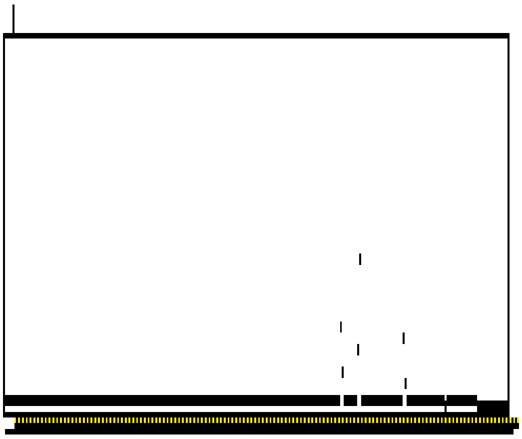


Fig2. 255 band “E.Q” noise filter.

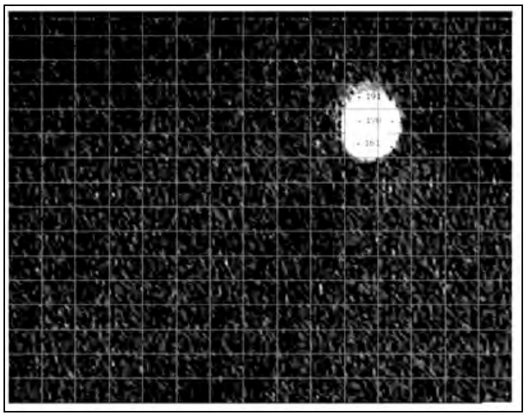


Fig3. Video image of bicycle light.

The two figures above represent two images of the same event generated by the Cyclops.02 patch. The image of a bicycle light in fig3 generates the filtering levels shown in fig2.

Note that the (non-zero) frequency bands are still disjunct. This problem is seemingly insurmountable given the paradigm of numbered zone-data as a method for mapping from image to an audio filtering system. Any array of zones where there are rows as well as columns will produce non-contiguous filter points from contiguous zones. There were other problems with the performance of the Cyclops object and these are dealt with in the appendix to chapter 3.

3.2.1 Implementation using MIDI

The nature of the list output by the Cyclops object necessitates a number of manipulations in order to tease out the various components. Each list contains the addresses and the values for all zones in a single line of data separated by commas after each zone value. In order to unpack the list, unlist and Lfilt objects by Peter Elsea were used in the new patches: Cyclops.04, .041 and .042. These objects enabled me to more easily transpose the data for MIDI processing though, as I discovered, I would have to also make a new grid with only 128 zones to scale to the available MIDI note numbers. (Each zone has to be placed in the grid individually so this can be a tedious process when dealing with a large number of zones).

In these patches, I used a simple piano sound and mapped the zone data to the chromatic scale (all pitches/MIDI note numbers 0 to 127). I encountered two major problems in this exercise. Each list representing the sum of all zone data, and there is one every 100 milliseconds in these examples, had to be converted to a format that represented a chord (pitch aggregate). Elsea’s unlist object worked for this problem. So as to reduce the amount of data per unit of time, zones returning a value of 0 needed to be filtered out otherwise there would be a pitch aggregate of 240 notes every 10th of a second - too much for MIDI to handle – with most notes at an amplitude of 0 anyway. Lfilt object resolved this problem in part.

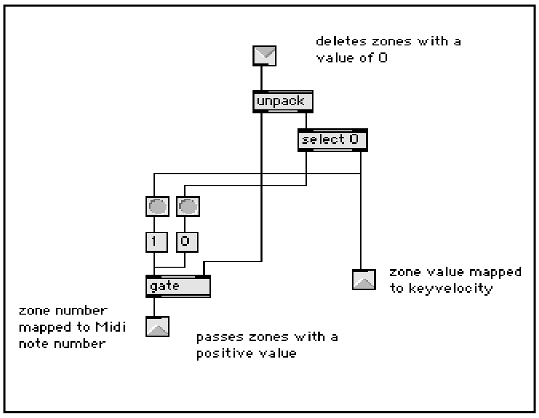
I found that with the video input I was using, which is of a lights moving in a darkened room passing across the camera's field of vision, the maximum number of zones that returned a non-zero value was about 20, so the maximum pitch aggregate would be 20. This would have implications for the patch variants of Cyclops.04.

Once the data had been filtered, I passed this straight on to the makenote object with key-velocity (volume) fixed at a moderate level and note duration of 200 milliseconds.

What I found was that sound and image did not correlate very well in terms of a one-to-one relationship. Two things were happening. Because of the zone configuration, the difference in pitch between vertically adjacent zones could be quite great. (If you take a central zone number – say 144, vertically adjacent zones are numbers 127 and 157: converted to pitch this represents intervals of over an octave.)

Secondly, even when there was a black screen a repeated low note was struck. This may have been due to some artefact of the patch or the limitations of computer clocking. The unrelated ness may also have been due to the possibility that zone number and zone value became scrambled.

Cyclops.042 fixes some of these problems in this way:



Each element of the zone data list is separated into its individual parts using the unpack object. Zones returning a zero value cause the gate to close and are thus filtered out. Zones with a non-zero value are passed. Zone number and zone value are passed out the sub-patch and mapped directly to MIDI note number and key-velocity directly.

fig6 sub-patch of Cyclops.042

The result is a much more successful mapping of the data to MIDI: light intensity directly affects the amplitude of the sound and the confusion between zone number and zone value that was a feature of Cyclops.04 is removed. Scaling of the 240 zones to 127 MIDI note numbers was simply achieved by dividing all zone numbers by 2 and rounding to whole integers. This effectively reduces the number of zones by half.

Issues to address in subsequent patches that use Cyclops are; 1. reducing the attack rate by deleting repeated notes and substituting a sustained note where zone data returns the same value from one list to the next, 2. slowing the rate of zone data output – this will necessarily give a “coarser grain”, but may provide more operating power that can be dedicated to a smoother interpolation between lists and reducing the number of zones to 127.

3.2.2 Implementation using *jitter*

In 2003, the studio acquired Jitter, a program that is fully integrated with *MAX/MSP* and supported by the makers and distributors, *Cycling’74*. This program, which includes a large number of objects for manipulating graphics, makes Cyclops redundant unless the numbered array of zones is a desirable method of working.

Jitter colour tracking

The primary difference between motion tracking in *Jitter* and Cyclops is that in *Jitter* the data describes the rectangular boundary of the body of (coloured) light that is tracked. The rectangle is described by a pair of xy co-ordinates giving the top left and bottom right of the bounded region of pixels. One can then choose which part of the object is to be tracked. It is relatively easy to extrapolate a central point within the region to track or simply use the individual elements of top, left, bottom and right coordinates to control aspects of sound. It is also significant when tracking in a hypothetical Z-axis. The total area of the region can be used as a rough indicator of position along the Z-axis providing the object being tracked is a uniform size refracting light in a uniform manner.

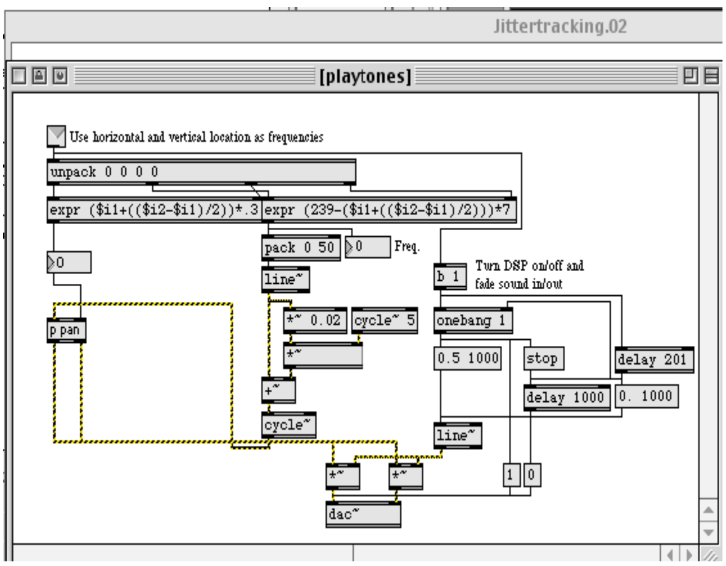


Fig. 7 horizontal and vertical locations used to determine pitch of sine tones

The patch Jittertracking.01 uses the bicycle light movie just as the Cyclops examples above. The horizontal and vertical locations are converted to sine-tone frequencies in a similar fashion to the *Jitter* tutorial on colour tracking[2].

The centre-horizontal and the centre-vertical positions are used to determine the pitches of two sine tones. The result in sound is a smooth portamento of a pair of pure tones in either similar or contrary motion depending on the direction of the moving object.

The patch can also function to convert colour position to MIDI notes – in this case a simple pentatonic scale where extreme left is the lowest note to extreme right – highest (like a keyboard) and volume, or key velocity, is scaled from bottom (softest) to top (loudest).

In this example, as with all others that use either a file movie or a live video feed, the incoming video is altered to emphasize the brightness, contrast and colour saturation necessary to successfully track the desired colours. The issue of stable colour tracking was to become a major one and led to a number of decisions about the kinds of environment most suitable for optimal results. (See below for further discussion).

The example Jittertracking.02 uses the same image and sine tones, but uses the x-axis position as panning data.

Jittertracking.03 uses a movie that has 3 different coloured lights – red, green and yellow. Each colour is associated with a different sampled sound and the xy coordinates determine the playback speed of the sample (pitch and duration) and pan position.

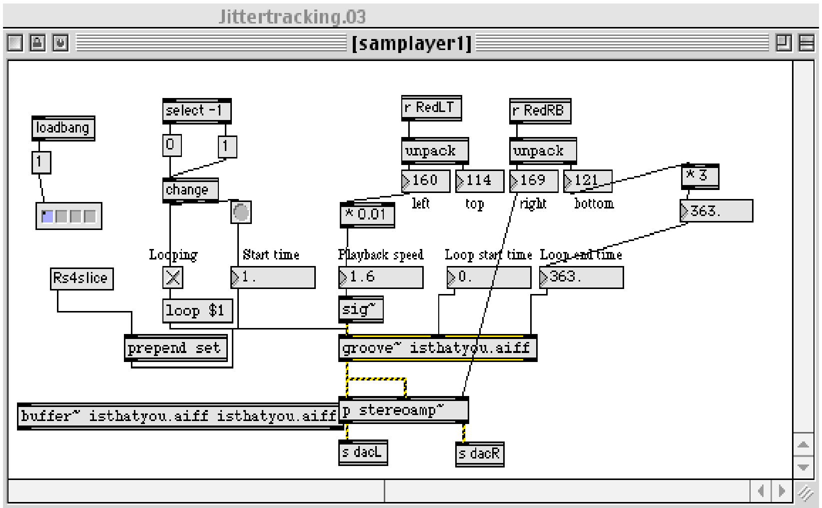


Fig.8 left and right coordinates used to determine playback speed and pan position of a sound sample

Jittertracking.04 & .14 use the movie two_balloons alternating over time with the movie maroon_balloon. The blue coloured balloon in the movie-clip is associated with the higher pitched voice singing a(n altered) line from the song "Blue Moon" and the Maroon coloured balloon in the clip affects the lower pitched voice. Again, the xy coordinates of each colour determine pan position and playback speed (pitch) respectively.

Jittertracking.15 again uses the coloured bicycle light movie and samples as in Jittertracking.03 and Jittertracking.05, but xy coordinates are used to modify a wave-table shift rate, affecting timbre, as well as playback speed and panning. Finding a useable scale to correlate an object's positional attributes to the combination of wave-table shift amount and wave-table shift rate is not easy. The shift rate affects the wave-table differently at different shift amounts because it adds

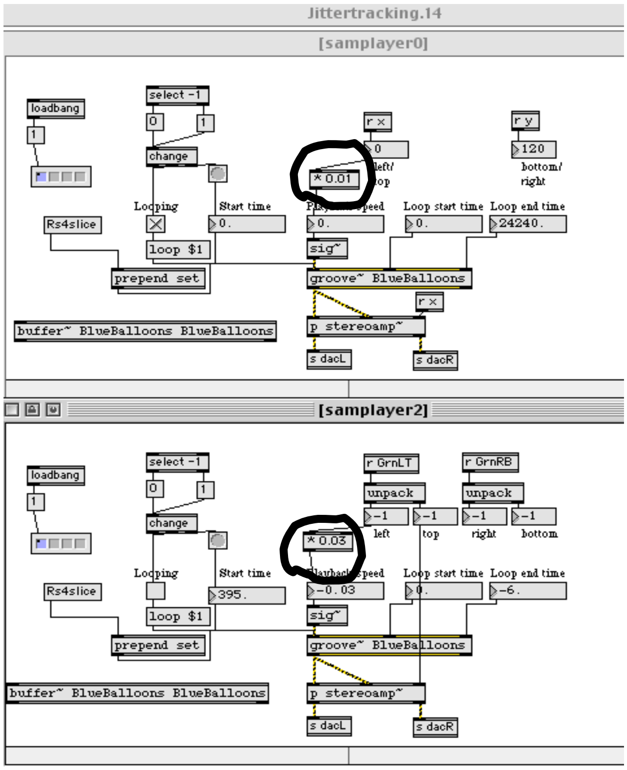


fig.9 a simple multiplication to achieve higher and lower pitched voices for the same sample.

> a sinusoidal position change to the wavetable. As the cosine wave rises and falls, the start and end times of the wavetable increase and decrease. As a result, the wavetable is constantly shifting its position in the [wavetable's] **buffer~**, in a sinusoidally varying manner. Sonically this produces a unique sort of vibrato, not of fundamental frequency but of timbre. The wavetable length and the rate at which it is being read stay the same, but the wavetable's contents are continually changing[3].

Jittertracking.09

Jittertracking.09 uses the movie ChrissiePLF.mov looping back and forth. ChrissiePLF.mov is a short movie-clip of captured *Life forms* animation. The figure in the clip is used to modulate white noise in a similar fashion to the patch Cyclops.02. Though three colours are tracked, only one has been used to modulate the white noise. Horizontal position is used to determine frequency and vertical position modulates the amplitude of the sound. Frequency is polled every 5 seconds resulting in a broadening band of frequencies sounding over this period as the colour moves from left to right or right to left on the screen.

Frequencies are then all returned to 0 amplitude ("cleared") and the next 5 second polling period begins.
This strategy might have been more successful if the polling periods and clearing periods were staggered so that at any one time, after an initial polling period, there is always a 5 second polling period sounding.

Jittertracking.10

Jittertracking.10 is similar to Jittertracking.05, but uses different samples, so that it is somewhat easier to follow the correlation between the motion of colour on screen and the individual sounds. The problem with sample usage in this example, as with others, is that unintentional clicks and glitches occur when visuals are disjunct. That is; when a colour suddenly disappears or jumps from one side of the screen to another, the corresponding change in sample frequency or cessation results in clicks and glitches that are less acceptable to the ear than corresponding visual disjunction for the eye.
In the Jitter examples above, the original conception of an analogous space-sound correlation was modified in order to explore the more immediate possibilities that the software suggested: that is, using a single sound object, like a sine tone or a sample, as a sound object to be manipulated by another object in space. The original conceit is not abandoned, but other paradigms are explored for their potential to yield techniques that may be applied to the original idea.
In the following examples, some of the techniques described above are used, modified and refined in the context of a more performative mode. Instead of using fixed movie files, a live video feed system is implemented. As hinted above, this presented new challenges in terms of controlling light in the environment and modifying the input in order to optimise the colour tracking function of the software.

3.3 Live Video Colour Tracking

JitterLIVEtracking.03

This patch replaces the jit.qt.movie object with the jit.qt.grab object that deals with live video input as opposed to a movie file. In this example, the colour yellow is tracked and the xy coordinates determine the centre-frequency and the bandwidth of a noise generator. The (MAX/MSP) object by Timothy Place, the tap.allpole~ object, was used as a noise filter.

The result is similar to the Cyclops patches, but without the errors associated with mapping zones to filter points. The scaling of positional attributes to audibly workable sound parameters is still some way off at this point in experimentation. The difficulty with this patch is that smaller bandwidths also reduce the amplitude and in combination with a centre-frequency at a lower pitch, results in a barely audible noise. Positional changes in the centre-frequency, where there is a wider bandwidth, tend to be obscured because the wider and louder noise dominates over small internal changes. Non-linear and sinusoidal scaling is implemented in later examples, and this assists with the above problem of perceiving smaller changes. The terms linear and non-linear here refer simply to lines being either straight or not straight. 'Not straight' (non-linear) may be curved, disjunct sequences of straight lines or disjunct sequences of curved lines. At other times, these terms may describe higher-level structural techniques and have a different meaning: i.e. referring to conjunct or disjunct progression from one thing to another, as in a web-based or hypertext situation.

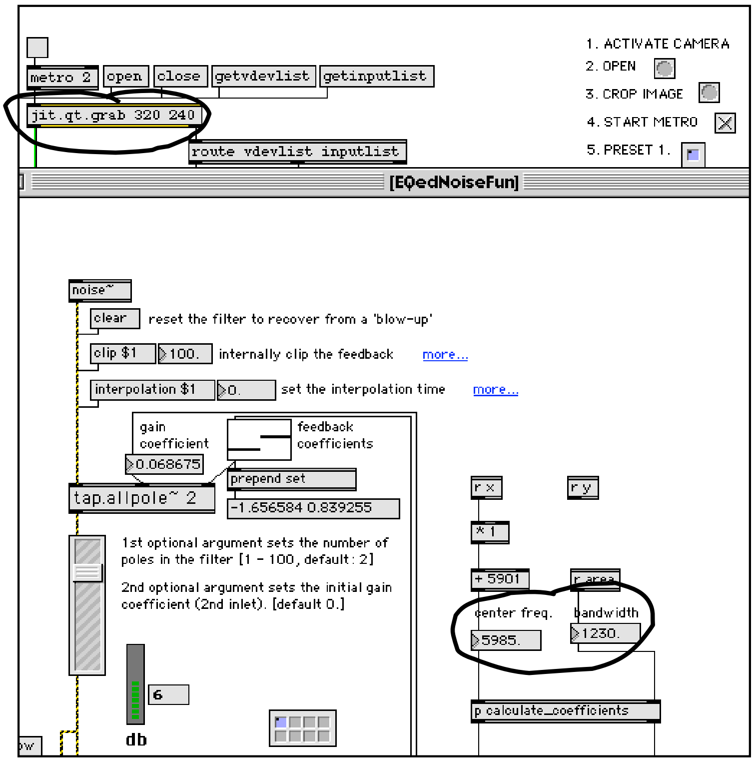


Fig.10 jit.qt.grab object and noise filter sub-patch

JitterLIVEtracking.07

This example implements a similar system as that used in Jittertracking.03, but for a live video situation. Xy coordinates determine the playback speed and pan position of a sound sample.

JitterLIVEtracking.08

Similar to a live version of Jittertracking.15, JitterLIVEtracking.08 correlates frequency and pan position to vertical and horizontal coordinates, but some variables that are not controlled by colour motion are varied step-wise using presets over time. The 3 samples (affected by either the colour red, green or purple) are varied at different rates so that they change out of phase with each other. When all samples are present (sounding) all the time, the tendency is that dissociation of sight from sound occurs. Re-association is enabled when just one colour moves while the others are stationary or when just one colour is present and affecting sound output.

TrioStudy.01

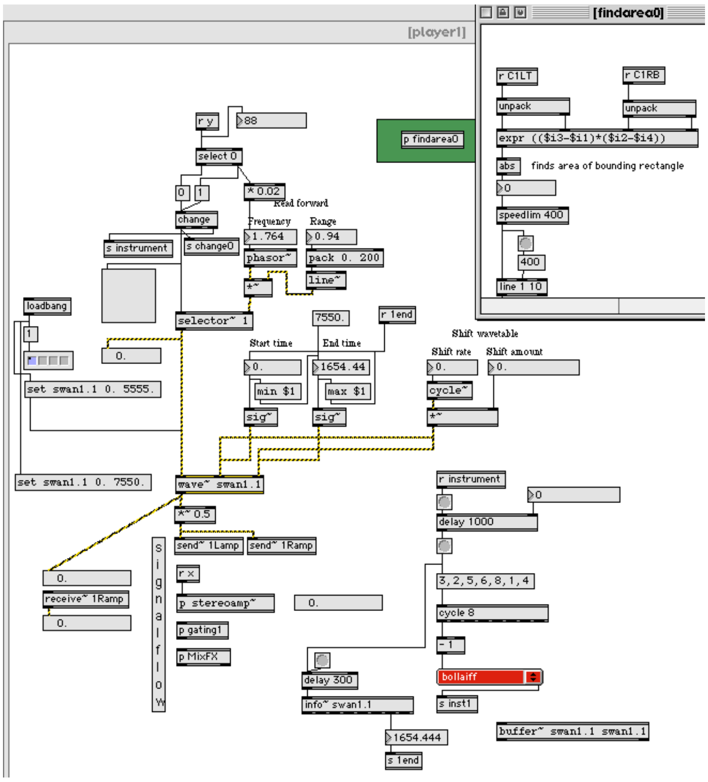


Fig.11 Sub-patch of TrioStudy.13 showing sample treatment of incoming video data

This patch tracks three colours, each of which is associated with a bank of samples that rotate over time. The spatial parameters of area, central height position and central vertical position are mapped to reverb mix, frequency and pan position. In addition, central height position and central vertical position are also mapped to the modulation frequency and modulation depth of reverberation. Area is also mapped to reverb decay time. Although the overall sound of this patch is somewhat more interesting than many of the previous examples, there is a tendency to dissociate image from sound in terms of a direct correlation. In spite of various attempts to avoid sample clipping, the problem remains. The general approach has been to use line objects to create envelops with a smooth decay. However, due to such variables as playback speed and sample end time being in constant motion, it is virtually impossible to consistently and correctly calculate the onset time of a line object in order for it to function optimally as an envelop preventing the clipping phenomenon.

Other problems arising from this patch were errors in audio processing where the right channel output often 'crashed' and the error message 'not a number' prevented further processing in that channel. This phenomena seems to be peculiar to patches using the tap.verb~ object.

Several more versions of this patch were attempted so that by version 5, TrioStudy.13, many of the errors and clipping were minimised. Unfortunately the samples used in this example were so obnoxious that it was difficult to listen to them after a short while.

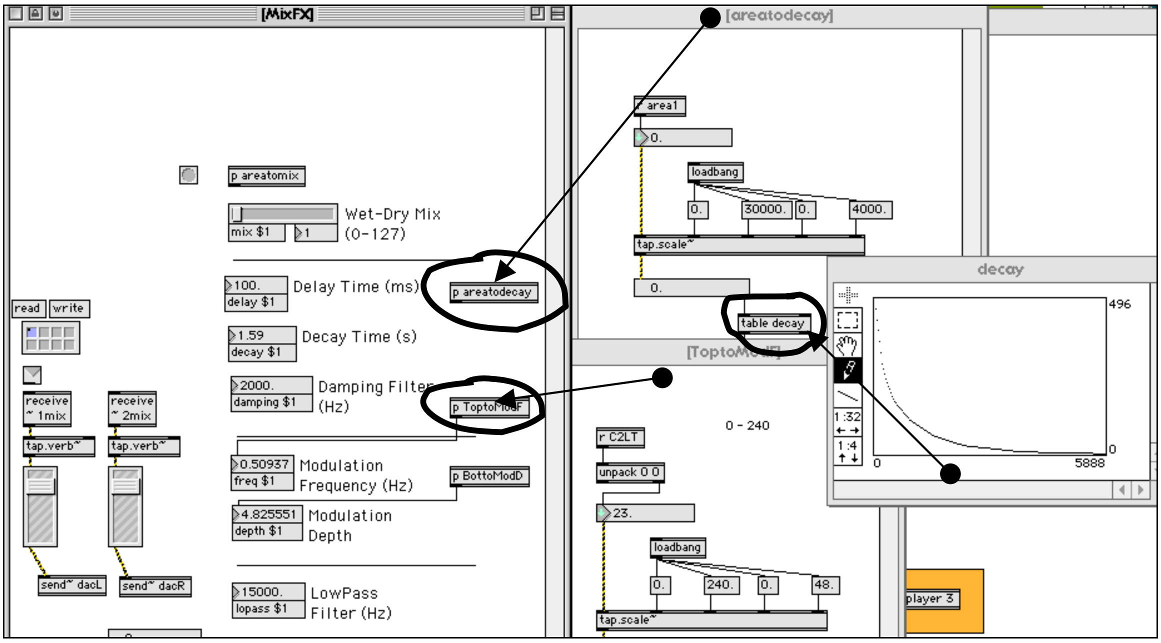


Fig.12 Sub-patches of TrioStudy.13 showing sample reverb treatment of incoming video data

3.4 Initial conclusions from the experiments.

The above examples seem to sustain some theatrical interest, but do not sustain musical/sonic interest for very long.

Even small degrees of complexity tend to dissociate or de-link sound from image in some circumstances. This is not necessarily a bad thing, but raises the issue of performer control. If a performer cannot perceive the effect that they are having at any given moment, they are unable to operate in the traditional performer mode that relies on instant feedback and adaptation of their physical movement to produce a desired outcome. Therefore some other performance paradigm must operate as soon as correlation becomes unclear. However, the boundary area of ambiguity is elastic. Depending on the level of skill and knowledge of the performer, direct correlation can be heard at a high degree of complexity. In this sense then, (level of skill and knowledge) these systems are instrument types as much as any other classical acoustic instrument, requiring the development of knowledge and physical dexterity peculiar to the system. (See above – in chapter 1.3 – for further discussion). What makes the system quite different from a classical instrument type is that each system is geared to play a very limited repertoire – in most cases – only one piece.

In terms of the effect on an audience, this dissociation or ambiguity of correlation provides a rich ground of opportunity for play, poetry, disturbance and interpretation to name a few areas in the relationship between the work and its audience. I liken this movement around concreteness and ambiguity to the experience of listening to a fugue, or similarly complex piece of music where, following the statement of the subject, a listener may easily trace subsequent iterations up to a certain level of polyphonic density whereafter tracing increasingly fractured elements of the fugue's subjects becomes irrelevant to the appreciation of the piece on another level. Having understood the connectedness of the material, the listener is asked to move beyond this kind of understanding to another level of knowing, where connectedness is taken as a given, perhaps, and the music's aesthetic, poetic, spiritual or other dimensions become the location for the listeners' appreciation.

3.5 Monody for Coloured Objects

Monody for Coloured Objects is the first attempt at a real work using techniques and methods explored in the experiments above. It grew out of the Jitter software learning process and as such represents a stage in learning that is perhaps equivalent to an intermediate level skill on an instrument. The piece is about eight minutes long and is for a single performer. The equipment requirements are modest. Just one DV camera is used for colour tracking and one MacG4 computer running MAX/MSP/Jitter and of course sound outputs to either two or four channels.

Fig.13 image from a performance of *Monody for coloured objects*.

The performer is a kind of puppeteer who manipulates various objects on and around a small table. One unusual aspect of the piece is that the tracking follows only the colours of the manipulated objects and not the human movement at all, but what the audience sees is the ‘puppeteer’ character as the primary focus and the objects as almost incidental to the character’s silent meanderings on stage.

The overall sense of this character is of someone isolated or alone who is involved in mundane almost meaningless tasks. He is incessantly checking the time on a small, solid brass clock (the primary object that is tracked). He gets up and makes tea (a red mug is a secondary tracked object). He contemplates eating a golden pear. Objects are re-arranged on and around a small red box on the table. He puts salt and pepper from S & P shakers (tertiary objects) on the pear.

What happens in the sight to sound system (Patch file name Monody.003) is that a single sample is triggered and manipulated by the presence and position of the brass clock. If the clock disappears from camera view the sound stops. If it reappears, another new sample comes into play. The further the clock is away from the camera, the more reverberation affects the sample. Essentially what is going on here is that the smaller the total area the colour takes up on screen, the greater the reverb, so when the golden pear is introduced into the scene there is more or less reverb depending on how close the pear is to the clock. The height of the pear/clock affects pitch and the left or rightness affects the pan position of the sound. The red mug and red box have a supplementary affect on the pitch of the pre-sampled sound. The smaller the bounding area of this colour on screen, the greater the lowering of pitch, so the further the mug and box are apart, the less affect they have unless one or both of them are very far from the camera. The blue S & P shakers affect a type of pitch modulation of the sample where a colour’s bounding area determines the pitch depth and left or righted-ness affects the rate of modulation. With this fairly simple set of correlates, I found that quite complex arrays of sound possibilities were represented in the system. Simply by the presence or absence of objects alone it was possible to build varied and interesting manipulations of sound samples, so the simple task of picking up a clock whilst holding a mug takes on a whole new sonic meaning when the relationships embedded in the system come into play. The sonification of object relations is used here to heighten the theatrical intent, which is to foreground the significance of the human-object connection. Exactly what that significance is, of course will vary from one viewer to another.

Depending on how the performer structures the performance determines how much, when or whether the sight to sound correlations become clear to the audience: at least to a large degree, because it is quite difficult when all the objects are in play to determine how each object is affecting which part of the sound. The traditional standpoint of one-to-one mapping relationships being aesthetically unpromising seems not to hold true given the addition of the camera’s sensitivities, the system’s artefacts, and the interplay of light in the environment.

In the above work, the difference in viewpoint of the camera to the viewpoint of the audience is another take on "human/machine relations" that extends my work in this field to include the manner of performance reception as well as systems for performance delivery[4].

On the one hand, as humans, generally we are more interested in what happens in performance with and to other humans and this is an entirely natural position, but on the other hand perhaps it is equally "natural" for machines and software to be more "interested" or more "receptive" to the non-human or the abstract (like a particular colour or shape) – a gold clock or a red box? While an audience is projecting themselves into the human in performance – a like-to-like relationship – they are hearing what the machine has to "say" about something that is quite abstract – an un-like relationship. However, as I suggested above, the psychology of the piece implies a like-to-*un*-like connection. So in one respect this gives prominence to the polymorphous-perverse response to the stress of the character's isolation.

3.5 Dismembered

Dismembered, the second work to implement and extend aspects of the experimental patches is a dance and puppetry piece that plays with the notion that the human and the abstracted are potentially illusory standpoints. The dancer's body is deconstructed and dispersed into abstract components. A leg-shape, a face, an arm, torso and a hand might come together in the usual way – a whole body – and dance a human dance. At other times they fly off and become "source files" that are "dispersed through the generative syntax of software", the focus shifting between the body and the constructed body, the body and the deconstructed body and the sound becomes the projection of a mediated body captured in performance and modulated and re-processed and re-transmitted via the camera, the computer and the speakers[5].

What happens on stage, which is lit with ultra violet light, is that the dancer has U.V. sensitive material attached to her limbs. The materials are sometimes moved and manipulated by the limbs in question and at other times they are detached from the dancer and manipulated by the dancer or the puppeteer. Various body parts are seen to "fly off" into space. They seem to have the ability to live separate lives from one another.

What happens in the system is similar to the *Monody* piece above, but visual data is bussed to different maps at different times, affecting multiple sounds in different ways simultaneously. The presence or absence of up to four different colours (associated with different body parts) re-maps the data flow from the visual input. Depending on which map is active, colours manipulate sounds in a primary, secondary or tertiary way similar to the way in which one might have an oscillator in a synthesiser being heard 'as is' or as a controller of another oscillator or as a controller of another effect (reverb for example).

In this system the performers have sixteen possible states to explore given all the permutations of presence or absence of four colours. The dancer and the puppeteers collaborate to invent their own narrative and movements within the general framework of the system and the themes of alienation and integration.

Fig.14 Image from a performance of *Dismembered*.

The use of ultraviolet light in this piece was devised to address the problems of colour tracking in natural light that was encountered in the *Monody* piece and in many of the experimental etudes. In theory at least, the U.V sensitive materials would reflect solid colours within a narrow colour spectrum thus eliminating cross talk. Colour tracking cross talk occurs when the bandwidth of alpha, red, green and blue in one colour overlaps the bandwidth of another colour. Cross talk can result in wild fluctuations in the boundary rectangle of a tracked colour. This translates in a number of ways to the sound depending on which method, either sampling or synthesis, is used. With sampling, the likelihood of clicks and glitches is multiplied ten-fold, especially if spatial parameters are mapped to loop start or end points. Similarly with a synthesis method, where spatial parameters are mapped to duration and/or attack onset, noisy envelope clipping will occur.

U.V sensitive materials under U.V light were successful in reducing cross talk, but not in eliminating it entirely. In addition; regardless of whether one, two or three colours were being tracked, boundary data tends to jump radically as portions of a three dimensional object are obscured from direct light, i.e. in partial or full shadow. While the human eye will compensate as an object passes in and out of shadow (the object is still recognized as having a uniform size and shape) a computer tracking a bandwidth of colour will not compensate in this way.

In constructing the performance for *Dismembered*, part of the role of director was to optimise the choreography for maximum reflectivity. Performers were asked to situate materials so that they faced the light as much as possible. Much of the detail of the choreography was left up to the performers, but the director set the overall aesthetic, primarily slow, smooth movements. Workshop and rehearsal collaboration between director and performers resulted in the set sequence of events, essentially a process of adding one limb after another to arrive at the semblance of a human form, followed by a more animated dismembering.

3.6 Killer

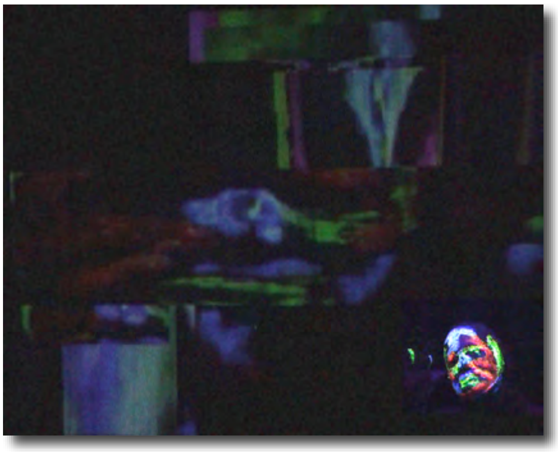


Fig.15 Live and projected image from a performance of *Killer*.

While *Monody* and *Dismembered* use sample manipulation and FM synthesis respectively, *Killer* maps video data to a virtual additive synthesizer. This work is for one performer only. The performer is placed in a puppet theatre and dressed in black, including a black latex mask for the face so that they are virtually invisible until, under U.V. light, they begin to paint their face using U.V. sensitive paints. The application and manipulation of the paint on the performer's face manipulates the sound. A complementary visual element amplifies the live image of a face that is confined to a space no bigger than a T.V monitor. The captured live image passes through the system and is itself manipulated before being projected onto a larger screen beside the performer. In this instance, an operator improvises using a number of pre-set algorithms that stretch and rotate the image in a variety of ways.

The performance ends when most of the face is covered in paint and the performer has explored the confines of the space by moving their head to different locations and revealing different profiles to the camera and audience.

Correlation of sound to image in this work is fairly stable. There is always a sense in which the movement of colour is perceived to control sound despite a high degree of complexity in the audio output resulting in part from a high degree of jitter. The piece differs from *Dismembered* in that only one map is active in the work. The presence or absence of one or another colour does not alter the way in which a tracked colour affects the sound. Though the patch itself has many of the objects required to do this, one aim of the work was to explore the limits of a single map in terms of musical and theatrical time.

One aspect of change that is affected by colour permutation is that slight variations in the affected timbre are enabled from one permutation to another. The presets shown in figure 16 recall variations of the shapes of the envelopes of six partials. This contributes to subtle differences from moment to moment in the piece while preserving the overall connectivity between a colour and its associated sound.

The sounds are all bell-like timbres where upper partials are in a mixture of rational and non-rational relationship to the fundamental. Each set of bells, controlled by one colour, are restricted to a frequency band functioning roughly as soprano, alto, tenor and bass. This conforms to the initial idea at least in that each colour-object moving in space has its own sound-object on which it acts as a kind of filter. The analogy is somewhat stretched, but holds true in many of its essential features.

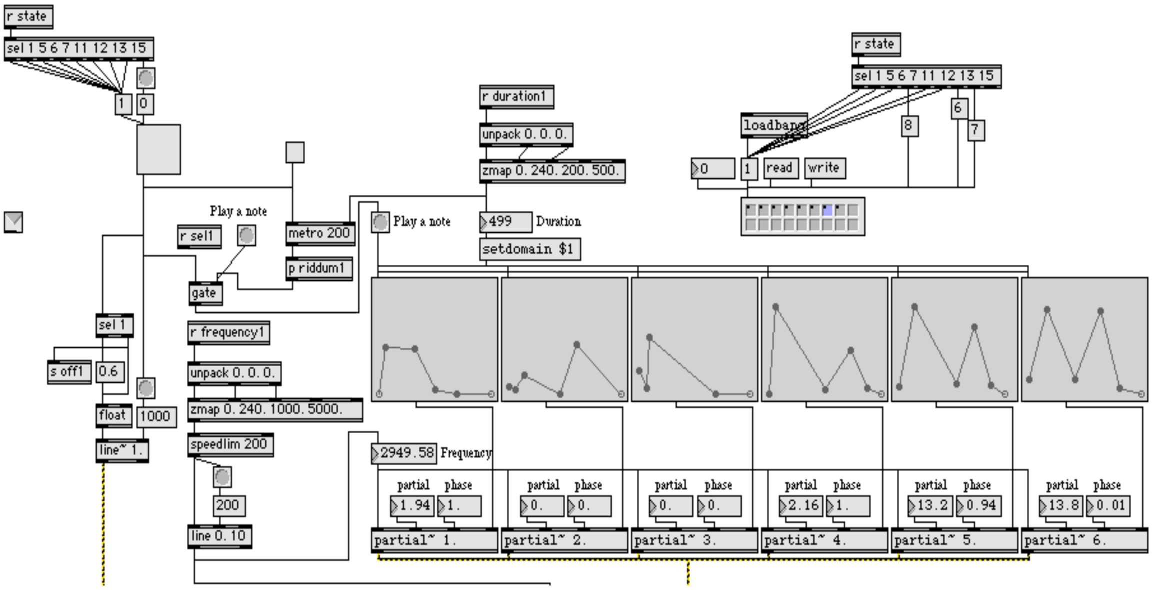


Fig.16 Additive synthesis processing in *Killer*.

3.7 Dis-Patch

Dis-Patch was first performed at the Loft theatre at Queensland University of Technology, Brisbane, in June 2004. It is a 40-minute work for three dancers, live video projection and music. The work was choreographed was by Chrissie Parrott. In this piece I returned to the use of an external MIDI device for sound synthesis and two cameras were used to track colours in the space. In the first section of the work, a solo dancer enters the “black box” space, which is lit only by ultraviolet light, entirely clad in black vinyl and a black latex mask, practically invisible to the audience. Using U.V sensitive paints, he gradually paints his face and body while moving through the space. The two other dancers, hidden from the audience, also participate in the painting process.

The second section of the piece is a duet featuring specially designed costumes by Shaaron Boughen in which various panels of colour are revealed by the dancers. Additional blue lighting increases the overall level of ambient light in this section of the work.

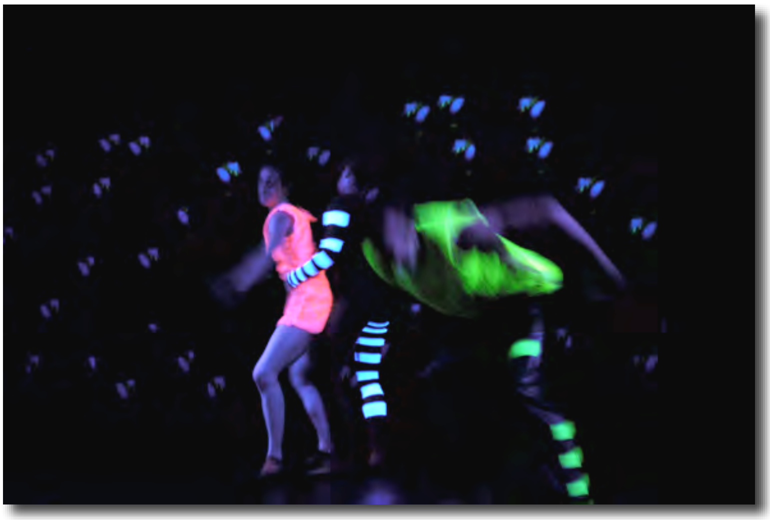


Fig.17 Live and projected image from a performance of *Dis-Patch.*

The third section is a combination of solos, trios and duets in which all the performers change their costumes from time to time. Each costume emphasizes either the Red, Green or Blue spectrum for manipulating sound and image. Again in this section, lighting levels are increased. The dancers' image is reconfigured using the scissors and glue object in *jitter* and a number of other processes by which the relationship between the various colours in the space determines the visual output parameters that are projected onto the back wall of the performance space. The effect is reminiscent of a kaleidoscope.

House lights and work lights are switched on so that chance lighting artefacts in the environment in the final section are just as likely to affect the colour-tracking system as the flesh-tones and rehearsal clothes of the dancers. The system is 'Dis-Patched', which refers also to the programming environment of 'patching' from which the system is created.

3.7.1 Mapping strategies in Dis-Patch

Mapping strategy, that is the mapping of topographical data from video input to sound parameter output, combines both the "one-to-one" and "divergent" or "one-to-many" paradigms[6]. This was a departure from most of the etudes above that use only the one-to-one strategy. In this piece, six output parameters are paired to three inputs and two other outputs are controlled singly by two inputs. This configuration was varied when portamento, as an output parameter, was turned off. Over the course of the work, five separate mappings were used. See table 3 below. One map predominated within each section of the work so that through each section, the audience could become acclimatized to a single mapping style. Audience perception of the precise correlations was not a goal per se, but simply the perception of difference. This difference was always reinforced by other changes in, for example, instrumental timbre, tempo, (pitch-) scale and the presence or absence of a portamento effect. These are referred to as global conditions for each section. See table 4 below.

Table 3. Five permutations of topographical data mapped to sound parameters

Width centre (of a colour)	Height centre	Width	Height	Area	Map #
Pan position	Frequency	Portamento Duration	Polyphony Pitchbend	Velocity Reverb	1
Velocity Reverb	Pan position	Frequency	Portamento Duration	Polyphony Pitchbend	2
Polyphony Pitchbend	Velocity Reverb	Pan position	Frequency	Portamento Duration	3
Portamento Duration	Polyphony Pitchbend	Velocity Reverb	Pan position	Frequency	4
Frequency	Portamento Duration	Polyphony Pitchbend	Velocity Reverb	Pan position	5

It is difficult to make definitive conclusions about which of these particular maps was the more effective. Certainly, the second of the maps above seemed to be the least dynamic and perhaps map one and three were the most dynamic. All of the maps have a useful function depending on the type and level of expressivity that is desirable. The decision as to which map was used at any particular time in the piece was made fairly intuitively, that is, on hearing the effect of each one during the rehearsal process, I decided which map was most appropriate for a section. Appropriateness was usually based on whether the sound had the right level of dynamism for a particular point in the overall structure of the work. Greater levels and rates of change in the sound define dynamism.

Table 4. Examples of types of global conditioning occurring in the different sections of *Dis-Patch*. (not in order of appearance)

Ensemble	Tempo	Scale	Port	Map
metallic percussion and harp	3 moderate	heptatonic	off	1
wooden percussion and pizz	2 quick	Heptatonic	Off	2

Metallic percussion	3 moderate	Heptatonic	On	3
Wooden percussion	1 very fast	Octatonic	Off	4
Reeds and plucked strings	5 very slow	Octatonic	On	5
Brass	5 very slow	Aeolian	Off	2
Strings	4 slow	Octatonic	On	4
Wooden percussion	1 very fast	Heptatonic	Off	1

A second camera and computer was used in this work to control reverb levels in a MIDI controllable effects unit. The output from the synthesizer was bussed directly to the effects unit and incoming MIDI messages from the second computer determined the level of wet to dry signal that was output. This meant that there were two levels of reverberation happening at any one time – or ought to have been as will be seen below.

- The first level was at the level of individual instrumental voices. The data associated with each colour in this system is used to manipulate a single instrumental timbre. There are up to three colours being tracked at any one time so there will be three different levels of reverberation.
- The second level is produced in the second camera-computer unit. In this system only one colour's data set is used to manipulate reverb mix, but this is applied to the sum of all the signals from the synthesizer and acts to integrate the overall sound.

Ideally I would have liked to have had the power of the external effects unit applied to individual voices separately, but the synthesizer (a Roland MOC-1) did not have separate audio outputs for each MIDI channel input and the effects unit would have required another four inputs – all independently controllable. This would have given the system the capability to place each instrumental timbre within its own reverberant space. The inclusion of the first level was an attempt to make this possible, but the reverb available on the synthesizer unit was not either powerful enough or modifiable at the level of character to make differences between the three reverbs easily discernable.

The reverb effect was also distorted in performance by jitter. Irregular jumps from low to high levels of reverb-mix resulted in unwanted glitches from time to time. This was modified in later performances so that the system on the second computer sampled video at a slower rate, while a smoothing algorithm interpolated between captured (non-contiguous) data points. While this introduces additional latency into this part of the system, its perceived effect was minimal as delay is a parameter of reverberation anyway.

Discreet reverb or any other effect for each timbre seems to require a significant expansion of outboard hardware in a MIDI set-up such as this. An audio “breakout box” for the synthesizer might be an inexpensive solution if it is possible in this scenario. Stereo channels for each timbre may then be bussed to the effects unit and to each computer for audio reprocessing in *MAX/MSP* or other digital signal processing (DSP) software.

3.7.2 Camera positioning

The position of cameras for capturing performance on stage is non-trivial for a number of reasons. Firstly, for audience perception of performer manipulation of sound and image it is easier to have at least one camera placed within the body of the audience however distracting this might be for an audience. Additional cameras that support secondary manipulations, like reverb, an overhead position has proven to be the most flexible. Where there is plenty of space to the side of the space, a side-placed camera may work well.

Secondly, the distance from the performance area to the camera needs to be quite large in order to capture the whole area. The camera angle needs to be carefully arranged so that the axis of tracked co-ordinates makes sense in terms of horizontal, vertical and depth planes. Too steep an angle will produce a distorted sense of space at odds with our sense of gravity. This might be a desirable effect in some circumstances.

Generally a larger auditorium is better suited to video tracking than spaces that have either low ceilings or a smaller depth to them. The Loft space at QUT with dimensions of 15m x 20m x 15 met most of the criteria for the performance of *Dis-Patch*. The raked seating necessitated placement of the primary tracking camera in the ‘bio-box’, fairly high in relation to the performance area, but not so high as to produce a distorted view and high enough so that the back wall projections could be eliminated from the camera lens. Even so, some cropping of the incoming picture was required to eliminate walls on both sides and the heads in the audience. The projector also needed to be tilted so that the image was placed quite high on the back wall, just above head-height. A lower image placement results in feedback through the system.

3.7.3 Environmental control and lighting

It is virtually impossible for three dancers to operate effectively under only ultraviolet light for forty minutes. Additional lighting is necessary for their sense of balance and orientation in relation to each other, the space and stage exits. Between six and ten minutes was the maximum time spent under U.V alone. Light levels increased throughout the rest of the performance, leading to full house and work lights at the end. This proved to be very challenging for the accurate calibration of colour tracking. Altogether five separate

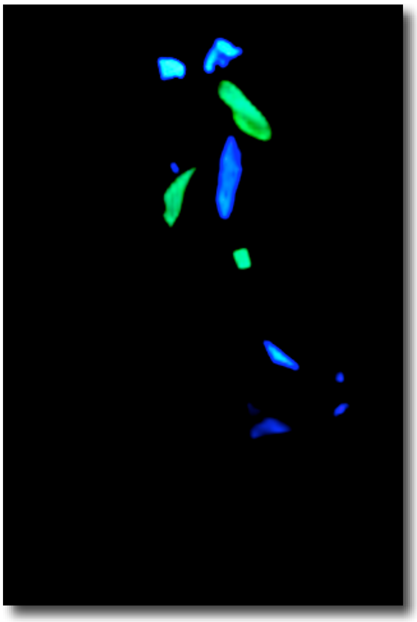


Fig.18 Image from section 2 of *Dis-Patch*.

calibration states had to be programmed for each main section and one transitional state for the piece.

The second section of the piece was perhaps the most challenging in this regard. ‘Surprise’ blue lights were added to the U.V. illumination and this altered the way in which colours were read by the camera and, indeed, how they were seen by the naked eye. The costuming for this section was perhaps the most impressive of all the costumes, consisting of full length black wool dresses that had a number of panels, zips and pleats that could be manipulated to reveal coloured sun-ray pleated silks. Unfortunately not all the colours turned out to be as fluorescent as they first appeared, which made them very difficult, in the low light, to track.

3.8 Initial conclusions from artistic practice

Despite some of the reservations expressed above about failings in various systems, all of the works achieved many of the goals set for them.

Dis-Patch in particular, achieved a significant number of goals and these are outlined in detail at the end of Chapter 4. Like other works that use a dancer’s movement to manipulate sound, the traditional relationship between music and dance is radically altered. The music tended to follow the dance, but through the filter of the composer’s system and the dancers responded in a continuous feedback loop to their own effect on the sound. Their response, though, was dance focused, rather than sound focussed. Their response was not an attempt to manipulate the sound, merely to enhance or modify the quality of the dance. They did not have to feel responsible for the sound in the way that some systems require a dancer to also be a kind of musician controlling a virtual instrument.

The dance did not always dominate the foreground of the visual field either, but rather, the combination of costume design, videography and dance were experienced as an interplay of events that often altered their positions in the visual depth of field – from foreground to middle-ground to background.

Chapter 3 References

1. Mann and Fletcher, para. 28, 30, 32; emphasis in quoted university documents is original. Cited in Milech, B & Schilo, A. (2004). "'Exit Jesus': Relating the Exegesis and Creative/Production Components of a Research Thesis." TEXT Special Issue No 3(April 2004). p. 13.

2. Singer, Eric. (2001). Cyclops. Vers. 1. Computer software. Eric Singer. Mac, plug-in for Max-Msp. p. 1.

3. Dobrian, C. (2003). MSP 45 Tutorials and Topics. California, Cycling '74. p. 114.

4. Vickery, Lindsay. The Robosax Project. Australiasion Computer Music conference. Melbourne: ACMC, 2002. p. 113.

5. Birringer, J. Suite Fantastique: Architecture and the Digital. 2002. Location One Gallery. http://www.location1.org/birringer_suite_fant2.html2002. Accessed 10.02.02.

6. Rovan, Wanderley, Dubnov & Depalle. Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance. Genova: The Technology of Emotion Workshop, 1997. p. 171.

Chapter 4

Different Lenses: Reflections from Installation, Instrumental and Dance Interactive Work.

"The computer program which interprets motion and translates it into sound makes perfect 'sense' at the foundation level. That level of sense is, however, not completely accessible through the experience of the piece. It is obscured by its own complexity. Any attempt to solely rationally understand the work is automatically doomed to failure." David Rokeby: Predicting the Weather[ix].

In this chapter I will explore some of the issues that appear to delineate cultural difference in three categories of interactive work in which technology is the mediator between human movement and sound in the work. In doing so I hope to clarify where in the spectrum my work, discussed in chapter 3, is placed, its relationship to other similarly made work and its relationship to work in other categories.

As mentioned in chapter 1, there has been enormous growth in the literature on alternative gestural controllers that reflects the parallel growth in the number and type of controllers being utilized for manipulating sound. The primary concern, in these musically oriented writings, has centred on the utility of the interface under discussion. In dance-interactive pieces, where there is more often than not, an interactive visual element as well, the concern has been focused on broader questions of when, why and what does the interactivity say about the body/human condition/status quo?

Installation art and its associated literature focus largely on aspects of culture, for example, the interplay of cultural modalities: learning, exploration, play and reflection or the relationships between human and machine systems. In addition, there is often a strong concern in installation artwork that addresses the question of how to engage an audience directly in an experience of the work – as active participants, who by their actions and reactions in the installation complete the work.

4.1 Installation

The quotation by David Rokeby, above, presents a common conundrum that practitioners of interactive work across all fields, dance, music and installation art, experience in the making and reception of their work. On the one hand, the artist has created a structure, often utilising a particular mapping procedure, that has its own internal logic and rationale, but an audience's experience of the piece does not necessarily recognise this procedure, logic or rationale in its entirety. In fact, procedures are often intentionally obscured in order to create a certain amount of mystery, intrigue or magic: in short – artifice. This artifice might be related to the subject matter of the work or to the mode of information transmission, for example, and at some point, conventional 'rationality' must be dispensed with in order to really understand the piece.

This 'dispensing with rationality in order to understand' sounds like a contradiction, but what Rokeby is getting at is a rationality that we all too often equate with predictability. Rokeby says that we feel cheated if an input stimulus to a system is repeated in what we perceive to be an exact replica, and the result from the system is anything but a slavish repetition of its initial response[2]. But this feeling, he says, is based on the misconception that reality ought to conform to some unalterable law of physics reduced to a lowest common denominator. Rokeby asks us to 'explore to discover' rather than 'exploring to confirm' so that questions of unpredictability, confusion and chaos disappear. They disappear in the sense that rather than being negative indicators in a proscribed reality they become positive indicators of the participant's relationship to the system, reinforcing "the hidden physiological and psychological resonances of the body and mind in physical interaction"[3].

Rokeby asks also that we act responsibly in our unknowing. He asks that we accept our inadequacy to control complex situations and that the universe will find its own equilibrium[4].

The difficulty with the assertion that the participant's physiological and psychological resonances are somehow exposed to them by this system is that it opens the way for almost any random or chaotic mapping of movement to sound correlates to make the same claim. So long as the results of a mapping are unpredictable, its internal logic, no matter how cogent, is irrelevant, so why spend the time making a logical set of correlates?

The answer may lie in the degrees of predictability and chaos that a system displays and the amount of time it takes for humans to perceive various levels of order. A logical set of correlates is more likely to set limits to control the level of chaos in some way so that surfaces may seem chaotic, but the variables are within predictable limits. The broader the limits, the longer it may take us to perceive what the system's behaviours are.

Another major theme that permeates much of Rokeby's work is the notion that interactive art is about how we reflect upon ourselves and how the medium refracts what is given to it and what is given back to us as an audience/participant. In many of Rokeby's pieces, the audience is required to engage physically with the work in order to activate the work. This requirement means that the participants must firstly see themselves as some kind of performer and therefore closer to authorship, which is a very different relationship to most traditional artwork. Thus the medium already presents the message “You (the audience) are implicated in the work’s outcome” and that “this outcome is a reflection of you, albeit refracted by another author – the author of the system in question[5].”

This refracted, or distorted reflection, tends to compound the sense of diminished control that the interactor feels in an experience of the work, which is quite different to the kind of control engendered in video game technology, or in many visually-based interactive works. The sense of empowerment in these other media is often really a false empowerment, because of the limited set of choices available to the interactor. Rokeby’s open-ended structures in his *Very Nervous System* pieces enable a wide range of choices, partly because there is no particular goal set for the interactor and partly because feedback is within the domain of sound. In the domain of sound, the interactor does not anticipate crashing into a (virtual) wall, running off the road, or making a wrong turn up a cul de sac. The dominant terrain of sound is time, not space, which is precisely the element that the interactor has most subjective control over in these works.

The sonic and visual outcomes of the works are almost secondary. These outcomes might be aesthetically worthy, but the primary interest lies in the composition of the relationship between the system and the participant. This composition determines their worthiness as art works.

Garth Paine’s sound installations, in particular *Map 1* and *Map 2*[6], are closer in conception to a musical instrument paradigm, albeit a virtual one. Paine uses Rokeby’s *Very Nervous System* as the “eyes” of the system, enabling him to follow human movement in the installation space. In *Map 1,* the participants’ sounds within the space are also captured above a certain threshold of volume, then reconfigured and dispersed via a multi-channel speaker system.

The participants’ positions and quality of movement in the space affect the usual sound parameters of pitch, timbre, volume, duration and density, etc. Visual cues, such as text placed on the walls, give the participant clues as to the type of parameter likely to be affected by being in a certain place, but generally they are expected to explore and develop a cognitive map of the space according to the system’s response to their behaviour in the space. In *Map 2,* ensemble playing is cleverly encouraged by the system’s partitioning of timbre and participant. A particular instrumental timbre has the tendency to “stick” to each performer in the space, allowing them to identify their movement with that sound and improvise with others in the environment without the confusion of wondering whose movement is causing a given result. This does not mean that any given movement will have an entirely predictable outcome, similar to Rokeby’s standpoint, but that the outcome at least is confined to a single class of sounds.

Both Rokeby and Paine's work, for the most part, require the participant to explore in a playful manner the system, sounds and spaces that have been designed for them. In Paine's work the players are more able to move towards a playfulness that has intentionality, that is, the player-participant can feel more able to manipulate specific aspects of the soundscape in order to execute a desired outcome. Rokeby's work, because it tends toward less predictable outcomes, resists participants' specific desires, but perhaps offers them an alternative that can satisfy other artistic urges: surprise, mystery, disturbance and a sense of loss of control that may be felt to be positive.

A point of interest, or perhaps a point of principle is that movement in these installations is never dancerly. When the quality of movement rises above the mundane, it is almost always in the sense that it has the quality of a conductor. The design of a system that encourages directorship of a sonic 'other' through whole-body movement is at least an interesting strategy to deal with some people's inhibition about being seen to 'dance' and it seems to work well in many installations. I have yet to find an installation that consistently elicits a dancerly style of movement from its participants, though anecdotal claims of some participants dancing in the above installations show that this kind of movement is not excluded either[7]. A definition of "dancerly", depends on how one categorises a movement style, but in my opinion, it is not usually a secondary outcome of the movement of a body in space. In these installations, the body's movement is firstly about the creation and manipulation of the sound-space, and not so focused on movement as a primary artistic goal in and of itself. This emphasis is not true of works whose manipulation of sound is acquired by choreographed or improvising dancers.

4.2 Choreographed Sounds

The concerns and modes of performance that are displayed in dance performances where the dancer affects sound are as varied as the number of works that employ these technologies. This makes it difficult to draw out generalizations about how the form addresses various questions. Questions such as "how does technology support and augment the choreographer's creative process", "is the technology leading the dance or visa versa", "what new territories, if any, are being explored in dance, by the use of these new techniques" and "what do these new works say about the (dancing) body and its relationship to real and virtual environments" tend to dominate discussion among dance practitioners. These questions are contrasted against the kinds of questions that computer engineers, programmers and electronic musicians, who are invariably involved in the implementation of these technologies, tend to ask. The programmers and musicians are usually tackling simpler questions of “what, how, and when”: “what aspect of the dance am I going to analyse”, “how am I going to transpose this information so that sound can be produced in such and such a manner” and “when am I going to use one or another method to do this?”. Beyond these utilitarian questions, some musicians and programmers have asked “what is the relationship between dance and music now that dancers are controlling the sound” and “what is my relationship to authorship in this situation”?

Without canvassing the entire debate, I would like to examine some of the discussion that surrounds dancers’ engagement with technology. The difficulty with attempting to focus on a discussion of technology that mediates between dancers and sound systems from the dancers’ perspective, is that dancers are rarely this specific, unless prompted by musicians or programmers. And usually this prompting, however subtle, ends up being directive, rather that allowing the dance practitioner to identify and explore what they perceive to be the important questions. I have tried to glean from more general discussions, initiated by dance practitioners, what dancers’ attitudes are to motion-driven music and other digitally mediated activities.

Susie Ramsay, in a 1994 discussion of virtual reality (VR) systems, raises the point that many systems make little use of the skill that dancers have in using their body, relying on gear that measures only the movements of head and hands. In addition, she questions the escapist promises that abound in the hype surrounding VR capability. On the one hand there is the notion that bodily experiences can somehow be intensified in these environments, or that the body can be freed from its usual physical restrictions, for instance by allowing it to fly, walk through walls or fall from a great height without being hurt. At the time the article was written, it was easy to suppose that what VR implied was a devaluing of natural physical experiences, and that it was a logical consequence of the “Judaeo-Christian, patriarchal, dualistic culture that marginalizes and degrades the physical body.” A VR that excludes the reality of pain, sickness, suffering, aging and dying is, therefore, incomplete[8].

On the positive side, Ramsay argues that some examples of new technology were attempting to address just these issues, particularly the issue of engaging more of the body in the exploration and manipulation of a system’s interface and therefore bringing more of the body’s skills into play. She also cites the example of Myron Kruger, who, in his explanation of one of his VR environments, says that he is not going to let you fall in this environment unless he can find a way of making you feel pain[9]. In the twelve years since this article was written, there has been an expansion in the number of interfaces that attempt to engage whole body movement and that try to preserve some notion of physical consequence, whether that consequence be haptic, sonic or imagistic in nature, but for many dancers, these questions still remain pertinent ones to ask, when contemplating creative work using interactive technology.

In 2002, the Ohio State University Department of Dance hosted a think tank on new performance tools "to explore the practical and conceptual implications of working with interactive tools, instruments and computer-controlled systems within performance conditions and exhibition-installation contexts."[10]

Attending the think tank were practitioners from many fields: choreographers, dancers, composers, media and installation artists, programmers, engineers, architects, writers, curators, DJs and cognitive psychologists.

The questions that seemed to have the most resonance for dancers and choreographers during the two days of discussions were generally the 'big questions', tending towards a philosophical or aesthetic enquiry. Bebe Miller was interested in how compositional processes might be deepened by an engagement with interactive technologies and telepresence. What is inherent in these technologies that make their use in dance qualitatively different from traditional methods for constructing a piece?[11] This line of thought would seem to resonate with Marshall McLuhan's dictum "The medium is the message"[12]. To use technology means that its history and relationship to people and their environments, as well as its function as a tool, are being used in the artwork's creation. Another implication here, is that any user of technology must also accept some responsibility in this history and these relations, in that the artist is an agent, or potential agent, for the continuance of one or another aspect of the technology's attributes. That so much technology has roots in the military-industrial complex comes with the caution that its use and acceptance by artists be informed.

Miller also presented the possibility that these technologies might allow a different kind of relationship with audiences; that different modes of practice enable dancers to get out of the studio in a way that bridges a gap between the makers of dance and non-makers. Miller's concern here was less about the ability of technology to broadcast and disseminate dance, than the possibility that audiences might one day be able to engage with the process of making dance through new interactive interfaces.

Lali Krotoszynski who created a work entitled "Dance Juke Box", designed like a cross between a video game and an image and sound-mixing console, presented an example of such alternatives. It is intended for use on the Internet by single users and enables them to compose a dance using twelve different animated dance sequences and five options for music. Mediated dance at this level offers new ways to collaborate, either with the public or a select team, generating a potential for large-scale accessibility, a combination that is rare in dance productions[13].

Several of the dance practitioners related the sensation of having their body “extended” in some way by technology either by “feeling different”, feeling “huge” or of the sense that they were dancing with a very different kind of partner; a partner made from data where the usual sense of gravity and other physical forces no longer apply[14]. These feelings confirm, in a very visceral way, some of the expectations of the technology that Ramsay mentions above, but perhaps in not quite the way anticipated by the makers and promoters of the technology. Feeling ‘huge’ or ‘different’ or being in control of some ineffable ‘datafied’ other, is not quite the same as the feeling of flying, floating, falling or being convinced one is inside an alternate environment.

“What are the connections between the use of the body as an [musical] instrument and how can the dancer have ‘control’ over the media and what are the ramifications for the training of the body in a media play or control environment?”[15]. These questions were a frequent leitmotif during the think tank. Given that most works, which could be said to address these questions, are iconic – one-of-a-kind – pieces, it would be difficult to imagine how to actually answer the question so that some general training program might be of any use and the idea that the body has a direct connection as an instrument, if posed to a musician would, no doubt, raise a smile as a statement of the obvious. The musician would simply ask, “is the body in question up to the task” and “is the interface [read instrument] any good?” (See “The Good Controller” Chapter 1.) This perspective also is dependant on whether the system is intended for use as a musical instrument and what your definition of an instrument is. The usual definition of an instrument as something that requires manual skill, rewards sustained effort and is capable of a wide variety of applications (repertoire), could not be applied to many dancer-controlled systems without significant modification as we have seen in chapter 2.

4.3 Cultural convergences between dance, music and installation

One of the connections between body and instrument is the mapping question. Which aspects of a body’s movement does one choose to map to sound parameters? This is no longer a technical question. Technology is capable of sensing almost any kind of movement, so the question is one of choice. Another connection is the conceptual one of similarity. In what ways does the (dancing) body resemble a musical instrument. If I was to take the definition of a “good controller” and try to apply it to a dancer, then I think we would find a very strong connection. Is it a body that requires skill to perform well? Does the body reward sustained effort and is capable of a wide variety of applications? Does an audience value the skill and effort of physical displays that go beyond the everyday use of the body? If all of these questions apply equally to dancing bodies as to musical instruments, then the connection between the two is clear.

There is a cultural convergence of dance and music, which is based on the way these art forms are taught, practiced, performed and received. The same cannot be said of sound installation works. There is some sense in which short-term effort can be rewarded, but the opportunity to “practice” on an installation for over ten years, as one would on an instrument, is unlikely to be very rewarding. There are no installations that are capable of a wide variety of applications and this is as it should be. The virtue of an installation, like any other artwork, is that it is specific. It is about something or says something in particular.

The appreciation of an installation is primarily aesthetic, not kinaesthetic, even when we are asked to engage physically with a work. It is possible to have a kinaesthetic appreciation of a sound installation, as in the Paine works above, but it is not a requirement. I use the phrase ‘kinaesthetic appreciation’ to mean an appreciation of the effort involved in moving the body *skilfully*, not simply in an everyday manner.

The convergence of installation work with dance or music may be seen, principally, in the way in which material elements are arranged. The design of a piece and the relationships that are established between the mediating system and other elements in the work, whether they be object-based or human, whether they interact with one another or whether static, are common considerations in all three practices. Structurally, these practices share the common ground of providing at least an open space for exploration, as opposed to linear or closed narrative environments. Where they diverge, structurally, is that dance and music have the option to be composed in a linear fashion and in fact may only be *experienced* linearly, whereas installation, by its very nature, is non-linear. An experience over time of an installation may be linear, but it may also be holistic, that is, some installations may be wholly experienced in a single moment, like some paintings or other static artwork. Interactive sound installations do not fit the usual mould. Because the experience and interaction is in relation to a time determinant medium – sound and dance – holistic experience, in the manner above, in sound installations, is forever illusive.

The question of the dancer’s control over the media is a thorny one, and one that the dance and technology community discussed at some length in their email list service in 2002-2003 [16]. Some practitioners felt that ‘control’ was not the best term to apply, and in this, the reasoning reflects Rokeby’s comments above. Control is not necessarily the point; rather, exploration of new relationships and possibilities within a system is a more apposite point of view.

Another issue with dancer control over other media, is that new skills are required for each new piece made in this way, and the skills can be very complex, requiring a dancer to master their own body, as well as manipulating sensors that have complex relationships to parameters in other media. The effort to acquire these skills is not always in proportion with the results attained. A dancer usually spends ten years acquiring the skills to execute a wide variety of movement approaches. To then be asked to invest large tracts of time to acquire new skills that may only be applied to one or two pieces might seem disproportionate to some.

A musician might argue that learning new skills is no more difficult than learning a new instrument, something that most experienced musicians are familiar with, but a dancer might respond that it is highly unlikely that a musician could be expected to perform at a high level on two instruments at once, or dance and play an instrument simultaneously. Having said this though, it is not entirely out of the question. The limits of human skill and coordination have not yet been exhausted, but is the extension of these limits what the author of a system would always want? Displays of (composed or choreographed) virtuosity have usually followed on the heels of extraordinary performers. Virtuoso J.S. Bach composed the "Well Tempered Clavier" after many years of performance and improvisation and later, works by Liszt and Paganini, for example, followed the same model. It will be some time, if ever, before the world develops a virtuoso multimedia performer who is skilled in manipulating their body and sound and image concurrently. The uncertainty as to whether such a performer could ever exist, stems from the non-standardisation of multimedia technology itself. Standardisation would seem to be a prerequisite condition for the evolution of a virtuoso in the same way that the standardisation of the keyboard's layout facilitated the evolution of so many keyboard virtuosi.

4.4 Conclusions

4.4.1 Pseudo-synaesthesia

The opening question from chapter one, "What aspects of sound forms correspondences with movement, force or position of a body or bodies in a space sensitised by devices for acquiring gestural and or topographical data" can not really be narrowed to one or another sound or movement parameter correlation. The correlations between sound and movement remain largely determined by the subjective factors of design. To this we must add the objective fact that there are many systems for implementing movement-generated sound and with each system, an even more numerous set of possibilities for expressive manipulation.

Each new work is itself a new attempt to solve the problem of correlation within the limits of the piece. The goals of the piece can also determine the strategies employed, so where Palindrome or Winkler's goals are to reflect the quality of the movement, a one-to-one and a one-to-many mapping strategy can be used with varying materials from one section of a work to another with effective results. Where an exploration of a system's possibilities is an aim, a one-to-many and a many-to-many strategy, either in sequence or alone, as in Paine, Rokeby or Mustard's pieces, is capable of providing rich material over substantial periods.

The objective fact that different synthesis methods enable access to their own set of modulating parameters also militates against a unitary solution to correlating movement to sound. While frequency modulation synthesis allows manipulation of one or more 'controlling' oscillators, granular synthesis does not. While granular synthesis allows manipulation of grain density, for example, additive synthesis does not. While additive synthesis is good for precise control of spectral parameters, granular is not.

While we could look at parameters common to all synthesis methods, there is no objective synaesthesia between movement and sound parameters. My experimentation with various parameters show that it is possible to make expressive sound by mapping any movement parameter to any sound parameter, provided that the scaling of one to the other is an effective scaling. Scaling itself can be linear, exponential, sinusoid, or some other asymmetric or subjective non-linear scale.

4.4.2 Effective technologies

An optimal system for analysing and translating movement in a scene to sound synthesis parameters needs to capture all of the elements of movement: gravity, pressure and tension, fine motor gestures, gross motor gestures, position of body in space, position of body parts in relation to one another, position of bodies in relation to one another and position of a body's parts in relation to another body's parts.

These elements seem to be indispensable capture parameters for a system that attempts a thoroughgoing approach to motion capture. A system that would enable this approach would be a combination of optical motion capture, data gloves, gyro' motion capture and numerous pressure sensors and muscle sensors distributed over the body. Many computers would be required to process the data from all of these sources and a team of programmers and operators needed to manage the flow of information from one part of the process to another. Such a system might seem overly complex, and it is difficult to imagine an outcome in sound alone that would justify the expense of gathering so much data. The problem with capturing every detail of movement and spatial relationships is that there could end up being more information gathered than can be used for translation into sound.

Therefore a minimal system would seem to be both more effective at the level of time and data management, cost and simplicity. A *Miburi* or a gyro motion suit coupled with a video motion tracking system enables some gestural and some positional data to be extracted from a scene with manageable data sets for translation to sound parameters.

Tracking multiple dancers in a scene adds to the complexity of the system and the amount of resources necessary for analysis and translation, so a simpler system to begin with is recommended.

4.4.3 Practice in context

Dis-Patch was the principal effort representing my research into ways in which movement in space might successfully correlate to the parameters of sound. The few pieces that preceded *Dis-Patch* were interesting studies and experiments, but will not be a feature of this section of the thesis.

The positioning of the three elements of choreography, design and sound in a balanced relationship counter-posed *Dis-Patch* to the traditional hierarchical relationship where one or another of the media drive the work. The element of visual design, including costume and scenic display, became integral to the process and the success of the performance and the way in which the represented media were alive to each other.

Fig1. Image from third section of *Dis-Patch*

As a translation of choreography into music, the work may be said to include artefacts of dance culture that have no equivalent in music in much the same way that a literal translation of words in one language to another may not convey the same meaning in the other language because of the changed cultural context. As a combination of disciplines within the same temporal frame, these cultural differences appeared to some to be reconciled. They were not necessarily resolved, but the experience of co-habitation through and over time made them less strange.

Although the dancers had a general understanding that their movements influenced the sonic outcome it was felt unnecessary and possibly even counterproductive to explain a high level of detail as to how their movements manipulated sound. This was a significant departure from current practice, where dancers are often required to learn the system in order to play it like a musical instrument[17]. The intention behind this approach was that almost any choreography that included the composition of movement in space over time as a primary consideration might be a suitable candidate for translation into sound. This approach was felt to enable the dancers the freedom to concentrate on what they do best; that is, move their bodies without feeling *responsible* for the sonic result. What was not anticipated was the degree to which the result influenced the dancers' subsequent movement, in particular, energy, tempo and articulation (legato/staccato movement etc). These influences multiplied the sensation of interactivity in the system and the perception of an intermedia conversation taking place before one's eyes and ears.

Despite my knowledge of the system, there were unexpected musical results from the distribution of the various colours in space. In both cases, our immediate subjective response was to try to fix the problem, but in some instances the unintended artefacts were left to 'play out' to see/hear whether they could be used and replicated in future. Thus the process of making the work revealed sounds that whilst alien to musically conventional ears, were nevertheless valid and interesting sonic translations of the choreography.

4.4.4 Shortcomings

Within this new framework, the shortcomings of the piece stem from, in my opinion, two sources. Firstly, the elements that combined to produce real time musical composition were underdeveloped and secondly: the choreography, as acknowledged by the choreographer, required more work.

With regard to the composition, the overall texture for the majority of the piece is very similar. Four or five-voice textures predominate from section two to the end of the work with very occasional respite. Although the instrumental timbre and tempo vary, the dynamic is fairly consistent – mostly loud – and the rhythmic relationship between voices and across tempi is also a constant. Rhythmic unison dominated in sections where the musical material had definite pitch and this was in most sections, so a thorough re-composition is needed to ameliorate this sense of sameness in the development of some of the basic musical building blocks of the work. This lack of musical development over time was felt to be a serious flaw, resulting in auditory fatigue.

One reason for this flaw slipping through to performance was that too much time was spent on the other technical issues of mapping and successful colour tracking and that too much reliance was placed on a single algorithm as the sole musical template to which movement parameters were mapped. Development of the algorithm, or the addition of several algorithms, as well as timbral and temporal development, is required for a more successful musical outcome.

Choreographically, the movement vocabulary tended towards a vocabulary that the choreographer was very familiar with and as such, felt to be less adventurous than it might otherwise have been. From my point of view, though, the choreography was somewhat misaligned with the basic premises built into the system's programming: that is, that the movement was more concerned with the composition of gestures and body shapes than the relationship of the body to the space in which it was moving. A more architectural, or topographical approach to choreographing within space might have been more successful in foregrounding the sonic differences conceived in the programming environment.

4.4.5 Global Position

One of the ways in which the music for *Dis-Patch* enters into the broader global discussion on, for example, the use of alternative controllers is its heavy reliance on sounds based on traditional western acoustic instruments. These sounds have a high indexicality, meaning that the gestures normally associated with producing them are relatively familiar – force-based gestures like pushing, pulling (on a violin bow) and blowing (down a length of tube), trajectory-based (point, put, glide) and pattern-based (chew, walk, repeat things) gestures[18]. Research in this field aims to replicate the general principles of the gestures involved in traditional sound production onto alternate controllers and to emulate the levels of control and complexity over sound that an instrumentalist attains through many years of practice. This same approach and with the same general aim of greater/finer expressivity is often used in the context of dancer-manipulated controllers. *Dis-Patch*, because of its use of topographical data, searches for its expressivity in the translation of large-scale choreographic structures rather than the movements of individual limbs. The "gestural primitives" at work in *Dis-Patch* are spatial and relational as well as trajectory and pattern responsive.

The way in which the spatial relationships between colours are mapped to sound alters from time to time. This is equivalent to exchanging, for example, a force-based gesture with a trajectory-based gesture or even two opposing force-based gestures. This hardly makes sense in a traditional instrumental paradigm, but makes for fascinating and cogent variations in sound for the same choreographic material. The sum of these performed attitudes however, results in a number of perceptual associations and dissociations.

Many of the sounds that are made have an association with gestures co-extensive with traditional orchestral instruments and, by association, the highly prescriptive cultural, historical and political framework of the orchestra. In addition, the trajectory of these sounds, i.e. what the sounds do to "make music", is predicated upon the relationship of colours in space; a dissociative phenomena decoupling sound from its historical gesture and re-coupling / reconfiguring the sounds within a new gestural context and by extension, a new historical context. However, what the sounds do to "make music" dissociates or de-couples the sound from its traditional context. This happens when a particular instrumental sound is forced to 'play' outside its traditional capabilities; beyond the limits of instrumental technique or physique. A trumpet sound played in the register of a tuba does not sound like a trumpet until we hear it rise into its natural register.

The transition from not-trumpet to trumpet produces, or can produce a quite sharp disturbance. The belief a listener had about the sound heard initially is undermined by subsequent events. The disturbance operates in both directions – that is, trumpet to not trumpet.

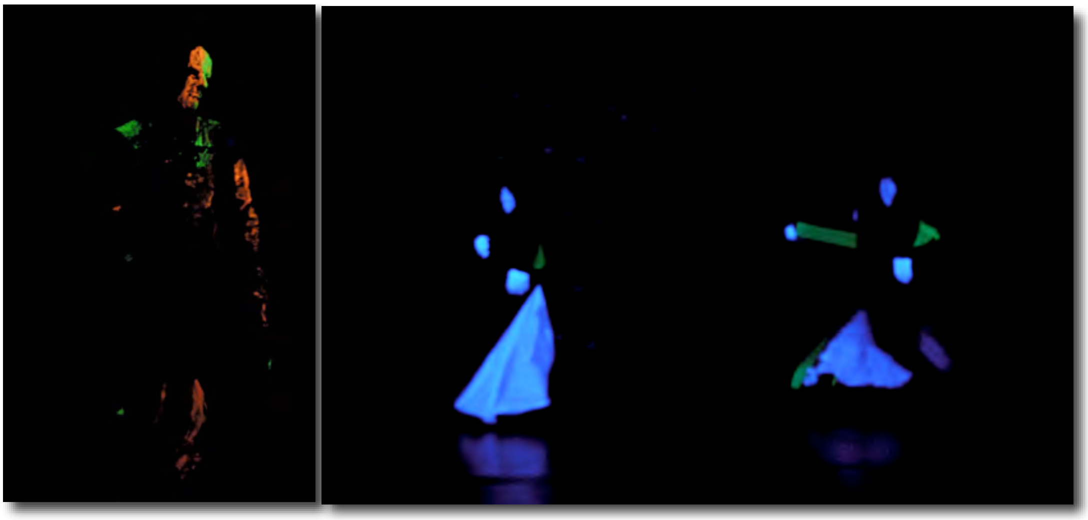


Fig2. Image from sections one and two of *Dis-Patch*

One of the aims of the work was to take a limited set of content from a single medium (choreography) and express that content across several other media (sound and video) simultaneously. It was not envisioned that these secondary media would be slaves to the principal medium, but would exert their own will on the content. Thus the translation of choreography into sound is translated via the filter of a compositional will and similarly the video projection reconfigures the choreographic content on its own terms. We have taken a traditional aesthetic – unity of content – and developed this content over diverse media. The work may be seen as re-problematising the relationship between music, dance and visual design in costume and videography.

Chapter 4 References

1. Rokeby, David (1985). Predicting the Weather. MUSICWORKS 33: Starting All Observations From Scratch. Winter 1985/6. http://homepage.mac.com/davidrokeby/.html Accessed 30.11.05.

2. ibid.

3. Rokeby, D. (1990). The Harmonics of Interaction. MUSICWORKS 46: Sound and Movement. Spring 1990. http://homepage.mac.com/davidrokeby/.html Accessed 30.11.05.

4. Rokeby, David (1985). Predicting the Weather. MUSICWORKS 33: Starting All Observations From Scratch. Winter 1985/6. http://homepage.mac.com/davidrokeby/.html Accessed 30.11.05.

5. Rokeby, D. (1995). Transforming Mirrors: Subjectivity and Control in Interactive Media. Critical Issues in Electronic Media, New York, State University of New York Press. http://homepage.mac.com/davidrokeby/mirrorsart.html. Accessed 30.11.05. p. 133-158.

6. Paine, Garth. (1988 and 2000 respectively). Artist's showreel DVD. Unpublished, courtesy of the artist.

7. ibid.

8. Ramsay, S. (1994). Bring Your Body: The Dance Community and New Technologies. Kunstforum International 133(Feb-April 1996). http://www.art.net/~dtz/theory.html. Accessed 13.08.04.

9. ibid.

10. Birringer, J. (2002). Think Tank Report: New Performance Tools: Technologies / Interactive Systems. Interactive Performance Series. Ohio, Ohio State University. p 1-11.

11. ibid. p13.

12. McLuhan, M. and Fiore, Q. (1967). The Medium is the Massage. Co-ordinated by Jerome Agel. Bantam Books / Random House, New York (1967).

13. Birringer, J. (2002). Think Tank Report: New Performance Tools: Technologies / Interactive Systems. Interactive Performance Series. Ohio, Ohio State University. p 8.

14. ibid. p 18.

15. ibid. p 6.

16. Various (2004). Dance Technology Archive, http://www.dancetechnology.com/dancetechnology/archive/. Accessed 15.04.03.

17. Coniglio, Mark. (1999). MidiDancer. Troika Ranch, website http://www.troikaranch.org/mididancer.html. Accessed 29.01.05.

18. Choi, I. (2004). Gestural Primitives and the Context for Computational Processing in an Interactive Performance System. 1st. 2000. CD Rom. Ircam. Wanderley, M and Battier, M. p. 207-208.

Appendix to chapter 2

[1] Pressure - measures the amount of tactile pressure exerted on a single point or discreet area. Usually this is a piezoresistive type of transducer where resistance is decreased when pressure is increased allowing greater levels of voltage through the resistor. Typical thresholds are from under 100g to 10 or more kilograms, which outputs a voltage from 0 to 4 or 5 volts. Dance company Scadada have also used simple binary pressure sensors custom made from tin foil and rubber matting, functioning similarly to the type of mat used in the doorways of retail shops to trigger a bell or buzzer announcing the arrival or departure of a customer. Scadada uses the mat to trigger sound samples and lights.
The pressure sensor can be used for both gestural and topographical motion capture depending on whether it is placed on the body or in space. Some types of tactile pressure sensors can be moulded to conform to complex surfaces such as the human body. [1].

[1] Flexion – operates on a similar principal to the pressure transducer – the greater the angle of flex the lower the resistance of the resistor. Typically operates from angles of 180^{o} (straight) to 0^{o} (fully flexed or closed). Some flexion sensors operate in both directions i.e. from 180^{o} increasing to 360^{o} or decreasing to 0^{o}. Recent developments in flexion sensors enable six degrees of freedom, but with output restricted to voltages representing 180^{o} to 0^{o} so rather than a flat tape-like structure that bends in only two directions, a flexible optical fibre is used and the amount of escaping light is a measure of the angle.

[1] Light level – often a photocell or phototransistor sensitive to light levels which can measure the level of ambient light or if of a more directional design, may be used to measure the level of light within a discreet beam of space. These are typically used in installations or instruments for triggering the presence of an object at a particular point in space or at a particular point on an instrument. Photocell sensors that measure the general ambient light can also be used dynamically in installation and performance offering a range of output voltage, again typically from 0 to 4 or 5volts. [2].

[1] Proximity – measures the closeness of an object to the sensor. There are several types of proximity sensor. Infrared sensors similar to the sensors found in security alarm systems operate by emitting a beam of infrared light, which bounces off a reflective object and is detected by an infrared detector usually placed next to the emitter. The closer an object is to the sensors the more infrared is reflected back determining the distance of the object. Some infrared systems detect only the presence of an object.
Ultra-sound emitters and detectors operate on a similar principle except that short bursts of sound are emitted and detected instead of light.

[1] Angle – a potentiometer that measures the angle of rotation between the sensor and an object fixed to it. A typical basic potentiometer operates through 10kΩ. Used in a wide variety of situations, this sensor is most commonly recognized as the volume knob on your average stereo system.

[1] "In a Hall effect sensor a current is passed through a semiconductor material. When a magnetic field is applied perpendicularly to the surface of the semiconductor, a voltage is developed. This Hall voltage is proportional to the applied field intensity, driving the magnetic speed sensor" [3].

[1] "Inertial gyros with a gyro reference have an inertial mass used as a reference for rotational movements. A vibrating tuning fork or plates measure Coriolis force in inertial gyros made using MEMS technology." [1]

[1] Full text of e-mail from David Chesworth regarding *The Light Room*

From: David Chesworth [redacted]
Subject: Re: your co.inspace collaboration
Date: 25 February 2005 9:36:47 AM
To: Jonathan Mustard [redacted]
hi Jonathan,
There wasn't any direct linkage to the motion suit or any video capture. We did discuss it but I have never been interested or convinced in the value of mirroring or interpreting actual visual or movement gestures in sound. For me it is too literal a starting point and presents me with a situation where the only creative path forward is to introduce distortions in that relationship to avoid banality, which kind of defeats the original purpose.

I have however been very interested in and aware of relationships between the various aural, visual and spatial worlds within a work including The Light Room. Sound can delineate and contextualize space. Sound can position both performer and audience, as participants within this space.

Plus the contextualized sound-space can become a site that has the potential to be read as literal or metaphoric.

In The Light Room many physical and sonically defined sites coexist. Some of them evolve and recede over time. Within these sites physical performances of the body, sound and voice can and indeed must interact.

I did use 'noise' in its many guises as a material to define different spatial arrangements within the work. We employed a three dimensional approach to sound dispersion. We used a system whereby we could arrange sound within a three-way axis - depths, width and height. This became a compositional parameter.

Hope that helps

regards

David

In response to . . .

From: [redacted]

Date: Thu, 24 Feb 2005 16:29:26 +0800

To: [redacted]

Subject: your co.inspace collaboration

Hi David,

I'm just looking for a bit more information on your Light Room collaboration with Co. in space for some research I'm doing on Motion Capture for sound manipulation.

I was wondering whether the mo cap suit was used to manipulate elements of your score / sound / composition and in what way.

Any organic link between say a performers position or orientation in space to various parameters of sound?

Any gestural link to sound?

Hoping you've got a few moments to answer. I did see the description of Light Room on companyinspace's web site, but it's not too detailed.

Thanks for any info

Jonathan Mustard

www.jambird.com

Appendix References

1. Globalspec (2005). About Pressure and Tactile Imaging Sensors. City unknown, Globalspec. **2005:** http://sensors-transducers.globalspec.com/LearnMore/

2. Axel Mulder, F. a. M. (1999). Design of Virtual 3D Instruments for Musical Interaction. Graphics Interface '99.

3. Globalspec (2005). About Speed Sensors, Magnetic. City unknown, Globalspec. **2005:** web page. http://speed-sensors.globalspec.com/LearnMore/Sensors_Transducers_Detectors/Velocity_Sensing/Magnetic_Speed_Sensors

Appendices to Chapter 3

There were problems capturing the whole image using the Cyclops object. I wrote to Singer and he advised that the capture resolution needed to be adjusted for certain video drivers.

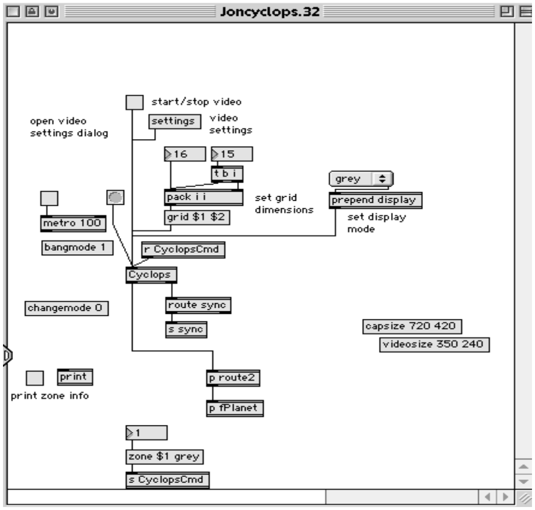


Fig4. Cyclops object and initialising commands

Joncyclops.32

A “Capsize” command set to 720x420 worked to capture the whole image for more accurate analysis. A further problem emerged in this version where the sound image seemed only to be shifting once every second instead of once every 100 milliseconds.

Fig4. is an image of the patch showing the Cyclops object and its relation to initialising commands and other sub-patches. (Message boxes eg. capsize 720 420 are commands and other objects have to do with processing or generating data. The number boxes above the command grid $1 $2 enable the division of the video image into a grid of 16 by 15. The number boxes can be “scrolled” to a higher or lower integer to vary the grid dimensions. The objects p_route2 and p_fPlanet are sub-patches that process data coming from the Cyclops object. See fig5 below)

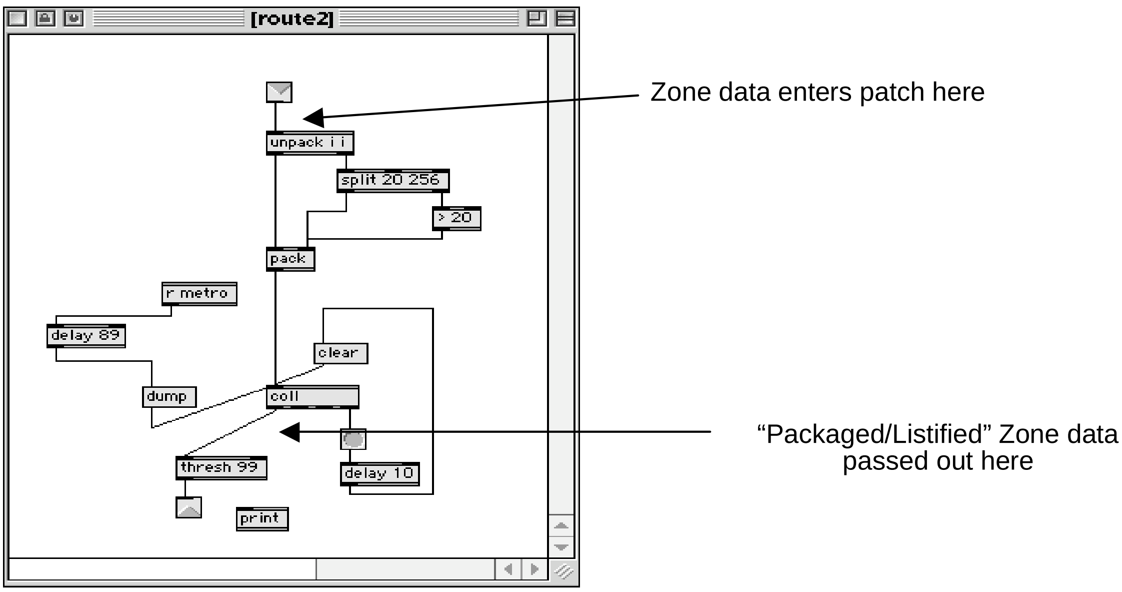


Fig5. sub-patch of Joncyclops.32

The sub-patch, route2, in fig5 shows initial processing of zone data from the Cyclops object. Two integers are output for each zone. One is its address and the other its value. In this example the value is a grey-scale value from 0 to 256 where 0 is black and 256 is white. The split object is a way of filtering out unwanted lower level light. Even in a very dark room the video camera registers some light level. 240 of these integer pairs are received every 100 milliseconds and are collected into a single list by the coll object then passed out to the next sub-patch to be turned into a filtered noise.

E-mail from Eric Singer.
"Hi Jonathon,
I've been away for the last week, so I'm just now catching
up on
email. Capturing only part of the image is an
incompatibility
between Cyclops and certain video drivers. As of yet, I
haven't been able to solve this problem. The workaround is
usually to find a capres that captures the whole image. With
one of my cards, it is
640 x 240 (strangely).

The zone data list is always output as one list per
zone. There may
be some other objects that can help you list-ify the whole
output...
possibly Peter Elsea's Lobjects?

Eric"

Jittertracking.06

Combines elements of Jittertracking.02 and Jittertracking.03. A sine wave ‘follows’ each coloured light. Frequency defaults to 1680hz when the vertical location of a colour is tracked to 240, the boundary of the video frame in this case. This is an unintentional artefact of the patch and distracts the viewer / listener from the perception of a one to one correlation slightly.

Bibliography

Ascension Technology Corporation. (2006). Products - MotionStar Wireless 2. Burlington, Ascension Technology Corporation. http://www.ascension-tech.com/products/motionstarwireless.php. Accessed 4.04.2006.

Battier, M. (2004). Electronic Music and Gesture. 1st. 2000. CD Rom. Ircam M.M. Wanderley and M. Battier.

Bevilacqua, F., Ridenour, J, and Cuccia, D. (2002). 3D motion capture data: motion analysis and mapping to music. Workshop/Symposium on Sensing and Input for Media-centric Systems, Santa Barbara CA, SIMS.

Birchfield, D. (1999). 1999 Columbia University Interactive Arts Festival. Computer Music Journal. 24.

Birringer, J. (2002). Think Tank Report: New Performance Tools: Technologies / Interactive Systems. Interactive Performance Series. Ohio, Ohio State University.

Birringer, J. Suite Fantastique: Architecture and the Digital. (2002). Location One Gallery. http://www.location1.org/birringer_suite_fant2.html. Accessed 10.02.02.

Blaxter, L., Hughes, C., and Tight, M. (2002). How to research (Second ed.). Berkshire, Open University Press.

Buchla, D., Sonami, L. and Machover, T. (2004). Round Table. 1st. 2000. CD Rom. Ircam. Wanderley, M. and Battier, M.

Burt, W. (1990). Fair Exchanges. Writings on Dance. 5 (Autumn).

Carter, C. (2001). Music and Movement. Sound on Sound. October 2001.

Choi, I. (2004). Gestural Primitives and the Context for Computational Processing in an Interactive Performance System. 1st. 2000. CD Rom. Ircam. Wanderley, M and Battier, M.

Coniglio, M. (2004). Mididancer, Troika Ranch. http://www.troikaranch.org/mididancer.html. Accessed 29.01.05.

Dobrian, C. (2001). Aesthetic Considerations in the Use of "Virtual" Music Instruments. Society for Electro-Acoustic Music in the United States. Spring 2003.

Dobrian, C. (2003). Gestural Control of Music Using the Vicon 8 Motion Capture System. New Interfaces for Musical Expression, Montréal, Québec, Canada, NIME.

Dobrian, C. (2003). MSP 45 Tutorials and Topics. California, Cycling ’74.

Doornbusch, P. & McIlwain, P. (2003). Converging Technologies. Australasian Computer Music Conference, Western Australian Academy of Performing Arts at Edith Cowan UIniversity, Perth Australia, Australasian Computer Music Association.

Durling, K. (2000). The Electronic Avant-Garde, Hearing Anew, Moving Ahead. Contemporary Music Review. May 6 2000.

Dyer, S. Martin, J., Zulauf, J. (1995). Motion Capture White Paper, Windlight Studios, Alias and Wavefront.

Furniss, M. (2000). Motion Capture. Massachusetts, MIT Communications Forum. http://web.mit.edu/comm-forum/papers/furniss.html. Accessed 04.04.07.

Godlovitch, S. (1998). Musical Performance: A Philosophical Study. London: Routledge.

Gurevich, M. (2001). The Accordiatron: A Midi Controller for Interactive Music. Proceedings of the 2001 New Interfaces for Musical Expression. Seattle, USA: NIME.

Highwater, J. (1978). Dance: Rituals of Experience. New York, A & W Publishers, Inc.

Hunt, A and Kirk, R. (2004). Mapping Strategies for Musical Performance. 2000. CD Rom. Ircam. Wanderley, M and Battier, M.

Hunt, A., Wanderley & Kirk (2000.). Towards a Model for Instrumental Mapping in Expert Musical Interaction. International Computer Music Conference, San Francisco, International Computer Music Association.

Laban, R. (1966). The Language of Movement: A Guidebook to Choreutics. Boston, Plays.

Laboratorio di Informatica Musicale. (2004). EyesWeb Open Platform. Genova, InfoMus Lab. www.eyesweb.org. Accessed 27.01.05.

Machover, M., Wanderley, M. and M. Battier. (2004). Round Table. 1st. 2000. CD Rom. Ircam.

McLuhan, M. and Fiore, Q. (1967). The Medium is the Massage. Co-ordinated by Jerome Agel. Bantam Books / Random House, New York (1967).

MetaMotion. (2004). Gypsy Gyro. USA, MetaMotion. www.metamotion.com. Accessed 20.01.05.

Merriam, S. B. (1998). Qualitative research and case study applications in education. San Francisco, Jossey-Bass.

Milech, B. & Schilo, A. (2004). 'Exit Jesus': Relating the Exegesis and Creative/Production Components of a Research Thesis. TEXT Special Issue No 3 (April 2004).

Modler, P. (2003). An Experimental Set of Hand Gestures for Expressive Control of Musical Parameters in Realtime. Proceedings of the 2003 Conference on New Interfaces for Musical Expression, Montreal, Canada, NIME.

Mulder, A. (1996). The I-Cube System User Manual. T. Sternberg. Los Angeles, Infusion Systems.

Mulder, A. Fels and Mase (1998). Human movement tracking technology: resources. Vancouver, Simon Fraser University. http://www.xspasm.com/x/sfu/vmi/HMTT.add.html. Accessed 15.05.03.

Mustard, Jonathan. (2002). Correlating Movement in Space to the Parameters of Sound. Proceedings of the 2002 Australasian Computer Music Conference. Melbourne: ACMA.

Paine, Garth. (1988 and 2000 respectively). Artist's showreel DVD. Unpublished, courtesy of the artist.

Raes, G. (2003). Gesture Controlled Musical Instruments. Gent, Hogeschool Gent - Music & Drama Department, Orpheus Institute, Logos Foundation.

Ramsay, S. (1994). Bring Your Body: The Dance Community and New Technologies. Kunstforum International. 133. (Feb-April 1996). http://www.art.net/~dtz/theory.html. Accessed 13.08.04.

Rokeby, D. (1983-4). Reflextions. Ontario, mac.com. Home pages. http://homepage.mac.com/davidrokeby/reflexions/. Accessed 19.10.04.

Rokeby, D. (1984). Body Language. Ontario, mac.com. Home pages. http://homepage.mac.com/davidrokeby/bodylanguage/. Accessed 19.10.04.

Rokeby, David (1985). Predicting the Weather. MUSICWORKS 33: Starting All Observations From Scratch. Winter 1985/6. http://homepage.mac.com/davidrokeby/.html Accessed 30.11.05.

Rokeby, D. (1986). Very Nervous System. Ontario, mac.com. Home pages. http://homepage.mac.com/davidrokeby/. Accessed 19.10.04.

Rokeby, D. (1990). The Harmonics of Interaction. MUSICWORKS 46: Sound and Movement. Spring 1990. http://homepage.mac.com/davidrokeby/.html Accessed 30.11.05.

Rokeby, D. (1995). Transforming Mirrors: Subjectivity and Control in Interactive Media. Critical Issues in Electronic Media, New York, State University of New York Press. http://homepage.mac.com/davidrokeby/mirrorsart.html. Accessed 30.11.05.

Rokeby, D. (2004). softVNS 2 real time video processing and tracking software for Max. Ontario, mac.com. http://homepage.mac.com/davidrokeby/softVNS.html. Accessed 19.10.04.

Rovan, Wanderley, Dubnov & Depalle. (1997). Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance. Genova: The Technology of Emotion Workshop, 1997.

Rovan, Joseph B., Wanderly, Marcello M., Dubnov, S., and Depalle, P. (1997). Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance. Kansei- The Technology of Emotion Workshop, Genova - Italia.

Sky, H. (2002). Company in Space. Melbourne, www.companyinspace.com. Accessed 24.02.05.

Van-Raalte, C. (2002). BodySynth. UK, BodySynth. www.synthzone.com/bsynth.html. Accessed 20.01.05.

Various (2004). Dance Technology Archive, http://www.dancetechnology.com/dancetechnology/archive/. Accessed 15.04.03.

Vickery, L. (2002). The Yamaha MIBURI MIDI jump suit as a controller for STEIM's Interactive Video software Image/ine. Australasian Computer Music Conference, Melbourne, ACMA.

Vickery, Lindsay. (2002). The Robosax Project. Australiasion Computer Music conference. Melbourne: ACMC.

Waisvisz, M. (2004). Round Table. 1st. 2000. CD rom. Ircam M.M. Wanderley and M. Battier.

Wanderley, M. and Rovan, J. (1997). Gesture Research in Music: Who, What, and Why.... IRCAM. Available: http://www.ircam.fr/equipes/analyse-synthese/wanderle/Gestes/Externe/people.html. Accessed 18.08.04.

Wanderley, M. (2001). Performer-Instrument Interaction: Applications to Gestural Control of Sound Synthesis. IRCAM – CENTRE POMPIDOU. Paris, UNIVERSITY PARIS 6.

Wanderley, M. (2004). Trends in Gestural Control of Music. Wanderley, M. and Battier, M. Paris, IRCAM.

Wechsler, R and Rovan, J. (2001). … Seine Hohle Form… - Artistic Collaboration in an Interactive Dance and Music Performance Environment. Proceedings of the Cosign Conference. Amsterdam: Cosign, 2001.

Winkler, T. (1995). Making Motion Musical: Gesture Mapping Strategies for Interactive Computer Music. Proceedings of the International Computer Music Conference, ICMC.

Winkler, T. (1997). Creating Interactive Dance With the Very Nervous System. Connecticut College Symposium on Arts and Technology, Connecticut, Connecticut College.

Winkler, T. (1998). Motion-Sensing Music: Artistic and Technical Challenges in Two Works for Dance. Proceedings of the International computer Music Conference, Bucharest, ICMC.

Winkler, T. (2003). Live Video and Sound Processing for Dance. Video, Technology and Performance Festival, Brown University.